CHAPTER

9 Ecologically Inspired Resource Distribution Techniques for Sustainable Communication Networks

Kumudu S. Munasinghe[1] **and Abbas Jamalipour**[2]

[1]*Faculty of Education, Science, Technology, and Mathematics, The University of Canberra, Canberra, Australia*

[2]*School of Electrical and Information Engineering, The University of Sydney, Sydney, NSW, Australia*

CHAPTER CONTENTS

9.1 INTRODUCTION

Future generations of communication networks, also known as next-generation networks (NGNs), have converged heterogeneous architectures (e.g., different access networks, traffic conditions, and user behavior). Hence, the core of an NGN can be seen as a converging point for heterogeneous classes of traffic. Therefore, it is absolutely essential to have appropriate controls for the NGN for optimizing resource utilization in an environment with multiservice session flows. The conventional method for guaranteeing fair allocation of resources is via admission control techniques. Regardless of the admission controlling scheme (e.g., measurement-based, transmission power-based, system load-based, throughput-based, and bandwidth-based algorithms (Tragos et al., 2008)), as the availability of resources diminishes with time, competition among the newly arriving flows become inevitable. Several

studies have been carried out for observing these competitive outcomes for classical Internet traffic (i.e., homogeneous flows) (Voorhies et al., 2006; Guo and Matta, 2001). This initially encouraged us to revisit this problem from the perspective of a heterogeneous NGN. For the first time, we carried out an investigation on resource competition that takes place in an NGN via an ecologically inspired approach (Munasinghe and Jamalipour, 2011). The conclusion of this study was that the session class with the lowest resource requirements competitively dominates its presence over other classes, despite the presence of admission control mechanisms. This motivated us to explore a solution for equitable resource distribution, which does not disadvantage either session class. Hence, this chapter presents an ecologically inspired graphical theory for finding an equilibrium outcome, such that all competing classes in an NGN may coexist, despite competition among themselves. This will be an important step toward modeling the behavior and conditions for the stability of an NGN ecosystem, which to the best of our knowledge has not been explored as yet.

Next, we will apply the aforementioned ecologically inspired graphical theory for developing a mobility load-balancing optimization (MLBO) scheme for the NGN. In this case, the MLBO is designed for a long-term evolution (LTE) system as one of its self-organizing network (SON) functions. The objective of MLBO is to balance the load of an eNodeB to maintain its stability, so that the number of handoffs can be minimized. As per the 3rd Generation Partnership Project (3GPP) specification (3GPP, 2011a,b), an MLBO must be able to successfully balance not only its radio load, but also its transport network load (i.e., S1 interface). To the best of our knowledge, an MLBO algorithm capable of handling both aspects has not been proposed as yet. Therefore, as a very first step toward achieving this goal, our distributed MLBO algorithm will be capable of simultaneously balancing the load of the radio and transport interfaces. This ecologically inspired graphical theory will successfully identify an optimized resource usage region in order to achieve a balanced load outcome between both radio and transport interfaces. It also has necessary coordination with other SON algorithms such as mobility robustness/handover optimization (MR/HO) and energy savings to achieve better system stability.

The remainder of this chapter is organized as follows: The next section introduces the consumer resource interaction theory. After this, the theory's applicability to the NGN is presented. The next sections provide a simulation-based validation of the aforementioned graphical equilibrium theory and its proposed applicability to the NGN. The next sections describe how this theory could be applied for developing a novel MLBO framework for the LTE system, followed by results and conclusion.

9.2 CONSUMER-RESOURCE DYNAMICS

Resources are entities that contribute positively to population growth, and are consumed in the process. We have identified that the dynamics of abiotic resources such as energy, bandwidth, content, and so on adequately describe the characteristics of the resources in an NGN. Hence, as per the consumer resource interaction theory, the following four pieces of information are needed for predicting the equilibrium of a

resource competition: the reproductive/growth response of the competing species against the resource, the mortality rate experienced by the same, the supply rate of the resource, and the consumption rate of each resource by each species. The two conditions of equilibrium to be satisfied are (1) the resource-dependent reproduction of each species must balance its mortality rate, and (2) a resource supply exactly balances total resource consumption for each resource. The relationship between the aforementioned four pieces for satisfying the conditions for equilibrium could be mathematically expressed as

$$\frac{dN_i}{dt} = \left[\mu_i(R_j) - m_i\right]N_i \quad \text{for } i = 1, \ldots, n \tag{9.1}$$

$$\frac{dR_j}{dt} = S_j - \sum\nolimits_{i=1}^{n} \mu_i(R_j) N_i Q_{ij} \quad \text{for } j = 1, \ldots, m \tag{9.2}$$

As per Equation (9.1), the resource-dependent population growth dN_i/dt, is equal to μ_i, the growth rate at resource availability R_j minus the mortality rate m_i, where N_i is the population density for a species i. Further, Equation (9.2) represents the resource availability. Here, S_j is the maximum available level for a particular resource j that can ever be available in a closed habitat, and Q_{ij} is the resource quota (i.e., amount of resource j contained in a unit of population i). Also, because the types of resources considered in our discussion are abiotic in nature, there is no supply rate. Therefore, at any given time, such a resource is either in a bound or a free state. In order to enable a clear and concise validation of this graphical theory, we will consider a case of two resources. This gives special insight on how to classify resources based on their joint effects on the per capita growth rate of a population.

