Chapter 1 – INTRODUCTION

Task Group IST-030/RTG-012 on 'Information Management over Disadvantaged Grids' was formed in January 2001 under the auspices of the Information Systems Technology Panel of the NATO Research and Technology Organization. The Task Group consisted of four countries: Canada, Germany, Poland and United States, with the Chairman being provided by Canada. The objective of the Task Group was the following:

> Investigation of adaptive information management schemes, implemented in the nodes of tactical command and control systems, to mitigate the effects of low bandwidth, variable throughput, unreliable connectivity and energy-constrained nodes imposed by the mobile wireless communications grid that links the command and control nodes.

The Task Group limited the scope of its study to the tactical wireless domain for a Land Force operating in a "national" context (i.e., issues related to multinational coalition interoperability were not addressed).

As part of its programme of work, the Task Group undertook to organize three informal workshops, each addressing a key issue in information management at a different level of C2IS architecture (application, middleware, communications network). The first workshop was held at DRDC – Valcartier in Canada in September 2002 on the topic of 'Data Replication over Disadvantaged Tactical Communication Links'. The second workshop was held at FGAN/FKIE in Germany in August 2003 on the topic of 'Role of Middleware in Systems Functioning over Mobile Wireless Networks'. The third workshop was held at Naval Research Laboratory in the US in June 2004 on the topic of 'Cross-Layer Issues in the Design of Tactical Mobile Ad Hoc Wireless Networks: Integration of Communication and Networking Functions to Support Optimal Information Management'.

This report discusses the third and final workshop in the series, which took place June 2-3, 2004 in Washington, DC, and was hosted by the Information Technology Division of the US Naval Research Laboratory. The technical programme of this workshop is provided in Annex A.

The objective of this workshop was to address the challenge of "top-to-bottom" information management in disadvantaged (unreliable, low bandwidth or energy-constrained) mobile wireless military communication networks. This document provides a technical evaluation of the workshop by the Task Group. Chapter 2 provides an overview of the Army tactical Command, Control, and Communications environment. Chapter 3 discusses the fundamental concepts of cross-layer networking, and indicates why such an approach to networking design is especially appropriate for wireless ad hoc networks. Chapter 4 addresses the use of cross-layer techniques in tactical military networks. Specific issues addressed include:

1) Special considerations introduced by the military environment;
2) The impact of energy-related considerations;
3) Cross-layering versus the conventional layered model; and
4) Similarities and differences between mobile ad hoc networks and sensor networks, from the cross-layer perspective.

Chapter 5 discusses the use of middleware in conjunction with (or as a possible alternative to) cross layering. Chapter 6 discusses some ways in which the application layer can benefit from information acquired at the lower layers. Finally, Chapter 7 presents a summary and conclusions derived from the major contributions of

the workshop, and highlights promising future research areas related to cross-layer network design in tactical ad hoc networks.

1.1 BACKGROUND

In contrast to the cellular networks that are the standard for commercial wireless applications, these military networks will typically be of an “ad hoc” or “infrastructureless” nature and will support a peer-to-peer, rather than client-server, computing model. Crucial issues relating to communication and networking in hostile environments include the characteristics of the physical channel, media-access mechanisms, routing, network control structures, and data structures. When viewed from the perspective of a layered network architecture, these issues span the entire protocol stack from the physical layer up through the application layer. Additionally, it is clear that the resources available to the application layer depend on the manner in which lower-layer functions are implemented. In addition to mobile ad hoc networks, wireless sensor networks were also addressed at the workshop.

The conventional way to implement networks is to use a strictly layered structure, in which issues associated with any particular layer are addressed separately, while defining interfaces to higher and lower layer functions. Although such an approach may be appropriate for wired networks, the characteristics of wireless networks suggest that improved performance may be obtained by addressing these issues in a coordinated (or “cross-layer”) fashion. The presentations in this workshop addressed alternative approaches to various aspects of the communication/networking problem, and did, in fact, demonstrate areas in which cross-layer approaches offer distinct advantages as compared to conventional strict layering. Nevertheless, it was also observed that care must be taken in assessing the compatibility of cross-layer approaches with legacy systems, as well as in anticipating potential unintended consequences of cross-layer network control.

Several of the presentations focused primarily on the “lower layers” of network design, and emphasized the impact of physical layer characteristics (and the need to adapt to them as they change) on media-access control and routing. Other presentations emphasized user applications, and how the application layer should be coordinated with the lower network functions to provide improved functionality in the presence of changing user demands as well as changing threats. Additionally, the role of middleware (the focus of the Task Group’s second workshop) to bridge networking functions with the needs of the application was addressed.

Chapter 2 – ARMY TACTICAL COMMAND, CONTROL AND COMMUNICATIONS ENVIRONMENT

2.1 MILITARY COMMAND AND CONTROL SYSTEM STRUCTURE

Command and control systems must support three types of relationships: command, support and proximity. The following extract is taken from [1]:

> "Command relationships exist whenever one unit or formation commander is a direct subordinate of another…Command requires a rich bi-directional exchange of information between the higher headquarters and the subordinate headquarters. The purpose of this exchange is to pass command information (plans, orders, task organization, battlefield geometry, alerts, warnings and status) between the two headquarters. This type of information exchange follows the parent-child relationship. The superior (parent) headquarters supplies directive information to the subordinate (child) headquarters – higher to lower; the subordinate (child) headquarters provides status information to the superior (parent) headquarters – lower to higher.
>
> Support relationships are a particular type of command relationship. Support relationships are established when one organization must aid, protect, complement or sustain another force. In the context of Command and Control (but not, for example, fire control) these organizations have the same requirements for information exchange as command relationship. Support relationships are of two types: Direct Support and General Support.
>
> Proximity relationships exist when units with no direct command or support relationship are operating in proximity to each other and must exchange non-command information in order to establish and maintain situational awareness. Examples of this type of relationship could be the flank coordination of adjacent tank and infantry battalions or the forward passage of lines of an armoured regiment through a mechanized infantry battle group. In proximity relationships, information flows horizontally between the headquarters of the units involved as peers, not parents/children. Units involved in proximity relationships may be subordinate to different higher headquarters."

