MULTI-ACCESS IN PACKET RADIO NETWORKS

by

ERDAL ARIKAN

B.S., California Institute of Technology

(1981)

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

MASTER OF SCIENCE

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

August, 1982

Signature of Author ________________________________

Department of Electrical Engineering and Computer Science, August 6,1982

Certified by ________________________________

Robert G. Gallager, Thesis Supervisor

Accepted by ________________________________

Arthur C. Smith, Chairman, Departmental Committee on Graduate Students

MULTI-ACCESS IN PACKET RADIO NETWORKS

by

ERDAL ARIKAN

Submitted to the Department of Electrical Engineering and Computer Science on August 6, 1982 in partial fulfillment of the requirements for the Degree of Master of Science.

ABSTRACT

A PRN (packet radio network) is a collection of geographically distributed, possibly mobile users where each user is capable of transmitting and receiving messages over a shared broadcast medium. In a PRN, messages are divided into packets, which may be fixed or variable in length, and each packet is transmitted through the network individually. Packets are assembled at their destinations to reconstruct the original messages.

The data traffic in a PRN is characterized by specifying the average message arrival rates to the network for each o-d (origin-destination) pair. A set of o-d rates is called feasible if there exist network protocols under which the number of packets in the network still not delivered to their destinations remains finite with probability one. The capacity region of a PRN is defined to be the set of all feasible sets of o-d rates.

In this thesis, PRNs are studied from the viewpoint of feasibility, i.e., we take an arbitrary set of message input rates as given and try to determine if it is feasible. Our main conclusion is that, unless P = NP, there exists no practical algorithm for characterizing the capacity region of a PRN, in the sense that the decision problem regarding the feasibility of a given set of o-d rates is NP-complete.

Thesis Supervisor: Robert G. Gallager

Title: Professor of Electrical Engineering and Computer Science

TO

MY PARENTS

ACKNOWLEDGEMENTS

I would like to thank my thesis supervisor Prof. Robert G.Gallager for his insightful guidance and encouragement throughout this work. His many helpful comments, corrections, and suggestions have greatly improved the substance and the presentation of this thesis.

I would also like to thank Prof. Michael Sipser for his help which led to the proof of Theorem 3.1.

Eli Gafni and Isidro Castiñeyra, my office-mates, deserve special thanks for their interest and patience in discussing many problems in the course of this work. Their generous help is greatly appreciated.

I also thank Mrs. Frantiska Frolik for her excellent typing of this thesis.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER I. INTRODUCTION

1.1 Background Information

Recently there have been attempts to extend the domain of data communications to networks of geographically distributed mobile radio users. Radio networks have traditionally been used for voice communications such as in police cars, emergency vehicles, etc.. We consider here radio networks designed mainly for computer data communications. Such networks present a multitude of problems which must be resolved to make the idea a feasible one.

One of the problems results from the bursty nature of computer data traffic. Typically, a long period of silence is followed by a sudden burst of huge amounts of data which must be transmitted through the network to the desired destination.

The fact that a station need not be (and typically is not) within the transmission range of every other station further complicates the matter. Unlike a wire network, transmissions may interfere with one another and this causes failure to detect messages correctly.

It is a challenging task to design network protocols which are distributed in nature and which satisfy the service demands of the users. In this thesis we study two well-known protocols, namely ALOHA and TDMA, under a model which is refined of all non-essential features so as to simplify the analysis.

For a general survey of packet radio networks we refer the reader to [T] and [KGBK 78].

1.2 The PRN Model

We assume that messages to be transmitted from one station to another are first framed into packets which may be of fixed or variable length. Subsequently, each packet is transmitted separately. The packets contain the information necessary for their travel through the network, such as the identities of the origin node, the destination node, etc. .

We shall represent a PRN (packet radio network) by a directed graph G = (N,A). (For definitions of graph theoretic terms used in this thesis, see e.g. [B].) To each station in the network there corresponds a node of G, and conversely. Throughout the thesis we use the words user, node, and station interchangeably.

For any two distinct nodes a, b ∈ N, there is a link (a,b) iff node b is within the transmission range of node a. We do not assume that (a,b) ∈ A implies (b,a) ∈ A.

We shall study only PRN's with finite numbers of users; accordingly, G = (N,A) will always be a finite graph.

The graphical representation of a PRN will be simple, that is, G = (N,A) will not contain self-loops or multiple links from one node to another.

Throughout the thesis L will denote the number of links in G = (N,A), i.e. $L = |A|$.

A PRN is said to be connected if packets can be routed from every user to every other user. Equivalently, a PRN is connected if the corresponding graph G is connected, i.e. if there exists a directed path from every node to every other node in G.

In our model, each user of the PRN has a single transmitter and a single receiver, thus, a user can transmit or receive only one packet at a time.

When a user transmits a packet, it is possible that all stations within its transmission range receive this packet without error; however, the packet may be intended only for a certain subset of them. In fact, in this thesis we shall assume that each transmitted packet has exactly one intended receiver.

If a packet is received without error by its intended receiver, the transmission is said to be successful. We assume that the only source of unsuccessful packet transmissions is interference. A packet which has not been received successfully is said to have suffered a collision.

Let (a,b) be a link; we say that a packet is transmitted over link (a,b) if the packet is transmitted by node a with the intended receiver being node b. We say that link (c,d) conflicts with link (a,b) if simultaneous transmissions on links (c,d) and (a,b) either are precluded (when c=a, d≠b) or cause a collision at node b (when c=b, or (c,b) is a link). We define C_{ab} to be the set of links which conflict with (a,b). Therefore, we have that

$$C_{ab} = \{(c,d) \in A: \ c \neq a, \ (c,b) \in A\} \cup \{(a,d) \in A: d \neq b\} \cup \{(b,d) \in A\}$$

Sometimes we shall consider the set

$$C^*_{ab} = \{(a,b)\} \cup C_{ab} = \{(b,d) \in A\} \cup \{(c,d) \in A : (c,b) \in A\}$$

Remarks:

1) $(c,d) \in C_{ab} \not\Rightarrow (a,b) \in C_{cd}$

2) $C^*_{ab} = C^*_{cb}$ for any two links (a,b), $(c,b) \in A$

This completes the description of the physical structure of the network. In the next section, we shall look at another aspect of the network; namely, the data traffic.

1.3 Data Traffic in a PRN

The origin of a packet is the node at which the packet enters the network; the destination of a packet is the node at which the packet leaves the network. In our model each packet has a unique destination.

With every o-d (origin-destination) pair we associate two random processes parametrized on the time interval $[0,\infty)$. The x-y arrival process $R_{xy}(t)$ is the number of packets which have arrived at origin x with destination y in the interval $[0,t]$, $t \geq 0$. The x-y departure process $D_{xy}(t)$ is the number of packets which have been delivered from origin x to destination y in the time interval $[0,t], t \geq 0$. We assume that prior to time 0, there are no packets in the network.

Let W be the set of all o-d pairs; let $|W| = w$; and let the o-d pairs be labeled by integers $1,\dots,w$.

We collect the arrival and departure processes associated with all o-d pairs into w-dimensional vectors $\vec{R}(t)$ and $\vec{D}(t)$, respectively. Notice that, $\vec{R}(t) \geq \vec{D}(t)$ and $\vec{1} \cdot (\vec{R}(t) - \vec{D}(t))$ is the number of packets in the network at time t.

The process $\vec{D}(t)$ depends both on $\vec{R}(t)$ and the decisions made by all the protocols that determine the flow of packets through the network.

The primary objective in a PRN is to find network protocols such that

$$\lim_{\alpha\to\infty} \lim_{t\to\infty} P\{\vec{1} \cdot (\vec{R}(t) - \vec{D}(t)) < \alpha\} = 1 \cdot \qquad (1.1)$$

In this thesis, we restrict our attention to the following class of o-d arrival processes.

1) $\vec{R}(t)$ is independent of the network protocols.

2) For all $i \in W$, there exists $r_i \geq 0$; and, given $\varepsilon > 0$, $\delta > 0$ we can find $T(\varepsilon,\delta) > 0$ such that

$$P\{\left|\frac{R_i(t)}{t} - r_i\right| > \varepsilon\} < \delta \qquad \forall t > T(\varepsilon,\delta).$$

Clearly, r_i with the above property is unique when it exists. We define $\vec{r}$ to be the column w-vector whose i^{th} row equals r_i; $\vec{r}$ is called the mean o-d (arrival) rate vector.

The mean o-d rate vector $\vec{r}$ will be used to characterize the arrival process $\vec{R}(t)$. We shall say that $\vec{r}$ is feasible if there exists network protocols such that (1.1) is satisfied.

We define the capacity region of $G = (N,A)$ to be the following set:

$$C(G) = \{\vec{r} : \vec{r} \text{ is feasible}\} \quad .$$

It is of extreme interest to determine the region C(G). A second problem, which is also interesting, is to determine whether a particular mean o-d rate vector is feasible under a particular set of network protocols. By network protocols we essentially mean the routing and link-level multi-access schemes. The interaction between routing and access schemes makes it difficult to study the feasibility problems. Therefore, we examine multi-access schemes in isolation from routing schemes by assuming that:

1) $r_{xy} = 0$ if $(x,y) \notin A$

2) Every (x-y) packet, i.e. a packet with origin x and destination y, is transmitted directly over link (x,y).

When we consider the feasibility problem under these assumptions, we shall use the column L-vector $\vec{f}$ instead of $\vec{r}$. The mean link (arrival) rate vector $\vec{f}$ has one element for each link $(a,b) \in A$, which equals r_{ab}. Thus, f_{ab} is the mean link rate of (a,b). Under the above assumptions, the analogues of $R_{xy}(t)$, $\vec{R}(t)$, $D_{xy}(t)$ and $\vec{D}(t)$ will be denoted by $F_{xy}(t)$, $\vec{F}(t)$, $S_{xy}(t)$ and $\vec{S}(t)$, respectively. For example, $F_{xy}(t) - S_{xy}(t)$ is the number of packets waiting at node x to be delivered to node y over the link (x,y) at t (including a packet in transmission, if any).

With this notation, we can now define the two main problems we shall study in this thesis.

Definition.

The $\vec{r}$ - feasibility problem

Given a PRN G = (N,A) and a w-vector $\vec{r} \geq \vec{0}$, is $\vec{r}$ feasible, i.e., does $\vec{r}$ belong to C(G)?

Definition.

The $\vec{f}$ - feasibility problem

Given a PRN G = (N,A), a multi-access scheme, and an L-vector $\vec{f} \geq \vec{0}$, is $\vec{f}$ feasible under the given multi-access scheme?

We shall study the $\vec{f}$-feasibility problem under two extreme ways of multi-accessing, namely ALOHA and TDMA. The $\vec{r}$-feasibility problem will be studied in connection with TDMA. Before we give an outline of our results on the feasibility problems, we shall review some of the relevant work in this area.

1.4 Review of Earlier Work

In a B.S. thesis Shapiro [S79] analyzes an Aloha scheme for PRNs and obtains conditions for minimizing average system delay in terms of routing variables.

A different approach to the analysis of Aloha type systems is taken by Sidi and Segall [SS(1)81] who consider the PRN as a network of interfering queues and obtain approximate results for the steady-state packet distribution in the buffers of the PRN.

PRNs have also been studied under different assumptions and multi-access schemes, e.g. [SS(2)81].

1.5 Thesis Outline

In Chapter 2, the $\vec{f}$-feasibility problem is studied under slotted Aloha. We examine conditions for stability from a feasibility viewpoint; routing variables and link delays do not enter into the analysis.

In Chapter 3, we formulate and state a number of facts about TDMA schemes, among these facts are the NP-completeness of the f-feasibility problem under TDMA, and the NP-completeness of the $\vec{r}$-feasibility problem.

The fourth chapter consists of NP-completeness proofs of the feasibility problems and an alternative formulation of TDMA schemes.

In Chapter 5, we show why not even a polynomial-time approximation algorithm is likely to be found for the NP-complete feasibility problems.

The main result of this thesis is the conclusion that, unless P = NP, there exists no practical algorithm for characterizing the capacity region of a PRN.

CHAPTER II. SLOTTED ALOHA IN PRNs

2.1 Introduction

In this chapter we study the $\vec{f}$-feasibility problem under slotted Aloha. We assume familiarity with Aloha as it is used in single-receiver multi-access communication systems. The scheme considered here is a simple extension of the well-known Aloha systems; we do not claim that it is original or practical. Our purpose is to study this scheme along the lines outlined in Chapter I.

The main difficulty in the study of Aloha in PRNs is encountered at the modelling stage. Our model must account for the essential features of Aloha and yet it must be tractable. We can make the model more and more sophisticated only to render it useless. On the other hand, an oversimplified model may lead to erroneous conclusions. The modelling issue is further complicated by the variability of network topology and data traffic.

Our analysis of Aloha is based on a model which is as simple as it can be. We introduce the equilibrium hypothesis and discuss the meaning of equilibrium. The analysis provides results which on the whole conform with intuition provided by the single-receiver Aloha systems. However, some of the results are non-obvious and seem to be useful.

2.2 The Aloha Model

We consider a slotted Aloha scheme where all stations are synchronized so that packet transmissions occur only in globally defined time slots. Packets are fixed in length and the duration of a slot is long enough to accommodate the time it takes to transmit a packet plus any delays associated with propagation and detection of packets. We assume propagation delay is negligible relative to the length of a slot, but we do not ignore it altogether.

In Aloha, transmitted packets may suffer collisions, and when they do, they must be retransmitted. Thus, there are three random processes associated with each link that we have to distinguish:

1) The process of new packet arrivals,

2) The process of successful packet transmissions,

3) The scheduling process of packet transmissions.

