Challenges in Multimedia Networking

Aurel A. Lazar

Department of Electrical Engineering
and
Center for Telecommunications Research
Columbia University, New York, NY 10027-6699
aurel@ctr.columbia.edu

Abstract

We analyze the challenges encountered in realizing *scalable* multimedia networks with quality of service guarantees. Our presentation centers around: (i) seamless connection management between the network and the multimedia devices, (ii) multimedia abstractions with QOS guarantees, and (iii) the integration of service, traffic control and network management architectures.

1. Introduction

The field of multimedia networks has literally exploded in the last couple of years. This can be seen not only from the increasing number of conferences and workshops organized in the past two years but also from the mergers and acquisitions fever in the media, telephone, cable TV and publishing industries. Given the enormous scope of multimedia networking and the many changes it undergoes, it would be a gargantuan task to present an all encompassing picture of the challenges that lie ahead. Hence, we will limit ourselves here to a scenario of multimedia networks with Asynchronous Transfer Mode (ATM) based transport, such as already seen in broadband ATM LANs that are being deployed today. Our presentation will heavily draw upon the Multimedia Systems Services (MSS) proposal currently under consideration by the Interactive Multimedia Association (IMA) [12].

We focus our discussion on a number of multimedia applications that we expect will drive the architecture of multimedia networks. These are interactive multimedia, such as teleconferencing, video on demand and virtual reality based network management. The fundamental requirement on multimedia networks supporting these applications is to guarantee quality of service (QOS).

What are the main challenges that lie ahead for realizing these networks? We believe that the following prerequisites will determine them: (i) seamless connection management between the network and multimedia devices, (ii) multimedia abstractions with QOS guarantees, and (iii) the integration of service, traffic control and network management architectures.

We will discuss the challenges in multimedia networking based on a reference model for broadband networks put forth in [13]. We will extend this model to incorporate not only broadband networks but multimedia networks in general.

