

◆ Applying Emergent Self-Organizing Behavior for the Coordination of 4G Networks Using Complexity Metrics

Lester T. W. Ho, Louis G. Samuel, and Jonathan M. Pitts

Future fourth-generation (4G) wireless networks are expected to have an ad hoc, dynamic structure with cheap, ubiquitous, low-powered nodes that are autoconfigurable and flexible. Controlling such a network means coping with uncertainty, not only of traffic demand, but also in the network structure. Because of this, a new approach to the control and coordination of 4G networks will be needed, one that replaces centralized with highly decentralized control. One promising approach is to view networks as self-organizing systems comprising simple interacting nodes that rely on emergent behavior to provide network-wide coordination. However, such networks are often difficult to predict or manipulate, due to their distributed nature. This paper describes the use of an entropy-based complexity metric to investigate and manipulate the behavior of such self-organizing systems in mobile networks. We introduce a self-organizing algorithm for cell dimensioning, and apply the complexity metric to extract information on network-wide behavior. We then introduce a framework for using the metric to manipulate emergent self-organizing behavior in 4G networks. © 2003 Lucent Technologies Inc.

Introduction

To achieve the capacity and high bandwidth anticipated in future-generation wireless networks, it is necessary to employ small-cell technologies, entailing a substantial increase in the number of base stations required by a network. In addition, the adoption of more dynamic network architectures requires networks to have some self-configuration abilities. However, the implementation of self-configuration in a centralized system, particularly one that contains a large number of base stations, presents problems of scalability, because of the large amount of signaling that would be involved. These requirements

have steered the approach to network control away from the centralized and rigid methods used in current networks toward more decentralized and flexible approaches using autoconfigurability and self-organization schemes. Self-organizing algorithms that operate on localized information are particularly useful in providing the scalability and flexibility that future-generation networks require.

The two main challenges involved in using self-organizing behavior are the unpredictability of the network (caused by phase transitions) and the lack of direct control over how the network behaves.

Understanding how to set the parameters of self-organizing algorithms is an important aspect of dealing with these challenges. Using the wrong parameter settings can result in suboptimal performance, or it can cause the network to behave in a totally inappropriate manner, but determining the proper settings is not simple, because they vary from network to network. Thus, it is important to have an easily adaptable and implementable method of evaluating network behavior in order to establish the correct parameters of self-organizing algorithms.

In this paper, we present an overview of the current trends in wireless networks and make a case for using decentralization and self-organizing behavior in them. We then describe a method of monitoring and analyzing the behavior of self-organizing networks using a complexity metric, and proceed to demonstrate the use of the metric in the analysis of a network that is using a distributed cell-dimensioning algorithm. Finally, we discuss the possibility of invoking self-organizing behavior for network coordination using simplified algorithms, with the complexity metric used as a monitoring tool to trigger such behavior.

Systems Beyond Third Generation

Mobile communication networks have been experiencing evolutionary change every decade. The first-generation (1G) networks, which appeared in the 1980s, and the second-generation (2G) networks, which appeared in the 1990s, were used mainly for voice applications and limited circuit switched type services. 2G systems, which are the current mainstream systems and are operated all over the world, are limited to data rates of less than 10 kb/s. Third-generation (3G) systems, which are expected to be deployed in the beginning of the 21st century, are expected to provide 2 Mb/s in an indoor environment, and at least 144 Kb/s in a vehicular environment [34]. It is clear that one of the drivers of mobile communication system evolution is the need to satisfy the demand for high data rates. This is consistent with market forecasts that predict the growing importance of mobile data services, the demand for which is expected to grow continuously over the

Panel 1. Abbreviations, Acronyms, and Terms

BSC—base station controller
 BSR—base station router
 CA—cellular automata
 CDMA—code division multiple access
 EWMA—exponentially weighted moving average
 GSM—Global System for Mobile Communications
 IP—Internet protocol
 LAN—local area network
 MANET—mobile ad hoc network
 MSC—mobile services switching center
 OFDM—orthogonal frequency division multiplexing
 PAN—personal area networks
 PULSERS—the Pervasive Ultra-wideband Low Spectral Energy Radio Systems project
 RAN—radio access network
 RNC—radio network controller
 UE—user equipment
 WAN—wide area network

next few years [1]. With traffic increasing, and with bit rates expected to reach 10 to 20 Mb/s, it will be a challenge to provide enough bandwidth to accommodate the growing data and multimedia traffic by 2010.

The demand for increased bandwidth necessitates the use of a higher-frequency band that offers channels with bit rates ten times higher than those of 3G systems [22]. However, the use of higher frequencies means that free-space propagation loss is increased, with the loss depending on the frequency f as shown in the following equation [17]:

$$\frac{P_{or}}{P_t} = \frac{1}{(4\pi df/c)^2}, \quad (1)$$

where P_{or} is the received power, P_t is the transmitted power, c is the speed of light, and d is the distance from the transmitter. The effect of this loss is that, as data rates increase, cell sizes decrease (Table I [20]).

Thus, it is apparent that one of the effects of achieving adequate data rates for fourth-generation (4G) applications is a reduction in cell size. If the current architecture—generally a centralized, vertical

Table I. Comparison of cell size and data rates.

	Data Rates	Cell Size
Macrocellular (UMTS-FDD)	≈384 kb/s	≈3-5 km
Microcellular (UMTS-TDD)	≈1 Mb/s	≈1 km
Picocellular (Bluetooth, Hyperlan)	≈1-20 Mb/s	≈100 m

FDD—Frequency division duplex
TDD—Time division duplex
UMTS—Universal Mobile Telecommunications System

spanning tree architecture—is used, the resultant increase in the number of cells needed to cover an area poses the following problems:

- The cost of the network infrastructure and the cost of network planning and deployment increase substantially [22].
- There is a problem of scalability, because of the increased number of base stations needed to cover a given area.
- The increased number of base stations makes the task of network planning more complicated by making it more difficult to obtain optimum configurations.
- It takes much longer to install base stations, and space limitations and building regulations cause significant problems, especially in urban areas.

Besides the conventional cellular network, there are an increasing number of smaller networks comprising communicating devices. Such networks are made possible by the advent of cheap, integrated radio transmitters, such as those described in the Pervasive Ultra-wideband Low Spectral Energy Radio Systems (PULSERS) project [13] and in Bluetooth [6]. These transmitters are intended to enable the linking of various electronic devices so that they are able to communicate with each other and also possibly serve as part of the cellular network itself. The cost of these transmitters is expected to drop dramatically when they are mass produced. It is envisioned that these system-on-chip transceivers will be used (like micro-processors) in many everyday devices. The use of these devices in wireless personal area networks (PANs)

[4, 15], as well as in office and home networks, will require the ability to cope with dynamic ad hoc structures. The increasing ubiquity of these devices in everyday life will make the issue of scalability critical.

These challenges, among others, have prompted several studies to suggest the use of a new approach to the structure, control, and operation of wireless networks [3, 11, 30]. A small-cell environment, especially in large, highly loaded networks, favors the use of a horizontal, distributed architecture rather than a more conventional, vertical architecture. (**Figure 1** depicts these contrasting structures and highlights the function of the base station router (BSR) in the distributed architecture, where it assumes the combined functionality of the base station and its associated controller and packet interface.) Such a network architecture would have a structure like that of a wireless local area network (LAN), with a more localized form of control. The base stations in such a network would have to be low-powered, cheap devices [8]. They would often have an ad hoc, dynamic character—materializing, disappearing, modifying themselves, or communicating as the need and the opportunity arise. The state of the nodes might change from active to stand-by to disconnected during the functioning of the network. The network topology, hierarchy, and constituent nodes might change continuously. Base stations would also need to have the ability to self-configure, at least to some extent. The process of installing a base station would be simple and straightforward (i.e., plug and play), with the base station self-configuring different aspects of its operation, such as its cell size and routing.

An understanding of the advantages of a horizontal, distributed architecture in a small-cell environment can be obtained by considering current radio access networks (RANs), which are not optimized for high-speed micro-cellular networks [34]. Current RANs have a vertical tree structure, with tens of base stations connected to a radio network controller (RNC). During a handover, all signals are sent to the RNC and combined for diversity handover. (Diversity handover schemes are used to improve signal quality at cell edges by communicating with adjacent base stations simultaneously.) In 4G networks, with reduced cell size and

