UACEE International Journal of Advances in Computer Networks and its Security - Volume 2: Issue 3 [ISSN 2250 - 3757]

Fault Management of Computer Networks based on Probe Station and Probe Set Selection Algorithms

Balaji Patil
Computer Engineering Dept.
Maharashtra Institute of Technology,
Pune -411038, INDIA.
balaji.patil@mitpune.edu.in

Shakti Kinger
Computer Engineering Dept.
Maharashtra Institute of Technology,
Pune -411038, INDIA.
shakti.kinger@mitpune.edu.in

Vinay Kumar Pathak
Uttarakhand Open University,
Haldwani -263139, INDIA.
vpathak@uou.ac.in

Abstract—**Fault diagnosis is a central aspect of network fault management. Since faults are unavoidable in communication systems, their quick detection and isolation is essential for the robustness, reliability, and accessibility of a system. In large and complex communication networks, automating fault diagnosis is critical. Traditionally, fault localization has been performed manually by experts but, as systems grew larger and more complex, automated fault localization techniques became important.**

Probing technique for fault localization involves placement of probe stations which affects the diagnosis capability of the probes sent by the probe stations and the overhead of instrumentation. In this paper we discussed a novel integrated approach of probe station and probe set selection for fault localization which outperforms the independent fault localization approaches.

Keywords—**active probing, fault detection, fault localization, probe set, probe station selection.**

I. INTRODUCTION

An effective network management is necessary as computer networks continue to grow not only in size but also in heterogeneity and complexity. A typical network management system ensures that the network is monitored and it runs as smoothly as possible. The network management system needs to obtain and process a large amount of diagnostic information to successfully manage the network. This information can be either acquired using certain monitoring tools [1, 2, 3, 4, 5], or received from network entities in the form of network alarms [6, 7, 8, 9]. Fault management systems can be categorized into two paradigms: (1) An active or probe-based system that actively captures sample performance data from the managed network; and (2) passive system that uses network alarms to manage network. Both methods have their own advantages and disadvantages.

One promising approach proposed in the past is based on probing [1][2]. In probing, messages (also called probes) are periodically sent across the network. The results of these probes decide the success or failure of the network components used by the probe. *Traceroute* and *ping* are some of the probes that are used to check the network latency and availability. Custom application-level probes can be used to test specific application performance. Probing based techniques have various advantages over the traditional passive monitoring based techniques [10], such as (1) less instrumentation, (2) capability to compute end-to-end performance, (3) quicker localization, etc. The two challenges involved while developing probing-based monitoring solutions are probe station selection and probe selection. The probe station selection addresses the problem of identifying nodes in the network where the probe stations should be placed. The probe stations are the nodes that send probes into the network and analyze probe results. The placement of probe stations affects the diagnosis capability of the probes sent by the probe stations. The probe station placement also involves the overhead of instrumentation. Thus it is important to minimize the required number of probe stations without compromising on the required diagnosis capability of the probes. Once the probe stations are selected, the probe selection problem addresses the task of identifying appropriate probes such that the failure can be detected and localized.

In this paper, we have address the probe station selection and obtaining the small probe set problem, that is small and has the same diagnostic power as original probe set for fault detection and localization. We present a novel approach for integration of probe station and probe set selection.

II. FAULT MANAGEMENT

In probing-based technique, there are two main problems to address probe station selection and probe set selection. The probe station selection involves selecting nodes in the network where the probe stations should be placed. The probe station nodes should be selected such that the required diagnosis capability can be achieved through probes. The probe station selection involves an additional instrumentation cost, hence the number of probe stations need to be minimized. After the probe stations are selected, the probe set selection problem addresses the task of selecting appropriate probes such that the failure can be detected and localized.

We have modeled the Fault Management System (FMS) into 3 modules as depicted in the Fig. 1.

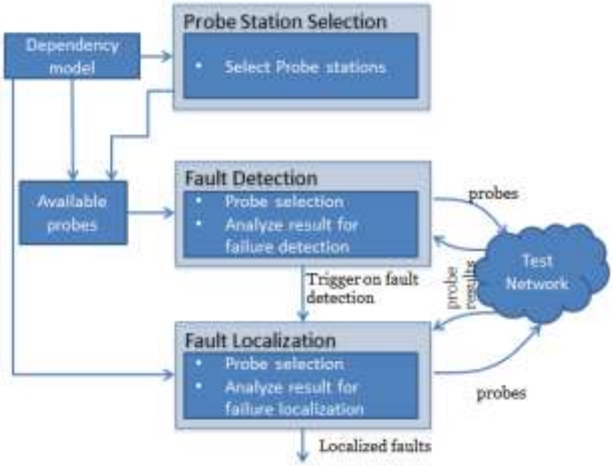

Figure 1: System architecture

Probe Station Selection: This module will determine which nodes in the network have more coverage and identify optimal set of probe stations.

