Providing Fault Tolerance in Wireless Access Networks

David Tipper, University of Pittsburgh
Teresa Dahlberg, University of North Carolina at Charlotte
Hyundoo Shin and Charlermpol Charnsripinyo, University of Pittsburgh

ABSTRACT

Research and development on network survivability has largely focused on public switched telecommunications networks and high-speed data networks with little attention on the survivability of wireless access networks supporting cellular and PCS communications. This article discusses the effects of failures and survivability issues in PCS networks with emphasis on the unique difficulties presented by user mobility and the wireless channel environment. A simulation model to study a variety of failure scenarios on a PCS network is described, and the results show that user mobility significantly worsens network performance after failures, as disconnected users move among adjacent cells and attempt to reconnect to the network. Thus, survivability strategies must be designed to contend with spatial as well as temporal network behavior. A multilayer framework for the study of PCS network survivability is presented. Metrics for quantifying network survivability are identified at each layer. Possible survivability strategies and restoration techniques for each layer in the framework are also discussed.

INTRODUCTION

Wireless networks have been growing rapidly in the past years to support increasing demands for mobile communications. The majority of recent wireless networks function as wireless access networks to provide mobile users with untethered access to resources that reside primarily in a wired network. Typical wireless access networks include analog and digital cellular phone/personal communications services (PCS) networks, wireless local area networks, and mobile wide-area data services (e.g., General Packet Radio Service, GPRS). Among those, mobile cellular and PCS networks currently represent the fastest growing sector with a current emphasis on mobile data services. In general, the flexibility provided by mobility has satisfied users of current wireless networks, despite the lower quality and reduced service offerings as compared to wired networks. Research is ongoing to extend the scope of services made available to mobile users to achieve the “anytime, anyplace, any form” communications vision. This vision is to provide voice, data, and multimedia services to users regardless of location, mobility pattern, or type of terminal used for access. As societal dependence on mobile terminals increases, users will demand the same system functionality, in terms of reliable service, that is characteristic of today’s wireline-based telecommunications and data networks. This implies that failures that inhibit communications or result in loss of critical data will not be tolerated.

The critical importance of providing communication service in the face of failures has been recognized in the public switched telephone network, and a great deal of attention has been paid to making these networks survivable and self-healing. However, little emphasis has been placed on understanding or improving the survivability of wireless access networks. The unique aspects of wireless access networks (e.g., user mobility, wireless channel, power conservation) suggest that survivability techniques for wired networks may not be directly applicable. Survivability is used to describe the available performance of a network after a failure. A survivability analysis measures the degree of functionality remaining in a system after a failure, and consists of evaluating metrics that quantify network performance during failure scenarios as well as normal operation. A variety of failure scenarios can be defined, determined by the network component that fails and its location. Examples of failure scenarios in cellular/PCS networks would include failure of a base station, loss of a mobile switching center, and loss of the link between a base station and a
Survivable network design refers to the incorporation of strategies into a network to mitigate the impact of failures. Strategies to improve network survivability can be classified into three categories:

- Prevention
- Network design and capacity allocation
- Traffic management and restoration

Prevention techniques focus primarily on improving component and system reliability. Some examples are the use of fault-tolerant hardware architectures in network switches and provisioning backup power supplies for network components (e.g., backup batteries at cell sites). Network design and capacity allocation techniques try to mitigate system-level failures such as loss of a network link by placing sufficient diversity and capacity in the network topology; for example, designing the topology and determining the capacity of links in a backbone network so that the network can carry the projected demand even if any one link is lost due to a failure. Traffic management and restoration procedures seek to direct the network load such that a failure has minimum impact when it occurs and that connections affected by the failure are restored while maintaining network stability. An example is the use of dynamic fault recovery routing algorithms to make use of the spare capacity remaining after a failure.

