US007782793B2

(12) **United States Patent**
Olesinski et al.

(10) **Patent No.: US 7,782,793 B2**
(45) **Date of Patent: Aug. 24, 2010**

(54) **STATISTICAL TRACE-BASED METHODS FOR REAL-TIME TRAFFIC CLASSIFICATION**

(75) Inventors: **Wladyslaw Olesinski**, Kanata (CA); **Peter Rabinovitch**, Kanata (CA)

(73) Assignee: **Alcatel Lucent**, Paris (FR)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 899 days.

(21) Appl. No.: **11/226,328**

(22) Filed: **Sep. 15, 2005**

(65) **Prior Publication Data**

US 2007/0076606 A1 Apr. 5, 2007

(51) **Int. Cl.**
H04L 12/26 (2006.01)
H04L 12/56 (2006.01)
H04L 12/28 (2006.01)

(52) **U.S. Cl.** **370/253**; 370/391; 370/395.43

(58) **Field of Classification Search** 370/252
See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

6,873,600	B1 *	3/2005	Duffield et al.	370/252
7,080,136	B2 *	7/2006	Duffield et al.	709/223
7,286,535	B2 *	10/2007	Ishikawa et al.	370/392
7,313,100	B1 *	12/2007	Turner et al.	370/253
7,376,085	B2 *	5/2008	Yazaki et al.	370/235
7,376,731	B2 *	5/2008	Khan et al.	709/224
2003/0012197	A1 *	1/2003	Yazaki et al.	370/392
2007/0214504	A1 *	9/2007	Comparetti et al.	726/23

FOREIGN PATENT DOCUMENTS

WO	WO 96/38955	12/1996

OTHER PUBLICATIONS

Sun et al., "Statistical Identification of Encrypted Web Browsing Traffic", Microsoft Research for Proc. IEEE Symposium on Security and Privacy, IEEE, May 2002.*

Zhang et al., "Detecting Backdoors", Proceedings of the 9th USENIX Security Symposium Denver, Colorado, Aug. 2000, p. 1-11.*

M. Roughan, S. Sen, O. Spatscheck, N. Duffield, "Class-of-service mapping for QoS: a statistical signature-based approach to IP traffic classification", 2004, pp. 135-148.

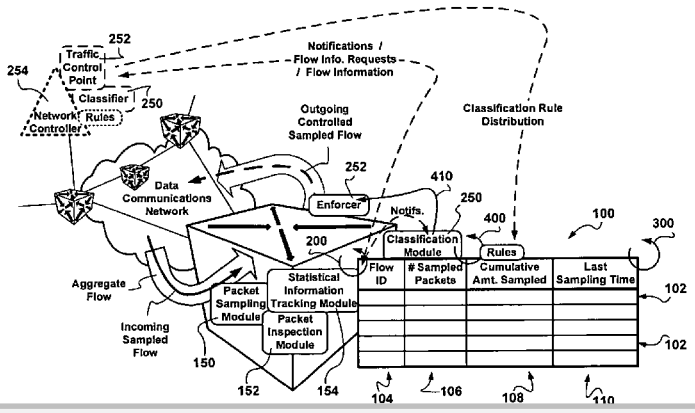
(Continued)

Primary Examiner—Daniel J. Ryman
Assistant Examiner—Cassandra Decker
(74) *Attorney, Agent, or Firm*—Kramer & Amado P.C.

(57) **ABSTRACT**

Apparatus and methods for real-time traffic classification based on off-line determined traffic classification rules are provided. Traces of real traffic are obtained and subjected to statistical analysis. The statistical analysis identifies the multidimensional domain space of characteristic traffic parameters. Classification rules associated with the identified domains are derived and provided to traffic classification points for real-time traffic classification. Traffic classification points, typically edge network nodes, sample packets in aggregate streams with a predetermined probability. Statistical information regarding the sampled flows is tracked in a table, the number of time a flow was sampled providing a probabilistic measure of the flow's duration before the flow terminates. The table entries, which predominantly track high bandwidth flows, are subjected to the classification rules for real-time classification of the sampled flows. Optionally, rules include an action to be taken in respect of flows having characteristics matching thereof. Advantages are derived from low overhead on-line real-time classification of high-bandwidth flows at low overheads before flow termination.

