Alarm Correlation for Congestion Diagnosis in ATM Networks

Ying-Dar Lin, Ren-Kuei Yang, Chi-Chun Lo

Institute of Computer and Information Science
National Chiao Tung University
Hsinchu, Taiwan, Republic of China

Tel: 886-035-715900
Fax: 886-035-721490
Email: ydlin@cis.nctu.edu.tw

ABSTRACT

In this paper, we study OAM measurement and alarm correlation methods in ATM networks. OAM fault management and performance management mechanisms are reviewed. We present the path intersection heuristics as the alarm correlation method to analyze the alarms collected by OAM measurements, and to find the congestion areas in ATM networks. We use simulation to evaluate the proposed heuristics. Besides, we analyze the congestion pattern in ATM networks by examing the simulation results. Finally, based on the proposed heuristics, we present three methods to choose the VPCs for rerouting in order to eliminate the congestion and we make a comparison among these three methods.

1. Introduction

Asynchronous transfer mode (ATM) is a promising concept in realizing broadband ISDN. A great deal of effort has been made by CCITT [1], ATM Forum [2] to specify and standardize operation functions for managing high-speed, high-performance ATM networks. Even though high-performance ATM-based networks will, at times, experience failures, congestion, and periods of degraded performance; and without the appropriate management tools to detect and overcome these problems, the viability of the broadband network would be in danger [3]. Network management system becomes an essential component which helps network managers to maintain the normal operation.

This paper concentrates on the congestion diagnosis problem. To find out the congestion areas in ATM networks, we first need to measure the status in a ATM network. Theoretically, we may conduct a fine-grade measurement to monitor the status of each small component and receive alarms with very specific information. However, the amount of measurements and alarms would be overwhelming, which means the processing load for the management system, both on the manager and agents, would be much increased. Thus we may want a coarse-grade measurement to limit the amount of measurements and alarms. This. in turn, may induce the problems of non-specific alarms. Here an alarm correlation [4,5] is required to further diagnose the problem.

In this paper, we present two alarm correlation heuristics to find the congestion areas in ATM networks. The method is based on a limited amount of OAM (Operation and Maintenance) [6,7,8] measurements, i.e., coarse-grade OAM measurement. By monitoring each VPC, network manager can detect all performance degraded VPCs. By intersecting the paths of these performance degraded VPCs, we suppose the intersection area as the congestion area and we want to find out the congested nodes in the congestion area. One of the heuristics is pure intersection: all the nodes, which are passed by more than one performance degraded VPC in the intersection area, are suspected congested nodes. Another heuristics is intersection with output link consideration: Only the nodes, which are passed by more than one performance degraded VPC and these congested VPCs have the same output port on the node, are suspected congested nodes. We use a simulation tool to evaluate these two heuristics. This is achieved by comparing the set of suspected congested nodes obtained from the heuristics with the set of real congested nodes which are measured in the simulated network. Besides, according to the data obtained from the simulation study, we may analyze the congestion pattern in the ATM networks to see that will the real congested nodes belong to one area or distribute at some areas? In addition, based on the heuristics, we present three methods to choose some performance degraded VPCs for rerouting in order to eliminate the congestion.

In section 2, we describe the congestion diagnosis problem. In section 3, we present our proposed alarm correlation heuristics. Section 4 gives the simulation results and simulation study. Finally, we conclude our research and suggest possible future work in section 5.

2. Problem Description

There are two approaches to solve the congestion diagnosis problem: fine-grade OAM measurement and coarse-grade OAM measurement. These two approaches are illustrated in Figure 1. The fine-grade OAM measurement approach monitors and reports the cell loss and cell transfer delay of each link for each VPC such that this approach may find out the congested nodes specifically. The coarse-grade OAM measurement can only monitor and report the end-to-end cell loss and cell transfer delay for each VPC such that this approach may not find out the congested nodes specifically. Thus, this approach need a heuristics alarm correlation method to locate the congested nodes.

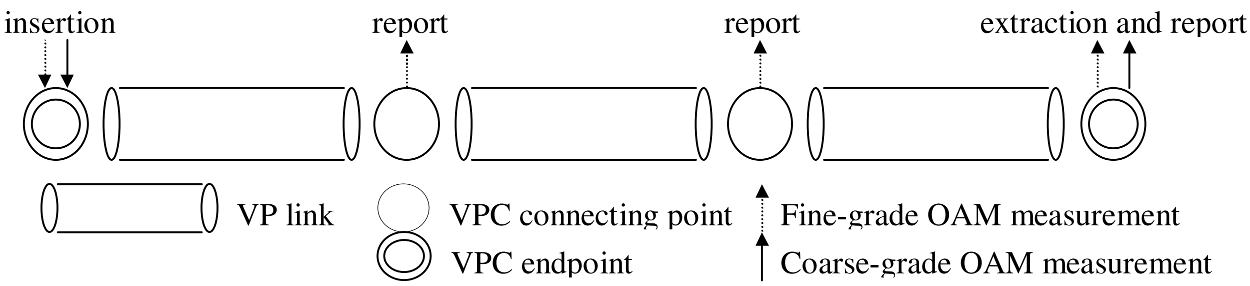


Figure 1. Fine-grade OAM measurement vs. coarse-grade OAM measurement

We choose the coarse-grade OAM measurement approach to solve the congestion diagnosis problem because that we want to limit the amount of OAM measurements, that is, to reduce the overload of agents. In addition, we propose intersection heuristics to find out the congested nodes which result in the performance degradation of VPCs.

