# Survivable Mobile Wireless Networks: Issues, Challenges, and Research Directions

James P.G. Sterbenz, Rajesh Krishnan,
Regina Rosales Hain, Alden W. Jackson, David Levin, Ram Ramanathan, and John Zao
BBN Technologies
10 Moulton Street, Cambridge, MA 02138, USA
+1 617 873 5063
{jpgs, krash, rrhain, awjacks, dlevin, ramanath, jzao}@bbn.com

## ABSTRACT

In this paper we survey issues and challenges in enhancing the survivability of mobile wireless networks, with particular emphasis on military requirements*. Research focus on three key aspects can significantly enhance network survivability: (i) establishing and maintaining survivable topologies that strive to keep the network connected even under attack, (ii) design for end-to-end communication in challenging environments in which the path from source to destination is not wholly available at any given instant of time, (iii) the use of technology to enhance survivability such as adaptive networks and satellites.

## Categories and Subject Descriptors

C.2.1 [**Computer-Communication Networks**]: Network Architecture and Design – *network communications, network topology, packet-switching networks, store and forward networks, wireless communication*; C.2.2 [**Computer-Communication Networks**]: Network Protocols – *routing protocols*; C.4 [**Computer Systems Organization**]: Performance of Systems – *fault tolerance, reliability, availability, and serviceability.*

## General Terms

Algorithms, Design, Performance, Reliability, Security.

## Keywords

Survivability, mobile wireless network, weak and episodic connectivity, disconnected, asymmetric channel, eventual stability, eventual connectivity, store and haul forwarding, low probability of detection (LPD), satellite, ad hoc routing, topology, security, fault tolerance.

## 1. INTRODUCTION TO SURVIVABILITY

Network survivability is an essential aspect of reliable communication services. Survivability consists not only of robustness against failures occurring due to natural faults, accidents, and unintentional operational errors or misconfigurations, but also failures that are induced by malicious adversaries, particularly in the context of military networks. Mobile wireless networks provide ubiquitous computing and untethered access to the Internet, but significantly challenge survivability, both because users are mobile and because the communication channels are accessible to anyone.

This paper is a survey of the issues, challenges, and proposed research directions in survivable mobile wireless networks, resulting from our participation in the DARPA survivable mobile wireless information networks study program. The rest of this paper is organized as follows: This first section introduces and defines survivable networking and its aspects. The next three sections outline major thrusts in achieving survivability: Section 2 discusses establishing and maintaining network connectivity for survivability; Section 3 argues that we should expect a challenging mobile wireless communication environment and design for survivability; Section 4 discusses adaptive networking and satellite technologies that can enhance survivability. Section 5 summarizes the paper.

### 1.1 Definitions of Survivability

Traditional security research is primarily focused on the detection and prevention of intrusions and attacks rather than on continued correct operation while under attack. Fault tolerance is usually concerned with redundancy that is required to detect and correct up to a given number of *naturally occurring* faults. Nature is not malicious, and conventional failure models make significant assumptions, in particular, assuming faults to be independent and random. The presence of intelligent adversarial attacks can significantly challenge these conventional models. Software and protocol vulnerability often become more important considerations in the presence of an adversary.

There are a number of definitions of survivability (e.g. [6,41]). The one we use here is from the Software Engineering Institute, which emphasizes timeliness, survivability under attacks and failures, and that detection of attack is a vital capability [35, 15]:

> *Survivability is the capability of a system to fulfill its mission in a timely manner, even in the presence of attacks or failures. Survivability goes beyond security and fault tolerance to focus on delivery of essential services, even when systems are penetrated or experience failures, and rapid recovery of full services when conditions improve. Unlike traditional security measures*

*This work was supported by DARPA under contract number F30602-99 C-0131 issued by Rome AFRL.


*WiSe'02,* September 28, 2002, Atlanta, Georgia, USA.

*that require central control and administration, survivability addresses highly distributed, unbounded network environments that lack central control and unified security policies.*

***The Three Rs: Resistance, Recognition, and Recovery*** *The focus of survivability is on delivery of essential services and preservation of essential assets. Essential services and assets are those system capabilities that are critical to fulfilling mission objectives. Survivability depends on three key capabilities: resistance, recognition, and recovery. Resistance is the capability of a system to repel attacks. Recognition is the capability to detect attacks as they occur and to evaluate the extent of damage and compromise. Recovery, a hallmark of survivability, is the capability to maintain essential services and assets during attack, limit the extent of damage, and restore full services following attack.*

We further extend this definition to require that survivable systems be able to quickly incorporate lessons learned from failures, evolve, and adapt to emerging threats. We call this survivability feature refinement.

We can classify survivable mobile wireless networking requirements into four categories based on [15]: (i) resistance requirements; (ii) recognition requirements; (iii) recovery requirements; and (iv) refinement requirements. We can also describe a requirements definition methodology for survivability that is based on software requirements definition processes [23]. This includes the definition of system and survivability requirements, legitimate and intruder usage requirements, development requirements, operations requirements, and evolution requirements. Essential services must be identified and resistance, recognition, and recovery requirements must be specified for the penetration, exploration, and exploitation phases of the attack.

These approaches have guided this work and are recommended for more detailed requirement analyses for future mobile wireless networks.

Ultimately, there are two distinct aspects of survivability that apply at all networking layers:

**Information access requirements:** Does the user have access to the information or services required to complete the task in the presence of failures or attack? For example, is it possible to replicate services or information and provide them locally when the network gets partitioned? End-to-end communication should not be mandated in these cases.

**End-to-end communication requirements:** On the other hand, there are interactive applications, inter-personal communications such as voice calls, or dynamically generated information such as current sensor data, which require true end-to-end connectivity. Do existing sessions survive? Can the user create new sessions to reach the intended communication end-point even in the presence of failures and attacks? This requires that the communication end-points themselves survive and that the adversary must not be able to permanently partition the network. Furthermore, the adversary must not be able to permanently disable access to required services such as authentication, naming, resource discovery, or routing.

## 1.2 Military Network Survivability

The use of wireless networking technologies to support military operations imposes stringent security and operational requirements on those technologies:

**Transmission Security (TRANSEC)** – the protection of wireless communication at physical, medium access and data link layers over wireless media. The services include countermeasures against radio signal detection, jamming, control/user data acquisition, and eavesdropping.