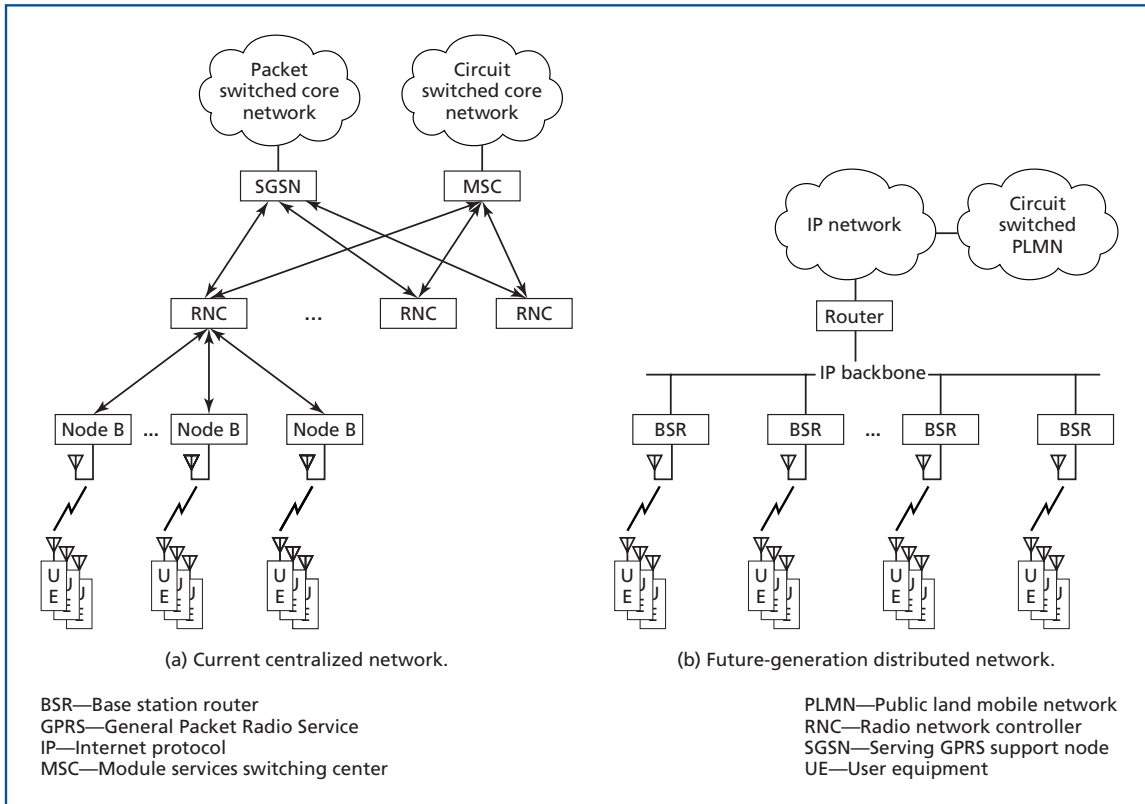


Figure 1.
High-level depiction of network structures.

reduced mobility, it is thought that still more frequent handovers will occur. This will place a high load on the entrance links and the signal processing equipment. A more horizontal, distributed approach, in which the processing of the handovers would be transferred from the RNC to the base stations, would dramatically reduce the loads on the entrance links and the RNC signal-processing equipment [34].

Distributed and Ad Hoc Architectures

A mobile ad hoc network (MANET) is an autonomous system of mobile nodes (and associated hosts) connected by wireless links, the union of which forms an arbitrary graph. The routers are free to move randomly and organize themselves arbitrarily; thus the network's wireless topology may change rapidly and unpredictably. Such a network may operate in a

standalone fashion, or may be connected to a larger network, such as the Internet. It can be used for such things as communications from office and home networks, battlefield networks in military operations, emergency networks, and PANs.

Research into the field of ad hoc networking for military applications has been going on for many years, but, recently, research in ad hoc networking in the commercial sector has increased, because of the growing importance of such networking in 4G systems. Because of the highly dynamic nature of ad hoc networks, a distributed approach to network control would provide many advantages over a centrally controlled approach. In ad hoc networks, all nodes are supposed to be identical, and they can leave the network at any time; hence there can be no one node that specializes in maintaining and coordinating the

network. Even if a node were dynamically assigned the task of controlling the network (i.e., becoming the center of a cluster), when it left the network, the task of control would have to be reassigned to another node. One way to control such a network without a central controller would be to make sure that all nodes have information on all the nodes in the network. While this would allow the network to function, the information at every node would have to be updated either periodically or dynamically. In a large network, this would result in high signaling traffic, and the issue of scalability would become a problem.

Hence there is a need for a different approach to controlling such dynamic networks. Such an approach would have to be highly distributed, with the nodes in the network acting on local information and information from nearby neighbors. Without complete knowledge of the network, the nodes would have to be able to adapt at both the physical and network layers to cope with changes in both traffic demand and network structure. This poses an obvious challenge, because incomplete knowledge may result in the nodes reacting wrongly, or not reacting at all, when a change is needed. Therefore, the main problem in implementing such a highly distributed approach is to determine how to make the network behave as desired, without having complete information about overall network status.

Self-Organizing Networks

Networks, such as those described above, that are able to adapt themselves to changes generally are called self-organizing networks. There has been little work on self-organization in networks [21], although recently a number of research activities have begun, particularly in the area of sensor networks [10, 23, 26, 28]. Furthermore, new paradigms for solving distributed problems are being investigated for their relevance to self-organization in ad hoc networks, among them ant-colony optimization [9] and swarm-intelligence [5]; both techniques are based on the behavior of natural systems consisting of many agents. These new approaches to the operation and control of networks should help to produce an intrinsically self-organizing network, in which the intelligence is inherent in the network's natural behavior.

Such intrinsic self-organizing behavior occurs in many different systems in many different fields. Examples include magnetic spin and crystallisation in physics; commodity markets in economics; and cell differentiation, population ecologies, and behaviors of social insects in biology. (Among of the most compatible systems—because its self-organizing behaviors can easily be transferred to wireless networks—are the simple mathematical models called cellular automata (CA) [25].) Such self-organizing systems have state evolutions that always evolve toward the same type of behavior or the same state, regardless of their initial configuration. These systems are made up of autonomous identical units that obey the same simple rules and have only basic knowledge of local neighbors. Despite the apparent simplicity of such a system, and despite the absence of a central coordinator in it, self-organizing behaviors emerge from the interactions among the units in the system. A 4G wireless network based on the common characteristics of these systems would have many advantages over a centrally controlled network.

Using the characteristics of such self-organizing systems for the coordination of wireless networks would allow the creation of a highly decentralized network, capable of self-configuration, having very low signaling overhead, and using very simple algorithms. These uncomplicated algorithms would make it possible to reduce the cost of network nodes significantly, an important consideration if wireless devices are to proliferate. Such a network would also have the flexibility, robustness, and scalability required to cope with the challenges mentioned above.

There are, however, several problems and challenges that can arise from the use of self-organization in wireless networks. For example, in self-organized systems, slight changes in the system parameters can bring about the onset of a phase transition, in which system behavior undergoes a sudden change. This phenomenon, called self-organized criticality, is very unpredictable, and its effect on a system is very dramatic [1]. Self-organized criticality has been observed in many different systems, including telecommunications networks [29]. Such behavior can be a problem, particularly when a network is working near critical

points, and a slight change may result in catastrophic failure. Nevertheless, the dangers of phase transitions in self-organized systems need not limit their use in wireless networks, because it is possible to minimize the risk of phase transitions occurring in a network by identifying the critical points of the network during its development and prior to its deployment. By keeping the range of system parameters within the safe operating region, the possibility of a phase transition occurring can be minimized. Networks operating close to critical points (i.e., just before a phase transition) are often operating under the most efficient conditions. For example, in the case of ad hoc network connectivity [16], there is a critical minimum transmit power required to achieve a high probability of connectivity. Similarly, in the case of network throughput in data networks [33], there is a critical traffic load, above which network throughput starts to decline. Ensuring that the network operates near a critical point, and avoids a phase transition, maximizes the performance of the network, but requires knowing which system parameter values to use and which to avoid.

When assessing the emergent behavior of these highly decentralized networks, behavioral information must be obtained not just for a single node, but for the whole network. Complexity metrics are able to provide this network-wide view by capturing the state space of the network as a whole; they can be used in two ways:

- To analyze the overall behavior of a network that is operating under a highly distributed algorithm. (This can help determine safe operating regions for the parameters or changes to the design of the algorithm, so critical phase transitions can be avoided.)
- To capture network states over a moving window online, in order to get advance warning of important behavioral changes.

In the rest of this paper, we introduce a method of analyzing the behavior of highly distributed networks by using a complexity metric. We then present a novel distributed algorithm for automatic cell dimensioning [14], as an example of how parts of a self-organizing network may be implemented. Next, we demonstrate

how the behavior of a network can be analyzed by applying the complexity metric to the algorithm. Finally, we describe a framework for using emergent behaviors to control a network.

The Complexity Metric

The term “complexity” is often used in the study of large, distributed systems. In this context, a system is defined as a collection of interacting, interconnected subsystems, each of which has a number of finite states, and the complexity of a system is defined as the minimal representation of its behavior. Various efforts at quantifying behavioral complexity have been made [7, 12, 19, 27], using different approaches to obtain the size of the minimal representation that describes system behavior. For example, in [7], a multilayered approach is proposed, incorporating a state transition model that is built up from the observed system state. However, less cumbersome approaches have also been developed, using the entropy of the system [19, 27], S :

$$S = - \sum_{i=1}^N p_i \log_{10} p_i, \quad \sum_{i=1}^N p_i = 1, \quad (2)$$

where p_i is the probability that the system is in state i of N possible states.

The complexity metric we shall use, which is derived from the system entropy, is defined as:

$$C = \frac{4S}{S_{\max}} \left(1 - \frac{S}{S_{\max}} \right), \quad (3)$$

where S is the entropy of the system, and S_{\max} is the maximum possible entropy of the system that occurs when all states are equiprobable, i.e., when $p_i = \frac{1}{N} \forall i$ and is given by $\log_{10} N$. This metric is similar to those used in [19] and [27], but it does not use the disequilibrium quantity used in [19], and it does not include the use of parameters to bias the complexity metric, as in [27].

The complexity metric provides us with a summary of the information on the network as a whole, but it uses only simplistic information on the state of the individual nodes. The generality of the metric means that it can be applied to various aspects of the

network, such as network resources (e.g., bandwidth and buffer space) and network operation (e.g., routing). The application of the metric is also fairly straightforward, once the state of the node has been defined.

The Cell-Dimensioning Algorithm

To exemplify some of the applications of the complexity metric in a decentralized network, we apply it to a novel distributed algorithm for automatic cell-dimensioning. The network that is considered is one similar to the one described in [24], using the BSR implementation. The algorithm uses only local information and simple local rules at each BSR.