To monitor the performance parameters [9] , one scheme being considered makes use of the periodically inserted OAM monitoring cells in the stream of user cells per VPC [8,10]. This scheme is described in the next section.

3. Heuristics for Congestion Diagnosis

In this section we present the heuristics to find the congestion areas in ATM networks based on the OAM performance monitoring mechanism. We describe our approach in the following five steps:

- OAM cells insertion at VPC source: OAM monitoring cells, carrying performance information, are inserted periodically at the source VPC endpoint for each VPC.
- Cell loss ratio calculation at VPC destination: these OAM monitoring cells are extracted at destination VPC endpoint and the performance information in the monitoring cells is retrieved to count the cell loss ratio.
- Congested VPCs Identification at the manager: we predefine the threshold of cell loss ratio for VPCs. If the cell loss ratio obtained from OAM measurement is higher than the threshold of cell loss ratio, the destination VPC endpoint sends an alarm carrying a congestion notification to the network manager and this VPC is identified as a congested VPC.
- Congested area diagnosis at the manager: when the network manager receives an alarm that carries a congestion notification, it begins to collect alarms for a given time period. Then the network manager marks all the congested VPCs indicated in the collected alarms. By intersecting the paths of these congested VPCs, we regard the intersection area as the congestion area.
- Congested area rerouting: now we know which are congested VPCs and where is the congestion area. We want to choose some congested VPCs to reroute out of the congestion area in order to eliminate the congestion. Three proposed methods are described in the next chapter.

The congestion area may contain exactly one node or several nodes. We define the nodes in the congestion area as congested nodes. To point out the congested nodes exactly, we propose two heuristics associated with the intersection of congested

VPCs. The heuristics are described below.

Heuristics 1: Pure Intersection

The nodes in the intersection area of all congested VPCs may be the sources which result in the congestion. We define degree of node *i* as the number of congested VPCs that pass node *i*. The network manager finds all the nodes which are passed by more than one congested VPC in the intersection area of all congested VPCs. That is, the network manager collects those nodes whose degrees are larger than one. These nodes are regarded as congested nodes. This is shown in Eq (3.1).

$$\{\text{congested nodes}\} = \bigcap_{i \in \{\text{congested VPCs}\}} \text{Path}(\text{VPC}i) \tag{3.1}$$

Heuristics 2: Intersection with Output Link Consideration

According to the characteristics of output queueing in ATM switches, VPCs will only congest if they compete the same output link in a switch. The congestion will not happen, however, at the nodes that no VPC competes the same output link. We propose another heuristics that only the nodes with more than one congested VPC competing the same output link are the sources which result in the congestion. That is, if a node is passed by more than one congested VPC and these congested VPCs have the same output port on the node, this node is included. These nodes will be regarded as the congested nodes.

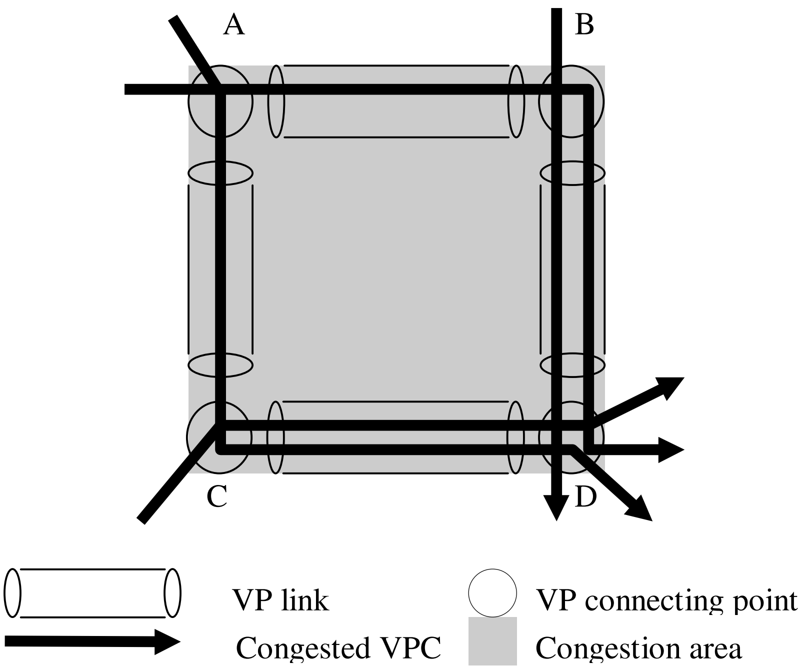


Figure 2. Congested VPCs vs. congested nodes: an example

Figure 2 shows an example of congestion area and how the heuristics work. We can see that there are four congested VPCs and these four congested VPCs construct a congestion area which contains four node A, B, C, and D. According to Heuristics 1, these four nodes in the congestion area are all congested nodes and their degrees are 2, 3, 2, and 3 respectively. According to Heuristics 2, B and C are congested nodes but A and D are not. This is because that there are more than one congested VPC competing the same output link on B and C, but no such condition happens at A and D.