**Communication Security (COMSEC)** – the protection of data and voice communications between designated endpoints. The services include message confidentiality, integrity, and end-point authentication. In addition, they may include optional non-repudiation, anti-replay protection, and traffic analysis countermeasures. Finally, military tactical networks often require rapidly supporting secure communications among dynamic groups of users or equipment, such as dynamically formed (or disbanded) coalitions.

**Authorization and Access Control** – the support of multi-level security measures by implementing identity or role based access control on applications, application servers, and their proxies. Multi-level security requires segregation of levels, possibly via cryptography. Authorization and access control require reliable authentication of human users and communication equipment.

**Network Infrastructure Protection** – the protection of routing and network management infrastructure against both passive and active attacks, such as rogue devices masquerading as switching elements, insertion, deletion, modification or replay of control messages, and introducing significant delays to message transport. This service may require strong authentication of switching equipment as well as confidentiality, integrity, non-repudiation, anti-replay protection and traffic analysis countermeasures for control traffic.

**Robustness** – the requirement to accommodate hardware and software failures, asymmetric and unidirectional links, or limited range of wireless communication. It includes the need for the networks to survive specific types of device overrun (physical seizure), network fragmentation and denial-of-service attacks.

**Efficiency** – Finally, even more than their commercial counterparts, military wireless networks are expected to be efficient in their use of electrical and computing power, silicon real estate, and communication bandwidth.

## 1.3 Cellular Telephone Network Survivability

Existing work on survivability in the context of cellular telephone networks concentrates primarily on infrastructure survivability (for example, see the *outage index* metric [41,14]) and does not consider adversarial attacks. However, they offer some insight on quantifying survivability and the role of network management tools.

Networks are vulnerable during upgrades, especially those involving software [39]. Furthermore, rapid evolution leads to learning-curve problems as well as over-concentration of traffic or services into single points of failure. This problem is exacerbated by deficits in network management tools to operate and maintain increasingly complex systems. Deployment errors (such as

backup circuits being carried on the same fiber as the primary) can defeat fault-tolerant designs.

Architectural improvements applicable to cellular telephony include the use of redundant networks (e.g. SONET rings), multimode handsets, and overlay networks to improve survivability [38]. In the past, landline and cellular carriers were different administrative entities, and radio links had poor reliability that set expectations low. End-to-end reliability and survivability issues will be increasingly important in future cellular telephone networks.

## 1.4 Ad Hoc Wireless Network Survivability

Tactical wireless networks are typically ad hoc networks with limited or no reliance on infrastructure. Thus, the focus on infrastructure survivability (which is of importance to cellular networks) is not appropriate for ad hoc networks. A different approach to survivability is needed in the context of tactical mobile wireless networks.

The wireless communication environment is harsher – issues include rapid attenuation with distance, multipath fading, weather effects, faraday cages, and obstructions in the terrain. Furthermore, wireless networks offer more opportunities to the attacker to eavesdrop (typically there is a lesser probability of detection than a wiretap), to do traffic analysis, or to jam.

Conventional techniques (for example, *k*-edge-connected or *k*-node-connected topologies, replication, and optimal replica placement) for fault tolerance and improving reliability are important to provide survivability. However, these must be hardened in the context of attack models, and adapted to the wireless context. Furthermore, the need for robust and survivable software cannot be overemphasized.

Techniques for providing data confidentiality, integrity, and authentication are necessary, but not sufficient, against denial of service attacks in the wireless context. Furthermore, in mobile ad hoc networks, access to a key infrastructure (located at a secure remote location) cannot be guaranteed, which poses challenges for security (e.g. delayed receipt of certificate revocation lists).

## 1.5 Research Pursuits towards Survivability

This following sections outline three major thrusts that have the potential to significantly increase the survivability of mobile wireless networks:

1. Establishing and maintaining network connectivity under adversarial situations.
2. Expectation of a challenging environment with design for survivability:
   - weak and episodic communication will be a common occurrence; communication should be possible even when end-to-end paths rarely exist.
   - mobility should be expected and exploited.
3. Exploit technology to enhance survivability:
   - network nodes and protocols should be adaptive.
   - satellites can mitigate the effects of mobility and enhance connectivity.

# 2. SURVIVABLE CONNECTIVITY

The first major goal in survivability is to establish and maintain a connected network, whenever practical. This allows conventional routing and end-to-end protocols to function with reasonable performance. Challenging this goal is the desire to remain stealthy.

## 2.1 Establishing the Network

Before a network can be used, it must be configured into a set of nodes that have network layer connectivity with one another, supporting addressing, routing, and signaling.

### *2.1.1 Infrastructure Assumptions*

There are two distinct schools of thought regarding infrastructure in the area of pervasive wireless networking. One approach (e.g. Mobile IP) relies heavily on available pre-configured infrastructure. The other approach (ad hoc networks) assumes that all nodes run a common ad hoc routing protocol and that no infrastructure is present.

There is limited support for combining heterogeneous networks of wired and wireless networks. One reason for this is the tradition of quasi-static addressing of nodes and subnetworks prevalent in IP internetworks. In general, mobile architectures do not support discovery and self-configuration of existing network infrastructure.

Some research systems have automatic fallback modes that allow for ad hoc networking of terminal Internet nodes when no connection to fixed infrastructure is available [46]. Such multi-modal operation, with automatic and seamless switchover between infrastructure-based and pure ad hoc modes, is critical to survivable mobile networking. By seamless, we mean that transport and application sessions must survive switchovers between infrastructure and ad hoc modes in either direction.

### *2.1.2 Network Layer Autoconfiguration*

Most work on self-configuration in heterogeneous networks focuses primarily on naming and service discovery issues. They assume that each network node has an address a priori[1] and also that a routing scheme is in place. They are concerned with application level overlays rather than on network bootstrap. For example [36] describes a CORBA-based management and self-configuration architecture for battlefield networks, in which, they require "…*a mechanism that can dynamically assign and release the network address associated with a network element, a group of network elements, or a subnetwork*."

Self-configuration in wired networks is usually limited to host autoconfiguration based on DHCP, Zeroconf [51], or based on pre-existing unique identifiers as with IPv6. These approaches require globally unique resource identifiers or the presence of infrastructure. Survivable nodes must address the problem of secure network layer autoconfiguration of addressing, routing, and signaling based on mission requirements.