Following the notation introduced in Chapter 1 in connection with the $\vec{f}$-feasibility problem, we define $F_{ab}(n)$ $n \in \{1,2,...\}$ to be the number of (a-b) packets that have arrived at origin a to be transmitted over $(a,b) \in A$ to destination b before the beginning of the n^{th} slot. Each arrival process $F_{ab}(\cdot) : (a,b) \in A$ is independent and Poisson with rate f_{ab}, i.e.,

$$P\{F_{ab}(n+1) - F_{ab}(n) = k\} = \frac{f_{ab}^k e^{-f_{ab}}}{k!} \qquad k \geq 0, \; n \geq 1.$$

We define $S_{ab}(n)$ $n \in \{1,2,...\}$ to be the number of (a - b) packets that have been successfully transmitted from origin a to destination b over (a,b) before the beginning of the $(n+1)^{st}$ slot.

Note that a packet leaves the network following its first successful transmission, so, $F_{ab}(n) \geq S_{ab}(n)$ for all $n \geq 1$, and all $(a,b) \in A$.

A packet is scheduled for transmission in the first slot following its arrival. Each packet which is waiting for retransmission over link (a,b) is scheduled for the current slot with probability $0 < u_{ab} < 1$ independently of all other packets. As a consequence, we admit the possibility of more than one packet being scheduled for transmission by the same node in the same slot. When this happens, all packets involved are treated as if they have suffered collisions and the corresponding node transmits a blank signal in that slot. Clearly, this system can be improved at no cost, but we wish to keep it this way to facilitate the analysis.

We shall assume that acknowledgements are available to the transmitting stations immediately following the transmissions. This assumption is made for definiteness and can be relaxed without altering our results.

We define the scheduling process of packets over link (a,b) as follows:

$$G_{ab}(1) = F_{ab}(1)$$

$$G_{ab}(n+1) = G_{ab}(n) + B_{ab}(n+1) \qquad n \in \{1,2,\ldots\}$$

where $B_{ab}(n)$ is the number of packets scheduled for transmission over (a,b) in the n^{th} slot.

2.3 The Equilibrium Hypothesis

Suppose the mean link arrival rates are sufficiently small so that, with high probability, packets are transmitted successfully at the first transmission attempt. Occasionally, there will be collisions, but if the retransmission probabilities are small enough, and if they are adjusted with respect to the prevailing traffic density in the network, we expect resolution of contentions for channel use and an eventual return to normal conditions.

The above argument is the basis of the equilibrium hypothesis, which approximates the probability assignment for $B_{ab}(n)$ by

$$P\{B_{ab}(n) = k\} \approx \frac{g_{ab}^{k} \; e^{-g_{ab}}}{k!} \tag{2.1}$$

where the constant g_{ab} is called the mean packet scheduling rate on link (a,b).

The equilibrium hypothesis does not hold for the Aloha model we have described in the previous section, because if $f_{ab} > 0$, then

$$B_{ab}(n) \to \infty \quad \text{as } n \to \infty \quad \text{with probability 1.}$$

Despite this inherent instability, the possibility of "stable" operation over long periods of time motivates the analysis we offer in the next section.

2.4 Analysis of Slotted Aloha Under the Equilibrium Hypothesis

Let s_{ab} be the probability that in an arbitrarily chosen slot there is a successful packet transmission over (a,b). Under the equilibrium hypothesis, we have that

$$s_{ab} = g_{ab} \, e^{-g_{ab}} \prod_{(c,d)\in C_{ab}} e^{-g_{cd}} \tag{2.2}$$

or

$$s_{ab} = g_{ab} \exp\left\{- \sum_{(c,d)\in C_{ab}^{*}} g_{cd}\right\} \tag{2.3}$$

The summation in the exponent depends only on node b (see the second remark following the definition of C^*_{ab} in Chapter 1). Therefore, we define

$$\lambda_b = \sum_{(c,d)\in C^*_{ab}} g_{cd}$$

and express (2.3) as

$$s_{ab} = g_{ab}\, e^{-\lambda_b} \qquad \text{for all } (a,b)\in A \ . \tag{2.4}$$

In equilibrium we must have $s_{ab} = f_{ab}$ for all $(a,b)\in A$. To see why this must be so, suppose $s_{ab} < f_{ab}$ for some link (a,b). Then, the number of packets waiting for retransmission over (a,b) increases continually, thus, increasing the value of g_{ab}, which in turn causes S_{ab} to decrease. As a result equilibrium cannot exist.

On the other hand, $s_{ab} > f_{ab}$ implies that the expected number of packets that are successfully transmitted over (a,b) in a time period exceeds the expected number of new packets entering the network for transmission over (a,b) in the same time period. This is clearly contrary to equilibrium.

Therefore, under equilibrium conditions we must have

$$f_{ab} = g_{ab}\, e^{-\lambda_b} \qquad \text{for all } (a,b)\in A \tag{2.5}$$

If we take $\{g_{ab}\}$ as given, then $\{f_{ab}\}$ is uniquely determined by (2.5). Therefore, there are sets $\{f_{ab}\}$ and $\{g_{ab}\}$ which satisfy (2.5), and the equilibrium hypothesis does not immediately lead to an inconsistent set of equations.

We would like to have necessary and sufficient conditions on

$\{f_{ab}\}$ which guarantee the existence of solutions to (2.5), and which are easily verifiable. We do not know of any such necessary and sufficient conditions short of solving (2.5).

However, a simple necessary condition on $\{f_{ab}\}$ for the existence of solutions to (2.5) is available through the following observation.

Suppose $\{g_{ab}\}$ is a solution to (2.5). Let us rewrite (2.5) as

$$g_{ab} = (f_{ab}\, e^{\lambda_b - g_{ab}})\, e^{g_{ab}}, \quad \text{for all } (a,b) \in A. \tag{2.6}$$

The term $\lambda_b - g_{ab}$ in the exponent does not explicitly depend on g_{ab}, thus, (2.6) is an equation of the form $x = \alpha e^x$, for which we have three cases.

i) $\alpha > e^{-1}$: there is no solution.

ii) $\alpha = e^{-1}$: $x = 1$ is the only solution.

iii) $\alpha < e^{-1}$: there are two distinct solutions; one greater than 1, and the other less than 1.

Therefore, we conclude that (2.6) has a solution only if

$$f_{ab}\, e^{\lambda_b - g_{ab}} \leq e^{-1} \tag{2.7}$$

or

$$-\ln f_{ab} \geq 1 - g_{ab} + \lambda_b \tag{2.8}$$

Since $g_{ab} \geq f_{ab}$ for all $(a,b) \in A$, we have obtained the following necessary condition:

(2.5) has a solution only if

$$-\ln f_{ab} \geq 1 + \sum_{(c,d) \in C_{ab}} f_{cd} \qquad \forall (a,b) \in A \text{ s.t. } f_{ab} > 0\,. \tag{2.9}$$

Notice that (2.9) can be satisfied only if $f_{ab} \leq e^{-1}$ for all $(a,b) \in A$.

In the rest of this chapter we shall investigate a computational algorithm which converges to a particular solution to (2.5), if one exists. For this purpose, we find it convenient to express our equations in vector form. We define $\vec{g}$ and $\vec{f}$ to be column L-vectors corresponding to $\{g_{ab}\}$ and $\{f_{ab}\}$, respectively, and consider the following iterative algorithm.

$$\begin{aligned} \vec{g}(n+1) &= \vec{h}(\vec{g}(n)) \qquad n = 1,2,\ldots \\ \vec{g}(1) &= \vec{f} \end{aligned} \tag{2.10}$$

where
$$h_{ab}(\vec{g}(n)) = f_{ab}\, e^{\lambda_b(n)}$$

$$\lambda_b(n) = \sum_{(c,d) \in C^*_{ab}} g_{cd}(n)$$

A vector $\vec{x}$ is called an equilibrium point of (2.10) if $\vec{x} = \vec{h}(\vec{x})$. The equilibrium points of (2.10) correspond to the solutions of (2.5).

We shall show that (2.10) converges to an equilibrium point if and only if (2.5) has a solution.

Let $\vec{x}$ and $\vec{y}$ be non-negative column L-vectors such that $\vec{x} \geq \vec{y}$. Then, since each row of $\vec{h}(\cdot)$ is a non-decreasing function in each variable, $\vec{h}(\vec{x}) \geq \vec{h}(\vec{y})$.

Thus, in (2.10), $\vec{g}(2) \geq \vec{g}(1) = \vec{f}$ because $\vec{g}(2) = \vec{h}(\vec{f})$ and $\vec{g}(1) = \vec{h}(\vec{0})$, and $\vec{f} \geq \vec{0}$. This argument also shows that (2.10) is non-decreasing.

Let $\vec{g}$ be any equilibrium point of (2.10), if one exists.

We have already argued that $\vec{g} \geq \vec{f}$. So, in (2.10), $\vec{g}(1) \leq \vec{g}$. Now suppose, for some $n \geq 1$, $\vec{g}(n) \leq \vec{g}$, then $\vec{g}(n+1) = \vec{h}(\vec{g}(n)) \leq \vec{h}(\vec{g}) = \vec{g}$. By induction, (2.10) is bounded above.

Therefore, (2.10) converges (because it is non-decreasing and bounded) if and only if (2.10) has an equilibrium point (equivalently, if and only if (2.5) has a solution).

Suppose (2.10) converges to $\vec{g}^*$. Note that $\vec{g}^*$ is an equilibrium point of (2.10). The above argument shows that if $\vec{g}$ is any equilibrium point of (2.10), then $\vec{g} \geq \vec{g}^*$. In other words, if (2.5) has solutions, there exists a smallest equilibrium point of (2.10) and (2.10) converges to this point.

In fact, if (2.10) converges to $\vec{g}^*$, then, starting from any initial point $\vec{g}(1)$ in the region $\vec{0} \leq \vec{g}(1) \leq \vec{g}^*$, $\vec{g}(n+1) = \vec{h}(\vec{g}(n))$ converges to $\vec{g}^*$. To see this, suppose $\vec{0} \leq \vec{x}(1) \leq \vec{g}^*$ is arbitrary and let $\vec{y}(1) = \vec{0}$. Since $\vec{y}(2) = \vec{h}(\vec{0}) = \vec{f}$, $\vec{y}(n) \to \vec{g}^*$ as $n \to \infty$. But $\vec{y}(n) \leq \vec{x}(n) \leq \vec{g}^*$ for all n; therefore $\vec{x}(n)$ converges to $\vec{g}^*$ as well.

Suppose again that (2.10) converges to $\vec{g}^*$. Let $\vec{x}(1)$ be such that $(1 - \epsilon)\vec{g}^* \leq \vec{x}(1) \leq \vec{g}^*$ where $0 < \epsilon < 1$. If ϵ is sufficiently small, we can approximate $\vec{x}(n)$, defined recursively by $\vec{x}(n) = \vec{h}(\vec{x}(n-1))$ $(n \geq 2)$, as follows:

$$\vec{x}(n+1) \approx \vec{g}^* + H(\vec{x}(n) - \vec{g}^*), \qquad n \geq 1, \tag{2.11}$$

where H is the Jacobian matrix of $h(\cdot)$ evaluated at $\vec{g}^*$.

By defining $\vec{\delta}(n) = \vec{g}^* - \vec{x}(n)$, (2.11) can be expressed as

$$\vec{\delta}(n+1) = H\,\vec{\delta}(n) \quad , \qquad n \geq 1 \tag{2.12}$$

Since $\vec{\delta}(n) \to \vec{0}$ as $n \to \infty$, the eigenvalues of H must lie within the unit circle in the complex plane.

The elements of H are as follows:

$$\left.\frac{\partial h_{ab}(\vec{x})}{\partial x_{cd}}\right|_{\vec{x}=\vec{g}^*} = \begin{cases} g^*_{ab} & (c,d)\in C^*_{ab}, \quad f_{cd} > 0 \\ 0 & \text{otherwise} \end{cases}$$

We observe that H is non-negative; therefore, by the Perron-Frobenius Theorem, H has a real and non-negative eigenvalue which is not smaller than any other eigenvalue of H in magnitude. The eigenvalue with this property, denoted by γ_H, is called the dominant eigenvalue of H. Thus, (2.5) has a solution if and only if $\gamma_H \leq 1$.

This argument is instructive and useful even though γ_H cannot be computed easily from the knowledge of $\vec{f}$ alone. For example, we can show that $\vec{g}^* \leq \vec{1}$ is satisfied by making use of some properties of non-negative matrices.

Consider the diagonal matrix $\text{diag}(\vec{g}^*)$ whose $(a,b)^{th}$ diagonal entry is g^*_{ab}; $\text{diag}(\vec{g}^*)$ is non-negative and $H \geq \text{diag}(\vec{g}^*)$. The dominant eigenvalue of $\text{diag}(\vec{g}^*)$ is

$\gamma_G = \max \{g^*_{ab} : (a,b)\in A\}$. It follows that $\gamma_G \leq \gamma_H$ (see, e.g., [KT] pages 542-551).

Therefore, (2.10) converges to $\vec{g}^*$ only if $\gamma_G \leq 1$, or in other words, if (2.10) converges to $\vec{g}^*$, then $\vec{g}^* \leq \vec{1}$.

We conclude the analysis of slotted Aloha by obtaining a lower bound on γ_H.

Define the matrix F as follows:

$$(F)_{ab,cd} = \begin{cases} f_{ab} & (c,d) \in C^*_{ab}\ ,\ \ f_{cd} > 0 \\ 0 & \text{otherwise} \end{cases}$$

F is obtained from G by replacing g^*_{ab} with f_{ab} for each $(a,b) \in A$. Thus, $0 \leq F \leq H$, and, if γ_F is the dominant eigenvalue of F, then $\gamma_F \leq \gamma_H$.

It is clear that (2.5) has a solution only if $\gamma_F \leq 1$.