In order to evaluate the above two heuristics, we need to compare the set of congested nodes obtained from the heuristics with the set of real congested nodes, which are measured in a network environment. We define hit ratio and miss ratio below:

H (by heuristics) = {suspected congested nodes obtained from the heuristics}

M (by measurement) = {real congested nodes measured in a network}

$$\text{Hit ratio} = \frac{H \cap M}{H}$$

$$\text{Miss ratio} = \frac{M - H}{M}$$

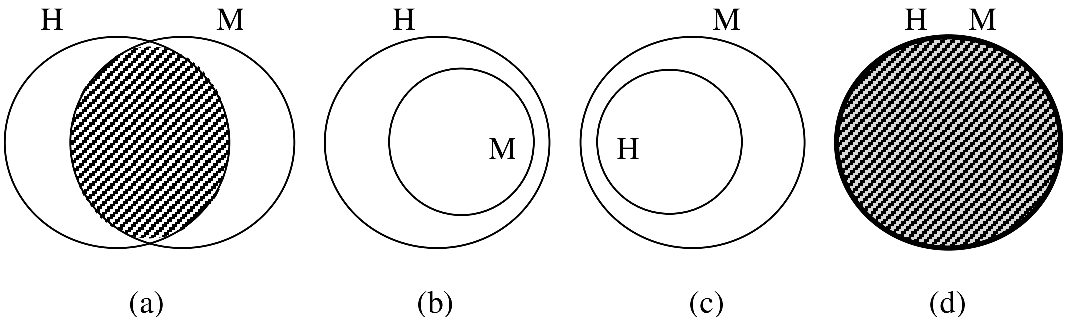


Figure 3. Relations between H and M

We illustrate the relations between H and M in Figure 3. Figure 3(a) shows a general condition. If the hit ratio is high and the miss ratio is low, we may say that the heuristics is good. Figure 3(b) shows the condition of miss ratio=0 and hit ratio<1; all real congested nodes are captured but some suspected congested nodes are not real congested nodes, i.e., the heuristics is not precise enough. Figure 3(c) shows the condition of hit ratio=1 and miss ratio>0; although all suspected congested nodes are real congested nodes, the heuristics doesn't capture all real congested nodes. Figure 3(d) shows the condition of hit ratio=1 and miss ratio=0, i.e., the best condition; the set of suspected congested nodes obtained from the heuristics exactly matches the set of real congested nodes. We would like to find out how these two heuristics perform by a simulation study.

Since the proposed heuristic alarm correlation for congestion alarm diagnosis doesn't need to monitor and report the performance at every VPC connecting point, it has the following advantages: (1) reduce the mount of OAM measurements, i.e., reduce the processing load of agents; (2) reduce the number of alarms, i.e., reduce the processing load of manager. These load produced by the management are overheads for a system. To minimize the overheads, we choose the proposed heuristic alarm correlation approach, which needs less processing load. Based on the heuristics, we can find all congested VPCs and congested nodes. We may choose some congested VPCs to reroute such that the congestion in the ATM network can be eliminated. The proposed choosing methods are described in section 4.4.

4. Simulation Study

In this section, we introduce a simulation tool: Broadband ISDN Simulation Testbed and the simulation cases used on the simulation. According to the simulation cases, we also present collected simulation results and analysis of the simulation results.

In this simulation study, we address the following three issues:

- Examine the effectiveness of the proposed heuristics
- Examine the congestion pattern in ATM networks
- Rerouting based on the proposed heuristics

We present these three issues in section 4.2~4.4.

4.1 Broadband ISDN Simulation Testbed

We use the Broadband ISDN Simulation Testbed which is designed by the NTUEE and NCUEE B-ISDN Research Group [11,12], sponsored by National Science Council to evaluate our heuristics. This simulation testbed provides capability for end-to-end performance evaluations, cell level simulation, and statistical measurement such that we can measure the cell loss ratio for each connection and node. According to the measured cell loss ratio and predefined threshold of cell loss ratio, we may apply our heuristics to find the suspected congested nodes. Besides, we may collect the real congested nodes in the simulated network by examing the statistical measurement of switches. Then we may calculate the hit ratio and miss ratio to evaluate our heuristics.

To test the heuristics, we need some network topologies for input data of the simulation testbed. We wrote a program to construct the network topologies randomly and build up connections manually. These topologies contain 8, 12, 16, 20 nodes and 20, 30, 40, 60 physical links respectively. We describe the algorithm used to construct the network topologies in the following: In the beginning, all nodes are in the unconnected set. We choose two nodes randomly from the unconnected set to connect and put them in the connected set. Then we iteratively choose one node from the connected node and one node from the unconnected set to connect and put the node from the unconnected set in the connected set till all nodes are in the connected set. At last, we iteratively choose two nodes randomly from the connected set to connect to fulfill the number of links.

We will examine how the heuristics works to answer the following questions:

- How does the heuristics perform if we alternate the traffic load from low to high, where all connections are of the same traffic load?
- How does the heuristics perform in different network topologies which contain different numbers of nodes?
- How does the heuristics perform under different types of traffic sources?
- How does the heuristics perform in unbalanced traffic comparing with balanced traffic, where balanced traffic means the traffic load of the connections are all the same and unbalanced traffic means the traffic load of the connections are not all the same.

We present the simulation results and analysis of the results in the next section.

