SPREAD: IMPROVING NETWORK SECURITY BY MULTIPATH ROUTING

Wenjing Lou Wei Liu Yuguang Fang

Department of Electrical and Computer Engineering
University of Florida
Gainesville, FL 32611

ABSTRACT

This paper considers the delivery of secret information across insecure networks. A novel end-to-end multipath secure data delivery scheme, Secure Protocol for REliable dAta Delivery (SPREAD), is proposed as a complementary mechanism for the data confidentiality service in the public networks. The idea behind SPREAD is to improve the confidentiality by enforcing the secret sharing principle in the network via multipath routing. With a (T,N) secret sharing scheme, the message to be protected can be divided into N shares such that from any T or more shares, it can easily recover the message, while from any T-1 or less shares, it should be impossible to recover the message. Then using multipath routing, the shares are delivered across the network via multiple independent paths. The destination node reconstructs the original message upon receiving T or more shares. This paper presents the system architecture of the SPREAD scheme, including how to divide the secret message into multiple shares using the secret sharing scheme, how to find the desired multiple secure paths, as well as how to allocate the message shares onto each selected path such that maximum security can be achieved. The discussion on the optimal share allocations reveals that redundant SPREAD scheme is not only more secure but also more error-tolerant and fault-tolerant. The simulation results show that significantly reduced message interception ratio can be achieved by SPREAD.

I. INTRODUCTION

With the emergence and popularity of applications such as world wide E-commence and Virtual Private Network (VPN), more and more important and confidential information are transmitted over the public Internet. There is a more pervasive need to protect the privacy and integrity of the transmitted messages. In the current network security solutions, cryptography provides the basis for the messages secrecy and integrity. However, it highly depends on the key distribution and management scheme, and carefully designed security protocols are required to exploit it. The information transmitted might be tapped by air or by line. With today's super computer, it is also possible to break any encryption algorithm when enough encrypted information has been collected. Thus far no absolute security is guaranteed in the network.

In this paper, we propose a novel scheme, Secure Protocol for REliable dAta Delivery (SPREAD), as a complementary mechanism to enhance the data confidentiality service in public networks. The SPREAD scheme is based on the *secret sharing* and *multipath routing*. The basic idea is as follows. Using a (T,N) threshold secret sharing scheme, we divide a secret message into N shares such that from any T or more shares, we can easily recover the message, while from any $T-1$ or less shares, it is computationally impossible to recover the message. Then using multipath routing, the N shares are delivered to the destination via multiple minimum overlapping (e.g. independent node-disjoint) paths. From network point of view, if a whole message follows a single path to its destination, a hacker can intercept all the necessary information to recover that message at any intermediate node. However, with the SPREAD scheme, the hacker has to compromise a number of nodes on a number of independent paths to obtain at least T shares. Improved network security can be expected from SPREAD.

There are two major parts in the implementation of the SPREAD scheme. The first is how to find the desired paths. The second is that given the available paths, how the shares may be allocated to the paths. In our previous work [1], we have developed a distributed multipath routing protocol that is capable of finding multiple node-disjoint paths for each source-destination pair in a network efficiently. In this paper, we first describe the overall system architecture of the SPREAD scheme; then we discuss the optimal share allocation schemes, which allocates shares onto each path such that the maximum security can be achieved; and finally we justify the feasibility and show the effectiveness of the SPREAD scheme by presenting the simulation results.

A few efforts have been made to improve the network

This work was supported in part by the Office of Naval Research Young Investigator Award under grant N000140210464 and the Office of Naval Research under grant N000140210554.

0-7803-8140-8/03/$17.00 (C) 2003 IEEE.

security by using multipath routing. Yang et al [2] proposed to improve the network security by traffic dispersion. They provided an analytical framework to study and evaluate the security performance provided by multipath traffic dispersion. However, their scheme did not integrate the secret sharing or any other coding scheme. Zhou and Haas [3] proposed to combine secret sharing and multipath routing to improve the availability and security of the certificate authority in a mobile ad hoc network (MANET). Tsirigos and Haas [4] provided an analytical evaluation for multipath routing in a MANET. They used a diversity coding scheme at the source node. The fundamental idea is similar to ours. However, the goal of their scheme is to maximize the number of successfully transmitted packets in an ad hoc network for reliability purpose in the presence of frequent topological changes, while our design goal is to improve the network security. The optimal share allocation we discuss here has not been discussed in the above works.

