Joint Demand Regulation and Capacity Management for Multi-cellular Clusters - a Stochastic Meanfield Control Approach

Abheek Saha
Hughes Systique Corporation
Email: abheek.saha@hsc.com

***Abstract*—Mean-field theory is a significant recent development in the field of stochastic optimal control. By allowing the optimal control functions to take into account not only the state of the individual agent, but also the common state of an entire ensemble of mutually inter-dependent agents, mean-field theory allows us to model ensembles of autonomous agents pursuing individually optimal trajectories in a shared environment. In this paper, the application of stochastic optimal control has been shown for a very standard problem of cellular networks, the optimum resource allocation problem. In modern cellular networks, the optimal resource assignment for individual cells has to take into account the loading of the entire network, since user stations are free to adjust transmission rates and migrate among cells and cells, while cells can also trade bandwidth between themselves. The problem is to achieve an optimal matching of available resources to the individual demands for capacity, taking into account the temporal and spatial variation in demand. By modelling the demand and capacity and their mutual interaction using mean-field theory, it has been shown that the matching problem can be cast as a distributed optimal control function. We have used a novel method to solve the corresponding mean-field game and demonstrated that the solution provides an effective mechanism for demand regulation and capacity assignment.**

***Keywords*—mean-field games; stochastic control; distributed resource allocation; distributed optimization; adaptive rate control.**

I. INTRODUCTION

Stochastic optimal control is a powerful technique to control time-varying systems with a random component in the inputs. Developed over the last fifty years from the base of variational inequality and dynamic programming, it has been applied in multiple disciplines ranging from finance to oil exploration and medical trials. By leveraging the fundamental strength of stochastic calculus and optimization theory, stochastic optimal control can help in developing the control law which is used to control live processes in the face of unknown, time varying inputs.

Wireless cellular networks have to deal with the problem of efficient resource allocation, and it is well known that this can be modelled as an optimal control problem [1]. In modern cellular networks the network resource allocation function has to deal with varying adaptive user demand as well which is also a stochastic processes. Application of stochastic optimal control to wireless networks, however, has been sporadic [2][3][4]. The immediate reason for this is that optimal control problems do not easily yield analytic solutions. In fact, most of the work in optimal control theory is limited to linear and linear quadratic state equations. Wireless protocol stacks are based on more complex laws.

However, there is a second, more difficult problem to solve. In the cellular wireless world, there is the problem of inter-node interdependence. This was recognized even in the 2nd generation wireless networks, which were designed to be interference constrained. Wireless networks of the 4th and 5th generation have taken this one step further. They are built around the principles of dynamic inter-cell coordination and cooperation for managing resources and matching them to demand. The need for cooperative resource management protocols is driven by two factors. The first is the ability of individual user terminals to use larger and larger bands of spectrum, while the overall network is spectrum-conservative; hence network nodes must be able to deploy spectrum flexibly in response to hotspots of demand, while minimizing the overall spectrum deployment. The second is the need for networks to dynamically adapt to large variations in demand, both spatially and temporally. Cellular networks are being moved towards newer and newer business cases such as wide-area connectivity for cellular networks supporting Internet Of Things, connected vehicles, etc. Most of these use cases are dependent on network nodes being able to flexibly adapt to new patterns in user behaviour. Hence, the paradigm of dynamically shared resources and network node cooperation is here to stay. There is wide-spread theoretical and practical interest in Coordinated Multipoint networks in 4G, Hetnets and Inter-Cell Interference Coordination (ICIC/eICIC). Indeed, the 3rd Generation Partnership Project (3gPP) has introduced the X-interface between network nodes as an explicit means of inter-node coordination in real-time, in order to make coordinated network operation possible.

A. Optimal Control for Wireless Networks

Network nodes are independent, yet coexisting agents, tied together by the constraints of shared resources and shared environments. In this situation, an optimal control law which treats the network node or cell as an individual agent and optimizes its performance in an isolated manner is not very useful and can result in severely degraded network performance. To apply optimal control theory to wireless network resource allocation problem, one would need to model the impact of resource usage by one node on the rest of the network. One solution is to try to solve the problem simultaneously by constructing as a N-dimensional control equation for all network nodes, i.e., the network state becomes a vector of states, one for each agent. This however leads to the dimensionality problem as the number of degrees of freedom increase as $\mathcal{O}(n^2)$. It also requires a degree of simultaneous coordinated control that is not feasible in most wireless networks.

A strictly adversarial approach (such as used in game theoretic techniques) has also been used with interesting results. In reality, network nodes cannot afford to be completely adversarial, since they are not operating in cooperation, not competition of each other. For example, it may make sense for a given node to hand over load to another node or to take over load from another node cooperatively. In these situations, the objective is to optimize overall network capacity, not individual node capacity. However, it is still possible to solve this problem as a game, by writing the utility function to take this into account. A second problem is that game theoretic approaches require inter-node negotiation, which requires a very large amount of signaling between nodes and their neighbours. This makes the game-theoretic solutions analytically difficult and hard to scale to a large number of nodes.

What is needed then, is a way to abstract the bulk state of the network and construct a common model for demand and capacity and their mutual interaction. The model should take into account both the effect of the network state on the individual agent, as well as the effect of the independent actions of the nodes and how it impacts the bulk network state in turn. As shall be seen below, the mean-field model provides exactly that.

B. The Mean-Field Extension to Stochastic Optimal Control

Mean-field theory originated in particle physics, where it is used to model the behaviour of a large number of particles within a single field. The states of the individual particle follows statistical laws, which is controlled by the field itself. However, its application to optimal control theory is relatively recent. It was only in the first decade of the 2000s that Lasry and Lions [5] and independently, yet nearly simultaneously, Minyi Huang and his team [6] kicked off a concerted research effort on optimal control of multiple interacting stochastic processes by modeling their interactions through a shared global variable called a mean-field constraint. Optimal control problems of this nature are called Mean-Field Games (henceforth MFG). The term field here is similar to the use of field in classical physics; a common, network-wide influence which modulates individual nodes in the field and is also modulated by them, under the control of a universal field equation.

The authors above showed that MFGs can be modelled as an extension to stochastic optimal control theory, by allowing the empirical distribution of individual network node states to be included in the transition and cost functions. This provides us a mechanism for incorporating the network state variables into individual node decision control algorithms. For example, Huang et al. in [7] use mean-field stochastic control as a way of optimal power control in wireless networks. Wireless nodes have to set transmission power so as to maximize the Signal to Interference Ratio (SIR), yet minimize cross-neighbour interference. In this case, the latter is modelled in terms of the empirical power distribution across the network.

C. Objective and Structure of the paper

In this paper, stochastic control with mean-field constraints has been applied to the problem of cellular resource allocation. The focus is to model the demand capacity gap in a multi-cellular environment and how this has to be incorporated into a dynamic model for aggregate demand in the per-cell level. By using rate adaptation models from the existing literature, this interaction and the resultant demand-capacity allocation problem has been modelled as a stochastic game, which can be solved to get a common optimal control law. A closed-form analytic solution has been worked out for a non-linear stochastic MFG which, to our knowledge is the first that has been presented in the existing literature.

The rest of the paper is organized as follows. In Section II, we introduce the problem in more detail. In the next Section III, we review the existing state-of-the-art in the area of cellular resource management and establish the contribution of this paper in contrast to the current literature. In Section IV a mathematical model has been defined for modelling adaptive user traffic, which is going to be the basis of the theoretical model. Section V contains an introduction to stochastic optimal control and its extension to the mean-field constrained game, along with the adjoint equation technique that shall be used as a basis for the solution of this game. In Section VI, the model demand regulation problem is introduced as a mean-field stochastic game and a closed form solution is presented. The success of the demand regulation algorithm is established through simulations in SectionVII. Finally, Section VIII contains the conclusion, analysis and a roadmap for future work in this area.

II. Problem Statement - Dynamic resource management in wireless cellular networks - a survey of existing literature

The resource management problem of cellular networks has been studied as part of dynamic network optimization since a long time and is seen as a fundamental component of the Self Optimizing Network (SON) [8],[9]. The problem is briefly described as follows: there is a network consisting of (possibly overlapping) cells in a given coverage area. Each cell is controlled by a network node (base-station). Within each cell, there are a number of active UEs (User Equipment), which have a requirement for network capacity; they communicate these resource requests to the base-station on a frame by frame basis. Each individual network node aggregates these per UE requests for capacity, into an aggregate demand which is the state variable of the network node, X_t^i, appropriately normalized. This term is frequently referred to the demand or the load of the cell. The network nodes also have a certain amount of capacity to handle this load, based on the resources available to it. A scheduling algorithm distributes the capacity to the individual UEs on a frame-by-frame basis.

The resources available in each cell are a combination of various different physical and computational resources, such as spectrum, power and backhaul capacity. All of these combine to determine the overall load handling capacity C_t^k

of a given cell. Clearly, one would like to minimize the gap between the available resources C_t^k and the resources which would be required to service the aggregate demand. The task of the resource optimization algorithm is to ensure an optimal deployment of resources per cell, so that the demand-gap (difference between capacity demand and capacity available) is minimized for all the cells in the network.

Since loading patterns vary dynamically in both temporal and spatial dimensions, the network has to be load-following. It must be able to use the reported aggregate load from each cell as a basis for spectrum allocation, reallocation and cooperative load-balancing. There are multiple mechanisms for this as has been reported in the literature. For example, in a dynamic network, real-time trading of spectrum may be possible within cells belonging to a group or cluster, under the control of a central controller. Alternately, cells may use power as a resource; in [10], the authors use a strategy of adjusting the transmit power of individual cells so as to let less loaded cells expand their coverage area to take more load. More cases shall be discussed in Section III-A. In a cellular network, the bulk of the cell-level resources (if not all) are shared resources, with either soft or hard constraints on their deployment in the network. No individual node can unilaterally change its deployment without affecting others. In many cases, the impact of reconfiguration of an individual resource may be network-wide. For example, a network node increasing its transmit power will cause interference among its immediate neighbours, whereas a network node deploying additional spectrum will cause co-channel interference with other nodes reusing that spectrum. Thus, a cell can only add spectrum or increase transmit power if other cells are willing to reduce the one or the other. Thus, adapting to load involves cooperation between network nodes.

However, solutions to the resource allocation problem are inadequate unless they are extended to incorporate the problem of modelling of behaviour at the user-level. This is a crucial, yet often overlooked factor. User applications, especially the data-hungry applications which dominate modern cellular usage, are autonomously adaptive, seeking to optimize their utility from the network by adapting to the environment that they experience. They continuously sense the ability of the network to service their requirements, and adapt their service requests to this, both at the micro and the macro level. At the macro level, they will move from heavily loaded cells to lightly loaded cells by initiating handovers. The handover triggering decision traditionally only used signal power; however, with the increased density of cells, actual network load is being increasingly used as an input. At the micro level, they implement sophisticated rate control algorithms so that the demand they generate will rise and fall based on the capacity available as measured by them individually. The mechanism for rate control is based on standard congestion control algorithms deployed in the transport layer. There is a near universal consensus, in modern congestion control theory on end-to-end congestion control algorithms similar to TCP or equivalents (Dynamic Adaptive Streaming over HTTP (DASH), TCP Friendly Rate Control (TFRC), etc.). This family of algorithms have widespread deployment in transmission stacks and have proven themselves over a long period of time extensively in a vast variety of environments [11]. While the basic algorithm is simple (computationally) to implement, its rather challenging to model; this shall be discussed further in Section IV-A.

In conclusion, the resource allocation problem can be cast as a dynamic optimization problem, but two separate issues have to be considered. The first is the problem of allocation of resources in a fair manner, which has been addressed in the literature. The second, relatively unsolved problem, is that of anticipating user behaviour, both in terms of mobility as well as in terms of demand variation in response to this allocation. In general, user behaviour is hard to model and takes time to converge to an equilibrium. On the other hand, capacity-allocation/de-allocation on the fly (by moving spectrum from one cell to the other, for example) is also complex problem with unexpected impacts network-wide. When the two occur together, they can have significant impacts on network stability and user perceived QoS. In this paper, the two variables, demand (load) and capacity are mutually interdependent. Only by considering the ones impact on the other and vice-versa, can a stable operating equilibrium be achieved and an optimal control law be derived. The need, then, is for a joint model of capacity allocation and demand regulation, which shall be described in subsequent sections.

III. Previous Work

The existing literature in multi-cellular resource management in wireless networks is focussed around the problem of resource allocation and load balancing. A relatively recent survey of the problem and analysis of the current status and open areas is given by Andrews et al. [12]. In this work, the authors also discuss the myths surrounding cell loading and QoS. One of the myths identified by the authors is that the capacity of a cell is rarely a property of the link SIR, but also has to take into account the loading of the cell itself. This underlies the need to do real-time resource planning as a network management strategy.

A. Survey of existing research in multi-cellular resource management

The approaches to multi-cellular resource management, as found in the existing literature, can be divided broadly into two categories. One set of research tends to focus on user distribution, using intra-cellular and inter-cellular handovers [13][14]. While handover optimization has been an area of study for a long time, the use of handovers as a strategic tool for resource optimization is somewhat more niche area. In these papers, the handover decisions are typically taken at the endpoints with the network nodes providing information about current loading. In [15], the authors provide a complete mathematical framework for this kind of re-direction, integrating both the equilibrium loading as well as the resources required for re-direction in the general analysis.

