

Optimizing Routes using Genetic Algorithms based on Throughput or Delay Sensitivity of Applications using SNMP for Automatic Discovery of Network Topology

M.Mohamed Surputheen
Research Scholar
Dr. M.G.R Educational and
Research Institute University
Chennai - 600 095, India.

G.Ravi
Research Scholar
Dr. M.G.R Educational and
Research Institute University
Chennai - 600 095, India.

R.Srinivasan, PhD.
Dean, Research and PG studies
RNS Institute of Technology
Bangalore– 560 061
India

ABSTRACT

Network Applications can be broadly classified as Throughput sensitive or Delay sensitive. Such applications require efficient routing mechanisms in order to work effectively. Genetic Algorithms can be used for defining the Best or Optimal Route based on its sensitivity and Various Constraints that the application imposes usually referred to as Constraint Satisfaction Problems (CSP).

The use of Genetic Algorithms for selecting an optimal route based on CSP requires a mechanism for automatic discovery of network topology and also mechanisms for learning the capacity of the network infrastructure. SNMP along with MIB provides the required data for Topology discovery and also to learn the information about the current network capacity along with various bottlenecks present in the topology. Since the shortest path is not always the best path, our genetic algorithm will provide the optimal route based on CSP and application sensitivity.

Keywords

Genetic Algorithm, Network Management, SNMP, Topology Discovery

1 INTRODUCTION

1.1 Traffic Sensitive Applications

Network Applications are of two broad types: Throughput Sensitive applications and Delay Sensitive Applications.

Throughput Sensitive Applications use TCP for transmission because they require acknowledgement and retransmission in case of packet loss.

Throughput-sensitive applications

- Prefer high-bandwidth paths
- Keep queues occupied

Example: Download Managers and Browsers

Delay Sensitive Applications usually prefer UDP over TCP because acknowledgement and retransmission is of no use there.

Delay-sensitive applications

- Prefer low-propagation delay paths
- Keep queues small

Example: Video Conferencing and Online Games

Since retransmission is of no use in delay sensitive applications, they must always choose the route with the low-propagation delay to reach their destination because a little delay will make them unusable. But in case of throughput sensitive applications, we need high bandwidth reliable transfer, because loss of packets could not be tolerated. Here the time it takes could be minimized if a proper route is selected between the source and destination that satisfies the constraints. We must know that the shortest path may not be the best or optimal path for all cases, under varying traffic conditions.

1.2 Simple Network Management Protocol (SNMP)

SNMP (Simple Network Management Protocol) manages all the SNMP supported devices based on TCP/IP protocol. Likewise the manager can communicate with all SNMP-supported devices to manage the network. In realization, SNMP provides the administrator a platform (NMS, also called a station) to take charge of network command sending, data store and data analysis. The managed devices would run a SNMP Agent to enable the communication between devices and stations. SNMP is becoming the most universal and popular network management protocol and the standard SNMPv3 for its easy realization. Its purpose is to enable network stations to effectively and easily supervise and manage devices. It consists of manager, management information database (MIB), agent and managed devices/objects. [1],[2]

An SNMP-managed network consists of three key components:

1. Managed device
2. Agent — software which runs on managed devices
3. Network management system (NMS) — software which runs on the manager

A managed device implements an SNMP interface that help in the sharing of node-related information. The information flow can be unidirectional or bidirectional. Every managed devices shares its information with the network management software.

An agent is a network-management software module that is present on a managed device. An agent has local knowledge

of management information and translates that information to or from an SNMP specific form.

An NMS (Network Management System) provides the processors and memory for managing the network. They execute applications and monitor and control the managed devices. Any number of NMS's can exist in a single network.

1.3 Genetic Algorithms

Genetic algorithm is a computational model simulating the process of genetic selection and natural elimination in biologic evolution. Pioneering work in this field was conducted by Holland in the 1960s [3], [4]. As a high efficient search strategy for global optimization, genetic algorithm demonstrates favorable performance on solving the combinatorial optimization problems.

A Genetic algorithm has the capability to automatically gather information about the search space in which it is implemented. Since this is the inherent property of Genetic algorithms, we can obtain a global optimal solution.

The genetic algorithm simulates the process of selection and elimination that forms a part of the evolution process. During every process, the resultant sets are analyzed and are ranked. This ranking process is used in the elimination of items in the list. Genetic operators, such as crossover, mutation, translocation and inversion, are then performed on those qualified solutions to estimate new candidate solutions of the next generation. The above process is carried out repeatedly until certain convergent condition is met [5].