*Secure* autoconfiguration of wireless networks remains a hard problem. No satisfactory approaches are known beyond using a

[1]With the exception of addressless routing approaches such as diffusion routing, which are designed for specific applications.

single shared secret or gossip-based probabilistic protocols. This area is in need of further research.

### *2.1.3 Anonymous Networks of Sensors*

Most current ad hoc networks do not have support for anonymous nodes. They assume unique identifiers for each node such as an Ethernet MAC address or an EUI-64 identifier for IPv6. Unique identifiers raise several concerns for privacy and anonymity. Note that knowing the identifier of a node does not necessarily reveal the identity of the user or owner, however it may provide hints that are unacceptable risks to topology or traffic analysis.

It is impossible to deterministically assign globally unique IDs in an anonymous network. The only ways to avoid this are if an initiator can be specified or if randomized tie-breaking strategies are allowed.

Some approaches, such as amorphous computing, address anonymous networks. For example, the Clubs algorithm specifically looks at very dense wireless networks [26]. Sparse wireless networks may require other clustering techniques.

## 2.2 Low Probability of Detection

For most military ad hoc networks, *low probability of detection, interception, and exploitation* (LPD/LPI/LPE) – that is, the ability of an enemy to detect and exploit radio energy – is of paramount importance. A number of techniques may be employed to mask the radio signature of a node, including covert waveforms, directional transmissions, and reduced transmission power.

Survivability is enhanced when the network is stealthy to potential adversaries. This stealth, however, makes it more difficult for legitimate communications; in general lower transmission power reduces the probability of detection to both adversary and legitimate nodes.

Furthermore, military networks must be able to deny topology knowledge to the adversary.

## 2.3 Survivable Topological Connectivity

Reducing transmission power has serious consequences in terms of *topological connectivity survivability* – that is, the ability of a network to be connected (or to not partition). This is because reducing the transmit power limits the range of inter-node links used for multihop routing (one may be able to maintain the range by reducing the data rate but that impacts the network capacity). This in turn increases the probability of the network becoming partitioned. Sustaining topological connectivity while maintaining LPD presents a problem – how can we reduce power while maintaining desired connectivity?

In a mobile ad hoc network, we must adaptively adjust transmit powers of nodes in response to mobility, so as to optimize certain *power metrics* while adhering to certain *connectivity constraints*:

**Power metrics:** *Maximum power and average power.* That is, minimize the maximum (or average) power used by the network nodes, where the maximum (or average) is taken over all nodes and during the duration of operation.

Maximum power is important because it determines the range at which a receiver can detect the presence of a transmitter. Average power is important when detectors can coherently combine signals. Furthermore, each may signify a different kind of vulnerability. For instance, consider two networks $N_1$ and $N_2$. $N_1$ has a single node transmitting at 30 dBm and four others at 5 dBm (maximum 30 dBm, average 10 dBm). $N_2$ has all nodes transmitting at 20 dBm (maximum 20 dBm, average 20 dBm). Then, $N_1$ likely has a high degree of vulnerability to the detection (and perhaps enemy attack using RF seeking missiles) of *one* node, whereas $N_2$ likely has a low degree of vulnerability to the detection of *many* nodes.

**Connectivity constraints:** A network is *connected* if there is a path between every pair of nodes. A network is *biconnected* if the loss of any one link leaves the network connected. Biconnected topologies are desirable for networks to survive the loss of individual links.

Some of the poorly understood problems in balancing power and connectivity are:

*What level of connectivity (e.g., biconnectivity, triconnectivity) yields the best tradeoff between robust connectivity and LPD?* For instance, in some very hostile environments, biconnectivity may not be sufficient for survivable topology. Also, can one adaptively adjust the level of connectivity depending upon the situation?

*Adaptive adjustment of the transmit powers and topology to evade jammers and interceptors.* Knowing, or having sufficient information to deduce the locations of jammers and eavesdroppers would allow for better evasion, but require more sophisticated protocols.

*Combining network-layer approaches with physical layer approaches.* Physical layer approaches such as covert waveforms have been used or been suggested for use for LPD. New challenges arise in combining these with network-layer approaches in a vertically integrated fashion. It is important to understand the relative benefits and tradeoffs of physical, MAC, and network layer techniques, in combination with one another.

### *2.3.1 Energy Management*

Conventional research into energy management in wireless networks focuses on topology design techniques that construct topologies that allow optimum power consumption for the node and/or network. Energy efficient protocols typically involve the design of protocols of lower message complexity.

Battery-powered wireless nodes are susceptible to a form of denial-of-service that involves depleting their energy by forcing them to transmit frequently (or at higher power by raising the noise threshold), or to remain active constantly due to continuous reception of malicious packets. Survivability in this context involves the ability of nodes to transfer roles or tasks when suspected of being under attack or to conserve battery for critical tasks.

# 3. SURVIVABLE COMMUNICATION

While it is important to establish and maintain survivable connectivity whenever practical (as described in the last section), there will be situations where the environment is so challenging that it is not possible to continuously keep nodes connected with one another, particularly when constrained by LPI/LPD requirements. This may be the case either due to challenging channel conditions (noise or jamming) or due to extremely high

mobility. Thus, we should *expect* asymmetric links, weak connectivity, episodic connectivity, and dynamic topologies to be common occurrences during which communication must proceed.

## 3.1 Weak and Episodic Connectivity

One of the key aspects that make it difficult to maintain a connected network is a channel that is noisy, jammed, or has eavesdroppers. Since we must assume that the conditions are time-varying (particularly as nodes move), the result is a channel that may be asymmetric, may be weakly connected, and may suffer episodic connectivity during which there are periods of disconnection. It is crucial for network survivability that the protocols and algorithms *expect* these conditions as part of their normal operation, and communicate in spite of them.

### *3.1.1 Asymmetric Channel Connectivity*

Conventional network and transport protocols have traditionally assumed bidirectional connectivity for proper operation. At the network layer, this means that routing protocols do not have to account for unidirectional disjoint paths; at the transport layer this means that a reliable back channel is assumed.

Several ad hoc routing protocols (with some exceptions such as DSR) expressly prohibit unidirectional routing. Performance reasons are often cited for why unidirectional links can be considered harmful. However, survivability in tactical networks may require the use of highly asymmetric and sometimes unidirectional links. For example, communications may be effectively jammed in one direction, due to the limited transmission power of particular nodes. A node may need to be radio-silent to prevent detection while still requiring frequent mission-specific updates for situational awareness.