2. The Integrated Reference Model

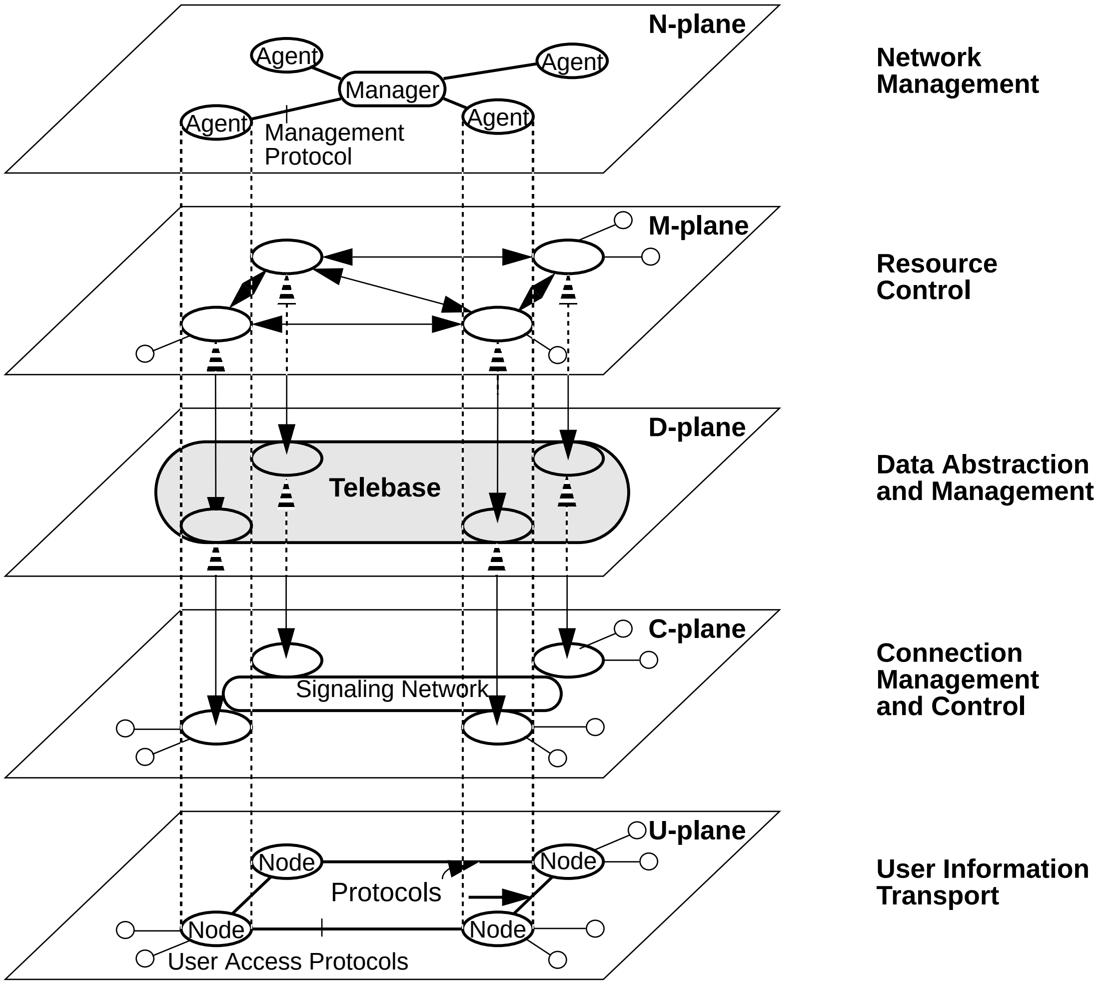

Figure 1. The Integrated Reference Model.

The IRM incorporates monitoring, control, communication, and abstraction primitives that are organized into the *Traffic Control Architecture*, the *Management Architecture*, the *Information Transport Architecture* and the *Telebase Architecture*, respectively. The subdivision of the IRM into the Management and the Traffic Control Architectures on the one hand, and the Information Transport Architecture on the other, is based on the principle of separation between communications and controls. The separation between the Management and the Traffic Control Architecture is primarily due to the different time-scales on which these architectures operate.

The Integrated Reference Model is organized into five planes that model the above architectures (figure 1). The Management Architecture resides in the network management or N- plane, and covers the functional areas of network management, namely, configuration, performance, fault, accounting and security management. Manager and agents, its basic functional components, interact with each other according to the client-server paradigm. The Traffic Control Architecture consists of the resource control, or M-, and the connection management and control, or C-, planes. The M-plane comprises the entities and mechanisms responsible for resource control, such as cell scheduling, call admission, and call routing; the C-plane those for connection management and control, and is based on a signalling network. The Information Transport Architecture is located in the user transport or U-plane, and models the protocols and entities for the transport of user information. The U-plane is horizontally layered, following the ISO Reference Model for Open System Interconnection. Finally, the Telebase Architecture resides within the Data Abstraction and Management or D-plane, and implements the principles of data sharing for network monitoring and control primitives, the functional building blocks of the N-, M-, and C-plane mechanisms. (A mechanism is a functional atomic unit that performs a specific task, such as setting up a virtual circuit in the network.)

The extension of this model to the multimedia networking architecture is simple. The functionality of the planes is extended to include Customer Premises Equipment (CPE), *i.e.*, multimedia devices. This simple extension of the IRM is mainly possible due to the fact that multimedia systems and broadband networks consist of producers, consumers and processors of media. The foundations for the operability of these devices in both, multimedia systems and broadband networks, is the same; the only difference appears to be in the overall goal that a group of devices is set to achieve in the network or the multimedia workstation.

Therefore, the N-plane will include system management functionality, and the M-plane will include scheduling, buffer management, routing (when applicable), admission control and flow control. The D-plane will contain objects modeling multimedia devices, the C-plane binding functionality, and the U-plane transport of user information within the CPE.

3. Connection Management (Binding) for Multimedia Networks

There is a realization in the IT industry that building open, programmable multimedia networks increases its capabilities and leads to innovation. As an example, the field of network management is often cited for making tremendous progress towards interoperability by applying a set of widely accepted standards to an open architecture. This is not yet the case in network control and in connection management.

Connection management traditionally has been dealt with in telecommunication networks, in particular telephone networks. The methodology in telephone networks is to approach the problem of complexity of connection management by defining two sets of interfaces. These are the so called UNI (User/Network Interface) and the NNI (the Network/Network Interface). The commonly talked about UNI is the Q.93B whereas the NNI is realized through the CCSS #7 (Common Channel Signalling System) [4].