For two noninteractive and essential resources, the population growth requires consumption of each of the two resources, say, R_1 and R_2. Consider the first equilibrium arising from the growth of a single population on two such essential resources. By setting $dN_i/dt = 0$ in Equation (9.1), the growth isocline where a species' growth equals its mortality rate is obtained. This is called the zero net growth isocline (ZNGI) for that species. The population density will remain unchanged for a habitat when resource availabilities lie on the ZNGI, hence the satisfaction of the following condition:

$$\mu_i(R_1, R_2) = m_i. \tag{9.3}$$

As previously mentioned, the ZNGI is only half of the condition for establishing equilibrium. The second condition is where resource consumption rates balance resource supply rates. This balance in resource dynamics can be expressed from the following vector equation:

$$\overrightarrow{u} = \begin{bmatrix} S_1 - R_1^* \\ S_2 - R_2^* \end{bmatrix} = \begin{bmatrix} N_i^* Q_{i1} \\ N_i^* Q_{i2} \end{bmatrix} = \overrightarrow{c_i} \tag{9.4}$$

where $\overrightarrow{c_i}$ is the consumption vector and $\overrightarrow{u}$ is the supply vector, which indicated the trajectory that (R_1^*, R_2^*) would take if consumption ceased (Tilman, 1980). The simultaneous solution of Equations (9.3) and (9.4) for unknown R_1^*, R_2^*, and N_i^* defines

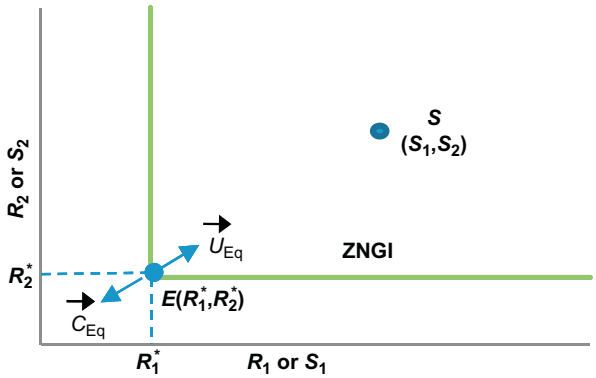


FIGURE 9.1

The equilibrium point, *E*, associated with a resource supply point, *S*, for two essential resource classes, R_1 and R_2.

the equilibrium for a single population (Leon and Tumpson, 1975). As per the graphical theory, out of all the points on the ZNGI, there is one point where Equation (9.4) will be satisfied—in other words, where $\vec{c_i}$ the consumption vector will be opposite in direction to $\vec{u}$ the resource supply vector. The slope of $\vec{c_i}$ the consumption vector is $-Q_{i2}/Q_{i1}$. The minus sign indicates that the consumption reduces the resource availability. The solution of Equation (9.4) also requires that $\vec{u}$ the supply vector has the same slope but points the opposite direction, in fact, directed toward the supply point. This is the resource equilibrium point, (R_1^*,R_2^*), as illustrated in Figure 9.1, directed toward $S(S_1,S_2)$ resource supply point. At this instance, N_i^* is the equilibrium density of species i and R^* is defined as the resource availability at which population growth rate equilibrates. Therefore, the graphical representation of the equilibrium of Equations (9.1) and (9.2) consists of finding the point lying on the isocline such that the supply vector of slope Q_{i2}/Q_{i1} also points toward the supply point. The final requirement is that the magnitudes of vectors $\vec{c_i}$ and $\vec{u}$ be equal.

Nevertheless, we wish to point out to the reader that there is much deeper meaning to the concept of R^*, which is governed by the R^* rule. Interested readers are referred to Grover (1997) for further information. Similarly, competition between two (or more) species for two essential resources could be represented by superimposing the graphical elements, as derived above. The condition for coexistence is where the ZNGIs of each competing population intersect (Phillips, 1973). This also implies that there is a set of resources available where all populations may coexist.

9.3 RESOURCE COMPETITION IN THE NGN

Similar adaptations of ecologically inspired solutions for the NGN are scarce in the literature. One such rare example is where an ecologically inspired approach is applied for developing an access selection algorithm and a price control algorithm

for optimizing network revenue for a heterogeneous network (Chen et al., 2006). Furthermore, although not directly related, there are two noteworthy instances where ecologically inspired concepts have been applied to the field of information technology (IT). In one instance, a conceptual view of a sustainable IT ecosystem was proposed (Bash et al., 2008). In the other, a conceptual model toward homogenizing between two Internet enterprises through an ecologically inspired approach was presented (Yingzi et al., 2007).

In the proposed approach, the NGN can be considered as an ecosystem of heterogeneous networks over which multiservice application traffic flows. Hence, we argue that, similar to the competition between species for a limited supply of resources in an ecological system, sessions belonging to different classes may also compete for limited resources within a converged heterogeneous networking ecosystem. Therefore, if an NGN traffic/session class can be considered analogous to a species' group in an ecosystem, competition and distribution of resources could be modeled similar to an ecological environment. As illustrated in Figure 9.2, let's assume that one of the gateways at the backbone network that connects access networks to the Internet (i.e., the core of the NGN) becomes a bottleneck. As in the case of any bottleneck, this too has limited resource availability. The remainder of this discussion considers two basic resources, bandwidth and power, which are abiotic in nature and often affect the performance of a network bottleneck. Hence, let's assume that the bottleneck has a maximum bandwidth and power capacity of, say, S_b and S_p. For clarity and convenience, all available resources will be referred to as R_j, where $j=1,\ldots,n$ (i.e., $j=2$ in this case). As session flows corresponding to different service classes increase, resource levels at the gateway start to decrease.

If we consider these different service classes as various species groups indicated by $i=1,\ldots,n$, the minimum required resource quota for each flow belonging to a service class can be denoted as Q_{ij}. Let the growth rate of flow arrivals attributing to service class (or species) i be denoted by the function f_i, where it is assumed that either $f_i>0$ (i.e., the arrival request rate increases), or there exist enough requests for exhausting the gateway's resources. Note that existing flows have limited lifetimes due to sessions being hung up, and therefore frequently change the state of the abiotic resource from bound to free state and vice versa over a period of time. Such behavior is also shown by f_i. Therefore, the ecologically inspired dynamics for such a competition can be yielded by

$$\frac{dN_i}{dt}=[f_i-d_i(R_1,R_2)]N_i \quad \text{for } i=1,\ldots,n \tag{9.5}$$

$$\frac{dR_j}{dt}=S_j-\sum_{i=1}^{n} f_i N_i Q_{ij} \quad \text{for } j=1,2 \tag{9.6}$$

where N_i is the population density of active flows attributed to class i, and function d_i is the drop rate for class i at resource availability R_j. It is obvious that flows of the service class i are dropped when the available resource $R_j<Q_{ij}$, as j falls below its