II. SYSTEM ARCHITECTURE

Several issues need to be addressed for SPREAD scheme in order to achieve the security enhancement. First, how do we apply the secret sharing to divide the secret message into multiple shares? Secondly, how do we find the multiple paths needed for the delivery of shares? And thirdly, how the message shares should be allocated onto each selected path such that the maximum security can be achieved? We discuss these issues briefly in this section.

To better understand the scheme, we give a brief introduction to the threshold secret sharing system, which is used to generate the shares from a message (messages). Details can be found in [5]. Suppose that we have a system secret K and we divide it into N pieces, $S_1, S_2, ..., S_N$, called shares or shadows. Each of N participants of the system, $P_1, P_2, ..., P_N$, holds one share of the secret respectively. The generation of the secret shares guarantees that any less than T participants cannot learn anything about the system secret K, while with an effective algorithm, any T out of N participants can reconstruct the system secret K. This is called a (T,N) *threshold secret sharing scheme* [6]. Secret sharing schemes consist of two algorithms. The first is called the *dealer*, which generates and distributes the shares among the participants. The second is called the *combiner*, which collects shares from the participants and re-computes the secret, i.e., it produces the secret K from any T correct shares. A combiner fails to re-compute the secret if the number of the correct shares is less than T.

For illustration purpose, we take the Shamir's Lagrange interpolating polynomial scheme as an example. The dealer obtains the ith participant's share by evaluating a polynomial of degree $(T-1)$

$$f(x) = (a_0 + a_1 x + ... + a_{T-1} x^{T-1}) \bmod p$$

at $x=i$:

$$S_i = f(i)$$

which is given to the participant P_i, where p is a large prime number greater than any of the coefficients and is made available to both the dealer and the combiner, and the coefficient $a_0 = K$ is the secret while other coefficients $a_1, a_2, ..., a_{T-1}$ are all randomly chosen. Then, at a combiner, once T shares have been obtained, the combiner can reconstruct the original polynomial by Lagrange interpolation. For example, assume that the received T shares are $S_{i1}, S_{i2}, ..., S_{iT}$, the original polynomial $f(x)$ can be recovered by Lagrange interpolation.

$$f(x) = \sum_{j=1}^{T} S_{i_j} \cdot l_{i_j}(x) \bmod p$$

where

$$l_{i_j}(x) = \prod_{k=1, k \neq j}^{T} \frac{x - i_k}{i_j - i_k}$$

Particually, the original secret K can be recovered by calculating $f(0)$. It is known that this equation has a unique solution over the finite field $GF(p)$. Efficient ($O(T \log^2 T)$) algorithms for polynomial evaluation and interpolation have been discussed in [7]. Even the straightforward quadratic algorithms are fast enough for practical implementation.

The SPREAD scheme works as follows: if a source node wants to send a message to a destination node securely, the source can use a multipath routing algorithm to find multiple paths from the source to destination with certain properties (for example, disjoint paths in certain sense), then, depending on the required message security level and the availability of the multiple paths, the source determines a secret sharing scheme, say, (T,N) threshold scheme, to generate the message shares and routes them to the destination through the selected multiple paths. The source will be the dealer in this case. Limited by the size of the chosen prime number p, the dealer will chop a long message into small blocks, which is similar to any block cipher used to encrypt a large message. In addition, depending on the number of paths used, the SPREAD seems to waste a lot of bandwidth. To save the network bandwidth, in SPREAD all the coefficients $a_0, a_1, a_2, ..., a_{T-1}$ can be assigned using message blocks. The destination will be a combiner, upon receiving T shares, it is able to recover the original secure message. Detailed description on how to apply secret sharing on the message can be found in [1].

Figure 1 illustrates the idea of multipath routing of message shares. It is clear that when the message shares

are transmitted over multiple independent paths, the adversary has to compromise more intermediate nodes to intercept enough shares necessary for message recovery. In our previous work, we designed an efficient distributed multipath routing algorithm which is able to find multiple node-disjoint paths from each node to any other node in the network [1]. If we use an appropriate security related cost function in the routing protocol, the protocol proposed in [1] can be easily applied to find the multiple secure paths desired in SPREAD scheme. We assume each node i in the network is associated with a security level q_i, e.g. the probability that the node might be compromised. Notice that this value could be estimated from the feedback of the some widely implemented security monitoring software and hardware such as firewalls and intrusion detection devices, as well as from the administrative decisions. Then for a s-t path $(s,i,j,\ldots,l,t)$, the security of the path (the probability that the path is compromised) would be

$$p = 1-(1-q_i)(1-q_j)\cdots(1-q_l)$$

Here we assume that the source and destination node are safe and we consider the protection of the messages while they are transmitted across the network.