The methodology developed for binding multimedia objects in Multimedia Systems Services (MSS) [8] is based on modern foundations of distributed algorithms and software engineering. The MSS proposal is object-oriented. Objects that participate in a multimedia session export three types of interfaces. The latter are specified in the IDL interface definition language. In order to support interaction among distributed objects, MSS depends upon the Object Management Group's (OMG) Common Object Request Broker Architecture (CORBA).

An ad-hoc approach for interconnecting MSS to a broadband network specified via a UNI would be to present it with the Q.93B interface. We feel that this approach has a number of limitations that we will discuss in the next section. We advocate a solution based on a different modeling paradigm. In this paradigm, network connection entities are modeled as communicating objects. As in MSS, CORBA will provide the high level location independent communication facilities. This model allows for a seamless binding environment between the network and the multimedia resources. Overall, binding operations will exhibit a much lower level of complexity.

3.1 Binding for Broadband Networks

Standards bodies [6] are considering the CCSS #7 together with the Q.93B interface as the basis for signalling in broadband networks. There are a number of problems with this solution, three of which are discussed below.

Firstly, the UNI and NNI concepts, introduced in the 60s, rightly recognized that the CPE had a low level of intelligence in comparison with the switching equipment. That has now changed as the customer might possess the latest power MIPS workstation or parallel machine. In fact, the customer equipment is often at least as intelligent as the switching processors.

Secondly, software technologies have made big strides with object-oriented programming as a prime example. The requirements for defining and manipulating virtual networks and multicasting are readily modeled as high level objects. It is natural, therefore, to provide higher level language constructs in describing connection management operations. Note that in this context, the UNI/NNI model is akin to a low level programming language.

Thirdly, the field of distributed systems has made tremendous strides in the last twenty years. New distributed programming paradigms, algorithms and systems have been discovered. These have been widely available for some time now as the client /server model exemplifies. Some of these paradigms have been also standardized, CORBA being a prime example.

We propose here a connection management architecture for the C-plane of the IRM that exploits the main capabilities and power of abstraction that distributed systems and object-oriented programming offer. The connection management architecture advocated here is open: all the network entities that participate in the binding process are modeled as objects with well defined interfaces. Communication among these objects is supported by CORBA. A set of well defined methods support the task of binding. The objects operated upon by the connection management algorithms will reside in the D-plane. The C-plane will support the specific connection management and other distributed algorithms. These algorithms would not be standardized.

Among others, the connection management architecture for broadband networks will:
- provide virtual device abstractions for producers, consumers, and processors of media;
- provide virtual circuit, virtual path and virtual network abstractions between end-to-end virtual devices; and
- provide binding between network devices with end-to-end QOS guarantees.

Connection management objects will reside in the D-plane. Object interfaces will support N-plane mechanisms to specify and query the QOS. The D-plane will provide mechanisms to register, locate and instantiate local or remote resources.

The reader has probably recognized by now the similarities between this proposal with the one already considered by the IMA. Also, there are substantial similarities with standardization efforts within OSI management. As in OSI management, we propose to have an object-oriented model for the entities of interest, a standard communication support and a well defined set of primitives that support basic (connection management) operations. There are of course a number of differences, the main one being that the OSI management architecture is centralized whereas the connection management architecture discussed here is entirely decentralized.

3.2 Binding for Multimedia

Multimedia System Services [8] constitutes a framework of “middleware” — system software components lying in the region between the generic operating system and specific applications. Its goal is to provide an infrastructure for building multimedia computing platforms that support interactive multimedia applications dealing with synchronized, time-based media in a heterogeneous distributed environment. It is under evaluation by the IMA and is expected to become a recommended practice within the computer industry.

Objects in MSS can play the role of clients or servers. In addition to the requirements on the binding model for broadband networks, the Multimedia System Services:
- supports grouping for stream control and resource acquisition to simplify the support of end-to-end QOS; and
- provides a framework for synchronization and degradation.

How does the MSS framework fit into the IRM? Here we distinguish between facilities for creating and removing objects as well as binding operations. In MSS a number of interfaces have been defined to enable both the creation and destruction of objects that participate in the binding process. Creation and destruction operations are, therefore, D-plane native. Binding operations on the other hand are C-plane visible. Sizing of virtual resources derived from QOS requirements is exported through M- and N-plane interfaces.