The second category focusses on dynamic resource deployment between cells. In some cases, the network nodes controlling the cells operate autonomously to learn the optimal loading limit individually and then act to achieve this, without needing any active inter-node coordination. For example, in [16], the authors propose reinforcement learning techniques for network nodes to tune specific configuration parameters to achieve the optimal load. The cooperative approach using explicit coordination between cells is discussed in a series of papers by Bigham and Lin [10],[17,[18], where coverage is used as a metric for load and power is the resource variable to be optimized. The authors formalize a method of structured direct negotiation between a network node and its neighbours. The decision function optimizes the allocation of coverage to individual network nodes by jointly selecting the appropriate transmission signal power for each network. The coverage area, broken into tiles is treated as the resource. The equivalent, but more modern challenge of coverage optimization for the multi-antenna case is treated in the work of [19],[20]. Other approaches to the dynamic resource planning problem involve migration of spectrum [21]. This work is interesting because it allows both hard and soft channel blocking strategies (channel blocking with and without locking). The approach is extended in [22] by incorporating support for variable demand multi-media traffic. In this paper, the cells with multimedia traffic are marked in terms of their potential peak traffic, not just the current demand. By marking a cell in this state, it is taken out of the borrowing/lending pool of cells, since this may cause thrashing between peak and safe states. Finally, there is a fair amount of literature, where resource management is not cooperative, but adversarial. For example, a game theoretic approach is found in papers such as [23]. Here the authors model the negotiation process as a game between an individual loaded cell and underloaded neighbour cells, with each cell autonomously trying to maximize its own utility.

B. Contribution of this paper

Optimal resource planning in cellular networks is about matching capacity to demand. Looking at a cluster of cells, which are under a common optimization framework, it is desirable to find those cells which have surplus capacity and match them against cells which have excess demand. Algorithms as available in the existing literature typically tend to focus on optimization of this demand-capacity gap and balancing it, either by moving demand to where capacity is available (handover) or by moving capacity to where demand exists (spectrum/power redeployment).

As discussed earlier, our focus, in this paper is on a different problem; that of demand regulation at the cell-level and how it interacts with optimal resource planning at the cluster or network level. The motivation for this arises from the fact that the demand for resources in a cell is not merely a function of the coverage or the number of UEs (above a certain limit), but the availability of capacity in the cell. In other words, while capacity follows demand, demand adapts to existing capacity (see the simulation results in Figure 3). Hence, the need for a closed-form model which incorporates the dynamics of both.

To this end, a demand-regulation algorithm has been proposed that operates at individual cells, yet takes into account the availability of capacity within and without the cell, at the cluster and possibly the network level. The algorithm has two simultaneous purposes. One is to provide feedback to the users within the cell optimally and the other is to model the true demand of each cell and incorporate it into higher level resource optimization algorithms. It has been then shown how this demand-regulation algorithm can fit into existing models of resource allocation to provide a seamless whole which manages a cluster of cells at one time.

The problem of demand regulation vis-a-vis capacity limitations has been studied for many years in the context of the Internet and there are well- understood models of rate adaptation, which have been used for many years in the wired world. There are also existing models for modelling the aggregate behaviour of bandwidth adaptive TCP based applications as a stochastic process, which has been used in this paper. The key novelty in our paper is the incorporation of network-wide capacity and adaptive demand as part of a single demand management algorithm, and the solution of the resultant optimal control problem as a stochastic mean-field game. By using the outcome to drive both user-level rate-adaptation as well as macro resource management, it will be shown that it is possible to deliver stable, controllable capacity levels in a multi-cellular cluster which adapts automatically to the available capacity. Further, this algorithm works smoothly with macro-resource optimization and UE initiated optimal network attachment strategies (such as handovers) to smoothen demand and allow load-balancing over an ensemble of cells.

A key term that that shall be using repeatedly through this paper is congestion. Congestion, in our approach is a key metric of the twin problems of demand management and resource allocation problem. It is equally applicable to the end-point users as a metric of the ability of the network to respond to variations in demand, as well as external network management or resource allocation functions, as a metric of utilization of resources. In other words, its scope is both local (within the cell) and global (across a cluster of cells). Our demand management algorithm provides a way to compute the congestion level which is both globally and locally applicable.

A model for solving the multi-cell resource management as a mean-field game has been described in a previous paper [1]. Using a simple linear model for demand and capacity, this previous work has shown how cells could choose a stable operating point for resource utilization, both at the network level as well as the cell-level, which would take into account both the variations in resource allocation, as well as the variations in demand. In this paper, the existing approach is extended by providing a practical usable close-form solution to the coordinated resource management problem. The solution moves from the simplified static demand models to widely used models for dynamic data traffic, which dominates wireless traffic today. To this end, the bandwidth adaptive endpoints have been modelled as demand generating agents using TCP-like bandwidth hunting algorithms and the solution framework has been reformulated as a mean-field stochastic optimization problem. The mean-field game solved in this paper is not of the standard linear quadratic form; to our knowledge, this is the first paper to provide an analytical solution for a non-linear mean-field game in the available literature.

IV. Modeling of user level demand

Modern communication networks are dominated by data-traffic. Analysis depends on an understanding of how data connections behave in a dynamic environment. Our task is simplified by the fact that the majority of modern data-based applications use the Internet Transport Control Protocol (TCP) as the backbone transport protocol. This is true for Internet browsing, as well as video streaming using Dynamic Adaptive Streaming over HTTP (DASH). To accurately model the traffic patterns seen in a modern wireless system, one has to start by understanding how TCP protocol stacks work. It is well known that TCP connections are explicitly designed to be simultaneously bandwidth hunting and bandwidth-conservative. A TCP data source constantly increases demand as its current demand is met and it senses that there is surplus bandwidth in the network. On the other hand, if it detects a lack of bandwidth, it reduces its bandwidth demand aggressively. The combined behaviour is approximated by Additive Increase and Multiplicative Decrease (AIMD). The multiplicative decrease ensures that TCP behaviour is cooperative. The success of the TCP bandwidth hunting algorithm is such that even non TCP connections are nowadays required to maintain TCP like transmission rate management protocols. For example, the TCP Friendly Rate Control [24] is now an Internet standard for bandwidth control of media flows such as those proposed in Web real-time communication (WebRTC).

A. Modelling TCP dynamics at the cellular level

In order to model the load balancing problem mathematically, there is a need to select a suitable model for TCP connection dynamics, which incorporates both the traffic model, as well as its reaction to network feedback. It would be preferable to have a model which can aggregate a number of TCP flows, approximating the aggregate demand seen by a single cell. After the landmark work done by Paxson et al [25], there has been a large amount of interest in the modeling of the kind offeedback-sensitive traffic seen in TCP and there is now a large corpus of work available for modelling both TCP [26],[27],[28],[29] and TCP like traffic [30],[31]. These models are complicated by the asymmetric nature of the AIMD algorithm and are much too complex for us to use directly. Further, most of these models are time-lag systems, because individual TCPs are crucially dependent on the round-trip time. Analysis requires transformation to the Laplace domain, and is highly complex.

However, if one switches to modelling aggregate behaviour, then the job becomes easier. In the modelling of the collective behaviour of a number of independent TCP connections, the discontinuities tend to smoothen out; further, the effect of round-trip time in the aggregate can be abstracted out. These are known as mul-TCP models [32],[33],[34], which aim at modeling bulk-traffic reacting independently to congestion signals.

1) Kelly's model for aggregate TCP flows: The rest of the paper shall use the model proposed by Kelly in his landmark work [32]. The Kelly model is simple enough to be used in analysis and has an intuitive structure. What it proposes is simply this. At any given point of time t, the demand is measured as an aggregate variable X_t, which reacts to a congestion state U_t of the network (which is treated as the control variable). At any instance of time, a fraction U_t of the members of the ensemble of users in the cell are given a congestion signal; these react by multiplicatively decreasing their transmission rate. The rest of the users, who do not receive a congestion signal, increase their transmission rate additively, by a single segment. If the total number of users is W, the resultant dynamics is given by (1).

$$\begin{aligned} d\tilde{X}^i_t &= \left(W^i * (1 - U^i_t) - \tilde{X}^i * U^i_t\right) dt + \sigma^i dB_t \\ &= \left(W^i - U^i_t(W^i + \tilde{X}^i)\right) dt + \sigma^i dB_t \qquad (1) \end{aligned}$$

It is to be noted that W^i is also a cell specific term (because each cell may have a different number of active users) and captures the elasticity of demand within the cell. $\tilde{X}^i$ can be replaced by $X^i = W^i + \tilde{X}^i$ to simplify the notation above.

$$dX^i_t = \left(W^i - U^i_t X^i_t\right) dt + \sigma dB_t \qquad (2)$$

The model in (2) is simple, yet rich with possibilities. Readers would notice that it matches the form of the classical Ornstein-Uhlenbeck diffusion process (3). A physical interpretation of Ornstein-Uhlenbeck diffusion is that of a spring with spring constant k and damping coefficient γ, where $\theta = k/\gamma$ and mean-position μ, starting from a rest position and moving under the effect of thermal fluctuations. In our particular case, one can see that due to U_t being an externally computed optimization variable, it is effectively manipulating both the mean rest position μ as well as the spring constant k. It can be argueid that this is because of the difference in the feedback between the two cases. In the case of a Hookean spring, the feedback for a given value of X_t is necessarily linear to $X_t - \mu$. In our case, the feedback will be a function, but it may not necessarily be linear.

$$dX_t = \theta\left(\mu - X_t\right) dt + \sigma^2 dB_t, X_t|_{t=0} = x_0 \qquad (3)$$

A standard Ornstein-Uhlenbeck diffusion is mean-reverting; in the long term, the value of X_t converges towards μ. Further, its probability distribution converges to a Gaussian distribution. These are very useful properties, because they show the way for us to stabilize our particular model. If it can be demonstrated that our model ((2)) behaves approximately like a Ornstein-Uhlenbeck diffusion in a suitably chosen domain of X_t, it is reasonable to assume that the convergence properties of the classical Ornstein-Uhlenbeck diffusion model should hold, intuitively. As it happens, this turns out to be correct in this case.

B. Congestion signaling and feedback in the context of TCP

The TCP protocol has a rate-adaptation mechanism which operates on the basis of a built-in mechanisms of measuring congestion. The primary among these are packet drops and variations in round trip time for acknowledgements from the receiver, which feed into its sliding window mechanism for controlling transmission rate. These mechanisms were introduced in TCP Reno and are universally acknowledged as having successfully solved the network congestion problem. The algorithms have been fine-tuned over the years and more modern versions of TCP (such as Vegas and Westwood) use more sophisticated functions of the round-trip time and other indicators to augment the basic feedback. However, selective packet dropping using dynamic buffer management has been the tool of choice for congestion signaling in the wired internet. Routers and switches in the WAN routinely use probabilistic packet dropping as a way of controlling end-point traffic; for example, Random Early Dropping (RED) is an Internet standard which selectively drops packets at a given target drop rate in order to force end-users to use rate-control on the incoming interfaces to stay below the capacity of its outgoing links.

In the wireless world, queue management is typically done in the core network, but it is directly effected by the bandwidth deployed in a cell and the manner of allocation. If there is a large amount of destination traffic d_t^i for a given cell, whereas the bandwidth deployed results in a capacity c_t^i, over-time the queue builds up as the accumulated difference of the two $q_t^i = \int_{t_0}^{t_1} \left(d_t^i - c_t^i\right) dt$. As the queue builds up, packet dropping takes place (RED packet dropping is assumed), which in turn acts as a congestion signal for the endpoint. Packet dropping also happens naturally in wireless communications, due to link errors. However, with modern encoding techniques such as Low Density Parity Check (LDPC) and Turbo, the incidence of packet dropping per flow is typically of the order of 10^{-5} or less and has very little impact on the TCP throughput.

In our particular situation, the congestion is dynamically controlled by introducing artificial packet drops at a computed rate $p(t)$. By doing so, the feedback to the end-point rate-adaptation protocols is controlled and hence, the demand at a cellular level. The feedback has to be continuous, not a discrete jump from congestion to no congestion. To do this, the base-station continuously tracks the expressed demand and sets the target congestion rate to the adaptive buffer management system, as the solution of our optimal control algorithm requires. This relation between demand, capacity and congestion is given by the concept of effective bandwidth [35],[36], [37]. The effective bandwidth of a channel of a given buffering capacity B in the face of a variable traffic source $X(t)$ is the packet clearing rate r required to limit the probability of buffer overflow to some value ϵ. This can be expressed functionally as in (4).

$$\begin{aligned} Q(t) &= \int_0^{\tau} (X(t) - c)^+ dt \\ \epsilon &= \Pr\left[Q(t) > B\right] \forall 0 \leq t \leq T \end{aligned} \tag{4}$$

A simple application of effective bandwidth in our case may work as follows. There are a number of agents N, a clearing capacity C and a buffer of size B (all units are in segment sizes). At a given point t, there is a feedback of p broadcast to the agents. Each agent takes a random decision whether to transmit or not, based on p. The number of outcomes in which the transmissions exceeds C at any given point of time is given by $\mathcal{X}_p = \sum_{k=C+B+1}^{N} N_{C_{N-k}} (1 - p)^k p^{N-k}$. This is precisely the tail probability of a binomial distribution. By the Chernoff-Hoeffding inequality, it can be that shown that the probability of $\mathcal{X}$ exceeding the buffer size can be approximated as in (5).

$$\Pr\left\{\mathcal{X}_p > (1+\delta)C\right\} \approx \left(1 - \exp\left\{-\frac{C(1+\delta^2)}{3}\right\}\right) \tag{5}$$

If the required clearing rate is set to C and the buffer to be δC, then the achieved drop rate is given by the equation above. In other words, by limiting the clearing rate and the drop rate, the system can drop packets at the appropriate rate. This gives us a simple method to relate the allocated capacity C and the congestion feedback p.