We define the cost function of link between node i and j as

$$c_{ij} = -\log\sqrt{(1-q_i)(1-q_j)}$$

Then the cost of the (s,t) path using shortest path algorithm is

$$\begin{aligned}
&cost(s,t) = c_{si} + c_{ij} + \cdots + c_{lt} \\
&= -\log(1-q_i) - \log(1-q_j) - \cdots - \log(1-q_l) - \log\sqrt{(1-q_s)(1-q_t)} \\
&= -\log\{(1-q_i)(1-q_j)\cdots(1-q_l)\} - \log\sqrt{(1-q_s)(1-q_t)}
\end{aligned}$$

With the shortest path algorithm, node s and t are fixed,

$$\begin{aligned}
&cost(s,t) \text{ is minimized} \\
&\Rightarrow -\log\{(1-q_i)(1-q_j)\cdots(1-q_l)\} \text{ is minimized} \\
&\Rightarrow (1-q_i)(1-q_j)\cdots(1-q_l) \text{ is maximized} \\
&\Rightarrow p = 1-(1-q_i)(1-q_j)\cdots(1-q_l) \text{ is minimized}
\end{aligned}$$

So the path found by the shortest path algorithm would be the most secure path when the proposed cost function is used. With this cost definition, the algorithm proposed in [1] can be directly used to multiple secure paths.

The third issue of SPREAD is how to allocate shares onto each selected path such that the maximum security can be achieved. We develop a redundant SPREAD scheme (where $T<N$), which can achieve the best security at the same time improve the data reliability. This will be discussed in the following section.

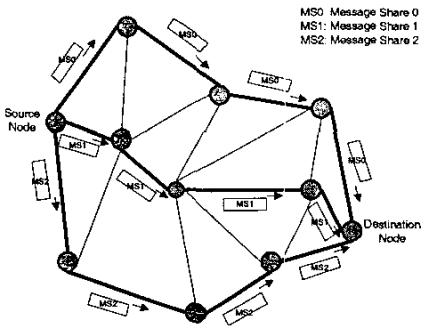


Figure 1 Multipath routing of the message shares

III. SHARE ALLOCATION

How to choose the appropriate values of (T,N) and how to allocate the N shares onto each selected paths is one important issue in SPREAD design. Provided the available paths and their corresponding security characteristics, the first objective is to maximize the message security. The fundamental idea is to allocate the shares in such a way that an adversary has most difficulty to recover the message, e.g. force the adversary to compromise all the paths to recover the message. In this section, we discuss the optimal share allocation schemes which exploit the redundant secret sharing scheme (where $T<N$) to achieve both data confidentiality and reliability.

A. Problem Formulation and Notations

Assume that (T,N) secret sharing algorithm is applied to the message to be protected at source node. In the network layer, we assume that there are totally M node disjoint paths, *path 1, path 2, ..., path M*, available from the source to the destination. We use vector $\underline{p}=[p_1,p_2,\ldots,p_M]$ to denote the security characteristics of the paths, where p_i (i= 1,2,...,M) is the probability that path i might be compromised. Without loss of generality, we further assume $p_1 \le p_2 \le \ldots \le p_M$ which means that the paths are ordered from more secure one to less secure one. Notice that this information is available at source from the multipath routing protocols we proposed in [1]. We assume that if one node were compromised, all the shares traveling through that node would be intercepted. Therefore, we define that a path is compromised as when any one or more of the nodes along the path is compromised. For each path, we consider that if it were compromised, all the shares allocated to it would be intercepted. Otherwise, if the path were not compromised, no share on that path would be lost. As those paths are node disjoint, we further assume that the probability that

one path is compromised is independent of others. As we pointed out in previous section, SPREAD scheme only enhance the data confidentiality statistically when the data are transmitted across the network. Thus the probability p_i does not include the probability that the source or the destination node is compromised, e.g., we assume source and destination are trustworthy.