4.2 Effectiveness of the Proposed Heuristics

We now present some results on hit ratio and miss ratio of the heuristics. For the topologies, we have five sizes: 8, 12, 16, and 20 nodes. For traffic sources, we have Bernoulli, DOO, and GOO sources. For Bernoulli sources, we use the parameters of traffic load: mean rate=0.4, 0.45, 0.5, 0.55, 0.6, 0.65, and 0.7. For Geometric On/Off sources, we use the parameters of traffic load: peak rate=0.8, 0.9, 1.0, 1.1, 1.2, and 1.3; Probability_of_On_to_Off=0.5, Probability_of_Off_to_On=0.5. For Deterministic On/Off sources, we use the parameters of traffic load: peak rate=0.5, 0.6, 0.7, 0.8, 0.9, and 1.0; On_Time=500, Off_Time=500. The simulation results is shown in Figure 5 - Figure 12. According to the simulation results, we obtain the following points.

- Miss ratio of heuristics 1 and heuristics 2 in any condition is 0. That is, the heuristics can capture all real congested nodes in the simulation network. All real congested nodes are also suspected congested nodes. But hit ratio may range from 0.5 to 1. This implies that the result of the heuristics contains more nodes than the set of real congested nodes. In other words, some suspected congested nodes are not real congested nodes.
- Hit ratio of the heuristics for Bernoulli sources and GOO sources are almost the same. Perhaps it's because we set the parameters of probability_of_on_to_off and probability_of_off_to_on for GOO sources equal to 0.5. The distribution of GOO sources and the distribution of Bernoulli sources are very alike.
- Hit ratio of the heuristics for Bernoulli sources is higher than DOO sources. That is, the heuristics works better when traffic sources are Bernoulli sources. Perhaps it's because the DOO sources are more bursty than the Bernoulli sources. This is shown in Figure 5, 6; Figure 7, 8; Figure 9, 10; Figure 11, 12.
- Heuristics 2 is always better than heuristics 1. This is obvious by observing Figure 5 - Figure 12. The hit ratio of heuristics 2 is always higher than or equal to heuristics 1. This is because heuristics 2 considers the characteristics of output queueing in ATM switches so that it may hit real congested nodes more precisely.
- Higher traffic load leads to higher hit ratio for heuristics 1. For heuristics 2, when the traffic load exceeds a certain value, higher traffic load also leads

to higher hit ratio. This is because when the traffic load exceeds a certain value, the congested VPCs are fixed such that the number of suspected congested nodes is fixed. And the number of real congested nodes goes up as the traffic load goes up.

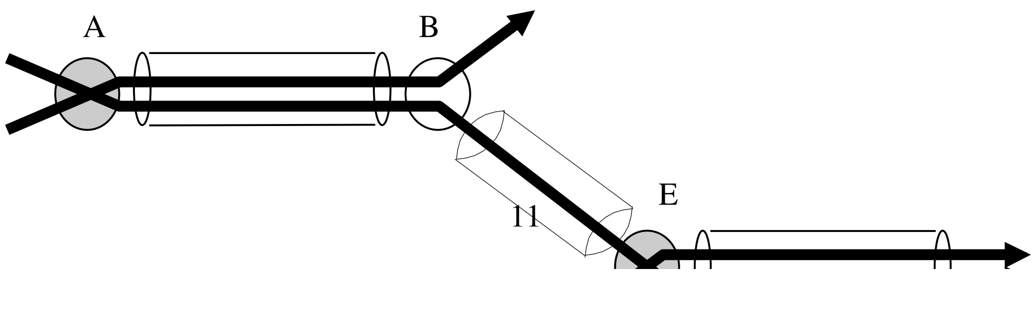
- The influence of the network size for the heuristics is inevident. This is because that the effectiveness of the heuristics is according to the number of congested VPCs and the number of real congested nodes. The network size do not have any evident relation with them.

We summarize the results according to the questions in section 4.1. This is listed in Table 1.

Table 1. Summarized results for heuristics 1 and 2

	traffic load	network size	types of traffic sources	unbalanced vs. balanced traffic
Heuristics 1	higher traffic load leads to higher hit ratio	no evident influence	hit ratio for Bernoulli sources≥hit ratio for DOO	no evident influence
Heuristics 2	when exceeds a certain value, higher traffic load leads to higher hit ratio	no evident influence	hit ratio for Bernoulli sources≥hit ratio for DOO sources	no evident influence

Besides, we can find that the curve of hit ratio of heuristics 2 may be lower at middle traffic load in some cases. As shown in Figure 4, heuristics 2 regard node A, C, E as congested nodes but measured data in the simulated network may show that only node A and C are real congested nodes. This is because that severe cell loss may only happen at node A and C, which reduces the traffic load of the congested VPCs on nodes B, D, and E and results in very low or no cell loss at node E. Due to this condition, heuristics 2 may perform badly if the connection path is very long.



