Thèse de Doctorat de l'université Paris VI
Pierre et Marie CURIE

Spécialité

SYSTÈMES INFORMATIQUES

présentée par

M. Artur ZIVIANI

pour obtenir le grade de

DOCTEUR de l'université Pierre et Marie CURIE

Qualité de Service et Localisation d'Hôtes

Présentée et soutenue publiquement le 8 décembre 2003 devant le jury composé de

M. Bernd E. WOLFINGER	**Président**
M. Walid DABBOUS	**Rapporteur**
M. Andrzej DUDA	**Rapporteur**
M. Mark CROVELLA	**Examinateur**
M. Éric HORLAIT	**Examinateur**
M. Otto Carlos M. B. DUARTE	**Directeur de thèse**
M. Serge FDIDA	**Directeur de thèse**
M. José F. de REZENDE	**Directeur de thèse**

Thèse de Doctorat de l'université Paris VI
Pierre et Marie CURIE

Spécialité

SYSTÈMES INFORMATIQUES

présentée par

M. Artur ZIVIANI

pour obtenir le grade de

DOCTEUR de l'université Pierre et Marie CURIE

Qualité de Service et Localisation d'Hôtes

Présentée et soutenue publiquement le 8 décembre 2003 devant le jury composé de

M. Bernd E. WOLFINGER	**Président**
M. Walid DABBOUS	**Rapporteur**
M. Andrzej DUDA	**Rapporteur**
M. Mark CROVELLA	**Examinateur**
M. Éric HORLAIT	**Examinateur**
M. Otto Carlos M. B. DUARTE	**Directeur de thèse**
M. Serge FDIDA	**Directeur de thèse**
M. José F. de REZENDE	**Directeur de thèse**

08/12/2003

A mes parents

et

à ma chérie Elisa.

Remerciements

Cette thèse, intitulée "Qualité de Service et Localisation d'Hôtes", a été soutenue le 8 décembre 2003 à l'université Pierre et Marie Curie (PARIS VI) devant le jury composé de M. Walid DABBOUS, M. Andrzej DUDA, M. Éric HORLAIT, M. Mark CROVELLA, M. Bernd E. WOLFINGER, M. Serge FDIDA, M. José F. de REZENDE et M. Otto Carlos M. B. DUARTE. Je tiens à remercier tous les membres de ce jury :

M. Walid DABBOUS, Chercheur à l'INRIA (Institut National de Recherche en Informatique et en Automatique), et M. Andrzej DUDA, Professeur à l'Ensimag (École Nationale Supérieure d'Informatique et de Mathématiques Appliquées de Grenoble), qui ont accepté d'être rapporteurs de ma thèse.

M. Éric HORLAIT, Professeur à l'université Pierre et Marie Curie (PARIS VI), et M. Mark CROVELLA, Professeur à l'université de Boston, États-Unis, pour avoir l'amabilité de participer à ce jury.

M. Bernd E. WOLFINGER, Professeur à l'université de Hambourg, Allemagne, pour son dévouement et son soutien dans notre travail de recherche coopératif qui compose le Chapitre 4 de cette thèse.

M. Serge FDIDA, Professeur à l'université Pierre et Marie Curie (PARIS VI), M. José F. de REZENDE et M. Otto Carlos M. B. DUARTE, Professeurs à l'université Fédérale de Rio de Janeiro (UFRJ), Brésil, qui ont dirigé mes travaux de recherche et m'ont soutenu pendant ces trois années.

Je tiens aussi à remercier la CAPES/Brésil pour le support financier de cette thèse.

Je voudrais exprimer ma gratitude envers les membres de l'équipe Réseaux et Performance du LIP6 et les amis que j'ai pu faire à Paris pour leur support au long de ces trois ans de travail.

Je remercie également mes parents qui m'ont toujours soutenu dans toutes les entreprises de ma vie. Finalement, j'aimerais beaucoup remercier Elisa Consalter qui m'a tendrement soutenu de très loin au début et de très près depuis récemment.

Résumé

Cette thèse se concentre sur deux problèmes : (i) le support de Qualité de Service (QdS) aux trafics de voix et de vidéo ; et (ii) la localisation d'hôtes dans les réseaux ad hoc mobiles et dans l'Internet.

L'Internet, basé sur le service "au mieux" (*best-effort*), n'offre aucune différenciation parmi les paquets. Ainsi, la qualité perçue par l'utilisateur d'une transmission audio ou vidéo peut être fortement dégradée. Dans ce travail, on évalue le support fourni par l'architecture de différenciation de services aux transmissions de voix et vidéo dont la besoin ne cesse de croître. Les résultats montrent que le mécanisme de différenciation de services adopté améliore la qualité de la distribution des flux de voix et de vidéo. Nous évaluons l'efficacité démontrée par différents modèles de trafic de voix en utilisant une allocation additionnelle de bande passante et l'impacte du *shaping* du trafic de voix. Nous démontrons aussi que l'utilisation conjointe d'une protection inégale des trames du flux vidéo et de la tolérance aux pertes de paquets s'avère une stratégie efficace pour la transmission de qualité de la vidéo MPEG.

La connaissance de la localisation physique d'un hôte basée sur son identifiant ouvre la voie pour une classe toute entière d'applications exploitant cette information. Nous proposons, pour les réseaux ad hoc mobiles, d'utiliser les premiers paquets d'un flux pour apprendre la position de son destinataire au lieu d'utiliser un paquet de recherche dédié. Nous démontrons que cette approche réduit les coûts de transmission comparés à ceux des approches traditionnelles pour les flux TCP. Nous avons aussi évalué la performance, dans le contexte de l'Internet, des procédures de localisation géographique d'un hôte basées sur des mesures. Ces mesures sont réalisées par des machines sondes vers l'hôte à localiser ainsi que vers des hôtes de références qui ont une localisation géographique connue. Nous proposons et évaluons une politique démographique de placement des hôtes de référence et des machines sondes. Nous évaluons aussi les modèles de similarité qui comparent les profils de délai pour estimer la localisation. Nos résultats démontrent que le placement stratégique des hôtes de références et des machines sondes ainsi que le choix adéquat du modèle de similarité à adopter contribuent à une estimation de localisation géographique plus précise des hôtes Internet.

Mots-clés :

Réseaux, qualité de service, trafic multimédia, connaissance de la localisation.

Abstract

This thesis, entitled "Quality of Service and Location-Awareness," focuses on two problems: (i) the Quality of Service (QoS) support for voice and video traffic in the Internet; and (ii) the location-awareness in both ad hoc networks and the Internet.

The best-effort service of the conventional Internet offers no differentiation among packets. Therefore, the quality perceived by the users of an audio or video transmission may be strongly degraded. In this work, we evaluate the support provided by the differentiated services architecture to transmissions of voice and video, which have an increasing demand. Results show that the adopted differentiated services mechanisms improve the delivery quality of the voice and video flows. We evaluate the performance of different models for voice traffic when using an extra allocated bandwidth and the impact of traffic shaping. We also show that the joint adoption of an unequal protection of frames and the tolerance to packet losses is an efficient strategy to improve the delivery quality of MPEG video streams.

Knowing the physical location of a host from an identifier of that host enables a whole new class of location-aware applications. Examples of these novel location-aware applications are the restricted content delivery based on location, the geographic routing of packets, and targeted advertising. In the case of mobile ad hoc networks, we propose to use the initial packet of the flow to learn the location of the destination instead of adopting dedicated location query packets. Results show that the proposed approach reduces the transmission costs in location management for the TCP flows with respect to the conventional schemes. In the context of the Internet, we focus on the measurement-based geographic location of an Internet host. These measurements are performed by probe machines toward landmarks, which are hosts with a known geographic location, and the host to be located. We propose and evaluate a demographic placement policy for landmarks and probe machines. We also assess similarity models that compare the measured delay patterns to indicate the location estimation. Our findings show that strategically placing landmarks and probe machines as well as using an adequate similarity model improve the accuracy of the location estimation in the measurement-based geographic location of Internet hosts.

Key Words:

Computer networks, quality of service, multimedia traffic, location-awareness.

Table of contents

Chapter 1

Introduction

The Internet is currently one of the most important infrastructures of the global information society. The growth of the Internet in the last decade has attracted a large number of new users that demand unforeseen new services. As a consequence, design decisions taken to meet the requirements of these new services no longer affect only a small technical community, but they can impact the public in general. These decisions, associated with social and economic factors, may influence the future of the telecommunication infrastructure. The knowledge of the limitations of the current Internet justifies the research concerning the provision of new services that are able to improve its functionality.

1.1 Motivation and Problem Statement

The Internet is a packet switched network that has been originally conceived to be as simple as possible at its core, pushing the complexity to the end points. Packets are forwarded at each intermediate router based on the destination address they carry and following a best-effort service. The best-effort service implies that every packet is handled by the network without any guarantees of bounded delay or limited packet losses. Therefore, packets from applications that tolerate some packet losses, delay, or jitter compete for resources with packets from applications sensitive to such parameters. Most current applications have been designed to run using the best-effort service. Nevertheless, new applications may have additional needs such as bounded delay and available bandwidth. Such needs are often referred to as Quality of Service (QoS) requirements. Examples of this category are real-time and multimedia applications. New functionalities to support QoS may be implemented at the end systems, either at the application layer or at a sub-layer between the network level and the application layer, but they can also benefit from new services deployed at the network layer of the Internet.

In general, network functionalities such as routing or service provision are decoupled from the geographic position of the network components. For example, nodes in an ad hoc network or routers in the Internet forward a packet based only on an address that identifies the packet destination. Traditional services are thus conceived disregarding the geographic location of the concerned participants. However, a step further in the provision of new services is the knowledge of the geographic position of nodes. Knowing the geographic position of a node from an identifier of that node allows a whole new class of location-aware applications. Nevertheless, if we take the Internet as an example, no direct relationship exists between an identifier of a host, such as a DNS name or an IP address, and the geographic location of that host. Therefore, a location service that maps host identifiers into host locations is of fundamental importance to location-aware applications.

We investigate in this thesis two problems: (i) the QoS support for voice and video traffic in the Internet; and (ii) the location-awareness in both ad hoc networks and the Internet.

A scenario that takes advantage of new functionalities concerning both QoS support and location-awareness in the Internet is described as follows. A content delivery network (CDN) replicates the pages of a Web site in a set of servers that are geographically dispersed, caching the contents of the pages. When a user requests some content from a CDN, the CDN redirects the request from the originating server to a server in the CDN that is the closest to the user and this server delivers the cached content to the user. This service is effective in speeding the content delivery of Web sites with high traffic that have global reach. The closer the CDN server is to the user, the faster the content is delivered to the user. CDNs also provide protection from large surges in traffic to a single server. If a CDN makes multimedia content available to its users, the delivery of this content is likely to demand a certain level of QoS to satisfy the end user. Furthermore, knowing the geographic location of the user allows a CDN to better select the server that will send the content to the user as well as eventually adapt the content to the user's location.

1.2 Contributions

In this work, we study different techniques that involve the QoS delivery of multimedia content and the location-awareness in networks, providing contributions in these two areas.

1.2.1 Quality of Service Support of Voice and Video

Under a best-effort service, such as the basic service of the Internet, multimedia packets, i.e. IP packets carrying voice or video traffic, compete head to head with data packets for link bandwidth. Data applications use TCP, which cuts back the transmission rate when packets are discarded at a router because of congestion, and most multimedia applications use UDP, which does not cut back transmission rates. This indiscriminate treatment of packets from applications with different requirements in terms of bounded delay, limited packet losses, or assured bandwidth, leads to a degraded service. Distinct applications present different QoS requirements and, as a consequence, specific solutions have different degrees of complexity. Therefore, to provide a generic solution for QoS in the Internet is a very difficult task. Nevertheless, the careful design of a QoS architecture may provide support for both voice and video traffic, taking into account bandwidth restrictions. Such restrictions can be due to economical issues or inherent limitations of the adopted technology (e.g. wireless networks). In this work, we investigate the QoS support for both voice and video traffic.

We aim at evaluating the delay and jitter experienced by voice traffic when handled by the Expedited Forwarding (EF) scheme of the Differentiated Services architecture. The analysis includes the effects of different packet scheduling mechanisms implementing EF and of the voice packet size. We also evaluate how efficiently each type of traffic uses an extra allocated bandwidth and the impact of traffic shaping. The results show that increasing the service rate share allocated to the EF aggregate does not significantly affect the competing best-effort (BE) traffic. This holds as long as the BE traffic can use the bandwidth left unused by the EF traffic during idle periods.

We also investigate different QoS schemes and their tradeoffs in improving the delivery quality of an MPEG video stream. Indiscriminate packet discards strongly degrade the quality perceived by end users of MPEG transmissions. From an analytical model, we evaluate the impact of frame losses on the quality of MPEG streams and on the waste of network resources. Our assessment considers issues such as the use of redundancy by applying a Forward Error Correction (FEC) scheme to tolerate losses, the changing of the compression factor in MPEG encoding, the unequal protection of MPEG frames in a Differentiated Services environment, and how to evaluate the impact of network losses onto application quality. Results provide predicted bounds on the quality to be expected by end users as well as guidelines on how to take the best advantage from the combined adoption of the investigated QoS schemes.

1.2.2 Location-Awareness

Location-aware applications depend on an efficient location service to provide the position of nodes. Different contexts, such as an ad hoc network or the Internet, have particular characteristics and thus require distinct solutions to the deployment of this location service.

In the context of ad hoc networks, we propose a TCP-tailored approach to the location management issue. The main location management proposals in mobile ad hoc networks have, as a common characteristic, two distinct phases: the location query of the position of a destination node and the transmission of a flow toward the destination node. We propose to send the initial packet of a flow to learn the position of its destination instead of adopting a dedicated query packet. Such an approach specially benefits TCP flows. This TCP-tailored approach can be applied to previous proposals in location management with minor changes in their particular features. We show that the proposed TCP-tailored approach reduces the cost of location management for TCP flows with respect to the conventional schemes. Furthermore, battery consumption also decreases as the location management costs decrease.

We also consider the measurement-based geographic location of Internet hosts. Location-aware applications are based on the knowledge of the geographic location of users and thereby can offer novel functionalities in the Internet. We focus on improving the accuracy of a geographic location service that relies on delay measurements to locate Internet hosts. Host locations are inferred by comparing delay patterns of geographically distributed landmarks, which are hosts with a known geographic location, with the delay pattern of the target host to be located. We deal with two problems that influence the accuracy of the resulting location estimation: (i) the placement of the landmarks and the probe machines that perform the delay measurements; and (ii) how to best measure the similarity between the delay patterns of the landmarks and the one observed for the target host. For the landmark placement problem, we propose a demographic approach to improve the representativeness of each landmark with respect to the hosts to be located. Given a limited number of landmarks, results show that a demographic placement provides closer landmarks and more accurate location estimations for most hosts. Concerning the placement of probe machines, we show that they have to be sparsely placed to avoid gathering redundant data. Furthermore, we define and evaluate three similarity models. Experiments show that the commonly adopted Euclidean distance is outperformed by other similarity models, resulting then in a more accurate geographic location of Internet hosts.