2. MOTIVATIONS

2.1 Traffic Sensitive Routing based on CSP

Using a shortest path approach to routing may not work well at all times. This is because the shortest path from a source to destination may have bottlenecks and also it is used by most of the applications because it is the shortest path. There may be another path that is a little longer than this but, without any bottlenecks and with much more bandwidth capacity to transfer huge amounts of data. Selecting the shortest path or the optimal path depends on the application requirements and also the amount of data that is to be transferred.

Based on these constraints, our genetic algorithm based approach, selects the route that is optimal for a particular type of application depending on its sensitivity.

2.2 Automatic Topology Discovery

The genetic algorithm based routing mechanism requires an accurate topology of the network and also the details about their capacity to determine the presence of bottlenecks. We can use ICMP, SNMP or ARP based Automatic Topology Discovery Mechanism.

2.2.1 Topology discovery method based on ICMP

Reference [8], [9] proposes a topology discovery method based on ICMP. This method depends on two basic ICMP tools.

Ping: Detect activity and attainability of the network devices by taking advantage of ICMP echo reply message.

Trace-route: Detect route information by taking advantage of ICMP time exceed message or the port unreachable message or TTL message.

Ping each IP address available in a network by polling way to find active devices in target network. Then trace-route every active device, record the route information of each trace-route operation so that we can find connection of devices. Finally, we conclude the whole network topology by analyzing information we got by ping and trace-route.

2.2.2 Limitations of ICMP based approach:

- Ping operation can easily find out active network devices, but it is not able to discover the devices in the path ICMP messages passed by.
- Trace-route operation can help us know which route devices ICMP messages passed by, but it is hard to discover connection of devices.
- Even if we get reply messages from targeted devices by ping, it is still hard to tell the structure inside the subnets.
- Because a lot of firewalls forbid ICMP message to pass through, this method is not able to discover the devices behind the firewall.

Topology discovery method based on SNMP

The SNMP [1], [6], [7], [12] (Simple Network Management Protocol) is a kind of protocol based on TCP/IP. We can use it as a tool to get the network information from MIB. Today, most of network devices support SNMP. We can get network topology information from MIB by SNMP command.

Limitations of SNMP based Approach

- The method can't discover the network devices which do not support SNMP.
- There is much redundant information in IP Router table.
- A route device always has several IP address (one interface binding one IP address), and route device always uses IP address as its ID. That means a route device can be several devices logically. As a result it is very difficult to reflect connections at port level. So this method is suitable to discover the topology of network backbone, but not suitable to reflect the whole network.

So SNMP based approach could be useful for determining the topology of the Network Core but not the Network Edge or Subnets.

Topology discovery method based on ARP

Ethernet network device supports ARP (Address Resolution Protocol) and maintains ARP table [10]. This ARP table is a list used for converting IP address to physical one. Each record in the ARP table contains a single IP address and its corresponding physical address. We can use this ARP Table to find the list of devices on a route or to determine the route to a particular device. After the identification of a

device, we can differentiate between whether the particular device is a node or a router by obtaining extra information available from the node. By continuing this process, we can obtain the topology of a complete network[11].

Limitations of ARP based Approach

- Devices must support ARP
- If network is too large, the ARP table becomes lengthier hence increasing the amount of memory consumed.

So this method is just suitable to discover the topology of LAN.

3. OUR APPROACH

Our aim is to find the optimal route based on the traffic sensitivity of the application. Throughput sensitive applications prefer high-bandwidth routes whereas delay sensitive applications prefer routes with low propagation delay. In order to optimize this search problem we rely on genetic algorithms using which we can arrive at a global optimal solution.

The various steps involved in our approach are as follows: Determine the Network Topology, Analyze network capacity and bottlenecks, use Genetic Algorithm to determine the optimal route based on constraints imposed by the application's traffic sensitivity.

3.1 Determine the Network Topology

In order to obtain the topology effectively we can combine SNMP based approach to detect the Network Core and then to use ARP based approach to determine the Network Edge. This way we can overcome the problems of both and also utilize their advantages synergistically to discover the network topology.

3.2 Analyze the Network Capacity and Bottlenecks

We must analyze the topology of the network to determine high bandwidth routes and routes with low propagation delay. The shortest path algorithms can then be fine tuned to suit our needs of finding the low propagation delay path for delay sensitive applications and to find the high bandwidth routes for throughput sensitive applications. We can retrieve the SNMP MIB database objects using GetBulkRequest PDU to identify the network capacity and also to find various bottlenecks in the network.

