

A Dynamic Polling Scheme for the Network Monitoring Problem

Feng Gao, Jairo Gutierrez*

Dept. of Computer Science

*Dept. of Management Science and Information Systems

University of Auckland, New Zealand

Email: fgao001@ec.auckland.ac.nz

* j.gutierrez@auckland.ac.nz

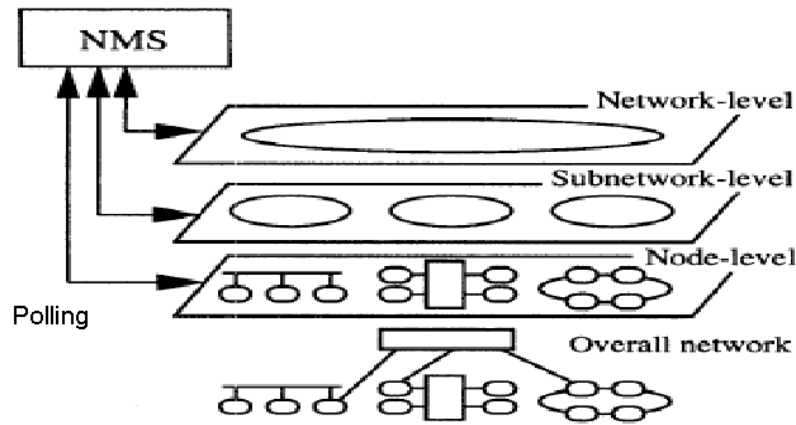
Abstract

Network monitoring is one of the most significant functions in network management. In order to measure network elements and services, normally a polling approach is used to help network professionals understand the state of the managed systems in real-time. Given the current rapid growing networks, the information exchanged between network nodes and a central network management system (NMS) could be substantial.

In this study, a dynamic polling scheme is proposed to transform the network monitoring problem into an optimization problem of polling frequencies assignment when network resource consumption becomes a constraint. The motivation for such scheme arises from the needs of network administrators who assign management efforts efficiently on the elements within the networked system. The proposed dynamic polling scheme determines the polling frequencies to each of the network nodes in accordance with the integration of the Analytic Hierarchy Process (AHP) method and Branch-and-bound techniques. Thus, the goal of the optimization is to trade-off between the management efforts on the managed elements and the performance of network behaviours.

The simulation results show that the proposed dynamic polling scheme is more efficient than a simple random polling scheme, and performs well across a wide range of traffic loads.

Introduction



1. Introduction

Efficient network monitoring assumes having reliable information about the managed systems. A typical set of network monitoring tasks used to determine the current state of the monitored systems are as follows. Intervals (frequencies) at which the state of the monitored systems need to be sampled are determined based upon the granularity of the control actions. After that, all the sample data are collected and transferred continuously for interpretation and action. In accordance with network management terminology, this activity is referred to as a polling approach.

As shown in the figure above, a network management system (NMS) monitors and controls a network by polling all network elements (could be in node-level, subnetwork-level, and network-level) in order to get the most recent performance metrics, such as loss, jitter, latency, throughput, etc. The value of any metrics for management decisions can be obtained only through polling actions issued by the NMS.

It is well known, however, that some limitations of polling hinders network management from efficiently sampling the state of the networked systems. For instance, under some scenarios, simple periodic polling may bring management overhead on some crucial links and thus reduce the efficiency of the entire system. The management efforts could also consume already limited network bandwidth. Besides, the increasing complexity of managed systems and services provided by them generates a need for monitoring of more and more network resources.

To balance the efforts of polling and the achievement of management efforts, the issue of efficient coverage and allocation of polling should be addressed. The coverage of polling and the volume of data directly impact the accuracy of the measurement and the network behaviors. Additionally, the selection of the polling frequency is very crucial.

The motivation of this study is to put more management efforts on more critical components while an adaptive polling technique is employed to dynamically adjust the polling interval on each individual network element so as to optimize management efficiency and overhead.