To obtain a lower bound on γ_F (and also on γ_H), let us define α_{ab} and β_{ab} to be the sum of the elements in the row and the column of F, both corresponding to link (a,b), respectively. That is,

$$\alpha_{ab} = \sum_{\substack{(c,d) \in C^*_{ab} \\ f_{cd} > 0}} f_{ab} = f_{ab}\, N_{ab}$$

where N_{ab} is the number of links in C^*_{ab} with positive arrival rates, and

$$\beta_{ab} = \begin{cases} \displaystyle\sum_{(c,d)\ :\ (a,b) \in C^*_{cd}} f_{cd} & f_{ab} > 0 \\ 0 & f_{ab} = 0 \end{cases}$$

Let

$$\gamma_o = \max\{\min\{\alpha_{ab} : (a,b)\in A,\ \alpha_{ab} > 0\}\ ,$$

$$\min\{\beta_{ab} : (a,b)\in A,\ \beta_{ab} > 0\}\}\quad .$$

By the general properties of non-negative matrices (see, e.g., page 194 of [L]), we have that $\gamma_0 \leq \gamma_F$. The computation of γ_o is easy and (2.5) has a solution only if $\gamma_0 \leq 1$.

The results of this section can be summarized as follows.

Proposition 2.1.

(2.5) has a solution if and only if (2.10) converges. (2.10) converges only if

$$-\ln f_{ab} \geq 1 + \sum_{(c,d)\in C_{ab}} f_{cd} \qquad \forall\ (a,b)\in A,\ \text{s.t. } f_{ab} > 0,$$

and $\gamma_F \leq 1$.

2.5 Discussion of Results

As we have remarked before, the Aloha system we have considered in this chapter is inherently unstable; the main reason for the analysis of Aloha was to understand the conditions under which the system could be expected to operate satisfactorily before the throughput collapsed.

The equilibrium hypothesis implies that stable operation is not possible, even temporarily, unless (2.5) has a solution. If (2.5) has a solution, stabilization of Aloha may be possible by the introduction of a control mechanism; in this sense we can say that $\vec{f}$ is feasible under slotted Aloha if (2.5) has a solution.

However, our results are subject to the limitations of the equilibrium hypothesis, which totally ignores the interdependencies among the packet scheduling rates of different links. It would be desirable to study the $\vec{f}$-feasibility problem on a model which at least partially accounts for these dependencies, but, mainly because of tractability problems, we do not pursue this matter any further.

CHAPTER III. TDMA in PRNs

In this chapter we consider TDMA (time-division-multi-access) schemes for PRNs After formulating TDMA in terms of the transmission vectors of a PRN, we address the feasibility problems, which were introduced in chapter 1.

The main result of this chapter is the NP-completeness of the feasibility problems. The reader, unfamiliar with NP concepts, is referred to [PS] or [GJ] for the essentials of algorithmic complexity.

3.1 Notation and Definitions

We change our notation slightly in this and the following chapters to avoid cumbersome subscripts.

As usual, $G = (N,A)$ will represent a PRN. Nodes will be denoted by the small case letters a,b,c,d and e. $L = |A|$ will be the number of links, and the links will be labelled by integers 1 through L. The letters i,j,k and ℓ are reserved for denoting the links.

With the new notation, f_i will stand for the mean packet arrival rate on link i. The collection of all arrival rates will be represented as a set $\{f_i\}$ and alternatively, as a column L-vector $\vec{f}$, the i^{th} component of which equals f_i.

TDMA analysis is based on the concept of collision-free transmission vectors. To every time instant we associate a column L-vector $\vec{t}$ with the following property: $t_i = 1$ if there is a transmission over link i at that time instant; $t_i = 0$ otherwise.

A vector $\vec{t}$ with the above property is called a transmission vector.

A collision-free transmission vector $\vec{t}$ is a transmission vector with the following additional property: $t_i + t_j \leq 1$ if $i \in C_j$ for all $i,j \in A$.

Since we are interested only in collision-free transmission vectors in our study of TDMA, from here on we shall refer to collision-free transmission vectors simply as transmission vectors.

We say that link $i \in A$ is enabled (or used) by a transmission vector $\vec{t}$ if $t_i = 1$. The transmission set of a transmission vector $\vec{t}$ is defined to be the set of links that are enabled by $\vec{t}$. Thus the transmission set of $\vec{t}$ is the set $\{i \in A: t_i = 1\}$.

By convention, the all zeroes vector $\vec{0}$ is a transmission vector, and the null set is a transmission set.

A maximal transmission set is a transmission set which is not properly contained in another transmission set. A maximal transmission vector is a transmission vector whose transmission set is maximal.

The total number of all transmission vectors, denoted by K, is obviously bounded by 2^L. We let $\vec{t}_1, \vec{t}_2, \ldots, \vec{t}_K$ be an ordering of all transmission vectors.

The transmission matrix T is defined to be an LxK matrix whose i^{th} column is $\vec{t}_i$ for $i = 1,\ldots,K$.

In the L-dimensional real space R^L, the transmission vectors can be thought of as points. The convex hull of all those points, denoted by CH(T), is a closed bounded polytope which lies in the non-negative orthant of R^L.

3.2 TDMA in PRNs

We are now in a position to introduce the TDMA scheme. For simplicity of exposition, in this chapter we consider variable slot length TDMA schemes. In the variable slot length case, each user has a sequence of positive numbers. The sequences are identical. The terms of the sequence represent the duration of slots. All users are synchronized so that the slots as perceived by different users always start and end simultaneously.

In TDMA, we associate a transmission vector with every slot. The sequence of transmission vectors associated with the slots is assumed to be the common knowledge of all users. In our study, we do not distinguish the order in which the transmission vectors are used. If in some period of time the same transmission vector is used more than once, we can combine all the slots in which it is used into a single larger slot whose duration is the sum of the durations of the individual slots.

Since we are studying the feasibility problems in a static setting, we restrict our attention to periodic sequences of transmission vectors. We let x_i denote the duration of time the transmission vector $\vec{t}_i$ is used in one period. The column K-vector $\vec{x}$ will denote the vector whose i^{th} row is x_i. We put a normalization constraint on $\vec{x}$ by demanding that $\vec{1} \cdot \vec{x} \leq 1$.

After this introduction, we give a formal definition of TDMA.

Definition. Let $G = (N,A)$ be a PRN and let T be the transmission matrix of G. A TDMA scheme (for G) is a system $\langle T,\vec{x}\rangle$, where $\vec{x}$ is a non-negative column K-vector such that $\vec{1} \cdot \vec{x} \leq 1$.

In a PRN, all links have the same capacity. We normalize link rates with respect to this capacity and require that $\vec{0} \leq \vec{f} \leq \vec{1}$.

Definition. Let $G = (N,A)$ be a PRN and let $\vec{f}$ be a link rate vector for G. $\vec{f}$ is called feasible under TDMA if there exists a TDMA scheme $\langle T,\vec{x}\rangle$ such that $T\vec{x} \geq \vec{f}$,

We will now state a series of straightforward facts about TDMA schemes. Facts 3.1 - 3.5 are stated in reference to a fixed but arbitrary PRN $G=(N,A)$.

FACT 3.1.

$\vec{f}$ is feasible under TDMA iff $u \leq 1$ where u is defined as follows:

LP1:

$$u = \min \vec{1} \cdot \vec{x}$$
$$\text{s.t.} \quad T\vec{x} \geq \vec{f}$$
$$\vec{x} \geq \vec{0} \quad .$$

FACT 3.2.

If $\vec{f}$ is feasible under TDMA, then there exists a TDMA scheme $\langle T,\vec{x}\rangle$ such that $T\vec{x} \geq \vec{f}$ and

$$|\{i : x_i > 0\}| \leq L.$$

Proof. If $\vec{f}$ is feasible under TDMA, then there exists a basic feasible solution which optimizes LP1 (the linear program defined in the statement of Fact 3.1).

FACT 3.3.

Let τ be the set of all transmission sets of $G = (N,A)$. The system (A,τ) is an independence system.

Proof. A system (E,S) is an independence system if S is a set of subsets of E with the following property: $E_1 \in S$ and $E_2 \subset E_1$ implies $E_2 \in S$.

Since a subset of a transmission set is a transmission set, (A,τ) is an independence system.

FACT 3.4.

Let u be as given in Fact 3.1, let v be as follows:

LP2:

$$v = \min \vec{1} \cdot \vec{x}$$
$$\text{s.t. } T\vec{x} = \vec{f}$$
$$\vec{x} \geq \vec{0}$$

Then, $u = v$.

Proof. Fact 3.4 follows from the independence system property of (A,τ).

FACT 3.5.

$\vec{f}$ is feasible under TDMA iff $\vec{f} \in CH(T)$.

Proof. $\vec{f} \in CH(T)$ iff $\vec{f}$ can be expressed as a convex combination of the extreme points of CH(T), which are, by definition, the transmission vectors of G.

FACT 3.6.

If $\vec{f}$ is feasible under TDMA, then there exists a TDMA scheme $\langle T,\vec{x}\rangle$ such that

(i) $T\vec{x} \geq \vec{f}$,

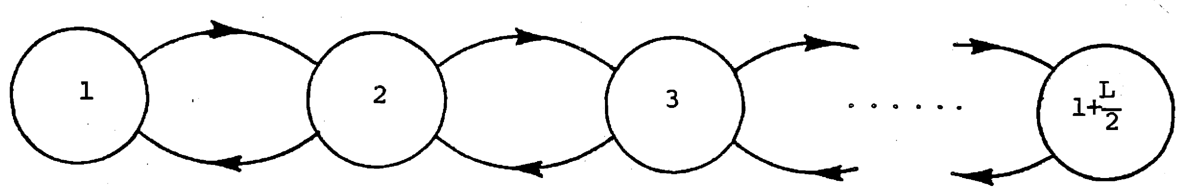
(ii) $j \in \{i : x_i > 0\} \Longrightarrow \vec{t}_j$ is maximal,

(iii) $|\{i : x_i > 0\}| \leq L$.

FACT 3.7.

Given a polynomial p(•), we can always find a PRN with the property that K > p(L).

Proof. An example suffices. Consider the PRN in Figure 3.1 with $(1 + \frac{L}{2})$ nodes and L links (L is even). There is a transmission set with L/4 elements. Therefore, the total number of transmission vectors is at least $2^{(L/4)}$.



A Chain PRN

Figure 3.1

As a result, linear programming as applied to LP1, or LP2 does not guarantee an efficient solution to the $\vec{f}$-feasibility problem.

Can we reduce the size of LP1 (or LP2) by discarding non-maximal transmission vectors, so that linear programming can be used to solve the $\vec{f}$-feasibility problem efficiently? The following fact suggests that we cannot.

FACT 3.8.

Given a polynomial $p(\cdot)$, we can always find a PRN with the property that $K' > p(L)$, where K' is the number of all maximal transmission vectors.

Proof. Consider a chain PRN G_m, such as the one in Figure 3.1 with $3m+2$ nodes and $L = 6m+2$ links for some integer $m \geq 1$. By induction we will show that G_m has at least $2 \cdot 3^m$ maximal transmission vectors.

G_{m+1} is constructed by adjoining three nodes to G_m, as in Figure 3.2.

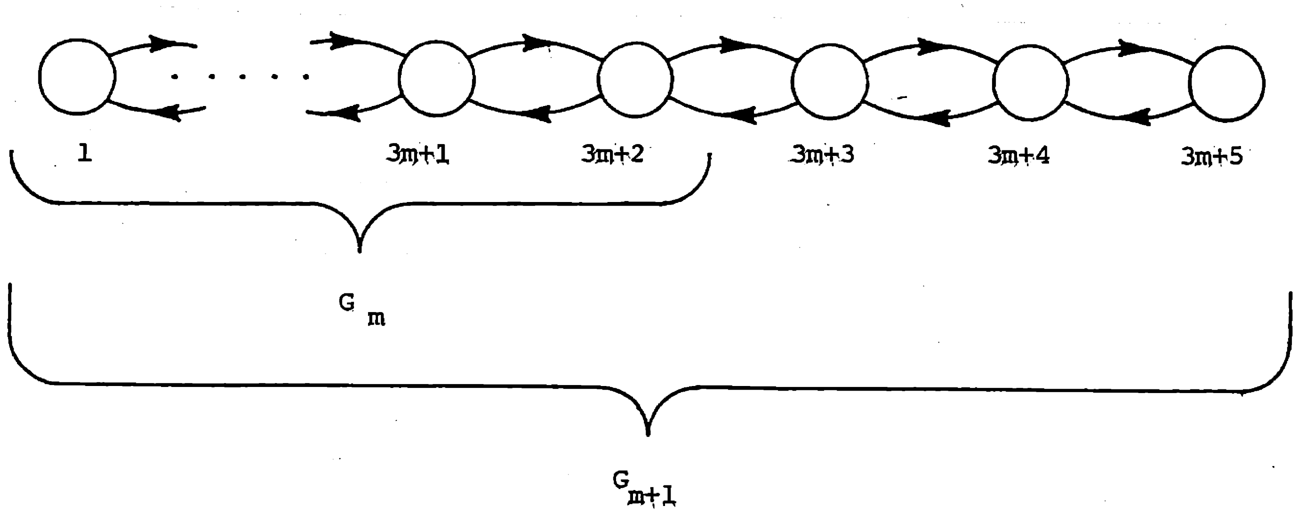


Figure 3.2

Let S be a maximal transmission set of G_m. There are four cases:

1) (3m+1, 3m+2) $\in$ S

In this case $S \cup \{(3m+4, 3m+5)\}$, $S \cup \{(3m+4, 3m+3)\}$, and $S \cup \{(3m+5, 3m+4)\}$ are maximal transmission sets of G_{m+1}.