The “ideal” survivability goal is to make a network failure imperceptible to the user by providing service continuity and minimizing network congestion. Cost is always an issue, and the challenge is to provide an acceptable level of service for a set of failure scenarios in a cost-effective manner. A body of literature exists on survivability techniques for circuit-switched networks, asynchronous transfer mode (ATM) networks, and wavelength-division multiplex (WDM) networks, and there is emerging literature on the survivability of next-generation Internet (NGI) networks. However, relatively little research has been done on survivability for wireless access networks. Survivable network architectures and fault recovery protocols must be developed specifically for wireless access networks to support reliable services in this emerging technology.

Initial papers have mainly focused on database survivability in mobile cellular networks [1–3]. This work develops checkpoint algorithms and database restoration techniques to minimize the impact of location database failures. In [4, 5], algorithms for the design of a survivable topology for the landline portion of PCS networks are presented. However, the approach and assumptions used in the article were identical to techniques used in wired backbone networks and did not include any of the unique aspects of wireless networks (e.g., user mobility).

In [6–8], various survivability issues in wireless mobile networks were studied, such as types of component failures and mitigation strategies, metrics for characterizing outages, and survivability frameworks for wireless networks.

In this article we present a general overview of survivability issues in PCS networks, a framework for the study of wireless access network survivability, and results of a sample survivability analysis of a Global System for Mobile Communications (GSM) network. The simulation results given here show that user mobility can significantly degrade network performance after failures occur and that one must include the effects of mobile users when developing survivability mechanisms.

A WIRELESS ACCESS NETWORK ARCHITECTURE

A generic second-generation wireless access network architecture for supporting mobile communications is illustrated in Fig. 1. The architecture shown illustrates what is typical of current cellular/PCS networks. A wireless access network usually covers a large geographical service area partitioned into many small regions called cells. Each cell is served by a base station (BS) that serves as a fixed access point for all mobile terminals (MTs) within the cell. The BS terminates the wireless communication links (or channels) to the user on the network side of the user-to-network interface. The wireless links between the BS and MTs within a cell are digital and employ either time-division multiple access (TDMA) or spread-spectrum code-division multiple access (CDMA) techniques. The network may include base station controllers (BSCs), which manage a group of BSs and doe
Survivability issues in wireless networks must take into account these unique characteristics, especially user mobility which can significantly degrade network performance after failures occur. Therefore, existing survivability techniques, which have been successfully developed for wired networks may not be directly applicable to wireless access networks.

A SURVIVABILITY FRAMEWORK

The typical wireless access network shown in Fig. 1 is quite different from the typical wired telephone network. First it has a root-branch-leaf topology, with the MSC at the root. For the network to be survivable, alternate routes must exist between the network components with appropriate traffic restoration methods, or intelligent spare components must be provisioned (e.g., spare link between BS-BSC with automatic protection switching at endpoints). Note that the network design and restoration protocols must take into account the unique aspects of wireless access networks. Unlike wired networks, survivability techniques in wireless access networks must deal with user mobility, power conservation in mobile terminals, security (encryption and authentication), the poor quality of radio links (in comparison to wired equivalents), and channel capacity that is limited by a regulated frequency spectrum. Survivability issues in wireless networks must take into account these unique characteristics, especially user mobility, which can significantly degrade network performance after failures occur. Therefore, existing survivability techniques that have been successfully developed for wired networks may not be directly applicable to wireless access networks. In order to facilitate survivable analysis and design of wireless access networks we have developed a survivability framework similar to the approaches of [7, 9] for wired backbone networks. Our survivability framework [6, 10] for wireless access networks consists of three layers with survivability strategies possible at each layer. The three layers are termed access, transport, and intelligent. Note that the logical layers defined here are independent of the physical implementation of the network. Each of the three layers are characterized by network functions, network components, and communication links, as illustrated in Table 1.

The access layer has two sublayers, the radio and link levels, in order to distinguish between the wireless component and the landline portion. The access layer at the radio level defines the physical interface for communication over their wireless links within a cell. This includes the mobile terminal and BS wireless communication.