1.3 Document Outline

This thesis is organized in two parts as follows.

The first part of this thesis investigates the QoS support to voice and video traffic. Chapter 2 provides a brief overview of the main proposals for QoS support in the Internet. We identify the characteristics of these proposals that are relevant to the support of voice and video traffic. Chapter 3 analyzes the support provided by the Expedited Forwarding (EF) service of the Differentiated Services architecture to voice traffic, considering different models for voice traffic as well as different scheduling policies to implement the EF service. In Chapter 4, we investigate the joint adoption of different QoS schemes to support the quality delivery of MPEG video streams. This investigation comprises the application of redundancy within the video stream and the unequal protection of video frames.

The second part of this thesis addresses problems related to the location of a host from an identifier of this host. In Chapter 5, we present the location-awareness problem in networks, elaborating on the motivation behind it. We formalize the location problem with a general view that comprises its applicability on different contexts. We also briefly review the main proposals of location services in both ad hoc networks and the Internet. In the case of ad hoc networks, we propose in Chapter 6 to use the first packet of a flow to learn the position of the destination node instead of adopting a dedicated query packet. This approach can be applied to most previous schemes of location management in ad hoc networks. We show that the proposed approach reduces the costs for TCP flows with respect to the conventional approach. Chapter 7 concerns the Internet context, where we focus on the measurement-based geographic location of an Internet host. The measures are carried out by probe machines toward landmarks, which are hosts with a known geographic location, and toward the host to be located. We address the problem of placing the landmarks and probe machines considering the distribution of users to be located. We also investigate similarity models to be used in the selection of the landmark that provides the location estimation of the host to be located. Both the placement and similarity issues are key points in the accuracy of the location estimation.

1.4 Résumé du Chapitre

L'Internet constitue actuellement une des structures de base de la société d'information globale. Étant conçu à l'origine pour être le plus simple possible, l'Internet laisse toute l'intelligence aux points d'extrémité. Les paquets sont acheminés par chaque routeur intermédiaire en prenant comme base l'adresse qu'ils portent et selon un service "au mieux" (*best-effort*). La plupart des applications à ce jour sont crées pour utiliser ce service. Cependant, de nouvelles applications, telles que les multimédias, peuvent être plus exigeantes en ce qui concerne la qualité de service (QdS) offerte par le réseau.

Différentes approches, telles que DiffServ (Services Différenciés), ont été proposées pour ajouter des fonctionnalités à la couche réseau de l'Internet. Dans ce travail, nous analysons le support de voix par le service *expedited forwarding* de l'architecture DiffServ. Cette analyse considère des modèles de sources de voix différents, des différents mécanismes d'ordonnancement ainsi que des différentes tailles de paquets de voix. Nous évaluons également la transmission des flux vidéo MPEG en utilisant conjointement différents schémas pour la QdS. Notre évaluation traite des questions telles que l'introduction de la redondance par un schéma de *Forward Error Correction* (FEC) pour la tolérance aux pertes, le changement du facteur de compression du codage MPEG, la protection inégale de trames MPEG et l'impacte des pertes au niveau réseau sur la qualité au niveau application.

Une fonctionnalité additionnelle visant le déploiement de nouveaux services concerne la localisation géographique d'hôtes. Cela ouvre le chemin pour une classe toute entière d'applications conscientes de la localisation des hôtes. Des exemples de ces applications sont, parmi d'autres, le contrôle d'accès au contenu basé sur la localisation, le routage géographique de paquets et la publicité ciblée. Pour la gestion de la localisation dans les réseaux ad hoc mobiles, nous proposons une approche adaptée au flux TCP. Cette approche s'applique à tous les schémas précédents de gestion de la localisation dans les réseaux ad hoc mobiles. Nous démontrons que l'approche proposée réduit dans tous les cas les coûts par rapport à l'approche traditionnelle qui ne prend pas en considération de caractéristiques particuliers des flux TCP. Dans le contexte de l'Internet, nous nous concentrons sur la localisation géographique d'un hôte basée sur des mesures. Ces mesures sont réalisées par des machines sondes vers des hôtes de références ayant une localisation géographique connue et vers l'hôte à être localisé. Nous abordons le problème de placement des hôtes de référence et des machines sondes en proposant une politique démographique de placement. Nous analysons des modèles de similarité utilisés pour choisir l'hôte de référence qui fournira l'estimation de localisation. Le placement des hôtes de références et des machines sondes ainsi que les modèles de similarité influencent la précision de l'estimation de localisation réalisée.

Part I

Quality of Service for Voice and Video

Chapter 2

Quality of Service in the Internet

An increasing number of private and commercial users rely on the Internet to communicate and exchange data. As a consequence of this diversity, the best-effort model of the conventional Internet has become inadequate to deal with the very diverse requirements on network Quality of Service (QoS) of an ever-increasing range of traffic types [1, 2, 3]. Such requirements imply that the network guarantees a certain bandwidth, maximum packet delivery delay, or minimum variations in the inter-arrival time of consecutive packets (jitter). Therefore, a growing need to extend the Internet architecture with QoS mechanisms arises to allow some users and applications to obtain a better service than others [4, 5]. This need is however controversial. Some people argue that the adoption of optical fibers and the evolution of wavelength-division multiplexing (WDM) would make bandwidth so abundant that QoS would be automatically delivered. Others argue that no matter how much bandwidth is added to the networks, new applications will be developed to consume them. We do not intend to develop a personal opinion on this issue. We assume that currently bandwidth is not an infinite resource and thus mechanisms to provide QoS in the Internet are needed at least in the foreseeable future.

In order to address the issue of QoS provisioning over the Internet, different approaches have been proposed by the Internet Engineering Task Force (IETF). The main approaches are the MultiProtocol Label Switching (MPLS) [6], constraint-based routing [7], the Integrated Services (IntServ) architecture [8], and the Differentiated Services (DiffServ or DS) architecture [9, 10]. Along this chapter we provide an overview of these proposals.

2.1 MPLS

The basic idea behind MultiProtocol Label Switching (MPLS) [6, 11] is to label packets as they enter an MPLS domain. Within an MPLS domain, Label Switching Routers (LSRs) perform the classification, forwarding, and services for the packets based on the assigned label. Ingress, egress, and interior LSRs of an MPLS domain use conventional routing algorithms to determine the existent routes inside the domain. A label in each ingress LSR is assigned to each route toward each egress LSR within the MPLS domain. Such a label is used as an index for the forwarding tables at each LSR along the path, which specify the next hop LSR and a new label. The new outgoing label substitutes the old incoming label and the packet is then forwarded toward the next hop LSR. Before leaving an MPLS domain, packets have their labels removed.

The paths taken by packets between the ingress and egress LSRs within an MPLS domain are called Label Switched Paths (LSPs). Each LSP may be viewed as a tunnel between an ingress LSR and the correspondent egress LSR. Whenever a packet arrives at the entrance of a tunnel, the label assigned to the packet by the ingress LSR completely determines its path toward the end of the tunnel, i.e. the egress LSR. The LSPs setup is done using a signaling protocol. RSVP (ReSource reserVation Protocol) [12, 13, 14, 15] or LDP (Label Distribution Protocol) [16] can be used to perform this task. The MPLS becomes strategically significant because it provides faster packet classification and forwarding as well as an efficient tunneling mechanism. This decoupling between forwarding and routing allows MPLS to efficiently support traffic engineering [17, 18, 19] and virtual private networks (VPNs) [20, 21].

2.2 Constraint-Based Routing

Constraint-based routing [7] computes routes that are subject to multiple constraints. It evolves from QoS routing [22, 23, 24, 25]. While determining a route, constraint-based routing considers: (i) the network topology like conventional routing algorithms; (ii) the QoS requirements and the availability of network resources like QoS routing proposals; and (iii) other constraints such as policies specified by the network administrator. For example, a policy may restrict the access to some routes to packets coming from a limited set of addresses. As a consequence, constraint-based routing may find a longer and lightly-loaded path better than a heavily-loaded shortest path. Network traffic is thus distributed more evenly. In order to perform constraint-based routing, compatible routers need to distribute new link state information to be used in route computation.

The advantages of adopting constraint-based routing are a better meeting of the QoS requirements of flows and an improved network utilization. Although constraint-based routing offers significant advantages, some drawbacks also exist such as increased overhead in communication and computation, increased routing table size, and potential routing instability. Such drawbacks are inherently linked to routing granularity. Routing can be destination based, source-destination based, class based, traffic trunk based, or flow based. Considering a route with finer granularity is more efficient in terms of network utilization and more stable. Nevertheless, overheads in communication and storage are higher.

The joint adoption of constraint-based routing and MPLS can be mutually beneficial [26]. MPLS allows to perform constraint-based routing at traffic trunk granularity without introducing per flow classification at core routers. The per LSP statistics from MPLS provide constraint-based routing with accurate information about amount of traffic between every ingress-egress pair. Based on such data, constraint-based routing is able to better compute routes for setting up LSPs.

2.3 Integrated Services

In the Integrated Services (IntServ) architecture [8, 27], a router provides a certain QoS for each flow according to its service model. This function is done by the traffic control that has three components: the packet scheduler, the packet classifier, and the admission control. The packet classifier performs the mapping of packets that arrive into a given predefined service class. The packet scheduler manages a set of queues deciding which queue (class) has to be served. This decision is based on priorities among the queues according to the adopted service model. All packets belonging to the same class receive the same treatment by the packet scheduler. The same packet may be classified in different ways by different routers along its path. The admission control verifies if it is possible to provide the QoS requirements to a certain flow without jeopardizing the contracted service to flows that are already active. In an IntServ environment, each data flow represents an open session. Two proposals of new services appear in the IntServ context: the Controlled-Load service [28] and the Guaranteed Service [29].

2.3.1 Controlled-Load Service

The Controlled-Load (CL) service [28] is intended for applications that can tolerate a certain amount of delay or packets losses if this amount is kept to a reasonable level. Adaptive real-time applications fall into this category. The CL service provides approximately the same quality of service under heavy loads as under light loads. If the flow is accepted for the CL service, then the router assures to offer the flow a service equivalent to that seen by a best-effort flow on a lightly-loaded network. The difference from the conventional Internet best-effort service is that the CL flow does not noticeably deteriorate as the network load increases. This will be true regardless of the level of load increase. By contrast, a best-effort flow would progressively experience a worse service, characterized by higher delay and eventually packet losses, as the network load increases.

2.3.2 Guaranteed Service

The Guaranteed Service (GS) [29] offers a deterministic guarantee to the contracting application in terms of a bounded end-to-end delay and no discard of packets that are in accordance with the contracted traffic profile. The GS model aims at serving applications with stringent real-time delivery requirements such as certain audio and video applications that consider useless any packet arriving after its playback point. Each router assures the GS contracted characteristic for a certain flow by allocating a bandwidth R and a buffer space B that the flow is allowed to consume. The goal is to approximate the fluid service model [30, 31] in a way that, from the viewpoint of the flow, the network presents itself as a dedicated link of bandwidth R. To achieve this goal, a message sent from the source to the receiver carries the traffic description and stores information about accumulated deviations from the fluid model that happened throughout the path. From these stored deviations and from the traffic description, the receiver is then able to compute the needed bandwidth R and buffer space B [30, 31, 32] to be reserved at each router to assure a bounded end-to-end delay and no packet losses. After performing this computation, the receiver sends its reservation request containing the parameters R and B toward the intermediate routers. The intermediate routers then reserve the needed resources to guarantee the QoS requirements of the flow.

2.3.3 Applicability of Integrated Services

The Integrated Services architecture recommends the adoption of RSVP for the signaling of the needed resources to be reserved to meet the QoS requirements of a particular session. The RSVP protocol is receiver-oriented, that is, it allows the receivers to decide the level of needed resources to be reserved, to establish the reservation, and to keep a reservation active as long as they want to receive data. The reservations at intermediate routers are kept in soft-states that need to be periodically refreshed to maintain them active. The soft-state approach provides to RSVP adaptivity and robustness since it allows the automatic reestablishment of reservations in the case of route changes. RSVP defines different reservation styles through the use of filters. The reservation of resources controls the quantity of reserved resources while the filter assigned to this reservation controls which packets can use the reserved resources. Since it is receiver-oriented, RSVP must assure that the resource reservation messages from the receiver follow exactly the same path taken by the data flows sent by the source. To perform such a task, the source sends path messages to the receivers. These messages store informations about the path taken toward the receivers. In the other hand, receivers send reservation messages signaling the needed resources to be reserved through the reverse path toward the source. The passage of such messages refreshes the soft-states at the intermediate routers.

The maintenance of one state per flow and the processing associated with them is unfeasible in routers with large traffic volume in a global scale network like the Internet [33]. This lack of scalability has limited the adoption of the IntServ proposal in the Internet. State aggregate proposals [34, 35, 36, 37] aim at reducing the state volume at the classifier, at the packet scheduler, and at the signaling protocol (for instance, RSVP). State aggregation implies partial loss of the isolation between flows and prevents the provision of deterministic guarantees like those of Guaranteed Services, thus compromising this service. Due to these scalability problems, the use of RSVP and IntServ is recommended only for small and confined networks [38].

2.4 Differentiated Services

From the preliminary studies of Integrated Services emerged the proposal of Differentiated Services (DS) [10, 39]. The DS scheme provides scalable discriminated services without the need of one state per flow and signaling at every router. The treatment offered by a DS compatible node is applied to an aggregate of flows, not more to individual flows. Individual flows are classified in aggregates that are previously set. These aggregates are differently served within a DS domain. The identification of a flow aggregate within a DS domain is performed by checking a new field called Differentiated Services (DS) in the IP header [9].

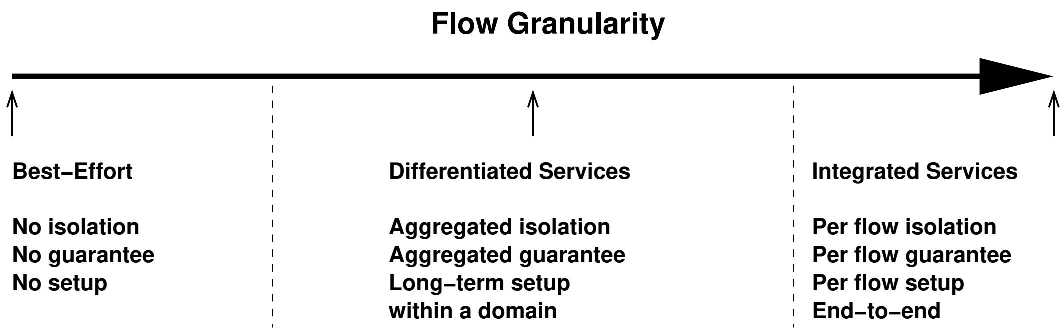


Figure 2.1: Comparison between service models.