Related Work

- Minimizing the monitoring cost associated with network monitoring
- Dynamic polling scheme based on time variations
- Distributed Network monitoring with Bounded Link Utilization
- Adaptive Sampling for network management
- A model for efficient configuration of management

2. Related Work

Despite the extensive work and contributions that have been made towards the monitoring of networks and systems, none of the existing approaches propose a single model that captures the need for optimizing management efforts, and maintaining efficient network behaviors by allocating different polling frequencies to different preferential network elements and services.

The need for optimized monitoring is the focus on [2]. Jiao etc. gave a definition of the network monitoring problem: the monitoring problem is to construct correct efficient monitoring algorithms for any given set of variables, integrity constraints and alarm conditions. Despite classifying the network monitoring problem in terms of the integrity constraints and proposing different algorithms for each of these classes, this model did not specify the polling strategies to reduce management traffic.

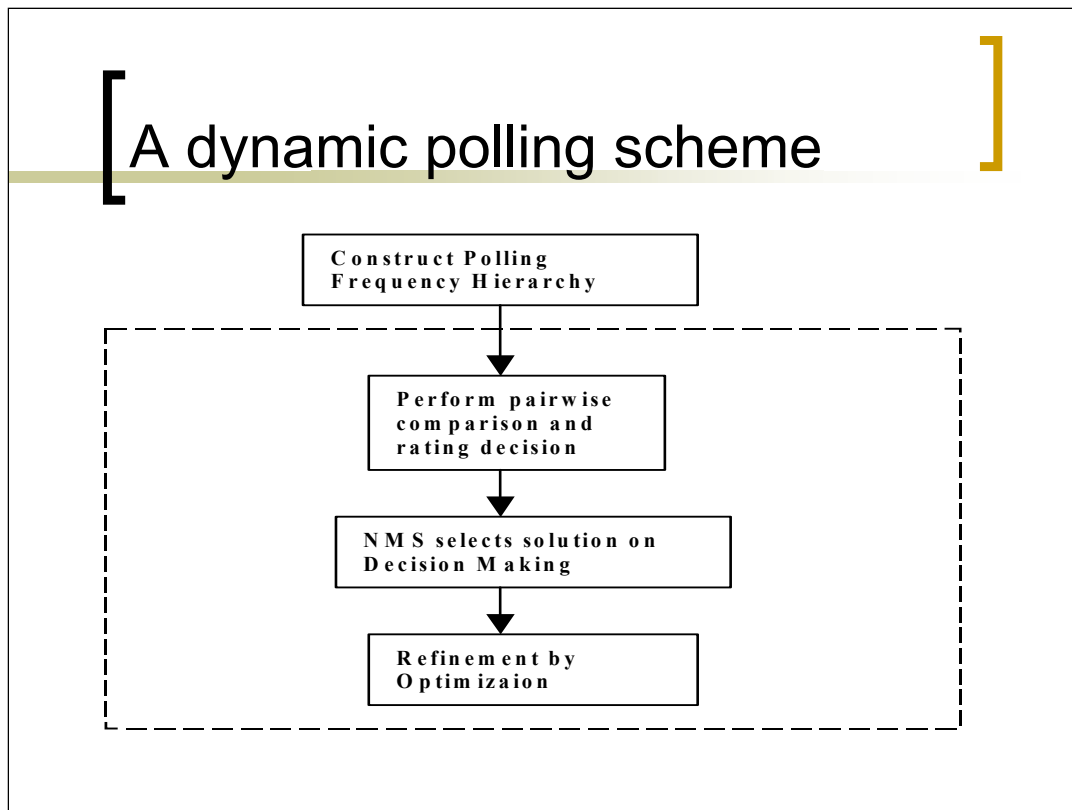
Many schemes have been proposed for controlling the overheads of dynamic polling by throttling polling rates. For instance, after having compared existing schemes, Yoshihara etc. [8] proposed a dynamic polling scheme which considers both the overhead of polling traffic and the message rates required for management tasks. However, this work did not address the optimization of network resource consumption, nor did they seem to address issues related to network behaviors.