The goal of the BSR project is to develop a platform capable of supporting an economical 4G architecture that can run 4G services. The architecture introduces the use of BSRs—low-powered devices operating over multiple platforms—that are structured like a wireless LAN (Figure 2). The BSR units are expected to have functionality (i.e., a simple, plug-and-play method of installation) similar to that of wireless LAN access points.

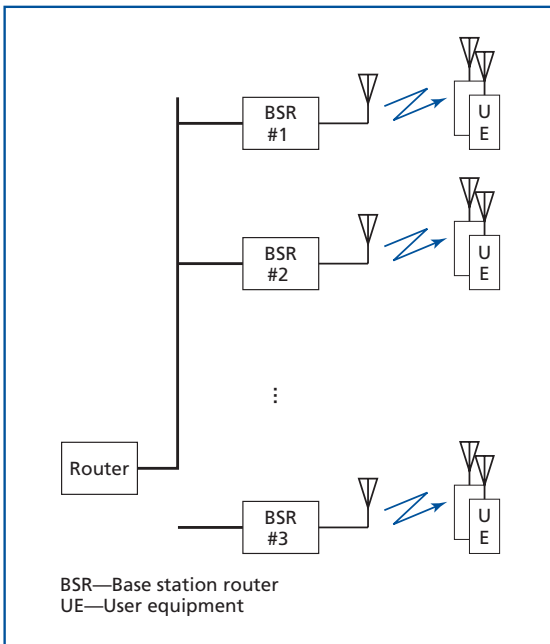


Figure 2. Logical description of BSR network.

The BSR approach differs from more conventional schemes, which have less flexible, centralized structures, with controlling elements in the network, such as the base station controllers (BSCs) or the mobile switching centers (MSCs) that are used in the Global System for Mobile Communications (GSM) network. The proposed BSR architecture has a flat hierarchical structure; therefore, it is a good candidate for distributed self-configuration. Indeed, one of the research areas highlighted in [24] is automatic configuration. A scenario mentioned in [24] seeks to reduce the number of elements required in the network, and allows the network to grow in an ad hoc fashion, which implies that the configuration and deployment of the base stations should be truly plug-and-play.

The main goal of the cell-dimensioning algorithm is to make it as easy as possible for the user to set up the system. Accordingly, the user need only:

1. Mount the BSR in the desired location and make the appropriate connections before powering up.
2. Place another BSR at an estimated rough minimum distance from the other BSR units.
3. Repeat the process throughout the area intended for coverage.

The user does not have to know the transmit power of the BSR, nor worry much about positioning it, because the algorithm is intended for use in a microcell or picocell environment, and the cells are fairly small.

In the distributed automatic cell-dimensioning algorithm, the base station is treated as an independent and totally autonomous entity that has three states, S :

$$S = \{A, I, C\}, \quad (4)$$

where A is the active state, I is the inactive state, and C is the configuring state. The state transition scenarios are illustrated in Figure 3. The BSR starts out in state C upon power-up. Upon completion of the configuration process, the BSR enters state I , in which it is inactive and ignores the state of its neighbors for a set period of time. Once that period is over, the base station goes into state A . The base station then remains in state A until one of its neighbors is in state C , at which point it, too, enters state C .

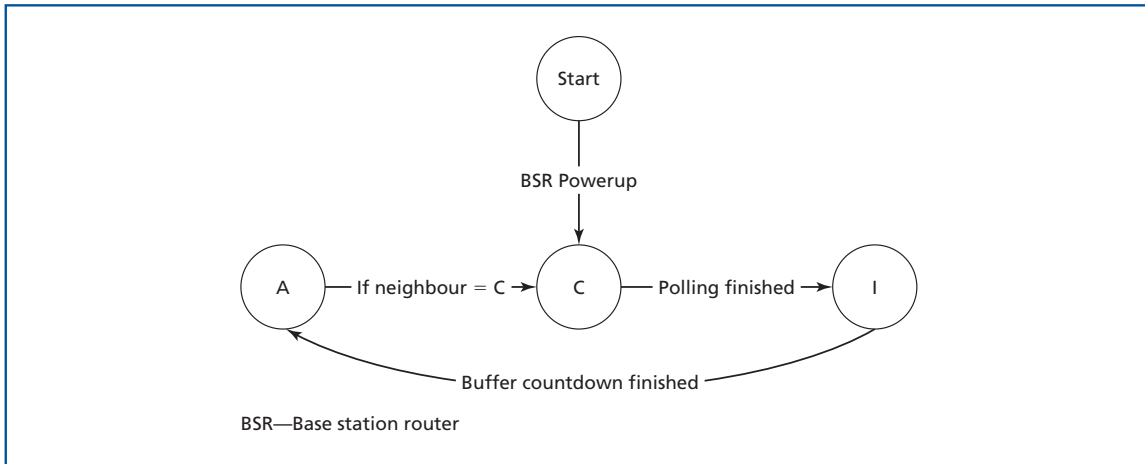


Figure 3.
State transition diagram of BSR using the algorithm.

Description of the Configuring State

The algorithm works by gathering information about the approximate positions of its neighbors, in order to determine its cell size. It does this during state *C*, the configuring state. While in this state, each base station creates a list containing the distances of all its neighbors (i.e., all base stations within its maximum range). It creates the list by polling its neighbors and receiving information from them. In the process of polling its neighbors, the BSR in the contacting cell gradually increases its cell size until it reaches its maximum.

The process of increasing the cell size requires careful implementation, with special consideration given to the power-control aspects of the network. In a code division multiple access (CDMA) system, power control has a significant impact on the performance and capacity of the network [18]. Similarly, in an orthogonal frequency division multiplexing (OFDM) system (which is used in high-rate wireless LANs [31] and is also a possible candidate for use in 4G systems), power control is needed to cope with the large peak-to-average power ratio [2]. For these reasons, a method of dynamically changing cell size without significantly affecting power control is required.

The signal transmitted by the cell during the configuring state contains timing information and the identification of the transmitting cell. Neighboring cells,

detecting the presence of the signal, send out a response to the transmitting cell over the Internet protocol (IP) backbone link between the two BSRs. The response signal contains the identification of the responding cell and the time at which it received the signal from the transmitting cell. The transmitting cell thus collects information about the base stations located within its range, and, based on the time each base station received its signal, is able to approximate its distance from that base station. The transmitting cell then calculates and determines its final cell size, using the information contained in the list it has compiled. The flowchart describing this process is shown in **Figure 4**.

This discovery process is similar to that used in wide area networks (WANs). However, in WANs, the discovery process can be done over the IP backbone, using established procedures to establish logical links between the nodes. But in this case, because the physical distances between the nodes are involved, the discovery process must be done over the air.

The discovery process we have described above works only if just two cells are involved. If there are more than two cells, the process may produce erroneous results, as illustrated in **Figure 5**. The error occurs when cells A and B adjust their cell sizes based on the distance between them. Meanwhile, a third cell, C, which has cell B as its nearest neighbor, adjusts its cell size based on the distance between them. This

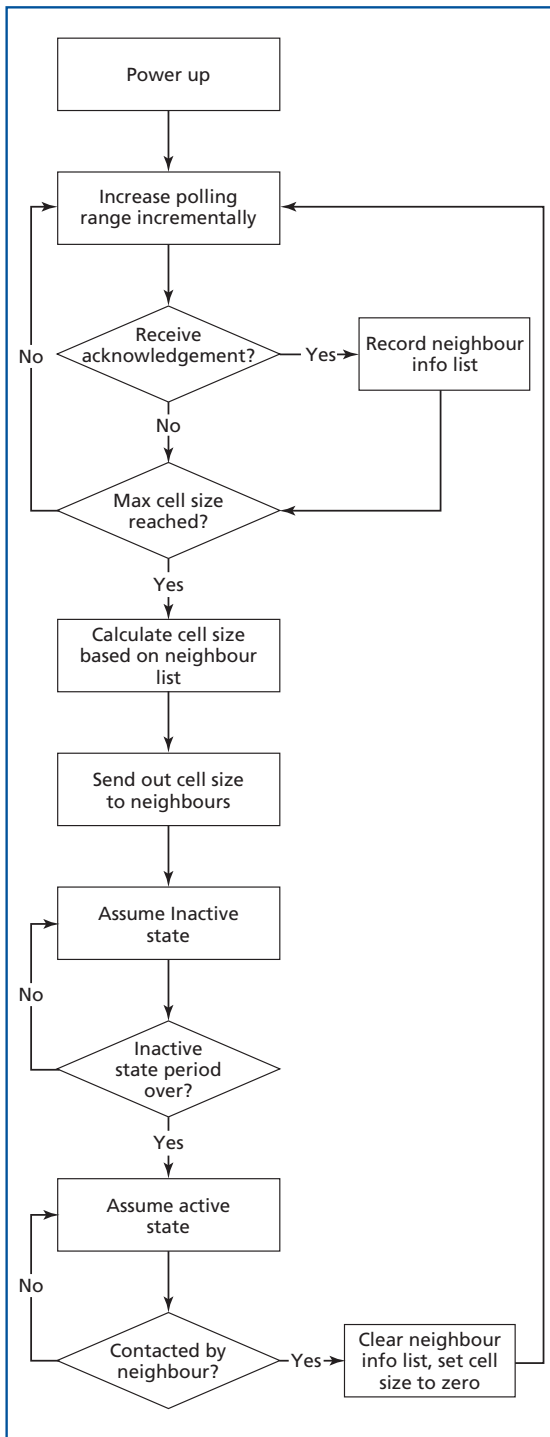


Figure 4.
Flowchart of configuration process.