A DS domain is composed by a contiguous set of nodes that operate a common set of service provisioning policies. At each DS domain, services are performed by traffic conditioning at the edges and simple forwarding mechanisms therein. Thus, individual flows needing a similar service are aggregated at the DS domain boundary and such aggregates receive differentiated services across the DS domain. The packets that do not belong to any contracted flow receive the standard service, equivalent to the best-effort service of the current Internet. Core routers maintain per aggregate state instead of one state per flow, pushing the complexity of handling individual flows to the domain edges [40]. The core mechanisms are simplified to enhance the scalability potential therein. In summary, there is one state per flow at the boundaries of the DS domain and one state per offered service class at the nodes located within the DS domain.

Figure 2.1 qualitatively compares the best-effort, the Integrated Services, and the Differentiated Services models based on the level of distinction adopted by each model when facing the diversity of flows. The comparison in terms of flow granularity goes from the best-effort model, where there is no flow granularity at all since there is no distinction among the flows, to the Integrated Services model, where the granularity is maximal since the flows are individually treated. As illustrated in the figure, the best-effort service provides neither isolation nor guarantee of any kind. The Integrated Services proposal is at the opposite extreme with a per flow individualization. The Differentiated Services proposal is at an intermediate point in terms of flow granularity. In this proposal, the flows are grouped onto flow aggregates and the guarantees are provided to those aggregates. The signaling and resource reservations are in the long-term within each DS domain since these configurations survive individual flows.

Individual flows needing a similar service are aggregated at the DS domain boundary and these aggregates receive differentiated services across the DS domain. At every DS-compliant node, aggregates are forwarded according to a Per-Hop Behavior (PHB). The PHB is selected through a mapping between a codepoint that identifies the aggregate and this behavior. The edge-to-edge service received by an aggregate results from the combination of all PHBs along its path within a DS domain. The efforts of the DS research community in the IETF have been focused mainly on two PHBs: the Expedited Forwarding (EF) PHB [41, 42] and the Assured Forwarding (AF) PHB [43]. Both are briefly described as follows.

2.4.1 Expedited Forwarding

The Expedited Forwarding (EF) PHB [41, 42] intends to allow the building of a low loss, low latency, and low jitter service within DS domains. The intent is thus to provide a PHB in which suitably marked packets usually encounter short or empty queues. Furthermore, if EF queues remain short relatively to the available buffer space, packet loss is also kept low. To achieve this behavior, the incoming EF traffic must be conditioned by ingress routers. This conditioning mechanism should assure the aggregate arrival rate at the interior nodes to be compliant with the contracted rate. Traffic conditioning includes traffic shaping, packet discarding, and packet marking in agreement with a service profile. A more rigorous definition of the described EF behavior is provided with additional mathematical formalism in [42]. The full rationale for this formal definition of the EF PHB is given in [44].

Due to the best-effort model of the current Internet architecture, queues along the path are the main reasons for packet loss, latency, and jitter of a traffic. The EF PHB intends to reduce the time spent in queues by an EF aggregate while within a DS domain. To meet this goal, nodes that implement the EF PHB must assure that, at every transit node, the aggregate maximum arrival rate is less than that aggregate minimum departure rate. This guarantee must prevail independently of the intensity of any other traffic attempting to transit the node. The EF service is specified by the desired peak bit-rate for a specific flow or aggregate of flows [45]. The user contract with the network is not to exceed the peak rate. The network contract is that the contracted bandwidth will be available when traffic is sent. Ingress nodes must strictly condition the incoming EF traffic. Packets that exceed the contracted profile are dropped or delayed until they are within the service profile.

An EF PHB implementation can be performed by different queuing scheduling mechanisms. A priority queue served in priority round-robin (PRR) gives the desired behavior. Thereby, the priority queue is served as long as it contains a packet to be forwarded. The other queue is only served when the priority queue is empty. In this case, the absence of traffic policing or traffic shaping at the DS domain boundary may allow the use of resources that are not contracted. This can harm other traffic aggregates or even cause starvation. Traffic shaping concerns the regulation of the transmission rate of a flow or of an aggregate of flows, whereby packets exceeding the allowed rate profile are delayed until they are in accordance with the contracted rate profile.

An alternative to PRR scheduling is to use a weighted round-robin (WRR) scheduler [46]. In this scheduling police, the relative weight assigned to each queue determines the proportional share of the output bandwidth each queue is served with. A single queue is allocated to the EF traffic. The share of the output bandwidth assigned to the EF queue must be at least the EF contracted rate. A misbehaviored user cannot jeopardize other aggregates because of the existing isolation between them. However, the EF PHB definition may be violated if a proper traffic shaping and reshaping is absent at the DS nodes. Quantitative compliance testing of the offered EF service [47, 48, 49] is needed to allow vendors, operators, service providers, and users to exploit adequately the EF service.

2.4.2 Assured Forwarding

The Assured Forwarding (AF) PHB [43] allows a DS domain provider to offer different levels of assurances for forwarding IP packets for a user of the domain. AF PHB provides the delivery of IP packets in up to four independent delivery classes, called AF classes, where each class receives a certain amount of resources, such as buffers and bandwidth, in each DS node. Within each AF class, an IP packet can be marked, either by the user or at the DS domain boundary, with up to three levels of loss precedence. In case of congestion within the AF class, the loss precedence of each packet determines the relative importance of that packet. A congested DS node thus preferentially discards packets with a higher loss precedence.

The unequal protection of packets against discards requires the adoption of a multiple level active queue management (AQM) algorithm instead of the conventional Drop Tail mechanism. An example of AQM algorithm is the Random Early Detection (RED) mechanism [50], which is the recommended mechanism for AQM in the Internet [51]. Nevertheless, a simple RED mechanism does not provide the necessary drop discrimination of packets since no explicit packet marking from different flows is used. However, a two level RED mechanisms, like RIO (RED with In and Out [52]), is capable of providing the needed unequal protection of packets. Within an AF class, packets are marked as being in or out of profile at the DS boundary and are differentially handled by the RIO mechanism at the DS-compliant routers within the DS domain. Extended RIO mechanisms can implement multilevel drop precedence considering multilevel packet marking such as TCM (Three Color Marking) [53, 54]. Alternative marking schemes to be used with the AF PHB have been recently proposed, such as Equation-Based Marking (EBM) [55] and Adaptive Marking [56, 57]. Some extensions to the AF PHB aim at either providing stronger guarantees to enforce the desired class level differentiation [58, 59] or at improving the weighted fairness among different classes [60].

2.4.3 Other Propositions

Apart from the DiffServ propositions of the IETF, there are other efforts to provide different levels of service to aggregate traffic. Examples of such efforts are: the Paris Metro Pricing (PMP) [61], the Alternative Best Effort (ABE) [62], and the proportional differentiation model [63, 64, 65, 66]. The basic idea of PMP is to have some channels that differ in price. These channels offer different expected QoS through the action of users who select the channel to send their data on based on their willingness to pay for a better service.

The ABE scheme relies on the idea of providing low delay at the expense of less throughput. The goal is to preserve the simplicity of the original single-class best-effort service while providing low delay to interactive adaptive applications. With ABE, every best-effort packet is marked as either green or blue. Green packets are guaranteed a low bounded delay in every router. In exchange, green packets are more likely to be dropped during periods of congestion than blue packets. Typically, an interactive application will mark most of its packets as green and applications that perform a bulk data transfer will send blue traffic. The difference from DiffServ or IntServ resides in that neither packet color can be said to receive better treatment. Therefore there is no need for reservations or profiles.

The proportional differentiation model is proposed as a way to control the quality spacing between classes at each hop, independent of the load at each class. With this model, certain forwarding performance metrics are ratioed proportionally to class differentiation parameters chosen by the network operator. The relative class differentiation is kept even when significant variations occur in the class load distribution. Authors argue that the proportional differentiation model is more predictable and controllable than the DiffServ model from the viewpoint of the user that needs to dynamically choose a class to be served with. First, the model is predictable, meaning that higher classes provide better performance than lower classes, independent of the aggregate load or the class load distribution. Second, the model is controllable, meaning that the network provider can adjust the performance spacing between classes based on the class differentiation parameters. The proportional differentiation model can be applied in both queuing delay [63, 67] and packet losses [68].

2.4.4 Applicability of Differentiated Services

An Internet Service Provider (ISP) should establish a Service Level Agreement (SLA) specifying the forwarding services offered to its clients [69, 70, 71]. Clients may be individual users or corporate organizations. The SLA defines the contracted service to be expected by a client and the set of measurable criteria to be used to verify the accomplishment of this contracted service. The SLA may include traffic conditioning rules to be applied to the incoming traffic, then constituting a Traffic Conditioning Agreement (TCA). A customer is obliged to adjust the generated traffic streams to a contracted profile. In order to force a customer's traffic conformance to the profile, particular metering, marking, discarding, and shaping rules are defined in the TCA. The treatment of out of profile packets is also specified by the TCA. As work progressed in the DiffServ as well as in the Policy working groups of the IETF [72, 73], it has been perceived that the notion of agreement comprised, besides the strictly technical considerations, some considerations of pricing, contractual, or other issues of business nature. Such additional commercial considerations are not addressed by DiffServ. Therefore, the notions of SLA and TCA are to be taken in the broader context.

Concerning the specific context of the DS architecture, the elements of service and traffic conditioning are addressed as follows. A Service Level Specification (SLS) establishes a set of parameters and their values which together define the service offered to a traffic stream by a DS domain. The SLS contains the technical part of a SLA. As an element of the SLS, a Traffic Conditioning Specification (TCS) defines the classifier rules and a traffic profile to be confronted with the incoming traffic and is also the technical part of the TCA. The interrelations between SLA, SLS, TCA, and TCS are shown in a simplified way in Figure 2.2 [3].

Within the DS architecture, the forwarding service is determined by the sequence of PHBs that handle an aggregate along the DS domain. This forwarding service is specified by a SLS described within a SLA. An ISP that controls a DS domain may offer to its clients SLAs determining a statistical guarantee of services within this DS domain. Serban *et al.* [74] propose a dynamic and self-tuning mechanism for allocating resources in the core routers of a DS domain. To extend the offered guarantee to different parts of

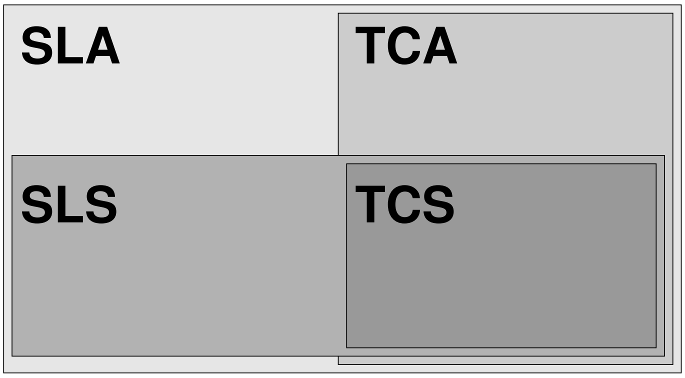


Figure 2.2: Interrelations between SLA, SLS, TCA, and TCS.

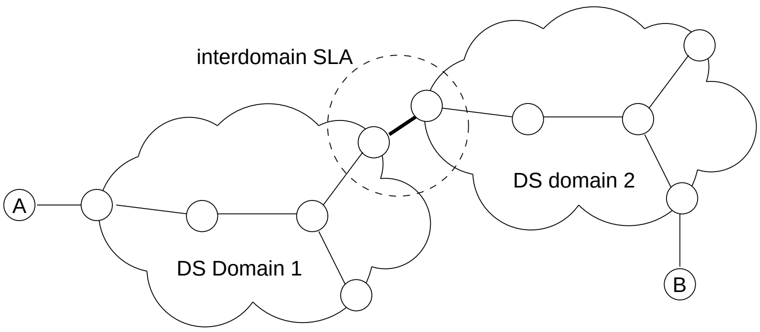


Figure 2.3: Differentiated Services provided by more than one domain.

the network, interdomain SLAs between the different DS domains are needed. Figure 2.3 presents an example of the interdomain relationship between two ISPs. If client A of DS domain 1 wants to send traffic with a differentiated service to node B, which is a client of DS domain 2, an interdomain SLA is needed between DS domains 1 and 2. This SLA determines the service expected by the traffic aggregate from DS domain 1 with respect to the offered services to DS domain 2. The established SLA of client A with DS domain 1 should define the expected service of this client including the interdomain SLAs held with other DS domains.

A key point for the success of the new multimedia applications is the network QoS provided to video and audio streams [75, 76, 77]. These QoS-sensitive streams point out in the emerging multimedia applications and have QoS requirements that are rarely achieved in the current Internet. The QoS support of voice and video applications constitutes a service that may be conceived applying concepts from the DS architecture. Therefore, we investigate different strategies to provide QoS support for voice and MPEG video traffic in Chapters 3 and 4, respectively.

2.5 Résumé du Chapitre

L'utilisation des applications multimédias dans l'Internet demande de nouvelles classes de service autre que le service "au mieux" (*best-effort*) d'IP [1, 2, 3]. Ces nouvelles classes de service doivent être créées pour fournir des garanties de Qualité de Service (QdS) aux applications. L'*Internet Engineering Task Force* (IETF) s'adresse au support de la QdS dans l'Internet par des approches différentes. Parmi elles, sont le *MultiProtocol Label Switching* (MPLS) [6], le routage basé sur des contraintes [7], les services intégrés (*Integrated Services – IntServ*) [8] et les services différenciés (*Differentiated Services – DiffServ*) [39].

L'idée de base dans la proposition MPLS consiste à étiqueter les paquets à l'entrée d'un domaine MPLS. À l'intérieur du domaine les paquets sont acheminés par les routeurs intermédiaires selon ces étiquettes. Le chemin pris par le trafic à l'intérieur du domaine MPLS peut être vu comme un tunnel entre les routeurs d'entrée et de sortie. La séparation entre le routage et l'acheminement fourni par MPLS permet le support efficace de l'ingénierie de trafic et des réseaux privés virtuels.

Le routage basé sur des contraintes détermine une route à partir : (i) de la topologie du réseau comme les algorithmes traditionnels de routage ; (ii) des requêtes de QdS et de la disponibilité des ressources du réseau comme les propositions de routage à QdS ; et (iii) d'autres contraintes telles que des politiques déterminées par l'administrateur du réseau. Une politique peut limiter quelques routes aux paquets provenant d'un certain ensemble d'adresses. Le routage basé sur des contraintes peut considérer une route plus longe et peu chargée meilleure que la route la plus courte si celle-ci se trouve très chargée.