In the case of intermediate links along a path, it is essential that the routing protocol support disjoint forward and reverse paths. In the case of an asymmetric end user, the routing protocol must support disjoint unidirectional paths and network layer signaling must not require a back channel.

Similarly, in such situations it is necessary to maintain end-to-end sessions even when the link shuts down in one direction. Closed loop control mechanisms (such as TCP error, flow, and congestion control) generally assume a reliable return channel for acknowledgements to properly function. Some work has been done on enhancing transport protocols for asymmetric channels (e.g. [8]). Additional investigation should consider how to apply open loop mechanisms when necessary for highly asymmetric and unidirectional paths.

As a variation, an alternative low capacity wired channel might be available for communications in the reverse direction. Therefore, survivable routing and transport protocols must support asymmetric and unidirectional links as well as polychannel architectures.

### *3.1.2 Unstable End-to-End Paths*

Routing protocols currently require that a route (complete path) exist from source to destination before communication is initiated.

The *eventual stability* model of ad hoc routing assumes that routing converges eventually after partitioning. Under this model, a complete path to destination must exist at a given time; otherwise, communication is not attempted. Note that this is true whether or not datagrams are to be forwarded along the path, or a connection is to be established. This is shown in Figure 1, in which intermittent sources of noise, including jammers, (gray circles) affect some of the links. Communication is only attempted along the path that has stable links that are strongly connected. Note that if another intermittent noise source (hashed circle) were to become active, communication would not be attempted *at all*.

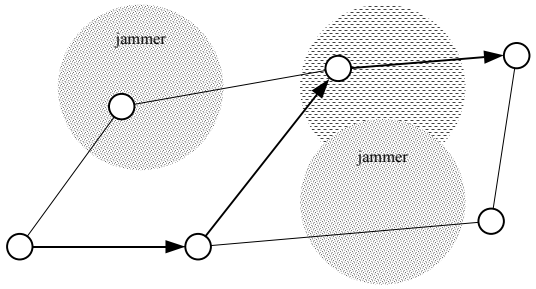


**Figure 1. Communication over stable paths**

This model is used by on-demand routing as well as table-driven protocols. While this is a traditional operating assumption, challenges in connectivity arising from new technologies such as directional antennas [33], and range limitations imposed by LPI/LPD concerns challenge this assumption. Furthermore, intermittent and/or targeted jamming can disrupt routing convergence in tactical wireless networks. High mobility is another aspect that challenges the ability for routing algorithms to converge. In tactical military networks, routing algorithms may rarely or never converge to stable end-to-end paths.

The *eventual connectivity* model relaxes the traditional assumptions so that communication can proceed along partial segments of paths between communicating nodes. The notion of eventual connectivity follows from the observation (in the distributed computing context) that there is no need to require that a complete physical path between communicating processes exist at a particular point in time to ensure delivery of information [48]. Algorithms to achieve end-to-end communications under eventual connectivity have been proposed [1,2] in this context.

Information progresses as far as possible, along whatever paths possible, until it reaches its destination. This extends the concept of store-and-forward, and requires modifying the typical forwarding behavior of dropping packets if an outgoing link to the next node becomes temporarily unavailable.

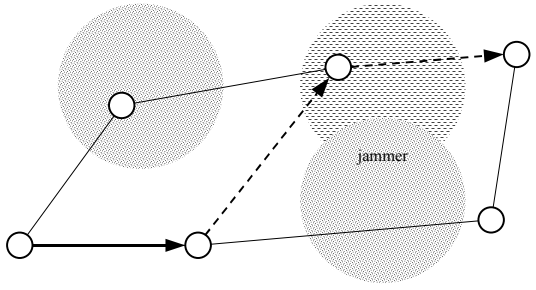


**Figure 2. Communication over intermittent links**

This is depicted in Figure 2, in which all links are intermittent such that there is never a complete path, but there are times in which partial segments of a path are available. Data can be moved as shown by the solid arrow. Then, when the next two intermittent links become temporarily available, data progresses along the path of the dashed arrows (and at this point the first link may become unavailable).

We recommend that a significant shift is required so that we *design for eventual connectivity*, but however leverage existing routing algorithms in order to *optimize for eventual stability*. Thus, when routing algorithms do converge, the efficiency of conventional eventual stability (and optimizations such as cut-through) can be exploited.

This requires changes to current forwarding mechanisms. For example, when a link becomes unavailable, it should be marked in a new field in the forwarding table, rather than the forwarding entry removed. This ability to forward under eventual connectivity has the benefit of dampening of the routing algorithm control loop, thus reducing instability and routing algorithm update traffic.

This also requires an analysis of the buffering required at nodes to store packets when necessary, and algorithms to determine which data is dropped when the buffers are full.

Furthermore, survivable mobile nodes must support multiple routing approaches at the same time [40]. For example, they should be able to switch between table-driven and on-demand approaches or use table-driven approaches within a cluster and use on-demand inter-cluster protocols. End-to-end communications protocols must not depend on a single path [16,28,7,13]. The early detection and location of (and recovery from) arbitrary communication failures including those due to the presence of malicious processors is vital to network survivability [19,28].

### *3.1.3 Hierarchical and Multipath Routing*

Support for hierarchy is missing from most ad hoc routing protocols [34] with the exception of a few approaches, such as CGSR and MMWN. MMWN provides a robust virtual gateway mechanism which keeps routes stable even under mobility or failure of gateway nodes [31].

A number of recent works consider the use of multipath routing to improve survivability under mobility and failure [30,50,22]. However they require that at least one complete path exists from the source to destination prior to attempting communication. The combination of these techniques along with routing for eventual connectivity has the potential to significantly improve survivability.

## 3.2 Mobility

Traditionally, mobility in networks has been handled as a necessary evil, with routing protocols adapting as best as possible to mobile and nomadic nodes. Just as survivable networks should expect challenging channel conditions as a normal mode of operation, they should be designed to *expect* and *exploit* mobility.

### *3.2.1 Nomadicity versus Mobility*

Wireless network architectures exhibit a dichotomy between nomadicity and mobility. *Nomadicity* assumes constant movement during communication, and anticipates disconnected operation as the norm. Thus, applications are expected to tolerate disconnection during movement.