A share allocation scheme is used to allocate the N shares onto the M available paths. Denote the share allocation as $\underline{n}=[n_1,n_2,...,n_M]$, where n_i is the number of shares allocated to path i, n_i is an integer, $n_i \geq 0$, $\sum_{i=1}^{M} n_i = N$. According to the secret sharing algorithm, the probability that the message is intercepted equals to the probability that T or more shares are intercepted. We denote the probability that the message is interceptd in terms of the share allocation $\underline{n}$ as $P_{msg}(\underline{n})$. Then the share allocation can be formulated to a constrained optimization problem

$$\begin{aligned} &\textit{minimize} \quad P_{msg}(\underline{n}) \\ &\textit{subject to} \quad \sum_{i=1}^{M} n_i = N,\ n_i \text{ is an integer, } n_i \geq 0 \end{aligned}$$

B. Maximum Security without Redundancy

Let define $r = 1 - T/N$ as the redundancy factor of the (T,N) secret sharing scheme. A non-redundant SPREAD scheme is one where $r=0$, e.g. $N=T$. It is easy to derive that given the number of available paths, M, and the corresponding path security characteristics $\underline{p}=[p_1,p_2,...,p_M]$, the non-redundant (N,N) $(N \geq M)$ secret sharing scheme would give the maximum security, e.g. minimum message interception probability, when at least *one* share and at most *T-1* shares are allocated to each of the available paths, i.e.

$$\begin{cases} 1 \leq n_i \leq T-1, & \text{i} = 1,...,\text{m} \\ \sum_{i=1}^{m} n_i = N \end{cases}$$

This share allocation requires the adversary to compromise all the paths to intercept the message. This probability equals to the probability that all the paths are compromised.

$$P_{msg}(\underline{n}) = \prod_{i=1}^{M} p_i$$

It is noted that the security provided only depends on the paths chosen. As p_i is a probability satisfying $0 \leq p_i \leq 1$, the more paths we use to distribute the shares, the less the probability is, and the more secure the message delivered. Thus, given required security level (in terms of message interception ratio) γ_{P_n}, the SPREAD scheme would chose the first m paths, path *1*, path *2*, ..., path m, which satisfying $P_{msg}(\underline{n}) = \prod_{i=1}^{m} p_i \leq \gamma_{P_n}$, to deliver the message.

C. Maximum Security with Redundancy

It is intuitive that non-redundant secret sharing scheme provides the maximum security to the message. However, with our SPREAD scheme, as the shares are spread onto multiple paths, long paths might be used. The reliability of the message, in terms of transmission error, packet lost ratio, etc., may be degraded. In the case that the reliability is also an issue, redundant secret sharing scheme will be desirable.

Redundancy is a common way to improve the reliability. It is based on the idea of sending more information than minimum requirement, so that the original message can be reconstructed in the event of loss in the network. It may be used independently of the multipath routing, by adding Forward Error Correction (FEC) code to each individual share. We denote this type of redundancy as serial redundancy. Serial redundancy is good for correcting noise-like random errors introduced to the bit stream, while helpless for the persistent errors or link failure. With a (T,N) secret sharing scheme, when $T<N$, we actually introduce redundancy from another dimension, the parallel redundancy. When this type of redundancy is used in combination with multipath routing, the system becomes more error tolerant, because a certain number $(N-T)$ of message shares can be corrupted or lost without affecting the reconstruction of the original message. The system also becomes more fault-tolerant because a certain fraction of the paths can be affected by failure without interrupting the flow of the information.

Using the same path chosen criteria i.e. chose the first m most secure paths which satisfy the required security level, it is intuitive to show that, in order to achieve the maximum security, the total number of shares allocated to any *m-1* or less paths should be less than T. Again, this share allocation forces the adversary to compromise all the m paths to intercept the message. This is also a necessary and sufficient condition to achieve the maximum security. It can be simplified as

$$\begin{cases} N - n_i < T, & \forall i \in (1,2,\cdots,m) \\ n_1 + n_2 + \cdots + n_m = N \end{cases}$$

Remember $r = 1 - T/N$ is the redundancy factor of the secret sharing scheme. Then we could derive a necessary condition for achieving the maximum security, i.e.