Further, this formula also gives us a way to compute the surplus capacity in a cell and associate it with the target congestion rate p. The surplus capacity is nothing but the difference between the allocated bandwidth C_{alloc}^k in each kth cell and the equilibrium demand rate X^k. This can be used as a way to redistribute resources. This shall be explored in the simulations below.

In real life, the TCP model is more complex than this, because a TCP endpoint will emit a number of packets depending on the current state of its window; an exact Chernoff bound for multiple TCP endpoints is hard to compute. However, it is strictly speaking not necessary; rather, the buffer management system can simply directly use RED to achieve the target drop rate, as long as it maintains a clearing rate greater than the anticipated demand rate X_t^k. The difference between X_t^k (or equivalent bandwidth thereof) versus the capacity actually available to the kth cell C_k is the surplus capacity and is what is available for redistribution to other cells.

V. Stochastic Optimal Control and Mean Field Games

This section covers the basics of mean-field stochastic optimal control. For the rest of this paper the following mathematical conventions are followed. Variables are in uppercase X_t and functions in lower case $b(X_t)$. The subscript B_t indicates a variable changing with time. Functions are Lipschitz continuous and adapted wherever applicable. $\mathbb{E}()$ represents the expectation function.

A. Fundamentals

The basic stochastic optimal control problem is as follows. Consider a system described by a state variable X_t, which is controlled by the transition function (6) given below

$$dX_t^k = b(X_t^k, U_t)dt + \sigma(X_t^k, U_t)dB_t^k \tag{6}$$

The variable $U_t = u(X_t^k)$ is the output of a control function $u()$, which only depends on the state variable X_t. As is standard for stochastic calculus, it is assumed that all processes are adapted adapted to the filtration generated by the stochastic process B_t^k, which is a Brownian motion. Further, $b()$ and $\sigma()$ are Lipschitz continuous bounded functions as required for the standard definition of a Wiener process. The system governed by this equation has a long term cost function as in (7), and the initial value of X_t is known.

$$\Phi\left(x^0, U_t, T\right) = \mathbb{E}\left[g(X_T) + \int_0^T f\left(X_t^k, U_t\right) dt\right]$$
$$X(0)= x^0 \tag{7}$$

The aim is to find the optimal control function $u^*(t)$ from a set of possible control functions $\mathfrak{u}$ of Lipschitz continuous, bounded and adapted functions, so as to minimize the expected minimum total cost $\Phi(x0, a, T)$, over the time period $[0, T]$, including the termination cost $g(X_t)$.

The general solution techniques are derived from the corresponding deterministic optimal control problem. There is, however, one important difference which increases the complexity of the problem. In a deterministic control problem, the state variable corresponding to each choice of $u()$ can be forecast. In a stochastic control problem, one is faced with uncertainty in the future. At each point t, the value of $u(t)$ has to be based on the information as known up to then, i.e., $X_s \ \forall 0 \leq s \leq t$. This is known as the filtration of the process variable, X_t. A solution of the form $u(X_t) = (1/2)\left(X_t + X_{T-t}\right)$, for example, is not acceptable, because X_{T-t} cannot be forecast at time t. In a deterministic setting, on the other hand, this would be perfectly acceptable. In other words, in a stochastic setting $u()$ has to be non-anticipatory.

There are two techniques which are used to solve stochastic optimal control problems, both of which have analogues from the world of deterministic optimal control. The first is the Hamilton Jacobi Bellman formulation, briefly described in Section V-C1, which extends the equivalent Bellman Ford technique of optimal control. The second is the adjoint equation technique, described in Section V-C2, which extends the Pontryagin Minimum Principle for the stochastic case.

B. Adding the Meanfield Constraint

The stochastic control problem is now extended by adding the mean-field constraint. This extends a one-off optimization problem to a multi-agent optimization which involves N agents, (where N is fairly large) trying to solve the same stochastic control problem in parallel. Each agent starts from a different starting value $x_0^i,\ 1 \leq i \leq N$. The N optimization problem becomes a game when their reward and cost functions are incorporate a common mean-field term, which is a function of the empirical distribution of the state-variable $X_t^k,\ 1 \leq k \leq N$ at each point t. Hence, the optimization problems are entangled with each other. Consequently, for each separate optimization problem, the equations (6), (7) change to the form given in (8). The term $\widehat{X_t} = h(\mu_t^X)$ is the mean-field term, where μ_t^X is the empirical distribution of X_t over all N agents participating in the game. In the simplest case, $\widehat{X_t}$ is simply the average value of X_t. However, more complex functions are also possible. In general, any integrable expression of the type $\int_0^T f(x)\mu_t^X$ is admissible.

$$dX_t^k = b(X_t^k, U_t, \widehat{X_t^k})dt + \sigma(X_t^k, U_t, \widehat{X_t})dB_t^k$$
$$\Phi\left(x^0, U\right) = E\{g\left(X_T, \widehat{X_t}\right) + \int_0^T f\left(X_t^k, U_t, \widehat{X_t}\right) dt\}$$
$$\mu_t^X(Y) = \frac{1}{N}\sum_{j=1}^N \mathcal{I}_{X_j=Y}$$
$$\widehat{X_{t}}_t = \int_0^T h(x) d\mu_t^X \tag{8}$$

Since the actions of each agent in the game impacts the others, solving the game requires taking into account the global ensemble of states. Thus, when computing the optimal strategy U^* for an individual agent, one has to forecast how this will affect the empirical distribution μ_t^X of the individual state variables $X_t^k, 1 \leq k \leq N$.

1) Solvability of a mean-field game: Before the actual solution technique, the solvability of the problem given in (8) has to be established. There are two specific considerations, each of which have been addressed in the literature. The first concerns the existence of a solution. The second is tied to its robustness.

In order to demonstrate the existence of a solution, it has to be shown that the equilibrium mean-field term μ_t^X and the optimal strategy are consistent with each and self-reinforcing. As per the expression in (8) the empirical distribution μ_t^X directly affects the cost $Phi()$ and hence the outcome of the transition function $b()$. In turn, because all the N agents all independently execute the same control function, the entire ensemble of X_t^k and consequently their empirical mean $\widehat{X_t}$ are driven by the choice of $u^*(X_t^k, \mu_t^X)$. In other words, there is a direct relationship between the choice of $u^*()$ and the consequent μ_t^X. Equilibrium is established when the partial derivatives of each respect to the other is zero. Thus, our optimal solution must anticipate the evolution of the mean-field distribution μ_t^X itself as a function of t. Ideally, a subspace of $\mathfrak{U}_{mfg} \subset \mathfrak{U}$ of the domain $\mathfrak{U}$ of possible optimal control functions can be found, such that a choice of optimal control strategy $u_t^*(X_t) \in \mathfrak{U}_{mfg}$ will drive the empirical distribution μ_t^X in such a way that $\widehat{X_t}$ stabilizes to an independent variable. At this point the mean-field optimization problem reverts to a standard stochastic optimization problem with an added variable $\widehat{X_t}$.

Typically, convergence in optimal control is solved by the variational inequality approach. An optimal control function u_t^* is a stable equilibrium if it can be shown that a small perturbation ϵ to u_t^* will lead to a linear degradation in the Hamiltonian of the cost function $\Phi()$ proportional to ϵ. The problem here is that this means differentiating the Hamiltonian with respect to the empirical distribution μ_t^X. The mathematical foundation of this has been discussed by Lasry [5], using the Wasserstein space of probablity measures. By using a suitably defined lifting function $\hat{\mu}()$ which can replace the term μ_t^X in the Hamiltonian, a derivative with respect to u_t is possible.

The second associated problem is that of robustness. Recall that the underlying assumption is that all N agents are executing the same optimal control function u_t^*. But this assumption only holds if the optimal solution itself is deviation proof. In other words, no individual agent can get better results by executing a separate strategy at any time-period $0 \leq t \leq T$. This is the Nash equilibrium or Nash Certainty Equivalence principle described in [6]. The authors demonstrate that NCE solutions are possible in mean-field games, which can be solved by taking the limiting value of N ($N \to \infty$) and converting the empirical distribution of the ensemble to the expected distribution of X_t^k for each k as the stochastic game evolves. It shall be shown in the case of the solution technique in the HJB-KFP approach Section V-C1 below. The NCE has very interesting properties. For example, the robustness property holds even in the case where the number of agents are small as long as the agents cannot track the individual states of other agents. In this case, it has been shown that it is the optimal strategy for each agent to follow the Nash equilibrium strategy, because deviation is punished as long as the others are following the same strategy.

C. Solution techniques for the Stochastic Mean-field Game - overview and comparision

Due to the above issues, solutions for stochastic optimal control problems with mean-field games are more complex than the standard stochastic control problem. There are two main techniques, both of which depend on solving Forward Backward Stochastic Differential Equations (FBSDE). To date, most of the research in solutions of MFGs pertain to a special class of MFGs, the so-called Linear Quadratic MFG [38][39][40]. There two main approaches that shall be discussed below, which have been studied mostly in the context of LQMFGs. Recently, a paper has been published by Pham and Wei [41], which discusses a dynamic programming solution to these games. However, it is not covered here.

Of the two widely used methods for solving general mean-field games, this paper shall focus more on the adjoint approach. A brief description of the HJB technique in Section V-C1 is provided for completeness. In both cases, the solution is in the form of an FBSDE. This is due to the aforementioned essential difference between the deterministic and stochastic problems, that a deterministic differential equation is time-reversible, but a stochastic one is not. For a deterministic differential equation, the backward equation can be re-cast as a forward equation, simply changing the sign of the variable. The resultant solution holds true for either case. However, the stochastic optimal control law cannot be anticipatory, i.e., it can only use the information about X_t upto the time t and no further [42]. Hence the forward and backward versions of the same stochastic differential equation may have different solutions.

1) The HJB-KFP approach: The classic way to solve a stochastic optimal control problem is to construct the Hamilton Jacobi Bellman (HJB) equation, which, for the above problem is given in (9).

$$
\begin{aligned}
&\frac{\partial \phi}{\partial s}(y,u) + b(y,u)\nabla_x \phi + \frac{\sigma^2(y,u)}{2}\nabla_x^2 \phi + f(y,u) = 0 \\
&\Rightarrow \frac{\partial \phi}{\partial s}(y,u) + \mathcal{H}(b, \nabla_x \phi, f, u) + \frac{\sigma^2}{2}(y,u)\nabla_x^2 \phi = 0 \\
&\phi(Y) = g(Y)
\end{aligned} \tag{9}
$$

The value of $u = u^*$ which solves this equation for all y gives the optimal value of u. Note the second derivative term, which makes the solution rather complex. The function $\mathcal{H}(y,b,f,x,u) = \langle y(x,u), b(x,u)\rangle + f(x,u)$ is called the Hamiltonian; the solution to the above equation depends, to a very large extent, on the structure of the Hamiltonian.

To extend the HJB technique for the stochastic mean-field case, one makes the fundamental assumption that as the number of agents becomes large, the empirical distribution for the states of the individual agents approaches the probability distribution for the state of each of the individual agents, i.e., instead of taking an empirical distribution over N players, the distribution of X_t^k can be modelled directly, based on the known starting distribution of the agents. In [6], Huang et al. have shown that this assumption leads to a Nash equilibrium. The probability distribution of the state variable X_t for an individual agent evolves according to the Kolmogorov Backward equation (sometimes called the Kolmogorov Fokker Planck or KFP equation). By combining this together with the HJB equation, as shown below (10), together with the probability distribution of the starting state, the optimal control u^* can be derived so as to drive the two equations into a mutual equilibrium. In a stable equilibrium, the long term probability distribution of X_t^k under the Fokker Plank equilibrium should match the empirical distribution of X_T^k as $T \to \infty$, leading to a stable solution for the HJB equation and thereby making the equilibrium self-sustaining. In this situation, it is postulated that $X_T^k \to \widehat{X_t}$ as the distribution evolves, for large values of T.

$$
\begin{aligned}
&d\phi_t = \frac{\sigma^2}{2}\nabla^2\phi + \mathcal{H}(\nabla_x\phi, b, f, X_t^k, u_t, \widehat{X_t}) \\
&d\widehat{X_t} = -b()\nabla_x\widehat{X_t} + \frac{1}{2}\sigma^2(x)\nabla^2\widehat{X_t},\ \widehat{X_t}|_0 = \mathbb{E}\left[X_0^k\right] \\
&\phi_T = g(X_T) \\
&\widehat{X_t} = \mathbb{E}\left[X_t^k\right],\ X_0^k = x_0
\end{aligned} \tag{10}
$$

The HJB equation is a backward stochastic differential equation, whereas the KFP is a forward equation. Once again, the value of u which solves both equations simultaneously is the optimal control function. The KFP-HJB technique has been used successfully for LQMFGs in many papers; a good example is that of Bardi [38].