The military command and control structure is hierarchical (Figure 1). A headquarters at a certain command level will be parent of one or more subordinate headquarters, and will itself be subordinate (child) to a higher headquarters. The communication infrastructure and flow of information over that infrastructure reflects this command hierarchy. A commander at a given level will generally be required to maintain information from one level up and two levels down in the command hierarchy, as well as from flanking formations with which he has a proximity relationship.

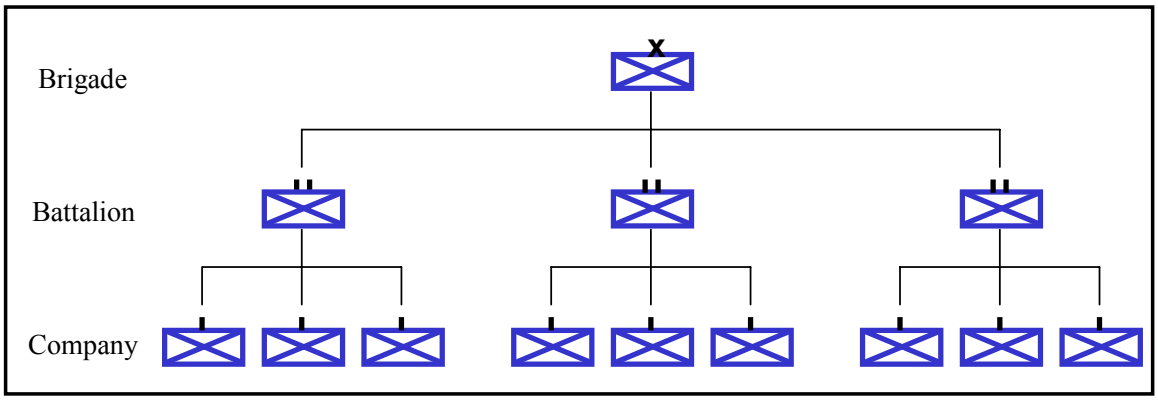


Figure 1: A Typical Hierarchy of Command Headquarters.

2.2 COMMAND AND CONTROL COMMUNICATIONS INFRASTRUCTURE

The command and control communications infrastructure is generally organized as a series of hierarchical subnetworks. A brigade command subnet will include a node at each of the three battalion headquarters under the brigade's command. A battalion command subnet may have a dozen or more nodes, but will include a node at each of the three company command posts under the battalion's command. Links from battalion headquarters to brigade or higher headquarters are typically provided by relatively reliable and high-bandwidth satellite links or dial trunk systems employing wire or wideband microwave links. The workshop focused its attention on the Army battlefield environment forward of a battle group or battalion[1] headquarters in which all communication occurs between mobile nodes equipped with a combat net radio operating in either the Very High Frequency (VHF) or Ultra High Frequency (UHF) band. A radio subnetwork consists of a set of radios tuned to a common assigned frequency. The subnet is linked to an adjacent subnet through a relay or gateway node that is common to both subnets. The relay node contains two or more radios, each tuned to a different subnet frequency. A verbal message received at the relay node from a sending node on one subnet is recorded in writing by a human operator, and then retransmitted verbally on the appropriate subnet to reach the destination addressee(s) on that subnet, if required. A gateway performs the same function for a data transmission, except that caching of the received transmission, and retransmission on the target subnet(s), are handled automatically by the gateway.

2.3 COMBAT NET RADIO COMMUNICATIONS ENVIRONMENT

Forward of battalion, communication occurs over line-of-sight (LOS) radios operating in the VHF or UHF bands. Most of these radios are vehicle-mounted, but man-portable versions are also employed. At the lowest echelon, such as an infantry section on foot engaged in urban warfare, soldiers may use short-range radios operating in the High Band UHF. These radios have ranges of 400 meters or less. UHF radios are used for medium-range LOS wireless communication from 400 meters to 15 kilometres. VHF Radios must be used where non-LOS communication beyond 15 kilometres is required.

[1] Battalions are either mechanized battalions (two armoured companies, one infantry company) or infantry battalions (two infantry companies, one armoured company). A battle group is a battalion augmented with other assets (usually engineer or artillery assets).

The vast majority of VHF combat net radios have base bandwidths of 16 kilobits per second (Kbps) half duplex[2]. UHF radios can have base bandwidths as high as 288 Kbps full duplex. High Bandwidth UHF radios can have base bandwidths as high as 11 Mbps. In spite of their limited bandwidth, combat net radios operating in the VHF band still predominate because of their greater range and beyond-LOS capability.

In practice, the useable throughput at the application layer is typically a small fraction of this base rate. Factors such as forward error correction (FEC), encryption overhead (e.g., crypto synchronization sequence), acknowledgement request and retransmission mechanisms at data link and transport layer, and media access control mechanisms, are responsible for this reduction. Moreover, the effective throughput at any time may fall well below this maximum value due to variations in the performance of the physical radio channel caused by real-world factors such as terrain interference, atmospheric interference, multi-path (reflections) and prolonged fading. Effective throughput can become zero for periods of time for certain links, or for the entire subnet (in the case of imposed radio silence).

A VHF combat net radio with a base rate of 16 Kbps may have a maximum effective throughput as low as 1 Kbps due to the factors quoted above. Consequently, if ten users share the radio subnet, the maximum effective throughput per user will be only 100 bps. This figure assumes a data-only network with no voice contention for the channel. The reality is that the residual bit-error rate (BER) (i.e., the BER after error-control coding) for these radio channels can be as high as 10^{-5}.

One known UHF combat net radio can have a base rate as high as 288 Kbps. However, UHF radios have shorter ranges than VHF radios since the higher frequency UHF waves are more susceptible to the real-world factors quoted above. For the case considered, field measurements suggest that the maximum effective throughput at the top of the data link layer in a tactical environment would be approximately 16 Kbps for a link with a 22 Kbps base rate and approximately 80 Kbps for a link with a 100 Kbps base rate. For a subnet with ten users, maximum effective throughput per user would be 1.6 Kbps and 8.0 Kbps respectively. Residual BERs are similar to rates experienced with VHF radios.