Li etc. [4] addressed the problem of optimizing a scalable distributed polling system. Their goal was to determine a minimum subset of nodes on which to place pollers such that the bandwidth constraint on each and every network link is the maximum bandwidth that can be used for polling on that link. They also showed that the problem of optimization is NP-hard and proposed three heuristics to obtain a solution. Nevertheless, this approach was interested in minimizing the coverage cost of monitoring, and neither efficient management effort allocation nor detailed polling strategy were considered.

In [10], the adaptive sampling proposed by Hernandez et al. dynamically adjusted the sampling (polling) rate based on the observed sampled data. Unfortunately, they only considered the prediction of future behavior based on the observed samples, neither existing interrelationship and dependencies of network nodes nor optimization of network resources were considered.

Furthermore, from the configuration and resource provisioning perspective, Abdu [9] modeled three main aspects of achieving efficiency in distributed systems management: (1) the need for minimizing resource consumption by management; (2) the need to satisfy system and application requirements; (3) and the need for a dynamic solution. This work can be used as a reference for the polling frequency allocation used in this study.

[A dynamic polling scheme]



4. A Dynamic Polling Scheme

For a solution to the problem of polling frequency optimization, a new dynamic polling scheme is proposed which not only controls the increase in the polling message traffic but also complies with the change of network performance.

4.1 Execution flow and functional components of the proposed scheme

The figure above shows the execution flow of the proposed scheme. After constructing a hierarchy which incorporates network components and associated network performance metrics, in each monitoring session:

- Compute an aggregation of global comparisons
- Select solution based on management objectives
- Refine the Solution from the optimization set

The proposed scheme consists of three main components: a polling frequency hierarchy, Selection on decision making, and Refinement by optimization.

Construct a polling frequency hierarchy

The fundamental problem of the proposed dynamic polling scheme is the selection of relevant parameters and the construction of an associated polling frequency model to reflect monitoring demands. This is addressed by a hierarchy which incorporates network performance constraints and network behaviours.

NMS decisions

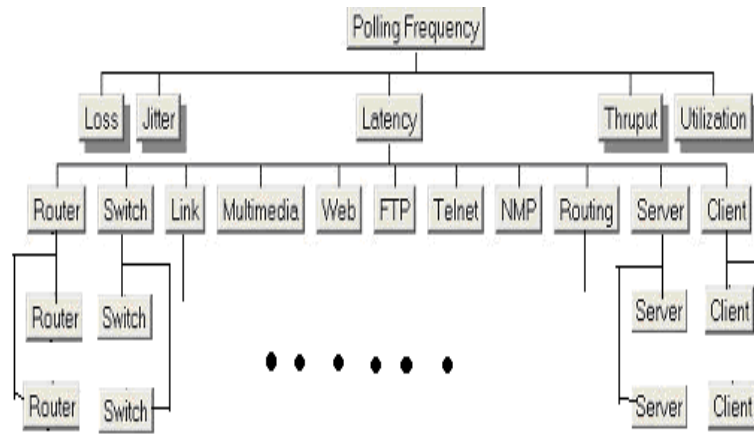
Once the global weights of each element in the hierarchy have been computed, NMS is going to select the solution based on management objectives (if existing resources are available) simply by associating weights with the polling frequency.

Optimization determines the most efficient combination

There are cases where network resources are limited (to be allocated for management purposes). Under such scenarios, by including the network constraints and transforming the resource allocation problem to a problem of optimization, the proposed scheme can provide better satisfaction to the management requirements.

In conclusion, the advantage of such integration of decision making and optimization is to filter out uncritical elements by using the appropriate weights if necessary.

[Polling Frequency Hierarchy]



4.2 Polling Frequency Model

In the proposed scheme, at first, a hierarchy of polling frequency model is constructed. Organizing all relevant network elements and associated performance in a hierarchy can simplify the complexity of management.