4. Modeling Multimedia Networking Abstractions with Quality of Service Guarantees

In the previous section we have shown how the methodology first applied to multimedia systems and network management can be extended to devise an open connection management architecture for broadband networks. In this section the reverse is the case. Abstractions with QOS guarantees are well advanced in the broadband networking arena. We will translate them to the multimedia arena, and thereby, come up with a uniform virtual resource characterization for QOS. This view towards modeling of resources substantially reduces the complexity of the overall multimedia network.

The characterization of the capacity of physical resources in broadband networks is based on the concept of schedulable region. This concept was first advanced for evaluating the capacity of multiplexers in broadband switching with QOS guarantees [9], [14]. The schedulable region is a stability concept. It represents the region in the space of calls for which QOS is guaranteed. As such it has solid foundations in queueing theory.

Why should the extension of the concept of schedulable region work for multimedia devices? As we already mentioned in the introduction, both networking as well as multimedia devices are either producers, consumers or processors of media. Stability is a general concept that applies to these irrespective of the size of data they manipulate. It is, therefore, not surprising that the same quantitative characterization of networking and multimedia devices is possible (and highly desirable).

The number of traffic classes with QOS guarantees that the network supports will not be the same as the number of service classes with QOS guarantees offered to users. We expect that the number of user service classes will be much larger then the number of network traffic classes. Therefore, there is a need to “bundle” service classes into network traffic classes in such a way that the end-to-end QOS experienced by the users corresponds to a set of a priori QOS guarantees. This issue will not be discussed further in this paper.

4.1 Modeling Network Abstractions with QOS Guarantees

A broadband switch accepts traffic from a set of input links, and routes each incoming cell though a nonblocking switch fabric to the appropriate output port, where it is queued in a link control unit, for transmission over the output link. This link control unit is essentially a multiplexer. It consists of a set of buffers, a buffer manager, and a scheduler, and it mediates the contention among cells from different input links and among those of different traffic classes. The capacity of the link, the size of the buffer, and the scheduling algorithm used will determine how many calls of a given class the link will be able to support [9], while guaranteeing quality of service to each class. The set of points in the space of possible calls for which QOS can be guaranteed on the cell level is called the schedulable region.

From the point of admission control, the schedulable region is a sufficient representation of the link, which summarizes the net effect of all these cell-level details. Note that while the schedulable region is depicted as a three dimensional volume, it is in fact an n-dimensional space, where n is the number of traffic classes recognized by the admission control entities.