In the Internet world, the transport protocol of choice for most applications is the Transport Control Protocol (TCP). TCP provides reliable packet delivery using a sequencing and positive acknowledgement scheme and is a connection-oriented protocol as it establishes a connection with each recipient. Although point-to-multipoint addressing schemes are being developed for TCP, in the radio domain the vast majority of systems still use a point-to-point addressing scheme for TCP. Therefore, in this domain, if a transmission is intended for N recipients, it must be sent N times. In a highly bandwidth-constrained wireless environment, the communication overhead associated with the use of a connection-oriented transport protocol like TCP is generally unacceptable. As well, the congestion-control mechanisms used by TCP were developed for the wired domain. TCP interprets latency and packet loss as evidence of congestion, to which its reaction is to lower the offered load to the network. Consequently, use of TCP results in a significantly lower throughput than the network can actually provide. In low-bandwidth, high-latency, and relatively high error-rate networks like those found in the tactical radio domain, many TCP connections would be terminated due to these effects. User Datagram Protocol (UDP) is a connectionless alternative to TCP. However, UDP is unreliable since it does not employ sequencing or acknowledgement. Custom middleware operating just above UDP may be required to optimise reliability of packet transmissions in a connectionless, high-BER tactical wireless environment.

[2] Half duplex means that the radio can either send or receive on the same channel, but cannot simultaneously send and receive. Full duplex means that the radio can simultaneously send and receive on the same channel.

The Internet model is based on a rigidly layered model in which functions at different layers are defined independently, and in which information at any individual node is exchanged only among adjacent layers. Such a layered model is appropriate for wired networks such as the Internet. However, as is explained in Chapter 3, alternative approaches may be appropriate for wireless networks. Such cross-layer approaches were the subject of this workshop.

Chapter 3 – FUNDAMENTAL ISSUES IN CROSS-LAYER NETWORK DESIGN AND CONTROL FOR WIRELESS NETWORKS

In the traditional view of networking, user applications view the network as a service provider, and are not concerned with the characteristics of the network, as long as it can support the desired user traffic. This approach led to the development of layered network architectures, which facilitate modular network design and interoperability. Such approaches have distinct advantages for wired networks such as the Internet. However, they may not be appropriate for ad hoc wireless networks such as tactical networks. This chapter discusses how the characteristics of ad hoc wireless networks differ from those of wired networks, and even from other wireless networks such as cellular networks. Consequently, it is appropriate to consider novel approaches to networking.

In addition to their obvious differences from wired networks, wireless ad hoc networks are also fundamentally different from the cellular systems and wireless local area networks (LANs) that have been developed in the commercial domain. The primary differences result from the lack of an infrastructure. For example, cellular systems have fixed base stations, which communicate among themselves using dedicated non-wireless lines; thus, the primary issues to be addressed in cellular systems involve tracking the mobile users. Otherwise, wireless communication is limited to that between mobile users and base stations. In fully connected wireless LANs, since there is single-hop connectivity among all the nodes, routing is not an issue.

However, unlike the case of wired networks, the set of network links and their capacities are not determined a priori, but depend on factors such as distance between nodes, transmitted power, error-control schemes, other-user interference, and background noise. Thus, even when the physical locations of the nodes are fixed, many of the factors that affect network topology (and hence network control schemes) are (at least partially) influenced by the actions of the network nodes. Therefore, the wireless networking environment poses many new challenges not encountered in non-wireless or cellular networks.

Section 3.1 addresses the conventional approach of layered network design that is used in wired networks such as the Internet. Section 3.2 discusses some characteristics of wireless networks that are markedly different from those of wired networks, and which suggest that novel approaches are needed to provide good performance in wireless networks. Finally, Section 3.3 provides examples of how cross-layer techniques can be used to approach the design and control of ad hoc networks. Many of the major issues in cross-layer issues in ad hoc wireless networks are discussed in [2].

3.1 LAYERED NETWORK DESIGN

A hierarchical, or layered, structure is typically used to reduce a network's complexity. Figure 2 shows one commonly used abstraction of the layered protocol stack; it is similar to the seven-layer OSI reference model, except that it does not include the session layer (OSI Layer 5) or presentation layer (OSI Layer 6).[3] In such an architectural design, each layer offers services to the layer above it. Thus, for example, the application layer, which represents the functionality the user would like to obtain from the network, interfaces directly only with the transport layer. It is shielded from the lower layer functions, and consequently it does not have to know how they are implemented. These lower layers are the network layer (which performs routing in multihop networks), the data link (or simply link) layer (which performs media access control, error control, etc.), and the physical

[3] These two layers are not included in the TCP/IP protocol stack either.

layer (which supports communication over the specific medium, which may be wire, fiber, wireless, etc.). An excellent description of layered network structures is provided in the textbook by Tanenbaum [3].

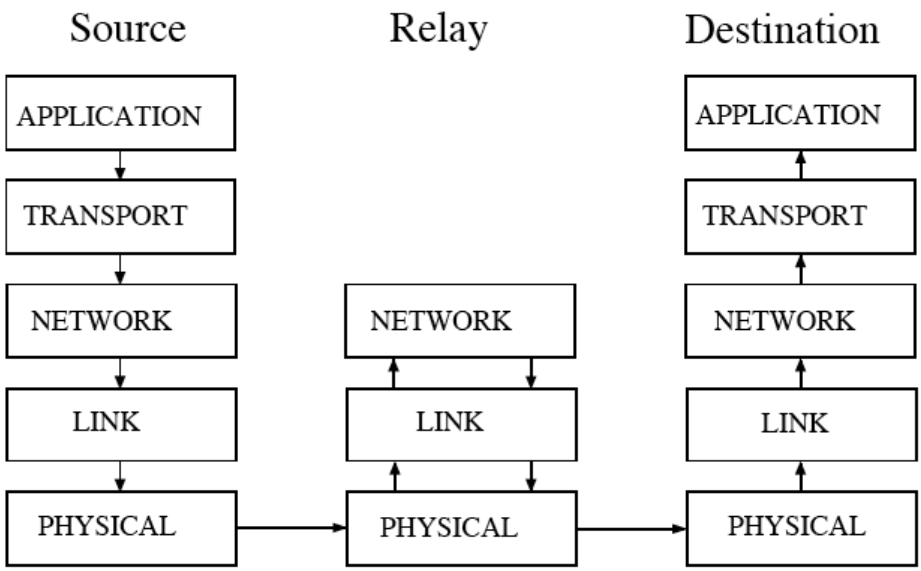


Figure 2: The Conventional Layered Protocol Stack.