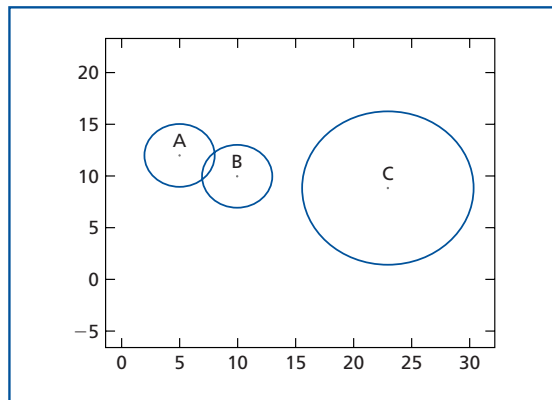


Figure 5.
Cell boundaries without correction.

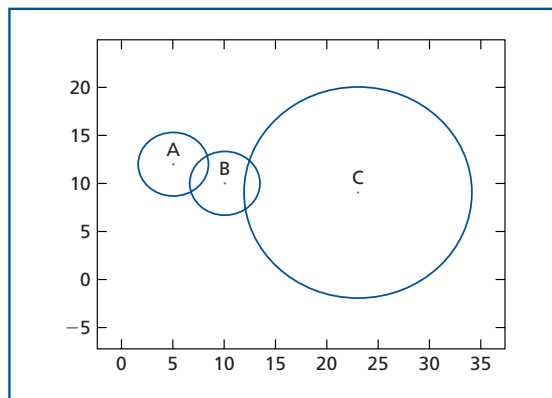


Figure 6.
Cell boundaries with correction.

causes an error, because, while cell C is configured to fit the distance between itself and cell B, cell B is not. To avoid this error, whenever a base station completes its initial polling state, it sends out its final cell size to its neighbours. Upon receiving this information, the neighbours update their information lists and make the appropriate adjustments. The configuration that results from this additional procedure is shown in **Figure 6**. The additional step ensures that a base station is able to realize that the calculations it has made, based on its distance from its nearest neighbor, have not been reciprocated, because its neighbor has based its calculations on its distance from a different base station, which is closer to it.

The base station also checks to see if a neighbor has been placed too close to it. If it has, then the base station waits a random period of time before shutting down and sending a message to its neighbors to clear their lists and begin the polling process again. (This pushes the neighbors into state *C*.) This feature was added to the algorithm because it was assumed that the placement of the base stations would be done in a relatively unplanned manner, and hence some base stations would be too close to others.

In shutdown mode, the base station waits a random period of time before initiating the shutdown process by sending a reset signal to all its neighbors. The reset signal triggers its neighbors to clear their neighborhood lists and go into polling mode. The neighbors then begin compiling a new list that does not include the base station that is shutting down. After it has sent out the reset signal, the base station powers down and stops transmitting, effectively removing itself from the network.

The ability of the algorithm to produce good overall coverage is determined by the placement of the base stations. Unless the base stations are perfectly evenly spaced over an area, there will be gaps in the coverage. To prevent the occurrence of very large gaps, care should be taken when placing the base stations. For example, placing groups of base stations close together in clusters results in inadequate coverage in areas between the clusters. To minimize this effect, the BSR is made to check its cell size and proximity to its neighbors. A base station that is located too close to its neighbor will activate its shutdown sequence. This feature is intended to make possible a flexible, plug-and-play, ad hoc strategy of node deployment and addition.

Calculation of the Distances between Neighbors

When a base station is polling its neighbors, it keeps track of its transmit power and of the time at which it was transmitting at that power level. The polling signal that it broadcasts contains identification information and timing information. When a neighbor detects this signal, it sends out a reply signal containing the timing information that it has received from the polling signal. The polling base station, upon receiving the reply, knows at what power level the transmitted

signal reached the neighbor, and thus can estimate—using a suitable propagation model, such as the one described in [35]—the distance between them.

Algorithm Simulation Results

The base stations are laid out over an area of 10 kilometers by 10 kilometers, with varying densities. Most of the base stations are laid out according to a grid, placed 1200 meters apart, but with a maximum deviation of ± 5 meters from the grid, so that their placements are slightly imperfect. In certain places in the middle of the area, the base stations are placed not according to the grid, but closer together or further apart. This arrangement is intended to reflect a placement procedure in which base stations are placed reasonable distances apart, but no exact pattern is followed. The base stations are then powered up at different times, in a random sequence.

Once the simulation has started, there is a setting-up period—during which all the base stations begin collecting and broadcasting information—before all the base stations settle down to stable mode. The results of the experiment, which are depicted in **Figure 7**, show that the algorithm is able to achieve coverage without causing too much interference (i.e., the handover boundaries of the cells do not encroach on those of their neighbors by more than a specified amount).

Another simulation was run with five base stations (see **Figure 8**) placed in the arrangement shown in Figure 8a. The base stations were then powered up and allowed to settle down before the base station in the middle (labeled *X*) was shut down. While shutting down, base station *X* sent out a reset signal to its neighbors, prompting them to empty their neighborhood lists and to enter into polling mode. Figure 8b shows the final configuration of the four remaining base stations. This result demonstrates the ability of the algorithm to make changes to compensate for the removal of a base station from the network.

Figure 9 shows the result of a simulation in which two base stations were added to an existing network of six base stations. The six base stations were placed approximately 200 meters apart in a grid, but with a random deviation of ± 10 meters from the

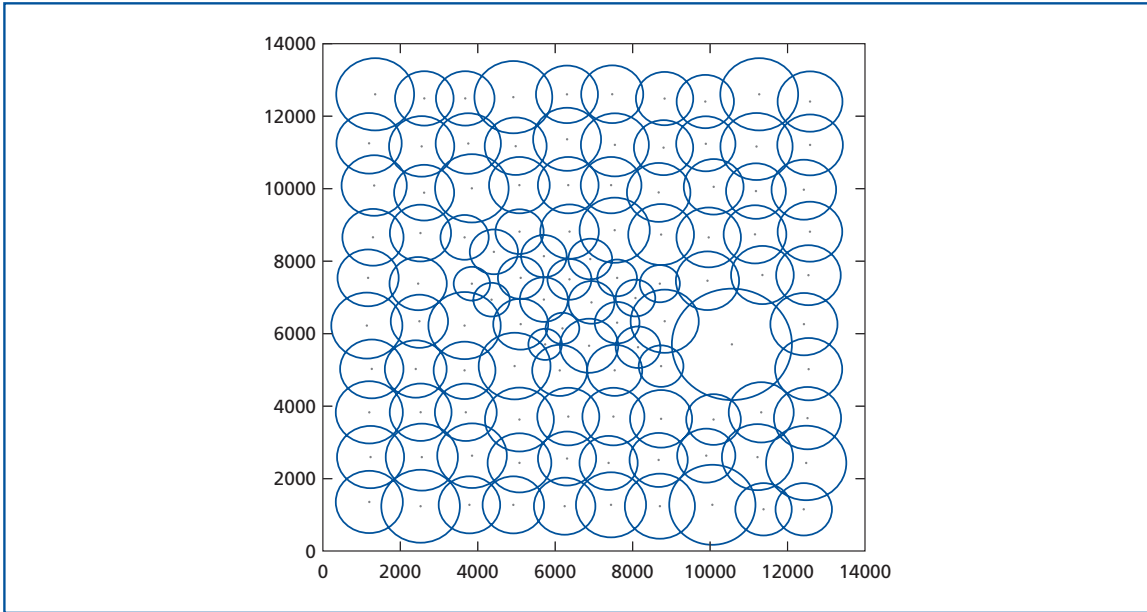


Figure 7.
Resulting configuration of network with base stations placed at different densities.

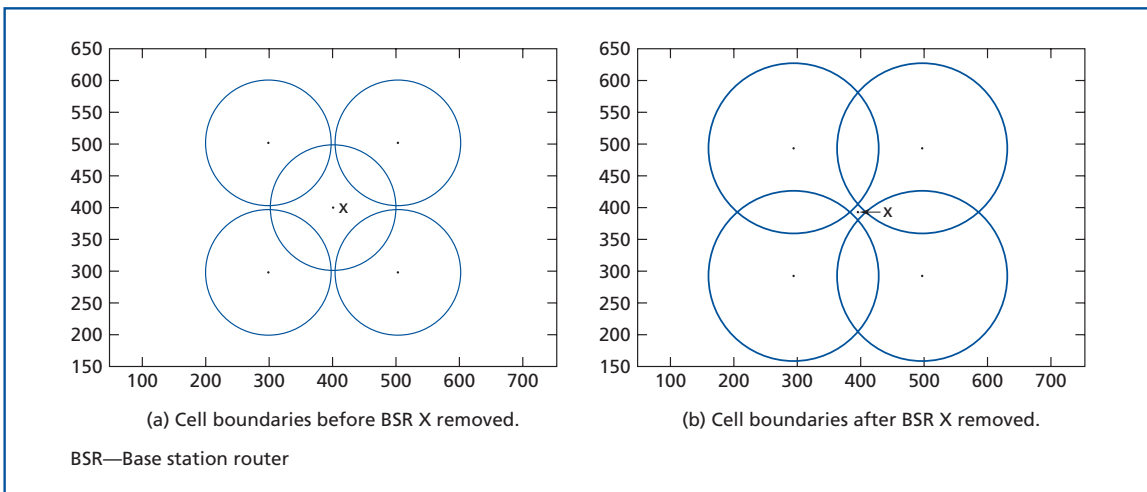


Figure 8.
Simulation with removal of BSR.

grid. The two base stations were added one after the other, at random times. The resulting configurations demonstrate the ability of the algorithm to make changes to compensate for the addition of new base stations to the network.