2) The Adjoint Equation Approach: A second method is the direct analogue of the Stochastic Maximum Principle [43], which in turn is conceptually similar to the Pontryagin maximum principle for solving deterministic optimal control problems. The approach requires us to convert the constrained optimization problem into a generic optimization problem involving the state variable and its dual. In the SMP case, there are two duals, one for the drift term and the other for the diffusion term. Hence, two stochastic variables Y_t, Z_t have to be found such that the equation pair (11) holds.

$$\begin{aligned} - dY_t &= \nabla_x \mathcal{H}\left(Y_t, Z_t, b, f, X_t^k, u\right) dt + Z_t dB_t^k \\ Y_T &= \nabla_x g\left(X_T\right) \end{aligned} \tag{11}$$

In this equation, the Hamiltonian takes the extended form defined as in (12).

$$\begin{aligned} &\mathcal{H}(Y, Z, b, f, x, u) \\ &= \langle Y, b(x, u)\rangle + \operatorname{tr}\left\{Z^T.\sigma(u)\right\} + f(x, u) \\ &\partial_u \mathcal{H}(p, q, b, f, x, u^*) = 0 \end{aligned} \tag{12}$$

The expression for $u(X_t)$ which gives a joint solution of X_t, Y_t, Z_t (if it exists), provides an optimal control function u^*. Note, once again, that the equation involving Y_t, Z_t is a backward stochastic differential equation since only a termination value of Y_t is provided. The expression above is true for a simplified version of the SMP, where the function $\sigma()$ is independent of X_t. If $\sigma()$ is a function of X, then one needs to add a second pair of variables to take care of the additional risk of modifying the diffusion term in dX_t^k. The interested reader should consult Yong [42, Section 3.1] for more information. The rest of this article only considers problems where $\sigma()$ is independent of both x and u.

By incorporating the two additional variables and the mean-field term μ_t^X, the extended Hamiltonian is as shown in (13). It is required to extend the adjoint equation approach to the mean-field case, in a way that takes into account the evolution of the mean-field term in (11), as shown in (13). Here μ_t^X is the distribution at time t for the state variable X_t.

$$\tilde{\mathcal{H}}\left(X, y, z, \tilde{X}, u\right) = \mathcal{H}\left(X, y, z, \mu_t^X, u\right) \tag{13}$$

As X_t changes, the nature of μ_t^X also changes and this has to be taken into account in the solution to the adjoint equation. In general, the problem may not require μ_t^X directly, but typically a moment of μ_t^X. For example, many stochastic control problems deal with the average values of X_t, which can be expressed as $\mathbb{E}\left(X_t\right) = \int X_t d\mu_t^X$. The existing literature offers multiple approaches within the general Stochastic Maximum Principle framework. The first method can be used if the mean-field term can be expressed as a simple integral of the form $\int f().dX_t()$. In this case, the Hamiltonian can be differentiated directly, by writing the extended Hamiltonian $\tilde{\mathcal{H}}$ as a lifted version of the standard Hamiltonian, allowing us to take the derivative with respect to the distribution μ_t^X. For example, if the function $\mathcal{H}(x, \mu_t^X)$ involves the distribution μ_t^X in the form of $\langle m(x)\mu_t^X\rangle = \int m(x) d\mu_t^X$, then the derivative of the Hamiltonian $\mathcal{H}(x, \langle m, \mu_t^X\rangle)$ with respect to μ_t^X becomes $\partial_{\mu_t^X}\mathcal{H} = \partial_m \mathcal{H}\mu_t^X$.

The full expression for the stochastic Maximum Principle is then an FBSDE as shown in (14).

$$\begin{aligned} dX_t^k &= b(X_t^k, U_t, \hat{X}_t^k)dt + \sigma dB_t^k, X^k(0) = x_0 \\ -dY_t &= \nabla_x \mathcal{H}\left(X_t^k, U_t, Y_t, Z_t, \mu_t^X\right) dt \\ &\quad + \mathbb{E}\left[\partial_\mu \mathcal{H}(X_t^k, U_t, Y_t, Z_t, \widehat{X_t})\right] + Z_t dB_t^k \\ Y_T &= \nabla_x g\left(X_T\right) + \mathbb{E}\left[\partial_\mu g\left(X_T\right)\right] \end{aligned} \tag{14}$$

D. Analytical approach to solving the adjoint problem

In general, FBSDEs are difficult to solve analytically, though numerical solution techniques exist. However, if one is just interested in finding the optimal control term, then an analytical approach is possible. The trick to an analytical solution is to make a good guess at the domain from which a possible mean-field optimal solution U^* must come, the afore-mentioned $\mathfrak{U}_{mfg}$. Recall that there are two conditions that must be satisfied; one that U^* must stabilize the mean-field term $\widehat{X_t}$ and two, that U^* must be optimal in the Nash sense. If not, it gives one of the agents an incentive to deviate (possibly after $\widehat{X_t}$ has stabilized), which violates the Nash equilibrium requirement.

One possible approach is a variational inequality technique as suggested by Bensoussan in [39]. This approach considers the situation that all of the agents, other than the kth agent, is using the optimal strategy and only the kth agent is deviating. The solution then looks at finding a deviation proof solution. The optimal strategy, if it is deviation proof, will lead to a fixed point solution, where by $X_t = \mathbb{E}\left[X_t | g(z_t)\right] = z_t$, where $g(z_t)$ represents the moment of the distribution that is incorporated into the value function. Because the problem is linear quadratic, the transition function for $\widehat{X_t}$ can be directly derived by taking the expectation of the transition function of X_t^k and removing diffusion term. This leads to the condition for existence of a stable equilibrium which is deviation proof.

This approach is extended in [44], which is the one followed in this paper. In this, first the class of the optimal control function u_t^* is computed, by solving the stochastic control problem assuming $\widehat{X_t}$ is an independent variable. By substituting the general element of this class back into the transition function and taking expectations, the specific members of the class for which a stable equilibrium arises can be isolated. Since all members of the class are optimal in the generic sense, this gives a range in which the mean-field game has a stable solution. Hence, this solution technique may be implemented in three steps. Only the first and second steps require an analytic solution.

1) First solve the stochastic optimal control problem (11), using a deterministic variable z_t in place of the mean-field term.
2) Use the solution for the optimal strategy U_t^* and rewrite the expression (11) as the evolution of the mean-field term by taking expectations over μ_t^X on both sides. The stochastic optimal solution becomes a deterministic solution, with the diffusion term gone and can be solved using standard numerical techniques. The outcome is an expression for the mean-field term as an evolution in time, given the known optimal strategy u_t^*. The advantage of having a deterministic differential equation for the mean-field term is that it can be computed individually by each agent, as long as the starting value is known. This crucial factor shall be used in the current implementation.
3) Substitute this in the first equation and now solve simultaneously for a final solution of X_t.

The technique can be extended to cases where $\widehat{X_t}$ is an arbitrary function of X_t, as long as the expectation of the terms on the right hand side of (14) is defined and computable.

VI. Application to the demand regulation problem

The analysis can now return to the original demand-regulation model, by casting it as an MFG. The overall system model is as follows. There is a cluster of cells, which are integrated into a common resource management unit as seen in Figure 1. The initial analysis is restricted to the single-cell case. Here, each cell has a set of active, independent end-users $\mathcal{N}_k$ of dimensionality N_k. These end-users are continuously requesting the network node within the cell for bandwidth allocations, so as to make forward or reverse transmissions. In this paper, the focus is on the forward link, which represents the bottleneck in most cellular systems. The results hold with appropriate modification, for the reverse link as well. The network node controlling the cell allocates resources based on an algorithm that adjusts measures the aggregate demand and provides congestion feedback to the end-users within the associated cell. The congestion feedback is automatically used by the end-user rate control algorithms to adjust their resource requirements. The purpose of the proposed algorithm is to ensure that the aggregate demand in the cell reaches a stable equilibrium at a set point relative to the capacity available to that cell.

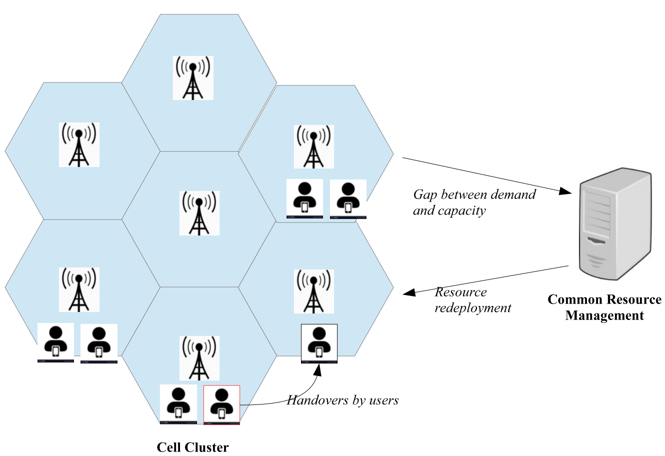


Figure 1. Cluster of cells under a common resource management framework

In the multi-cell case, external resource allocation algorithms can compare the congestion in different cells to identify over-loaded and under-loaded cells and then reconfigure the resources appropriately. Thus, if capacity is suddenly made available to one cell, the congestion metric in all cells will reflect the change and can be used to trigger handovers or resource reallocation. The core aspect of the proposed algorithm is the common function, which computes the congestion metric at the cellular level. As seen from the literature survey in Section III-A, this kind of joint optimization that takes into account the user behaviour as well as the availability of resources, is novel in the published literature.

A. Demand regulation as an optimal control problem

The aggregate demand in each cell is due to the collective dynamics of a large number of TCP connections, coexisting in a shared channel of fixed capacity. This is modelled at the cell-level by the state variable X_t^k for the kth cell. The TCP and TCP-like connections enter and leave at various times; however, since they share the same resources, their collective behaviour is affected by the overall capacity of the system. The coupling between the endpoints and the system comes through packet drops. The packet drop rate is under the control of the cell-level demand regulation mechanism.

The objective is to create an algorithm for regulating aggregate demand as the cell level by setting the optimal packet drop rate as a function of aggregate demand X_t^k and cell-capacity C_t^k. The packet dropping is a congestion signal, which acts to expose the state of network resources to the UEs. By adjusting it as needed, and knowing how the endpoints react to this signal, the aggregate demand (and consequent throughput) X_t^k is driven to its desired value. It is to be noted that packet dropping is processed individually at each endpoint. However, it is still possible to control it so as to influence the aggregate behaviour, as shall be shown.

The optimal level of X_t^k is a function of how much capacity C_t^k is available and how close the aggregate demand is allowed to come to it, i.e., the cell-level load factor. Ideally, it should be allowed to be as close to the available capacity as possible, without exceeding it, so as to keep utilization high. However, allowances have to be made for the variation in demand. The variation in demand is based on two parts. One is the natural variation (connections terminating and new ones arriving), captured through a random diffusion term. The second is the variation of demand due to packet dropping; this is captured in the model in (2). High level of packet dropping can cause significant oscillations in X_t^k due to simultaneous back-off by a number of endpoints; this is something to be avoided by appropriate algorithm design.

1) Setting an optimal reward function: For a TCP connection, the instantaneous bandwidth is not of interest. Rather, the total number of bytes transferred is what counts, which is the integral of the instantaneous bandwidth variable. However, the cost of the deployed resources has also to be incorporated. The optimization problem, as given in (15), hence becomes choosing the optimal congestion response $u(X_t)$ to the current state, so as to minimize the integrated cost function $\Phi(X_t, u(X_t))$. The cost function will have two parts, one of which handles the capacity demand gap and the second the congestion signal. It is to be noted that congestion signaling is a real cost, in the sense that it comes from the deliberate packet drops in the core. Hence, it has real consequences in terms of resource utilization. The termination cost $g(X_t)$ is of relatively little significance and can be selected so as to enable us to solve this FBSDE, while still being intuitively valid.

$$
\begin{aligned}
dX_t^i &= \left(W^i - U_t^i X_t^i\right) dt + \sigma dB_t \\
U_t^{*,i} &= \arg_{0 \le U_t \le 1} \min \Phi(x_0, U_t) \\
\Phi(x_0, U_t, r_t) &= \int_0^T \left\{\phi(X_t^i . U_t^i, r_t)\right\} dt + g(X_T^i)
\end{aligned}
\quad (15)
$$

The variable r_t is a set-point that is used as the target capacity, and as of now, it is assumed that it is an independent external variable supplied by the network management function. It has to be incorporated into $\phi()$ in such a way that cost rises with $|X_t - r_t|$. In the mean-field case, this variable will contain the mean-field term, as shall be seen in Section VI-D. The selection of the appropriate form $\Phi(x, u, r)$ is deferred, till the form of the differential equation to be solved has been established.