3.3 Genetic Algorithms to obtain Optimal Route

In order to optimize our search problem, genetic operators and coding methods are designed in advance so that the individuals may satisfy the constraints. In contrast, the objective of constraint satisfaction problems (CSPs) is to find an individual that satisfied constraints as the fitness value. Search in usual GA, based on neo-Darwinian evolutionary theory is conducted by crossover and mutation.

Crossover is generally considered a robust search means. Offspring may inherit partial solutions without conflict from their parents, but no information to decide which genes are

partial solutions is available. From a search strategic point of view this means that variables are randomly selected and then values which will be assigned to them are also randomly selected from genes contained within a population.

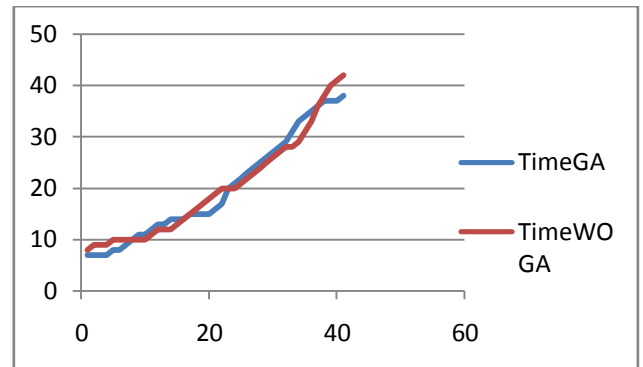


Fig 1: Time and No of Packets with and without Genetic Algorithm

Instead of just finding a high bandwidth route or low propagation delay based route, GA's can further be optimized if we classify applications based on their relative degrees of sensitivity. We can include the Sensitivity information in the TCP/UDP header fields and thus enable the classification of applications into various groups instead of two. We can use a **4 bit** sensitivity value to classify the applications into 16 groups with varying degrees of sensitivity requirements for transmission.

The optimal routes between source and destination, for each of the 16 groups could be calculated based on the Genetic Algorithm. Recently used routes could be cached in the routers' interfaces.

To obtain the optimal path we can consider the Dijkstra's Algorithm fine tuned for Genetic Approach.

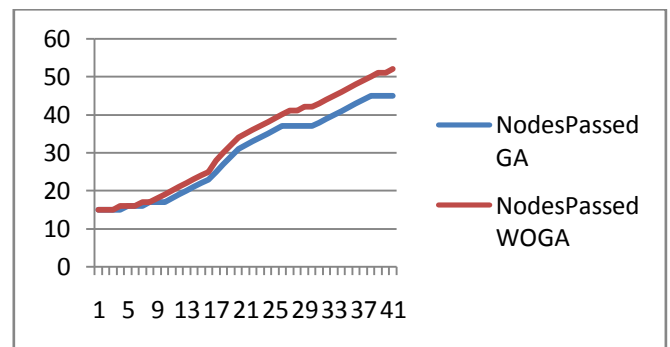


Fig 2: Time and No of Nodes Passed with and without Genetic Algorithm

Strategy:

A (directed) graph $G = (N, E, C)$ consists of a node set N , a cost set C and an edge set E . Each item in the set E represents the edges or connections between two nodes. If an element (u, v) exists in the edge set E then it represents that a connection exists between the two nodes u and v . Each item in the edge set has a direct correspondence to an item in the cost set C . A path is said to exist between two nodes u and v if a direct

entry exists in the edge set or a path can be made from u to v using the existing edges.

Algorithm:

```

Dijkstra (N,E,s,d);
foreach u in N do {
C(s,u) = ;
C(u,u)=0;
path(u,v)=null;
}
fs=[s];
es=[];
while not_empty(fs) do {
select u from fs with min C(s,u);
fs=fs - [u];
es=es+[u];

if (u=d)
exit;
else {
fetch_neighbours(u);
foreach (v, C(u,v)) in n {
if (C(s,v)>C(s,u)+C(u,v)) {
C(s,v)=C(s,u)+C(u,v);
path(s,v)=path(s,u)+(u,v);
if not_in(v, fs U es)
fs=fs+[v];
}}}}
fetch_neighbours(u)
{
calculate fitness value of
individuals
select two individuals at random
single point cross-over
return high fitness individual
}

```

Here we have used a modified form of Dijkstra's algorithm for finding the optimal path from a source to a destination. The usual Dijkstra's algorithm finds the shortest path without considering other parameters such as bandwidth or traffic. In our approach, the Dijkstra's algorithm modified using a Genetic Approach is followed. This algorithm considers all other parameters that influence the transmission of packets. Hence even if the returned path is not the shortest, it will prove to be the most optimal path for packet transmission.