Traditional mobility has tried to maintain active sessions to mobile nodes as long as possible until a network partition or routing failure occurs. In order to do so, continuous access to pre-configured infrastructure is assumed (for example foreign agents and home agents in mobile IP).

A specific ramification of this dichotomy is how addressing is handled, in particular, whether the address is assigned once and held under mobility as long as possible, or the node acquires new addresses as it moves to a different subnetwork. While each approach has relative benefits, neither should be the only supported. The ability to support multiple addresses on the same network interface allows operation in at a middle ground between nomadicity and mobility.

When multiple addresses are available, the issue becomes whether we can seamlessly and securely migrate sessions when we readdress due to mobility (see for example TCP migration [37]).

### *3.2.2 Routing under very High Mobility*

High mobility often poses challenges to conventional ad hoc routing protocols especially after they reach their reactive limit. In this case it is necessary to use knowledge of the location and trajectories of nodes to predict future location without requiring rapid convergence of routing algorithms. Trajectory routing [42] uses trajectories to compute destination node locations. In the case of predictable motion this is sufficient, for example low earth orbiting (LEO) satellites and racetrack-path unpiloted aerial vehicles (UAVs).

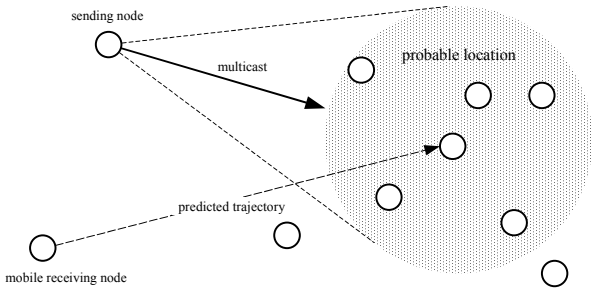


**Figure 3. Trajectory and spray routing**

In the case where the trajectory is not deterministic, the need for current location can be mitigated by multicast *spraying* of information into a region that the node can be expected with high probability [42], as shown in Figure 3. Hierarchical clustering can be exploited to multicast to all of the relevant lowest level clusters, broadcasting to all of the cluster members.

Note that high mobility can invalidate the eventual stability assumption by preventing routing convergence (discussed previously), thereby providing additional motivation to look at communication algorithms that can work under eventual connectivity.

In cases where there is an architectural requirement to affiliate with a base station within the current cell before communication can commence, a serious bottleneck results in high mobility environments if the cell footprints (and therefore dwell times) are small. Techniques that combine trajectory or spray routing with

architectures that require less frequent or no reaffiliation/handoffs must be explored. An example is a technique that allows affiliation to any base station in the neighborhood and not necessarily the one in the current cell, while still being able to use the current base station for communications to promote spatial re-use.

### *3.2.3 Exploiting Mobility to Achieve Connectivity*

It is possible to *exploit* mobility to communicate when otherwise impossible. In the worst case, eventual connectivity routing will store data until a promising outgoing link becomes available. Proactive control can be used in two ways to expedite the transfer of data. *Movement control* can be used to exert control on other nodes to move them into range such that a path toward the destination exists.

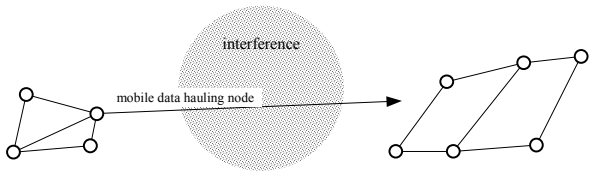


**Figure 4. Store-and-haul data forwarding**

Alternatively, mobile node can *store-and-haul* packets toward their destination by physically transporting the data, as shown in Figure 4.

# 4. SURVIVABILITY TECHNOLOGIES

We have discussed the need for mechanisms to establish and maintain connected networks to the degree possible, and to expect and survive weakly connected networks with high mobility. In this section, we will describe two technology centric thrusts to adapt to dynamic environments and to achieve connectivity: adaptive networks and satellites.

## 4.1 Adaptive and Agile Networking

Even if application and mission scenarios were uniform and known in advance, mobile wireless networks are inherently dynamic, with unpredictable time-varying channel conditions. Thus, survivable networks need network nodes and protocols that are aware of, and adapt to their environment.

### *4.1.1 Link Layer Agility*

Research prototype radios offer agility in terms of frequency bands of operation, modulation techniques, choice of MAC protocols, and power levels. These can be used to enhance LPI/LPD as well as to augment network layer survivability. An example application of this is topology control via power management [32]. Another example is dynamically switching to a different frequency band to evade jammers. A third example involves adaptive MAC layers – more survivable networks can be constructed by providing the network layer control over how the link activation is scheduled. An impediment to using link layer agility to improve survivability is the lack of standard interfaces (link and physical layer APIs) that can make this functionality available to the network layer, but some research has been done [47]. Software radios [25,9] are an important enabling technology for link and MAC adaptation.

### *4.1.2 Topological versus Geographical Routing*

Recently, geographical routing techniques have been proposed for wireless and sensor networks [24]. GPSR (greedy perimeter stateless routing) [21], Cartesian routing [17,18], and diffusion routing [20] are examples of these newer geographic routing approaches.

Static wired infrastructure tends to be better suited to topological approaches (which are already widely in use) whereas some wireless networks can benefit from geographical knowledge. Therefore, depending on the requirements, survivable nodes must be capable of supporting both of these strategies in particular, and multiple simultaneous routing and forwarding modes in general [40].

A specific example is the case in which wireless nodes choose backup or alternate paths for routes that are not only node/edge disjoint but also are diversified *geographically*. This is particularly important since physical and physical layer attacks are likely to be geographically localized.

### *4.1.3 Adaptive Nodes and Networks*

Active networking technology [43,11] provides a basis for dynamic deployment of protocol mechanisms and adaptation to traffic in the context of the wired Internet, and has been the subject of considerable research. The application of this technology to mobile wireless networking allows the dynamic selection of not only MAC and network layer parameters previously discussed, but also the ability to dynamically provision and negotiate algorithms and select entire protocols based on application requirements and the communication environment [29]. For example, sets of communicating nodes may wish to change from a simple efficient MAC protocol and routing algorithm to more sophisticated and survivable, as the environment becomes more challenging.