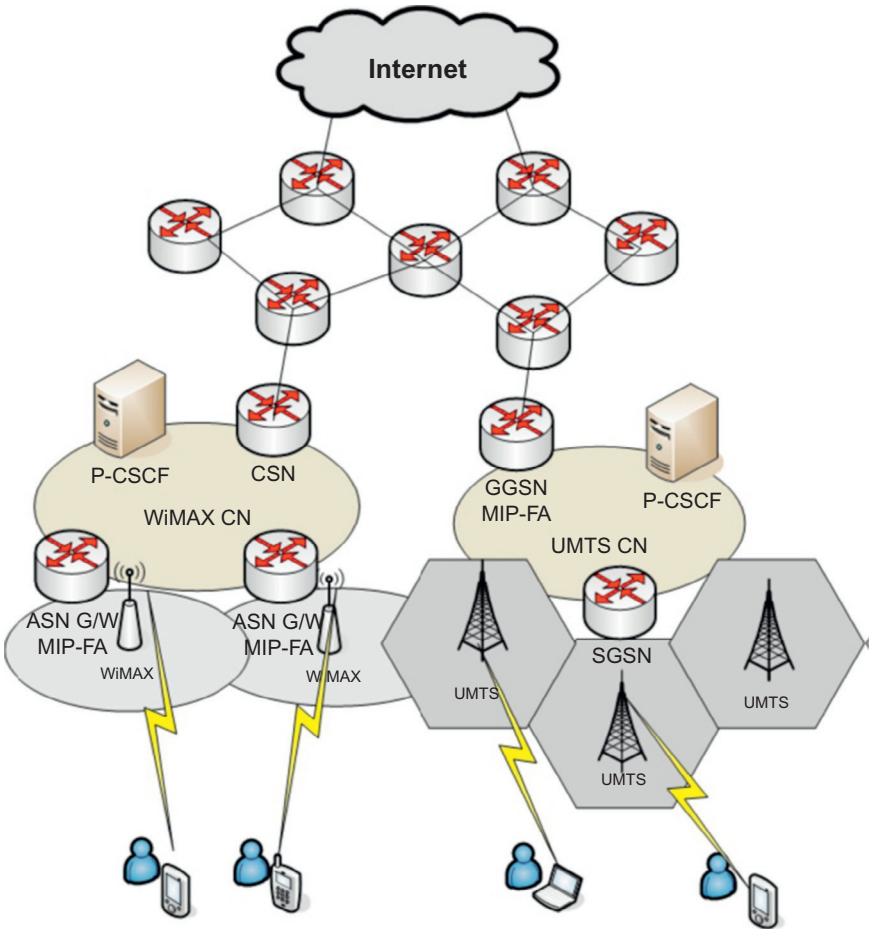


FIGURE 9.2

NGN core network architecture.

minimal quota. Similar to the scenario in Equations (9.1) and (9.2), in the case of an NGN, different classes have mutually negative effects on each other due to the consumption of limited resources. In other words, session classes in Equations (9.5) and (9.6) do not have any direct impact on each other, except through the competition for the diminishing resources. For an NGN where only one class exists, the population density (N_i) increases for $f_i > 0$. As the population grows, the available resource level declines. This level continues to decline as the number of flows increases, until the remaining resource R_j reaches R_j^*, which is considered the equilibrium (stable) point. At R_j^* the growth rate of the flows' arrival would exactly balance its drop rate (i.e., $[f_i = d_i(R_1, R_2)]$), and therefore the number of active sessions neither increases nor decreases.

This is similar to $dN_i/dt = 0$, in which case the population growth of a species stabilizes. This is called the ZNGI for that particular species group. As previously stated, the population density will remain unchanged for a habitat when resource availabilities lie on the ZNGI. It can be argued that R_j^* depends on the characteristics of a flow (e.g., duration, data rate, delay, resource consumption level, and so on) as well as its underlying transport protocol [Transmission Control Protocol (TCP) and User Datagram Protocol (UDP)]. For available resource levels smaller than R_j^*, flows are not admitted, as not enough resources are available for their Quality of Service (QoS) requirements. This trend continues until the resource level reaches an adequate level for accepting a new arrival.

Now for the case of two essential resource groups, the competition for i species groups is represented by superimposing the graphical elements, as derived above. Coexistence requires that the isoclines of the competing population intersect, as illustrated in Figure 9.3, implying that there is a set of resource availabilities for each population alone that can increase (regions 2 and 6), thus satisfying Phillips' coexistence conditions (Phillips, 1973). As per Phillips' conditions, one population ($i = 1$) must be a superior competitor for one resource ($j = 1$) and an inferior competitor for the other ($j = 2$), and the second population ($i = 2$) must be a superior competitor for the second resource ($j = 2$) and an inferior competitor for the first ($j = 1$).

Intersection of ZNGIs guarantees that there is an equilibrium point for both species. However, to see whether this point represents feasible and stable coexistence, consumption and supply vectors must be examined. Therefore, for a stable coexistence, the overall consumption vector $\overrightarrow{c}$, which is the resultant of single population vectors $\overrightarrow{c_1}$ and $\overrightarrow{c_2}$, and the supply vector must have the same magnitude and slope but point in opposite directions. As illustrated in Figure 9.3, for a habitat with two service

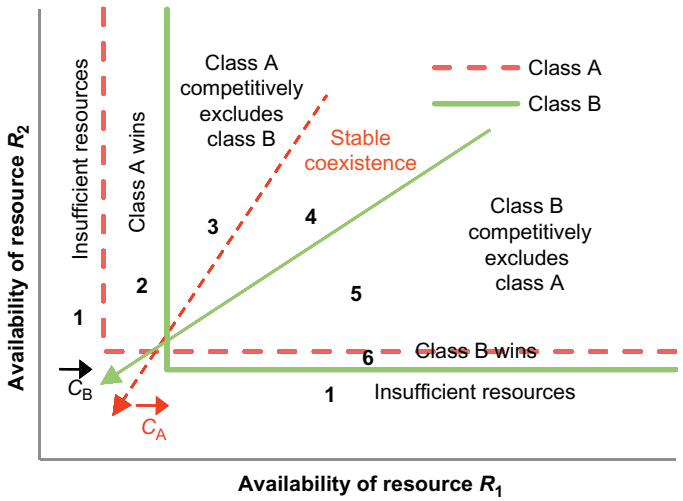


FIGURE 9.3

Competition between two service classes for two resources.

classes (say, A and B) and two resources (say, R_1 and R_2), resource ZNGIs and consumption vectors define six regions. Habitats with resource supply points falling in region 1 have insufficient resources for the survival of neither service class. Next, habitats with supply points falling in region 2 will only have sufficient resources for the survival of class A and insufficient resources for the survival of class B. Although it may seem that there are sufficient resources for the survival of both classes in region 3, class A reduces the resources down to the ZNGI of class B, thus competitively excluding class B.