B. Existence of a solution - independent cells with target congestion

The analysis starts with the non mean-field case, and is the resultant solution is then extended to the mean-field case, as suggested by [45]. The initial algorithm starts with the standard form of the stochastic Hamiltonian (16). The suffix ^{i} is dropped in the rest of this section, because the entire analysis is only in the context of a single cell.

$$\mathcal{H}(x, u, y) = y.b + f = y(W - X_tU_t) + \phi(X_tU_t, r_t) \quad (16)$$

There are two conditions that have to be fulfilled by the optimal control function u. The first one, in (17), simply requires that the derivative of the Hamiltonian is zero.

$$\begin{aligned} \partial_u \mathcal{H} = 0 &\Rightarrow -Y_tX_t + X_t\partial_z\phi(z)|_{z=X_tU_t} = 0 \\ &\Rightarrow Y_t = -\partial_z\phi(z)|_{z=X_tU_t} \end{aligned} \quad (17)$$

The second one is the condition for the dual of X_t. Specifically, it dictates that the optimization problem in (15) can be solved if the variables Y_t, Z_t can be found such that (18) holds [42].

$$\begin{aligned} dY_t &= -\partial_x\mathcal{H}(X_t, U_t) + Z_tdB_t \\ &= -\left(Y_tU_t + U_t\partial_z\phi(z)|_{z=X_tU_t}\right)dt + \\ &\quad \partial_x\psi(X_t, r_t) + Z_tdB_t \\ Y_T &= \partial_x h(X_t)|_{t=T} \end{aligned} \quad (18)$$

Substituting the value of y from (18) in (18), we get (19).

$$dY_t = -\partial_x\psi(X_t, r_t)dt - Z_tdB_t \quad (19)$$

Our method takes a novel approach in the search for solutions of Y_t, U_t. Instead of solving the stochastic PDE directly, it is converted to the equivalent functions in X_t, i.e., $Y_t = y(X_t)$, $U_t = u(X_t)$, By Ito's formula, $y(X_t)$ can be differentiated directly as given in (20). The advantage of this approach is that it moves from derivatives in t to derivatives in terms of X_t and it is possible to choose appropriate forms of $y()$ and $u()$. Direct solution of the FBSDE is not required.

$$\begin{aligned} dY_t &= \partial_x y(X_t)dX_t + \frac{1}{2}\partial_x^2 y(X_t)(dX_t)^2 \\ &= \partial_x y(X_t)\left((W - X_tu(X_t))\,dt + \sigma dB_t\right) + \frac{\sigma^2}{2}\partial_x^2 y(X_t)dt \\ &= \left[\partial_x y(X_t)\left(W - X_tu(X_t)\right)) + \frac{\sigma^2}{2}\right]dt + \left(\sigma\partial_x y(X_t)\right)dB_t \end{aligned} \quad (20)$$

Comparing the terms with dB_t from (19) with the above, $Z_t = \sigma\partial_x y(X_t)$. Collecting the terms for dt from (20) and (19), we get the identity in (21).

$$\partial_x y(x)\left(W - X_tu(X_t)\right) + \frac{\sigma^2}{2}\partial_x^2 y(x) = -\partial_x\psi(x, r) \quad (21)$$

1) Appropriate solutions for Y_t: First, it is to be emphasized that the expression in (21) has no direct dependence on t. Hence, it is only necessary to solve Y_t and U_t as a function of X_t. For this purpose, the previous relation between Y_t and U_t^* can be used, as $y(x) = -\partial_z\phi(X_t, u(X_t))|_{z=X_t.u(X_t)}$ from (17). There is another consideration, however, that needs to be taken care of. The chosen solution should degenerate to the non-stochastic case as $\sigma \downarrow 0$. One way of achieving this is to ensure that $\partial_x^2 y(x)$ to be of the same form as $\partial_x y(x)\,(W - xu)$. Essentially, this means that the expression on the left hand side of (21) becomes $\left(1 - \sigma^2/2\right)\partial_x y(x)$. Obviously, if $\sigma \downarrow 0$, only the constant multiplier changes.

To provide a solution, an appropriate form for the two components of the reward function $\phi()$ and $\psi()$ is proposed as in (22). As discussed before, the first term incorporates the aggregate demand and the congestion as the product X_tU_t and the second term incorporates the gap between X_t and r_t as a cost;

$$\begin{aligned} \phi(X_tU_t) &= -\frac{1}{2}X_t^2U_t^2 \\ \Rightarrow Y_t &= -X_tU_t(X_t)((17)) \end{aligned} \quad (22)$$

How can the form of $\phi()$ be justified as a cost function as given in (22)? It can be seen that $\phi()$ decreases as X_tU_t rises. In other words, if a high value of achieved bandwidth X_t can be maintained in the face of high congestion U_t, the solution is preferable. Since reduction of the congestion metric can only be achieved by adding to the available resources within the cell/network, it makes sense to reward the combination of high congestion U_t and high bandwidth X_t. For the second reward, the corresponding form of $\psi(X_t, r_t)$ is as given in (23).

$$\begin{aligned} \psi(X_t, r_t) &= \alpha\left(\eta - e^{-\frac{X_t}{r_t}}\right)^2 \\ \partial_x\psi(X_t, r_t) &= \frac{\alpha}{r_t}e^{-\frac{X_t}{r_t}}\left(\eta - e^{-\frac{X_t}{r_t}}\right) \end{aligned} \quad (23)$$

The expression $\psi(X, r)$ is designed to penalize deviation of (X_t/r_t) from the fixed term η. X_t/r_t is the classical utilization term and $\log(1/\eta)$ becomes the target value. β is a discount factor with respect to the growth term W. W and β are taken common in the rest of the expansion, but actually can be scaled on a per-cell basis, as long as the ratio W/β is maintained. As shall be seen subsequently, β can be expressed in terms of η. η becomes the crucial operator supplied constant, which controls the demand regulation function. The higher its value, the tighter the regulation. The system response to set values of η shall be demonstrated in the Section VII when the simulation results are presented.

2) Analytic solution of the FBSDE: It is now necessary to use the forms of $\phi()$ and $\psi()$ as given in (22) and (23) to solve the stochastic PDE equation in (21). The deterministic case is solved first and then the stochastic case is solved as an extended form of the deterministic case. Hence the first step to propose a form of $y(x)$ such that $\partial_x^2 y(x)$ is of the same form as $\partial_x y(x)$ or $y(x)\partial_x y(x)$.

In this article, the latter approach is chosen; $y(x) = -\beta\left(1 - e^{-\frac{x}{r_t}}\right)$ as given in (24). Recall that r_t is an independent, but deterministic variable with known values at each time t. By the second equation in (22), $Y_t = -X_t u(X_t)$. By taking Y_t as proposed in (24), the value of $u(x) = y(x)/x$ is well defined for small values of x. In fact, as $x \to 0$, the congestion term becomes $\approx \beta/r_t$.

$$\begin{aligned} Y_t &= -\beta\left[1 - e^{-\frac{x_t}{r_t}}\right] \\ \partial_x Y_t &= -\frac{\beta}{r_t} e^{-\frac{X_t}{r_t}} \\ \partial_x^2 Y_t &= \frac{\beta}{r_t^2} e^{-\frac{X_t}{r_t}} \end{aligned} \tag{24}$$

Substituting this against the identity in (21), and taking advantage of the fact that $X_t u(X_t) = -Y_t$, the expression in (25) are derived.

$$\begin{aligned} &-\partial_x \psi(X_t, r_t) \\ &= \left[\left(W - \beta\left(1.0 - e^{-\frac{X_t}{r_t}}\right)\right)\left(-\frac{\beta}{r_t} e^{-\frac{X_t}{r_t}}\right) + \frac{\sigma^2}{2}\frac{\beta}{r_t^2} e^{-\frac{X_t}{r_t}}\right] \\ &= -\frac{\beta}{r_t} e^{-\frac{X_t}{r_t}}\left[W - \beta - \frac{\sigma^2}{2r_t} + \beta e^{-\frac{X_t}{r_t}}\right] \\ &\Rightarrow \partial_x \psi(X_t, r_t) = \frac{\beta^2}{r_t} e^{-\frac{X_t}{r_t}}\left[\left(\frac{\beta - (W - \frac{\sigma^2}{r_t})}{\beta} - e^{-\frac{X_t}{r_t}}\right)\right] \end{aligned} \tag{25}$$

By taking the corresponding form of $\psi(X_t, r_t)$ as given in (23), the relationship between the operator supplied control factor η and the corresponding values of β, α is as in (26).

$$\begin{aligned} \eta &= \frac{\left(\beta + \frac{\sigma^2}{2r_t} - W\right)}{\beta}, \\ \beta &= \frac{\left(W - \frac{\sigma^2}{2r_t}\right)}{1 - \eta},\ \alpha = \beta^2 \end{aligned} \tag{26}$$

Hence, the appropriate form of the optimal control function for the stochastic optimal problem as posed in this section, as given in (27).

$$\begin{aligned} u_t^* &= \frac{\beta}{X_t}\left[1 - e^{-\frac{x_t}{r_t}}\right] \\ \beta &= \frac{\left(W - \frac{\sigma^2}{2r_t}\right)}{1 - \eta} \end{aligned} \tag{27}$$

C. A practical justification of the proposed model

While there exists a solution to the given problem, it is still necessary to justify the solution in practical terms. Starting with the form of u_t^* as given in (27) above, the original state equation as given in (2) is examineded by taking the computed form of y_t and substituting the optimal value of $X_t u_t^*(X_t) = -Y_t$. On doing this operation, we get the transition function as given in (28).

$$dX_t^i = \left(W^i - \beta\left(1.0 - e^{-\frac{X_t^i}{r_t}}\right)\right) dt + \sigma dB_t \tag{28}$$

It is to be noted that the term $W^i - \beta(1.0 - \exp -\frac{X_t^i}{r_t})$ is just W^i for $X_t^i = 0$ and reduces gradually, till it changes sign for a particular value of X_t^i, the transition value ((29)). If the capacity available in a cell is known, it can be used in the formula in (29) to set the appropriate value of β. Even though W^i can usually not be directly controlled, it can be estimated from the number of UEs in the cell, since each UE, in the uncongested case, simply increases transmission rate by 1 segment per round trip time. Equilibrium in a cell can be maintained simply by controlling the transition point, at which the capacity change term in (28) turns negative. This in turn gives us the practical capacity limit per cell (29).

$$X_t^{c,i} = r_t \log\left(\frac{\beta}{W - \beta}\right) = r_t \log\left(\frac{1}{1 - \eta}\right) \tag{29}$$

In other words, the terms W^i and β play no role in the equilibrium; they only control the elasticity of demand as congestion changes. This is very important for the subsequent analysis, because each cell will have a different combination of users and the value of W^i will vary from cell to cell.

Let us now consider the form of the optimal congestion function $u^*(X_t^i)$ as given in (27). For small values of X_t/r_t, $u_t^* \approx \beta/r_t$ and the transition function starts behaving like a standard Ornstein-Uhlenbeck diffusion problem.

Once again, η turns out to be the crucial external parameter that controls the equilibrium operating point. From the expression of $\phi()$ and the relation between β and η in the second equation in (26), it can be seen that η is a measure of how close the current load X_t can be allowed to approach the effective capacity r_t. Adjusting this value means adjusting value of β, which in turn controls the discounting of the congestion term in (27) to the growth term W.

D. Extension to multiple cells - the mean-field problem

The control problem is extended in this section to incorporate the multi-cell case. The multi-cell case works on a cluster of cells N_c whose capacity allocations are integrated with each other, and allocations/deallocations of capacity are done for the cluster as a whole. Further, the cells in the cluster are geographically close to each other, and all users have visibility of all the cells in the cluster; this is frequently how hotspots are configured in urban areas. This assumption allows us to implement joint allocation of resources and allow movement of users between cells within the cluster. There is no restriction to the size of the cluster, or the number of cells in it, as long as the above restrictions hold.

Within the cluster, the demand regulation algorithm must ensure that the cells are utilized evenly (load balancing) and the congestion metric u_t^k gives both the UEs and the resource management algorithm a good approximation of the overall utilization within the cluster as well as the utilization within each cell. This is necessary to provide appropriate feedback both to the UEs and to the resource management operating for the cluster as a whole. The individual UEs freely move within the cluster to optimize their QoS, based on the congestion feedback. External resource management in turn uses the congestion metric to compare relative utilization of capacity and urgency of demand within the cells of the cluster.