Fault Detection: This module will identify any failures in the network and trigger next module to narrow down the exact failure. Select a minimum probe set from available probes to detect any failure in the network.

Fault Localization: This module will send additional probes, if required, to localize the failure in the failed network path.

III. PROBE STATION SELECTION

Location and responsibilities assigned to probe stations must be decided while building an active probing solution. These decisions are based on nature of routes, nature of targeted failures, availability of dependency information etc. [11]. Below we discuss various such factors that contribute to the overall decision making of probe station selection:

- Nature of targeted failures: Probe station selection depends on the nature of faults to be diagnosed viz. a node failure or an edge failure. A single probe station might not be sufficient to detect all of node and edge failures.

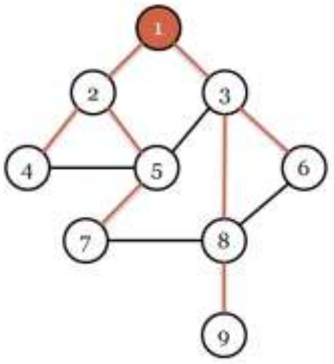

Figure 2: Link failures not being covered by Probe station 1

For instance, consider the network shown in Fig. 2. Consider node 1 to be a probe station; it can detect any single node failure in this network. However, it can detect failure of only those links that are used in reaching other nodes in the network, i.e., the links shown in red.

- Maximum numbers of failures: In a connected network consisting of *k* failures, a set of probe stations can localize any *k* non-probe-station node failures if and only if there exists *k* independent probe paths to each non-probe-station nodes.

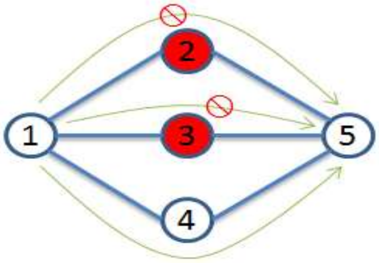

Figure 3: *k* Independent paths allow detection of *k* node failures

Fig. 3 shows three independent (node disjoint) paths to node 5 from probe station 1. Even if there are failures in two paths, node 5 can still be probed.

- Probe station failure: The assumption of fault tolerant probe station may not be practical and hence probe station selection problem becomes even more challenging. In case of probe station failure, probe stations are selected such that there exists *k* independent paths to each of probe station as well.

- Topological constraints: Another important criterion involved in probe station selection is the topological constraint. The node with less connectivity needs special treatment. Special topology structures like chains and rings also demand specific probe station placement requirements. One approach to simplify this problem could be to devise a solution by reducing the network into smaller sub-networks connected by such specific network structures like rings, chains, leaves, etc.

A. *Probe Station Selection Algorithm*

As explained earlier probe station selection criteria requires multiple factors to be considered. For our work we have used algorithm that covers some of these factors. We assume a limit *k* on the maximum number of faults that need to be diagnosed in the network.

We model the network by an undirected graph $G(V,E)$, where the graph nodes, *V*, represent the network nodes(routers, end hosts) and the edges, *E*, represent the communication links connecting the nodes, We use $P_{u,v}$ to denote the path traversed by an probe from a source node *u* to a destination node *v*.

Probe Station Selection: find the set $Q \subseteq V$ of least cardinality such that every node $u \in \{V - Q\}$ has *k* independent paths from the nodes in *Q*.

Initially the selected probe station set is empty and all nodes belong to the uncovered node set. The first probe station is selected based on the node degree. The node with the largest number of neighboring nodes can remove maximum number of nodes from the uncovered node set during the first probe station selection and hence is a good candidate to be selected as the first probe station.

When only one probe station has been selected, all nodes that are not neighbors of the selected probe station belong to the set of uncovered nodes. All the nodes that do not belong to the selected probe station set are candidates for the next probe station selection. For each candidate probe station, the algorithm determines how the uncovered node set would change if the candidate was selected as a probe station. This uncovered node set will consist of

i) nodes that are not neighbors of selected probe stations, and
ii) nodes that do not have *k* unique paths from the selected probe stations.

of all the candidate probe station nodes, the node that produces the smallest set of shadow nodes is selected as the next probe station node. The algorithm iteratively adds a new node to the probe station set till the desired capacity of diagnosing k faults is achieved. The algorithm terminates when no shadow nodes are present or the probe station set size reaches the maximum limit.