32 Claims, 2 Drawing Sheets



OTHER PUBLICATIONS

Konstantinos Psounis, Arpita Ghosh, and Balaji Prabhakar: "SIFT: A Low-complexity Scheduler for Reducing Flow Delays in the Internet", 2004, pp. 1-13.

Duffield N et al.: "Estimating Flow Distributions From Sampled Flow Statistics", vol. 33, No. 4, October 20036, pp. 325-336.

Karagiannis, T., et al., Transport Layer Indentification of P2P Traffic, ACM, 2004.

Prabhakar, B., Network Processor Algorithms: Design and Analysis, Stochastic Networks Conference, 2004.

Roughan, M., et al., Class-of-Service Mapping for QoS: A Statistical Signature-Based Approach to IP Traffic Classification, ACM, 2004.

Sen, S., et al., Accurate, Scalable, In-Network Identification of P2P Traffice Using Application Signatures, ACM, 2004.

* cited by examiner

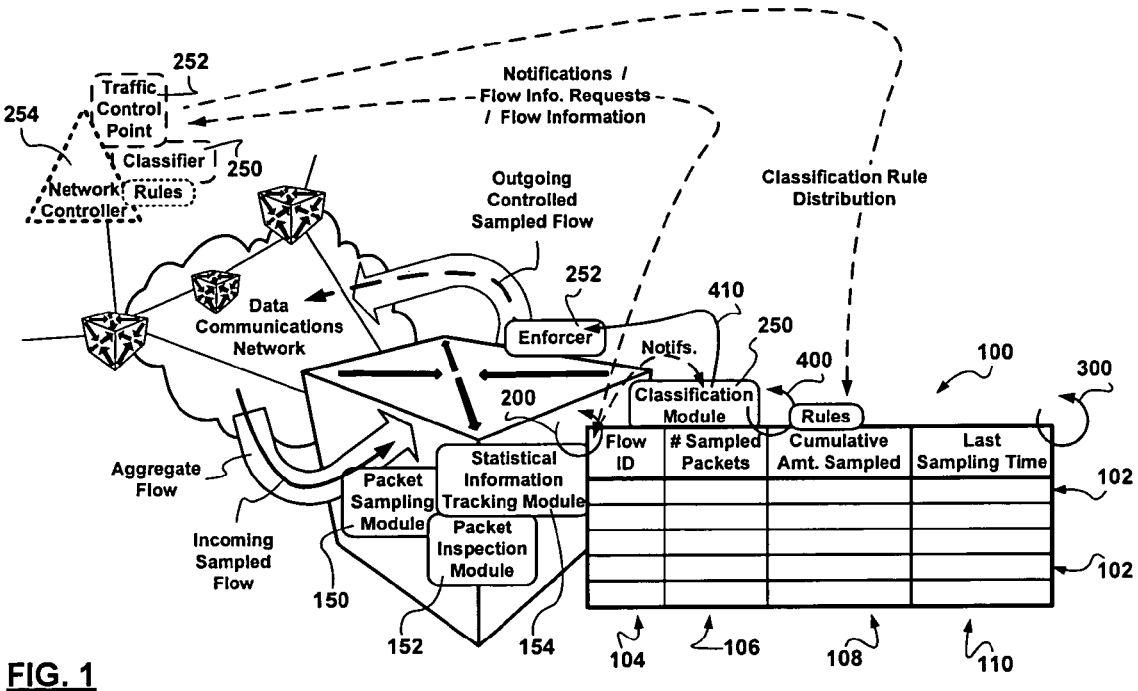


FIG. 1

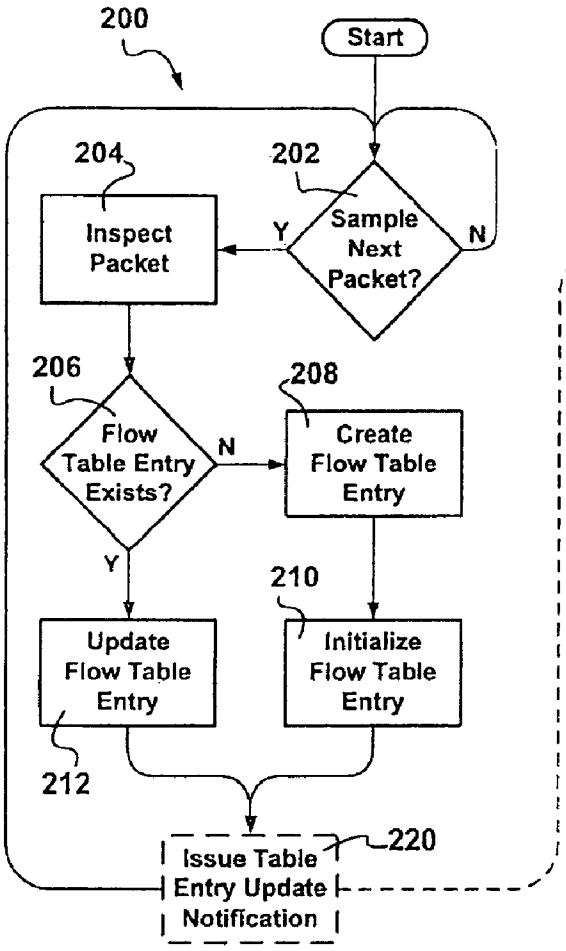


FIG. 2

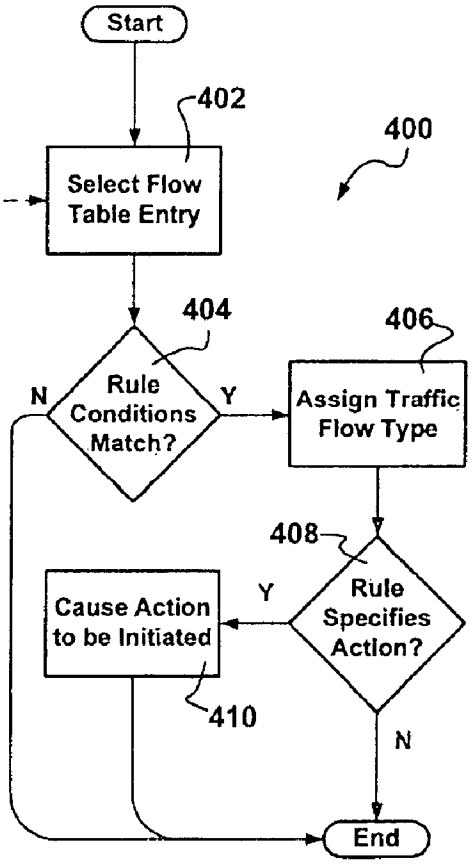


FIG. 4

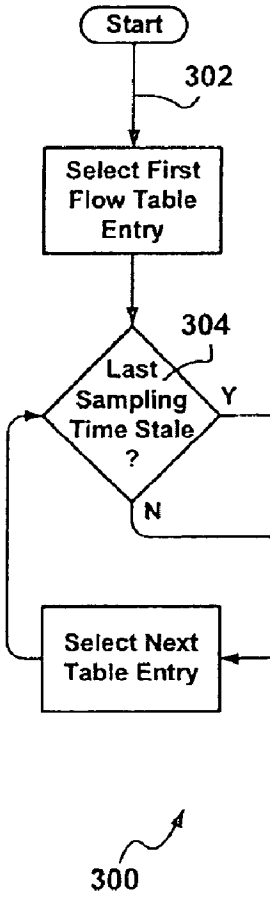


FIG.

STATISTICAL TRACE-BASED METHODS FOR REAL-TIME TRAFFIC CLASSIFICATION

FIELD OF THE INVENTION

The invention relates to content delivery at the edge of communications networks, and in particular methods and apparatus providing real-time trace-based traffic classification.

BACKGROUND OF THE INVENTION

Traffic classification is important for many reasons in delivering content to customers at the edge of communications networks. For example, Quality of Service (QoS) requires the traffic to be segregated first in order to assign packets to particular Classes of Service (CoS). A network operator can provide a different level of service to each class as well as a pricing structure.

Knowledge of traffic characteristics can help optimize the usage of the communications network infrastructure employed, and can help ensure a desired level of performance for applications/services important to the customers. The intention has always been that application requirements be considered in offering a level of service. Traditional methods of traffic detection and classification rely on monitoring logical port specifications typically carried in packet headers as, in the past, applications and/or services were, in a sense, assigned well known logical ports.