$$r < 1/m \quad (m \geq 2)$$

This is a useful condition as it defines the maximum

redundancy we can add to the SPREAD scheme without sacrificing the security. We could claim that to maintain the maximum security achievable from the chosen path set, the maximum redundancy we can add to the secret sharing algorithm is bounded by $r < 1/m$, where m is the number of chosen paths $(m \geq 2)$. In other words, we could claim that for a r-redundancy SPREAD scheme, the maximum security can be achieved only if the redundancy factor r satisfies $r < 1/m$ $(m \geq 2)$. Then by choosing an appropriate (T,N) value which satisfies

$$T \geq N\frac{m-1}{m}+1 \qquad (m \geq 2)$$

an optimal share allocation can be designed such that the maximum security can be achieved while at the same time certain (r) redundancy can be provided. Any allocation that conforms to the constraints

$$\begin{cases} N-T+1 \leq n_i \leq T-1, & i = 1,\ldots,m \\ \sum_{i=1}^{M} n_i = N \end{cases}$$

is an optimal share allocation in terms of security. The optimal share allocation is not unique. Other optimization objectives, such as the minimal delivery cost, balanced bandwidth usage, or maximum reliability, might be set to further optimize the share allocation for other purposes.

IV. SIMULATION RESULTS

In this section we present the simulation results to show the effectiveness of the SPREAD scheme in terms of the enhancement to the data confidentiality. The multipath routing algorithm proposed in [1] is used to find the desired multiple node-disjoint paths. Five types of 20-node networks, with the node degree equals to 4, 5, 6, 7, 8 respectively, are evaluated. For each type of network, the simulation results are averaged over 20 random networks which are generated using a random graph generator based on Waxman's generator [8]. Two sets of simulation are executed. In the first set, each node is assumed to be independently and equally likely be compromised with probability 0.152. In the second set, we assume nodes with different probability be compromised. The probability that a node might be compromised is selected from 4 values: 10% of nodes with probability 0.50 being compromised, 30% of nodes with probability 0.20, 40% of nodes with probability 0.10, and 20% of nodes with probability 0.01. For both sets, we use the proposed link cost function to define the link cost based on the node security level so that the most secure paths are selected. For each network, we run the multipath routing algorithm to find the maximal number of most secure paths between any source-destination pair which is not directly connected. Then we calculate the message interception ratio when different number of paths $(m=1,2,3,4,5)$ is used.

Table 1 summarizes some parameters of the simulated networks, including the node degree and the average network diameter.

Table 1 Network Parameters

Number of Nodes	20	20	20	20	20
Node Degree	4	5	6	7	8
Average Diameter	4.25	3.80	3.40	3.15	3.00

Figure 2 shows the probability that multiple node disjoint paths are found by the distributed multipath routing algorithm in the simulated networks. We observe that for a well connected network, e.g. node degree equal to or larger than 4, the chances to find 2 or more paths are pretty high, in both simulation sets. Since our SPREAD scheme depends on the availability of multiple paths, the existence of such multiple paths and the capability of finding such paths justify the feasibility of our SPREAD scheme.

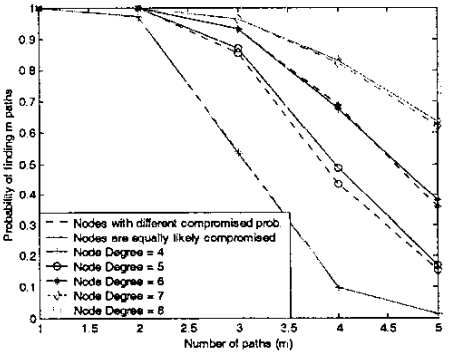
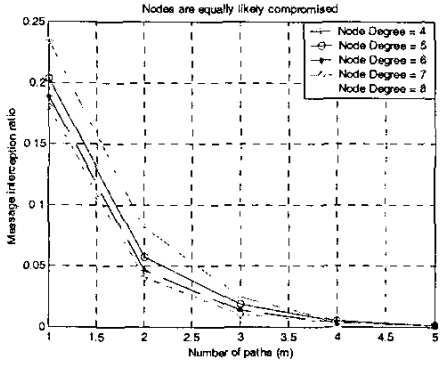