Algorithm: MinPS

```
input: MAXFAULTS
output: Probe station set

1. define:  $N$= Number of nodes in the network
            $UN$ = Uncovered nodes set
            $PS$ = Probe station set
            $V$ = Set of nodes in the network
2. initialize $PS \leftarrow$ NULL, $UN \leftarrow V$
3. select node $u$ with highest node degree as first probe
   station
4. add node $u$ to $PS$ and remove $u$ from $UN$
5. remove neighbors of $u$ from $UN$
6. for each node $c \notin PS$, compute uncovered node set $S(c)$
   such that there are k independent paths from these probe
   stations to remaining uncovered and non-neighbor nodes
7. select node c with smallest $|S(c)|$ as next probe station
8. add $c$ to $PS$ and set $UN \leftarrow S(c)$
9. repeat step 6 thru 8 until $|UN|$ = 0
```

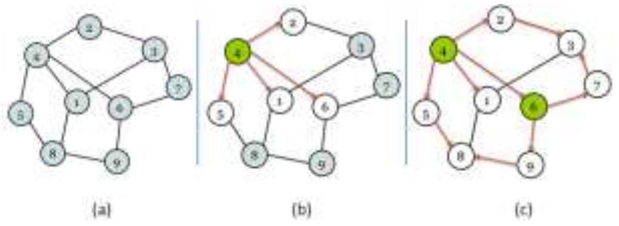

Figure 4: Probe station selection

Fig. 4 presents an example of how the probe station selection algorithm selects probe stations to detect any two node failures in the network. Fig. 4a shows a network topology with nine nodes considering all nodes as uncovered nodes. Fig. 4b shows node 4, being the node with largest degree, as the selected probe station removing neighboring nodes 2, 6, 1, and 5 from the uncovered node set. Fig. 4c shows node 6 as the next selected probe station, which removes neighboring nodes 9 and 7 from the uncovered node set. Nodes 3 and 8 are not neighbors of any probe station, but they have two independent probe paths from probe station 4 and 6 as shown in the Fig. 4c. Thus both nodes 3 and 8 are also removed from the uncovered node set. Thus the probe station placement at nodes 4 and 6 can detect any two node failures in the network.

IV. PROBE SET SELECTION

For fault detection and localization requires finding the smallest probe set. Probe set selection criteria for fault detection and localization is different. Probe set for failure detection consists of probes that cover all network elements. However probe set for fault localization consist of probes that uniquely diagnoses the suspected network elements.

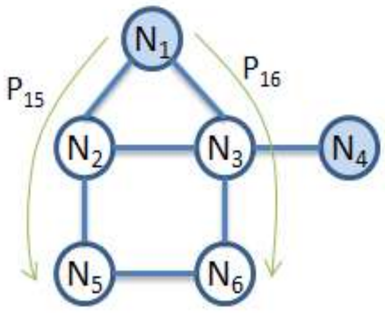

Figure 5: Sample network

A. *Fault Detection*

The task of fault detection is to find the smallest subset *P'* of the probe set *P* such that, if any (non-empty) $f \in F$ occurs, there is some probe $p \in P'$ that is affected by *f*. This can be formulated in terms of the dependency matrix:

Detection: Given $D_{P,F}$, find P^* that minimizes $|P'|$, where $P' \subseteq P$ such that there is at least one 1 in every column of $D_{P',F}$. by monitoring the probes we will know, as soon as a probe fails to return, that there is a problem somewhere in the network, but we may not know exactly what the problem is.

B. *Fault Localization*

Fault localization requires finding the smallest probe set such that every fault has a unique probe signal, since in that case exactly which fault has occurred can be determined from the probe results. Since the probe signal of fault f_j is the column c_j of $D_{P,F}$, each fault has a unique probe signal if and only if each column in $D_{P,F}$ is **unique**; i.e. differs from every other column. Since two columns c_i, c_j differ if and only if there is some entry where one of them has the value 1 while the other has the value 0 (i.e. there is some probe which is affected by one of the faults but not the other), fault localization can be expressed using the number of non-zero elements, denoted by n_{ij}, in $c_i \oplus c_j$, where $\oplus$ denotes exclusive-OR:

Localization: Given $D_{P,F}$, find P^* which minimizes $|P'|$, where $P' \subseteq P$ satisfies $\forall f_i$, $f_j \in F$, $n_{ij} \geq 1$.