The process of reconfiguring a network—adding BSRs to it or removing BSRs from it—while it is in operation has an impact on both the performance and the capacity of the network. (During a reconfiguration, the BSRs must increase the size of their cells until

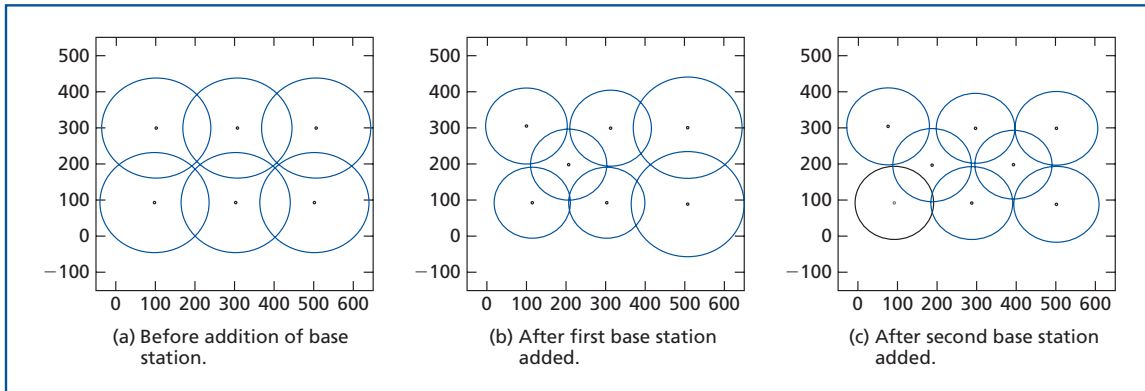


Figure 9.
Changes in cell boundaries when base stations are added.

they significantly overlap neighboring cells, and this naturally increases interference.) Therefore, if possible, the reconfiguration of a network should be done during off-peak periods (e.g., late at night), when the impact on users is minimal. However, if network reconfiguration must be done during peak periods (e.g., because of an emergency), the disruption will only be temporary, and it will be over as soon as the network completes its reconfiguration cycle.

Post-Deployment Optimization

One of the problems of the autoconfiguration process is the occurrence of gaps in coverage. Although perfect coverage is achievable if the placement of the BSRs is optimal, such placement is not possible in real-world situations. Furthermore, the purpose of autoconfiguration is to make the installation of a network as simple as possible, which inevitably results in non-optimal BSR placements. Therefore, a second stage has been added to the configuration process to deal with this problem. This stage involves using feedback from the mobiles or user equipment (UE) to detect gaps in coverage.

During this second stage, the BSR keeps track of the UEs that are connected to it. Each UE monitors the signal it receives from the BSR that it is connected to. When the signal begins to go below a predetermined threshold, and the UE cannot find a signal from a neighboring BSR, the UE sends out a signal to the base station to which it is connected to indicate a

possible gap in coverage. When a UE reports a possible gap in coverage, the BSR to which it is connected increases its cell size by an increment. Periodically, the BSR checks the status of the UEs in its cell and increases its cell size by an amount that depends on how many mobiles have reported gaps in coverage. The cell size is increased by a factor of F :

$$F = ne^{-2d}, \quad (5)$$

where n is the number of UEs that have reported coverage gaps, and d is the difference between the current cell size and the cell size that was established during the initial deployment stage. This factor ensures that a given BSR does not increase its cell size too much, and that both the BSR and its neighbors increase their cell sizes to cover the gap.

This process was simulated in a network containing 100 BSR units placed in a loose grid with a deviation of ± 250 meters from the grid. This placement was intended to reflect a placement procedure that did not require extensive planning beforehand. The initial deployment stage was then initiated, resulting in the first configuration shown in **Figure 10**, in which gaps in coverage exist. Five hundred UEs were then placed randomly in the network, each moving with a mean velocity of 5 meters per second and sending back reports to the BSR every second. The simulation was then run for 10,000 seconds. The result—obtained after UE feedback—was the second configuration shown in Figure 10, in which gaps in

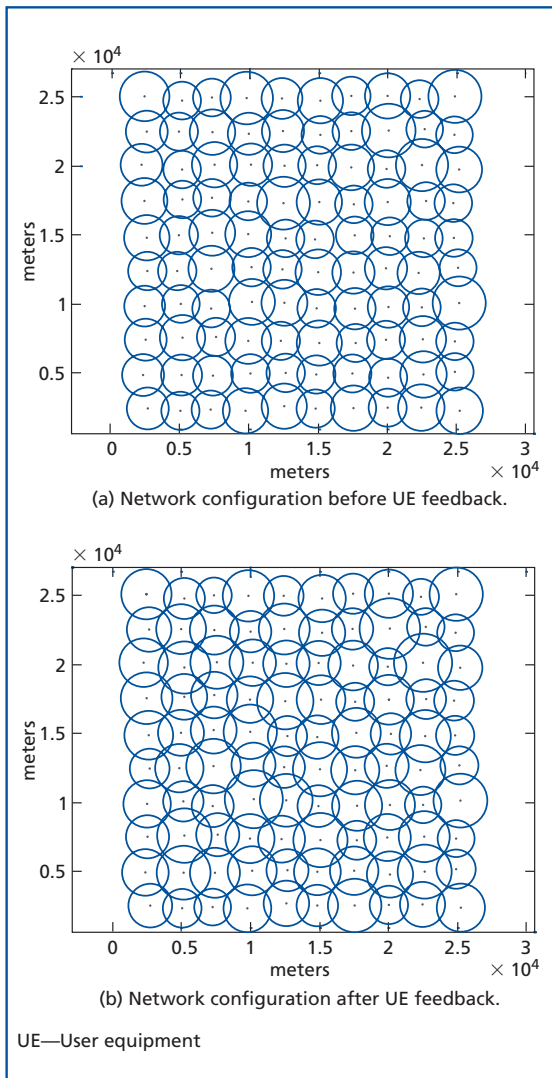


Figure 10. Simulation showing effect of UE feedback on network configuration.

coverage no longer exist. Also, as shown in **Figure 11**, the number of UEs dropped because of lack of coverage was reduced to zero.

Cost-Benefit Analysis of the Algorithm

The use of a cell-dimensioning algorithm in the process of deployment greatly simplifies the task of cell planning. The conventional method involves making a survey of the intended coverage area and calculating

suitable placements for the base stations prior to the deployment of a network. After the placements have been established, they are tested, either by making onsite measurements or by using simulation software. Any changes that must be made to the network (e.g., the addition, removal, or relocation of base stations) require that this involved process be repeated to reconfigure the network nodes. Considerable resources are required to perform this labor-intensive exercise, and the amount of time required is also substantial. In a network with a large number of base stations (e.g., the anticipated 4G networks), the cost of network configuration and reconfiguration can be very high.

The cell-dimensioning algorithm is designed to transfer this task from the network designers to the network nodes. Since the algorithm is not computationally intensive, it can be implemented fairly easily, because it does not involve extensive communication among the nodes and it reduces the amount of computational power required at each base station. The highly distributed nature of the algorithm also obviates the need for costly, specialized network elements when implementing the autoconfiguration aspect of the network. The simple, plug-and-play process we have outlined above (which eliminates the need for highly skilled labor) reduces both the cost and the time required for network planning, deployment, and maintenance.

Application of the Complexity Metric to the Algorithm

One of the parameters that we will examine using the complexity metric is the period the BSR remains in the inactive state (I), or the back-off time. The back-off time should be set to be as short as possible, so that the base station will be as receptive as possible to either additions or removals of neighboring base stations. Setting the back-off time to be long is an easy way to avoid the problem of two base stations continuously triggering each other off, but it increases the likelihood that the base station will not be able to detect any changes made to the network. The back-off parameter has a significant impact on the behavior of the network, but, in a large network, it is difficult to judge its precise effect. We examine the effect here by using the

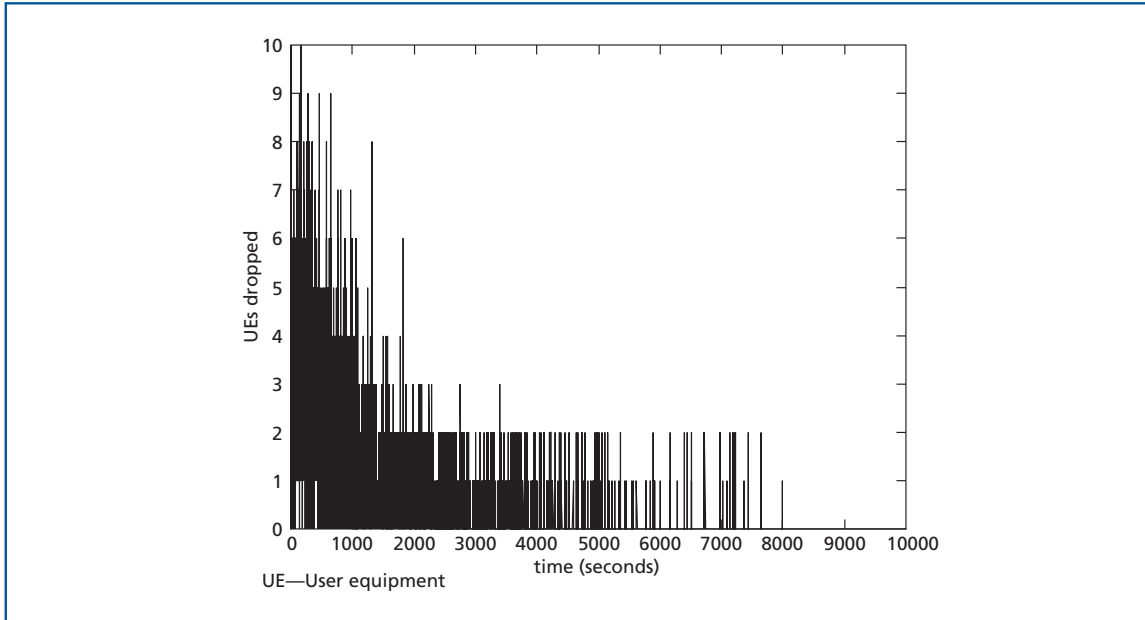


Figure 11.
Number of UEs dropped over time because of gaps in coverage.

complexity metric, which has the advantage of obtaining a network-wide view of the behavior of the network in a straightforward manner, by using simplistic state information from the nodes. The 100 base stations in the network model that we use in the simulations are loosely arranged in a 10 by 10 grid at approximately 200-meter intervals, and are powered on at randomly set times. The base stations are recorded as being in one of three states: polling, idle, and frozen.