<table>
<thead>
<tr>
<th>Layer</th>
<th>Failure scenario</th>
<th>Potential impact</th>
<th>Possible metrics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Access</td>
<td>Loss of BS</td>
<td>Partial/full service loss in cell,</td>
<td>Call blocking probability,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>increased traffic in cells adjacent</td>
<td>forced call termination probability</td>
</tr>
<tr>
<td></td>
<td></td>
<td>to failure</td>
<td></td>
</tr>
<tr>
<td>Transport</td>
<td>Loss of BSC-MSC</td>
<td>Partial/full service loss in a cluster</td>
<td>Call blocking probability,</td>
</tr>
<tr>
<td></td>
<td>Link</td>
<td>of cells</td>
<td>forced call termination probability</td>
</tr>
<tr>
<td></td>
<td></td>
<td>increased traffic in cells adjacent</td>
<td>call setup delay, call release delay,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>to failure</td>
<td>paging/location update/registration delays</td>
</tr>
<tr>
<td>Intelligent</td>
<td>Loss of VLR</td>
<td>Loss of roaming service in a MSC</td>
<td>Lost user load (Erlangs),</td>
</tr>
<tr>
<td></td>
<td></td>
<td>coverage area</td>
<td>database access delay,</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>information accuracy probability</td>
</tr>
</tbody>
</table>

Table 1. Wireless access network survivability layers.
ation scheme for multiple access, modulation, error correction, control channels, and so on. The access layer at the link level includes the BSs, BSC, and radio resource management schemes (e.g., channel allocation and handoff). The transport layer supports call management functions (e.g., connection setup/teardown) and mobility management (e.g., location tracking) functions using the landline interconnection of BS, BSC, and MSC, with the MSC as the primary controller. The MSC at the transport layer uses the signaling network and services provided by service data management functions, implemented at the intelligent layer, to support call and mobility management. The intelligent layer supports service data management functions to provide the transport layer access to system databases (HLR, etc.) using SS7 signaling protocols. Together the three layers of Table 1 support network mobility with respect to terminals, users, and services.

Given the framework above to conduct a survivability analysis, one must identify performance-oriented survivability metrics along with techniques for evaluating the metrics over various modes of operation. The modes of operation include normal, single-failure, and multiple-failure modes. Table 2 lists examples of possible survivability metrics and sample failure conditions at each layer in the framework, along with some of the potential impacts of a failure in terms of the area affected and network service disruption.

At the access layer, a typical failure would be the loss of a BS, with appropriate survivability metrics of call blocking probability and forced call termination probability. The call blocking probability measures the percentage of call requests turned down due to lack of resources, where as the forced call termination probability measures the percentage of calls that are prematurely terminated, including those dropped at handoff. At the transport layer a typical failure would be the loss of a BSC-MSC link, resulting in loss of service to a cluster of cells. Appropriate metrics include call blocking probability and forced call termination probability, as in the access layer case. Since a large number of users are affected by the failure and may attempt to reconnect, one must also consider metrics such as the call setup delay, call release delay, and location update delay among other metrics listed in Table 2. Such metrics are defined for an entire MSC/VLR coverage area and have target mean and .95 percentile values recommended by the International Telecommunication Union (ITU). At the intelligent layer a possible failure scenario would be the loss of a VLR database, resulting in partial or complete loss of roaming service in a VLR/MSC coverage area. Possible survivability metrics would include the lost user load (i.e., user lost Erlangs) and the information accuracy probability at the HLR. The information accuracy probability measures the percentage of queries to the HLR that result in accurate responses (e.g., location information request).

To illustrate the difficulties encountered in designing survivable wireless access networks given the framework above, we report some sample results of a simulation-based survivability analysis. In the survivability analysis we considered a variety of failure scenarios at each layer in the framework above, and measured both the transient and steady state impact of the failure on the survivability metrics of Table 2. Note that in order to see the area affected by a failure one must consider both the steady state and transient network performance of the network after a failure. Transient conditions occur after a failure due to a combination of delays in detecting a fault, reporting it, and invoking restoration algorithms; coupled with increased call initiation requests from disconnected users attempting to reconnect in circuit-switched networks or dropped packets needing retransmission in packet networks. The importance of transient conditions after a failure has been documented for circuit-switched, packet-switched (both connectionless and connection-oriented), and signaling networks. In wireless access networks, user mobility only worsens transient conditions as disconnected users move among geographical areas to attempt to reconnect to the network.