A set of meaningful equilibrium solutions for both service classes will be possible in region 4 when a resource supply point lies between $\overrightarrow{c_1}$ and $\overrightarrow{c_2}$. This two-service class equilibrium point is locally stable because each species consumes proportionately more of the resource that limits its own growth (Phillips, 1973). This can be confirmed by noting that at the two-class equilibrium point, class A is limited by resource R_2 and class B is limited by resource R_1. Hence it is clear that stable coexistence is feasible only for a restricted set of supply points falling in region 4 of Figure 9.3. Opposite to what was observed for habitats in region 3, class B competitively excludes class A for habitats in region 5. Also, opposite to region 2, for habitats in region 6, available resources are only sufficient for the growth of class B. Then again, for supply points lying in region 1, neither species persists because the habitat is resource poor.

9.4 CONDITIONS FOR STABILITY AND COEXISTENCE

This section provides a simulation-based validation of the aforementioned graphical equilibrium theory of resources and its proposed applicability to the NGN. The aim of this simulation is to identify the level of resource supplies required for a stable coexistence for a given set of service classes. In this case, the two types of NGN core resources considered in this model are bandwidth and power. Similar to Figure 9.1, the basic concepts relating to coupling heterogeneous networks and vertical session handoff used for the simulation setup have been taken from Yingzi et al. (2007) (Munasinghe and Jamalipour, 2007). The MatLab-based simulation topology consists of four UMTS cells, six WLAN hotspots, five WiMAX cells, and a backbone mesh network consisting of nine routers (see Figure 9.1).

For the purpose of analyzing heterogeneous traffic classes at the NGN core network, a flow-level characterization has been used. A flow is defined as a unidirectional sequence of packets having the same identifier and transmitted with an interpacket interval smaller than a certain threshold. The WINNER project (WINNER, 2005) segregates these flows into 18 service classes based on their corresponding application types. Tragos et al. (2008) further summarizes this into eight representative classes by taking their QoS as well as resource requirements into consideration. Out of these eight classes, we have carefully chosen three classes to represent a fair characterization of low, medium, and high resource-consuming flows. As a result, simple telephony and messaging (64 kbps), multimedia telephony

(512 kbps), and HQ video streaming (1024 kbps) have been chosen to represent low, medium, and high resource-consuming classes of flows.

Next, flows belonging to these classes are generated via different networks according to a Poisson distribution, where the mean arrival rate grows over time. Durations of heterogeneous sessions are distributed between 30 and 120 s. This is based on the widely accepted assumption that human-initiated traffic is best modeled as a Poisson distribution (Roberts, 2004). Further, for simulation purposes, a measurement-based admission control algorithm is implemented at gateways (Tragos et al., 2008). Under this admission control algorithm, a flow is rejected if the estimated available resource is below a certain threshold (e.g., bandwidth and power in this case). The adopted routing mechanism for the mesh backbone is the Open Shortest Path First protocol.

Because resource competition at the NGN core mostly happens at gateway routers, as cited in Tucker et al. (2008), Cisco's recommendations are followed when estimating the required energy per transmitted bit at a bottleneck router. Hence, the average total per-bit energy consumption is assumed to be 10 nJ, which is a combination of processing energy, ingress/egress storage energy, transmit/receive energy, switch control energy, routing engine energy, and finally the energy overheads of the power supply inefficiency, including fans and blowers. Therefore, on average, a simple telephony session would consume 0.64 mW of energy from a gateway router. Similarly, an HQ video streaming would consume 5.12 and 20 mW of energy, respectively. Further, each of these routers has a resource availability of 100 Mbps bandwidth and 200 W of power.

We assume that at the NGN core network, streaming flows are handled by network nodes using a nonpreemptive priority queuing. Via application-specific prioritization, maximal responsiveness is ensured for streaming flows, which is otherwise not supported by its transport protocol UDP. Elastic-type flows have purposely been omitted because TCP plays a major role in manipulating the degree of fairness, which is beyond the scope of this discussion. Next, the arrival rates are increased for identifying the bottleneck nodes/routers in the backbone network.

As a result of heavy competition between arriving sessions, the resource availability at the bottleneck may rapidly diminish with time. The first step in this case is the determination of positions of ZNGIs of the flow classes. As previously mentioned, the ZNGIs are determined by the corresponding R^* values of the flow classes for essential resources (i.e., bandwidth and power). According to Section 9.2, R^* is defined as the resource availability at which population growth rate equilibrates (Grover, 1997). This is obtained by individually simulating bandwidth-limited and power-limited growth curves for each of the flow classes at the bottleneck node. Figure 9.4 illustrates the superimposed ZNGIs and the consumption vectors for the three flow classes.

As mentioned in Section 9.2, in order to achieve a locally stable equilibrium point, each competing class must consume relatively more of the resource that limits its growth at equilibrium (Tilman, 1980). This is also known as the coexistence condition for the considered classes. In other words, the ZNGIs of the flow classes must