Such a modular design permits the development of "open standards" that facilitate both hardware and software development. Updates to individual layers are possible without disturbing the overall network structure. Such updates may be introduced as a consequence of improved network design (e.g., improved equipment or improved algorithm design) or updated user requirements.

In the example of Figure 2, the network provides a service in which the application layer at the Source node communicates with the application layer at the Destination node. It does so by using the functions of the transport layer, which in turn employs the network layer, and so on down the stack. Note that the functions of relay nodes involve only the network layer and those below it.

3.2 CHARACTERISTICS OF AD HOC NETWORKS

In his keynote address, Prof. Michael Pursley set the stage for the workshop by presenting an overview of the fundamental issues associated with tactical ad hoc wireless networks and by indicating how cross-layer approaches can provide improved performance. This section, which highlights some of the fundamental differences between wireless ad hoc networks and wired networks, is based largely on his presentation [4]. As a consequence of these differences, novel approaches that deviate from the rigid layered structure discussed in Section 3.1 may provide improved performance. Moreover, in addition to their obvious differences from wired networks, wireless ad hoc networks are also fundamentally different from the cellular systems and wireless local area networks (LANs) that have been developed in the commercial domain. The primary differences result from the lack of an infrastructure. For example, cellular systems have fixed base stations, which communicate among themselves using dedicated non-wireless lines; thus, the primary issues to be addressed in cellular systems involve tracking the mobile users. Otherwise, wireless cellular communication is limited to that between mobile

users and base stations. In fully connected wireless LANs, since there is single-hop connectivity among all the nodes, routing is not an issue.

In ad hoc wireless networks it is possible to establish a link between any pair of nodes, provided that the signal-to-noise ratio at the receiving node is sufficiently high. Thus, unlike the case of wired networks, the set of network links and their capacities are not determined a priori. Factors relating to the existence of a link include:

- Distance between nodes;
- Transmitted RF power;
- Background noise;
- Data rate;
- Error-control code rate;
- Modulation scheme;
- Other-user interference; and
- Quality of service (QoS) requirements.

Thus, even when the physical locations of the nodes are fixed, many of the factors that affect network topology can be (at least partially) influenced by the actions of the network nodes. While some of the issues listed above are obvious, others may not be. For example, the interference level at a node depends not only on background noise (and possibly jamming) levels, but also on the interference caused by other nodes; thus, the mechanism used to schedule transmissions affects the interference level at nearby nodes. The specification of data rate and error-control code rate (along with the modulation scheme) affect the bit-error rate (BER), and hence impact on whether or not the desired QoS requirement is satisfied. Perhaps more subtle is the fact that the specified QoS level determines whether or not a link is present; reduction in the acceptable level of QoS permits the use of a link, but would be appropriate only if the user application can tolerate such a reduced QoS.

Furthermore, in ad hoc networks no distinction can be made between uplink and downlink traffic[4], thus greatly complicating the interference environment. Therefore, the wireless networking environment poses many new challenges not encountered in either wired or cellular networks.

3.3 EXAMPLES OF POTENTIAL CROSS-LAYER RELATIONSHIPS IN TACTICAL AD HOC NETWORKS

Figure 3, taken from Prof. Pursley's presentation [4], illustrates many of the potential interactions among communication and networking functions at various layers of the protocol stack. The most obvious interactions are among the lowest three layers (physical, data link, and network). However, the impact of the higher layers is apparent as well, once the applications that must be supported are addressed.

[4] Cellular networks are characterized by a hierarchical structure, in which, there is an uplink between mobile users and the base station and a downlink between the base station and the mobile users. The uplink and downlink communication generally use distinct (orthogonal) channels. However, in peer-to-peer architectures such as those of ad hoc networks, a single channel is typically shared by all users, which may use multihop routes from the source node to the destination node.

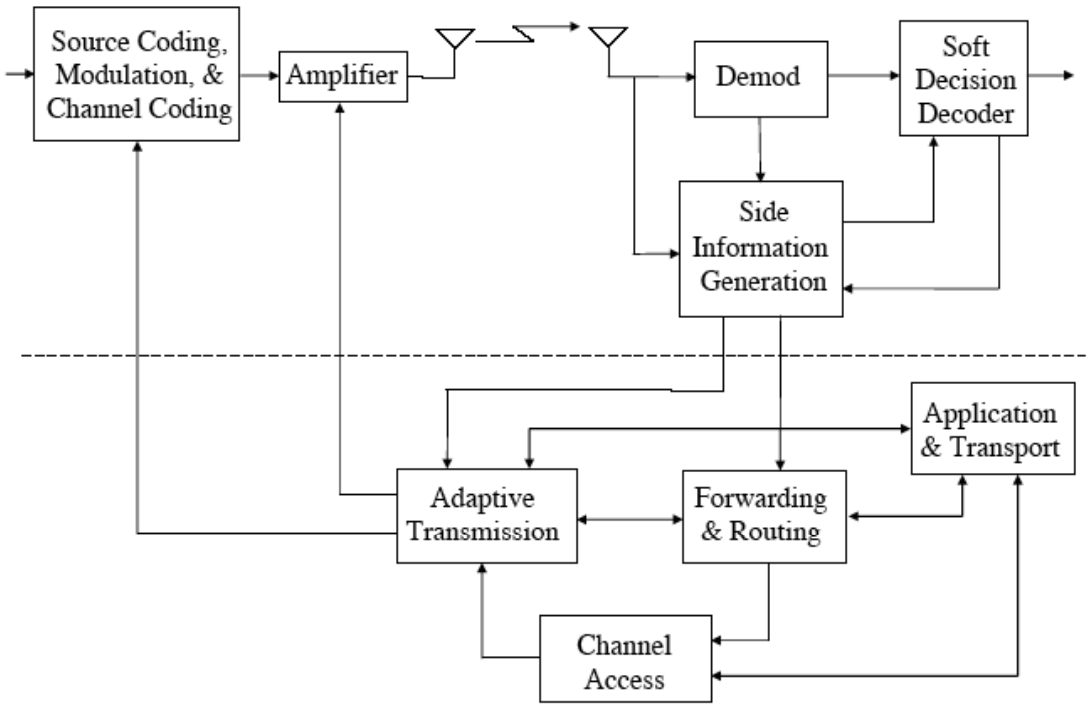


Figure 3: Some Protocol Interactions in Wireless Networks.