La proposition IntServ donne des garanties par flux. Pour cela, des ressources sont réservées sur les routeurs intermédiaires en utilisant un protocole de signalisation tel que RSVP. La maintenance d'un état par flux n'est pas faisable dans les routeurs qui acheminent des volumes importants de trafic dans un réseau au niveau global comme l'Internet. À cause de ces problèmes de résistance au facteur d'échelle, l'utilisation de RSVP et IntServ n'est recommandée qu'aux petits réseaux.

L'architecture DiffServ se propose de surmonter les problèmes de résistance au facteur d'échelle rencontrés par la proposition IntServ. Le traitement offert par un nœud DiffServ s'applique aux agrégats de flux, non plus aux flux individuels. Le service bordure à bordure fourni par DiffServ à un agrégat est une fonction des PHBs (*Per-Hop Behaviors*) de chaque nœud composant son chemin à l'intérieur du domaine DiffServ. Deux PHBs sont définis : l'*expedited forwarding* (EF) et l'*assured forwarding* (AF). Un fournisseur de services doit établir un accord de niveau de service (*Service Level Agreement – SLA*) qui détermine les services offerts à ses clients. Dans le cas où plusieurs fournisseurs de services sont traversés par le flux demandant un service différencié, des SLAs inter-domaines doivent être également établis.

Un point clé pour la réussite des nouvelles applications multimédias est la QdS fournie aux flux audio et vidéo. Ces flux possèdent des besoins de QdS qui sont rarement satisfaits dans l'Internet d'aujourd'hui. Par conséquent, le support de la QdS aux flux de voix et de vidéo s'avère très important pour les applications multimédias. Nous réalisons une analyse des différents schémas pour le support de QdS à la voix et aux flux vidéo MPEG dans les Chapitres 3 et 4 respectivement.

Chapter 3

Supporting Voice

Addressing the issue of voice traffic evaluation in a Differentiated Services (DS) environment becomes specially important as the DS architecture is expected to be one of the building blocks for Internet Telephony deployments [78, 79, 80, 81, 82]. This chapter focuses on how voice traffic may benefit from being treated by the Expedited Forwarding (EF) Per-Hop Behavior (PHB) across a DiffServ (DS) domain [83]. The voice traffic sources are either a constant bit rate (CBR) source or an On-Off source, which are both often used to model voice traffic. The assessment covers the different algorithms that could implement the EF service. For the considered scheduling mechanisms, we identify the sources of delay and jitter. The analysis covers the impact of traffic shaping, the influence of the size of voice packets, and how efficiently each type of traffic uses an extra allocated bandwidth.

3.1 Related Work

A lot of research effort concentrates on several aspects of differentiated services, such as provisioning [84, 85], delay and loss differentiation [86, 87], server selection [88], admission control [89], DiffServ multicasting [90, 91], pricing [92], and DiffServ extensions for wireless networks [93]. Nevertheless, the transmission of voice traffic over the DS architecture remains not well explored. Jacobson *et al.* [94] focus on the jitter analysis of voice traffic transported across a DS domain where all nodes have a priority queue implementation of the EF PHB. The work presented in [41] reports a comparison of the jitter observed by traffic transported by different possible implementations of the EF PHB definition. The results demonstrate that a priority queue provides the lowest jitter when compared with weighted round-robin (WRR) [95]. Bounds on jitter across DS domains implementing the EF PHB are investigated in [49]. The effects of different scheduling schemes and of different packet sizes on voice delay are studied in [96]. Although the obtained results focus on delay, additional delays introduced by a possible shaping of the incoming voice traffic are not considered. In [97, 98], we evaluate the performance of voice sources when handled by the EF PHB, not taking into account the effects of traffic shaping and considering fixed size voice packets. We also analyze [97, 99] the performance of voice flows transmitted through a DS domain implementing both EF and AF PHBs.

3.2 Evaluation Metrics for Voice Traffic

The most relevant QoS parameters for voice traffic include delay, jitter, and packet loss rate. These metrics reflect the continuity of a voice flow [100, 101]. For interactive voice traffic, however, packet loss is not as critical as delay or jitter, whereas many voice coders can tolerate some small loss without several penalty. A maximum end-to-end delay requirement must be satisfied in order to keep a voice conversation interactive. There should be no or little impact if one-way end-to-end delay is below 150 ms [102, 103]. A small jitter (i.e. 50 ms at most) is also suitable to reduce the buffer capacity needed by the receivers to compensate these variations in delay and keep the playout at the receiver smooth.

This chapter aims at evaluating the delay and jitter performance of the EF PHB when transporting voice traffic modeled by CBR and On-Off sources. The analysis compares two implementations of the EF PHB. One is a priority queue served in Priority Round-Robin (PRR) and the other a Weighted Round-Robin (WRR) scheduling mechanism. We also evaluate how efficiently each type of traffic uses an extra allocated bandwidth and the influence on performance of the voice packet size. Moreover, we study the effects of traffic shaping. The shaping of the EF aggregate at the ingress node allows the maintenance of a smaller reservation of bandwidth across the DS domain. Such an action is done at the cost of an additional delay for the voice traffic. We analyze the trade-offs between this additional delay and the bandwidth allocated inside the DS domain for different possible implementations of the EF PHB and different models for voice traffic.

3.3 Simulation Model

We implement our simulation model in a modified ns-2 simulator [104]. An extension to the basic ns-2 package was developed to support Differentiated Services. The implementation of different PHBs is achieved through a combination of a variety of packet classification, packet marking, packet scheduling, and traffic conditioning mechanisms. We use some of such mechanisms from our modified ns-2 to implement the EF PHB.

3.3.1 Simulated Network Scenario

Figure 3.1 presents the network scenario used in our simulations of voice over the EF PHB. A number of voice flows are forwarded as an aggregate through a DS domain by the EF PHB. Along its path, the EF aggregate is shaped at the ingress node of the DS domain and re-shaped at the interior nodes in order to be at any moment in accordance with the contracted rate and not violate the EF definition. These shapers are implemented by a token bucket algorithm. The token bucket depth is always set to be equivalent to the size of one voice packet, that is, no burst exceeding the contracted rate is allowed into the DS domain. The token rate may have values that range from the average rate of the EF aggregate to the peak rate of this aggregate depending on the simulation.

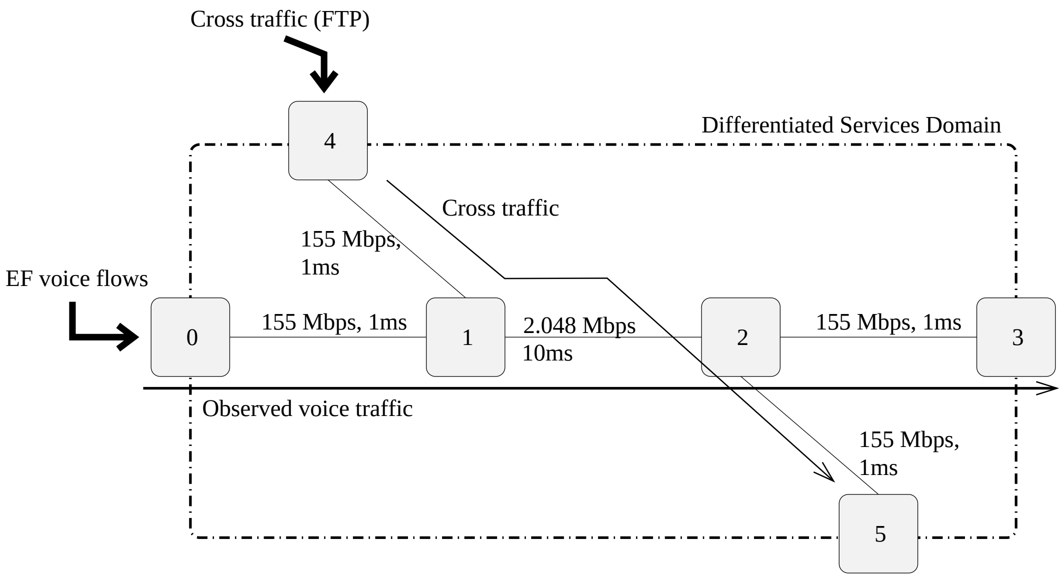


Figure 3.1: Network scenario for voice over EF.

3.3.2 CBR and On-Off Voice Sources

The EF voice sources generate voice traffic following either the On-Off model or the CBR model. When an On-Off source is in the On state, it generates voice traffic at a peak rate of 64 kbps. The On periods are determined by a random variable with exponential distribution with mean 400 ms. The Off periods are also exponentially distributed by a random variable with mean 600 ms. Therefore, the average transmission rate for each On-Off source is 25.6 kbps. Those mean values for the On and Off periods are in accordance with the bounds to model voice traffic presented in [105]. Since the allocated share of the bottleneck link to EF traffic is 512 kbps, there are 20 On-Off sources, leading to a worst-case situation as their average rates exactly fill up the EF allocation.

The CBR sources have rates of 64 kbps with a small variation around this value. Such sources are called jittered CBR sources [41, 106]. This kind of sources represents better a real implementation of a digitized voice generated by an application running on a desktop computer. We use 8 CBR sources to fill up the 512 kbps share of the bottleneck link allocated to EF traffic. The voice packets may have different sizes ranging from 40 to 160 octets. Besides the EF voice flows, the bottleneck link receives cross traffic composed of FTP connections that have their packets forwarded as best-effort packets. The TCP cross traffic sources use the TCP-Reno implementation which adds the fast recovery mechanism to the original congestion control mechanisms of the TCP. The cross traffic packets have 1500 octets which leads to a worst-case scenario for delay and jitter. Over 80% of Internet traffic packet sizes are much smaller than 1500 octets [107]. Thus, a more realistic traffic mix would give a better performance than the ones observed in our simulations. The TCP cross traffic is generated by FTP applications, active during the whole simulation. As the goal of this work is to evaluate the delay and jitter performance of the voice traffic, the buffers of the shapers and of the DS-compliant nodes are large enough so that EF packets are never discarded due to lack of buffering resources.

3.4 Simulation Results

The results are measured from one EF reference voice flow. That is, on all simulations the delay and jitter are from packets belonging to the reference flow within the DS domain. The delay of a single packet is the interval between the arrival of the packet at the ingress node and its departure from the egress node, including any delays introduced by traffic shaping. The number of simulations are large enough to provide a 95% confidence interval of 5% around the presented averages of delay and jitter. As in [41, 108], we assume the definition of jitter as the absolute value of the difference between the arrival times (a) of two adjacent packets i and j minus their departure times (d), then

$$jitter = |(a_j - a_i) - (d_j - d_i)|. \tag{3.1}$$

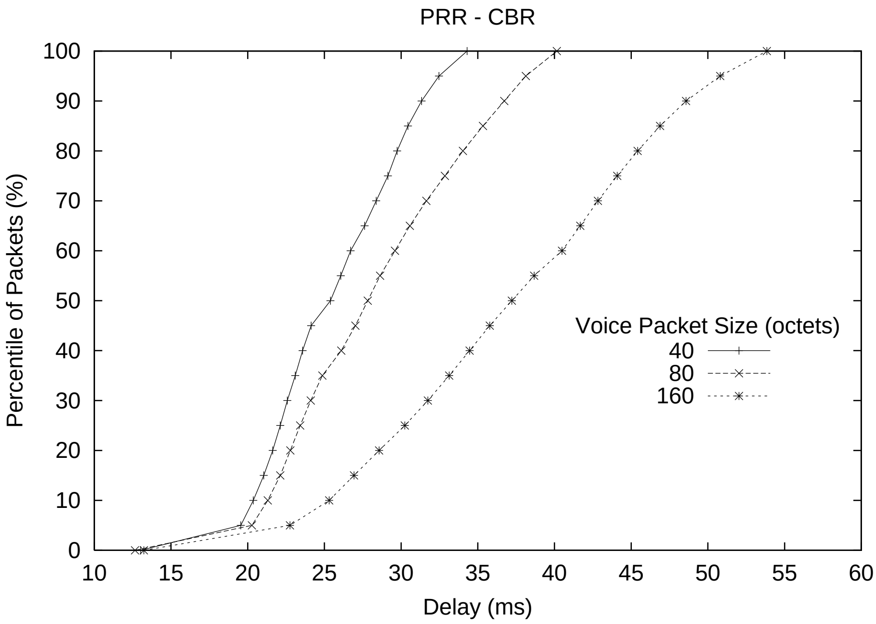
3.4.1 PRR Implementation of the EF PHB

In a Priority Round-Robin (PRR) implementation of the EF PHB a priority queue is allocated to the EF aggregate. The remaining traffic is sent through a low priority queue. The PRR scheduler serves the priority queue as long as there are packets to be served. The low priority queue is served if there are no packets in the priority queue in a non-preemptive way. The EF traffic is shaped at the entrance of the DS domain and reshaped therein in order to assure that it is in accordance with the contracted rate. This action avoids the starvation of the low priority traffic if there are incoming bursts in the EF traffic arriving at the DS domain. One of our goals is to evaluate the impact of this traffic shaping on the QoS support of voice flows.

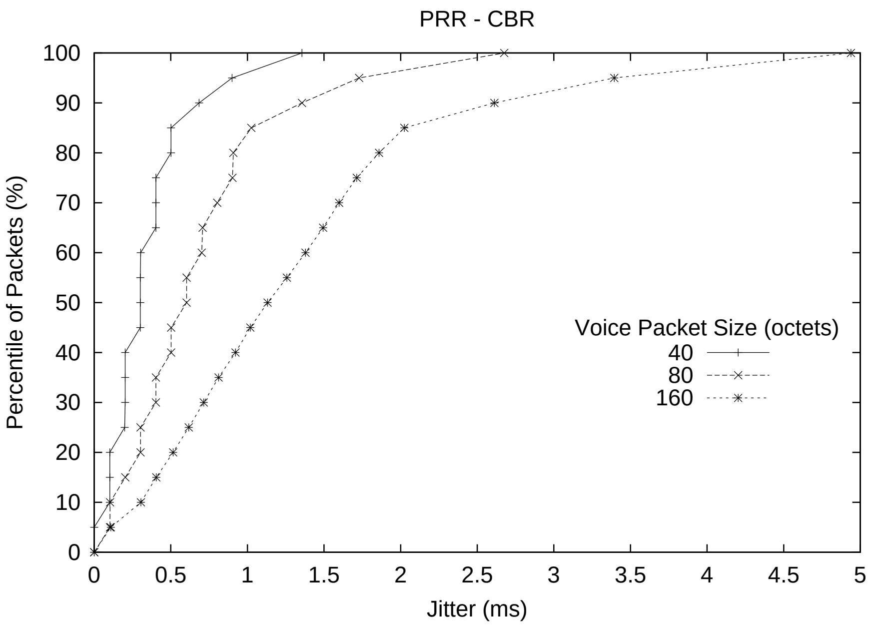
3.4.1.1 CBR Sources

First we study the service offered by a PRR implementation of the EF PHB to voice traffic modeled by the jittered CBR sources (10% of variation around an average rate). The voice traffic is shaped at the ingress node of the DS domain to the rate of 64 kbps. Figure 3.2(a) presents the percentile delay for three different voice packet sizes: 40, 80, and 160 octets. Figure 3.2(b) shows the corresponding results for jitter. The smaller the voice packet, the better the delay and jitter performance of voice flows. Moreover, a smaller voice packet also implies a smaller delay for the packetization and the depacketization of the voice samples at either end. However, there is a trade-off between the voice packet size and the overhead of the header.