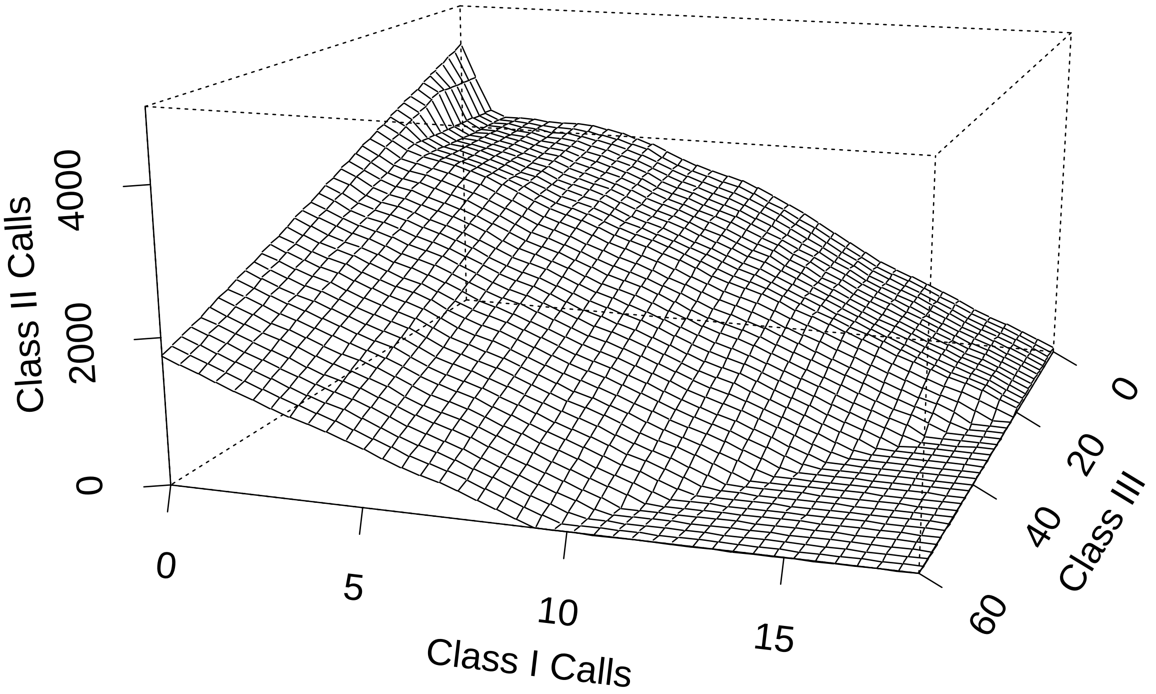

Figure 2. The Schedulable Region of a Multiplexer with Three Traffic Classes.

Based on the schedulable region broadband network abstractions such as virtual circuit, virtual path and virtual network with QOS guarantees can be readily defined. This entails the concepts of admissible load and contract regions. See [10], [11] for further details.

4.2 Modeling Multimedia Abstractions with QOS Guarantees

What are the multimedia abstractions that correspond to the networking abstractions discussed above? In order to understand these we need to digress somewhat and take a look at the limitations of current workstations architectures. A number of bottlenecks can be identified: I/O bottlenecks, processing bottlenecks and overall performance bottlenecks. I/O bottlenecks are due to slow peripherals that were designed for small volumes of data. Processing bottlenecks are due to the master/slave relationship between the CPU and peripheral units. Finally, performance bottlenecks are due to the excessive emphasis on system throughput without QOS considerations.

In order to address these problems a multiprocessor architecture for multimedia workstations seems appropriate [7], [5]. By employing a number of separate processors for various media, audio and video streams are supported by the first processor, the usual data or best effort applications are supported by the second processor and finally storage by the third processor. We see an advantage in this configuration because the three different applications are using different granularity of data. A video conferencing application might require processing of frames, the best effort data swapping of pages and finally data stored on the disc transfer of segments. Execution and thus integration of these various data chunks on the same processor appears difficult to realize under QOS constraints. Note that, the architecture of the multimedia workstation described here runs counter to the developments of broadband networks where the integration requirements of video, voice and data led to an integrated architecture supporting a single data size, the cell.

By isolating the execution of various multimedia applications to specialized processors, a characterization of the Audio Video Unit (AVU), Main Processing Unit (MPU) and Storage Unit (SU) as a device abstraction that guarantees QOS becomes possible. As in [1], [9], real-time scheduling and memory management algorithms running on these processors support such an abstraction. Note that the concept of schedulable region in the domain of multimedia devices can also be applied to protocol stacks running multiple connections on the same processor under QOS requirements. The same applies to the disc as a device abstraction supporting multiple streams of real-time traffic with QOS constraints.

5. The Integration of Binding, Resource Control and Management Architectures

In section 3 we proposed an open architecture for connection management. This architecture is based on modeling entities that take part in the connection management task as communicating objects in the D-plane. The connection management algorithms operating on these objects reside in the C-plane. Communication among objects is supported by CORBA.

The network management architecture residing in the N-plane is designed around the basic manager agent interaction. Information about managed resources is stored in the Managed Information Base (MIB). The MIB is distributed throughout the D-plane. Information about network resources can be accessed remotely; this information is used to execute management applications.

When setting up connections that guarantee QOS, however, a characterization of resources is needed. Based on this characterization, a set of distributed algorithms will decide the optimal policies for allocating resources. These algorithms represent M-plane functionality and are part of the Traffic Control Architecture. Since integration between connection management and resource control is essential to the functioning of a network with QOS guarantees, it is natural to consider an open architecture for the resource control architecture as well. One obvious way to achieve this goal is to model the set of objects that participate in the resource control task. Exchange of information among these objects will be supported again by CORBA. The resource allocation algorithms will not be standardized. However, M-plane visible interfaces will be standardized using IDL and the message passing paradigm will be standardized around CORBA.

What are the advantages of this approach? First, we achieve integration among connection management, resource control and network management via data sharing. Note that our approach to integration also includes the IDL-based MSS multimedia objects at the periphery of the network. The latter can be dynamically bound via CORBA.