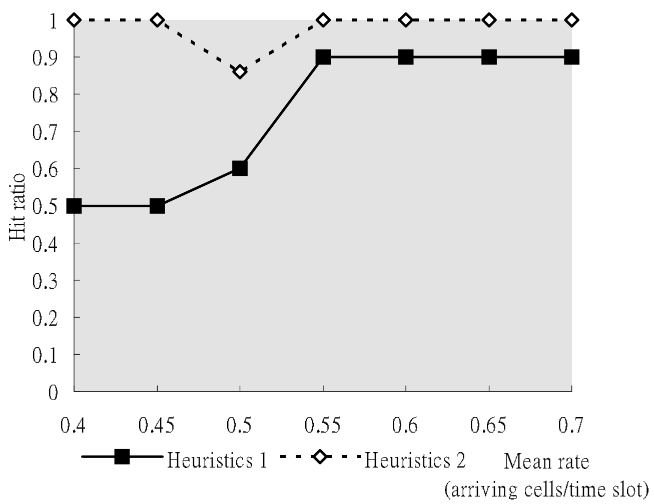


Figure 5. Hit ratio for 20 nodes with uniform Bernoulli traffic

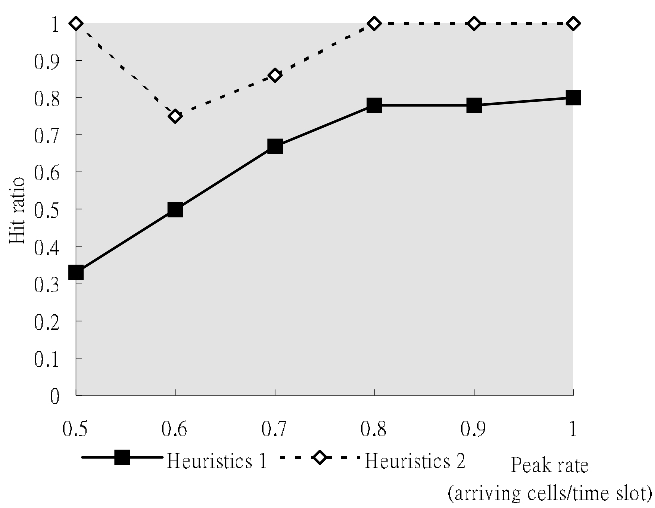


Figure 6. Hit ratio for 20 nodes with uniform DOO traffic

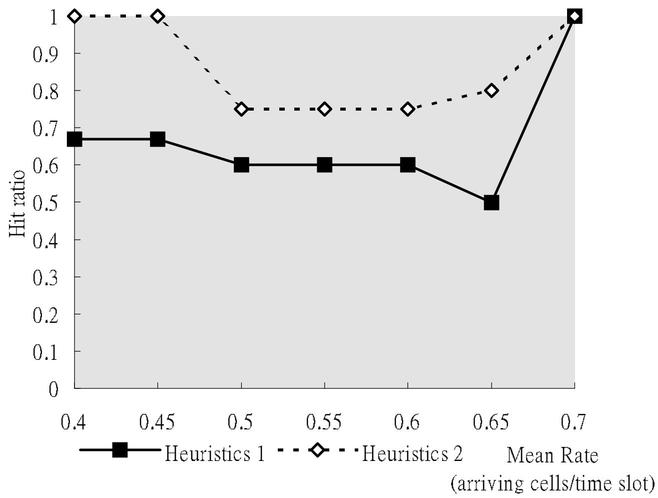


Figure 7. Hit ratio for 16 nodes with uniform Bernoulli traffic

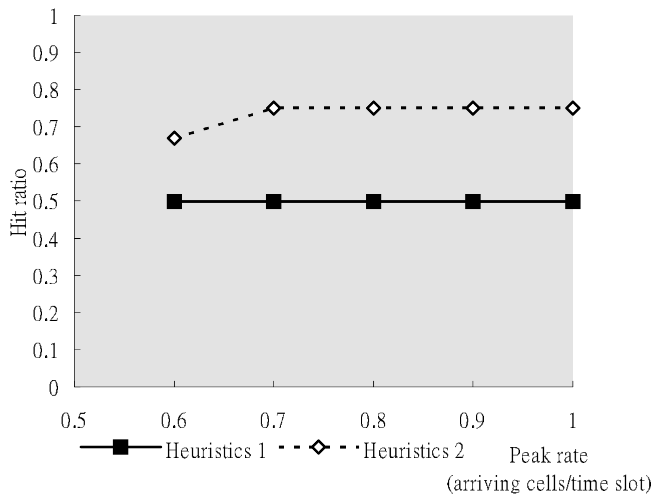


Figure 8. Hit ratio for 16 nodes with uniform DOO traffic

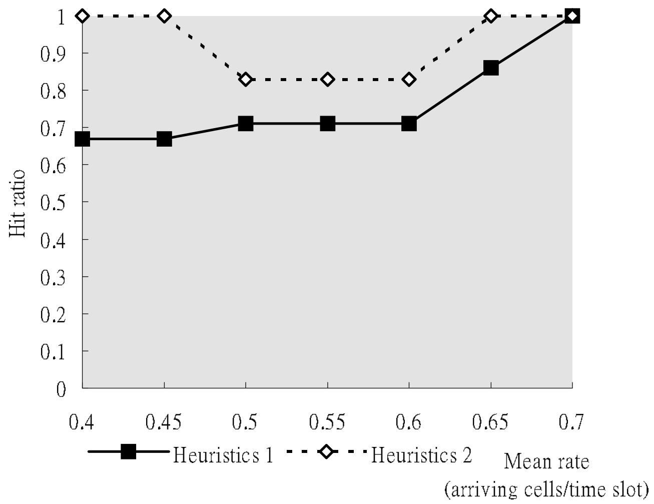


Figure 9. Hit ratio for 12 nodes with uniform Bernoulli traffic

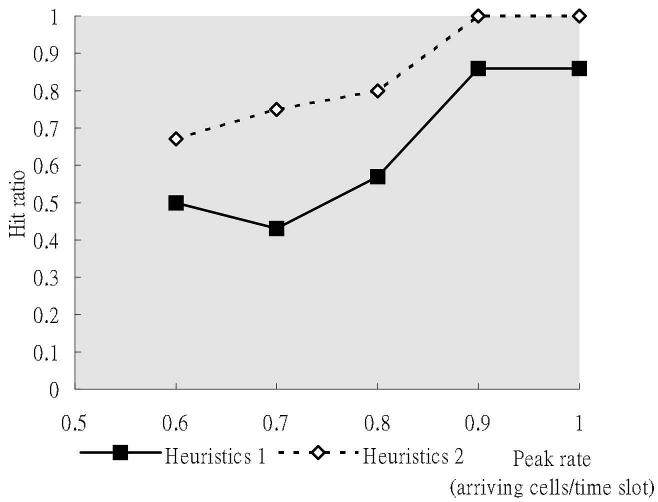


Figure 10. Hit ratio for 12 nodes with uniform DOO traffic

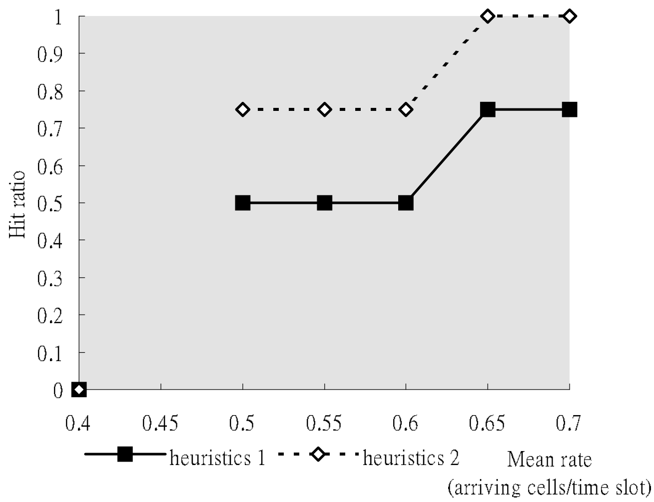


Figure 11. Hit ratio for 8 nodes with uniform Bernoulli traffic

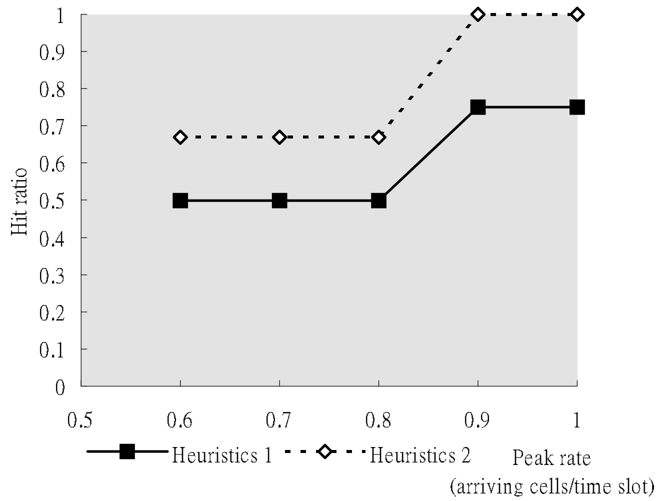
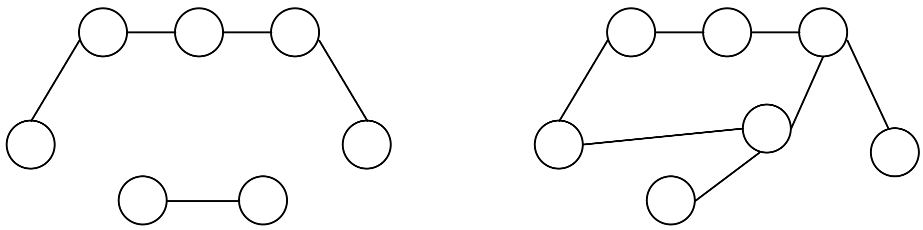


Figure 12. Hit ratio for 8 nodes with uniform DOO traffic

4.3 Congestion Pattern

According to the data obtained from the simulation study, we may analyze the congestion pattern in the ATM networks. We have collected the results of five cases of unbalanced traffic, i.e., different traffic load distributions for the connections, with Bernoulli sources and DOO sources. We regard the real congested nodes as belonging to an *area* if these nodes and their links form a connected graph. Figure 13 shows an example of one congestion area and two congestion areas formed by seven congested nodes (case 3 and case 2 for 20 nodes in Table 2).



(a) (b)

Figure 13. One congestion area vs. two congestion areas: an example

The results of congestion areas are listed in Table 2 and Table 3. For five cases in table 2 and table 3, we change the traffic load of each connection and set the values randomly such that the traffic load for each connection in each case are different.

We can see that in most cases, the number of congestion areas is 1. From this observation, we conclude that the congestion area has the tendency of extending from one node to its neighbors. This implies that the congestion area may be eliminated if the congestion at the bottleneck node can be cleared. In the next section, we investigate the effective ways to eliminate the congestion by rerouting VPCs passing through the bottleneck node.