When one EF voice packet is enqueued at the shaper, it must wait for the forwarding of all EF voice packets ahead of it in the queue before it is allowed to proceed. Depending on the size of the EF voice packets ahead of the new packet in the shaper's queue, it takes



(a) Delay (ms)



(b) Jitter (ms)

Figure 3.2: CBR traffic handled by a PRR implementation of the EF PHB.

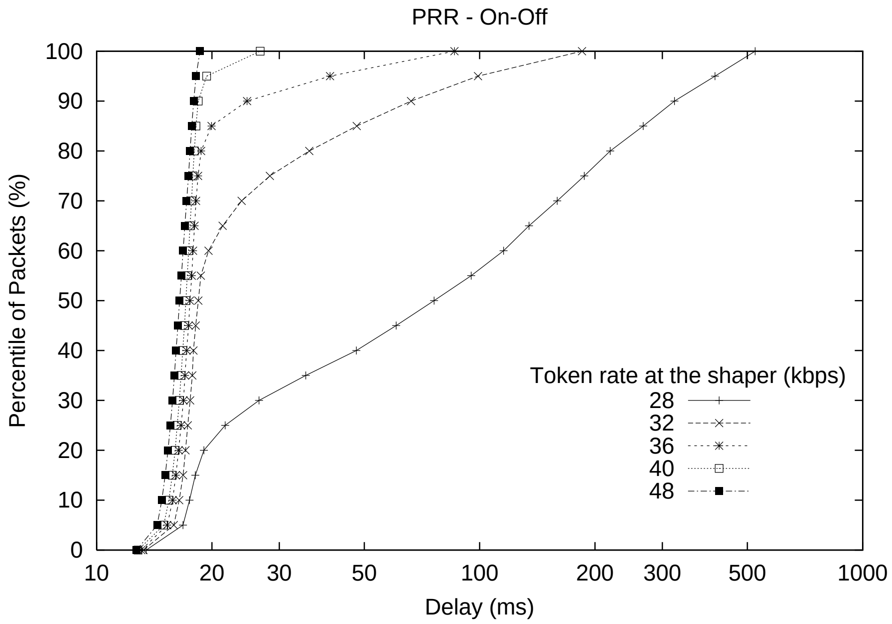


Figure 3.3: On-Off traffic handled by a PRR implementation of the EF PHB.

a larger or a smaller time for this packet to be forwarded. Also, if there is a low priority packet being transmitted when an EF packet arrives, it must wait until the low priority packet is completely served since the scheduling mechanism is nonpreemptive. It should be noted that the low priority packets are the TCP cross traffic packets having 1500 octets, thus leading to a worst-case scenario for delay and jitter. The difference in the individual delay of packets also influences on the jitter of the CBR traffic, as shown in Figure 3.2(b).

3.4.1.2 On-Off Sources

Next we evaluate the performance of a PRR implementation of the EF PHB when handling voice traffic generated by On-Off sources. The voice packets have 40 octets. As in the last simulation, the EF aggregate is shaped at the DS domain boundary. We define the *individual flow shaping rate* as the resulting rate that would be obtained by each component of a flow aggregation if they were equally affected by the shaping of the aggregation. Therefore, the actual token rate of the EF aggregation shaper is the product of the individual flow shaping rate and the number of EF active flows. Figure 3.3 presents the percentile delay as a function of the individual flow shaping rate, which ranges between 28 kbps and 48 kbps for each flow. Similarly, Table 3.1 provides the average delay for the same range of individual flow shaping rates. In these simulations, the EF aggregate is composed of 20 individual flows for a worst-case situation.

The larger the individual flow shaping rate, the smaller the time spent by packets waiting to be in accordance with the contracted rate. That is, as the individual flow shaping rate

Table 3.1: On-Off traffic delay in ms for a PRR implementation of the EF PHB.

Individual flow shaping rate at the shaper (kbps)	Average delay (ms)
28	123.82
32	31.68
36	20.53
40	17.25
48	16.42

increases, more bursts of the EF aggregate are allowed into the DS domain. The DS-compliant interior nodes forward these voice bursts with priority over the best-effort traffic. Nevertheless, the total goodput of the best-effort traffic composed of 8 FTP connections are not significantly affected by the variation of the allowed voice bursts, as shown is Figure 3.4. This is because the best-effort traffic uses the idle periods of the bursty traffic generated by the On-Off sources. For an EF aggregate composed of 20 voice flows, the total goodput of the best-effort traffic is slightly above 1.5 Mbps, which is reasonable since the share of the bottleneck link of 2.048 Mbps allocated to the EF aggregate is 512 kbps. For a similar worst-case situation, the allocation of a larger share of the bottleneck link to the EF aggregate decreases the total goodput of the best-effort traffic. We observe this result for EF aggregates composed of 40 and 60 voice flows, which correspond to approximately 50% and 75% of the bottleneck link, respectively. As the capacity of the bottleneck link allocated to the EF aggregate increases, the best-effort traffic tends to starve.

3.4.2 WRR Implementation of the EF PHB

A Weighted Round-Robin (WRR) implementation of the EF PHB is achieved by allocating to the EF aggregate a portion of the output bandwidth that is larger than the contracted rate to the EF aggregate at any moment in the DS domain. If the WRR weight is chosen to exactly balance arrival and departure rates the results are not stable [41]. A service-to-arrival rate ratio r is defined as

$$r = \frac{WRR_{share} \cdot BW_{output}}{PR_{EF}} = \frac{reserved\ bandwidth}{requested\ bandwidth}, \tag{3.2}$$

where WRR_{share} is the fraction of the output link bandwidth (BW_{output}) allocated to the EF aggregate, and PR_{EF} is the peak rate of the EF aggregate entering the DS domain.

A service-to-arrival ratio greater than 1 in every link within a DS domain means that more bandwidth was actually reserved than it was requested. The minimum service-to-arrival ratio to have stability is around 1.03 [41], and the value used in our simulations is initially set to 1.06, i.e., the actually reserved bandwidth is 6% larger than it has been requested.

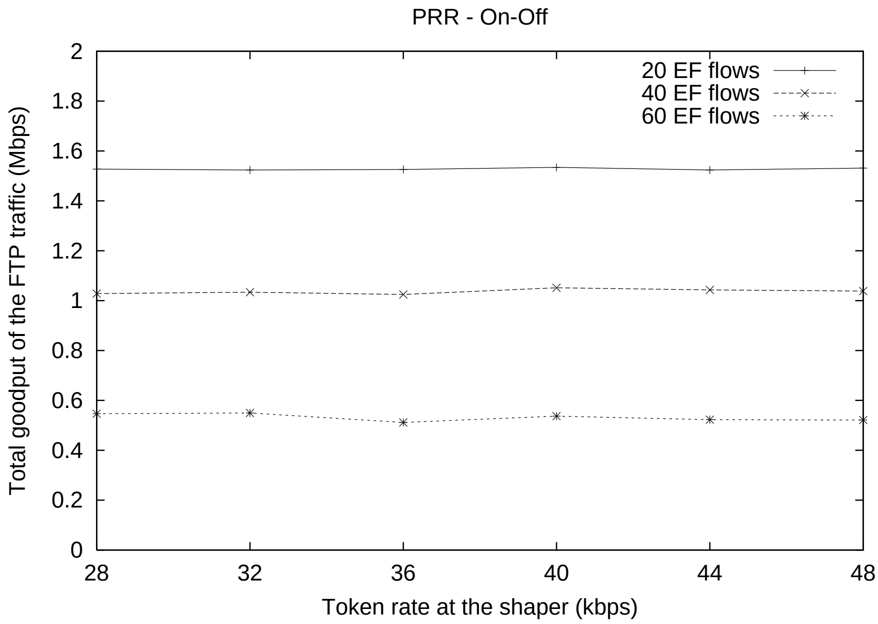


Figure 3.4: Goodput of the FTP cross traffic in a PRR implementation of the EF PHB.

In contrast to the PRR, where packet delays are mainly due to traffic shaping, the WRR implementation may have an additional source of delay and jitter. At any DS node, even if an EF packet is in accordance with the contracted rate, it may have to wait for its queue to be served by the WRR scheduler. If an EF packet arrives just after the EF turn in the scheduler has passed, the EF packet has to wait the whole round allocated to the best-effort traffic to be forwarded. Any EF packets arriving while the best-effort traffic is being served are delayed. Furthermore, when the EF traffic is served the accumulated EF packets are forwarded in a burst, which has to be conditioned by the downstream intermediate shapers. This reshaping also introduces additional delays to the voice traffic, resulting in a cascade delay effect.

With the variation of r, it is possible to allocate a greater WRR service share to the EF aggregate. This reduces the average time spent by EF packets waiting for being scheduled at each DS-compliant node. That is, by increasing the WRR service-to-arrival ratio r, an extra bandwidth is allocated to the EF aggregate. Increasing the service-to-arrival ratio r means that the traffic must always be shaped upstream, otherwise misbehaving EF traffic could take more than its fair share.

3.4.2.1 CBR Sources

The average delay of the CBR traffic composed of 40 octets packets and served by a WRR implementation of the EF PHB for different values for the WRR service-to-arrival ratio r is shown in Figure 3.5. The curves show results for jittered CBR sources with $\pm 1\%$ and $\pm 10\%$

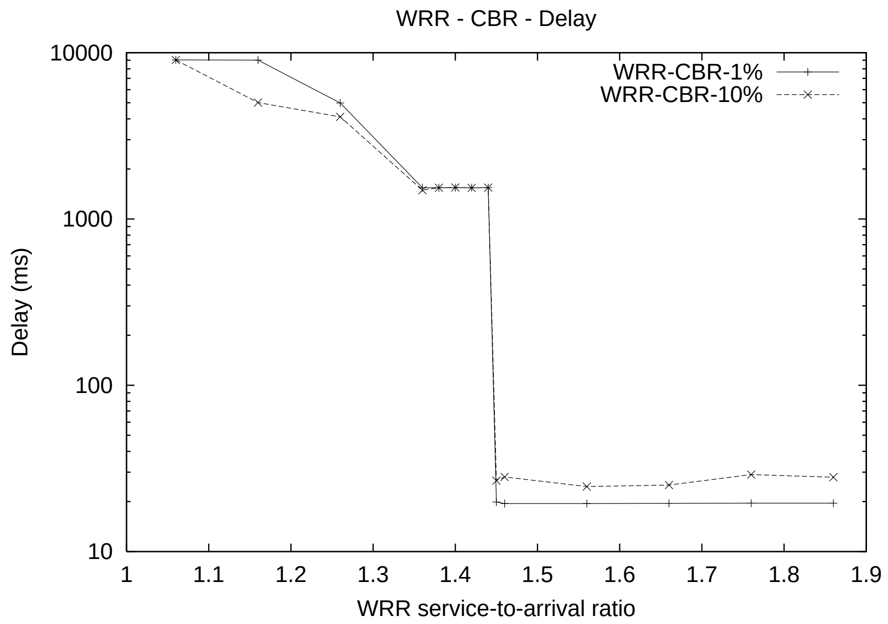
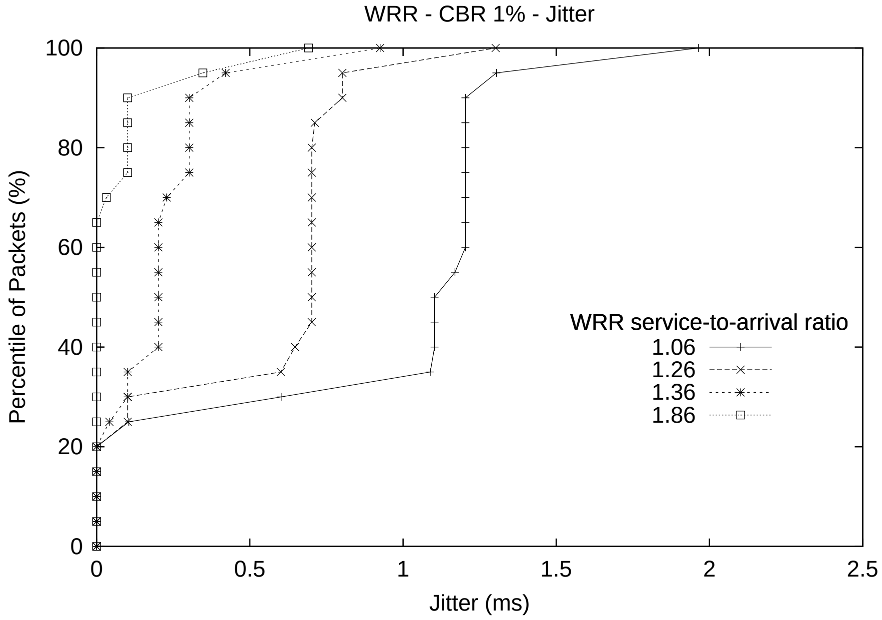


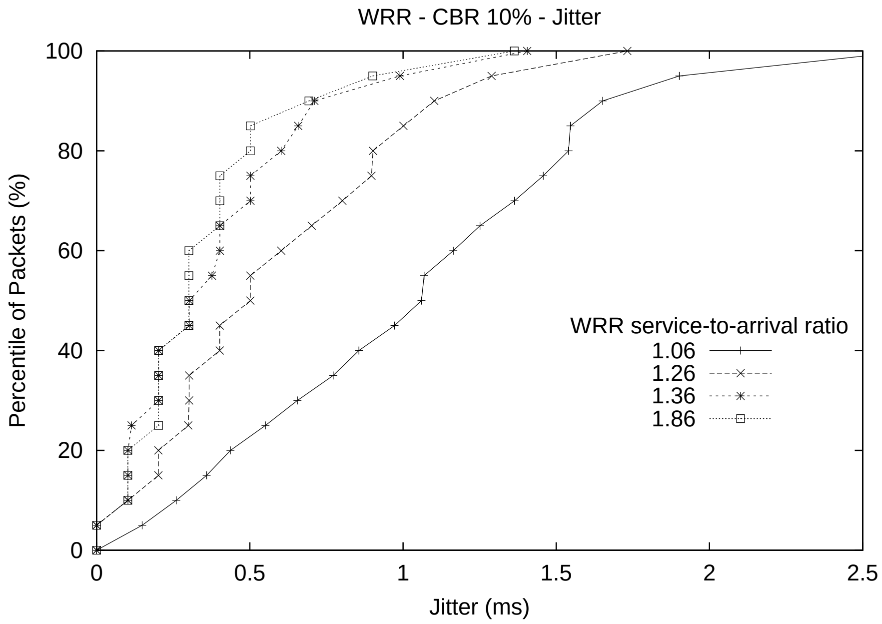
Figure 3.5: Average delay of CBR traffic in a WRR implementation of the EF PHB.

variation. Both aggregates are shaped to the rate of $N \times 64$ *kbps*, where N is the number of flows that compose the aggregate. For both values of variation, the average delay is too high to interactive voice traffic when $r < 1.45$ because of the cascade delay effect with delays being introduced by the waiting in the scheduler and the traffic reshaping. When $r \geq 1.45$, the delay levels off around 20 ms for the 1% variation jittered CBR sources and around 27 ms for the 10% variation jittered CBR sources. Under the considered conditions, a 45% over-provisioned share in the WRR scheduler is needed to mitigate the delay imposed on voice traffic by the non-voice traffic, thus avoiding the cascade delay effect. Nevertheless, for instance, the goodput of the cross traffic only decreases from 1.645 Mbps for $r = 1.06$ to 1.528 Mbps for $r = 1.86$. This is possible because the best-effort traffic uses the bandwidth of the EF aggregate when there are idle periods due to the work-conserving nature of the WRR scheduler.