To illustrate the need for cross-layer protocols, Prof. Pursley presented an example based on voice messaging. Voice traffic typically requires low delay to preserve its intelligibility and timeliness, but can tolerate significantly higher error rates (frame erasures) than data traffic. Moreover, voice traffic typically consists of relatively long sessions, rather than individual packets. The impact of these requirements at the various layers is summarized as follows:

- Application layer: Speech compression must match available routes and links and satisfy QoS needs (e.g., intelligible speech vs. speaker recognition).
- Network layer: Routing should emphasize the need for low delay; high-quality (low BER) routes are not needed.
- Data link layer: To accommodate a voice session, it is necessary to reserve multiple time slots on each link via the media access control (MAC) protocol. Detected packet errors may not require retransmission (because some errors can be tolerated).
- Physical layer: Code rate should adapt to link quality. Low-rate codes should be used on poor links (to avoid need for retransmissions), and high-rate codes on good links (to reduce delay). Energy conservation is secondary to need for timely delivery when considering voice traffic.

In addition to illustrating the relationships among functions at various layers, this example highlights the particular dependence on an application that involves session-based voice messaging. For example, if packet-oriented data communication were considered, it might be possible to consider the imposition of less-stringent delay requirements; however, more-stringent BER requirements would be appropriate. Therefore, the cross-layer dependencies depend strongly on the particular application that is being supported.

When considering multimedia traffic, the various classes of traffic (e.g., voice and data) should be handled differently through the use of adaptive transmission and routing protocols, based on trade-offs such as those discussed here.

The potential benefits of cross layering are greatest in ad hoc wireless networks because of the strong interrelationships among the physical, data link, and network layers. However, some benefits may be possible in wired networks as well. For example, Prof. Klara Nahrstedt claimed during the panel discussion [5] that if the Internet had exploited cross-layer design, application QoS would already be supported.

Chapter 4 – CROSS-LAYER ISSUES IN TACTICAL MILITARY NETWORKS

This section is based largely, although not exclusively, on the expert panel discussion held at the end of the workshop. A list of panellists is provided in the technical programme (see Annex A).

4.1 SPECIAL CONSIDERATIONS INTRODUCED BY THE MILITARY ENVIRONMENT

Some characteristics of wireless networks that distinguish them from wired networks, and which suggest the potential benefits of the use of cross-layer techniques were discussed in Section 3.2. The panellists addressed some unique considerations imposed by the tactical military communications environment, which make it significantly different not only from wired networks, but from typical wireless networks as well. The primary issues identified by the panel are the following:

- Tactical military equipment can support only low data rates.
- Heterogeneous equipment with different capabilities must function in the same network.
- Environment is hostile (e.g., jammers, node destruction).
- Applications with very different requirements and priorities must be supported.
- Widely varying communication conditions and network topologies must be supported.
- Legacy systems must be supported while transitioning to future systems.

The general consensus among the expert panel members, and most attendees, was that cross-layer approaches can, in fact, provide improved performance relating to at least some of these problems. For example:

- Adaptation and diversity can provide robustness to jamming and destruction or compromise of nodes.
- Cross layering can support different requirements (e.g., voice, data) and priorities across all layers of the network protocol stack.
- Cross layering can adjust higher layer protocols to the capabilities of underlying equipment.
- Cross layering can adapt to and provide robustness against variations in the communication capabilities and network topology.
- Cross layering can allow nodes to use information obtained by one layer at a higher or lower layer as well (particularly important to permit exploitation of network status information by the application/middleware layers).

However, it will be difficult to overcome some obstacles, such as the need to support communication with legacy systems that cannot provide the necessary degree of adaptivity. For example, new military radio systems such as the USA's Joint Tactical Radio System (JTRS) will be able to use adaptive cross-layer protocols to save energy and to improve QoS performance. However, legacy SINCGARS radios do not have the necessary degree of adaptivity to support cross-layer operation. For example, SINCGARS radios have only manual power settings, and SINCGARS ACKs do not provide the necessary physical-layer information to permit JTRS equipment to adapt intelligently in joint JTRS-SINCGARS networks.

4.2 THE IMPACT OF ENERGY-RELATED CONSIDERATIONS

Energy-awareness is a crucial aspect of ad hoc networks because the nodes are typically powered by batteries. For example, the batteries carried by a soldier may constitute a significant fraction of the overall load that must be carried. Therefore, means to reduce energy consumption are extremely desirable. The traditional approaches to energy reduction involve the development of energy-efficient electronics and energy-efficient modulation schemes, as well as the use of directional antennas (which focus the beam in the desired direction, thereby eliminating the transmission of energy in unnecessary directions). Additionally, energy awareness can be viewed from the networking perspective, which involves multiple layers in the protocol stack.

Two basic forms of energy-aware network operation can be considered. Under "energy-efficient" operation, the goal is to maximize the number of bits that are delivered per unit energy over a period of time. In this mode of operation, energy use may be considered as a cost (e.g., the cost of replacing the batteries). However, in some applications batteries cannot be replaced during the course of a mission. Such a situation can arise when soldiers are unable to return to base, or alternatively in the case of sensor networks. This case is referred to as that of "energy-constrained" operation.

The introduction of hard constraints on the total amount of energy available at each node results in a problem that is very different from that in which unlimited energy is available (although energy efficiency still may be desired). Under such hard constraints on energy, the network is capable of operation for a limited period of time. A node dies (and hence can no longer transmit) when its energy is depleted, and the network dies when it is no longer capable of providing a minimum acceptable level of service. By contrast, when the goal is energy efficiency, it is implicitly assumed that ample energy is available as long as one is willing to pay the cost.

It is important to note that energy-efficient operation does not ensure good performance in energy-constrained applications. For example, use of the most energy-efficient routes may result in premature depletion of energy at some nodes.

The issue of energy awareness (in both its energy-efficient and energy-constrained forms) crosses several layers of the protocol stack. One obvious trade-off is that of energy used for signal processing versus that used for communications. For example, signal-processing algorithms that significantly compress the data can provide benefits to overall network operation by reducing the number of bits that must be transmitted, thereby saving RF energy and reducing bandwidth requirements. However, the reduction in energy may be outweighed by the increased energy needed for data compression and decompression operations. Energy consumed by the nodes' hardware is an especially significant component of overall energy consumption in short-range networks, where RF transmission energy is relatively low. Furthermore, the amount of energy consumed by hardware can be reduced by reducing the bit duration of the transmitted symbol, whereas RF energy can be reduced by doing the opposite. Therefore, trade-offs between energy and delay must be considered.