This eliminates the need to standardize and decide on the entire range of protocols and algorithms, and to hard-code them into nodes before they are deployed. Rather, only a framework for node discovery and protocol negotiation need be pre-determined; software radios are a key enabling technology.

Cognitive networking makes the next leap, with the potential for nodes and networks to learn about their environment, and take actions to enhance survivability.

## 4.2 The Role of Satellites

Satellites and UAVs (unpiloted aerial vehicles) can serve an important role in mitigating the effects of weakly connected channels and node mobility.

Satellites and other airborne nodes have a set of unique characteristics when compared to ground based communication nodes. The high altitude of a satellite enables it to have a very large terrestrial footprint, within which any ground node can receive communications from, and optionally communicate to the satellite. This advantage, coupled with the inherent broadcast capability of the satellite channel, enables the satellite to communicate to a large number of ground nodes, giving it a larger range than ground based nodes. On the other hand, spot beam technology, which is used on NASA's ACTS and TDRSS and the commercial Thuraya [44] satellites, can support localized

communications (and in some instances high bandwidth) between disjoint sets of ground nodes.

To a mobile node, a satellite appears in a relatively predictable point in space. Geostationary (GEO) satellites are fixed in location, while medium- (MEO) and low earth orbit (LEO) satellites have computable trajectories. UAVs may have predictable trajectories (e.g. in a racetrack formation). When the cluster or cell size are equivalent to the satellite footprint, handoffs between the node and the satellite are more infrequent than those between the node and base stations in ground-based cells, while re-acquisition and registration delays are minimized. The altitude that protects the satellite from overrun (physical attack) also has the effect of mitigating the mobility of nodes.

As a non-local resource to the mobile ground nodes, satellites have a much lower vulnerability to the threats that face ground units. Given that it is miles above the ground, a satellite is much harder to attack than any ground based communication system. Furthermore, capturing a satellite is extremely difficult and expensive, thus minimizing the chance that tactical or strategic information could be compromised due to overrun.

Satellites require high cost of deployment and operation, although picosatellites may mitigate this problem in the future. It takes a ground node considerably more power to communicate with a satellite than it takes to communicate to other nearby ground nodes. Given the power budget of a mobile wireless node is a deciding factor to its survivability [45], two-way communications through the satellite may be prohibitive for frequent transmissions. However, infrequent communications through the satellite, e.g. to register a topology after a significant change has occurred, can offer significant benefits to the overall system. Furthermore, the nodes in the wireless network could share the cost of the uplink to the satellite.

The characteristics of satellite links can significantly impact transport layer and application performance. The four major characteristics of satellite links that affect transport layer performance are: large round trip times (RTT), large bandwidth-´-delay product, burst errors on coded satellite links, and variable RTT (due to satellite movement and the handoff process) [27, 4, 5, 12, 3]. The inter-satellite links in many MEO and LEO constellations change too quickly for timely convergence of wireless network routing protocols if they are included in the topology database [49], but trajectory routing can solve this problem.

In spite of these challenges above, satellites can have significant roles in enhancing survivability.

### *4.2.1 Enhancing Connectivity*

The high altitude of the satellite also gives it different modes of communications based on obstruction than ground-based nodes. While subject to frequency dependent foliage and rain attenuation and view of the sky obstruction, satellites are not subject to the line of sight range limitations of ground-based nodes.

Satellites can be used to stitch together isolated islands of nodes that cannot directly communicate to one another, as shown in Figure 5.

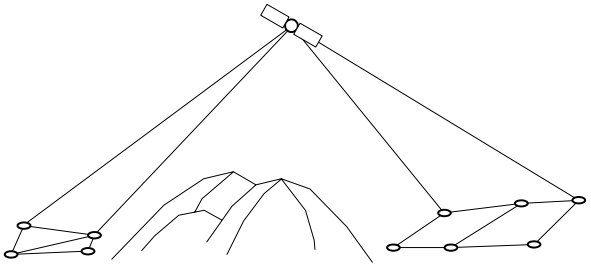

**Figure 5. Communication between disconnected islands**

### *4.2.2 Data and Control Information Dissemination*

One of the original advantages of satellite technology is that it permits the timely dissemination of information to large numbers of nodes in non-collocated areas via broadcast, as shown in Figure 6. The dissemination of topology and routing information is fundamental to the operation of a mobile wireless network. A natural one-to-many communications medium enables the information update portion of protocols, such as topology discovery, resource discovery, or routing, to have performance similar to ARP. Since topology state advertisements have to be flooded (at least per cluster), satellites can offer a distinct advantage. Advances in communication and transponder technology, such as WCDMA, now permit the establishment of multicast groups that enable directed communication to subsets of the nodes in a single area or non-collocated areas.

### *4.2.3 Support for Radio Silence*

Tactical situations can require nodes to be radio-silent. However, these nodes still need to receive situation and topology updates and new tactical information. Datacycle is the repetitive transmission of information on a channel [10]. Since the information repeats continuously, a node just needs to listen long enough to find what it needs without explicitly requesting information. The repetition of the information also minimizes an adversary's ability to ascertain receiving nodes.

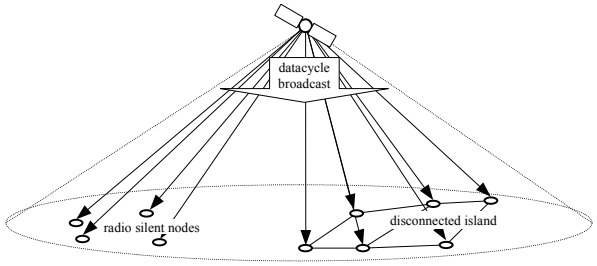


**Figure 6. Broadcast and datacycle to silent nodes**

### *4.2.4 Certificate/CRL Distribution*

The unique ability of a satellite (particularly GEO) to be both highly available and nearly omnipresent makes it a natural candidate to distribute certificate revocation lists (CRLs) and handle queries for certificates. Wide and timely dissemination of CRLs is crucial if communication to nodes with revoked certificates is to be prohibited. Besides distributing CRLs, the satellite can act as a conduit for certificate queries to a registration authority in a secure location far away from the tactical theater.

