A Study on Real-Time Measurement in the Internet

著者 | 小出 和秀
号 | 
学位授与番号 | 

URL: http://hdl.handle.net/10097/37815
論文内容の要旨

1. Introduction
To understand the network dynamics is a matter of great interest from the management, operations, planning and security point of view. Administrators want to provide stable network reachability and service to end-users. They have to monitor their network and look for problems. If a problem is detected, it has to be diagnosed and solved as soon as possible. So they require real-time measurement technologies. Real-time measurement realizes "observation of current value of a particular statistic" effectively. The concept of current will vary in depending on the site and the mission. Practically current should be within several minutes (it will be up to 5 minutes in general). This type of measurement brings out information about abnormal changes in link traffic and node activities.

For achieving real-time measurement, administrators have to continuously collect the most recent data about network and nodes. But the continuous growth of the Internet and the rapid development of various connection technologies make it difficult in many cases. The first problem is that it is not known a priori what is the significant statistic that should be monitored in real-time. This is a major problem particularly when the number of statistics related to traffic is quite massive. It is impossible to monitor all traffic statistics, simply because of the huge number of the statistics. The second problem is the failure in data collection from nodes. SNMP(Simple Network Management Protocol) uses the unreliable UDP protocol as transport. In general wireless connections are prone to disconnection, and the disconnection is not predictable. This leads to failure in data collection from the node when SNMP-like unreliable transport protocol is used. An increased number of failures in the data collection have a disastrous effect on real-time node monitoring mechanisms.
2. Real-Time traffic data summarization techniques

Traffic data summarization is a challenging issue for realizing real-time measurement. Because, except for the routine cases, aggregation parameters, or the flows that need to be observed are not known a priori. Dynamic adaptive aggregation algorithms adapt to the network traffic to detect the important flows. But present day algorithms are inadequate as they often produce inaccurate or meaningless aggregation. We propose a Dynamic Constrained Adaptive Aggregation algorithm that produces meaningful aggregation by using information about the network's configuration. We compare the performance of this algorithm with the erstwhile Dynamic (Unconstrained) Adaptive Aggregation algorithm and show its efficacy.

The tool called Aguri implements the traffic data summarization algorithm. IP addresses with traffic amount less than threshold will be aggregated into the parent node in the IP address tree. This method has an unconstrained aggregation strategy. It generates addresses that do not always represent logical network entities. More intelligent algorithms will be needed for effective, dynamic information aggregation.

We propose the 'DCAA' (Dynamic Constrained Adaptive Aggregation) strategy. The core of DCAA is to controlling aggregation in order to keep correctness of the generated outputs when identifying significant flows adaptively. IP addresses whose prefix exists as network/sub-network in the real Internet world are defined as SBlock (Significant Block) nodes. An ABlock (Associated Block) characterizes a logical group of SBlocks that are closely related, for example belonging to the same administrative domain. We propose to constrain the IP address based aggregation process explained above using SBlock and ABlock information. In essence, the idea is to cease node aggregation at a SBlock. Here, nodes along the arc from the root to a SBlock and, belonging to a different ABlock each other will be called as non-SBlock.

We aggregated traffic traces using Aguri and proposed method, DCAA. We checked the correctness of results of them. We counted the number of IP addresses in the trace that aggregated to prefixes shorter than that of the corresponding SBlock(non-SBlock) and used this count of 'Error' as a measure of correctness. The greater the Error count, the lower in the correctness. Figure 1 shows the correctness for different days. The left column shows the total number of IP addresses in the base data, the middle column shows the count of Error by Aguri, and the right column shows the count of Error by DCAA. In Aguri case, almost all of the traffic got aggregated into Error whereas in DCAA the traffic that got aggregated into Error was much less. From these results, it is apparent that DCAA seems to be able to maintain the correctness of outputs.
3. Real-Time tracing mechanisms of link failure in the Internet

It is difficult to analyze the pathologies of a network. Active probing e.g. using tools like 'ping' and 'traceroute' have limitations. These are constrained by policy restrictions of local and remote networks. On the other hand, passive observation has its merits - it is not constrained by the policy of external networks and does not generate traffic. But passive observation is generally limited by its scope - a monitor's view is limited to the traffic that is 'seen' at the observed link. We propose a mechanism which is passive and yet has a wide range of coverage. We focus on TCP retransmissions and show that passively monitored data can be analyzed using a tree-based fault-tracing method in conjunction with packet path information to tell us about the faults of remote network links in (near) real-time. We examine the related algorithms and discuss the results of our experiments on real networks.

In our approach, retransmission detection is carried out on micro flows but, analysis is carried out on aggregated flows, called torrents. Each torrent will contain more than 1 micro flows. We extend the definition of retransmission flow and normal flow to torrents. If at least one micro flow contains no retransmitted packet (that is, normal flow) in the torrent destined to the aggregated endpoint, the torrent is a normal torrent. If every micro flow in the torrent is a retransmission flow, the torrent is a retransmission torrent, that is, eventful torrent.
Our approach is similar to 'Network Tomography'. They select a 'back-to-back' pair of packets sent to two different receivers from the passively monitored traffic. It then attempts to trace the point at which one of the packets is lost. A drawback of the passive tomography approach is that the passively monitored traffic may not always have candidate packets for analysis. Further, many packets are lost due to non-network related problems e.g. due to heavily-loaded or non-existent hosts. Our proposed method can eliminate non-network related problems from traffic data, and we effectively detect network fault.