2) (3m+2, 3m+1) $\in$ S

In this case $S \cup \{(3m+3, 3m+4)\}$, $S \cup \{(3m+4, 3m+5)\}$, and $S \cup \{(3m+5, 3m+4)\}$ are maximal transmission sets of G_{m+1}.

3) (3m, 3m+1) $\in$ S

In this case $S \cup \{(3m+3, 3m+2), (3m+4, 3m+5)\}$, $S \cup \{(3m+3, 3m+4)\}$, and $S \cup \{(3m+5, 3m+4)\}$ are maximal transmission sets of G_{m+1}.

4) (3m+1, 3m) $\in$ S

In this case $S \cup \{(3m+2, 3m+3), (3m+5, 3m+4)\}$, $S \cup \{(3m+3, 3m+4)\}$, $S \cup \{(3m+4, 3m+3)\}$, and $S \cup \{(3m+4, 3m+5)\}$ are maximal transmission sets of G_{m+1}.

Since every maximal transmission set of G_m must fall into one and only one of the four cases listed above, we have shown that G_{m+1} has at least three times as many maximal transmission sets as G_m.

For m = 1, by enumeration, we see that there are at least $2 \cdot 3^1$ maximal transmission sets. Therefore, we conclude that G_m has at least $2 \cdot 3^m$ maximal transmission sets.

In order to discuss the complexity of the $\vec{f}$-feasibility problem under TDMA, we need to define a size for the problem. This forces us to make the following assumption. The link arrival rates are rational numbers such that there exists a fixed (but arbitrary) positive integer P with the following property: $P \cdot f_i$ is integer for all $i \in A$. It is important to note that P stays fixed over all instances of the problem, but otherwise, P is arbitrary.

Definition.

FF (TDMA - $\vec{f}$ - Feasibility Problem)

Instance: A PRN G = (N,A), a link arrival rate vector $\vec{f}$.

Question: Is $\vec{f}$ feasible under TDMA?

We define the size of an instance of FF to be L.

We now state the major result of this section as a theorem, the proof of which is deferred to the next chapter.

Theorem 3.1.

FF is NP-complete.

3.3 The $\vec{r}$-Feasibility Problem

Let $\vec{r}$ be an o-d arrival rate vector of a PRN, G = (N,A). We assume that $\vec{r}$ is rational; that is, there exists a fixed, but arbitrary, integer P such that $P\vec{r}$ is integer.

Definition.

RF ($\vec{r}$-Feasibility Problem)

Instance: A PRN G = (N,A), an o-d arrival rate vector $\vec{r}$.

Question: Is $\vec{r}$ feasible, that is, is it true that $\vec{r} \in C(G)$?

The above problem is the same as the problem defined in section 3 of Chapter 1 except for the assumption that $\vec{r}$ is a rational vector.

We shall discuss the significance of this assumption in the next chapter. We close this section by stating a theorem which will be proved in Chapter IV.

Theorem 3.2.

RF is NP-complete.

3.4 Discussion of Results

The fact that such fundamental problems as FF and RF are NP-complete is discouraging, because the lack of adaptability of the network protocols to unexpected changes in network topology or in data traffic makes proper operation of the PRN uncertain.

CHAPTER IV. NP COMPLETENESS PROOFS AND AN ALTERNATIVE FORMULATION OF TDMA

In this chapter we shall prove the theorems we stated in Chapter III, and give a formulation of TDMA for the case where all slots are the same in length.

The transformation algorithms we use in this chapter and the next require constructions of certain sets from other sets. The set notation that we use in these algorithms is not a standard one, so, we should explain the notation first.

Let V be a set, say $V = \{x,y,z\}$, then V' denotes the set $\{x',y', z'\}$. In general, we write $V' = \{a' : a \in V\}$. Likewise, V'' is the set $\{a'' : a \in V\}$.

Often, we use indices instead of primes. If, for example, $Y = \{a, b\}$, then $Y_1 = \{a_1, b_1\}$, $Y_2 = \{a_2, b_2\}$, and so on. In general, we write $Y_i = \{a_i : a \in Y\}$.

This notation applies to sets of ordered pairs as well. For example, let $E = \{(a,b), (c,d), (a,d)\}$, then the set $\{(x_i,y_i) : (x,y) \in E\}$ written by its elements is $\{(a_i, b_i), (c_i,d_i), (a_i,d_i)\}$. Similarly, $\{(x_i,y_j) : (x,y) \in E\} = \{(a_i,b_j), (c_i, d_j), (a_i, d_j)\}$.

Finally, let us note the following type of equivalences:

$$\{a, a' : a \in V\} = V \cup V' ,$$

$$\{(a,b'), (a',b) : (a, b) \in E\} =$$

$$\{(a,b') : (a, b) \in E\} \cup \{(a',b) : (a,b) \in E\} .$$

In this chapter, $H = (V,E)$ denotes an undirected, finite, and simple graph.

4.1 NP-Completeness of FF

The NP-completeness proof for FF will be given in two steps. First, we shall consider a related problem and prove that it is NP-complete. Then, it will be shown that this problem is polynomially transformable to FF.

4.1.1 NP-Completeness of the Maximum Transmission Set Problem

Definition. MTS (The Maximum Transmission Set Problem)

Instance: PRN $G = (N,A)$ set $B \subseteq A$, positive integer $k \leq |B|$.

Question: Does there exist a transmission set S of G such that $|S \cap B| = k$?

We will prove the NP-completeness of MTS by polynomially transforming the following NP-complete problem to MTS.

Definition. MC (The Maximum Clique Problem)

Instance: Graph $H = (V,E)$, positive integer $k \leq |V|$.

Question: Does G contain a clique of size k ?

(A clique of $H = (V,E)$ is a set $Q \subseteq V$ such that $(a,b) \in E$ for every pair of distinct nodes $a, b \in Q$).

Fact 4.1. MC is NP-complete.

Proof. See, e.g., page 360 of [PS].

MC is polynomially transformable to MTS by using the following algorithm.

<u>Algorithm A1.</u>

Input: Graph $H = (V,E)$

Output: PRN $G = (N,A)$, set $B \subseteq A$ where

$N = V' \cup V''$

$A = \{(a',b''), (a'',b') : a \in V, b \in V, (a,b) \notin E\}$

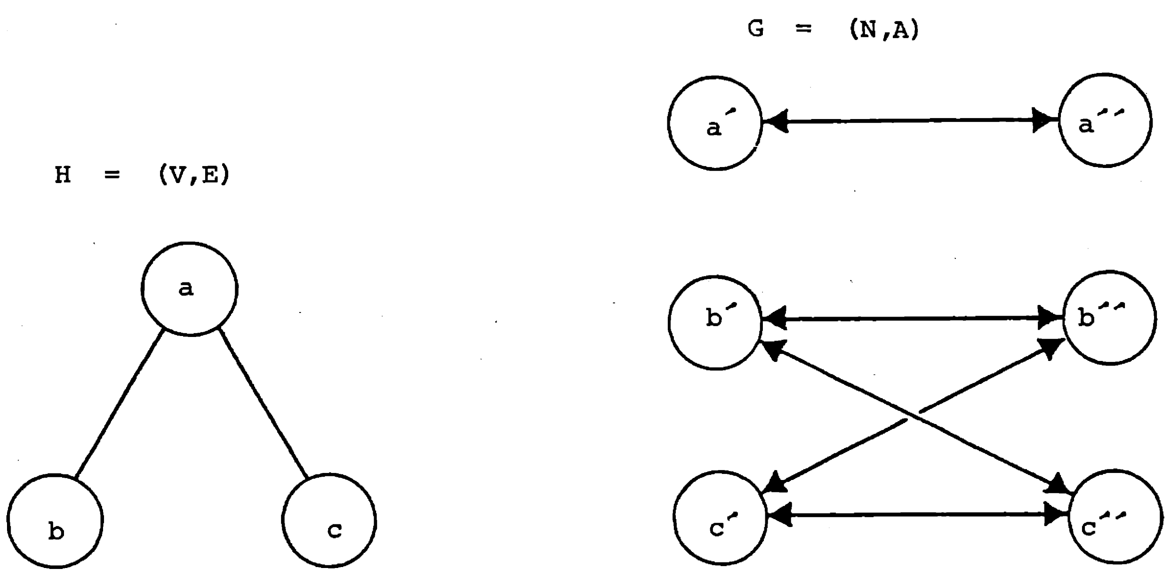
$B = \{(a', a'') : a \in V\}$

We illustrate the transformation by an example.

<u>Example.</u> Let $H = (V,E)$ be as in Figure 4.1.a.

$G = (N,A)$, the output of A1, is shown in Figure 4.1.b.

$B = \{(a',a''), (b',b''), (c',c'')\}$.



(a) (b)

Figure 4.1

Fact 4.2. Let $H = (V,E)$ and $G = (N,A)$ be the input and the output of A1, respectively. H has a clique of size k iff there exists a transmission set S of G such that $|S \cap B| = k$.

Proof. (i) Suppose Q is a clique of H with $|Q| = k$.
Then, $S = \{(a', a'') : a \in Q\}$ is a transmission set of G and $|S \cap B| = k$.

(ii) Suppose S is a transmission set of G with the property that $|S \cap B| = k$. Then, $Q = \{a \in V : (a', a'') \in (S \cap B)\}$ is a clique of H and $|Q| = k$.

Fact 4.3. MTS is NP - complete.

Proof. It is clear that the set S is a concise certificate that can be checked in polynomial time for validity for a YES instance of MTS. By Facts 4.1 - 4.2, we know that MTS is as hard as any problem in NP. Therefore, MTS is NP-complete.

4.1.2 MTS is Polynomially Transformable to FF.

The algorithm that polynomially transforms MTS to FF makes use of the following special device.

Definition. The m^{th} power of $G = (N,A)$ is $G^m = (N^m, A^m)$

where $N^m = \bigcup_{i=1}^{m} N_i$

$N_i = \{a_i : a \in N\}$ $\qquad i = 1,\ldots,m$

$A^m = \bigcup_{i=1}^{m} \bigcup_{j=1}^{m} A_{ij}$

$$A_{ii} = \{(a_i,b_i) : a_i \in N_i,\ b_i \in N_i,\quad a_i \neq b_i\} \qquad i = 1,\dots,m$$

$$A_{ij} = \{(a_i,b_j) : a_i \in N_i,\ b_j \in N_j,\quad a = b \ \text{ or } (a,b) \in A\} \qquad \begin{array}{l} i \neq j \\ i = 1,\dots,m \\ j = 1,\dots,m \end{array}$$

The best way to understand the structure of G^m is to consider an example.

Example. Let G be as in Figure 4.2a. G^2 is the graph in Figure 4.2.b.

Verify the following on Figure 4.2.

1) (N_1, A_{11}) and (N_2, A_{22}) are complete directed graphs.

2) $(b,a) \in A$ does not conflict with $(c,d) \in A$ in G; $(b_1,a_1) \in A_{11}$ does not conflict with (c_2,d_2) in G^2. More generally, if $(x,y) \in A$ does not conflict with $(u,v) \in A$, then $(x_1,y_1) \in A_{11}$ does not conflict with $(u_2,v_2) \in A_{22}$; it is also true that (x_2,y_2) does not conflict with (u_1,v_1).

With the above observations, we can prove the following general result about the powers of G.

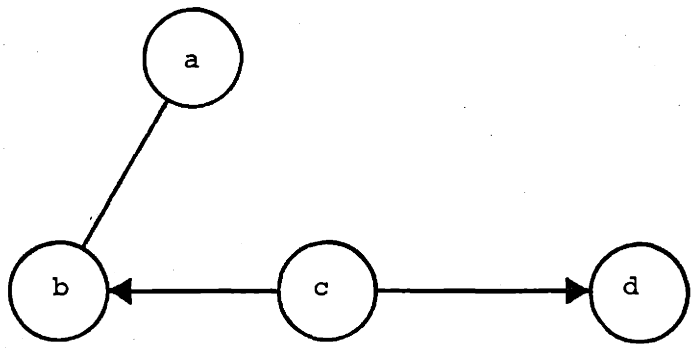
Fact 4.4. Let $G = (N,A)$ and $B \subseteq A$ be given, let G^k be the k^{th} power of G, and let $B^k = \bigcup_{i=1}^{k} B_i$ where

$$B_i = \{(a_i,b_i) : (a,b) \in B\}$$

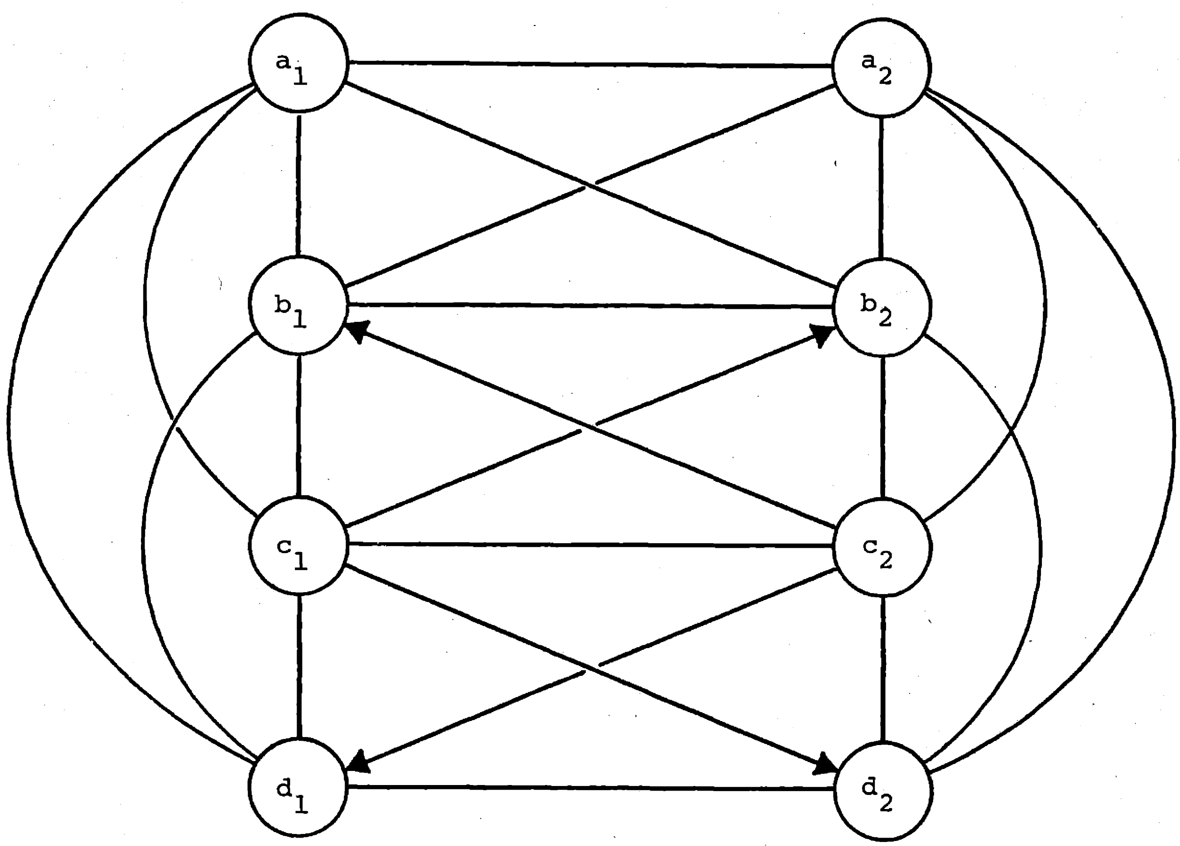
Then, the following statements are equivalent.