Experimental Results:

Network Simulator 2 (NS2) is used for performing the simulation. All the nodes that are present along with their neighbors, bandwidth and distance are recorded in the MIB, which serves as the base data for this approach. The obtained results show that the number of packets that can be transmitted from a source to destination shows a considerable reduction in time (Figure 1) while using the modified Dijkstra's algorithm when compared to the algorithm that does not use the Genetic approach. Our approach also shows an increase in the number of nodes passed during a given

period of time (Figure 2). Hence we can conclude that our approach provides faster transmission periods and optimal paths when compared to the ordinary approach.

4. COMPARITIVE STUDY

We cannot rely on ICMP based approach as mentioned in [8][9] because it works on a higher layer of the TCP/IP Stack and also imposes heavy load on the network infrastructure. Since ICMP is usually used for Denial of Service (DoS) attacks, it is blocked by most network firewalls, which prevent us from determining the structure of the network beyond the firewall. Hence our usage of SNMP provides a simple, efficient and cost effective methodology. It works on a lower layer of the TCP/IP Stack and also doesn't impose heavy load on the infrastructure. But we cannot determine the topology beyond the network core, where the devices may not support SNMP. So this method is suitable to discover the topology of network backbone, but not suitable to reflect the whole network. So we also combine the ARP based approach [10] to determine the topology. The IP address in ARP refers to the IP of active devices, and ARP table almost has no redundant data. So the topology discovery method based on ARP is very efficient.

Hence our usage of SNMP compared with ARP provides us the best solution for finding the network structure rather than single approaches.

5. CONCLUSION

Our approach for finding the optimal route based on genetic algorithms demonstrates favorable performance on solving the combinatorial optimization problems in Traffic Sensitive Routing with varying degrees of sensitivity. We show that the optimal route selection problem in network analysis can be effectively solved with genetic algorithm through efficient encoding, selection of fitness function and various genetic operations. Further efforts can be made on enhancing the adaptability of the algorithm under dynamic constraints and expanding the applications of the algorithm into large scale Ad-Hoc networks where cache consistency could be a problem.

6. REFERENCES

- [1] "Network topology discovery algorithm based on SNMP", DENG Ze-lin, ZHANG Li-fang, LIU Yi-nan, FU Ming, Journal of Changsha University of Science & Technology(Natural Science), 2007,4(4):68-72.
- [2] "Algorithm and realization of network topology discovery", Huang Xiaobo, Pan Xuezheng, Computer Applications and Software,2008,21(3):64—66.
- [3] "Adaptation in Nature and Artificial Systems", Holland J. H.,
- [4] "An Introduction to Genetic Algorithms for Scientists and Engineers", 1992. World Scientific, Coley, D.A.
- [5] "Genetic Algorithms in Search, Optimization and Machine Learning", Goldberg, D.E., 1989, Addison Welsey Publishing Company.
- [6] "Algorithm Research of Topology Discovery on SNMP", Yang Qiuxiang, 2010

- [7] "Analysis and Research of Network Topology Discovery Method", He Peng, Qiu Jian-lin, Gu Xiang, 2010
- [8] "Traceroute Probe Method and Forward IP Path Inference", Luckie M, Hyun Y, Huffaker B., In Proceedings of the 2008 ACM SIGCOMM Internet Measurement Conference: 311- 323,2008.
- [9] "An intelligent self-learning algorithm for IP network topology discovery.", Najeeb Z, Nazir F, Haider S, Suguri H, Ahmad H.F, Ali A., 2005 14TH IEEE Workshop On Local & Metropolitan Area Networks (Lanman):60-65, 2005.
- [10] "Research and implementation of topology discovery algorithms in campus networks", Liu YH, Yang JG, Xiao DB. In the Proceedings of DCABES 2002:33-35,2002.
- [11] "Hierarchical broadcast ring architecture for highspeed Ethernet networks", Jang H, Kim HS. In the proceedings IEEE INFOCOM 2006, IEEE INFOCOM SERIES: 3027-3031, 2006.
- [12] "Research and Application on automatic Network Topology Discovery in ITSM System", Yide Wang, Dancheng Li, Chunyan Han, Zhiliang Zhu