The combination of high altitude, high availability, and insensitivity to node mobility makes satellites appear to be a near-omnipresent communications resource to nodes in a mobile wireless network. Incorporating satellites into the architecture and design of the survivable mobile network offers significant advantages over deploying a terrestrial-based network alone.

## 5. SUMMARY

In summary, survivable networks require more than conventional reliability and fault tolerance. While significant progress has been made toward establishing and maintaining connected networks, further work needs to be done to understand the tradeoffs against stealth requirements (LPI/LPD/LPE).

Survivable mobile wireless networks require that asymmetric, weakly connected, and episodically disconnected links be considered as first class citizens, rather than faults that must be occasionally repaired. Similarly, mobility must be expected and exploited to enhance survivability. We propose a significant change in the way routing algorithms interact with forwarding, supporting eventual connectivity so that communication is possible in environments where it is currently not possible. Research in these areas has just begun to scratch the surface.

Since it is not possible nor practical to predict the communication environment a priori, it is critical that network nodes and protocols be able to adapt to their environment and communication scenario or mission. The support for dynamically adaptive protocols, algorithms, and parameters using active network and software radio technology are key enablers of this capability.

Finally, airborne nodes such as satellites and UAVs provide a promising infrastructure to help mitigate the effects of disconnected and asymmetric links and mobility.

## 6. ACKNOWLEDGMENTS

Rob Ruth of DARPA initially encouraged this work as part of investigating a follow-on to the GloMo program. Doug Maughan continued this path and directly supported this work as part of the SUMOWIN seedling program. Craig Partridge and the anonymous reviewer comments lead to significant structural improvements in the paper. While the co-authors were directly responsible for the text in this paper, we would like to additionally recognize the work of Isidro Castiñeyra, Martha Steenstrup, Fabrice Tchakountio, and Greg Troxel in this work.

## 7. REFERENCES

[1] Afek Y., and E. Gafni, "End-to-End Communication in Unreliable Networks," *Proc. Seventh Annual ACM Symp. on the Principles of Distributed Computing*, Toronto, Ontario, Canada, 1988, pp. 131–148.

[2] Afek, Y., B. Awerbuch, E. Gafni, E. Rosen, and N. Shavit, "Slide – the Key to Polynomial End-to-End Communication," *J. Algorithms*, vol. 22, 1997, pp. 158–186.

[3] Akyildiz, I.F., G. Morabito, and S. Palazzo, "Research Issues for Transport Protocols in Satellite IP Networks," *IEEE Personal Communications*, vol. 8, no. 3, Jun 2001, pp. 44–48.

[4] Allman, M., D. Glover, and L. Sanchez, "Enhancing TCP over Satellite Channels Using Standard Mechanisms," *RFC 2488*, Jan 1999.

[5] Allman, M., ed., "Ongoing TCP Research Related to Satellites," *RFC 2766*, Feb 2000.

[6] Arquilla, J., and D. Ronfeldt, (eds.), "In Athena's Camp: Preparing for Conflict in the Information Age," *Rand Technical Report MR-880-OSD/RC*, ISBN: 0-8330-2514-7, 1997.

[7] Awerbuch, B., O. Goldreich, and A. Herzberg, "A quantitative approach to dynamic networks," *Proc. Ninth Annual ACM Symp. on Principles of Distributed Computing*, Quebec city, Quebec, Canada, 1990, pp. 189–204.

[8] Balakrishnan, H., V. Padmanabhan, S. Seshan, and R.H. Katz., "A Comparison of Mechanisms for Improving TCP Performance over Wireless Links," *IEEE/ACM Trans. Networking*, vol. 5, no. 6, Dec 1997, pp.756–769.

[9] Bose, V.G., "The Impact of Software Radio on Wireless Networking," *Mobile Computing and Communications Review*, vol. 3, no. 1, Jan 1999, pp. 30–37.

[10] Bowen, T.F., G. Gopal, G.E. Herman, T.M. Hickey, K.C. Lee, W.H. Mansfield, J. Raitz, and A. Weinrib, "The Datacycle Architecture," *Communications of the ACM*, vol. 35 no. 12, Dec. 1992, pp. 71–79.

[11] Calvert, K, S. Bhatacharjee, E.W. Zegura, and J.P.G. Sterbenz, "Directions in Active Networks," *IEEE Communications*, vol. 36 no. 10, Oct. 1998, pp. 72–78.

[12] Chotikapong, Y., H. Cruickshank, and Z. Sun, "Evaluation of TCP and Internet Traffic via low Earth Orbit Satellites," *IEEE Personal Communications*, vol. 8, no. 3, Jun 2001, pp. 28–34.

[13] Cidon, I., and R. Rom, "Failsafe end-to-end protocols in computer networks with changing topology," *IEEE Trans. Communications*, vol. 35, 1987, pp. 410–413.

[14] Dahlberg, T., S. Ramaswamy, and D. Tipper, "Issues in the Survivability of Wireless Networks," *Proc. IEEE Mobile and Wireless Communication Networks Workshop*, May 1997.

[15] Ellison, R.J., D.A. Fisher, R.C. Linger, H.F. Lipson, T. Longstaff, and N.R. Mead, "Survivable Network Systems: An Emerging Discipline," *Technical Report CMU/SEI-97-TR-013 and ESC-TR-97-013*, Carnegie Mellon University, Software Engineering Institute, Nov 1997, Rev. May 1999.

[16] Finn, S.G., "Resynch procedures and failsafe network protocol," *IEEE Trans. Communications*, vol. 27, 1979, pp. 840–846.

[17] Finn, G.G., "Routing and addressing problems in large metropolitan-scale internetworks," *ISI/RR-87-180*, ISI, Mar 1987.

[18] Hughes, L., O. Banyasad, and E. Hughes, "Cartesian routing," *Computer Networks*, vol. 34, no. 3, Sep 2000, pp. 455–466.

[19] Herzberg, A., and S. Kutten, "Early Detection of Message Forwarding Faults," *SIAM J. Computing*, vol. 30, no. 4, 2000, pp. 1169–1196.

[20] Intanagonwiwat, C., R. Govindan and D. Estrin, "Directed diffusion: A scalable and robust communication paradigm for sensor networks," *Proc. Sixth Annual Intl. Conf. Mobile Computing and Networking*, Aug 2000, Boston, MA, pp. 56–67.