When calculating the state probabilities, we record the state of the base stations at fixed intervals, and calculate the number of times the network has been in a certain state. The state probability is then calculated by dividing the number of times the network is in a certain state by the total number of times the system state has been recorded. With three states and 100 base stations, there are 3^{100} , or approximately 10^{47} , system states. Because of this huge number of possible states, we introduce a state definition filter that captures the basic trend of network behavior, rather than the raw states. The filter removes any slight fluctuations in the system state when we

are recording network behavior, by maintaining a weighted moving average of the frequency of the state space of the network. We use this filter because we are interested only in the basic overall trend of network behavior, and do not want to consider insignificant changes in system behavior.

To filter out the effects of these slight system changes, the exponentially weighted moving average (EWMA) of the network state is used when recording the state of the system. The EWMA is calculated using the formula:

$$A_{jt} = k(M_{jt} - A_{j(t-1)}) + A_{j(t-1)}, \quad (6)$$

where A_{jt} is the EWMA of one of the possible states of base station j at time t , and M_{jt} is a vector composed of ones and zeros that reflect the occurrence of one of the states at base station j at time t . (For example, the value of M_{jt} would be 1 if base station j was currently in the polling state, and 0 if it was not.) k is the smoothing constant, $k = \frac{2}{1+N}$, and N is the period in the EWMA (i.e., how many previous values of M_{jt} are

considered). Using a high value of N increases the amount of filtering.

The EWMA basically provides a record of the network state with a certain amount of memory. When applying this filter to the states recorded by the cell-dimensioning algorithm, three separate EWMA's are maintained, one for each state. The values of the three EWMA's are then compared to determine which state the base stations are in, and the state having the highest EWMA is selected.

Results of Complexity Measurement

Figure 12, which shows the complexity measured for different back-off times, displays three distinct operating regions: a region in which complexity is zero, when the back-off time is between 5 and 14.8 seconds; a region in which there is an increase and subsequent decrease in complexity, when the back-off time is between 14.8 and 24.9 seconds; and a region of zero complexity, when the back-off time is above 24.9 seconds. These results help us make decisions on setting the parameters. In this case, the cell-dimensioning algorithm was iterated 3000 times for each back-off time value; back-off times ranged from 5 to 37 seconds in 0.1 second increments. The results indicate that if the back-off time is not set long enough, the network tends to behave randomly and never self-organizes. Specifically,

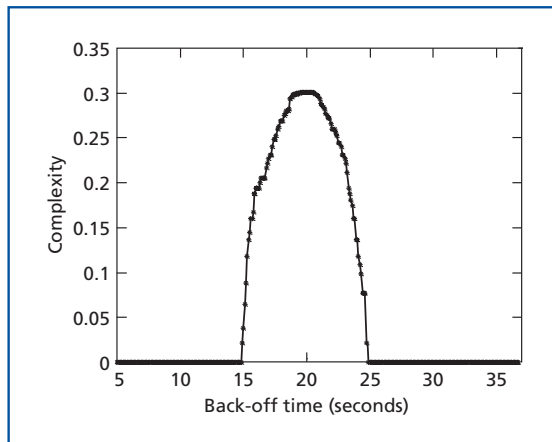


Figure 12.
Relationship between complexity of network behaviour and back-off time.

when the back-off parameter is set to a value below 14.8 seconds, the network never self-organizes; it remains in a random or periodic state. But when the back-off parameter is set to a value above 14.8 seconds, the network begins to exhibit self-organizing behavior, and the chance of its reaching a self-organized state increases as the back-off time is increased. Finally, when the back-off time is set to a value above 24.9 seconds, the network always becomes self-organized.

Figure 13 shows the use of online measurements. Here the data are collected through a time window and displayed as the algorithm is running. Because the moving window restricts the number of states captured, the maximum achievable entropy, S_{\max} , is also limited; it is dependent on the size of the moving window used.

With the online measurements, the network nodes are able to get some form of advance warning of unexpected behavior. When the state space starts to increase quickly (reflected in a jump in the metric), it indicates the start of a phase transition. Figure 13 shows the online measurements, displaying different behaviors when the back-off time is changed to show the effects of the critical points obtained in Figure 12. The three different behaviors are observed when the back-off time is set to 15, 20, and 25 seconds. At 15 seconds, the network exhibits chaotic behavior, while at 20 seconds, the network begins to self-organize, but isolated random behavior starts to ripple out to the whole network, causing periodic behavior. Finally, at 25 seconds, all the network nodes are able to organize themselves at the first try, with no undesirable interruptions. Figure 13 also shows the cell boundaries resulting from the cell-dimensioning algorithm. It is important to note that, with back-off times of 15 and 20 seconds, the boundaries are still pulsing, and only snapshots of the cell-boundaries are shown, while with a back-off time of 25 seconds, the cell boundaries are static.

Figure 14 gives an indication of how fast the network is able to adapt to changes in topology. The online measurements illustrate how quickly the network goes back to its organized behavior. The first hump in the graph shows the initial self-organization that occurs when the base stations are switched on, before

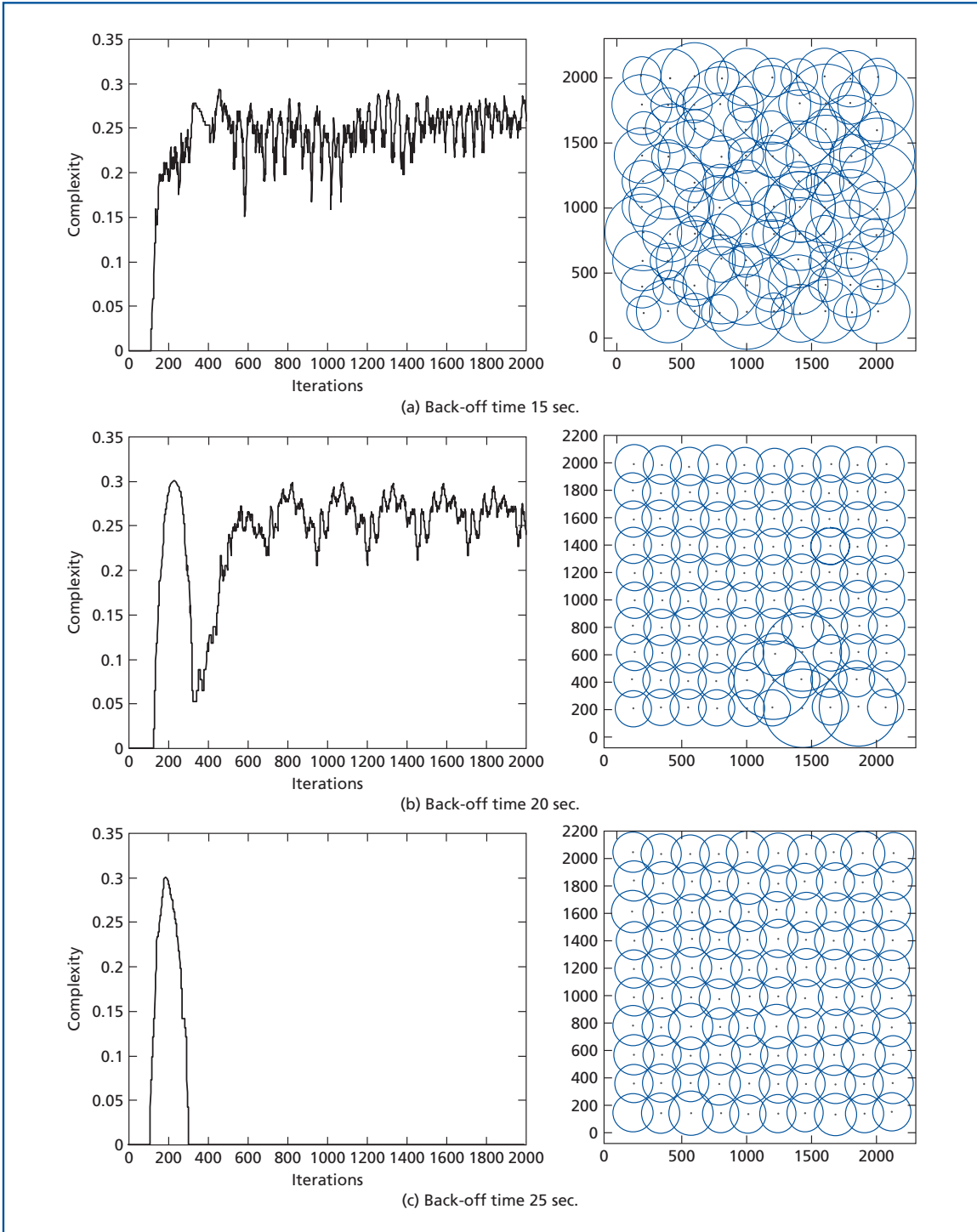


Figure 13. Online complexity measurement graphs and the corresponding cell boundary diagrams for different back-off times.