Figure 2 Probability of finding multiple paths

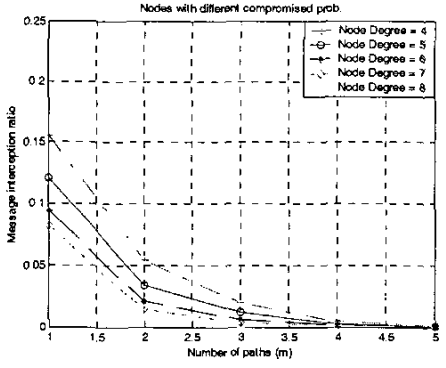
Figure 3 shows the probability that the message might be intercepted when different number of paths is used. Figure 3(a) is for simulation set 1, i.e. each node is equally likely compromised. Figure 3(b) is for the simulation set 2 in which nodes have different probability being compromised. It is clear that, in both sets, the message interception ratio drops significantly, actually exponentially, with the increasing of the number of paths used. This result verifies the effectiveness of our SPREAD idea. Notice that the message interception ratio here is the additional security achieved on top of the cryptographic scheme. The adversary still needs to decrypt the message

after intercepting all the necessary shares.

The average node compromised probability in simulation set 1 and 2 is made equal. However, we observe significant differences in the achieved message interception ratios in figure 3(a) and 3(b). The simulation set 2 actually achieves better security than simulation set 1. This is because when nodes are equally likely compromised, the multipath routing algorithm actually finds the minimum hop paths. While when nodes have different security levels, by incorporating the security link cost function proposed in this paper, the multipath algorithm will find paths according to their security levels. Those paths are more secure thus achieve better security. This trend is clearer in the networks where node degree is higher, which basically provide more choices from which the routing algorithm can select more secure paths.



(a) Nodes are equally likely compromised



(b) Nodes are compromised with different probabilities

Figure 3 Message interception probabilities

V. CONCLUSION

In this paper, we propose a novel scheme, SPREAD, as a complementary mechanism to secure data delivery across insecure networks. The basic idea and system architecture of SPREAD scheme is presented in this paper. The effect of share allocations on the security and reliability is discussed. The optimal share allocation scheme is discussed for redundant SPREAD scheme which could maintain the maximum security while at the same time provide a certain degree of redundancy for reliability purpose. Finally using simulation, we show the effectiveness of the SPREAD scheme - the message interception ratio during the transmission is significantly reduced because the secure information is distributed among several independent paths.

A few remarks are in order. First, the SPREAD scheme considers the confidentiality enhancement when messages are transmitted across the network, assuming the source and destination are trusted. The protection of each particular node is a separated issue and is out the scope of this paper. Secondly, the SPREAD scheme cannot address the confidentiality alone, it could be built on top of any cryptographic scheme, it only statistically enhances such service. For example, it is still possible for adversaries to intercept all the shares, e.g. by collusion. In this case, encryption algorithm might be used to further protect the message. Due to the salient feature of the secret sharing, only part of the message shares $(N-T+1)$ need to be encrypted to achieve the security from encryption. Finally, the SPREAD can be made adaptive in the sense that the source node could make final decision whether a message is delivered at certain time instant according to the security level and the availability of multiple paths. Moreover, the chosen set of multiple paths may be changed from time to time to avoid any potential capture of those multiple paths by adversaries. More variation of SPREAD scheme is under active research.

REFERENCE

[1] W. Lou, Y. Fang, "A multipath routing approach for secure data delivery", *IEEE Milcom '01*, Oct 2001

[2] J. Yang, S. Papavassiliou, "Improving network security by multipath traffic dispersion", *IEEE Milcom '01*, Oct 2001

[3] L. Zhou, Z.J. Haas, "Securing ad hoc networks", *IEEE Network Magazine*, Nov 1999

[4] A. Tsirigos, Z.J. Haas, "Multipath routing in the presence of frequent topological changes", *IEEE Communication Magazine*, Nov 2001

[5] G. J. Simmons, "An Introduction to Shared Secret and/or Shared Control Schemes and The Application", in *Contemporary Cryptology: The Science of Information Integrity*, IEEE Press, 1992, pp.441-497

[6] A. Shamir, "How to Share a Secret", *Communications of the ACM*, 22(11):612-613, Nov 1979

[7] T. Cormen, C. Leiserson, R. Rivest, *Introduction to algorithms*, MIT Press, 1990

[8] B. M. Waxman, "Routing of multipoint connections," *IEEE Journal on Selected Areas in Communications*, 6(9):1617-1622, Dec. 1988