Figure 2 is the flow-annotated network map generated by prototype system. Each circle represents an AS and the arc between ASes represents a link between ASes. The observation point in this case is inside the AS named TOPIC. TOPIC's ISP, SINET has some external links and one of them is from KDD. We infer that under normal circumstances there is traffic between SINET-KDD link as well as between other ASes, ODN or WIDE. At some point a change was observed in the network map. The corresponding image is shown in Figure 2. It appears that one of the neighboring ASes namely SINET was experiencing some problem at that time. It reveals that the component of the flow between the ASes SINET and KDD disappeared. The other components continued as before. From these visual images, it seems that there was a network problem on the link between SINET and KDD while the other external links of SINET functioned normally. This inference corresponded well with the facts in this case. This is one example of the potential to detect network faults.

**Figure 2** Visualization result at network failure period
4. Real-Time data collection on wireless environment

The advent of mobile IP communication has opened up several new areas of mission critical communication applications. But the bandwidth and reliability constraints coupled with handover latency are posing some hurdles which need to be overcome before real world MobileIP applications, with low tolerance for data loss, can be deployed. We analyze the unreliability of existing data collection methods in the real-world wireless MobileIP environment. We focus on this problem and propose a novel monitoring model that anticipates the wireless mobile nodes and uses SNMP. The key idea of this model is the introduction of a store-and-forward type Managed Object (MO). During the unreachable period between the SNMP-manager and the SNMP-agent, the data is cached at the agent until the connectivity recovers. In our experiment we used a prototype implementation in real-world wireless communication field, and showed the effectiveness of our proposed method.

The existing standard remote management/monitoring model is the agent-manager model adapted in SNMP network management framework. This model is simple and flexible but unreliable. If the connectivity to the agents goes down for some reason, the information collection by managers will fail and there is no way to recover the lost data. We propose a robust information collection model for an unreliable communication environment. Here we adapt a store-and-forward information collection strategy. The innovative point of this model is that an agent contains three new components, Data Collector, Data Storage and Forwarding Controller. Manager can access to data source directly, and also can fetch the stored data when the agent is reachable because the stored data contains time-stamp.

We deploy the SNMP agent with our proposed system on the car-computer with wireless link. When the car moves, the location information will be updated and can be remotely monitored using SNMP. We collected normal polling data from the probe computer on the car for nearly 1 month. It shows that in normal cases a considerable amount of data is lost. On the other hand Figure 3 shows the loss rate of polling with proposed method. We make an experiment of polling with normal polling and recovery method using proposed system in parallel. Our system succeeded to recover almost half of failure cases' data. In the latter part of Figure 3 (17:30:00 onwards) shows our method's remarkable robustness in recovering lost data, when the car passes through the tunnel, i.e. for a certain period of time without connectivity. This results clearly shows that our proposed model achieved sufficient improvement in reducing polling loss rate in the actual real-world situation.
5. Conclusion

We developed methods for real-time measurement in the Internet. These will help network administrators to understand the network dynamics, and enable early detection of network problems. That will lead to quick diagnosis and solutions to network problems. It will lead to more stable and effective network management, operations, planning and security. This research will contribute to the realization of more stable Internet for many users.
論文審査結果の要旨

インターネットの安定した管理・運用のためには、管理者が障害などの問題を出来るだけ早期に発見し解決する事が重要である。このため、ネットワークの観測を継続的に行い、最大でも5分程度の遅延で観測結果を処理し異常を検知する、リアルタイム計測の実現が必要となる。しかしながら、インターネットにおけるトラフィックの多様化や無線技術の普及に伴い、既存技術ではリアルタイム計測を実現できない状況が増えている。そこで著者は、インターネットにおけるリアルタイム計測を実現する新しい観測方式、および観測結果の処理方式に関する詳細な研究を行った。本論文はその成果をまとめたものであり、全編5章からなる。

第1章は序論である。
第2章では、IPアドレスの省略と併合、すなわちアドレス集約において適切な粒度の集約を保証するため、IPアドレスに関する属性を利用して集約を行う手法を考案し、さらに同手法を用いて、リアルタイム計測の対象とする多様なトラフィックを処理し、管理上重要な意味を持つトラフィックパターンを抽出する方式を提案している。そして本方式が、適切なアドレス集約を行うために有効であることを示している。

第3章では、第2章で提案した処理方式を応用し、インターネットの障害をリアルタイムに検出する方式を提案している。本方式は再送パケットを含むTCPフローの宛先IPアドレスを集約することにより、検出ミスが少なく、かつローカルネットワークでの計測のみで実現できることが特徴である。これは、多様なトラフィックが流れるインターネットのリアルタイム計測を効率的に実現するための重要な成果である。

第4章では、ノードのリアルタイム計測を対象とし、現在位置や通信状況などのノードの状態情報を、不安定な通信環境においても継続的に観測可能とする方式を提案している。この方式の基本は、通信の切断中にノードから取得できなかった状態情報を観測対象ノード内に保持するバッファを設け、これを用いて通信回復後に状態情報を再取得可能とする機構を構成したことである。また実証試験により、実環境での本方式の有効性を確認している。これは、無線接続において一時切断が頻発する接続リンクのリアルタイム計測を実現するための有用な成果である。

第5章は結論である。
以上要するに本論文は、インターネットにおけるリアルタイム計測方式の提案・実装・評価を行い、その有効性を示すことにより、インターネットの安定した管理・運用のための基礎を与えたものであり、情報基礎科学の発展に寄与するところが少なくなってある。
よって、本論文は博士（情報科学）の学位論文として合格と認める。