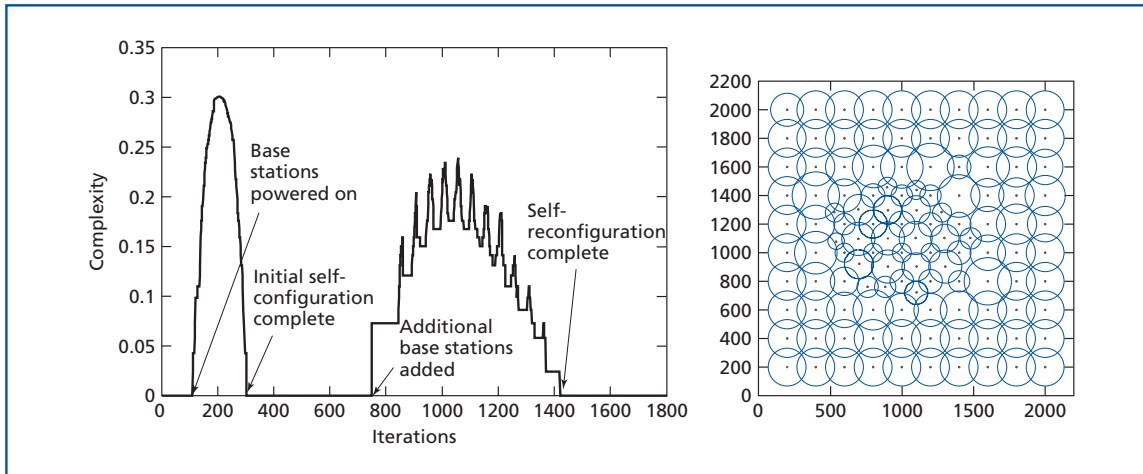


Figure 14. Effect of the addition of new base stations on complexity measurements and configuration.

they reach the desired state. At 750 iterations, base stations are added to the network (e.g., to cope with hot spots in demand, or to increase coverage), triggering the existing base stations to change their configurations. The graph clearly shows the dynamics of the reconfiguration, while the cell boundary diagram shows the resulting configuration of the cell boundaries with the added base stations.

Generalizing the Emergent Behavior Approach to Control

As stated earlier, a network that operated on the basis of emergent self-organizing behavior and used an organic approach would be both low-cost and highly flexible. However, to make the transition from current network control methods requires some kind of framework. This section provides such a framework.

An obvious method of bringing about self-organizing behavior in a wireless network is to view and design the network in such a way that it has characteristics similar to those of well-known self-organizing systems. One of the most compatible and convenient of self-organizing systems is CA. CA are abstract models that have been used to study and model many different aspects of the physical world [25]. The simplest description of a CA is a lattice, with each lattice site, or cell, having a number of finite

states. The array can have many different dimensions, ranging from the simplest one-dimensional array to a three-dimensional matrix. Time is discrete, and at each clock tick the cells change state. The new state is completely determined by the present states of the cell and its neighbors. The function—called the local rule—that determines the change of state is the same for all cells:

$$s_i(t + 1) = f(s_{i-r}(t), s_{i-r+1}(t), \dots, s_i(t), \dots, s_{i+r}(t)), \quad (7)$$

where the state, s_i , in cell i is changed according to a rule that depends on the states in a finite neighborhood of cell i . The local rules are generally simple, and usually involve logical functions (e.g., AND, OR, and XOR).

A well-studied type of CA is the simple one-dimensional CA, an example of which is shown in **Figure 15**. Here, in the rule table, neighborhood 111 means that the current state is 1, and that the state of the neighbors on either side is also 1. The output bit of one associated with neighborhood 111 means that, at the next clock tick, the state will remain 1. While the local rules of the one-dimensional CA are simple, the global behavior of the CA that results from the application of these rules is quite complex. CA are capable of displaying a wide array of behaviors, which

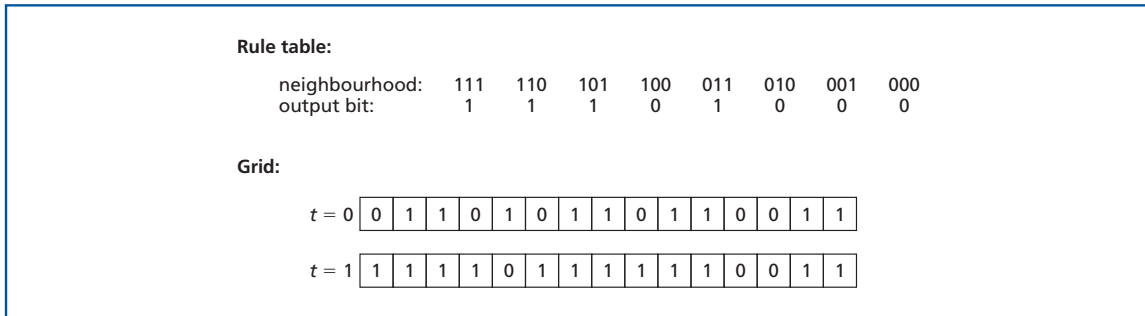


Figure 15.
One-dimensional, three-neighborhood, binary cellular automata.

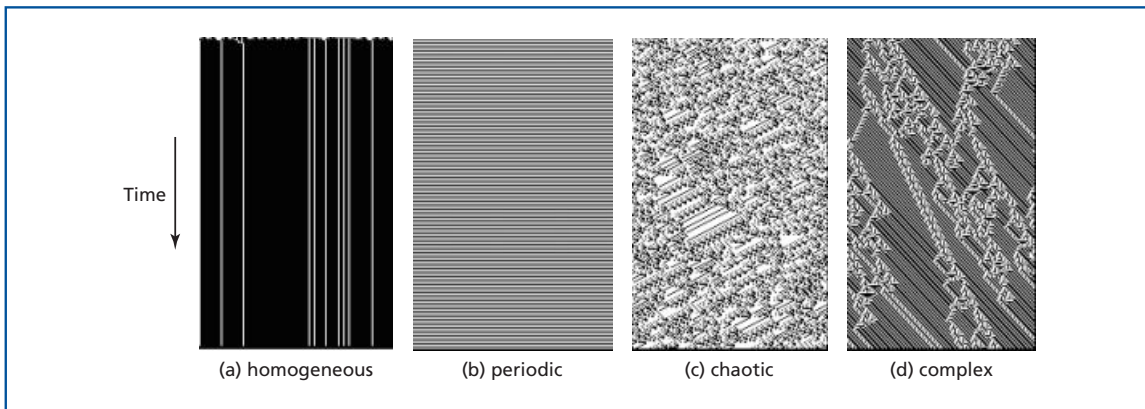


Figure 16.
Time evolution of a one-dimensional CA, with four different rules causing four different types of behavior.

have been grouped into four general classes (see **Figure 16**) [32]:

- Homogeneous state,
- Separated simple, stable, or periodic structures,
- Chaotic patterns, and
- Complex localized structures.

Which class the CA will be in depends on the local rules of the system, so the CA will always evolve to exhibit the same behavior type no matter what the initial starting configuration is. This self-organizing property can be beneficial, if applied to the problem of autoconfiguration in wireless networks. Each BSR can be viewed as a cell in the CA, having a finite number of states, interacting only with the BSRs in its limited neighborhood, and operating on a set of local rules. Depending on the rules that we apply, the behavior

toward which the network will evolve—no matter what the initial configuration is—can be homogeneous, periodic, chaotic, or complex. Different layers of CA can be used for different aspects of the network. For example, there may be a CA with its own set of rules and states that deals with routing, and another that deals with cell-dimensioning.

Using such self-organizing algorithms can be limiting, because one rule will tend to produce just one behavior, as illustrated by the CA classes [32]. However, it is possible to trigger different emergent behaviors in the network simply by changing the rules obeyed by the nodes. In such a scheme, each node would have a library of predetermined rules that are known to trigger different emergent behaviors. The rules could then be selected and applied to obtain

a variety of emergent behaviors during the operation of the network. For example, during an initial configuration during network deployment, a rule could be applied to trigger the network to exhibit random or periodic behavior while searching for an appropriate configuration. When the search is over, a different rule could be applied to bring the network back to a static, self-organized state. Thus, there could be different rules to be used in different scenarios, and they could be triggered whenever they were needed. This multi-rule approach would provide more flexibility in the manipulation of network behavior than could be obtained if only one self-organizing algorithm were used.

A simple application of the complexity metric to such a multirule network could enable a node to gather information in order to determine which rule in its library is needed, and also to check that the desired behavior has emerged. The development of a more distributed form of the complexity metric (e.g., a form in which the complexity is calculated locally at each node, using the states of its neighboring nodes) would result in a coupling of the simple control algorithm libraries with the metric, used as a monitoring method that triggers the invocation of different behaviors in the network.

Discussion and Conclusions

In this paper, we have emphasized the importance of the use of self-organization in 4G networks, because, in order to make 4G systems flexible, we need to move away from centralized architectures and control methodologies. To achieve a high degree of decentralization, we have had to look at structures which naturally exhibit decentralized properties, and we have found that CA are such structures. CAs show potential as analogues for self-contained 4G network elements such as BSRs; an association can be made between the automatic configuration characteristics of the network element (i.e., the BSR) and the behavioral characteristics of a CA. We have modelled the radio autoconfigurability of BSRs as CAs, and, in order to understand the global interactions of these structures, have developed a metric that is based on the system entropy and found that this metric gives forewarning of unwanted system behavior in the form

of unwanted phase transitions. Finally, we have proposed the notion of behavioral libraries for the decentralized control of cellular automatized BSR architectures. The emergent self-organizing behavioral techniques that this paper has investigated show promise in decentralized small-cell architectures that are being proposed for 4G networks.