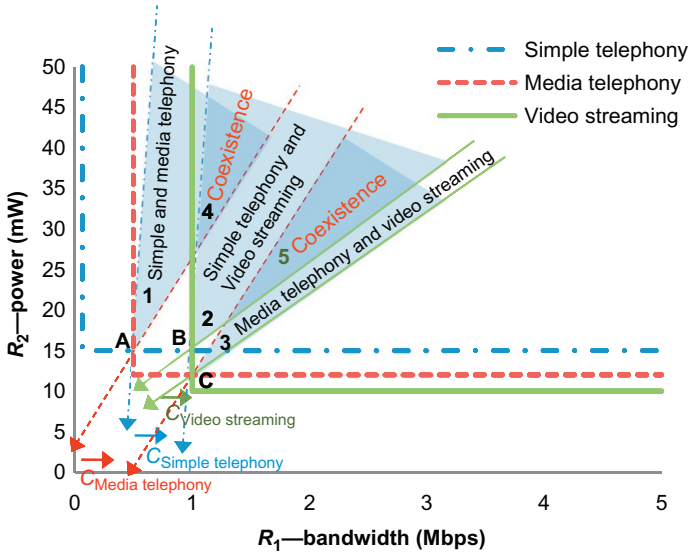
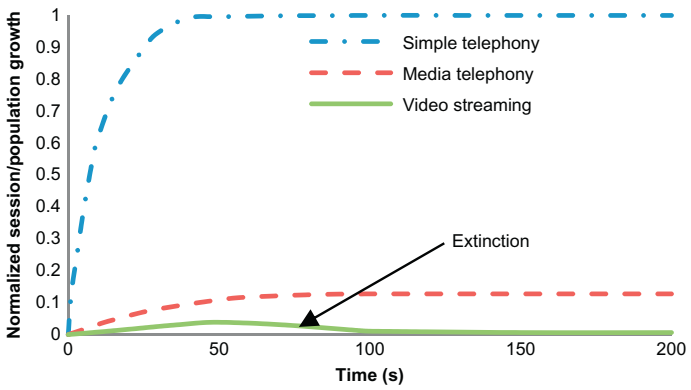


FIGURE 9.4

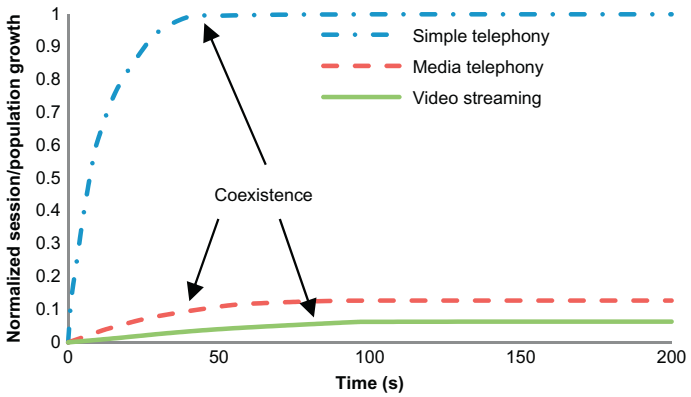
Competition between three service classes for two resources.

intersect, and their resource supply points must "focus" toward an equilibrium point for them to coexist. According to Figure 9.4, three points are noted (A, B, and C) where ZNGIs intersect each other. For example, at point A, media telephony is limited by R_1 (bandwidth) and simple telephony is limited by R_2 (power).

For example, the stable coexistence of simple and media telephony classes would be possible for a habitat with its resource supply point falling in region 1. In this case, the resource levels would eventually reduce down to A, at which point the growth of each flow class would stabilize. It further confirms with the equilibrium theory, because at point A, both of these classes consume proportionately more of the resource that limits their own growth (Bash et al., 2008). This can be seen by noting that at point A, the growth of the simple telephony class is limited by power, and the growth of the media telephony class is limited by bandwidth. In this region, the video-streaming class will be competitively excluded. This is illustrated in Figure 9.5a. Similarly, simple telephony and video-streaming classes coexist in region 2, and video telephony and video-streaming classes coexist in region 3. A further analysis into Figure 9.4 reveals two overlapped regions. Regions 4 and 5 correspond to the overlapping of regions 1 and 2, and 2 and 3, respectively. By applying the aforementioned argument, regions 4 and 5 can be identified as habitats for stable coexistence for all flow classes. Hence, fixing resource supply points within these regions could guarantee stable coexistence of all flow classes. Figure 9.5b illustrates how these three classes coexist when the resource supply point is fixed to region 4. A similar observation is noted when the resource supply point is fixed to region 5.



(a) Resource supply point in region 1—instability of the system.



(b) Resource supply point in region 4—stability ofthe system.

FIGURE 9.5

Dynamics of competition at different resource supply regions.

9.5 APPLICATION FOR LTE LOAD BALANCING

Next, we will apply the aforementioned ecologically inspired graphical theory for developing an MLBO scheme for an LTE system. The aim of MLB is to distribute user traffic across the system (i.e., between adjacent cells) in such a way that quality end user experience and higher system capacity is achieved. To this end, there have been two main approaches proposed thus far: distributed and centralized load balancing (LB). In the case of LTE, distributed LB approaches are better suited. For example, in the case of Lv et al. (2010), Zhang et al. (2011a,b), Kwan et al. (2010), Lobinger et al. (2011), and Zhang et al. (2011a,b), the algorithms run locally at eNodeBs and load information is exchanged via X2 interfaces. On the other hand, centralized LB approaches such as Suga et al. (2012) also exist, where the algorithm runs in a core network element. Nevertheless, a common deficiency in all of these

algorithms is that the LB algorithm only takes into account the radio resource usage or simply the radio interface. Therefore, there is a need for designing a novel LB algorithm that is capable of optimizing multiple load aspects (e.g., radio and transport loads in this case).

The proposed distributed MLB framework resides at each eNodeB of the LTE Evolved - Universal Terrestrial Radio Access Network (E-UTRAN). The eNodeBs are interconnected with each other via the X2 interface and are connected to the LTE evolved packet core via the S1 interface. The functionality of the proposed MLB framework for an intra-LTE scenario is illustrated in Figure 9.6. First, the load information function senses and calculates the resource/load usage of the eNodeB for all interfaces (i.e., radio and transport in this case). Next, it exchanges cell-load information between adjacent eNodeBs via the X2 interface. This exchange of information happens periodically. However, under extreme load conditions, it could be an event-triggered operation. Then, this load information is fed into the load eco-inspired optimization function. This is where the stability of the actual eNodeB is evaluated. This algorithm defines three states for the eNodeB: stable, unstable, and overload.

If our proposed eco-inspired optimization function finds the eNodeB to be stable, it would then trigger the energy-saving SON function (beyond the scope of this chapter). If it finds the cell to be unstable, it would inform the call admission control (CAC) algorithm to selectively/partially block certain incoming session classes until the eNodeB regains its stability. Next, the LB hand-over process starts. Finally, if the cell is overloaded, it will first inform the CAC to fully block all sessions and initiate the LB hand-over process.