Energy may be saved by the use of sleep modes, because nodes consume energy even when they are not transmitting or receiving. However, use of such sleep modes complicates many aspects of networking, including synchronization, routing, channel access, sensing functionality, etc. Therefore, there is a need to coordinate functions at several layers of the protocol stack.

One example of a networking scheme in which cross-layer techniques are used for energy-aware networking is that of Wieselthier et al's Broadcast Incremental Power (BIP) algorithm [6], which produces energy-

efficient broadcast trees. BIP is based on the joint choice of transmission power and tree structure, and thus represents joint control of physical and network (routing) layer functions. Although BIP was originally developed for energy-efficient operation, it was later extended to energy-constrained operation as well. It was demonstrated that incorporating a measure of the "residual energy" of each node (i.e., the remaining energy in the node's battery) into the cost function used in the tree-construction algorithm can provide a measure of load sharing that can extend network lifetime greatly.

The bottom line is that the introduction of energy considerations, especially in energy-constrained operations, has a major impact on network design and control. Most importantly, it introduces trade-offs among performance measures such as delay, throughput, and node/network lifetime, and necessitates tight coupling among the layers if near-optimal performance is to be obtained. In fact, in addressing energy considerations Prof. Andrea Goldsmith [7] suggested that it is appropriate to take a novel view of the protocol stack in which a "hardware layer" is inserted below the physical layer.

4.3 CROSS-LAYERING VS THE CONVENTIONAL LAYERED MODEL

There was a strong consensus at the workshop that the use of cross-layering techniques does not mean that the layered architecture should be abandoned. To the contrary, it has worked well in the Internet, and its modular structure (see Section 3.1) provides an efficient and scalable framework for network design. Nevertheless, significant performance improvement can be expected if cross-layer techniques are implemented in ad hoc networks.

Thus, cross-layer design is not about eliminating layers, but is rather about designing across them. Wireless networks can benefit most from cross-layer design, but benefits are possible for wired networks as well. For example, Prof. Klara Nahrstedt claimed during the panel discussion that if the Internet had exploited cross-layer design, application Quality of Service (QoS) would already be supported. The degree of improvement that can be achieved depends strongly on the type of network. For example, sensor networks are expected to benefit more from cross layering than general mobile ad hoc networks (see Section 4.4). The research community is only beginning to understand the nature of cross-layer design, and still needs to determine where significant cross-layer gains are possible.

Despite the potential benefits of cross layering, a cautionary note was sounded concerning the possible unintended consequences of some cross-layer interactions.[5] For example, the tight coupling of layers may lead toward a tendency to develop proprietary protocols, and hence the need to redesign a new system for every application, thereby eliminating many of the benefits originally obtained by layering. In addition, tight coupling of layers may lead to "spaghetti code" in which patches are continually added to improve performance, resulting in a system that is difficult to understand and hence difficult to update to accommodate changing requirements. Furthermore, the performance of an optimally designed system may be highly sensitive to the operating point, and minor errors in system parameters or minor environmental or topological changes may result in significant performance degradation from that at the optimal operating point. Finally, it was suggested that an optimized cross-layer system might be a vulnerable target for an intelligent adversary; however, others present suggested that cross-layering might improve network robustness against such threats. It should be emphasized that this cautionary message was intended not to discourage the study of cross-layer techniques, but rather to warn of the unintended consequences of protocol interactions that are not yet well understood.

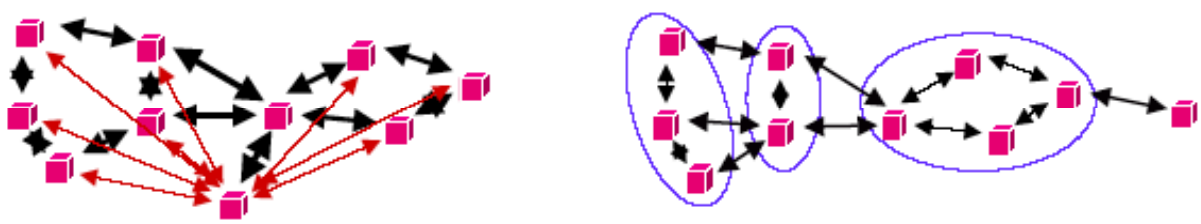
[5] See a recent paper by Kawadia and Kumar [8] for a discussion of such issues.

Key issues in cross-layer research and development raised at the workshop include:

- Development of the right framework for cross-layer design;
- Determination of information to be exchanged across layers, and how to use it;
- Balancing of adaptivity, diversity, and scheduling;
- Identification of the key cross-layer synergies, and which layers should be involved;
- Avoidance of unexpected interactions across layers;
- Management of cross-layer complexity; and
- Accommodation of legacy systems and protocols.

4.4 SIMILARITIES AND DIFFERENCES BETWEEN MOBILE AD HOC NETWORKS AND SENSOR NETWORKS

Wireless ad hoc networks and sensor networks are similar, in that both classes of networks are infrastructureless and use the wireless channel, as shown in Figure 4. In fact, in many ways sensor networks may be viewed as a special case of ad hoc networks. However, there are some significant differences between these types of networks, typical characteristics of which are summarized in Table 1. Although the primary focus of the workshop was ad hoc networks, rather than sensor networks, three of the oral presentations did, in fact, address ad hoc networks. Furthermore, a comparison of ad hoc and sensor networks facilitates the understanding of the role of cross layering in ad hoc network design and control.



(a) Wireless ad hoc network (b) Wireless sensor network

Figure 4: Examples of Ad Hoc and Sensor Networks.

Table 1: Typical Characteristics of Ad Hoc Networks and Sensor Networks

Ad Hoc Networks	Sensor Networks
• Peer-to-peer with no backbone infrastructure.	• Data flows to a centralized location.
• Nodes typically mobile.	• Nodes typically stationary.
• Network size: up to tens of nodes.	• Network size: hundreds or thousands of nodes.
• Nodes can be well-equipped.	• Nodes typically have limited capability.
• Nodes generate independent information.	• Node information correlated in time and space.
• Can require high data rates.	• Low per-node rates, but large number of nodes.
• Typically support multiple applications.	• Typically support a single application.
• Batteries can usually be recharged or replaced.	• Nodes typically energy-constrained.