References

- [1] P. Bak, C. Tang, and K. Wiesenfeld, "Self-Organized Criticality," *Physical Review A (General Physics)*, 38:1 (1988), 364–374.
- [2] A. N. Barreto, "Controlling the Peak-to-Average Power Ratio in OFDM-based Wireless LAN Systems," *Proc. 55th IEEE Vehicular Tech. Conf. (Birmingham, AL, 2002)*, Vol. 2, pp. 762–766.
- [3] A. Bira, F. Gessler, O. Queseth, R. Stridh, M. Unbehaun, J. Wu, and J. Zander, "4th-Generation Wireless Infrastructures: Scenarios and Research Challenges," *IEEE Personal Commun.*, 8:6 (2001), 25–31.
- [4] C. Bisdikian, P. Bhogat, and N. Golmie, "Wireless Personal Area Networks," *IEEE Network*, 15:5 (2001), 10–11.
- [5] E. Bonabeau, M. Dorigo, and G. Theraulaz, "Swarm Intelligence: From Natural to Artificial Systems," *Santa Fe Studies in the Sciences of Complexity*, Oxford University Press, New York, 1999.
- [6] B. Chatschik, "An Overview of the Bluetooth Wireless Technology," *IEEE Commun. Mag.*, 39:12 (2001), 86–94.
- [7] J. P. Crutchfield, "The Calculi of Emergence: Computation, Dynamics and Induction," *Physica D*, 75:1–2 (1994), 11–54.
- [8] S. H. de Groot, I. G. M. M. Niemegeers, D. R. Bueno, and R. Vlssers, "Self-Organizing Networks," *Contributions to the World Wireless Research Forum Book of Visions 2001*, <http://www.wireless-world-research.org/BoV1.0/Contributions%2002/02_3_09-self-organizing-networks.pdf>.
- [9] M. Dorigo, V. Maniezzo, and A. Colomi, "The Ant System: Optimization by a Colony of Cooperating Agents," *IEEE Trans. on Systems, Man, and Cybernetics-Part B*, 26:1 (1996), 29–41.
- [10] ESP project at Personal Information Architecture Group MIT, <<http://www.media.mit.edu/pia/Research/ESP/index.html>>.

- [11] B. G. Evans and K. Baughan, "Visions of 4G," *Electronics and Commun. J.*, 12:6 (2000), 293–303.
- [12] P. Grasberger, "Toward a Quantitative Theory of Self-Generated Complexity," *International J. of Theoretical Physics*, 25:9 (1986), 907–938.
- [13] W. Hirt and D. Porcino, "Pervasive Ultra-wideband Low Spectral Energy Radio Systems (PULSERS)," *Proc. Wireless World Research Forum Sixth Meeting (London, United Kingdom, 2002)*.
- [14] L. T. W. Ho, L. G. Samuel, and J. M. Pitts, "A Method of Selecting Maximum Transmission Power Level to Be Used by a Radio Telecommunications Base Station or Base Stations in a Network, a Radio Telecommunications Base Station and a Radio Telecommunications Network," *European Patent Pending, Application No. 02254380.5, June 24, 2002*.
- [15] R. Jordan and C. T. Abdallah, "Wireless Communications and Networking: An Overview," *IEEE Antennas and Propagation Mag.*, 44:1 (2002), 185–193.
- [16] B. Krishmanachari, S. B. Wicker, R. Béjar, and M. Pearlman, "Phase Transition Phenomena in Wireless Ad-Hoc Networks," *Symposium on Ad-Hoc Wireless Networks (San Antonio, TX, 2001)*.
- [17] W. C. Y. Lee, *Mobile Communications Design Fundamentals*, 2nd ed., John Wiley, New York, 1993.
- [18] W. C. Y. Lee, "Overview of Cellular CDMA," *IEEE Trans. on Vehicular Technology*, 40:2 (1991), 291–302.
- [19] R. Lopez-Ruiz, H. L. Mancini, and X. Calbet, "A Statistical Measure of Complexity," *Physical Letters A*, 25 (1995), 321–326.
- [20] P. Mankiewicz and R. Howard, "3G Mobility: The 'Evolutionary' Route to Wireless Data," *Sixth Annual Workshop on Frontiers in Distributed Information Systems FDIS (Aspen, CO, 2000)*, < http://www.fdis.org/00/fdis_2000-Howard_3G_Mobility.ppt>.
- [21] I. G. Niemegeers and S. H. deGroot, "Personal Networks: User-centric Networking," *Proc. Wireless World Research Forum Sixth Meeting (London, United Kingdom, 2002)*.
- [22] T. Otsu, I. Okajima, N. Umeda, and Y. Yamao, "Network Architecture for Mobile Communications Systems Beyond IMT-2000," *IEEE Personal Commun.*, 8:5 (2001), 31–37.
- [23] J. M. Rabaey, M. J. Ammer, J. L. da Silva, and D. P. S. Roundy, "PicoRadio Supports Ad-Hoc Ultra-Low Power Wireless Networking," *IEEE Computer Mag.*, July 2000, 42–48.
- [24] A. Rouz, M. J. Wilson, and B. G. Marchent, "Broadband Interworking Architecture (BRAIN) for Future Mobile Multimedia Systems," *IEEE Int. Conf. on Personal Wireless Communication (Jaipur, India, 1999)*, pp. 279–283.
- [25] P. Sarkar, "A Brief History of Cellular Automata," *ACM Computing Surveys*, 32:1 (2000), 80–107.
- [26] Sensit Project on Self-Organizing Sensor Networks with Reconfigurable Smart Components, Auburn University, <<http://www.darpa.mil/ito/psum2000/K286-0.html>>.
- [27] J. S. Shiner, M. Davison, and P. T. Landsberg, "Simple Measure for Complexity," *Physical Review E*, 59:2 (1999), 1459–1464.
- [28] K. Shohrabi, J. Gao, V. Ailawadhi, and G. J. Pottie, "Protocols for Self-Organization of a Wireless Sensor Network," *IEEE Personal Commun.*, Oct. 2000, 16–27.
- [29] R. V. Sole and S. Valverde, "Information Transfer and Phase Transitions in a Model of Internet Traffic," *Physica A*, 289:3–4 (2001), 595–605.
- [30] J.-Z. Sun, J. Sauvola, and D. Howie, "Features in Future: 4G Visions from a Technical Perspective," *Proc. Global Telecommunications Conference, 2001 (San Antonio, TX, 2001)*, Vol. 6, pp. 3533–3537.
- [31] R. Van Nee, G. Awater, M. Morikura, H. Takanashi, M. Webster, and W. Halford, "New High-Rate Wireless LAN Standards," *IEEE Commun. Mag.*, 37:12 (1999), 82–88.
- [32] S. Wolfram, "Universality and Complexity in Cellular Automata," *Physica D*, 10:1–2 (1984), 1–35.
- [33] M. Woolf, D. K. Arrowsmith, R. J. Mondargón, and J. M. Pitts, "Optimisation and Phase Transitions in a Chaotic Model of Data Traffic," to appear in *Physical Review E*, 66:4 (2002), 46106-1-11.
- [34] Y. Yamao, H. Suda, N. Umeda, and N. Nakajima, "Radio Access Network Design Concept for the Fourth Generation Mobile Communication System," *IEEE 51st Vehicular Technology Conference Proceedings, 2000, 3 (Tokyo, Japan, 2000)*, pp. 2285–2289.
- [35] X. Zhao, J. Kivinen, P. Vainikainen, and K. Skog, "Propagation Characteristics for

Wideband Outdoor Mobile Communications at 5.3 GHz," IEEE J. on Selected Areas in Commun., 20:3 (200), 507–514.

(Manuscript approved March 2003)

LESTER T. W. HO is currently a research student, working toward a Ph.D. degree in electronic engineering at the Department of Electronic Engineering at Queen Mary, University of London, United Kingdom, in a studentship sponsored by Lucent



Technologies. His research has been done in collaboration with the Global Wireless Systems Research Group at Bell Labs in Swindon, United Kingdom. His research interests include wireless networks, self-organization, and complex behavior. He received his B.Eng. degree (first class honors) in electronic engineering from Queen Mary, University of London.

LOUIS G. SAMUEL is currently a technical manager in the Wireless Lab of the Global Wireless Systems Research Group at Bell Labs in Swindon, United Kingdom. The lab is active in both layer 1 and layer 2 issues, such as MIMO, IA, HSDPA, and wireless architectures. His professional interests are in wireless network architectures, wireless protocols, non-linear dynamics, and complexity. He received his Ph.D. in the application of nonlinear dynamics to teletraffic modelling from Queen Mary, University of London.



JONATHAN M. PITTS is a professor at Queen Mary, University of London, where he is currently the head of the Telecommunications Research Group at the Department of Electronic Engineering. His professional interests are in resource management, performance evaluation, traffic engineering, accelerated simulation, and queuing theory. Professor Pitts received his M.Eng (first class honors) in communication engineering and his Ph.D. in cell rate simulation modelling of ATM telecommunications networks, both from Queen Mary, University of London. ♦



Copyright of Bell Labs Technical Journal is the property of Lucent Technologies, Inc. Published by Wiley Periodicals, Inc., a Wiley Company. Content may not be copied or emailed to multiple sites or posted to a listserv without the Publisher's express written permission. However, users may print, download, or email articles for individual use.