In each of these last two cases, the next step is to identify sessions for performing LB handoff. This will first require obtaining neighbor cell-load levels from the load information function and identifying target cells to perform load-balance handoff.

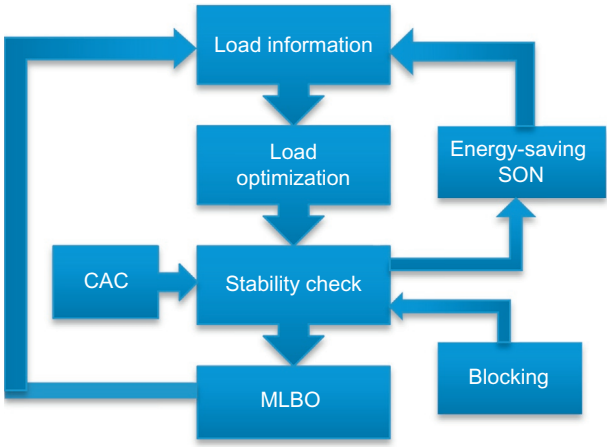


FIGURE 9.6

MLB framework for LTE system.

Once the source cell initiates handover to distribute some of its load, the target cell may perform CAC for the LB handovers. Then the handoff is performed. The key point to note is that a LB handoff must be clearly distinguishable from a normal handoff for the target cell. Finally, the updating of handover and/or reselection configuration settings between the source and target cells takes place (Lv et al., 2010).

9.6 VALIDATION AND RESULTS

This section provides a simulation-based validation of the proposed MLB framework for an intra-LTE scenario. The LTE E-UTRAN setup consists of a 37 hexagonal-cell network connected via X2 interface. Because the proposed mechanism is distributed, each eNodeB will execute the MLB framework for balancing load between two of its interfaces (i.e., radio and transport) and then offloading to its neighboring cells. For the purpose of analyzing heterogeneous traffic classes at the eNodeB, a flow-level characterization has been used. A flow is defined as a unidirectional sequence of packets having the same identifier and transmitted with an interpacket interval smaller than a certain threshold. Based on the recommendations by the WINNER project (2005) and Tragos et al. (2008), we have carefully chosen three classes to represent a fair characterization of low, medium, and high resource-consuming flows. As a result, simple telephony and messaging (64 kbps), multimedia telephony (512 kbps), and HQ video streaming (1024 kbps) have been chosen to represent low, medium, and high resource-consuming classes of flows.

In the case of radio resource usage, the total (uplink and downlink) physical resource block (PRB) usage is considered (3GPP, 2011a,b). Therefore, the load is defined as the fraction of used PRBs in a cell. Constant bit rate is used for maintaining simplicity for this simulation. Based on the achievable throughput at 5 dB signal-to-interference noise ratio, the required number of PRBs for the above flow classes are calculated with the use of the look-up table in 3GPP (2009) for the best coding schemes (PRB bandwidth in LTE is 180 kHz). Hence, the required PRBs for 64, 512, and 1024 kbps flows are 1, 3, and 5, respectively. We consider each radio interface to have a 20 MHz channel bandwidth, and therefore 100 PRBs.

On the other hand, the transport interface is considered to have a fixed-link bandwidth of 10 Mbps connecting it to the backbone via the S2 link. Further, for simulation purposes, a measurement-based admission control algorithm is implemented at the transport interface (Voorhies et al., 2006). Under this admission control algorithm, a flow is rejected if the estimated available resource is below a certain threshold (e.g., bandwidth in this case). Next, a normalized user traffic pattern $A(t)$ is generated. This pattern is approximated by a simple sinusoidal profile, can be considered as a theoretical model, and is given by $A(t) = \frac{1}{2^b}\left[1 + \sin\left(\frac{\pi t}{12}\right)\right]^b + k$, where $A(t)$ is the instantaneous normalized traffic in Erlang, k is a constant, and $b \in \{1,3\}$, which determines the abruptness of the traffic profile (Marsan and Meo, 2009). A Poisson-distributed random process is added with both the patterns, modeling the random fluctuations of the total traffic. This is based on the widely accepted

assumption that human-initiated traffic is best modeled as a Poisson distribution (Roberts, 2004).

As a result of heavy competition between arriving sessions at the eNodeB, the resource availability at the radio and transport interfaces may rapidly diminish with time. The first step of the optimization algorithm is the determination of positions of ZNGIs of the flow classes. As previously mentioned, the ZNGIs are determined by the corresponding R^* values of the flow classes for essential resources (i.e., bandwidth and power). R^* is defined as the resource availability at which population growth rate equilibrates (Grover, 1997). This is a predefined parameter for the eNodeB by individually simulating bandwidth-limited growth curves for each of the flow classes for each interface. Figure 9.7 illustrates the superimposed ZNGIs and the consumption vectors for the three flow classes.

As explained previously, in order to achieve a locally stable balanced-load condition, each competing class must consume relatively more of the resource that limits its growth at equilibrium (Phillips, 1973). According to Figure 9.7, under the given conditions, two stable regions can be identified for an eNodeB to perform. Hence, fixing resource supply points within these regions could guarantee stable coexistence of all flow classes. Further, Figure 9.7 shows three unstable regions. In these regions, one class may end up being competitively excluded by the others. In this case, the load-optimization algorithm will initiate selective blocking for overcoming this condition. Should the resource supply point fall on any other region, the optimization algorithm will trigger the overload case and hand off excess sessions until the resource supply points arrive at a stable region.

Figure 9.8 illustrates the number of unsatisfied users in the system against LB time for a 24-h period. Obviously, for the case of no MLB, a relatively higher number of unsatisfied users can be observed. Then, we compare our proposed model with a closely similar reference model (Zhang et al., 2011a,b) to evaluate

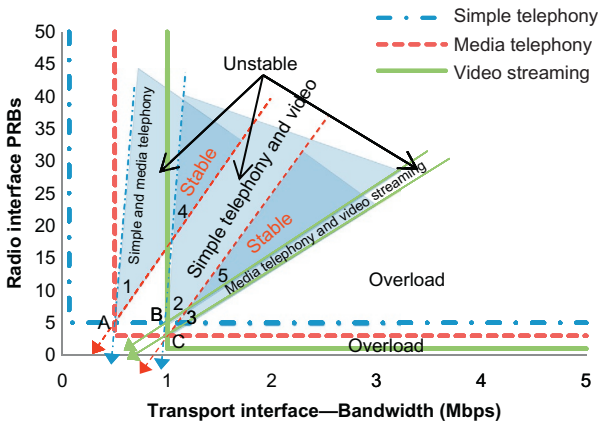


FIGURE 9.7

Decision making of the optimization algorithm.