Mathematically, the demand regulation algorithm has to be modified so that the congestion metric tracks both the cell-specific resource loading as well as the average loading within the cluster. A cell with high load, but relatively less metric is a signal to the UEs that there is surplus capacity in some other cell within the cluster. On the other hand, a cell with high load and a high metric indicates to UEs that there is no surplus capacity in the entire cluster and that UEs should pull back their demand so as to reduce loading within the immediate cell. It is proposed to make this happen by adding a cluster average term to the reward function. Recall that the consolidated stochastic control problem is given as in (30), for each jth cell in the cluster, and the corresponding optimal congestion function is given in (31).

$$\begin{aligned} &\text{Minimize } \int_0^T f(X_t^j, r_t, u_t^j)dt + g(X_T^j, r_T^j) \\ &\text{where } dX_t^j = \left(W - X_t^j u_t^j\right)dt + \sigma dB_t \\ &f(X_t^j, r_t^j) = \beta^2\left(\eta - e^{-\frac{X_t^j}{r_t^j}}\right)^2 - \frac{1}{2}(X_t^j u_t^j)^2 \qquad (30) \\ &g(X_t^j, r_t^j) = \beta^2\left(\eta - e^{-\frac{X_t^j}{r_t^j}}\right)^2 \\ &X_t|_{t=0} = x_0 \\ &u_t^{*,j} = \frac{\beta}{X_t^j}\left(1 - e^{-\frac{X_t}{r_t}}\right) \qquad (31) \end{aligned}$$

The term $\beta\left(\eta - \exp\{-\frac{X_t^j}{r_t^j}\}\right)$ penalizes the deviation of the fraction X_t^j/r_t^j from the target value η, where r_t^j denotes a kind of threshold level of resource usage within the cell. A second factor is now introduced into r_t^j, which is the ratio between the empirical average of X_t among all the cells in the target cluster $\mathcal{C}$, as shown in (32).

$$\begin{aligned} \widehat{X_t} &= \frac{1}{N_C}\sum_{k=1}^{N_C} X_t^k \\ r_t &= \frac{C_1 C_2}{\widehat{X_t}} \\ \widehat{X_t} &= \frac{1}{N}\left(\sum_{j \in \mathcal{C}} X_t^j\right) \\ \Rightarrow \frac{X_t^j}{r_t^j} &= \frac{X_t^j}{C_1}\frac{\widehat{X_t}}{C_2} \end{aligned} \qquad (32)$$

Effectively, problem has been converted in to a mean-field optimization problem, with the mean-field term being the

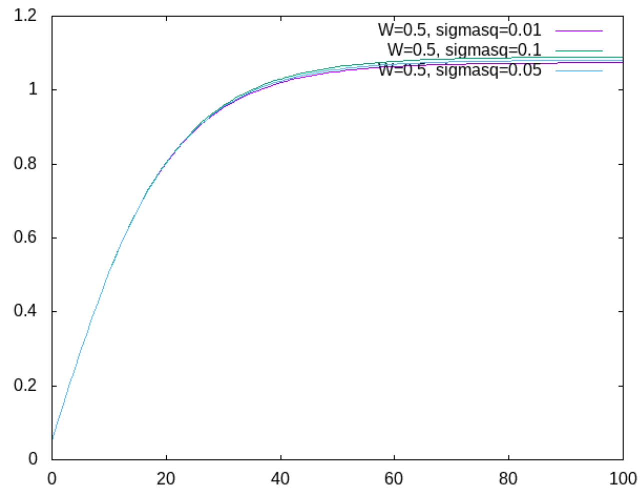


Figure 2. Evolution of the empirical mean for different values of W, η and C

empirical average $\widehat{X_t}$. For any individual jth cell, the congestion signal u_t^j is then affected by both the cell's own bandwidth level X_t^j/r_t^j and the empirical average value. Since r_t is a part of the reward function, this turns the N_c separate optimization problems (one per cell) into a game. It is clear that the game has to be solved cooperatively. For example, any individual cell raising X_t^k will cause the empirical average to rise; this, in turn may cause one or more of the other cells to cross the threshold term in the term $\psi()$, which will force it to reduce its own $X_t^j, j \neq k$.

1) Solving the mean-field demand regulation problem: The next step in the mean-field analysis is to compute the function $\widehat{X_t}\forall t$. To do this, it is necessary to solve the expectation of the original diffusion (2), substituting the optimal value of u_t as computed in (31). The corresponding deterministic differential equation is given in (2). Note that the diffusion term is no longer present; however, r_t is also no longer an independent variable.

$$\begin{aligned} d\widehat{X_t} &= \left(W - \frac{\beta}{\widehat{X_t}}\widehat{X_t}u(\widehat{X_t})\right)dt \\ &= \left(W - \frac{1}{\widehat{X_t}}\left(\frac{W}{1-\eta} - \frac{\sigma^2\widehat{X_t}^2}{2C_1C_2(1-\eta)}\right)\widehat{X_t}u_t^*(\widehat{X_t})\right)dt \\ &= W dt - \left(\frac{W}{1-\eta} - \frac{\sigma^2}{2C_tC_2(1-\eta)}\widehat{X_t}\right)\left(1 - e^{-\frac{\hat{X}_t^2}{C_1C_2}}\right)dt \end{aligned} \qquad (33)$$

The differential equation in (33) can be solved numerically. However, it is easy to see that a stationary point exists, where $d\widehat{X_t} \to 0$. For small values of σ^2, the stable point should be at $\hat{X} \approx \sqrt{(C_1.C_2.\ln(1/\eta))}$. The parameter W dictates the rate of convergence. This is demonstrated in Figure 2, which shows the evolution of $\hat{X}$ for different combinations of W, η and C. As noted earlier, the value of W has no impact on the final equilibrium achieved, only the equilibrium rate. Hence, one can also model heterogeneous populations within cells, where different cells have different numbers of UEs and further, different UEs have different traffic generating patterns facing different round trip times. At the top level, the combination of these just add up to different values of W for each cell. Since the value of W plays no part in the optimal control term given in (27), it is also not something that needs to be estimated.

In the final step, the mean-field term is inserted back into optimal transition function to get the mean-field optimal control law. The optimal control thus occurs in two steps at each time t at each cell. First the current value of $\widehat{X_t}$ is estimated, either by directly sampling it from all cells within the cluster, or by each cell computing the current value from (33). It is then inserted back into (27) so as to get the optimal congestion metric u_t for each cell. The congestion metric is used to set the buffer value based on the dynamic given in (4) and (5).

VII. Simulation Results

In this section, the performance of the mean-field demand regulation algorithm through simulation is shown. There are two primary purposes to the simulation. One is to show how the simulation stabilizes the performance of the entire system by providing appropriate feedback in each cell. The second is how the simulation allows cooperative resource re-distribution to take place, by providing the appropriate signals to the end-user and the network resource allocation function.

The mean-field based demand regulation algorithm has been simulated on a six-cell cluster, each cluster being allocated a set of UEs and a set of resources Each UE is downloading data using a simplified TCP-Reno protocol stack, with a Maximum Segment Size (MSS) of 1200 bytes and a maximum window size of 64. It is assumed that the round-trip time is approximately 50ms, which means that each UE has a saturation transmission rate of 12.3 Mb/s. The cells each have a capacity of about 50Mb/s. Congestion signalling is performed by each cell by fair dropping (as used in RED), using the dropping rate computed by the demand regulation algorithm. In the case where there is no dynamic congestion signaling, a buffer of approximately 1MB is available to each cell, subject to Random Early Dropping. The base stations/eNodeBs use a simple fair-sharing scheduling, using Scott Shenker's fair sharing protocol, with bandwidth delay trade-off. There is a baseline packet drop rate of 10^{-5}, with minor or no impact on the system.

A. Comparison against the default situation - no demand regulation

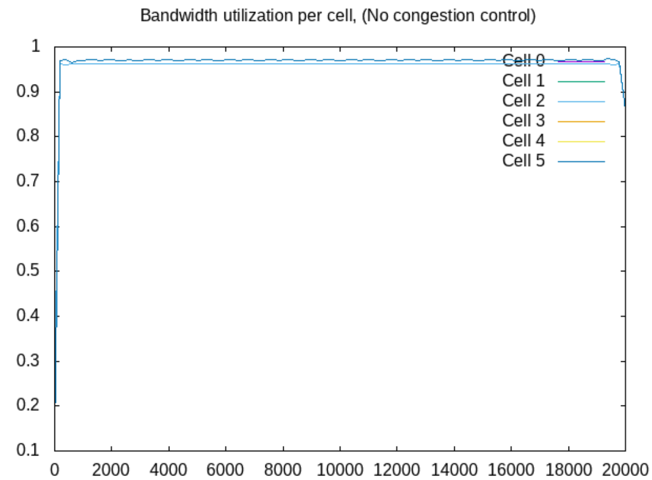
The first scenario compares the performance of the proposed algorithm against the situation where there is no demand regulation algorithm working. The first three cells start out with 24 active UEs each and the second three cells with 12 active UEs each. Each cell is allocated capacity equivalent to 100Mb/s. The comparative results for aggregate bandwidth utilization are presented in Figure 3. There are three congestion cases presented with operator supplied η value of 0.3, 0.5 and 0.6 corresponding to the graphs in Figures 3b, 3c and 3d respectively.

Two things stand out very clearly from the above. First, it is clear that bandwidth hunting TCPs operating in a cell with limited capacity will drive the bandwidth utilization to saturation, regardless of the number of users [Figure 3a]. Second, congestion feedback is an effective tool to regulate the behaviour of TCPs in this regard. Recall from the discussion in VI-B that the parameter η regulates the impact of the mean-field term on the generic control function. For small values of η, mean-field demand regulation is switched off. As η approaches 1, the effect of mean-field demand regulation starts dominating overall performance. The result can be seen in the demand regulation cases shown in the Figures 3b, 3c), 3d. As the value of η rises, sharply rising congestion feedback forces users in congested cells to back-off in order to avoid saturating the cell. Further, recall that β is approximated by $W/(1-\eta)$. Hence, for larger values of W, the demand regulation algorithm is more aggressive in communicating congestion to the end-users as is expected from the model.

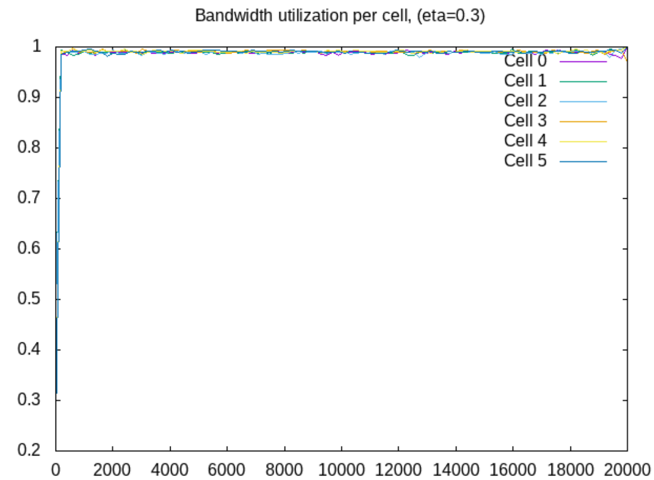
An alternative scenario is given in the second set of graphs in Figure 4, where the number of users in each cell is the same (20 users per cell), but the capacity allocation is different. The first three cells have 50Mb/s capacity each, whereas the last three cells have 36Mb/s. The unregulated demand case and the case where demand is regulated by setting η to 0.3 is presented in Figures 4a and 4b respectively. The congestion metric for the second case is presented in Figure 5. It is clearly seen how the congestion metric clearly separates out the two different cells, so that the lower capacity cells see a higher congestion metric, even though the capacity utilization value is the same. This allows the system to handover cells from lower to higher capacity cells, using the congestion metric as a guide.

B. Performance in conjunction with macro level resource management

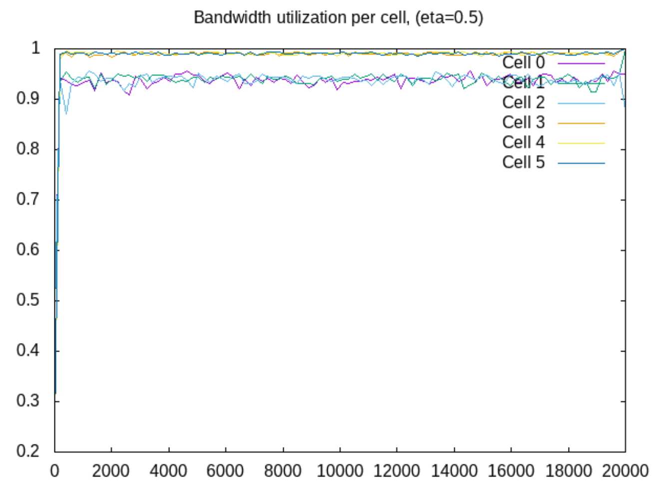
In this section, it is shown how the demand regulation algorithm works when coupled with external resource management. The simulation is extended to include both user initiated network assisted handover (based on congestion) and dynamic resource swapping between cells. The macro-resource management entity is treated as an external module, which has visibility into the basic runtime statistics of each cell; number of UEs, allocated channels and (in the cases where demand regulation mechanism is active) the computed congestion metric. The resource management as well as handover is continuously active. At each frame it selects pairs of cells and moves a unit of spectrum from one to the other with a probability proportional to the gap between the load metrics. The load metric for a cell is either the computed congestion metric, if available, or the number of UEs in the cell. The same metric is communicated to UEs that are considering handover, by network signaling. The UEs use the same probabilistic approach to choose when to handover. It may be argued that this is an overly simplistic mechanism for macro-resource management. Our position is that the specifics of the mechanism may be kept simple, because the



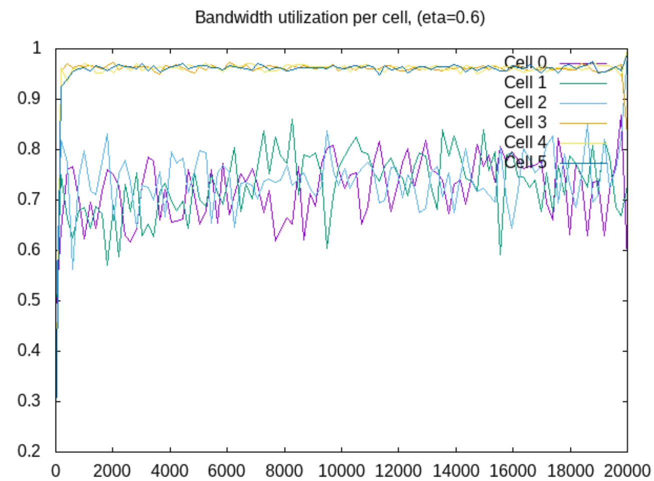
(a) No demand regulation



(b) Regulated demand, $\eta = 0.3$



(c) Regulated demand, $\eta = 0.5$



(d) Regulated demand, $\eta = 0.6$

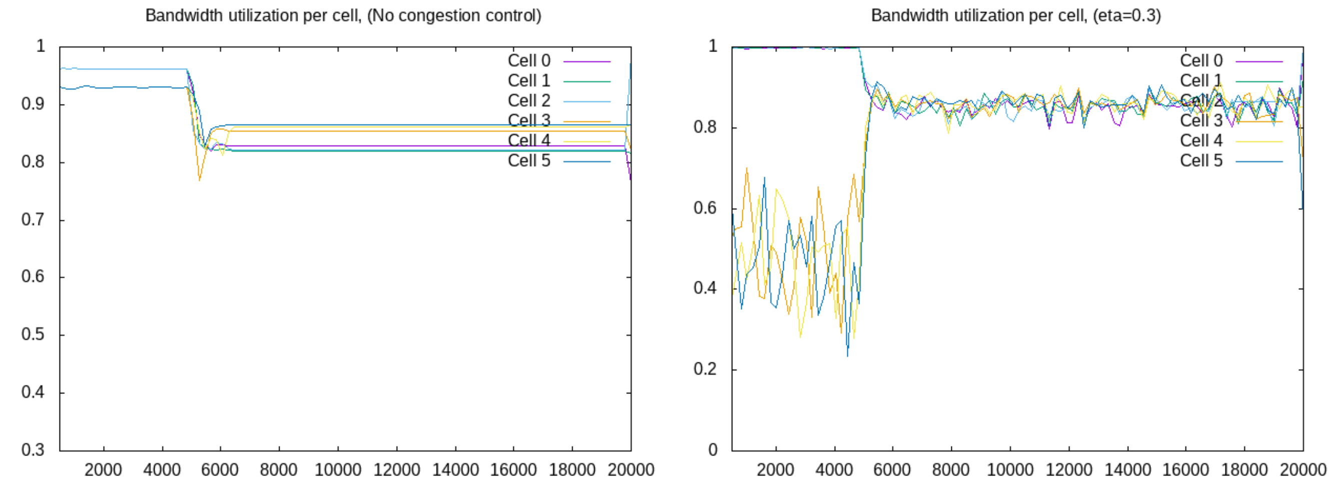
Figure 3. Simulation results - six cell cluster, comparative capacity utilization, variable users per cell, no handovers

focus is on how useful the proposed demand regulation is in working with this or any other macro-resource management algorithm.