(i) There exists a transmission set S of G such that $|S \cap B| = k$.

(A link without an arrow stands for two oppositely directed links in this figure.)



(a) G = (N,A)



(b) $G^2 = (N^2, A^2)$

Figure 4.2

(ii) There exists a transmission set S^k of G^k such that $|S^k \cap B^k| = k$.

<u>Proof.</u> Suppose (i) is true. Let

$S \cap B = \{\ell_1, \ell_2, \ldots, \ell_k\}$. Let ℓ_{ij} be the copy of ℓ_i in A_{jj}, that is, if $\ell_i = (a,b)$, then $\ell_{ij} = (a_j, b_j)$ for $j = 1,\ldots,k$.

Then, $S^k = \{\ell_{ii} : i = 1,\ldots,k\}$ is a transmission set of G^k, because, for any two distinct links ℓ_{ii}, $\ell_{jj} \in S^k$, if ℓ_{ii} conflicts with ℓ_{jj}, then $\ell_i \in S$ conflicts with $\ell_j \in S$; thus, S cannot be a transmission set contrary to the hypothesis.

S^k contains one link in B_i $(i = 1,\ldots,k)$; hence $|S^k \cap B^k| = |S^k| = k$.

Suppose (ii) is true. Then we note that S^k contains one link in B_i $(i = 1,\ldots,k)$, because any two distinct links in B_i $(i = 1,\ldots,k)$ conflict with each other. In fact, S^k cannot contain more than one link in A_{ii} $(i = 1,\ldots,k)$. Thus,
$S = \{(a,b) : (a_i, b_i) \in S^k \text{ for some } i \in \{1,\ldots,k\}\}$ has exactly k elements. Moreover, $S \subseteq B$ and S is a transmission set of G.

This completes the proof.

MTS is polynomially transformable to FF by using the following algorithm.

<u>Algorithm A2.</u>

(The graph $\tilde{G} = (\tilde{N}, \tilde{A})$ which appears at the output of this algorithm has a special structure which can be visualized easily by first skimming over the following definitions, skipping the part about $\tilde{f}_{xy}$, studying the example, and then carefully reading the algorithm again.)

Input: PRN $G = (N,A)$, set $B \subseteq A$, positive integer $k \leq |B|$.

Output: PRN $\tilde{G} = (\tilde{N},\tilde{A})$, set $\{\tilde{f}_{xy} : (x,y) \in \tilde{A}\}$ where

$$\tilde{N} = N^k \cup P \cup Q$$

$$N^k = \bigcup_{i=1}^{k} N_i$$

$$N_i = \{a_i : a \in N\} \qquad i = 1,\ldots,k$$

$$P = \{p_1,p_2,\ldots,p_k\}$$

$$Q = \{q_1,q_2,\ldots,q_k\}$$

$$\tilde{A} = A^k \cup Y^k \cup Z^k$$

$$A^k = \bigcup_{i=1}^{k} \bigcup_{j=1}^{k} A_{ij}$$

$$A_{ii} = \{(a_i,b_i) : a_i \in N_i,\ b_i \in N_i,\ a \neq b\} \qquad i = 1,\ldots,k$$

$$A_{ij} = \{(a_i,b_j) : a_i \in N_i,\ b_j \in N_j,\ a = b \text{ or } (a,b) \in A\} \qquad i \neq j ;$$

$i = 1,\ldots,k$; $j = 1,\ldots,k$

(Note that (N^k, A^k) is just the k^{th} power of (N,A) decribed before.)

$$Y^k = \bigcup_{i=1}^{k} \bigcup_{j=1}^{k} Y_{ij}$$

$$Y_{ij} = \{(p_i,q_j),\ (q_j,p_i) : p_i \in P,\ q_j \in Q\} \qquad i = 1,\ldots,k ;$$

$j = 1,\ldots,k$

$$Z^k = \bigcup_{i=1}^{k} \bigcup_{\substack{j=1 \\ j\neq i}}^{k} Z_{ij}$$

$$Z_{ij} = \{(p_i,a_j),\ (a_j,p_i) : p_i \in P,\ a_j \in N_j\} \qquad j \neq i \ ;$$

$$i = 1,\ldots,k \ ; \quad j = 1,\ldots,k$$

$$\tilde{f}_{xy} = \begin{cases} \delta & (x,y) \in B^k \\ (|B| - 1)\,\delta & (x,y) \in D \\ 0 & \text{otherwise} \end{cases}$$

where

$$\delta = \frac{1}{1 + k\,(|B| - 1)}$$

$$B^k = \bigcup_{i=1}^{k} B_i$$

$$B_i = \{(a_i,b_i) : a_i \in N_i,\ b_i \in N_i,\ (a,b) \in B\} \qquad i = 1,\ldots,k$$

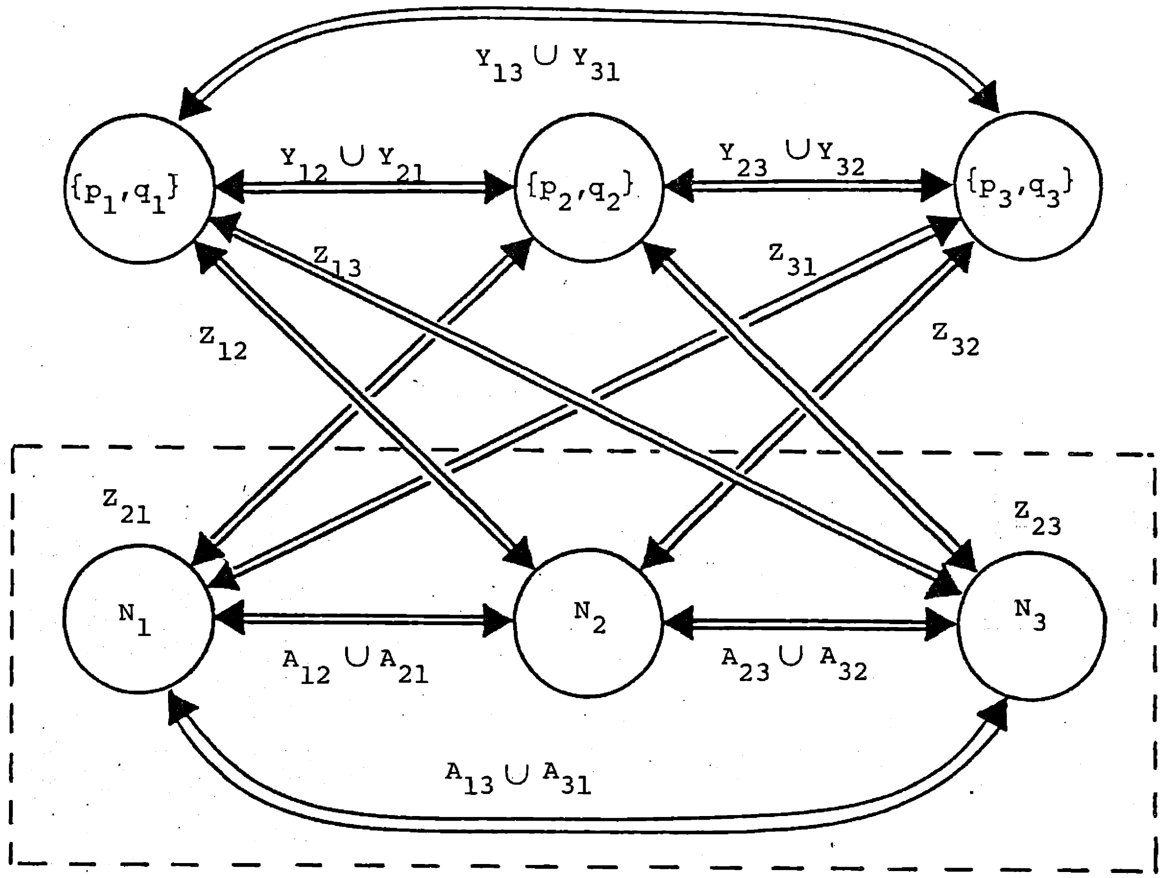
$$D = \{(p_i,q_i) : i = 1,\ldots,k\}$$

Example. Consider an instance of MTS <G,B,k> where $G = (N,A)$ is a PRN, $B \subseteq A$ is a subset of links, and $k \leq |B|$ is a positive integer. Suppose $|B| = 4$ and $k = 3$.

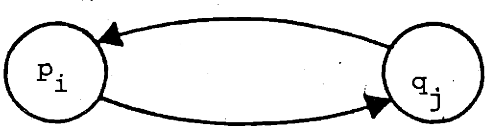
With <G,B,k> as the input, the output of A2 is illustrated in Figure 4.3.

In Figure 4.3.a, we see the general structure of $\tilde{G} = (\tilde{N},\tilde{A})$. The part of $\tilde{G}$ that is enclosed in the dashed rectangle is the same as G^3.

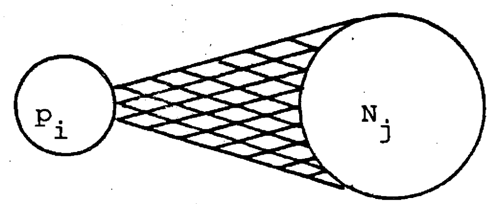
Y_{ij} is shown in Figure 4.3.b; we note that (P,Q,Y^k) is a complete bipartite directed graph, that is, $(x,y) \in Y^k$ iff either $x \in P$ and $y \in Q$, or $x \in Q$ and $y \in P$. In other words, each node in P is connected to each



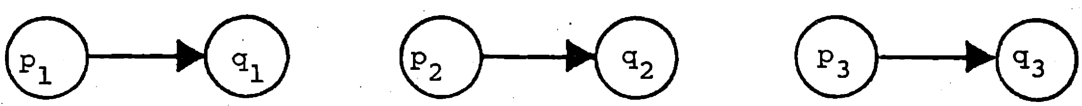
(a) $\tilde{G} = (\tilde{N},\tilde{A})$



(b) Y_{ij}



(c) Z_{ij} $(i \neq j)$



(d) D

Figure 4.3

node in Q, and vice versa. Thus, a transmission set $\tilde{S}$ of $\tilde{G}$ can have at most one link in common with Y^k.

The cross-hatched area in Figure 4.3.c represents the links in Z_{ij} $(i \neq j)$; we note that $(\{p_i\}, N_j, Z_{ij})$ is also a complete bipartite directed graph. Hence, a transmission by node p_i interferes with a simultaneous transmission over any link in A_{jj} $(j \neq i)$. So, if a transmission set $\tilde{S}$ of $\tilde{G}$ contains a link $(p_i, x) \in \tilde{A}$, then $\tilde{S}$ cannot contain any link in A_{jj} $(j \neq i)$.

Finally, in Figure 4.3.d, we see the links in D. Each of the links in D is assigned a mean link arrival rate of $(|B| - 1)\ \delta = 0.3$ units. (Here we do not use packets/slot as the unit of arrival rates, because slot, as a time unit, does not have a meaning in a variable slot length TDMA scheme. Instead, we interpret the arrival rates as service requirements; thus, a mean link arrival rate of, say, 0.1 units implies that the link in question must be enabled 0.1 fraction of time to satisfy its service requirement; sometimes, we express the same fact by saying that the link has a service requirement of 0.1 units.)

The following result is based on the properties of transmission sets of $\tilde{G}$ which have been pointed out in the above example.

<u>Fact 4.5.</u>

Let the instance $\langle G,B,k \rangle$ of MTS be the input of A2, and let $\langle \tilde{G},\tilde{f} \rangle$ be the corresponding output. Then, the following statements are equivalent.

(i) There exists a transmission set S of G such that $|S \cap B| = k$.

(ii) $\{\tilde{f}_{xy} : (x,y) \in \tilde{A}\}$ is feasible under TDMA.