Table 2. Congestion areas for Bernoulli source

	number of congestion areas/number of congested nodes				
nodes	case 1	case 2	case 3	case 4	case 5
6	-	-	1/2	1/1	-
8	1/2	1/3	1/3	-	-
12	2/3	1/5	1/4	1/2	1/2
16	1/4	1/4	1/5	1/2	1/2
20	1/6	2/7	1/7	1/4	1/2

Table 3. Congestion areas for DOO source

	number of congestion areas/number of congested nodes				
nodes	case 1	case 2	case 3	case 4	case 5
6	-	-	1/2	-	-
8	1/1	1/3	1/3	1/1	-
12	1/2	1/7	1/4	1/4	1/2
16	1/2	1/3	-	-	1/2
20	1/4	1/6	1/6	1/4	2/3

4.4 Rerouting based on the Proposed Heuristics

By utilizing our heuristics, we may choose some congested VPCs passing the congested node to reroute such that the congestion will be eliminated. We present three methods to pick out those VPCs for rerouting. In the simulation, we do not reroute the VPCs but choose which VPCs to reroute.

Large Capacity First (LCF): For every congested node, remove the congested VPC with the largest capacity, then remove the congested VPC with the second large capacity, and the same process goes on until the node is not congested.

Small Capacity First (SCF): For every congested node, remove the congested VPC with the smallest capacity, then remove the congested VPC with the second smallest capacity, and the same process goes on until the node is not congested.

Less Summed Capacity (LSC): In some cases, a congested VPC may pass through more than one congested node. By comparing the capacity of this specified VPC and the sum of the smallest capacities from each congested nodes that the specified VPC passes, we choose the one which needs to remove less capacities.

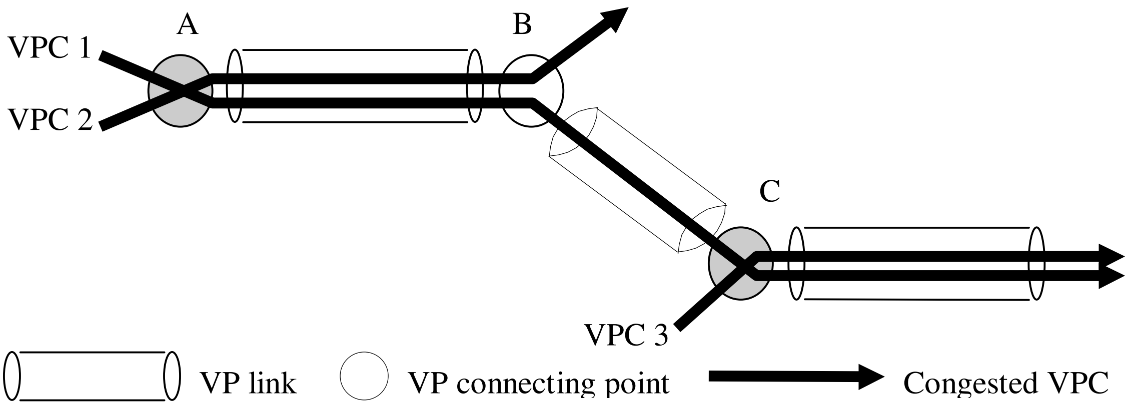


Figure 14. An example of congested VPCs which need to be rerouted

As shown in Figure 14, there are three congested VPCs: VPC 1, VPC 2, VPC 3 and two congested node A, C, where congested node A is caused by VPC 1 and VPC 2, and congested node C is caused by VPC 1 and VPC 3. We assume the following two conditions: (a) capacity of VPC 1 = 0.9, capacity of VPC 2 = 0.1, capacity of VPC 3 = 0.2; (b) capacity of VPC 1 = 0.5, capacity of VPC 2 = 0.3, capacity of VPC 3 = 0.4. We now apply three proposed methods to remove congested VPCs. LCF removes VPC 1 in both conditions; SCF removes VPC 2 and VPC 3 in both conditions; Since VPC 1 passes through two congested nodes A, C, and VPC 2 is the VPC with smallest capacity on node A and VPC 3 is the VPC with smallest capacity on node C, LSC compares the capacity of VPC 1 with the sum of the capacities of VPC 2 and VPC 3. In condition (a), LSC removes VPC 2 and VPC 3 because the sum of the capacities of VPC 2 and VPC 3 is smaller; In condition (b), LSC removes VPC 1 because the capacity of VPC 1 is smaller. In this example, average capacities of removed VPCs for three methods are: LCF = 0.7, SCF = 0.5, LSC = 0.4, where LSC needs to remove the least VPC capacities.

We have also collected some simulation results of average traffic load of removed connections for three proposed methods, which are shown in Figure 15. We can see that LSC removes the connections of the least traffic load on the average and LCF removes the connections of the most traffic load on the average. This is because LSC would not remove more capacities than SCF by comparing the capacity of the congested VPC, which passes several congested nodes, with the sum of the capacities of the smallest VPCs which pass through these nodes. And usually, LCF needs to remove more capacities than SCF and LCF needs to remove very large capacities in the worst case. Thus the capacities needed to be removed for three methods are ordered as: LSC≤SCF≤LCF. That is, LSC is the best among these three proposed methods to choose the congested VPCs for rerouting and LCF is the worst.