The jitter performance, however, continually improves as the WRR service-to-arrival ratio r increases. Figure 3.6 shows the jitter of the CBR sources while an extra share of bandwidth is allocated to the EF aggregate. For a variation of 1% around the average rate and for all applied values of r, most voice packets present a narrow range of jitter variation, causing thus the cliffs presented in Figure 3.6(a). Meanwhile, for a variation of 10%, the jitter of voice packets presents a worse performance and has a wider range of variation. How much the traffic is jittered at the source directly influences the jitter variation. This effect may be observed comparing the range of jitter variation in Figures 3.6(a) and 3.6(b).



(a) 1% variation



(b) 10% variation

Figure 3.6: Jitter of CBR traffic in a WRR implementation of the EF PHB.

Table 3.2: On-Off average delay in ms for a WRR implementation of the EF PHB.

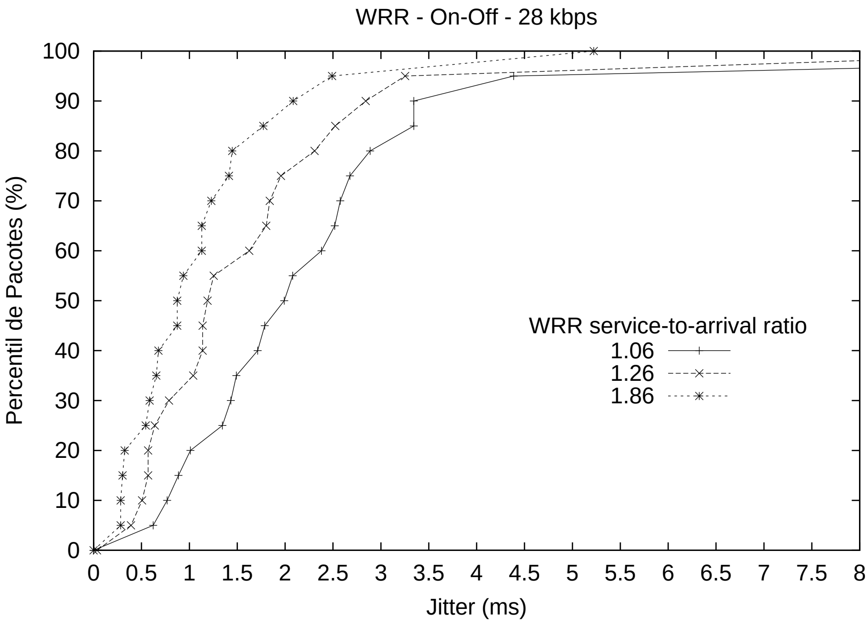
Individual flow shaping rate at the shaper (kbps)	WRR service-to-arrival rate 1.06	1.46	1.86
28	4,989.37	184.38	105.52
32	1,930.45	39.90	31.32
36	453.93	26.19	20.23
40	121.87	17.95	17.27
44	46.54	17.05	16.76
48	31.29	16.43	16.44

3.4.2.2 On-Off Sources

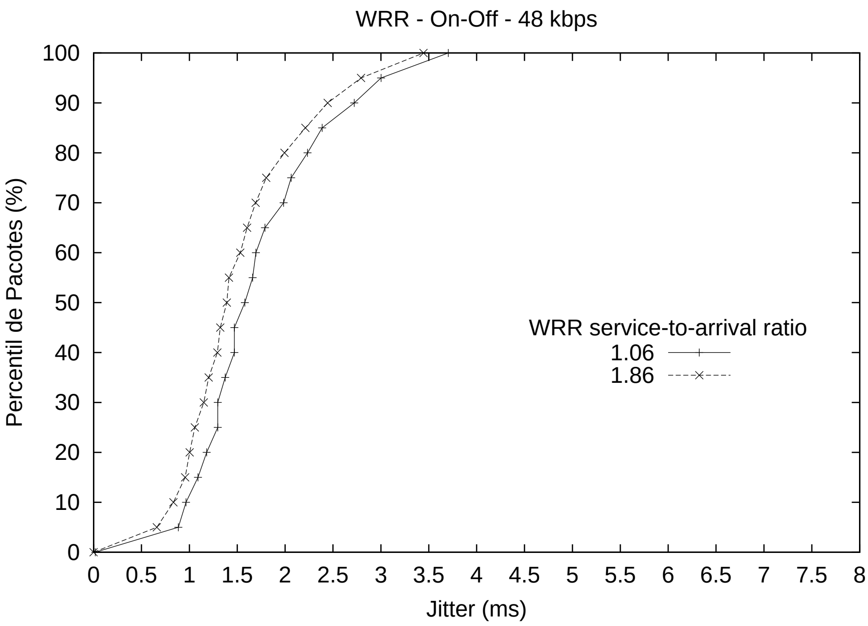
The delay and jitter performance for the voice traffic generated by On-Off sources and handled by a WRR implementation of the EF PHB is analyzed next. In this case, voice packets are delayed by two main reasons: shaping delay because of the bursty nature of On-Off sources and the waiting time in the queue served by the WRR scheduler. Hence, the delay performance may be improved by increasing the individual flow shaping rate at the shaper, the WRR service-to-arrival ratio, or both. By adjusting these parameters, an administrator may control the trade-off between the delay introduced by traffic shaping and the bandwidth allocated to the EF aggregate.

Table 3.2 provides the results for the On-Off traffic composed of 40 octets packets as we increase both the individual flow shaping rate of the shaper and the WRR service-to-arrival ratio r. From Table 3.2, the configuration with 28 kbps for the individual flow shaping rate at the shapers and $r = 1.06$ presents a prohibitive delay for interactive voice traffic. This configuration allows a goodput of 1.645 Mbps for the FTP cross traffic. At the other extreme of Table 3.2, where we have 48 kbps as the individual flow shaping rate of shapers and $r = 1.86$, the delay becomes low enough to allow the feasibility of interactive voice traffic. It should be noted that just for a service-to-arrival rate ratio $r = 1.86$, the WRR gets a delay performance similar to PRR (see also Table 3.1). Although the allocated bandwidth for the EF aggregate inside the DS domain is almost 1.8 Mbps (48 kbps $\times$ 20 EF flows $\times$ 1.86), which is nearly 87% of the bottleneck link, the goodput of the FTP cross traffic remains equal to 1.528 Mbps. This is possible because the best-effort traffic uses the bandwidth of the EF aggregate when there are idle periods and the EF aggregate is limited to a 512 kbps share of the bottleneck link. Note that the average delay decreases from 4,989.37 ms to 16.44 ms at the cost of just 100 kbps from the goodput of the best-effort traffic.

The jitter performance of the On-Off traffic handled by the WRR implementation of the EF PHB is shown in Figure 3.7. With a larger individual flow shaping rate at the shaper, there is a smaller influence on jitter of the WRR service-to-arrival ratio. This is illustrated when we compare Figures 3.7(a) and 3.7(b).



(a) Individual flow shaping rate = 28 kbps



(b) Individual flow shaping rate = 48 kbps

Figure 3.7: Jitter of the On-Off traffic for a WRR implementation of the EF PHB.

3.5 Conclusion

In this chapter, we have presented a simulation study of voice traffic over a Differentiated Services architecture for the Internet. We evaluated the service offered to voice traffic by the Expedited Forwarding PHB. The voice traffic was modeled by CBR and On-Off sources. We compared the performance obtained considering two different implementations of the EF PHB. One is a priority round-robin (PRR) and the second a weighted round-robin (WRR) scheduling mechanism. The analysis focuses on the delay and jitter observed by voice flows across a DS domain.

The traffic shaping of an EF aggregate composed of individual voice flows is fundamental to keep the delay and jitter within an acceptable bound. It also provides control over the incoming rate of the EF aggregate in order to guarantee that the EF definition is valid at any DS-compliant node. Although the traffic shaping introduces an additional delay, it allows the limiting of the contracted rate for the EF aggregate within a DS domain. The EF PHB is expected to constitute an expensive service. Thus, a trade-off may be established between allowing an additional but tolerable shaping delay and contracting a less costly amount of resources. The results show that as one increases the individual flow shaping rate at the shaper while keeping the same amount of EF flows, voice flows experience a smaller delay and jitter. Such an increase reduces the service rate share allocated to the best-effort aggregate. Nevertheless, the results also show that this does not significantly influence the best-effort traffic performance. This behavior is due to the work-conserving nature of the schedulers, which allows the BE traffic to use the bandwidth left unused by the EF traffic in idle periods.

The PRR implementation of the EF PHB provides the best results for delay as well as for jitter. Nevertheless, the study of alternatives is needed since this scheduler has some undesirable features. The PRR cannot avoid low priority traffic starvation and does not offer a good isolation among different aggregates. Thus, improving the performance of a WRR implementation of the EF PHB is justified knowing that this scheduling mechanism is already available on commercial routers [109, 110]. Besides the shaping delay, the WRR implementation may introduce an additional delay to the EF voice flows due to the waiting it may impose on each queue to be served. To improve the performance of voice flows, the individual flow shaping rate can be increased in order to decrease the shaping delay. Similarly, the service-to-arrival rate ratio can be increased in order to decrease the retention delay at the WRR scheduler. These two parameters may be viewed as tuning knobs for adjusting the configuration of a WRR implementation of the EF PHB within a DS domain.

3.6 Résumé du Chapitre

L'architecture de Différentiation de Services (DS) offre une solution pour fournir de la Qualité de Service (QdS) aux nouvelles applications multimédias dans l'Internet. Les paramètres de QdS les plus importants pour le trafic de voix sont le délai, la gigue et le taux de pertes des paquets. Ces métriques jouent un rôle sur la continuité du trafic de voix [100, 101]. Dans le cas de trafic de voix interactif, le taux de pertes de paquets n'est pas aussi critique

que le délai ou la gigue vu que plusieurs décodeurs de voix peuvent tolérer quelques pertes de paquets. Ce chapitre évalue le délai et la gigue observés par le trafic de voix lorsqu'il est traité par le schéma *expedited forwarding* (EF) de l'architecture DS [41, 42]. L'évaluation du service offert par un environnement DS au trafic de voix devient particulièrement importante, car cet environnement se présente comme une solution raisonable pour le déploiement de la téléphonie Internet [78, 79, 80, 81, 82].

Nous réalisons une évaluation par simulation en utilisant une version modifiée du simulateur ns-2 [104]. L'analyse inclue les effets des différents mécanismes d'ordonnancement qui peuvent implémenter le service EF, telles que l'usage d'une fille d'attente prioritaire ou d'un schéma *weighted round-robin* (WRR) [95], ainsi que différentes tailles du paquet de voix. Nous considérons deux modèles de sources de voix : le débit constant (*constant bit rate* – CBR) et l'activité-silence (*On-Off*). Nous évaluons aussi l'efficacité démontrée par chaque modèle en utilisant une allocation additionnelle de bande passante et l'impacte du *shaping* du trafic de voix.

Le *shaping* de l'agrégat EF qui se compose de flux de voix individuels est fondamental pour mantenir le délai et la gigue dans des limites acceptables. Malgré l'introduction d'un délai additionnel dû au *shaping* du trafic, le *shaping* permet la limitation du taux de transmission contracté par l'agrégat EF dans un domaine DS. Nous pensons que le service EF sera un service de coût elevé. Par conséquent, un compromis peut être établi entre l'introduction d'un délai additionnel, mais tolérable, et l'allocation d'un volume de ressources ayant un coût inférieur. Les résultats montrent que l'augmentation de la fraction de service attribuée à l'agrégat EF n'affecte pas fortement le trafic concurrent appartenant au service "au mieux" (*best-effort*). Cette situation se maintient tandis que le trafic "au mieux" utilise la bande passante qui a été laissée sans utilisation par le trafic EF pendant des périodes d'inactivité.

Chapter 4

Supporting MPEG Video

The wide adoption of the MPEG encoding turns it into an attractive way for the distribution of audio and video over the Internet. In the case of a video transmission over the Internet, both the encoding and the transmission processes affect the QoS [111, 112]. The MPEG encoding is pointed out as a standard for future networked interactive video applications [113]. Nevertheless, the hierarchical structure of MPEG encoding with a possible error propagation through its frames imposes a careful analysis on sending MPEG video streams over lossy networks. Small packet loss rates may translate into much higher frame error rates. For example, a 3% packet loss percentage could translate into a 30% frame error probability [114]. This situation seriously degrades the perceived quality by a user at the video reproduction. Moreover, network resources are wasted with the transmission of information that becomes useless to the receiver [115]. Some portion of the received data becomes useless to the decoder as insufficient frame data are available for decoding the frame because of losses in the network or because a frame, in which the current frame depends on, is undecodable. Thus, the transmission of video streams with a certain level of QoS is a challenging problem [116], possibly requiring the combined adoption of multiple strategies [117, 118].

The transmission of an MPEG video stream adopting a combination of different QoS strategies over a lossy Internet-like network is of significant importance because of the large number of IP networks. Quality video delivery over these widely deployed IP-based networks is required by several applications such as distance learning and collaboration, video distribution, video conferencing, and interactive virtual environments.

This chapter evaluates different possibilities to improve the delivery quality of an MPEG video stream crossing a lossy network like the Internet or an intranet[1]. We evaluate the joint adoption of such QoS strategies as well. Results cover issues as how to:

- Apply redundancy, either by using Forward Error Correction (FEC) to tolerate some losses or by using a different compression factor in the MPEG encoding. FEC schemes protect video streams against packet losses up to a certain level at the expense of data

[1]The work in this chapter has been done in collaboration with Prof. Bernd E. Wolfinger from the University of Hamburg, Germany.

redundancy. Adopting different frame patterns allows a MPEG stream to better adapt itself to the available transmission conditions.