[21] Karp, B., and H.T. Kung, "Greedy Perimeter Stateless Routing for Wireless Networks," *Proc. Sixth Annual Intl. Conf. Mobile Computing and Networking*, Boston, MA, Aug, 2000, pp. 243-254.

[22] Lee, S.-J., and M. Gerla, "Split Multipath Routing with Maximally Disjoint Paths in Ad hoc Networks," *Proc. ICC 2001*, Helsinki, Finland, Jun 2001, pp.3201–3205.

[23] Linger, R.C., N.R. Mead, and H.F. Lipson, "Requirements Definition for Survivable Network Systems," *Proc. 1998 Intl. Conf. Requirements Engineering (ICRE'98)*, Colorado Springs, CO, 6–10 Apr 1998.

[24] Mauve, M., J. Widmer, and H. Hartenstein, "A Survey of Position-based Routing in Mobile Ad-Hoc Networks," *IEEE Network*, vol. 15, no. 6, Nov/Dec 2001, pp. 30–39.

[25] Mitola, J. "The Software Radio Architecture," *IEEE Communications Magazine*, May 1995, pp. 26–38.

[26] Nagpal, R., and D. Coore, "An Algorithm for Group Formation and Maximal Independent Set in an Amorphous Computer," *AI Memo 1626*, MIT, 1997.

[27] Partridge, C., and T.J. Shepard, "TCP/IP Performance over Satellite Links," *IEEE Network*, vol. 11, no. 5, Sep/Oct 1997, pp. 44–49.

[28] Perlman, R., "Network Layer Protocols with Byzantine Robustness," *PhD Thesis*, MIT Laboratory for Computer Science, Cambridge, MA, 1988.

[29] Plattner, B. and J.P.G. Sterbenz, "Mobile Wireless Active Networking: Issues and Research Agenda," IEICE *Workshop on Active Network Technology and Applications (ANTA) 2002*, Tokyo, Mar. 2002, pp. 71–74.

[30] Raju, J., and J.J. Garcia-Luna-Aceves, "A New Approach to On-demand Loop-Free Multipath Routing," *Proc. 8th Annual IEEE Intl. Conf. Computer Communications and Networks (ICCCN)*, Boston, MA, Oct 1999, pp. 522–527.

[31] Ramanathan, R., and M. Steenstrup, "Hierarchically-Organized, Multihop Mobile Wireless Networks for Quality-of-Service Support," *Mobile Networks and Applications*, vol. 3, no. 2, Aug 1998, pp. 101-119.

[32] Ramanathan, S., and R. Rosales-Hain, "Topology Control of Multihop Radio Networks using Transmit Power Adjustment," *Proc. IEEE INFOCOM*, Tel-Aviv, Israel, 2000, pp. 404–413.

[33] R. Ramanathan, "On the Performance of Ad Hoc Networks Using Beamforming Antennas," *Proc. ACM Mobihoc 2001*, Long Beach, CA, USA, Oct 2001, pp. 95–105.

[34] Royer, E.M., and C.-K. Toh, "A Review of Current Routing Protocols for Ad Hoc Mobile Wireless Networks," *IEEE Personal Communications*, Apr 1999, pp. 46–55.

[35] Software Engineering Institute, Carnegie Mellon University, "Survivable Network Analysis," *http://www.sei.cmu.edu/programs/nss/analysis-method.html*

[36] Shen, C.-C., D.-P. Hsing, T.-H. Wu, and Y. Tsai, "A Network Management Architecture for Battlefield Networks," *Proc. MILCOM'97*, Monterey, California, Nov 2–5, 1997, pp. 1226–1231.

[37] Snoeren, A., and H. Balakrishnan, "An End-to-End Approach to Host Mobility," *Proc. Sixth ACM/IEEE Intl. Conf. Mobile Computing and Networking*, Boston, MA, USA, Aug 2000, pp.155-164.

[38] Snow, A.P., U. Varshney, and A.D. Malloy, "Reliability and survivability of wireless and mobile networks," *Computer*, vol. 33, Issue 7, Jul 2000, pp. 49–55.

[39] Snow, A.P., "Network reliability: the concurrent challenges of innovation, competition, and complexity," *IEEE Trans. Reliability*, vol. 50, Issue 1, Mar 2001, pp. 38–40.

[40] Sterbenz, J.P.G., and R. Krishnan, "Multimodal Routing and Switch Architecture," *DARPA/ DOE/ NASA/ NIST/ NLM/ NSF Workshop on New Visions for Large-Scale Networks: Research and Applications*, Mar 2001.

[41] Committee T1, "T1A1.2 Working Group," *http://www.t1.org/t1a1/_a12-hom.htm*

[42] Tchakountio, F., and R. Ramanathan, "Tracking Highly Mobile Endpoints," *Proc. ACM Workshop on Wireless Mobile Multimedia (WoWMoM)*, Jul 2001, Rome, Italy.

[43] Tennenhouse, D., and D.J. Wetherall, "Toward and Active Network Architecture," *ACM Computer Communication Review*, vol. 26 no. 2, April 1996, pp. 5–18.

[44] Thuraya Satellite Communication Company of United Arab Emirates, *http://www.thuraya.com/*

[45] Toh, C.-K., "Maximum Battery Life Routing to Support Ubiquitous Mobile Computing in Wireless Ad Hoc Networks," *IEEE Communications*, vol. 39, no. 6, Jun 2001, pp. 138–147.

[46] Troxel, G.D., Personal Communication.

[47] UDAAN Transceiver API, BBN Technologies, *http://www.ir.bbn.com/projects/udaan/udaan-index.html*

[48] Vishkin, U., "An Efficient Distributed Orientation Algorithm," *IEEE Trans. Information Theory*, vol. 29, no. 4, Jul 1983, pp. 624–629.

[49] Wood, L., G. Pavlou, and B. Evans, "Effects on TCP of Routing Strategies in Satellite Constellations," *IEEE Communications*, vol. 39, no. 3, Mar 2001, pp.172–181.

[50] Zaumen, W.T., and J.J. Garcia-Luna-Aceves, "Loop-Free Multipath Routing Using Generalized Diffusing Computations," *Proc. IEEE INFOCOM '98*, Mar 1998, pp.1408–1417.

[51] IETF ZEROCONF Working Group, *http://www.ietf.org/html.charters/zeroconf-charter.html*