The figure above shows the hierarchical structure of a polling frequency model, which is the aggregation of all the related network performance metrics, the relevant network devices and services, and the existing network elements which would be selected.

The hierarchy comprises performance metrics associated with network transmission: loss, jitter, latency, throughput, and utilization. These metrics are grouped in a level because they share the properties of being the most fundamental factors for determining network behaviour.

Since the network devices and services influence the representations of the performance metrics, these network elements are layered underneath. The network element level can be tailored and scaled to a wide range of networks and applications, ranging from network devices (router, switch) to network services (server, web), and clients.

Finally, in accordance with the AHP methodology, a combination of decision elements, the real combination of all existing network elements to be polled, such as router1, route2, switch1, and so on, are deployed at the bottom.

By applying a multiple decision making technique, the Analytic Hierarchy Process (AHP) [6], this hierarchy casts polling frequency model into a multiple choice decision making problem. With the AHP, the relative importance, i.e. global weight, of each bottom decision element is judged by pairwise comparison. For comparison considerations, preferences are assigned based on the network activities and on the influence of the associated network components. Factors such as for example: the importance of the network component, its failure characteristics, etc.

After having obtained the global weights of the decision elements, the relative ranking of the decision elements is computed. Then, a synthesis of a linear weighted measure is aggregated by the multiplicity of the global weights and relative rankings of each decision element in the alternative solution. In fact, the output is a linear weighted measure that combines managed variables in a decision. For more details, readers are referred to [6].

[Decision making and Optimization]

- Multiple choice decision making by using the AHP
- Selection based on management objectives
- Optimization by Branch-and-Bound

4.3 Decision making

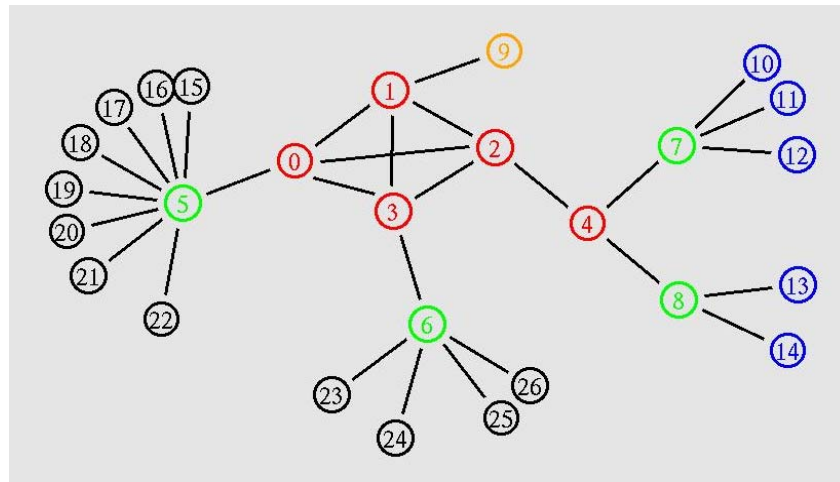
Besides taking into consideration a wide range of factors including objective and subjective aspects for evaluating monitoring effectively, this approach also examines all relevant factors, balances various objective criteria in terms of influence associated with locations, impact on network behaviour, and failure percentages. The selection of the polling frequencies based on management objectives is therefore a combination of the cumulative actions imposed on network behaviours.

A multiple choice decision making technique, the AHP, develops a utility function so that values of trade-offs are indicated and the viewpoints of various groups within the organization are reflected. It is anticipated that the AHP will be applied to balanced scores to determine values of trade-offs between various alternatives. In this way, Polling is based on preferences by weighting. It is possible to perform numerical analyses on various comparisons in order to make decisions.