A large percentage of the traffic conveyed by communications networks today consists of peer-to-peer (P2P) traffic. Because peer-to-peer traffic is conveyed between pairs of customer network nodes, it is not necessary that a well known logical port be allocated, reserved, and assigned to traffic generated by applications generating peer-to-peer traffic and/or applications retrieving peer-to-peer content. Therefore known approaches to traffic classification are no longer valid as logical ports are undefined for peer-to-peer applications and/or logical ports may be dynamically allocated as needed such in the case of the standard File Transfer Protocol (FTP) and others.

Peer-to-peer content exchange techniques are increasingly being used to convey without permission content subject to intellectual property protection, such as music and movies. Network operators are under an increasing regulatory pressure to detect peer-to-peer traffic and to control illicit peer-to-peer traffic, while rogue users are seeking ways to defy traffic classification to avoid detection.

Besides peer-to-peer traffic detection, means and methods are being sought on a continual basis for detecting short duration traffic flows to help identify possible intrusions such as, but not limited to, Denial of Service (DOS) attacks.

Statistical billing is another domain in which knowledge of traffic characteristics is necessary. Network operators increasingly employ resource utilization measurements as a component in determining customer charges.

Returning to peer-to-peer traffic detection, not all peer-to-peer traffic is illicit: in view of the high levels of resource utilization demanded by peer-to-peer traffic, network operators may want to charge customers generating peer-to-peer traffic and retrieving peer-to-peer content more for their high bandwidth usage. Resource utilization alone is not always an adequate traffic characteristic differentiator as in many instances content conveyed to, and received from, multiple

Attempts to characterize traffic, to detect traffic types, with a view of classifying traffic, include Deep Packet Inspection (DPI) techniques. Deep packet inspection techniques are described by Sen S., Spatscheck O. and Wang D. in "Accurate, Scalable In-Network Identification of P2P Traffic Using Application Signatures", Proceedings of the 13th international conference on World Wide Web, New York, N.Y., 2004; and by Karagiannis T., Broido A., Faloutsos M., Claffy K. in "Transport layer identification of P2P traffic", Proceedings of the 4th ACM SIGCOMM Conference on Internet Measurement, Taormina, Sicily, Italy, 2004.

Proposed deep packet inspection techniques, as the name suggests, assume the availability of unlimited resources to inspect entire packets to the perform packet characterization. Therefore deep packet inspection incurs high processing overheads and is subject to high costs. Deep packet inspection also suffers from a complexity associated with the requirement of inspecting packet payloads at high line rates. For certainty, deep packet inspection is not suited at all for typical high throughput communications network nodes deployed in current communications networks. Deep packet inspection also suffers from a high maintenance overhead as the detection techniques rely on signatures, peer-to-peer applications, especially, are known for concealing their identities—a deep packet inspection detection signature that provides conclusive detection now may not work in the future, and another conclusive signature would have to be found and coded therein.

Traffic classification means and methods are being actively sought by network operators in order to determine the types of traffic present in a managed communications network for traffic and network engineering purposes, on-line marking of packets, quality of service assessment/assurance, billing, etc. In view of impending regulatory pressures, efficient detection and classification of peer-to-peer traffic is especially desired, as peer-to-peer traffic consumes large, disproportional percentages of bandwidth and other communication network resources. Network operators have to employ a combination of: peer-to-peer traffic control in order to reserve network resources for other types of traffic, charge peer-to-peer users different rates to curb behavior, and/or even block peer-to-peer completely in accordance with regulations imposed on network operators. There therefore is a need to solve the above mentioned issues to provide traffic classification means and methods which avoid the complexities of deep packet inspection and the pitfalls of logical port based packet classification.

SUMMARY OF THE INVENTION

In accordance with an aspect of the invention, a packet flow classification apparatus for on-line real-time traffic flow classification at a communications network node is provided. Packet sampling means randomly selects packets from an aggregate flow with a pre-determined sampling probability. Packet inspection means determines the packet size of each sampled packet and obtains the flow identification of the sampled traffic flow with which the sampled packet is associated. A sampled flow information table has sampled flow table entries for storing real-time sampled flow statistical information. Flow information tracking means maintain the flow information table in real-time. And, a packet classifier classifies sampled traffic flows on-line in real-time based on a group of classification rules trained off-line on statistical trace traffic flow information.

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS
Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS
Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS
Sync your system to PACER to automate legal marketing.