For the simulation, the six cell scenario is retained, but now each cell start with a variable number of users and a variable amount of bandwidth per cell. To make matters interesting, the cells with the maximum capacity available have the minimum number of users to start with. The first three cells have 100 Mb/s of capacity and 10 users each. The last three cells have 50Mb/s of capacity and 20 users each. The comparison is between the default case (no demand regulation) and the cases where the proposed demand regulation algorithm is running with η values of 0.3 and 0.5. The simulation starts off with the macro level resource management disabled for the first 4000 frames and then is switched on. The transition can clearly be seen in the two figures.

Figure 6 shows the comparative bandwidth utilization between the default and the demand regulation algorithm. It can be seen that macro-resource optimization is working in both the cases; however, whereas the default case still shows different levels of utilization in the cells after rebalancing (Figure 6a), the managed demand/congestion cases provide an extremely tight balancing with respect to the default case. In both the $\eta = 0.3$ case (Figure 6b) and $\eta = 0.5$ (Figure 6c) cases, the capacity utilization stabilizes to a similar level of 90% after the rebalancing takes place.

To understand the reason for the difference, note how the congestion signal is computed in Figure 7a and Figure 7b below, corresponding to the $\eta = 0.3$ and $\eta = 0.5$ cases respectively. The default case does not use the congestion metric for feedback; rather it uses the number of UEs in each cell. In this situation, the number of UEs per cell is quickly balanced. However, each UE has its own rate-control and when one moves to another cell, it takes a long time to adjust its current state based on the new environment. Hence, even within the same cell, there is large variation in the performance between UEs. On the other hand, the congestion-metric computed by the demand regulation metric already takes network state into account. This means that



(a) No demand regulation for variable capacity cells

(b) Demand regulation for variable capacity cells, $\eta = 0.3$

Figure 4. Simulation results - six cell cluster, comparative capacity utilization with variable spectrum allocation per cell

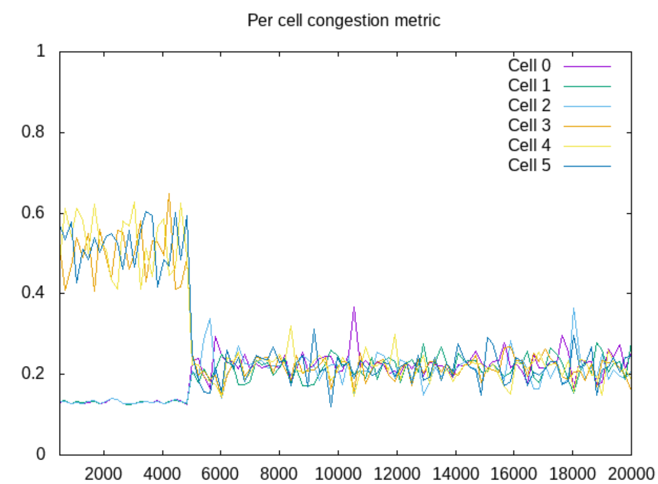


Figure 5. Congestion metric value, $\eta = 0.3$

all UEs have similar rate-control states once equilibrium is reached and they tend to have very similar performance subsequently. One of the purposes of demand regulation is provide fair allocation of capacity. The fairness of allocation is measured in terms of a metric γ that measures the spread of allocations (achieved throughput) per UE in a given cell. The metric is computed as in the equation (34) where $t(m)$ measures the throughput of the mth UE and $\mathcal{N}_k$ is the set of UEs in the kth cell.

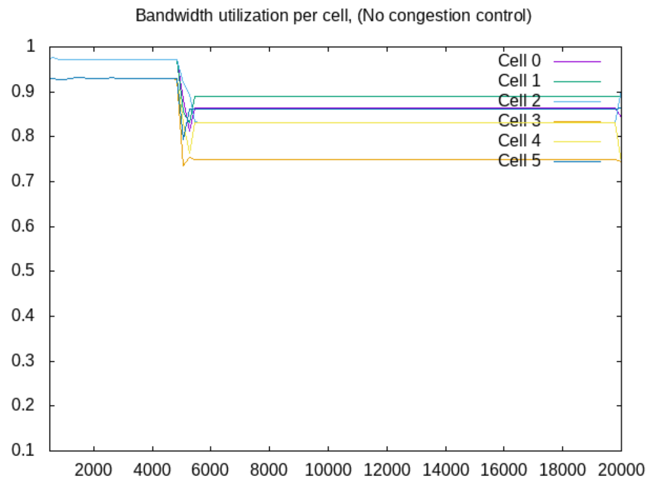
$$\gamma_k = \frac{\max_{j \in \mathcal{N}_k} t(j) - \min_{j \in \mathcal{N}_k} t(j)}{\max_{j \in \mathcal{N}_k} t(j)} \tag{34}$$

Figure 8 shows the allocation fairness for each of the simulated cases. Since fair scheduling has been implemented within each cell, the difference is purely because of the dis-balance between allocated capacity and per-cell demand. This simulation shows the difference in the starkest terms. Because the congestion metric takes into account the global state, it forces the per-UE allocation into a tight band, with complete fairness between cells.

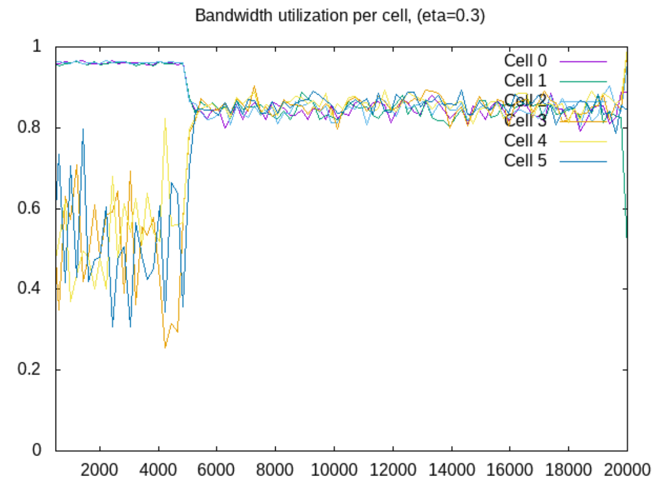
VIII. Conclusion and Future Work

In this paper, we have demonstrated the application of mean-field stochastic optimal control to the problem of optimal resource management in a wireless cellular network, one of the most demanding problems of wireless network control. We have demonstrated that we can apply mean-field control to regulate demand within individual cells, in keeping with cell-specific and network wide capacity levels. Consequently we have demonstrated that an integrated load-balancing and resource deployment solution can be achieved, with significantly superior performance, as opposed to the standard case where demand is not regulated and resource allocation uses external metrics such as the number of UEs in a cell. As part of the theoretical framework, we have also derived a closed form analytical solution for a non-linear mean-field model, which is novel in the MFG literature. As we have shown, this algorithm is implementable by incorporating it directly the dynamic buffer management at the cellular and system level. This is sufficient to provide the appropriate feedback to the end-users. We also show how the same dynamic buffer management algorithm works with macro-resource optimization algorithms to achieve stability and fair load-balancing, across multiple cells. The application of stochastic control techniques, and stochastic games in general, has rich potential application in the wireless and cellular domain.

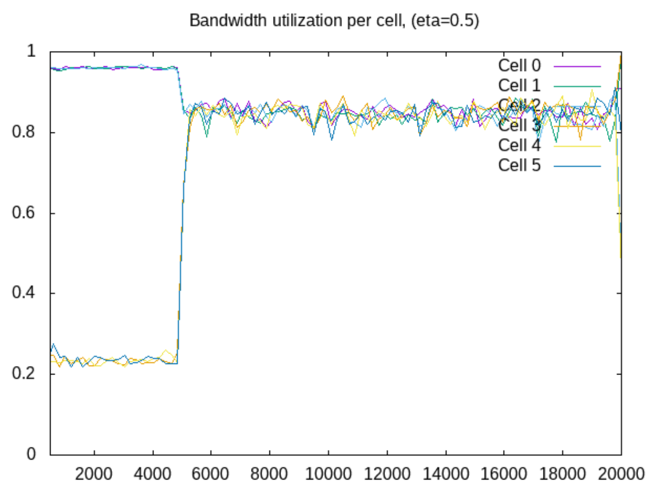
In future work, this shall be extended to more advanced problems, such as coordinated multi-point, multi-user Multi-Input Multiple Output (MIMO) and dynamic network slicing for 5th generation networks. Here, a key extension will be to track the UE when it is a member of more than one cell at any point of time. In this situation, the mean-field control will apply not just to the cell, but to the UE as well. A second possible area of work is in application to



(a) No demand regulation



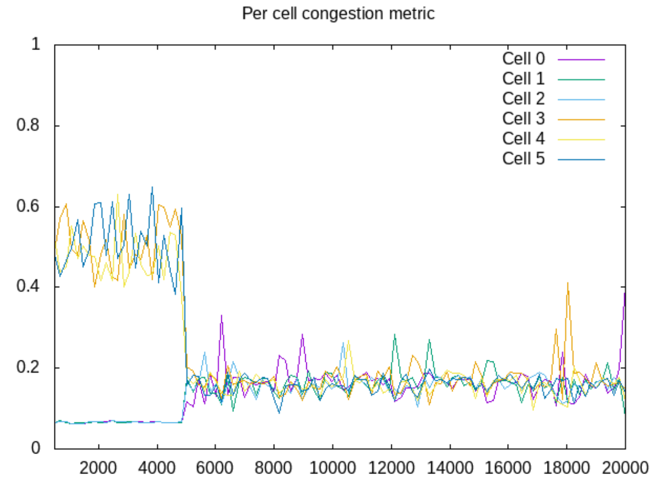
(b) Regulated demand, $\eta = 0.3$



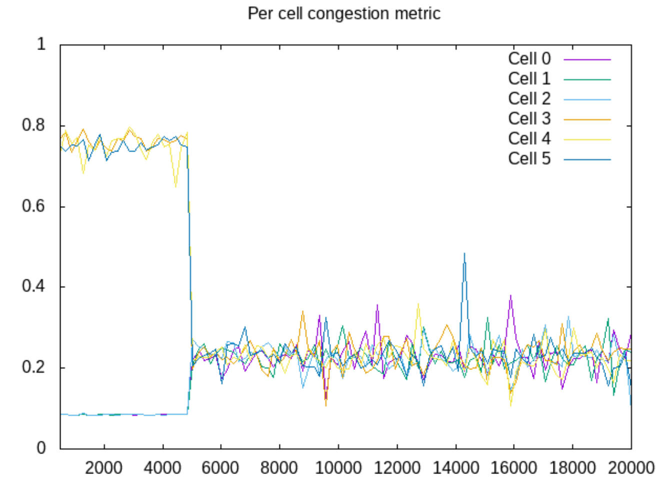
(c) Regulated demand, $\eta = 0.5$

Figure 6. Simulation results - six cell cluster, comparative capacity utilization with handovers and channel swapping

heterogeneous environments, such as leader-follower environments, as studied in [7]. This kind of environment occurs, for example, when there is a single macro-cell and multiple small-cells; the macro-cell follows its own individual control law, to which the small cell has to adapt. This gives rise to interesting one-sided equilibriums, which are not readily solved using the Nash Certainty Equivalence principle. These shall be considered in future work.