A Sample Survivability Analysis

We have developed a simulation model to study the impact of various failure scenarios in the wireless access network of Fig. 1 as follows. Consider a typical GSM network serving a medium-sized city; we assume that a mobile network has 100 cells/MSC with 1 VLR, 20 BSCs, and 9 location areas (LAs), as shown in Fig. 2. In the figure the locations of the BSCs are denoted by the

![Figure 2. The simulation model.](image-url)
### Table 3. Mean results for 10 minutes post failure.

<table>
<thead>
<tr>
<th>Metric</th>
<th>No failure</th>
<th>Four cells failure</th>
<th>BSC-MSC link failure</th>
</tr>
</thead>
<tbody>
<tr>
<td>MOC blocking (%) $, P_o$</td>
<td>1.64</td>
<td>9.57</td>
<td>15.5</td>
</tr>
<tr>
<td>MTC blocking (%) $, P_f$</td>
<td>7.29</td>
<td>16.3</td>
<td>22.6</td>
</tr>
<tr>
<td>Location update delay (s), LD</td>
<td>0.257</td>
<td>5.23</td>
<td>3.85</td>
</tr>
</tbody>
</table>

20 numbers in larger font size, and the cells controlled by the BSC are outlined with a heavy black line. For example, BSC 5 controls the set of cells numbered {36, 26, 31, 41, 46}. The location areas are marked by the two long horizontal and vertical lines in Fig. 2. The cell radius for a BS is 3 km. In North America, the GSM system has 124 radio channels that are equally divided into two set of channels for different network providers in the same service area. Thus, we assume the system in the simulation model has 62 radio channels, with a frequency reuse cluster size of 7. Therefore, in a cluster of cells, there are 6 cells with 9 radio channels/cell, and one cell with 8 radio channels/cell. A GSM radio channel has eight time slots, and there is one control channel/cell. This results in an average of 70 traffic channels/cell.

To represent user mobility, we adopted the random mobility model. Specifically, the users are randomly placed in the network, and the distance from the base station in a cell and direction of movement are randomly selected. The speed of a user is constant within a cell and is uniformly distributed between 0 and 80 km/h. When a user crosses a cell boundary their direction of movement is again chosen randomly, and a new speed is selected randomly as well. In the simulation model, two types of calls are generated. They are mobile-originated calls (MOCs) and mobile-terminated calls (MTCs). The percentage of each call type is 70 percent MOC and 30 percent MTC. We assume that calls arrive to the system according to a Poisson process, and the calls have an exponentially distributed holding times with a mean of 120 s. For 2 percent call blocking with 70 traffic channels, each cell can support a load of about 59.1 Erlangs. The total number of subscribers in the system is set at 100,000. In order to meet the target ITU benchmark mean delays [11] in processing a call handling request (1 s) and a location update (2 s), we scale the processing time and set related parameters in this simulation model as follows: the post-selection delay is 58.4 ms, the location id query processing time is 8 ms, and the location update processing time is 9 ms.

Using the simulation model, a variety of failure scenarios have been studied with detailed results given in [10]. Typical results for the case of the failure of four disconnected cells and a BSC-MSC link failure, which results in the failure of a cluster of seven cells, are shown in Table 3. The mean performance results for the network 10 min post failures are given in the table. Some of metrics used to evaluate the effects of failures are the MOC blocking probability, the MTC blocking probability, and the mean location update time for the entire MSC service area. In our simulation after network failures, both the mobile user and PSTN user will make one attempt to reconnect interrupted calls.