4.4 Optimization

This part considers polling frequency assignment optimization given network resource constraints. The objective is to find a combination of polling frequencies from each individual network elements under the constraint, so that the management traffic consumption is not more than a given level of network bandwidth utilization. A decomposition method, based on branch-and-bound (BB) techniques [11], is used for solving the problem. In order to speed up the procedure, the depth first approach of BB is applied. The procedure is outlined here: after initialization, compute upper bound U_1 , then perform a forward step. Whenever a forward move is exhausted, the upper bound U_1 corresponding to the current solution is computed and compared with the existing best solution so far, in order to check whether further forward moves could lead to a better one: if so, a new forward move is performed, otherwise a backtracking step follows. For time saving, the optimization can be stopped when a close to optimal solution is found, that is, when further forward steps could not bring a better solution.

Simulation setup



5. Experimental Study and Results

This section describes a set of simulation experiments used to evaluate the influence of the dynamic polling scheme on the system performance. The simulation model building procedure used in this study is as follows: (1) Construct preliminary simulation model diagrams and develop data flow diagrams; (2) Collect system data and test and analyze the models for the required performance.

5.1 Simulation setup

A discrete event simulation package (ns-2) is used for this study. Input data for simulation include the topology of the network, workload placed on the network, faults injected in the network, and management structure.

The Figure above shows the simulated network connections. The core routers consist of nodes 0, 1, 2, 3 and 4. four edge routers connected to these core routers are nodes 5, 6, 7, and 8. Nodes 10 to 14 are servers, and nodes from 15 to 26 are clients. Servers provide network services such as ftp, http, multimedia. Each client node in the network may have a number of customers connected to it.

The simulated network was fed with various kinds of traffic: ftp, multimedia, web, web traffic models, etc. The main workload is generated by a set of applications running on client sides. The traffic load can be adjusted by varying the number of access users and the type of requests. Three network load conditions are studied: light load, medium load, and heavy load.

A Fault model is used to simulate a link/node failure or delay. The network fault model includes faults possible for network elements. This is realized in two ways: by being embedded within a traffic load provided by ns-2, or by being injected by specifying a down time of nodes and/or links.

For simplicity, there is a centralized network manager. That is, there is a single administrative domain where management applications reside on node 9 and they poll all nodes and links within the topology.

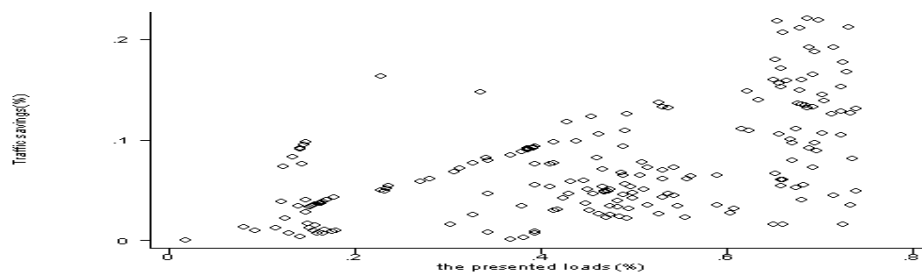
Within the experimental study, the simulation lasts 200 seconds. The data input are the same for both schemes.

For the sake of simplicity, the polling interval base for the dynamic polling scheme was set to 2 seconds. Performance metrics are collected at each network node and link.

Simulation Results

. oneway DP RP

Source	Analysis of Variance SS	df	MS	F	Prob > F
Between groups	.479397368	95	.005046288	6.83	0.0683
Within groups	.002217719	3	.00073924		
Total	.481615087	98	.00491444		



5.2 Simulation results

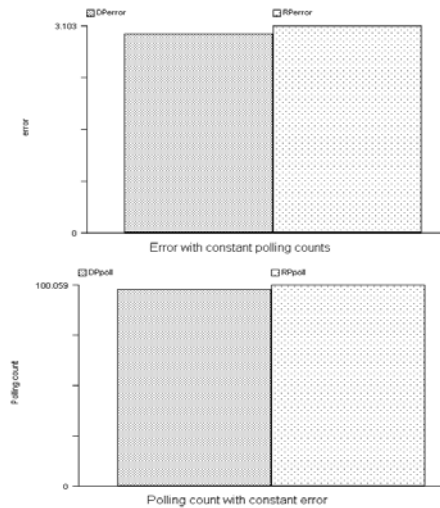
A performance analysis of the proposed scheme is conducted by studying the simulation results for accuracy, efficiency, and the feasibility of the fault findings. After that, a comparison with the random polling scheme was carried out.