It was concluded that both types of networks can benefit from cross-layer design, but that sensor networks typically will benefit more. One of the primary concerns in sensor networks is that nodes are typically energy constrained. (Some examples of mobile ad hoc networks, such as those in which the nodes consist of individual soldiers, are also energy constrained because it may not be possible to replenish batteries during the course of a mission.) Therefore, it is essential that energy use be optimized across the protocol stack. Additionally, the fact that sensor networks are generally designed for one dedicated purpose (e.g., target detection, surveillance) permits a tightly coupled design in which only those functions that are necessary need to be supported. By contrast, ad hoc networks may be more constrained by existing standards and the need for interoperability, which may make cross-layer design more difficult.

Chapter 5 – THE ROLE OF MIDDLEWARE IN CROSS-LAYER ARCHITECTURES

Middleware, the subject of the second workshop held by Task Group 12, and one of the topics addressed in Prof. Klara Nahrstedt's presentation [5] at the panel discussion of the 'Cross-Layer' workshop, is generally seen as a software system that connects (user) applications to service providers in a manner transparent to the application. The middleware establishes and manages the connection. Therefore, the application has the same interface, independent of whether the service is on the same machine or on a remote one. Ideally, the same interface could be used independent of the characteristics of the underlying network, which could be, e.g., the Internet, a local area network, or a tactical wireless ad hoc network. Thus, only the middleware (but not the application-layer software) would be required to adapt to the specific network environment.

An important goal of middleware is to enable the management and sharing of network resources by several applications, which may be located at either the same or at different network nodes. Although considerable effort will be needed for the design of effective middleware, it is believed that the benefits will be significant, especially when such resource sharing among applications is necessary. All applications using the middleware can benefit, independent of the percentage of the resources they need. However, in networks in which middleware is used to support only a single application, it is possible that the effort to produce cost-effective middleware may be justifiable only if the application is known to produce a major part of the resource allocation.

The advantages of middleware are greatest in environments where different network technologies are used, or in which the technology changes frequently. Since only the middleware has to adapt to the network technology, it is not necessary to adapt each application programme for every possible network interface. Thus the middleware would shield the applications from such variations in the networking environment. Moreover, middleware may be useful in enabling application-layer software to adapt to the time-varying conditions that characterize disadvantaged tactical networks.

For networks with ample resources, the use of well-designed middleware may be all that is needed to obtain nearly optimal performance of the deployed applications. However, the situation is different in tactical networking environments, which are characterized as disadvantaged because of the limited resources that are available (e.g., bandwidth, energy). In such networks, well-designed middleware is certainly needed to address crucial resource management issues, but it is not sufficient. Chapter 4 illustrated the potential benefits that can be obtained in disadvantaged wireless networks by using cross-layer techniques. These benefits cannot be achieved simply by inserting a middleware layer on top of a rigidly layered network protocol stack.

In conclusion, middleware is certainly an important part of the communication architecture in future wireless tactical networks. However, in disadvantaged environments such as these, it is not a substitute for cross-layer design. Middleware complements cross-layering designs, but cannot replace them.

Chapter 6 – THE APPLICATION PERSPECTIVE

It is important to distinguish between (1) the "user applications" that employ the network and (2) the application layer within the seven-layer OSI model, which contains applications that provide protocols and services to support the user applications. Examples of application-layer services are File Transfer Protocol (FTP) and Hypertext Transfer Protocol (http). These application-layer services generally provide a function that is generic and useful to a wide variety of user applications. User applications include custom-designed applications that incorporate business rules specific to an enterprise (e.g., an intelligence analysis application) and generic applications that can be adapted to support an enterprise's objectives and business rules (e.g., office applications like Word or Excel with the ability to create custom functionality through macros).

The previous chapter discussed the use of middleware to eliminate the need to design special-purpose application-layer software for each networking environment. Moreover, middleware may be useful in enabling application-layer software to adapt to the rapidly changing network conditions that characterize disadvantaged tactical networks. However, independent of whether the user application accesses the network directly or via a middleware, the question of primary interest to the Task Group was the following:

> 'How can the user application benefit most from the information available in the different layers of the network?'

This question is large in scope and difficult to answer adequately in a few lines because, as Klara Nahrstadt noted at the workshop [5], different application level objectives make different demands on the network layers, viz:

- **A Timing Objective** places demands on lower level scheduling (transport, connection and MAC scheduling) and network routing.
- **A Loss Rate Objective** places demands on error detection and control in the data link/physical, transport and network (routing) layers.
- **A Bandwidth Objective** places demands on transport flow, congestion control, and channel access.
- **An Energy Objective** places demands on the transport, network and data link (MAC) layers.
- **A Mobility Objective** places demands on the network (routing) and data link (MAC) layers.

The important information for an application to know is the extent to which the network can, or cannot, satisfy its objective(s). In a wireless network, this information is typically not made explicitly available to the application, leaving the application to infer indirectly the network state. For example, an application on a particular tactical node may infer, from the slow depletion of its output buffer, that there is an elevated level of channel contention, and/or high channel utilisation, but cannot know how congested the network really is. As various workshop presentations made clear, information about the present offered load to the network, degree of channel utilisation, network topology, throughput at the application layer, bit-error rate, or battery state could, depending upon the scenario and the application's objectives, be very useful to the application in adapting both its behaviour and user expectations to changing network conditions.

Nahrstadt mentions two specific examples:

- Database replication, where connectivity concerns may be paramount because the objective is to replicate the same data to many nodes reliably and simultaneously; and

- Real-time applications like Voice-over-IP applications where information about delay and reliability of the communication path are required. Such information can be used to adapt parameters like vocoder quality, error-correction code rate, or packet size and transmission frequency.

Assuming that cross-layered network design techniques are available to make some of this information available to the application, the question arises as to how the interface between the application and the network should be designed. Where support for legacy applications is required, a traditional Application Programming Interface (API) that respects the traditional layered structure will be required. For future applications seeking to exploit cross-layered design, one or more APIs could be designed which make information about the network (or even lower layers) available to the application. However, the raw information may have to be analyzed, collated or reformatted to render it in a form that the application can use. It is logical to employ middleware to provide this function.