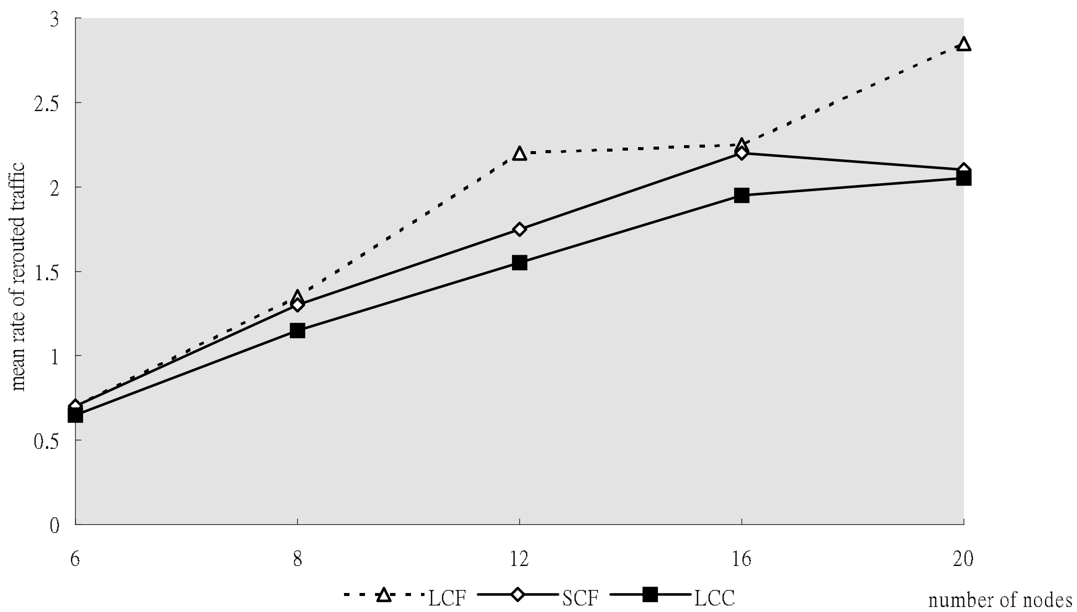


Figure 15. Average traffic load of rerouted VPCs

5. Conclusion and Future Work

In this paper, we have reviewed the OAM measurement methods and provided a survey on fault and performance management mechanisms of ATM layer.

Based on the OAM measurements, we propose two heuristics to find the congestion areas in ATM networks and we evaluate the heuristics by simulation. The simulation results show that the heuristics of intersection with output link consideration may locate the congestion nodes more precisely. Also, according to the simulation results, we analyze the congestion pattern and find that the congestion area has the tendency of extending from one node to its neighbors. In addition, to utilize our heuristics on rerouting, we propose three methods to choose the congested VPC for rerouting in order to eliminate the congestion and the result shows that the method of Less Summed Capacity is the best among these three methods.

We have evaluated our alarm correlation heuristics by simulation and found that the heuristics is applicable to locate the congested nodes in ATM networks. In the future, we plan to implement the heuristics to the network management systems in practical ATM networks. This includes two parts: (1) The agents measure the performance for each component (i.e., switch or virtual connection) in the ATM network by OAM measurement methods. When the performance degradation occurs, it sends an alarm to the manager. (2) The manager collects the alarms and correlates them to localize the source of congestion. We plan to evaluate the proposed heuristics in a real system and compare the tradeoff between the performance and overhead of fine-grade with coarse-grade OAM measurements.

References

[1] CCITT, "OAM Principles of the B-ISDN Access." CCITT Recommendation I.610, June 1992.

[2] ATM Forum, "ATM User-Network Interface Specification", Version 3.0, ATM Forum, October 1993.

[3] Stephen C. Farkouh, "Managing ATM-based Broadband Networks", IEEE Communications Magazine, pp. 82-86, May 1993.

[4] Gabriel Jakobson and Mark D. Weissman, "Alarm Correlation", IEEE Network Magazine, pp.52-59, November 1993.

[5] A. T. Bouloutas, S. Calo, and A. Finkel, "Alarm Correlation and Fault Identification in Communication Networks", IEEE Transactions on Communications, Vol.42, No.2/3/4, pp.523-533, February/March/April 1994.

[6] Hans-Jurgen Breuer, “ATM-Layer OAM: Principles and Open Issues”, IEEE Communications Magazine, pp.75-78, September 1991.

[7] Hiroshi Uno, Youichi Maeda, and Yukiharu Kanayama, “OAM Mechanisms for BISDN UNI and Access Networks”, International Journal of Digital and Analog Communication Systems, Vol.4, pp.161-168, 1991.

[8] Jon Anderson and Matt D. Nguyen, “ATM-Layer OAM Implementation Issues”, IEEE Communications Magazine, pp. 79-81, September 1991.

[9] CCITT, “Boardband ISDN Performance” CCITT Recommendation I.35B, June 1990.

[10] Naoshi Sato, Tsukasa Okamoto, and Tadahiro Yokoi, “Performance Monitoring Method Using OAM Cells in B-ISDN”, Proc. ICC ‘91, pp.264-268, June 1991.

[11] “Broadband ISDN Simulation Testbed - User‘s guide (Version 1.0)”, Tech. Report, Dept. of Electrical Engineering, National Taiwan University, April 1994.

[12] “Broadband ISDN Simulation Testbed - Reference manual (Version 1.0)”, Tech. Report, Dept. of Electrical Engineering, National Taiwan University, April 1994.