Referring to sample network in Fig. 5, fault detection requires finding the smallest number of rows such that every column (excluding, of course, f_7) has at least one 1. In this example, this means the smallest set of probes which pass through every node, so that, no matter which node fails, there is a probe that will detect it. The following set of 2 probes suffices:

	f_1	f_2	f_3	f_4	f_5	f_6
p_{16}	1	0	1	0	0	1
p_{45}	0	1	1	1	1	0

Since no single probe passes through all the nodes, this is clearly a smallest subset for fault detection. However this set fails for the task of fault localization because, for example, failures in nodes *N*1 and *N*6 cannot be distinguished from each other - they generate the same signal, since their columns are identical. However the following set of 3 probes is a minimal set for fault localization:

	f_1	f_2	f_3	f_4	f_5	f_6	f_7
p_{15}	1	1	0	0	1	0	0
p_{16}	1	0	1	0	0	1	0
$p42$	0	1	1	1	0	0	0

Note that the "fault" f_7, denoting no failure anywhere in the network, is included here, because we want its column to be unique, as well as the columns of each individual node failure. Since all 7 columns are unique, the results of these 3 probes allow us to determine exactly which node has failed. For example, if p_{15} and p_{16} both fail but p_{42} succeeds, then we infer that node N_1 has failed.

C. *Probe Set Selection Algorithm*

After the deployment of probe stations, appropriate probes need to be selected such that the required diagnosis capability can be obtained. As probes involve sending additional network traffic, it is important to minimize the number of probes to perform fault diagnosis. We use a form greedy search algorithm where each probe is evaluated in terms of their localization quality. Localization quality of a set of probes is defined as amount of information provided by a probe set for faults in a network.

The localization decomposition $S_{P,F}$ is a collection of groups $\{G_1,...,G_k\}$, where each group G_i contains the faults $f_i \in F$, that cannot be distinguished from one another by *P*. Then localization quality of *P* is defined as the conditional entropy $H(\mathcal{F}/\mathcal{G})$, where $\mathcal{F}$ is random variable denoting fault and $\mathcal{G}$ the random variable denoting which group of $S_{P,F}$ contains the fault.

$$Q(P,F) = H(\mathcal{F}/\mathcal{G})$$

If the faults are independent and equally likely, then

$$Q(P,F) = \sum_{i=1}^{k} \frac{n_i}{n} \log n_i$$

Where n_i is the no. of faults in group G_i of $S_{P,F}$ and $n=|F|$

Algorithm: Greedy Search

```
input: Dependency matrix $D_{P,F}$, with rows $p_1,p_2,...,p_r$
output: Probe set $P'$ (possibly non-minimal size)

  $P' = \phi$ = empty set
  While $S_{P',F} \neq S_{P,F}$
      $p^* = argmin_{p \in P \setminus P'} Q(P' \cup \{p\}, F)$
      $P' \leftarrow P' \cup \{p^*\}$
  Output $P'$
```

V. EXPERIMENTAL EVALUATION

In this section, we present the experimental evaluation of the proposed algorithm. We apply algorithms to select minimal set of probe station followed by minimal set of probe set for fault localization.

A. *Experiment Setup*

We are using OMNET++ as simulation tool to simulate network, test our algorithms and capture results. We produce different scale networks using OMNET++ random network generator. Given a network topology the simulation proceeds with

- Selecting probe stations using MinPS algorithm
- It next generates dependency matrix for the network
- Using Greedy search algorithm it selects probe set

B. *Simulation Results*

We have studied results of our algorithm with different size of networks and compared it with results obtained from random probe selection algorithm and proposed MinPS algorithm. The Random selection algorithm randomly selects node as a probe station. The process is repeated until probe station resulting into null shadow nodes is achieved.

We conducted experiments with network size varying between 10 and 50 nodes. The results show that the proposed algorithm MinPS provides better results as compared to random algorithm as network size increases.