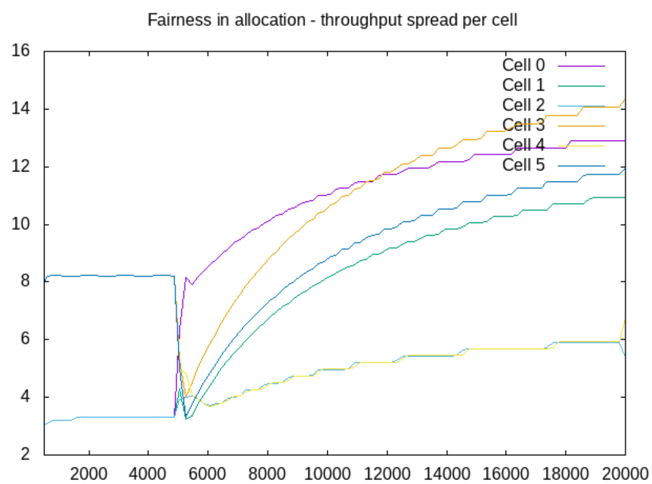


(a) Regulated demand, $\eta = 0.3$

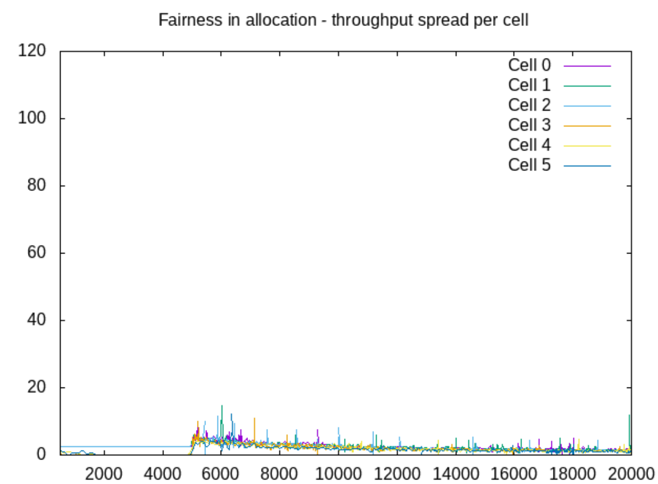


(b) Regulated demand, $\eta = 0.5$

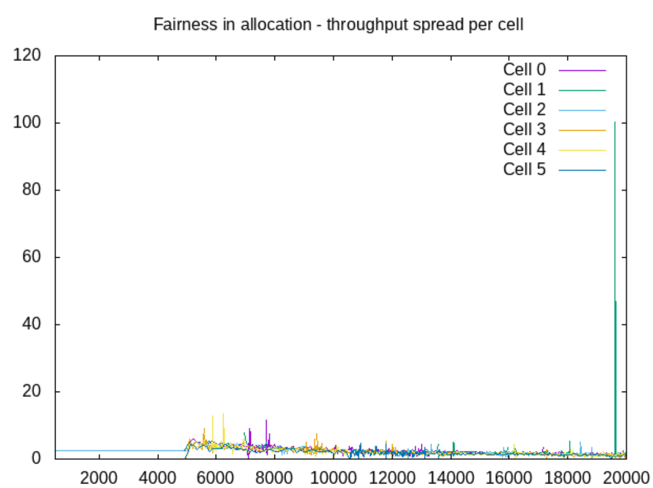
Figure 7. Simulation results - six cell cluster, congestion metric feedback to macro resource controller



(a) No demand regulation



(b) Regulated demand, $\eta = 0.3$



(c) Regulated demand, $\eta = 0.5$

Figure 8. Simulation results - six cell cluster, allocation fairness with handovers and channel swapping

References

[1] A. Saha, "Dynamic resource allocation and balanced cell loading - a stochastic meanfield approach," in Proceedings of the 14th Advanced international conference on Telecommunications. IARIA, July 2018, pp. 69–74.

[2] L. Georgiadis, M. J. Neely, L. Tassiulas *et al.*, "Resource allocation and cross-layer control in wireless networks," Foundations and Trends in Networking, vol. 1, no. 1, 2006, pp. 1–144.

[3] M. J. Neely, E. Modiano, and C.-P. Li, "Fairness and Optimal Stochastic control for Heterogeneous Networks," IEEE/ACM Transactions On Networking, vol. 16, no. 2, 2008, pp. 396–409.

[4] T. Holliday, A. Goldsmith, N. Bambos, and P. Glynn, "Distributed power and admission control for time-varying wireless networks," in Proceedings of the International Symposium on Information Theory. IEEE, 2004, pp. 352–352.

[5] J.-M. Lasry and P.-L. Lions, "Mean field games," Japanese journal of mathematics, vol. 2, no. 1, 2007, pp. 229–260.

[6] M. Huang, R. P. Malhamé, P. E. Caines *et al.*, "Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle," Communications in Information & Systems, vol. 6, no. 3, 2006, pp. 221–252.

[7] M. Huang, P. E. Caines, and R. P. Malhamé, "Individual and mass behaviour in large population stochastic wireless power control problems: centralized and Nash equilibrium solutions," in Proceedings of the 42nd IEEE Conference on Decision and Control. IEEE, 2003, pp. 98–103.

[8] O. G. Aliu, A. Imran, M. A. Imran, and B. Evans, "A survey of self organisation in future cellular networks," IEEE Communications Surveys & Tutorials, vol. 15, no. 1, 2013, pp. 336–361.

[9] D. Liu, L. Wang, Y. Chen, M. Elkashlan, K.-K. Wong, R. Schober, and L. Hanzo, "User association in 5G networks: A survey and an outlook," IEEE Communications Surveys & Tutorials, vol. 18, no. 2, 2016, pp. 1018–1044.

[10] L. Du, J. Bigham, L. Cuthbert, P. Nahi, and C. Parini, "Intelligent cellular network load balancing using a cooperative negotiation approach," in IEEE Wireless Communications and Networking, March 2003, pp. 1675–1679 vol.3.

[11] J. Widmer, R. Denda, and M. Mauve, "A survey on tcp-friendly congestion control," IEEE network, vol. 15, no. 3, 2001, pp. 28–37.

[12] J. G. Andrews, S. Singh, Q. Ye, X. Lin, and H. S. Dhillon, "An overview of load balancing in hetnets: old myths and open problems," IEEE Wireless Communications, vol. 21, no. 2, April 2014, pp. 18–25.

[13] H. Kim, G. De Veciana, X. Yang, and M. Venkatachalam, "Distributed alpha-optimal user association and cell load balancing in wireless networks," IEEE/ACM Trans. Netw., vol. 20, no. 1, Feb. 2012, pp. 177–190. [Online]. Available: https://doi.org/10.1109/TNET.2011.2157937

[14] K. Son, S. Chong, and G. D. Veciana, "Dynamic association for load balancing and interference avoidance in multi-cell networks," IEEE Transactions on Wireless Communications, vol. 8, no. 7, July 2009, pp. 3566–3576.

[15] O. K. Tonguz and E. Yanmaz, "The mathematical theory of dynamic load balancing in cellular networks," IEEE Transactions on Mobile Computing, vol. 7, no. 12, Dec 2008, pp. 1504–1518.

[16] P. Muñoz, R. Barco, J. M. Ruiz-Avilés, I. de la Bandera, and A. Aguilar, "Fuzzy rule-based reinforcement learning for load balancing techniques in enterprise lte femtocells," IEEE Transactions on Vehicular Technology, vol. 62, no. 5, 2013, pp. 1962–1973.

[17] J. Bigham and L. Du, "Cooperative negotiation in a multi-agent system for real-time load balancing of a mobile cellular network," in Proceedings of the 2nd International Joint Conference on Autonomous Agents and Multiagent Systems. New York, NY, USA: ACM, 2003, pp. 568–575. [Online]. Available: http://doi.acm.org/10.1145/860575.860666

[18] L. Du, J. Bigham, and L. Cuthbert, "Towards intelligent geographic load balancing for mobile cellular networks," IEEE Transactions on Systems, Man and Cybernetics, Part C (Applications and Reviews), vol. 33, no. 4, 2003, pp. 480–491.

[19] A. K. Gupta, H. S. Dhillon, S. Vishwanath, and J. G. Andrews, "Downlink multi-antenna heterogeneous cellular network with load balancing," IEEE Transactions on Communications, vol. 62, no. 11, Nov 2014, pp. 4052–4067.

[20] D. Bethanabhotla, O. Y. Bursalioglu, H. C. Papadopoulos, and G. Caire, "User association and load balancing for cellular massive mimo," in Information Theory and Applications Workshop (ITA), Feb 2014, pp. 1–10.

[21] S. K. Das, S. K. Sen, and R. Jayaram, "A dynamic load balancing strategy for channel assignment using selective borrowing in cellular mobile environment," Wireless Networks, vol. 3, no. 5, Oct 1997, pp. 333–347.

[22] S. Kim and P. K. Varshney, "Adaptive load balancing with preemption for multimedia cellular networks," in IEEE Wireless Communications and Networking, March 2003, pp. 1680–1684 vol.3.

[23] A. Awada, B. Wegmann, I. Viering, and A. Klein, "A game-theoretic approach to load balancing in cellular radio networks," in Personal Indoor and Mobile Radio Communications (PIMRC), IEEE 21st International Symposium on. IEEE, 2010, pp. 1184–1189.

[24] M. Handley, S. Floyd, J. Padhye, and J. Widmer, "TCP Friendly Rate Control (TFRC): Protocol Specification," Internet Requests for Comments, RFC Editor, RFC 3448, January 2003. [Online]. Available: http://www.rfc-editor.org/rfc/rfc3448.txt

[25] V. Paxson and S. Floyd, "Wide area traffic: the failure of poisson modeling," IEEE/ACM Transactions on Networking (ToN), vol. 3, no. 3, 1995, pp. 226–244.

[26] J. Padhye, V. Firoiu, D. Towsley, and J. Kurose, "Modeling tcp throughput: A simple model and its empirical validation," ACM SIGCOMM Computer Communication Review, vol. 28, no. 4, 1998, pp. 303–314.

[27] S. H. Low, "A duality model of tcp and queue management algorithms," IEEE/ACM Transactions on Networking (ToN), vol. 11, no. 4, 2003, pp. 525–536.

[28] J. Padhye, V. Firoiu, and D. Towsley, "A stochastic model of tcp reno congestion avoidence and control," University of Massachussetts, Tech. Rep., 1999.

[29] R. Shorten, F. Wirth, and D. Leith, "A positive systems model of tcp-like congestion control: asymptotic results," IEEE/ACM transactions on networking, vol. 14, no. 3, 2006, pp. 616–629.

[30] J. Padhye, J. Kurose, D. Towsley, and R. Koodli, "A model based tcp-friendly rate control protocol," in Proceedings of NOSSDAV. Citeseer, 1999.

[31] X. Yin, A. Jindal, V. Sekar, and B. Sinopoli, "A control-theoretic approach for dynamic adaptive video streaming over http," in ACM SIGCOMM Computer Communication Review. ACM, 2015, pp. 325–338.

[32] F. Kelly, "Mathematical modelling of the internet," in Mathematics unlimited, 2001 and beyond. Springer, 2001, pp. 685–702.

[33] F. Paganini, J. Doyle, and S. Low, "Scalable laws for stable network congestion control," in Proceedings of the 40th IEEE Conference on Decision and Control, Dec 2001, pp. 185–190 vol.1.

[34] S. H. Low, F. Paganini, and J. C. Doyle, "Internet congestion control," IEEE Control Systems Magazine, vol. 22, no. 1, Feb 2002, pp. 28–43.

[35] F. P. Kelly, "Effective bandwidths at multi-class queues," Queueing systems, vol. 9, no. 1-2, 1991, pp. 5–15.

[36] A. I. Elwalid and D. Mitra, "Effective bandwidth of general markovian traffic sources and admission control of high speed networks," IEEE/ACM Transactions on Networking (TON), vol. 1, no. 3, 1993, pp. 329–343.

[37] C.-S. Chang and J. A. Thomas, "Effective bandwidth in high-speed digital networks," IEEE Journal on Selected areas in Communications, vol. 13, no. 6, 1995, pp. 1091–1100.

[38] M. Bardi, "Explicit solutions of some linear-quadratic Mean field games," Networks and heterogeneous media, vol. 7, no. 2, 2012, pp. 243–261.

[39] A. Bensoussan, K. Sung, S. C. P. Yam, and S.-P. Yung, "Linear-quadratic Mean field games," Journal of Optimization Theory and Applications, vol. 169, no. 2, 2016, pp. 496–529.

[40] O. Guéant, "Mean field games - equations with quadratic Hamiltonian: a specific approach," Mathematical Models and Methods in Applied Sciences, vol. 22, no. 09, 2012, p. 1250022.

[41] H. Pham and X. Wei, "Dynamic Programming for McKean-Vlasov dynamics," Control and Optimization, SIAM Journal on, vol. 55, no. 2, 2017, pp. 1069–1011.

[42] J. Yong and X. Y. Zhou, *Stochastic controls: Hamiltonian systems and HJB equations*. Springer Science & Business Media, 1999, vol. 43.

[43] S. Peng, "A general stochastic maximum principle for optimal control problems," SIAM Journal on control and optimization, vol. 28, no. 4, 1990, pp. 966–979.

[44] R. Carmona, F. Delarue *et al.*, "Forward–backward stochastic differential equations and controlled McKean–Vlasov dynamics," The Annals of Probability, vol. 43, no. 5, 2015, pp. 2647–2700.

[45] R. Carmona, F. Delarue, and A. Lachapelle, "Control of McKean–Vlasov dynamics versus Mean field games," Mathematics and Financial Economics, 2013, pp. 1–36.