Chapter 7 – SUMMARY AND CONCLUSIONS

The overwhelming consensus among the expert panel members and participants in the workshop was that, to fully exploit the nature of the wireless channel, some use of cross-layer techniques would be necessary in future wireless network applications, both in commercial and military environments. Appropriate use of cross-layer techniques would not involve the abandonment of the layered protocol structure; rather, cross layering would be used to augment the network's capability by sharing information among the layers and by jointly optimizing their performance. There appears to be considerable potential for performance improvement. Nevertheless, this field is still in its early stages of development, and the research community does not yet have sufficient insight to understand the big picture. One fundamental question, which has not yet been answered, is that of which layer interactions provide the best opportunities for performance improvement. The research community is just starting to ask the right questions, and there is now a basis for fruitful research and development.

The wireless networking environment, particularly in ad hoc network examples, is quite different from that of wired networks. Consequently, the properties of the physical layer play a large part in ad hoc network design and performance, and cross-layer design techniques are especially well suited for them. Some potential benefits of the use of cross-layer techniques for wired networks, such as the Internet, were mentioned at the workshop as well.

Energy concerns are extremely important in ad hoc networks, and especially in sensor networks (which may be viewed as a special case of ad hoc networks). The fact that a finite quantity of energy must be shared among all of a node's functions (e.g., transmission, reception, and signal processing) strongly links virtually all layers of the protocol stack. In fact, it was stated that such energy constraints change *everything* about network design. It is now appropriate to consider a "hardware layer," which functions under the physical layer. The fact that sensor networks have severe constraints on energy makes them especially good candidates to benefit from cross-layer design.

Despite the overall positive assessment of the value of cross-layer approaches, a cautionary message was also delivered at the workshop. An obstacle to the use of cross-layer approaches is the need to accommodate legacy systems, which may not have the capabilities to implement such control functions. Additionally, excessive coordination among the layers may result in the design of special purpose systems, thereby eliminating some of the advantages of layered design. The result could be unwieldy systems that are difficult to understand, and hence difficult to update to accommodate changing requirements. Furthermore, attempts to optimize performance may result in a system that is overly sensitive to parameter values, thereby running the risk of poor performance unless all parameters are perfectly tuned (which is virtually impossible in practice). It should be emphasized that this cautionary message was intended not to discourage the use of cross-layer techniques, but rather to warn of the unintended consequences of protocol interactions that are not yet well understood.

7.1 FUTURE RESEARCH AGENDA

To realize the potential benefits of cross-layer design, considerable further study is needed, both at the basic research level and development of practical systems as well. Some of the future research directions suggested by the panel members are summarized below, with an emphasis on those that are related to tactical networks.

Some fundamental issues:

- Further explore adaptivity, diversity, and scheduling in a cross-layer context.
- Networks with energy-constrained nodes:
    - Determination of appropriate abstractions, especially for hardware;
    - Establishment of the right cross-layer framework; and
    - Incorporation of application layer/compression into networking design and control.
- Collaborative transmission and signal processing for sensor networks.
- QoS Management:
    - Delay Differentiated Services;
    - Dynamic Bandwidth Management; and
    - Prioritized Services vs. Statistical Services in Cross-Layer Design.
- Quality of Protection:
    - Tunable Security (encryption, watermarking) approaches to trade-off/balance performance and security; and
    - Digital Rights Management.

Some specific issues related to tactical ad hoc networks:

- Cross-layer protocols for tactical direct-sequence spread-spectrum networks with multimedia traffic.
- Application of cross-layer protocols to software defined radios.
- Cross-layer adaptation of modulation and coding:
    - Adaptation within a set of radically different waveforms (e.g., includes phase-shift keying (PSK) waveforms, non-binary orthogonal and quasi-orthogonal coded waveforms, and quadrature amplitude modulation (QAM) waveforms); and
    - Adaptation over a wide range of codes (e.g., includes selected turbo codes, standard convolutional codes, and Reed-Solomon (RS) codes).
- Cross-layer protocols for networks in which one or more terminals have directional antennas.

Chapter 8 – REFERENCES

(Task Group workshop presentations referred to below can be accessed through workshop programme in Annex A.)

[1] Multilateral Interoperability Programme Tactical C2IS Interoperability Requirement, MIP Operational Working Group document MTIR-OWG, draft for Edition 2.0, iteration 3, dated 27 September 2003, p. 17.

[2] Goldsmith, A.J. and Wicker, S.B., "Design Challenges for Energy-Constrained Ad Hoc Wireless Networks", IEEE Wireless Communications Magazine, Vol. 9, pp. 2-22, August 2002.

[3] Tanenbaum, A.S., Computer Networks, Fourth Edition, Upper Saddle River, NJ: Prentice Hall PTR 2003.

[4] Pursley, M.B., "Integrated Cross-Layer Protocols for Adaptive Transmission and Routing of Multimedia Traffic in Tactical Spread-Spectrum Networks", keynote presentation at NATO IST-030/RTG-012 Workshop on 'Cross-Layer Issues in the Design of Tactical Mobile Ad Hoc Wireless Networks', Naval Research Laboratory, Washington, DC, USA, June 2-3, 2004.

[5] Nahrstedt, K., Panel discussion presentation at NATO IST-030/RTG-012 Workshop on 'Cross-Layer Issues in the Design of Tactical Mobile Ad Hoc Wireless Networks', Naval Research Laboratory, Washington, DC, USA, June 2-3, 2004.

[6] Wieselthier, J.E., Nguyen, G.D. and Ephremides, A., "Energy-Efficient Broadcast and Multicast Trees in Wireless Networks," Mobile Networks and Applications (MONET), Vol. 7, pp. 481-492, December 2002.

[7] Goldsmith, A.J., "Can Cross-Layer Techniques Enhance the Performance of Tactical Military Networks?," Panel discussion presentation at NATO IST-030/RTG-012 Workshop on 'Cross-Layer Issues in the Design of Tactical Mobile Ad Hoc Wireless Networks', Naval Research Laboratory, Washington, DC, USA, June 2-3, 2004.

[8] Kawadia, V. and Kumar, P.R., "A Cautionary Perspective on Cross-Layer Design", IEEE Wireless Communications Magazine, Vol. 12, pp 3-11, January 2005.