In order to compare the performance of the proposed dynamic polling scheme with the random polling scheme, a measure of accuracy is needed, therefore the Analysis of Variance (Anova) procedure is used to evaluate that measurement. The idea of using this is that these two schemes are estimating the same quantity of the network monitoring and should be close in network performance variance. With Anova, the null hypothesis is that the variance of these two schemes is the same. As shown in the table above, the ratio of the “Between groups” variability is $F=6.83$. Thus, the null hypothesis is rejected at the 90% level but not at the 95% level. That is, there is a slight difference (at the 90% level, but not at the 95% level of confidence) between these two schemes. In other words, the proposed scheme can basically reflect the accuracy of the random polling scheme.

In addition, the experimental simulations were run at the fixed polling frequency for both the proposed scheme and the random polling scheme, so as to determine how the variability of the proposed scheme affects the management efficiency. The plot on the right above compares the difference of the numbers of transmitted polling packets under the different monitoring demands. The x-axis is the presented load, and the y-axis is the percentage of differentiation. At the worse cases, which are the cases when the presented loads are very low, less than five percent in the diagram, the proposed scheme produces almost the same number of packets as random polling. Under a medium presented loads between twenty percent and sixty percent loads, the diagram registered up to fifteen percent traffic savings, with a clear trend of increasing traffic savings. Finally, in the case of heavy traffic loads of over eighty percent bandwidth utilization, the application of the proposed scheme results in up to twenty-six percent of traffic savings.

From these results, it can be found that given sufficient system resources, both schemes are functional similar. However, if resources are scarce, or a heavy system load is causing resource contention, the proposed scheme is more efficient in terms of traffic savings.

Simulation results (Cont'd)



Fault time	Mean Diff of RP (ms)	Mean Diff of DP(ms)
25s	.39724	.57933
50s	.80924	.24953
75s	.40115	.31045
100s	.83784	.17162
125s	.83743	.16946
150s	.10576	.22602
175s	.10151	.37832

It is often difficult or impossible to identify the exact fault events in the simulation, in terms of the locations and the durations. Therefore, for the sake of simplicity, this study injects several fault events to check the difference between the proposed scheme and the random polling scheme in terms of performing the task of fault finding.

Furthermore, It is difficult to quantitatively compare the fault findings of the proposed dynamic polling scheme to an equivalent random polling approach when there is a difference in the value of the polling frequency assignment. Therefore, two forms of experiments were conducted to compare these two schemes. First, using a fixed number of polling intervals and then with a fixed number of error injections.

Two charts of relative error versus polling counts on particular link were made for both schemes above on the left. The upper one is the case where the polling count is held constant and the error is measured on the average. The lower plot is the case where the amount of errors is held constant and then the polling count is measured on the average. As the results indicate, the proposed scheme outperforms its corresponding random polling approach by using less polling counts to find more errors.

To represent the task of monitoring, a measurement that collects the time difference in response to the fault event situations is decided upon. As a profiled fault event occurs, information is recorded to gather the time difference between the time when the event occurred and the following exact time of polling. By comparing the time difference, it is possible to deduce the responsiveness of the algorithms to the network behaviours, hence, evaluating the accuracy of the algorithm.

In this study, information is aggregated to summarize the behaviour of fault finding activities. The difference is calculated base on 10 simulation runs. Overall, it can be concluded that the proposed scheme practically achieves enough precision compared with that of the RP approach.

It should be mentioned that it is possible that multiple presented workloads may be loading simultaneously in the network. However, for the sake of simplicity, the simulated fault occurs at one site, thus, the data in the table above can only reflect a single fault finding at a time.