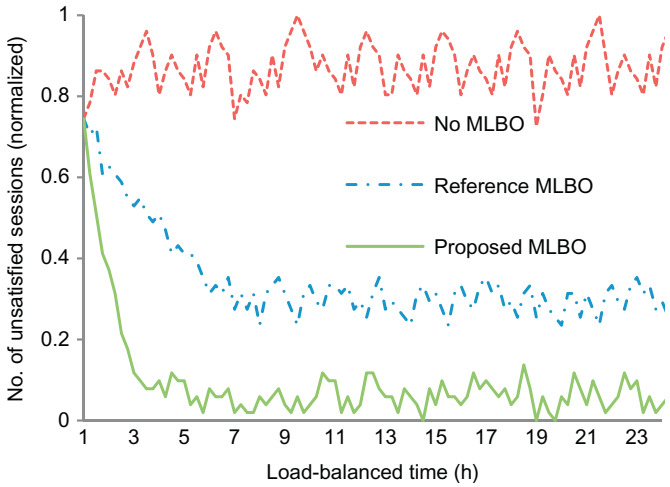


FIGURE 9.8

The number of unsatisfied users.

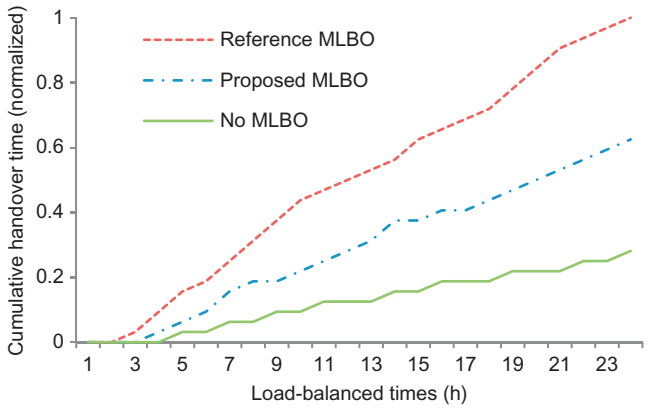


FIGURE 9.9

Cumulative handover times.

its performance. Our proposed model proves to be much more effective than Zhang et al's (2011a,b) in reducing the number of unsatisfied users. It also proves to be effectively stabilizing a load scenario before the system becomes overloaded. For example, when the system identifies an unstable condition, the eco-inspired optimization identifies specific flow classes that need to be blocked and handed off for the system to gain stability. This can also be noticed from Figure 9.8. Next, Figure 9.9 illustrates the cumulative hand-over time against the LB time. According to the results, it is clear how the proposed MLB framework introduces a relatively lower number

of handoffs in comparison to the reference model in Zhang et al. (2011a,b). Obviously, when no MLB is applied, the handoffs only take place because of to user mobility, so they are minimal. Figure 9.10 illustrates the percentage of Load Balancing Handoffs (LBHOs) of total HOs. Due to the stability of the proposed optimization method, the total number of HOs is reduced by approximately 12-13% more than the reference method. Figure 9.11 illustrates how the load-distribution index varies against time. This reflects the degree of distribution among cells, where the most effective LB (i.e., the proposed) shows a distribution closest to one. Figure 9.12 shows

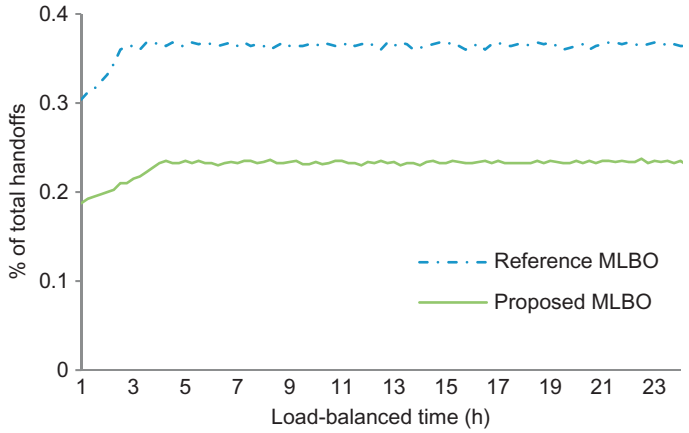


FIGURE 9.10

Percentage of LBHOs of total HOs.

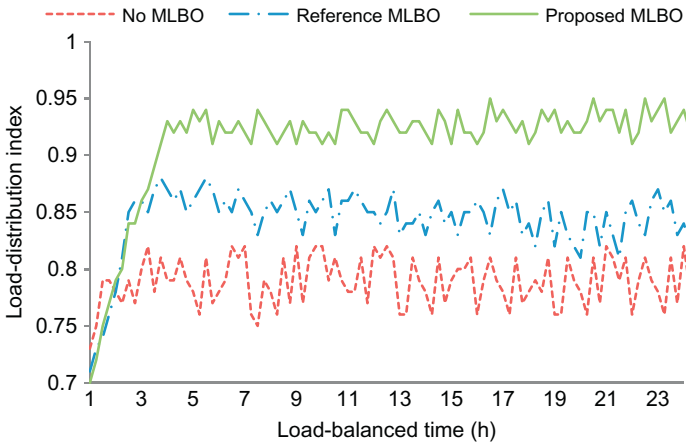


FIGURE 9.11

Load-distribution index.