- Adequately protect the different types of MPEG frames adopting a DiffServ architecture. Unequal protection based on frame type avoids quality degradation due to the loss of one particularly important frame and the possible propagation effect of such loss throughout the hierarchical structure of MPEG video streams.
- Analytically do the mapping of network QoS (in terms of packet loss probabilities) onto application QoS (in terms of decodable MPEG frames).

Furthermore, we evaluate how to best combine the investigated techniques to improve the delivery quality of the MPEG streams.

4.1 Related Work

In the MPEG video literature, common strategies for QoS provisioning include the adoption of different GOP structures, different frame dropping priority mechanisms, or FEC schemes. Changing the GOP pattern is a strategy to better adapt the MPEG stream to the available transmission conditions. Adjusting the adopted GOP structure to the image contents taking into account scene changes is studied in [119]. Detection of scene changes is used to optimize the GOP pattern for the actual video sequence. The impact of using different GOP structures on video quality is investigated in [120]. The work focuses on the tradeoff between storage benefits due to more compression and picture quality improvements. Turaga *et al.* [121] propose models for variable bit rate (VBR) video traffic that allow for different frame types and variable GOP patterns.

Indiscriminate packet discards may seriously degrade an MPEG video stream because of its hierarchical structure. Due to this feature of MPEG systems, different frame dropping priority schemes are proposed to protect different types of frames against such indiscriminate discards. In [122], a method to provide deterministic service guarantees to MPEG video streams is proposed. Alam *et al.* [123] study the effects of a traffic shaper at the video source on the transmission characteristics of MPEG video streams while providing delay and bandwidth guarantees. Hemy *et al.* [124] propose the use of filters in some nodes within the network to selectively drop frames of an MPEG stream, which must be partially decoded at each filter node.

A content-based packet video forwarding mechanism based on a relative priority score that is mapped onto a proportional loss-rate differentiation mechanism is proposed in [125]. Bounds on loss and delay of video-on-demand traffic crossing a DS domain are investigated in [126]. Shin *et al.* [127, 128] propose a framework to map categorized video packets (in terms of loss and delay) onto relative DiffServ levels, while taking into account a given pricing model. A similar technique is proposed in [129], whereby the decoder behavior is replicated and those packets that if lost would generate high distortion levels are assigned to classes expected to provide high QoS.

Video delivery can significantly be improved by the adoption of FEC mechanisms. Basically, redundancy is added to the data so a receiver may recover from losses or errors

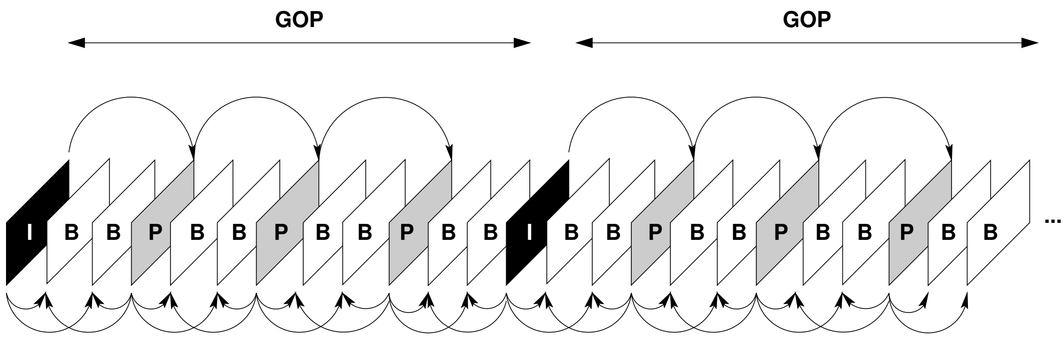


Figure 4.1: GOP structure ($N = 12$ and $M = 3$).

in the transmission without further intervention from the sender. A media-independent FEC scheme is the focus of recent efforts [130]. Another alternative is the development of FEC mechanisms that take into account the characteristics of the data being protected. The Priority Encoding Transmission (PET) scheme [131] has been developed based on the particular features of MPEG video streams and allows for FEC including the support of priorities. Wolfinger [132] studies the transmission of MPEG video streams with a FEC mechanism. Park *et al.* [133] propose an adaptive packet-level FEC protocol called AFEC suitable for the QoS-sensitive transport of real-time MPEG video streams.

We investigate the unequal protection of different types of frames using mechanisms based on the DiffServ architecture in [134, 135]. Although QoS schemes for MPEG video streams have been widely investigated, commonly QoS strategies are only studied individually. This chapter analyzes the joint adoption of different QoS approaches to enhance the quality of MPEG streams. We investigate the mutual influence these approaches exert on each other when jointly adopted. Moreover, we evaluate how to best combine the analyzed techniques to improve the delivery quality of an MPEG video stream.

4.2 Evaluation Metrics for MPEG

MPEG encodes video as a sequence of three types of compressed frames (I, P, or B) [136]. The whole video sequence is decomposed into smaller units which are coded together and called GOPs (Group of Pictures). The GOP pattern specifies the number and temporal order of P and B frames between two successive I frames. Such a GOP pattern is characterized by two parameters: the I-to-I frame distance (N), and the I-to-P frame distance (M). This hierarchical structure with dependencies for the decoding of some frames in each GOP is illustrated in Figure 4.1. For a given pattern (N,M), the number of frames from each type ($N_{\{I,P,B\}}$) in a single GOP are $N_I = 1$, $N_P = \frac{N}{M} - 1$, and $N_B = \frac{N}{M}(M - 1)$.

In order to decode a P or B frame, it is necessary to also decode the frames they depend on. Thus, an incorrect I or P frame propagates problems to all dependent frames and corrupts these frames as well. In the worst case, a whole GOP is considered undecodable due to an incorrect I frame because all other frames in the GOP depend directly or indirectly on the I frame. For example, the effects of an undecodable I frame of a GOP with $N = 12$

Table 4.1: Adopted notation to model the impact of losses on MPEG video quality.

Variable	Definition
ξ_I, ξ_P, ξ_B	probability for a frame to be directly lost, but not indirectly lost
$\alpha_I, \alpha_P, \alpha_B$	packet loss probability of packets containing data from each type of frame
C_I, C_P, C_B	mean number of packets to transport the data of each type of frame

will persist for 12 frames, or almost 500 ms at a typical frame rate of 25 frames per second, which is quite noticeable to a user.

To evaluate the transmission of an MPEG video stream, we consider the following metrics: *Fraction of decodable frames* and *fraction of useless data received*. The former reports the number of decodable frames over the total number of frames sent by a video source. The latter reports the fraction of the data received that is useless to the decoder. The decoder has useless data from partially received frames or from frames that can not be decoded because they depend on other undecodable frames.

The adopted criterion to consider a frame decodable defines that at least a fixed fraction dt (decodable threshold) of the data in each frame must be received to consider the frame decodable. Further, a frame is considered decodable if and only if all the frames it depends on are also considered decodable. Therefore, when $dt = 1.0$, the decoder is completely intolerant to losses, that is, one lost packet is enough to render a frame undecodable. Similarly, with $dt = 0.75$, 25% of the data from a frame may be absent due to losses in the network and the frame is still considered decodable. If the decoder can tolerate a certain level of losses ($dt < 1.0$), then we assume the use of a FEC scheme in the video transmission. The tolerance to losses is obtained at the expense of additional data introduced into the video stream (FEC redundancy).

4.3 Analytical Model

We adopt the notation described in Table 4.1 to model the impact on quality considering the defined evaluation metrics. From the hierarchical structure of MPEG encoding, a frame may be lost, i.e. considered undecodable, directly or indirectly. Direct loss of a frame indicates that there are not enough frame data to decode it. We assume direct frame losses to be mutually independent. Indirect loss of a frame happens when a frame is considered undecodable because some frame it depends on is directly undecodable. We observe also that consecutive B frames have the same dependency throughout the hierarchical structure and may be considered as a B group. Each GOP pattern has $(M-1)$ B frames in each B group. Thus, the expected number of correctly decodable frames in a GOP (cf. Appendix A) is

$$
\begin{aligned}
N_{\text{dec}} = {} & (1-\xi_I) + (1-\xi_I)\sum_{j=1}^{N_P}(1-\xi_P)^j \\
& + (M-1)(1-\xi_I)(1-\xi_B) \\
& \times \left((1-\xi_I)(1-\xi_P)^{N_P} + \sum_{j=1}^{N_P}(1-\xi_P)^j\right).
\end{aligned} \tag{4.1}
$$

4.3.1 Objective Metric for MPEG Video Quality

To evaluate the quality of a video stream, we adopt as a criterion the fraction of decodable frames per GOP. Therefore, we define the picture quality Q of an MPEG video stream as

$$Q = \frac{N_{\text{dec}}}{N}, \quad 0 \leq Q \leq 1. \tag{4.2}$$

The measure Q is an objective metric to assess video quality. Evidently, in video communications, such as in communications in general, we would rather be interested in evaluating QoS as observed by (human) end users, taking into account a measure such as MOS (mean opinion score) [137]. As our study is based on a modeling approach, contents of video streams, i.e. the semantics of video communications, are not taken into account. An earlier study [132] observes that a quality measure such as Q is slightly pessimistic as compared to the quality observed by humans, because MPEG streams tend to recover from partial losses of frames thanks to their temporal and spacial redundancy. Nevertheless, the measure Q tends to be a good approximation to subjective video quality assessments.

4.3.2 Metric on Useless Data Received

At the receiver of a video transmission, the mean number β of correctly received packets for each frame type are $\beta_I = (1-\alpha_I)C_I$, $\beta_P = (1-\alpha_P)C_P$, and $\beta_B = (1-\alpha_B)C_B$. Thus, for an entire GOP the expected number of correctly transmitted packets is

$$P_{\text{corr}} = \beta_I N_I + \beta_P N_P + \beta_B N_B. \tag{4.3}$$

From the total number of correctly received packets, not all of them are useful to decode a frame. Useful packets are those which belong to a decodable frame. Using Equation 4.1 and the mean number β of correctly received packets for each frame type, the expected number of useful packets in a GOP is

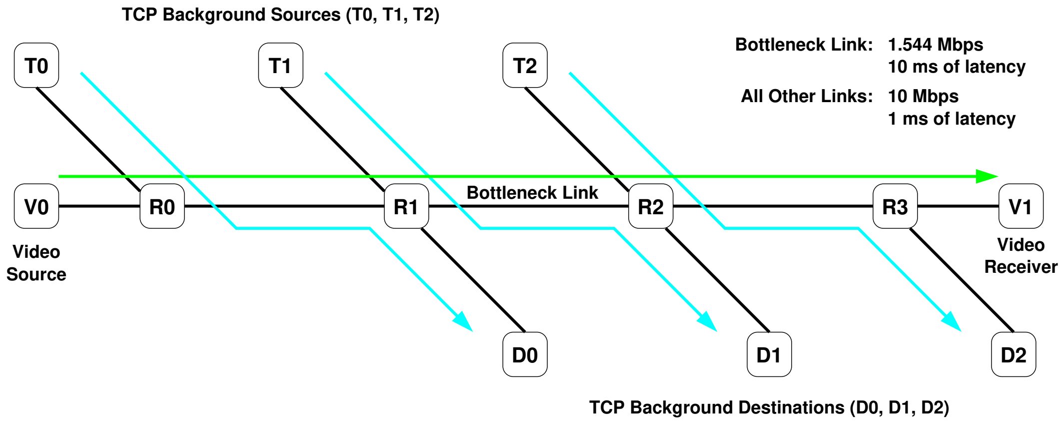


Figure 4.2: Network scenario for the MPEG simulation.

$$
\begin{aligned}
P_{\text{useful}} = {} & (1-\xi_I)\beta_I + (1-\xi_I)\beta_P \sum_{j=1}^{N_P} (1-\xi_P)^j \\
& + (M-1)(1-\xi_I)(1-\xi_B)\beta_B \\
& \times \left((1-\xi_I)(1-\xi_P)^{N_P} + \sum_{j=1}^{N_P} (1-\xi_P)^j \right).
\end{aligned} \tag{4.4}
$$

A fraction of the correctly received packets is not useful to the decoder. This fraction U of useless data received is defined as

$$
U = \frac{P_{\text{corr}} - P_{\text{useful}}}{P_{\text{corr}}}, \qquad 0 \leq U \leq 1. \tag{4.5}
$$

4.3.3 Model Validation

From Q and U we can evaluate the delivery quality of an MPEG video stream and the waste of resources due to some amount of useless data received.

To validate the described analytical model, we carried out simulations in the ns-2 simulator [104] using the network scenario depicted in Figure 4.2. A video source generates packets based on a public available MPEG frame trace file (see [138] and references therein for details of the generation of the frame trace file). We used the Starwars-IV movie sequence, encoded with high quality, which has a mean bit rate of 0.28 Mbps and peak rate of 1.9 Mbps. Each frame is fragmented into packets of 200 octets. The video stream is transmitted from node V0 to node V1 over UDP. The TCP background sources use the TCP-Reno implementation. The TCP cross traffic is generated by FTP applications, which are active during the whole simulation, with packets of 1500 octets. A number of TCP sources are uniformly distributed among the nodes T0, T1, and T2, and send packets to nodes D0, D1, and D2, respectively. Doing so, independent TCP background traffic interferes with the video stream at each hop. The TCP sources start transmitting before the video sources. Thereby, video streams already find loaded the intermediate routers with background traffic, avoiding thus initial transient periods. Such actions provide a conservative and pessimistic scenario that reflects realistic traffic conditions.

Figure 4.3 shows the fraction of decodable frames Q and fraction of useless data received U in face of an increasing load in the network. In the simulation results, the error bars show the 99% confidence interval. We compare simulation results to the results obtained from our analytical model that has been fed with the parameters ξ and α extracted from the simulations. Besides validating the model, the results make evident how strongly an MPEG stream may be degraded in a best-effort network. Although Figure 4.3 indicates that a certain level of loss tolerance ($dt < 1.0$) may significantly improve the quality of the video transmission, we investigate hereafter the effects of the compression factor, the tolerance to losses, and the introduction of drop precedence levels.