Proof. There are $k|B|$ links in B^k each with δ units of service requirement, i.e., each link in B^k must be enabled δ fraction of time on the average; otherwise, packets start piling up. There are k links in D each with a service requirement of $(|B| - 1)\delta$ units. So, the sum of the service requirements on all links of $\tilde{G}$ is $k\delta + 2k(|B| - 1)\delta$ units (or $k\delta + 2(1 - \delta)$ units).

Now, suppose (i) is true. Then, there exists a transmission set $\tilde{S}$ of $\tilde{G}$ such that $|\tilde{S} \cap B^k| = k$. $\tilde{S}$ can be used δ fraction of time to satisfy the service requirements of k links in B^k. So, if, in the remaining $(1 - \delta)$ fraction of time, we can complete the service of $k(|B| - 1)$ links in B^k each with a requirement of δ units, and k links in D each with a requirement of $(|B| - 1)\delta$ units, then statement (ii) will be true. We shall show that this is indeed the case.

We know that $\tilde{S}$ has one link in each B_i $(i=1,\ldots,k)$. So, let us write

$B_i - \tilde{S} = \{\ell_{ji} : j = 1,\ldots,k-1\}$ $i = 1,\ldots,k$. The set $S_{ij} = \{(p_i,q_i), \ell_{ji}\}$ is a transmission set of $\tilde{G}$ for each $\ell_{ji} \in (B_i - \tilde{S})$, $i = 1,\ldots,k$.

It is easy to see that the TDMA scheme which uses $\tilde{S}$ and S_{ij} $(i = 1,\ldots,k\ ;\ j = 1,\ldots,k-1)$ each for δ fraction of time satisfies the service requirements on all links of $\tilde{G}$ just on time.

Suppose (ii) is true. We note that exactly $k(|B| - 1)\delta$ (or $(1-\delta)$) fraction of time must be spent in satisfying the service requirements of the links in D. In the remaining δ fraction of time, the TDMA scheme under which $\{\tilde{f}_{xy}\}$ is feasible must satisfy a total of at least $k\delta$ units of service requirements on the average. (This last sentence follows by noting that the TDMA scheme cannot use transmission sets of size larger than 2 in the $(1-\delta)$ fraction of time that it spends while serving the links in D). Therefore, the average number of links in B^k that must be enabled by the transmission sets used in the remaining δ fraction of time is at least k. This implies that there exists a transmission set $\tilde{S}$ of $\tilde{G}$ such that $|\tilde{S} \cap B^k| = k$. But, the existence of $\tilde{S}$ implies that there exists a transmission set S of G such that $|S \cap B| = k$. This completes the proof.

To prove that FF is NP-complete, we need to show that every YES instance of FF has a concise certificate which can be checked in time bounded polynomially in the size of the instance.

Let $\vec{f}$ be feasible under TDMA. There exists linearly independent transmission vectors $\vec{t}_1, \vec{t}_2, \ldots, \vec{t}_L$ such that, for some $\vec{x} \geq \vec{0}$, $B\vec{x} = \vec{f}$ and $\vec{1} \cdot \vec{x} \leq 1$ where B is the LxL matrix whose i^{th} column is $\vec{t}_i$ $(i=1,\ldots,L)$. We propose B and $\vec{x}$ as a concise certificate for FF. First, let us see if $\vec{x}$ can be encoded concisely.

We have $\vec{x} = B^{-1} \vec{f}$.

Let $(B^{-1})_{ij}$ be the i-j^{th} element of B^{-1}, that is,

$$(B^{-1})_{ij} = \frac{\text{adj}(B_{ji})}{\det B}$$

Since $\text{adj}(B_{ji})$ is the determinant of an $(L-1)\times(L-1)$ 0-1 matrix, $|\text{adj}(B_{ji})| < (L-2)!$ Likewise, $|\det B| < (L-1)!$. Therefore, each of $|\det(B)|$ and $|\text{adj}(B_{ij})|$ can be encoded in fewer than L^2 bits. Since $\vec{f}$ is assumed to be rational, we conclude that $\vec{x}$ can be encoded in space bounded polynomially in L.

The columns of B can be checked in polynomial time for being transmission vectors. The operation $B\vec{x}$ can also be carried out in polynomial time. Therefore, if $\vec{f}$ is feasible under TDMA, then there exists a concise certificate. This proves that FF belongs to NP.

Notice that we have not used the assumption that $\vec{f}$ is rational in proving the fact that MTS is polynomially transformable to FF. This assumption was only necessary to show that FF belongs to NP.

4.2. NP - Completeness of RF

We define the total o-d rate associated with node x of PRN G = (N,A) as follows:

$$R_x = \sum_{y \in N} r_{xy} + \sum_{y \in N} r_{yx} \quad .$$

R_x can be interpreted as the fraction of time node x is receiving or transmitting packets which either originate from x or leave the network at x. If $\vec{r}$ is feasible, then $R_x \leq 1$ for all $x \in N$.

The following algorithm polynomially transforms FF to RF.

Algorithm A3.

Input: PRN $G = (N,A)$, set $\{f_{ab} : (a,b) \in A\}$

Output: PRN $\tilde{G} = (\tilde{N},\tilde{A})$, set $\{r_{xy} : x \in \tilde{N}, y \in \tilde{N}\}$

where $\tilde{N} = \{a, a', a'' : a \in N\}$

$\tilde{A} = A \cup A_1 \cup A_2$

$A_1 = \{(a', a'') : a \in N\}$

$$A_2 = \{(a,a''),\ (a'',a),\ (a',a),\ (a,a'),\ (a'',a') \ :\ a \in N\}$$

$$\tilde{r}_{ab} = f_{ab} \qquad\qquad (a,b) \in A$$

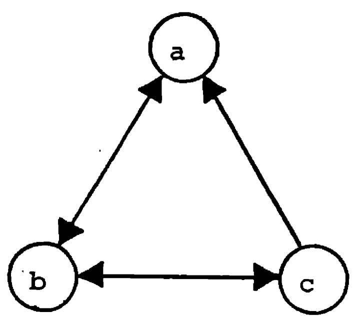
$$\tilde{r}_{a'a''} = 1 - \sum_{b} (f_{ab} + f_{ba}) \qquad\qquad (a',a'') \in A_1$$

$$\tilde{r}_{xy} = 0 \qquad\qquad (x,y) \notin \tilde{A} \text{ or } (x,y) \in A_2$$

Note that $\tilde{r}_{a'a''} = 1 - \tilde{R}_a$. Without loss of generality, we shall assume that $0 \leq \tilde{R}_x \leq 1$ for all $x \in \tilde{N}$. When $\{f_{ab} : (a,b) \in A\}$ is feasible under TDMA, this assumption will always be true.

We illustrate the algorithm by an example.

Example. Let $G = (N,A)$ be the PRN in Figure 4.4.a. $\tilde{G} = (\tilde{N},\tilde{A})$ is shown in Figure 4.4.b.



(a) $G = (N,A)$

$f_{ab} = 0.2$

$f_{ba} = 0.3$

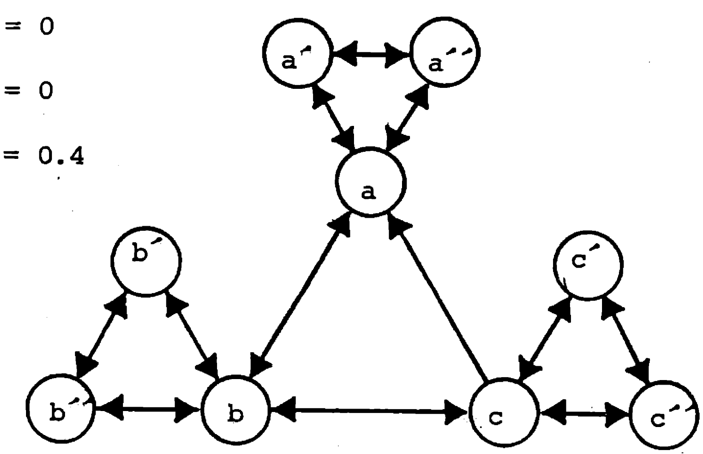
$f_{bc} = 0.1$

$f_{cb} = 0$

$f_{ca} = 0.5$

$\tilde{r}_{ab}$	$= 0.2$	$\tilde{r}_{ba}$	$= 0.3$
$\tilde{r}_{bc}$	$= 0.1$	$\tilde{r}_{cb}$	$= 0$
$\tilde{r}_{ca}$	$= 0.5$	$\tilde{r}_{a'a''}$	$= 0$
$\tilde{r}_{b'b''}$	$= 0.4$	$\tilde{r}_{c'c''}$	$= 0.4$

All other o-d rates are zero.



(b) $\tilde{G} = (\tilde{N},\tilde{A})$

Figure 4.4

Fact 4.6. Let the PRN $G = (N,A)$, and the set of mean link arrival rates $\{f_{ab} : (a,b) \in A\}$ be the input of A3, and let $\tilde{G} = (\tilde{N},\tilde{A})$ and $\{\tilde{r}_{xy} : x \in \tilde{N}, y \in \tilde{N}\}$ be the corresponding output. The following statements are equivalent.

(i) $\{f_{ab} : (a,b) \in A\}$ is feasible under TDMA

(ii) $\{\tilde{r}_{xy} : x \in \tilde{N}, y \in \tilde{N}\}$ is feasible.

Proof. Suppose there exists a TDMA scheme for G under which $\{f_{ab} : (a,b) \in A\}$ is feasible. We shall show that $\{\tilde{r}_{xy} : x \in \tilde{N}, y \in \tilde{N}\}$ is feasible in $\tilde{G}$. Since $\tilde{r}_{xy} = 0$ when $(x,y) \notin \tilde{A}$, each packet in the network can be transmitted directly from its origin to its destination. Let this be our rule for the assignment of paths to packets; that is, every $(x - y)$ packet is transmitted directly over link (x,y) for all $(x,y) \in \tilde{A}$. With this path assignment, we have $\tilde{f}_{xy} = \tilde{r}_{xy}$ for all $(x,y) \in \tilde{A}$, where $\tilde{f}_{xy}$ is the mean arrival rate of link (x,y).

If S is a transmission set of G such that

$$S \cap \{(x,y) \in A : x = a \text{ or } y = a\} = \emptyset \quad ,$$

then $S \cup \{(a',a'')\}$ is a transmission set of $\tilde{G}$. Thus, given a TDMA scheme for G under which $\{f_{ab} : (a,b) \in A\}$ is feasible, we can augment the transmission sets used by this TDMA scheme to obtain a TDMA scheme for $\tilde{G}$ under which $\{\tilde{f}_{xy} : (x,y) \in A\}$ is feasible.

Suppose (ii) is true. Consider nodes $a, a', a'' \in \tilde{N}$. Node a' must be transmitting at least $\tilde{r}_{a'a''}$ fraction of time, and node a cannot transmit or receive from nodes other than a' during this time. Node a must spend at least $1 - \tilde{r}_{a'a''} = \tilde{R}_a$ fraction of time receiving from and transmitting to nodes other than a' and a''. Thus, all $(a' - a'')$ traffic must be transmitted directly over (a', a''). But, if all $(a' - a'')$ traffic is sent directly over (a', a''), node a cannot handle any packets other than those which either originate at node a or leave the network at node a. Therefore, one must be able to send each packet directly from its origin to its destination if $\{\tilde{r}_{xy} : x \in \tilde{N}, y \in \tilde{N}\}$ is feasible. This completes the proof.

A concise certificate for a YES instance of RF is a set of routing variables and a TDMA scheme under which the induced set of mean link arrival rates are feasible.

Thus, we have proved that RF is NP-complete.

4.3 Fixed Slot Length TDMA Schemes

Until now, we have considered variable slot length TDMA schemes. Sometimes, it may be desirable to have a scheme with all slots fixed in length. In this section, we shall show that this can be done without loss of capacity, in the sense that, if a set of mean link arrival rates are feasible under TDMA, then there exists a fixed slot length (FSL) TDMA scheme under which the same set of mean link arrival rates are feasible. The converse of this statement is also true.

Definition. Let $G = (N,A)$ be a PRN, and let $\vec{f}$ be a mean link arrival rate vector. A FSL - TDMA scheme with period J is a system $\langle T, \vec{u}, J \rangle$ where T is a transmission matrix of G, $\vec{u}$ is a column k - vector with non-negative integer elements and J is integer such that $J \geq \vec{1} \cdot \vec{u}$. We say $\vec{f}$ is feasible under FSL-TDMA if there exists a FSL-TDMA scheme $\langle T,\vec{u},J\rangle$ such that $T\,\vec{u} \geq J\,\vec{f}$.

Fact 4.7. If for some P, $P \cdot \vec{f}$ has integer elements, and if $\vec{f}$ is feasible under TDMA, then there exists a FSL-TDMA scheme $\langle T,\vec{u},J\rangle$ such that $T\,\vec{u} \geq J\,\vec{f}$ and $J < P \cdot (L-1)!$.

Proof. If $\vec{f}$ is feasible under TDMA, then, as we have shown in the NP-completeness proof of FF, there exists a TDMA scheme $\langle T,\vec{x}\rangle$ such that $T\vec{x} = \vec{f}$ and $P \cdot M\vec{x}$ has integer elements, where M is the determinant of an LxL 0-1 matrix. Since the determinant of an LxL 0-1 matrix must be less than $(L-1)!$, we have that $M < (L-1)!$. Let $\vec{u} = P \cdot M\,\vec{x}$ and $J = P \cdot M$. Then, $T\,\vec{u} = P \cdot M \cdot T\,\vec{x} = P \cdot M\,\vec{f} = J \cdot \vec{f}$ and $\vec{1} \cdot \vec{u} = M \cdot P \cdot \vec{1} \cdot \vec{x} \leq MP = J$. Therefore, $\langle T,\vec{u},J\rangle$, as defined above, is a FSL - TDMA scheme under which $\vec{f}$ is feasible and $J < P \cdot (L-1)!$.