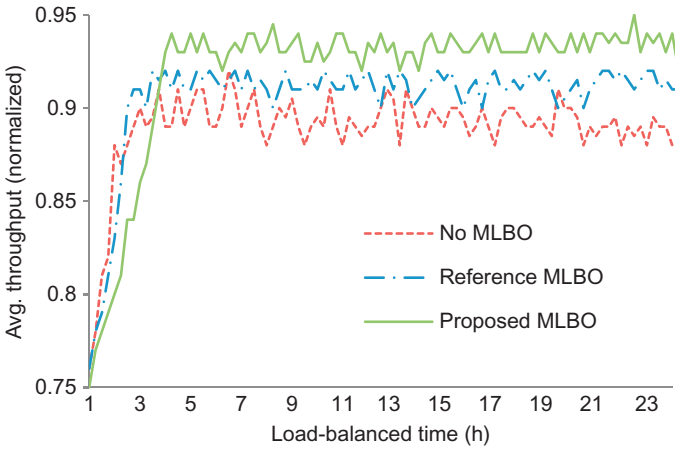


FIGURE 9.12

Average throughput.

the average of each cell's total throughput. After LB, some of the previous unsatisfied user's traffic can be successfully transmitted; hence the overall throughput increases, as illustrated.

9.7 CONCLUSIONS

This chapter develops an ecologically inspired graphical theory for equitable resource distribution in a multiresource, multiclass, heterogeneous NGN. With the proposed ecological graphical theory, optimal resource supply levels required for a stable coexistence could be accurately predicted for a heterogeneous network. Analytical proof and simulation results are provided on how a stable resource supply point could be determined for guaranteeing the coexistence of all flow classes in a closed system, thus achieving sustainability. The chapter further applies the aforementioned graphical theory to develop an MLB framework for LB in a multiresource, multiclass LTE E-UTRAN environment (a component of the NGN). The novelty of the proposed distributed MLB algorithm is that, for the first time, it is capable of simultaneously balancing the load of the radio and transport interfaces. With the help of the ecologically inspired load-optimizing algorithm, optimal load levels required for an eNodeB could be accurately predicted. Analytical proof and simulation results are provided on the effectiveness of our proposal against a conventional MLB, where the former outperforms the latter.

REFERENCES

3GPP, 2009. E-UTRA RF system scenarios (release 8), 3GPP TR 36.942 v8.2.0.
3GPP, 2011a. E-UTRA and E-UTRAN overall description, 3GPP TS 36.300 v10.3.0.

3GPP, 2011b. Self-configuring and self-optimizing networks (SON) use cases and solutions (release 9), 3GPP TR 136.902 v9.3.1.

Bash, C.E., Patel, C.D., Shah, A.J., Sharma, R.K., 2008. The sustainable information technology ecosystem. In: Proceedings of the 11th Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm 08).

Chen, J., et al., 2006. An ecology-based adaptive network control scheme for radio resource management in heterogeneous wireless networks. In: Proceedings of Bio-Inspired Models of Network, Information and Computing Systems, Madonna Di Capiglio, Italy.

Grover, J.P., 1997. Resource Competition. Chapman and Hall, London, UK.

Guo, L., Matta, I., 2001. The war between mice and elephants. In: Proceedings of Conference on Network Protocols, pp. 180–188.

Kwan, R., et al., 2010. On mobility load balancing for LTE systems. In: Proceedings of IEEE 72nd Vehicular Technology Conference Fall (VTC 2010-Fall), Ottawa, Canada.

Leon, J., Tumpson, D., 1975. Competition between two species for two complementary or substitutable resources. J. Theor. Biol. 50, 185–201.

Lobinger, A., Stefanski, S., Jansen, T., Balan, I., 2011. Coordinating handover parameter optimization and load balancing in LTE self-optimizing networks. In: Proceedings of IEEE Vehicular Technology Conference Spring (VTC 2011-Spring), Budapest, Hungry.

Lv, W., et al., 2010. Distributed mobility load balancing with RRM in LTE. In: Proceedings of 3rd IEEE International Conference on Broadband Network and Multimedia Technology (IC-BNMT), Beijing, China.

Marsan, M.A., Meo, M., 2009. Energy efficient management of two cellular access networks. In: Proceedings of Green Metrics Workshop, Washington, USA, pp. 1–5.

Munasinghe, K.S., Jamalipour, A., 2007. A unified mobility and session management platform for next generation mobile networks. In: Proceedings of IEEE Global Communications Conference, Washington, DC, USA.

Munasinghe, K.S., Jamalipour, A., 2011. Resource competition at the NGN core network: an ecologically inspired analysis. In: Proceedings of International Conference on Telecommunications (ICT2011), Cyprus.

Phillips, O.M., 1973. The equilibrium and stability of simple marine biological systems. I. Primary nutrient consumers. Am. Nat. 107, 73–93.

Roberts, J., 2004. Internet traffic, QoS and pricing. Proc. IEEE 92 (9), 1389–1399.

Suga, J., Kojima, Y., Okuda, M., 2012. Centralized mobility load balancing scheme in LTE systems. In: Proceedings of 8th International Conference on Wireless Communication Systems, Limassol, Cyprus.

Tilman, D., 1980. Resources: a graphical mechanistic approach to competition and predation. Am. Nat. 116, 362–393.

Tragos, E.Z., Tsiropoulos, G., Karetsos, G.T., Kyriazakos, S.A., 2008. Admission control for QoS support in heterogeneous 4G wireless networks. IEEE Netw. 22 (3), 32–37.

Tucker, R.S., et al., 2008. Energy consumption in IP networks. In: Proceedings of the European Conference and Exhibition on Optical Communications, Brussels.

Voorhies, S., Lee, H., Klappenecker, A., 2006. Fair service for mice in the presence of elephants. Inf. Process. Lett. 99 (3), 96–101.

WINNER Deliv. D4.4, 2005. Impact of cooperation schemes between RANs, IST-2003-507581 WINNER.

Yingzi, X., Zhenyu, L., Yupei, Z., Jie, L., 2007. Lotka-Volterra model: new perspectives in the study of homogeneity among Internet enterprises. In: Proceedings of International Conference on Wireless Communications (WCNC2007), Networking & Mobile Computing.

Zhang, L., et al., 2011a. A two-layer mobility load balancing in LTE self-organization networks. In: Proceedings of 13th International Conference on IEEE Communication Technology (ICCT), Jinan, China.

Zhang, M., Li, W., Jia, S., Zhang, L., Liu, Y., 2011b. A lightly-loaded cell initiated load balancing in LTE self-optimizing networks. In: Proceedings of 6th International Conference on Communications and Networking in China (CHINACOM), Harbin, China.