From the mean results in Table 3, it is obvious that the MOC and MTC blocking rate increase as the number of failed cells increases because of the greater number of users affected. In general, the MTC blocking probability is higher than the MOC blocking probability because there are other factors involved (e.g., the paging operation failed to reach the called terminal). For the mean location update delay, the time in case of failures is greater than that of no failure. However, the location update delay for BSC-MSC link failure is less than the delay in the case of four cells failed. This is because the number of requests to VLR in the case of BSC-MSC link failure is less than in the other. When mobile terminals enter the failed cells, they do not trigger a location update. Therefore, the location update delay at VLR in these cases increases from 0.25 to 5.23 s for the entire MSC service area after the failure.

### Table 4. Typical survivability strategies.

<table>
<thead>
<tr>
<th>Layer</th>
<th>Robustness and redundancy</th>
<th>Traffic restoration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Access radio level</td>
<td>Spare RF components, overlapping/scalable cells</td>
<td>Load sharing protocols, dynamic channel allocation, adaptive channel quality protocols</td>
</tr>
<tr>
<td>Access link level</td>
<td>Spare BS-BSC link, multihoming BS to BSCs</td>
<td>Automatic protection switching, dynamic rerouting protocols, self-healing rings</td>
</tr>
<tr>
<td>Transport</td>
<td>Spare BSC-MSC link, ring topology for BSC-MSC interconnect</td>
<td>Automatic protection switching, self-healing rings, dynamic rerouting, call gapping</td>
</tr>
<tr>
<td>Intelligent</td>
<td>Physical diversity in signal networking links, physical database diversity</td>
<td>Dynamic routing, checkpoint protocols</td>
</tr>
</tbody>
</table>
The transient behavior, which takes place during a time period immediately following a failure, is important. During the transient period, those users whose calls were prematurely terminated attempt to reestablish the calls at the same time. This incident will cause network congestion and increase the call blocking probability. The performance metrics can exceed the ITU recommended benchmarks for long periods of time.

The shape of the failed area contributes to the effects of the failure in the network. We found that failures of separated cells are worse than failures in a cluster of adjacent cells. For example, the failure of seven cells, cells 18, 23, 47, 49, 63, 79, and 82 in Fig. 2, have handoff call blocking rate $P_h = 13$ percent and location update delay $LD = 4.16$ s, while a BSC-MSC link failure that affects a cluster of seven cells has $P_h = 6.6$ percent and $LD = 3.85$ s.

The location of failures also matters. For example, the failure of nonboundary location area cells is worse than those of location area boundary cells. Consider the case of 4 nonboundary location area cells that fail at cells 47, 49, 53, and 89 in Fig. 2; this results in a MTC setup time of 7.2 s, while the case of 4 boundary cells that fail at cells 32, 42, 72, and 74 results in a MTC setup time of 3.5 s.

User movement is very important. For example, deterministic movement was found to be worse than random movement. Deterministic movement is the case where all users disconnected by a failure move in a fixed pattern to adjacent cells. This is consistent with a highway movement pattern and can result in longer transient periods after failures. In addition, the speed of user movement can significantly worsen the performance metrics. For example, in the case of random movement with two cell failures, the user speed of 0–80 km/h has a MTC blocking rate $P_t = 11.8$ percent, whereas a user speed of 10–100 km/h results in $P_t = 34.8$ percent.

User behavior matters. In our previous study, we assumed that each party of the failed connection attempts to reconnect one time. However, behavior in real networks indicates that both parties try two to three times to reconnect before giving up.

From the simulation results and observations described above, several unique characteristics, especially user mobility, in wireless mobile networks play an important role in the event of failures. Therefore, survivability strategies must be incorporated into the network to minimize the effects of failures. In the following section, survivability strategies and restoration techniques for each layer of the framework are presented.

**Survivability Strategies**

In general, survivability techniques can be deployed at each layer of the framework of Table 1 for specific failure scenarios. Examples of the types of survivability strategies possible at each layer are listed in Table 4.