Fact 4.8. If $\vec{f}$ is feasible under a FSL-TDMA scheme $\langle T,\vec{u},J\rangle$, then there exists a TDMA scheme $\langle T,\vec{x}\rangle$ under which $\vec{f}$ is feasible.

Proof. By letting $\vec{x} = \frac{1}{J}\vec{u}$, we obtain the desired TDMA scheme.

Thus, we have shown that if $\vec{f}$ is a rational vector, then feasibility under TDMA is equivalent to feasibility under FSL-TDMA.

Sometimes, we may wish to find a FSL-TDMA scheme with a fixed J. In this case the feasibility of $\vec{f}$ is equivalent to the non-emptiness of the region defined by the following inequalities

$$\sum_{i=1}^{J} x_{ij} \geq J\,f_j \qquad k = 1,\ldots,L$$

$$x_{ij} + x_{ik} \leq 1 \qquad j \in C_k;\ i=1,\ldots,J;\ j = 1,\ldots,L;\ k = 1,\ldots,L$$

$$x_{ij} \in \{0,1\} \qquad i = 1,\ldots,J;\ j = 1,\ldots,L$$

We interpret the binary variables x_{ij} as follows:

$$x_{ij} = \begin{cases} 1 & \text{if a transmission over link } j \\ & \text{is to be made at slot } i \\ 0 & \text{otherwise} \end{cases}$$

If the above region is empty for a certain J, $\vec{f}$ may still be feasible under FSL-TDMA for some $J' \neq J$.

We finish this chapter by stating the NP-completeness of a restricted version of the $\vec{f}$-feasibility problem under FSL-TDMA.

Instance: PRN G = (N,A), mean link arrival rate vector $\vec{f}$ such that $P \cdot \vec{f}$ has integer elements for some fixed - but arbitrary - positive integer P, integer J_o such that $3 \leq J_o \leq P \cdot (L-1)$!

Question: Does there exist a FSL-TDMA scheme $\langle T,\vec{u},J \rangle$ such that $J \leq J_o$ and $T\,\vec{u} \geq J\,\vec{f}$?

Comment: This problem remains NP-complete for all fixed $J_o \geq 3$. (As stated above, J_o is part of the input and is arbitrary in the specified range.)

We omit the proof of this fact, but let us note that a simple proof is possible by a transformation from the PARTITION INTO CLIQUES problem (see page 193 of [GJ] for the definition).

CHAPTER V. POLYNOMIAL TIME APPROXIMATION ALGORITHMS FOR FF AND RF

In this chapter we discuss polynomial time approximation algorithms for the feasibility problems, and show why it is difficult to solve them efficiently, even in an approximate sense.

5.1 Terminology [PS]

Let P be an optimization problem with a positive cost function; let I be an instance of P; let $\hat{V}(I)$ be an optimal solution to I; and let $c(\hat{V}(I))$ be the cost of $\hat{V}(I)$.

Let A be an algorithm which returns a feasible solution $V_a(I)$ when supplied with an instance I; and let $c(V_a(I))$ be the cost of $V_a(I)$.

The algorithm A is called an ϵ - approximate algorithm if, for any instance I of P, we have

$$\frac{|c(V_a(I)) - c(\hat{V}(I))|}{c(\hat{V}(I))} \leq \epsilon \quad .$$

An ϵ - AA (ϵ-approximate algorithm) which operates in polynomial time is called an ϵ - PTAA (ϵ - polynomial time AA).

The solution $V_a(I)$ returned by an ϵ-AA A is called an ϵ - approximate solution.

A polynomial time approximation scheme (PTAS) for P is an algorithm which, when supplied with an instance I of P and a number $\epsilon > 0$, returns an ϵ - approximate solution in time bounded polynomially in the size of the instance, where the polynomial depends on ϵ.

A PTAS is called a FPTAS (fully PTAS) if its running time is bounded polynomially in $1/\epsilon$ and the size of the instance.

5.2 Negative Results About FF

The optimization version of FF, which we shall still denote by FF, is the problem of finding a vector $\vec{x} \geq \vec{0}$, such that, for a given mean link arrival rate vector $\vec{f}$, $T\vec{x} \geq \vec{f}$ and $\vec{1} \cdot \vec{x}$ is minimized. This problem has a solution whether or not $\vec{f}$ is feasible. We shall denote the cost of an optimal solution to an instance $I = \langle G,\vec{f}\rangle$ of FF by $\hat{Z}(I)$; the cost of an approximate solution will be denoted by $Z_a(I)$.

We now list the main results about FF and prove them in the following sections.

Theorem 5.1.

For FF, there are two possibilities: either there is no ϵ - PTAA for any value of ϵ, or there exists a PTAS for all $\epsilon > 0$.

Theorem 5.2.

If there is a PTAS for FF, then there is a PTAS for MC (i.e. the optimization version of the maximum clique problem).

Theorem 5.3.

Unless $P = NP$, there is no FPTAS for FF.

We are unable either to find an ϵ - PTAA for FF or to prove that, unless P = NP, there exists no ϵ - PTAA for FF. However, in view of Theorem 5.2 and the fact that a PTAS for MC - an extensively studied problem - is yet to be found, the evidence about the difficulty of finding an ϵ-PTAA for FF is conclusive.

5.3 Proof of Theorem 5.1

Let $I = \langle G, \vec{f} \rangle$ be an instance of FF. We define αI to be the instance $\langle G, \alpha\vec{f} \rangle$ where $\alpha \geq 0$.

The instance $\tilde{I} = \tilde{I}(I) = \langle \tilde{G}, \vec{\tilde{f}} \rangle$ is defined as follows:

$$\tilde{G} = (\tilde{N}, \tilde{A})$$

$$\tilde{N} = \{a_i : a \in N,\ i \in \{i \in A : f_i > 0\}\}$$

$$\tilde{A} = \{(a_i, b_i) : a_i \in \tilde{N},\ b_i \in \tilde{N},\ (a,b) \in A\} \cup$$

$$\{(a_i, b_j),\ (b_j, a_i) : i \neq j,\ a_i \in \tilde{N},\ b_j \in \tilde{N},\ i \in C_j\}$$

$$\tilde{f}_{a_i b_j} = \begin{cases} f_i \cdot f_{ab} & i = j \\ 0 & i \neq j \end{cases}$$

(Note that f_i is the mean link arrival rate on link i of G; f_{ab} is the rate on link (a,b) of G.)

Since the case $\vec{f} = \vec{0}$ is trivial, we shall always assume that there is at least one link in G with a positive mean packet arrival rate, and we shall let k denote the number of links in G with positive mean arrival rates. The links will be labeled such that $f_1 > 0, \ldots, f_k > 0$, and $f_{k+1} = \ldots = f_{|A|} = 0$.

The subgraph

$$G_i = (\{a_i \in \tilde{N} : a \in N\}, \{(a_i,b_i) \in \tilde{A} : (a,b) \in A\})$$

of $\tilde{G}$ will be referred to as the i^{th} copy of G in $\tilde{G}$, because G_i is isomorphic to G. Thus, there are k copies of G in $\tilde{G}$.

Before we proceed with the proof, let us illustrate the graph $\tilde{G} = (\tilde{N},\tilde{A})$ by an example.

Example. Let $I = \langle G,\vec{f}\rangle$ be as in Figure 5.1.a. The instance $\tilde{I} = \langle \tilde{G},\tilde{\vec{f}}\rangle$ is shown in Figure 5.1.b.

The cross-hatched areas between the copies represent links ; thus, e.g., each node in copy G_1 is connected to each node in copy G_2 by directed links, and vice versa.

In general, for any two copies G_i and G_j, there are two cases:

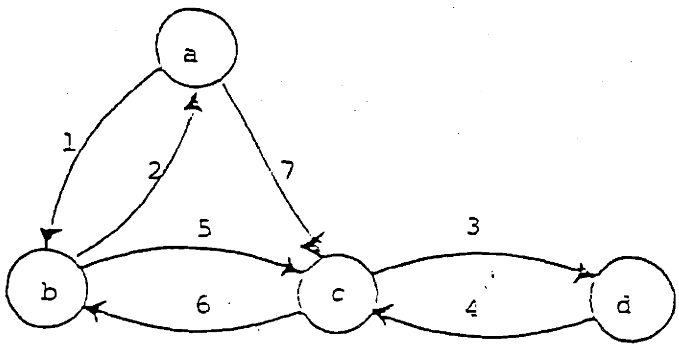
(1) If $i \in C_j$ in G, i.e. if link $i \in A$ conflicts with link $j \in A$, then each link in copy G_i conflicts with each link in G_j;

(2) If $i \notin C_j$ in G, then there is no link in G_i that conflicts with any link in G_j.

The mean link arrival rate vector associated with the links in G_i will be denoted by $\vec{f}_i$. The instance $\tilde{I}$ contains k instances $I_1,\ldots,I_k$ where $I_i = \langle G_i,\vec{f}_i\rangle$. Notice that I_i and $f_i I$ are identical.

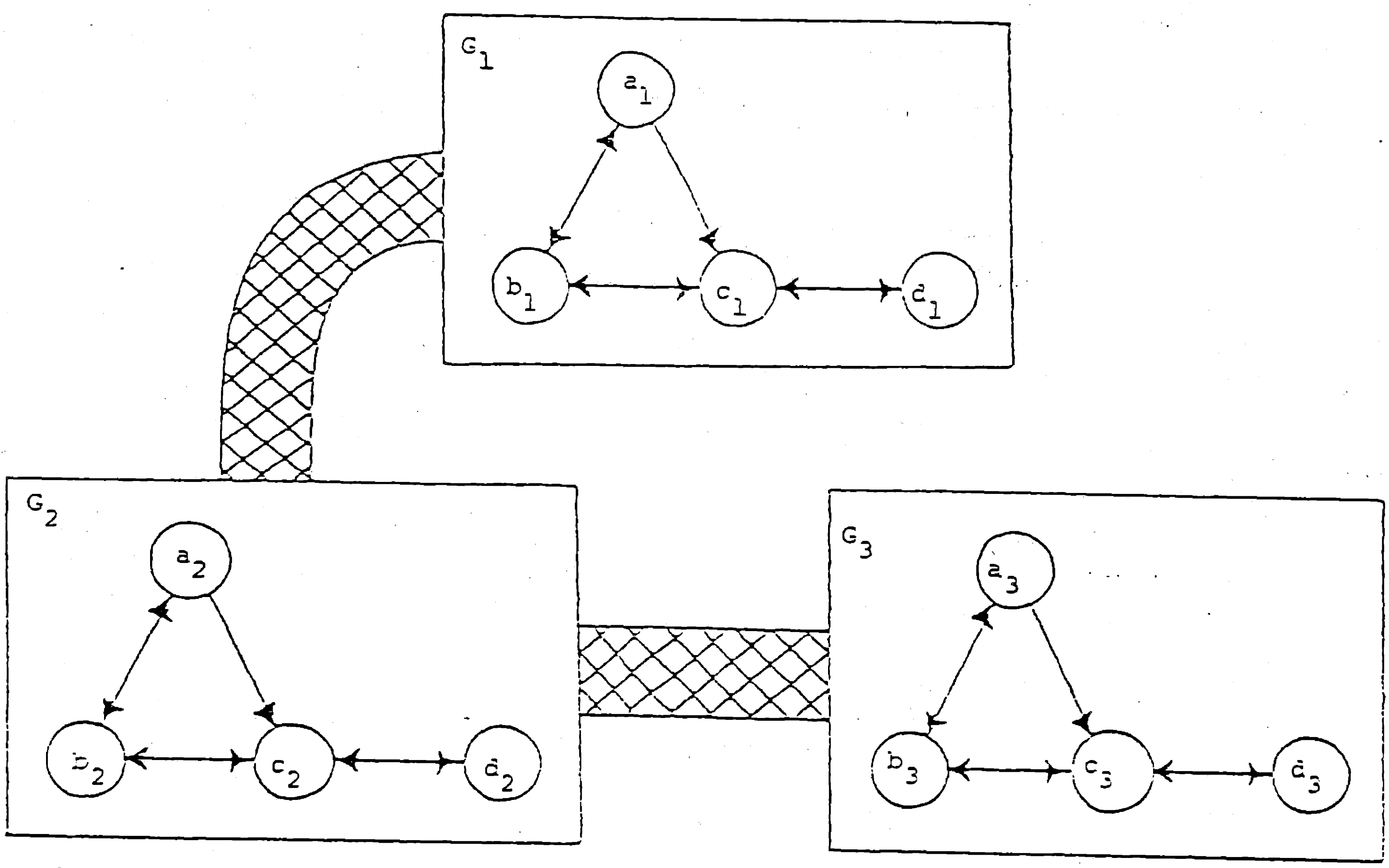
We shall first argue that $\hat{Z}(\tilde{I}) \leq (\hat{Z}(I))^2$. Let $\hat{V}(I) = \langle \vec{u}_i, y_i : i=1,\ldots,|A|\rangle$ be an optimal solution to I where $\vec{u}_i$ is a transmission vector of G and y_i is the time duration associated with $\vec{u}_i$.

(a) $G = (N,A)$



$f_1 = .4, \quad f_2 = .3, \quad f_3 = .2, \quad f_4 = f_5 = f_6 = f_7 = 0$

(b) $\tilde{G} = (\tilde{N},\tilde{A})$



$\tilde{f}_{a_1b_1} = .16$
$\tilde{f}_{a_2b_2} = .12$
$\tilde{f}_{a_3b_3} = .08$

$\tilde{f}_{b_1a_1} = .12$
$\tilde{f}_{b_2a_2} = .09$
$\tilde{f}_{b_3a_3} = .06$

$\tilde{f}_{c_1d_1} = .08$
$\tilde{f}_{c_2d_2} = .06$
$\tilde{f}_{c_3d_3} = .04$

Figure 5.1

The system $f_j\hat{V}(I) = \langle \vec{u}_i, f_j y_i : i = 1,\ldots,|A|\rangle$ is an optimal solution to the instance $f_j I$ and has a cost of $f_j\hat{Z}(I)$ for each $j = 1,\ldots,k$.