Taken together, the results of the experiments and the analyses confirm that the proposed dynamic polling scheme can play an important role in decreasing the load on the network and on the network operator/manager, increasing the accuracy of the measurements. Thus, the dynamic polling scheme can provide the potential for better monitoring, control, and management of networked systems.

[Discussion]

- Improving of the AHP in real-time decision making
- Flexibility and scalability of the dynamic polling scheme
- Implementation issues
- A trade-off between the reduction of management efforts and the cost of additional computation

6. Discussion

A simulation approach to evaluate the performance of dynamic polling schemes is presented in this study.

The AHP methodology used in the decision making has its strength on solving difficult, complicated problems with multiple criteria. However, for a real time action such as the proposed scheme, it is impractical to use human judgment interactively. Therefore, in this study, pairwise comparisons are conducted by using ratings decision comparisons which seek to obtain relative measurements in real time.

At the most fundamental level, the design of such a dynamic polling scheme can be divided into a hierarchy of levels or layers. Thus, a networked system can be easily matched into the hierarchy of polling frequency model. Therefore, this scheme is flexible and scalable.

At this moment, the comparison and decision making period of the scheme is done off-line in the simulation. However, it is not difficult to improve the usefulness and applicability of the AHP model in real-time decision making. This issue can be solved by querying MIB values during implementation, as each MIB defines the data that it makes accessible to agents and applications via a management protocol.

Finally, it should be emphasized that the reduction of data exchange overhead between managers and agents using the dynamic polling scheme comes at the cost of additional computation by the management nodes. Given that modern network devices have large processing capabilities, this appears to be a reasonable trade-off.

Conclusion and Future Work

- An optimization of network monitoring
- Future work
 - How would this scheme be implemented on real time networked systems

7. Conclusion and Future Work

This paper proposes a new scheme for dynamic polling that considers the polling frequencies required for network management tasks, the overhead of management traffic, and the associated performance impact on the network. The proposed scheme determines the polling frequencies with respect to the impact of the monitored elements on the network behaviour, and the optimization of the required management tasks.

Future work involves issues such as how would this scheme be implemented on real time networked systems.

Reference

- [1] Takano, M., Fujita, K. (1995). "Multilevel Network Management by Means of System Identification". In Proc. of the Third International Telecommunications Symposium.
- [2] Jiao, J., Naqvi, S., Raz, D., Sugla, B. (2000). "Toward efficient Monitoring." IEEE Journal on Selected Areas in Communications (JSAC), special issue on recent advances in network management and operations 18(5): 723 -- 732.
- [3] Gao, F., Gutierrez, J. (2002). "A Trade-off Analysis Model: Towards Polling Efficiency on Network Monitoring Problem". In Proc. of the eighth IEEE International Conference on Communication Systems, Singapore.
- [4] Li, L., Thottan, M., Yao, B., Paul, S. (2003). "Distributed Network Monitoring with Bounded Link Utilization in IP Networks". In Proc. of IEEE INFOCOM.
- [5] Martello, S., Toth, P. (1990). "Knapsack Problems: Algorithms and Computer Implementations", John Wiley & Sons.
- [6] Saaty, T. L. (1994). "The Analytic Hierarchy Process", University of Pittsburgh.
- [7] ns-2 (2001). "the Network Simulator", retrieved Sep., 2001, from www.isi.edu/nsnam/ns/.
- [8] Yoshihara, K., Sugiyama, K., Horiuchi, H., Obana, S. (1999). "Dynamic Polling Scheme Based on Time Variation of Network Management Information Values". In Proc. of Integrated Network Management Symposium (IM'99).
- [9] Abdu, H., Lutfiyya, H., Bauer, M. (2001). "Towards efficient resource allocation in distributed systems management". In Proc. of 15th International of Parallel and Distributed Processing Symposium.
- [10] Hernandez, E., Chidester, M., George, A. (2001). "Adaptive Sampling for Network Management." Journal of Network and Systems Management 9(4).
- [11] Martello, S., Toth, P. (1990). "Knapsack Problems: Algorithms and Computer Implementations", John Wiley & Sons.