At the access layer radio level the primary failure to be guarded against is failure of the wireless link to the user. Due to the constraints of a regulated frequency spectrum, allocation of spare radio channels for use in case of failure decreases the radio channel capacity available during normal operating modes, and such an approach is not economically feasible. A possible approach discussed in detail in [12] is to design the network with an overlapping cell site architecture along with frequency reuse partitions, a dynamic channel allocation algorithm, and adaptive power control to provide dual homing at the wireless link level. A cell site architecture with overlapping coverage area for radio-level survivability is shown in Fig. 3a. Each hexagon represents a cell with the BS in the center of cell. Each BS supports two groups of radio channels, namely short-haul and longhaul channels. The short-haul channels are used within the small circle, while the long-haul channels cover areas of the larger circle. Ideally, the cell size and coverage areas of both radio channel groups are selected so that all MTs can access at least two channel groups. This means each mobile terminal can access either long-haul channels from at least two BSs or short-haul and long-haul channels from the same BS. Therefore, the greater overlap provides more access channels to each MT, but it also increases co-channel interference.

At the access layer link level and transport layer, the primary concern is component/link failure in the landline portion of the network. Traditional survivability strategies such as a mesh-type architecture, automatic protection switching, and self-healing rings can be applied.
With emerging 2.5G and 3G data services in wireless mobile networks, the survivability framework will need to be modified to consider characteristics of mobile data networks.

CONCLUSIONS

In this article we present a sample survivability analysis of a PCS network. Unique characteristics in wireless mobile networks, especially user mobility, can significantly worsen network performance after failures. Unlike wired networks, the impact of a failure in a wireless network depends on a variety of factors like the location and shape of the failed area, user mobility, and user behavior. A multilayer survivability framework is presented to facilitate survivable wireless network design. This framework includes metrics for quantifying network survivability, possible survivability strategies, and restoration techniques for each layer. Additional work is needed on survivable network topology design and specific fault recovery protocols. Note that the survivability framework and analysis presented here focus on voice service networks. With emerging 2.5G and 3G data services in wireless mobile networks, the survivability framework will need to be modified to consider characteristics of mobile data networks. For example, the rate of packet loss can be used as a survivability metric for data services instead of call blocking rate in voice services, and the retransmission of lost packets must be considered in sizing spare capacity.

REFERENCES


BIographies

David Tipper (tipper@tele.pitt.edu) is an associate professor of telecommunications in the Department of Information Science and Telecommunications at the University of Pittsburgh. Prior to joining Pitt in 1994, he was an associate professor of electrical and computer engineering at Clemson University, Clemson, South Carolina. He is a graduate of the University of Arizona (Ph.D. electrical engineering 1988, M.S. systems engineering 1984) and Virginia Tech (B.S. electrical engineering 1980). His current research interests are network design and traffic restoration procedures for survivable networks, network design for wireless and wired networks, and network control algorithms. His research has been supported by grants from various government and corporate sources such as NSF, DARPA, NIST, AT&T, and IBM.

Teresa A. Dahlberg (tdahller@uncc.edu) is an assistant professor of computer science in the College of Information Technology at the University of North Carolina at Charlotte. She holds a B.S.E.E. from the University of Pittsburgh, and M.S. and Ph.D. degrees in computer engineering from North Carolina State University. She previously worked in the ECE department and was with IBM Corporation for several years. Her research, within the Multimedia Computing and Networking laboratory, is supported by the NSF. Her current funded projects include wireless access network survivability, interactive visualization of mobile network simulations, development of a secure, multimedia, wireless network testbed, and integrated solutions for fault-tolerant middleware and networking.

Hyundoo Shin received his B.S. in electronic engineering from Sung-Kyun-Kwan University, Korea, in 1995, and his M.S. in telecommunications from the University of Pittsburgh in 1999. His research work involved the study of PCS network survivability.

Chalermpol Charnsripinyo (charnsri@tele.pitt.edu) is a research assistant and Ph.D. student at the Department of Information Science and Telecommunications, University of Pittsburgh. He received his B.S. in computer science from Thammasat University, Thailand, in 1992, and his M.Sc. in electrical engineering from the George Washington University in 1998. His current research focuses on network survivability and wireless network design.