Second, the work on integrating IN and TMN indicates that [2], [3], the task of integration exhibits a tremendous amount of complexity. We believe that this complexity is a direct result of the lack of prior design and implementation principles that allow for a natural integration of the management and services architectures. We hope to have shown that the extended IRM facilities, in principle, low complexity solutions to key questions of integration.

The approach outlined here raises a number of issues that need to be resolved. OSI managed objects are defined in GDMO. Therefore, the D-plane of the IRM will contain both IDL as well as GDMO object representations. This can lead to added complexity because IDL-based and GDMO-based object representations are not identical; however, these representations are widely overlapping. We do not foresee any difficulties for the management system to extract data about D-plane services as the question of how to access this data has several solutions.One natural solution is to make IDL-based services available to the N-Plane via a “plane agent”. A plane agent gives access to distributed information that resides in a subset of an IRM plane.

The advantages of integration become more apparent when new capabilities or services are added to the multimedia network. Consider the question, for example, of how to manage a newly installed Video on Demand (VOD) service as part of the general network management system. After the service was installed exceptions will be generated that are made visible to the network management system. Passing all the data arriving to the manager through a filter allows for extracting the relevant VOD service events. Note that this example shows, more generally, that one can specifically monitor a subsystem of the multimedia network through the overall network management system. Hence, management capabilities can be tailored towards specific applications in a natural way.

References

[1] Amenyo, J.T., Lazar, A.A. and Pacifici, G., ‘‘Proactive Cooperative Scheduling and

Buffer Management for Multimedia Networks'', *ACM/Springer Verlag Multimedia Systems,* Vol. 1, No. 1, May1993, pp. 37-49.

[2] Appeldorn, M., Kung, R. and Sarraco, R., "TMN + IN = TINA", *IEEE Communications Magazine*, March 1993, pp. 78-85.

[3] Barr, W.J., Boyd, T. and Inoue, Y., "The TINA Initiative", *IEEE Communications Magazine*, March 1993, pp. 70-76.

[4] Crutcher, A. L. and Waters, A. G., "Connection Management for an ATM Network", *IEEE Network*, November 1992, pp. 42-55.

[5] Gopalakrishnan, R. and Bovopoulos, A.D., "A Protocol Processing Architecture for Networked Multimedia Computers", *ACM Operating Systems Review*, Vol. 27, No. 3, July 1993, pp. 19-33.

[6] Haendel, R. and Huber, M.N., *Integrated Broadband Networks: Architectures, Protocols and Standards*, Addison-Wesley, Workinham, United Kingdom, 1991.

[7] Hayter, M. and McAuley, D., "The Desk Area Network", *ACM Operating Systems Review*, Vol. 25, No. 4, October 1991, pp. 14-21.

[8] HP, IBM and SunSoft, "Multimedia Systems Services Architecture", Response to the *Multimedia System Services Request for Technology*, June 1, 1993.

[9] Hyman, J., Lazar, A.A. and Pacifici, G., "Real-Time Scheduling with Quality of Service Constraints", *IEEE Journal on Selected Areas in Communications*, Vol. SAC-9, No. 7, September 1991, pp. 1052-1063.

[10] Hyman, J., Lazar, A.A. and Pacifici, G., "A Separation Principle between Scheduling and Admission Control for Broadband Switching", *IEEE Journal on Selected Areas in Communications*, Vol. 11, No. 4, May 1993, pp. 605-616.

[11] Hyman, J.M., Lazar, A.A. and Pacifici, G., "VC, VP and VN Resource Assignment Strategies for Broadband Networks", Proceedings of the *Workshop on Network and Operating Systems Support for Digital Audio and Video*, Lancaster, United Kingdom, November 3-5, 1993, pp. 99-110.

[12] Interactive Multimedia Association*, "*Request for Technology: Multimedia System Services Version 2.0*", Interactive Multimedia Association Compatibility Project*, 1993.

[13] Lazar, A.A., "A Real-Time Control, Management and Information Transport Architecture for Broadband Networks", Proceedings of the *1992 International Zurich Seminar on Digital Communications*, March 16-19, 1992, pp. 281-296.

[14] Lazar, A.A. and Pacifici. G., ''Control of Resources in Broadband Networks with Quality of Service Guarantees'', *IEEE Communications Magazine*, Vol. 29, No. 10, October 1991, pp. 66-73.