Thus, $\hat{V}(I)$ can be used to solve each of the instances I_i $(i=1,\ldots,k)$ optimally. Moreover, the solutions of these instances can be performed concurrently since certain collections of copies of G are independent of each other. To be more precise, let $\vec{u}$ be a transmission vector of G such that $u_i = 0$ for $i > k$, i.e., $\vec{u}$ uses only those links of G with non-zero mean packet arrival rates. Consider the set $\{G_i : u_i = 1\}$; copies of G in this set are independent in the sense that, if s_i is a transmission set of G_i, then

$$\bigcup_{i\in\{i:u_i=1\}} s_i \qquad \text{is a transmission set of } \tilde{G}.$$

It is now easy to see that there is a feasible solution to $\tilde{I}$ with a cost of $(\hat{Z}(I))^2$. Therefore, $\hat{Z}(\tilde{I}) \leq (\hat{Z}(I))^2$.

Now, we shall give a reverse argument. A feasible solution $V(\tilde{I})$ to $\tilde{I}$ induces a solution to each instance I_i $(i=1,\ldots,k)$. More precisely, if $V(\tilde{I}) = \langle \vec{t}_j, x_j : j = 1,\ldots,r\rangle$, then the induced solution $V(I_i)$ to I_i is the system $\langle \vec{u}_{ij}, y_{ij} : j = 1,\ldots,r\rangle$ where $\vec{u}_{ij}$ is the $0-1$ column $|A|$ - vector obtained by deleting all rows of $\vec{t}_j$ that correspond to links which do not belong to G_i and

$$y_{ij} = \begin{cases} x_j & \vec{u}_{ij} \neq \vec{0} \\ 0 & \text{otherwise.} \end{cases}$$

The reader should verify that $\vec{u}_{ij}$ is a transmission vector of G_i $(i=1,\ldots,k;\ j=1,\ldots,r)$, and $\sum_{j=1}^{r} \vec{u}_{ij}\, y_{ij} \geq \vec{f}_i$ $(i=1,\ldots,k)$.

The cost of $V(I_i)$, denoted by Z_i, is $\sum_{j=1}^{r} y_{ij}$.

Let us associate a column 0-1 $|A|$ - vector $\vec{q}_i$ with transmission vector $\vec{t}_i$ of $\tilde{G}$ as follows:

$$q_{ij} = \begin{cases} 1 & j \in \{1,\ldots,k\} \text{ and at least one link in } G_j \text{ is used by } \vec{t}_i. \\ 0 & \text{otherwise} \end{cases}$$

It should be clear $\vec{q}_i$ is a transmission vector of G.

Now, consider the instance $I' = \langle G, \vec{f}' \rangle$ where

$$f_i' = \begin{cases} Z_i & i = 1,\ldots,k \\ 0 & i = k+1,\ldots, |A| \end{cases}$$

The system $V(I') = \langle \vec{q}_i, x_i : i = 1,\ldots,r \rangle$ is a feasible solution to I' and the cost of $V(I')$, denoted by $Z(I')$, equals $Z(\tilde{I})$.

Let us define $\alpha = \min\{(Z_i/f_i): i = 1,\ldots,k\}$, and compare the instance I with $\frac{1}{\alpha} I'$. Since the mean link arrival rates of $\frac{1}{\alpha} I'$ dominate the corresponding mean link arrival rates of I, the system $\frac{1}{\alpha} V(I') = \langle \vec{q}_i, \frac{x_i}{\alpha} : i = 1,\ldots,r \rangle$ is a feasible solution to I with a cost of $\frac{1}{\alpha} Z(I')$.

Therefore, any feasible solution $V(\tilde{I})$ to $\tilde{I}$ induces k+1 feasible solutions to I; k of which are on the copies of G, and one is on the instance $\frac{1}{\alpha} I'$. The solution induced on G_i is the system $\frac{1}{f_i} V(I_i) = \langle \vec{u}_{ij}, (y_{ij}/f_i) : j=1,\ldots,r \rangle$, and has a cost of (Z_i/f_i).

The induced solution with the least-cost has a cost of

$Z(I) = \min\{\alpha, (Z(I')/\alpha)\}$ since $\alpha = \min\{(z_i/f_i) : i=1,\ldots,k\}$. Therefore, $Z(\tilde{I}) = Z(I') \geq Z(I)\,\alpha \geq (Z(I))^2$.

The optimal costs associated with I and $\tilde{I}$ must also satisfy $\hat{Z}(\tilde{I}) \leq (\hat{Z}(I))^2$. Thus we conclude that $\hat{Z}(\tilde{I}) = (\hat{Z}(I))^2$.

Now suppose we have an ϵ - PTAA for FF; when we apply this algorithm to an instance I of FF, we are guaranteed a solution $V_a(I)$ such that $\frac{Z_a(I)}{\hat{Z}(I)} \leq 1 + \epsilon$, where $Z_a(I) = c(V_a(I))$.

If we apply this ϵ - PTAA to $\tilde{I}$ instead of I, we still obtain a solution such that $\frac{Z_a(\tilde{I})}{\hat{Z}(\tilde{I})} \leq 1 + \epsilon$. But then, by searching through the induced solutions, we can obtain a solution $V(I)$, whose cost satisfies $\frac{Z(I)}{\hat{Z}(I)} \leq \sqrt{1 + \epsilon}$. Thus, the existence of an ϵ - PTAA implies the existence of a $(\sqrt{1 + \epsilon} - 1)$ - PTAA. (We leave it to the reader to prove that the new algorithm is still polynomial time.)

Suppose, for an instance I of FF, a δ- approximate solution is desired ($\delta>0$). How do we use the ϵ - PTAA - which we assume exists - to obtain the desired accuracy? If $\delta \geq \epsilon$, then there is no problem. Let us consider the case $\delta < \epsilon$, and let m be the smallest integer such that $1 + \delta > (1 + \epsilon)^{2^{-m}}$. Then, we apply the ϵ - PTAA to the m^{th} order instance $\tilde{I}^{(m)}$, which is obtained from I by recursively applying the construction of $\tilde{I}$ m times, and descend to a $((1 + \epsilon)^{2^{-m}} - 1)$ - approximate solution to I by following the least-cost induced solutions at each stage.

This completes the proof of Theorem 5.1.

5.4 Proof of Theorem 5.2

The proof is based on two observations:

(1) There exists a PTAS for MTS if there exists a PTAS for FF,

(2) There exists a PTAS for MC if there exists a PTAS for MTS.

The second observation readily follows from the transformation algorithm A1 of Chapter IV; we leave out the proof.

To prove the first observation, we shall consider the transformation algorithm A2 of Chapter IV.

Suppose there is a PTAS for FF and we wish to obtain a γ - approximate solution to an instance $I = \langle G,B \rangle$ of MTS, where $G = (N,A)$ is a PRN, and $B \subseteq A$. By using A2, we can transform I to an instance of FF. But, there is a difficulty since we do not know what value of k to use in the transformation. The difficulty is overcome by trying all values of k between 1 and $|B|$. For each value of k, we transform the instance I of MTS to an instance of FF, and we operate the PTAS for FF with an accuracy requirement of ϵ such that $0 < \epsilon \leq \frac{\gamma}{(1-\gamma)(1-|B|+|B|^2)}$. Each time we obtain an approximate solution to MTS by simply searching through the transmission vectors used by the PTAS. We claim that the best of these approximate solutions is a γ- approximate solution to the instance I of MTS. To see why, suppose the optimal value of the instance I is k^*. When the PTAS is operated on the instance of FF with $k = k^*$, an ϵ - approximate solution must have a cost less than or equal to $1 + \epsilon$, because the optimal cost equals 1. But any feasible solution must spend at least $\delta k^*(|B| - 1)$ fraction of time in satisfying the service requirements of the links in D. In the remaining $(1 + \epsilon) - \delta k^*(|B| - 1) = \delta + \epsilon$ fraction of time the PTAS must use transmission vectors which on the average enable at least $\delta k^*/(\delta+\epsilon)$ links of B^k. Therefore, the PTAS uses a transmission set which constitutes a $\epsilon/(\delta+\epsilon)$ - approximate solution to the instance I of MTS. If we solve $\frac{\epsilon}{\delta+\epsilon} \leq \gamma$ for ϵ, we obtain $\epsilon \leq \frac{\gamma}{1-\gamma}\,\delta$. Replacing δ by its smallest value as k ranges from 1 to $|B|$, i.e. by $\frac{1}{1+|B|(|B|-1)}$, we see that

$$\epsilon \leq \frac{\gamma}{(1-\gamma)(1-|B|+|B|^2)}$$

guarantees a γ - approximate solution to the instance I of MTS.

5.5 Proof of Theorem 5.3

In section 5.4, we have seen that a PTAS for FF can be used to obtain approximate solutions to instances of MTS.

In this section we shall prove that if there is a FPTAS for FF, then we can obtain an optimal solution to any instance of MTS in polynomial time.

Suppose $I = \langle G,B \rangle$ is an instance of MTS. Let k^* denote the optimal value of I.

As in the previous section, we shall transform I to an instance of FF for all values of k between 1 and $|B|$. When $k=k^*$, the FPTAS for FF, which we assume exists, guarantees a solution V(I) such that $\frac{c(V(I))}{k^*} > \frac{\delta}{\delta + \epsilon}$ where $\delta = \frac{1}{1 + k^*(|B| - 1)}$.

We would like to choose ϵ small enough so that $\frac{k^* - 1}{k^*} < \frac{\delta}{\delta + \epsilon}$,
because this would guarantee that $c(V(I)) = k^*$.

Solving for ϵ, we obtain the following condition:

$$\frac{1}{\epsilon} > (k^* - 1)(1 + k^*(|B| - 1)).$$

Since we do not know the value of k^* in advance, we choose

$$\frac{1}{\epsilon} = (|B| - 1)(1 + |B|\ (|B| - 1)) + 1 .$$

We note that $\frac{1}{\epsilon}$ required for obtaining an optimal solution is polynomially bounded in the size of the problem. This completes the proof.

5.6 Negative Results About RF

Theorem 5.4

There exists a PTAS for FF if there exists a PTAS for RF.

There exists an ϵ - PTAA for FF if there exists an ϵ-PTAA for RF.

Unless P = NP, there exists no FPTAS for RF.

The proof of this theorem follows immediately when one considers the transformation algorithm A3 of Chapter IV.

The implications of Theorem 5.4 about the existence of approximation algorithms for the $\vec{r}$-feasibility problem should be clear in view of Theorem 5.2.

CHAPTER VI. CONCLUSIONS AND SUGGESTIONS FOR FURTHER RESEARCH

At the end of this thesis, we can say with some certainty that PRNs are complex; even the most fundamental problems are not likely to be solved in a practical way. For example, if NP $\neq$ P, there exists no polynomial time algorithm for deciding whether a given set of mean o-d arrival rates belongs to the capacity region of a given PRN.

We have proved several negative results about TDMA schemes; let us note as an extension that similar results are readily obtained for FDMA (frequency-division-multi-access) schemes by using a PRN model in which each station is capable of transmitting and receiving simultaneously as many packets as desired, provided, of course, the frequency bands that are used do not overlap. We should remind the reader that the above conclusions are based on worst-case considerations and should be interpreted accordingly.

Our initial interest in studying TDMA schemes in the course of this thesis work was to obtain a benchmark against which the performance of other multi-access schemes could be compared. As is implicit in the proof of Theorem 3.2, no multi-access scheme is superior to TDMA as far as throughput is concerned. However, there are situations where TDMA is undesirable, because it leads to extremely long delays; e.g. a single-receiver PRN with many small users is one such case. When TDMA is undesirable, a multi-access scheme such as Aloha becomes attractive. In this sense TDMA and Aloha are complementary to each other, but this tradeoff still remains to be quantified.

As a second suggestion for research, TDMA schemes can be studied in PRNs where connectivity between the nodes of the network depends on the distance between them.

REFERENCES

[B] B. Bollobás, Graph Theory. New York: Springer-Verlag, 1979.

[GJ] M.R. Garey and D.S. Johnson, Computers and Intractability. San Francisco: W.H. Freeman and Company, 1979.

[KGBK78] R.E. Kahn et al, "Advances in Packet Radio Technology", Proceedings of the IEEE, Vol. 66, No. 11, November 1978.

[L] D.G. Luenberger, Introduction to Dynamic Systems. John Wiley and Sons, Inc., 1979.

[PS] C.H. Papadimitriou and K. Steiglitz, Combinatorial Optimization. Englewood Cliffs, New Jersey 07632: Prentice-Hall, Inc., 1982.

[S79] P.D. Shapiro, "Conditions for Optimal Routing in a Radio Network", B.S. Thesis Report, M.I.T., May 1979.

[SS(1)81] M. Sidi and A. Segall, "Two Interfering Queues in Packet Radio Networks", Technion Report, Haifa, Israel; Dec.1981.

[SS(2)81] M. Sidi and A. Segall, "A Busy-Tone-Multiple-Access-Type Scheme for Packet Radio Networks", M.I.T. Report, LIDS-P-1103, June 1981.

[T] A.S. Tanenbaum, Computer Networks. Englewood Cliffs, New Jersey 07632: Prentice-Hall, Inc., 1981.

[KT] S. Karlin and H.M. Taylor, A First Course In Stochastic Processes. New York: Academic Press, Inc.,1978.