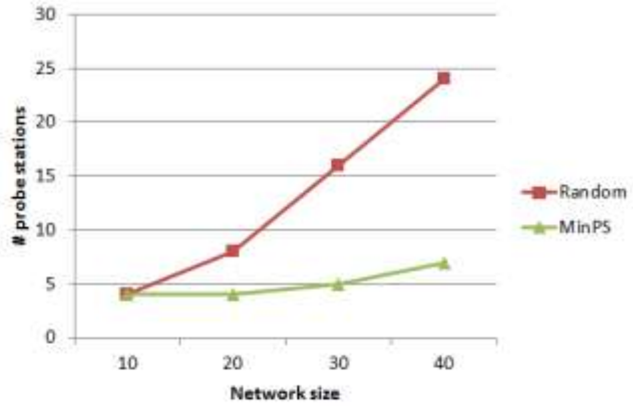

Figure 1: Number of probe stations with different network sizes

The results of experiments with integrated probe station and probe set selection algorithm reveals that probe station selection plays a pivotal role in identifying minimal set of probes. It is more evident as network size increases than with smaller networks.

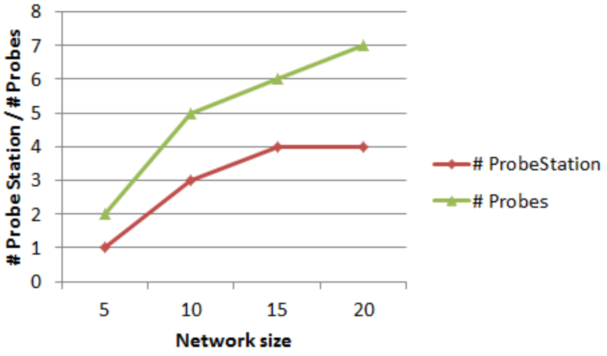

Figure 2: Number of probe stations and probes

We plan to continue with our experiments to study the impact of node degree along with network size on probe station, probe set selection through integrated algorithm.

VI. CONCLUSION

In this paper we have developed active probing solution for fault localization in computer networks. We presented architecture for building such fault management system. We have successfully integrated probe station selection, probe set selection for fault detection and localization. We have presented experimental analysis of the algorithms through simulation results.

ACKNOWLEDGMENT

We would like to thank Mr. Maitraya Natu, Scientist, TRDDC, Pune, for guidance. We also thank for all the encouragement, motivation and support provided by Dr. V. D. Karad, Executive President, MIT, Pune, and Dr. L. K.Kshirsagar, Principal, MIT Pune.

REFERENCES

[1] M. Brodie, I. Rish, S. Ma, Optimizing probe selection for fault localization, In the 12th International Workshop on Distributed Systems Operations Management, 2001.

[2] M. Brodie, I. Rish, S. Ma, G, Grabarnik, N. Odintsova, Active probing, Technical Report IBM, 2002.

[3] M. Natu, A. S. Sethi, Active probing approach for fault localization in computer network", In E2EMON'06, Vancouver, Canada, 2006.

[4] M. Natu, A. S. Sethi, Efficient probing techniques for fault diagnosis, Second International Conference on Internet Monitoring and Protection, IEEE, 2007.

[5] M. Brodie, I. Rish, S. Ma, N. Odintsova, A. Beygelzimer, G. Grabarnik, K. Hernandez, Adaptive diagnosis in distributed systems, Technical Report IBM, 2002.

[6] S. A. Yemini, S. Kliger, E. Mozes, Y. Yemini, D. Ohsie, High speed and robust event correlation, IEEE communications Magazine 34 (5) (1996) 82-90.

[7] R. Gardner, D. Harle, Alarm correlation and network fault resolution using Kohonen Self-Organizing map, Globecom 97 proceedings, pp. 1398-1402, 1997.

[8] A. T. Bouloutas, G. W. Hart, M. Shwartz, Fault identification using a FSM model with unreliable partially observed data sequences, IEEE Transactions on Communications, 41(7):pp. 1074-1083, 1993.

[9] C. Wang, M. Schwartz, Identification of faulty links in dynamic-routed networks, IEEE Journal on Selected Areas in Communications,11 (3) 1449-1460, 1993.

[10] M. Stemm, R. Kaltz, and S. Seshan. A network measurement architecture for adaptive applications. In IEEE INFOCOM, Mar 2000.

[11] M. Natu and A. S. Sethi. Probe station placement for robust monitoring of networks. Submitted to Journal of Network and Systems Management.

[12] Yalian Pan, Xuesong Qiu, Shunli Zhang, Probe Station Selection in Non-deterministic and Dynamic Virtual Network Environment *Journal of Information & Computational Science (2012)*.