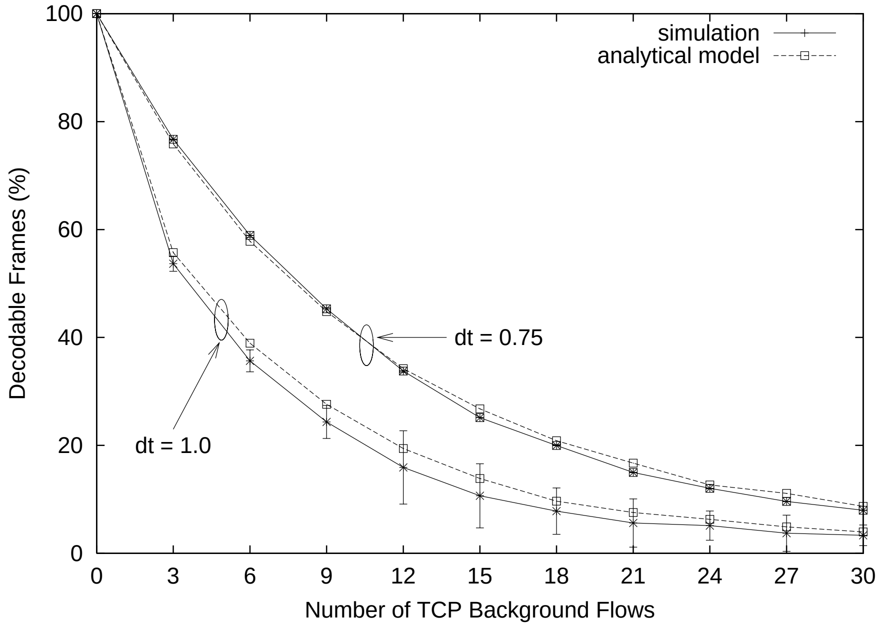
4.4 QoS and Video Communications

Our objective is to evaluate the video transmission of an MPEG stream considering the compression factor of different GOP patterns, the level of tolerance to losses, and the mapping of different types of frames onto different drop precedence levels. Furthermore, our investigation covers the effects of simultaneously adopting such QoS strategies. To perform the evaluation, we consider two classes of GOP patterns:

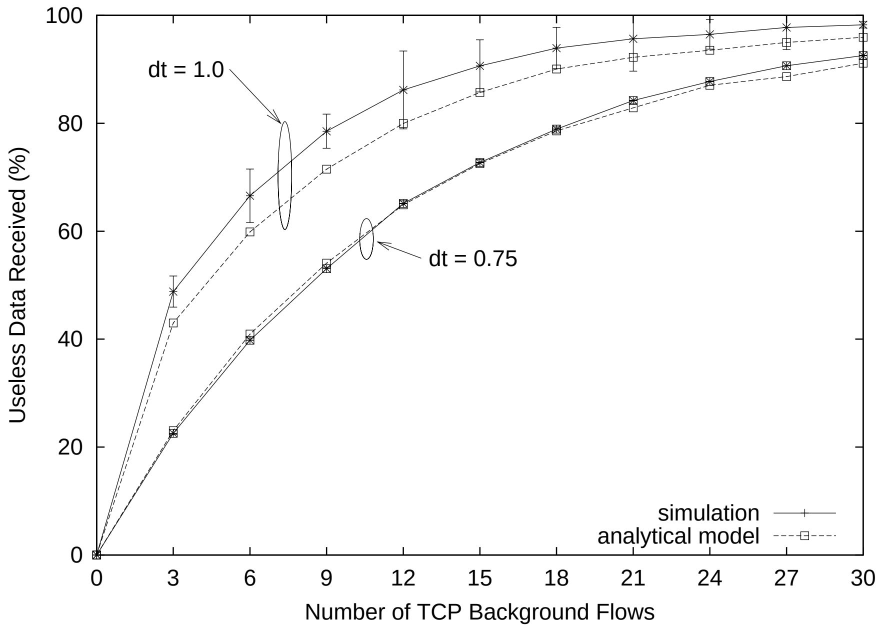
Class 1: $I(BP)^kB \quad ; k \in \{1, 2, \ldots, 7\}$
Class 2: $I(BBP)^kBB \quad ; k \in \{1, 2, 3\}$

GOPs from Class 1 have less B frames than GOPs from Class 2 for the same GOP size (N). Therefore, GOPs from Class 2 are expected to have on average a higher compression rate than comparable GOPs from Class 1 when applied to the same video sequence using a constant quantizer. By comparable GOPs, we mean two GOP patterns that despite belonging each one to one of the defined classes have the same GOP size (N).

To illustrate the quality degradation because of the hierarchical structure of an MPEG stream and the effects on quality of the compression factor on different GOP patterns, we consider the following situation. Suppose the probability ξ of directly losing a frame being the same whatever the type of the frame is ($\xi = \xi_{\{I,P,B\}}$). In this environment, the quality offered by GOPs belonging to Class 1 degrades slightly stronger when facing an increasing frame loss probability ξ than the quality offered by comparable GOPs from Class 2, as shown in Figure 4.4. As opposed to intuition, this happens even if GOPs from Class 2 are expected to be more compressed on average than comparable GOPs from Class 1. The explanation to this seemingly contradictory finding is that GOPs from Class 1 have more P frames for the same GOP size (N). As the different frames share the same probability ξ for direct frame loss, GOP patterns belonging to Class 1 are expected to have a greater probability of directly losing a P frame and thereby indirectly losing the frames that depend on the directly lost P frame as well. The effects on quality of the compression factor of



(a) Decodable frames



(b) Useless data received

Figure 4.3: MPEG video stream through a best-effort network.

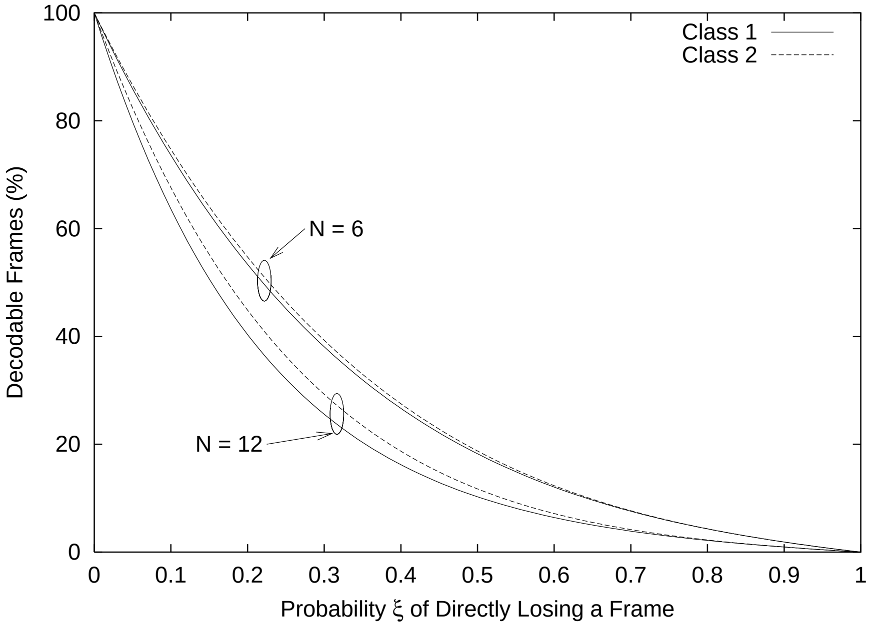


Figure 4.4: Effects of the probability ξ on comparable GOPs from Classes 1 and 2.

different GOP patterns are further discussed in Section 4.4.1.

For a given frame type $F \in \{I, P, B\}$ and considering no loss tolerance ($dt = 1.0$), the probability ξ_F of directly losing a frame F is mapped from the packet loss probability α_F by

$$\begin{aligned} \xi_F &= \sum_{k=1}^{C_F} \binom{C_F}{k} \alpha_F^k (1-\alpha_F)^{C_F-k} \\ &= 1-(1-\alpha_F)^{C_F}. \end{aligned} \tag{4.6}$$

In a non-differentiated environment, like a best-effort network, the probability α of losing a packet is the same whatever the frame content of the packet is ($\alpha = \alpha_{\{I,P,B\}}$). In such an environment, Equation 4.6 may be simplified to $\xi_F = 1-(1-\alpha)^{C_F}$. We assume packet losses to be mutually independent. Figure 4.5 illustrates the effect of the packet loss probability α on comparable GOPs from Class 1 and 2 for GOP size $N = 12$. In the case the number of packets containing data from each type of frame is the same whatever the type of frame is ($C_I = C_P = C_B = 1$ in Figure 4.5), we have the same probability ξ of directly losing the frame for all types of frames. Since less compressed frames take up more data and thus need to transmit more packets than the more compressed frames ($C_I \geq C_P \geq C_B$), if the same packet loss probability α applies to all packets, then $\xi_I \geq \xi_P \geq \xi_B$. In Figure 4.5, two particular cases ($C_I = 4, C_P = 2, C_B = 1$ to consider larger packets and $C_I = 20, C_P = 10, C_B = 5$ to consider smaller packets) are shown to depict the quality degradation of an MPEG stream in a non-differentiated environment with no loss tolerance. The most

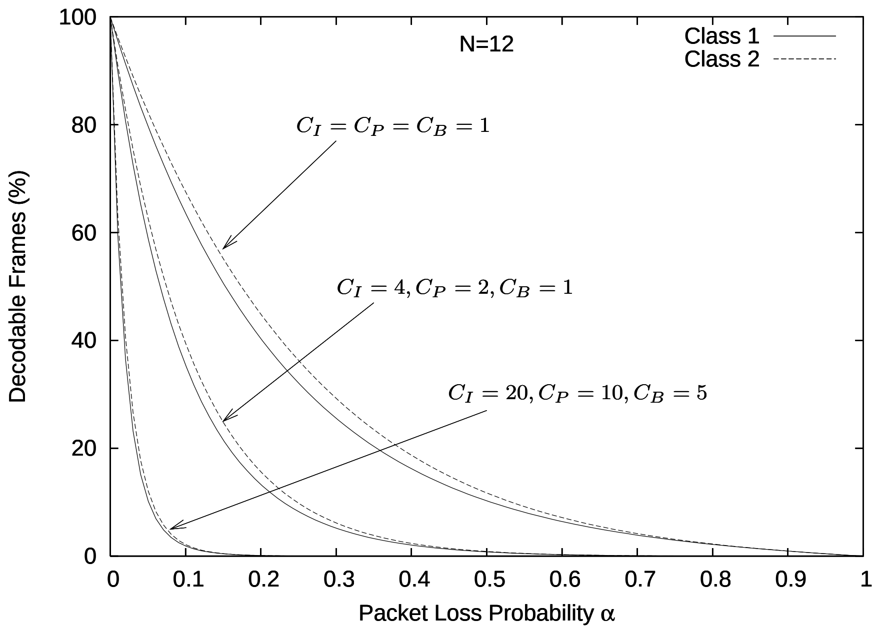


Figure 4.5: Effects of the probability α on comparable GOPs from Classes 1 and 2.

important frames suffer higher direct frame loss rates because they have more packets and the packet loss rate is the same for all packets independent of the frame information they carry. This higher probability for direct frame losses of the most important frames implies indirect losses of dependent frames, strongly compromising the delivery quality in a non-differentiated environment.

Differentiation among packets from distinct frames may be achieved by mapping those packets onto different drop precedences within an Assured Forwarding (AF) class [43]. The mapping of frames onto network priorities P_h (high priority) and P_l (low priority) covers three possibilities as shown in Table 4.2. Background traffic is always mapped onto P_l. Network priorities P_h and P_l are implemented as two different virtual queues managed in a RIO (RED with IN and OUT) scheme [52], reflecting two different levels of drop precedence within an AF class. The RIO parameters are min, max, and P_{max} (see [52, 50] for further details about the meaning of each parameter). Issues as how to best configure the RIO parameters are outside the scope of this work. Our goal when adopting the RIO mechanism is to provide a certain level of discrimination between the P_l and P_h packets that would be consistently effective, but not necessarily optimal, in the general case for any MPEG video stream.

For the cases that require a simulation applying the mapping variants of Table 4.2, routers R0, R1, and R2 (see Figure 4.2) adopt a RIO mechanism for active queue management instead of the conventional Drop Tail policy. The threshold parameters of the RIO mechanism for the two drop precedence levels are configured as percentages of the total queue length ($qlen$), which is kept equivalent to 50 packets. The RIO parameters (min, max, P_{max}) we use are ($0.2 * qlen$, $0.5 * qlen$, 0.1) and ($0.5 * qlen$, $0.8 * qlen$, 0.02) for

Table 4.2: Mapping of frames onto network priorities.

Mapping	I	P	B
M1	P_h	P_h	P_h
M2	P_h	P_h	P_l
M3	P_h	P_l	P_l

P_l and P_h packets, respectively. This set discriminates P_l packets, progressively avoiding the discard of P_h packets. We further discuss the effects of different drop precedence levels in Section 4.4.3.

For the experiments, we first take two GOP patterns, named GOP1 and GOP2:

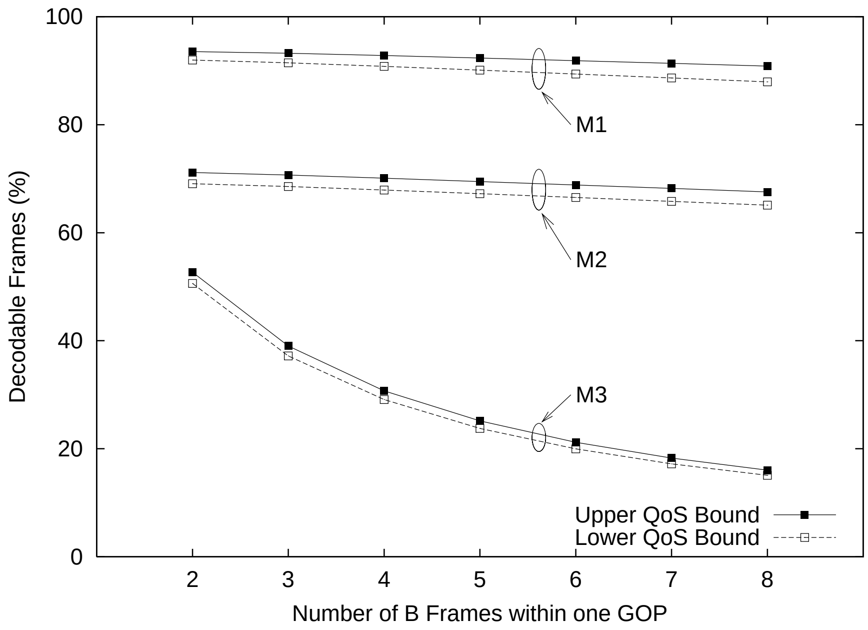
GOP1: $IBPB$
GOP2: $IBBPBBPBBPBB$

GOP1 and GOP2 correspond, respectively, to the lowest and highest compression for GOP patterns belonging to Class 1 and to Class 2. For each proposed experiment, we extract through simulation over the scenario of Figure 4.2 the parameters ξ and α for GOP1 and GOP2 when facing 30 TCP background flows, the case where quality is already heavily degraded in Figure 4.3. Then, these parameters are used to feed our analytical model and obtain the results for the large variety of remaining patterns. It should be noted that these analytical QoS predictions are done without any support by additional simulation experiments. Results are shown as a range between the highest and lowest expected picture quality Q (as defined in Equation 4.2) for the MPEG video stream under the given conditions, i.e. the analytical model provides an upper and a lower bound for the video quality to be expected by the receiving end user. The level of quality predicted analytically in this way is valid (i.e. very close to actual system behavior) indeed as we could successfully verify thanks to additional simulation experiments.

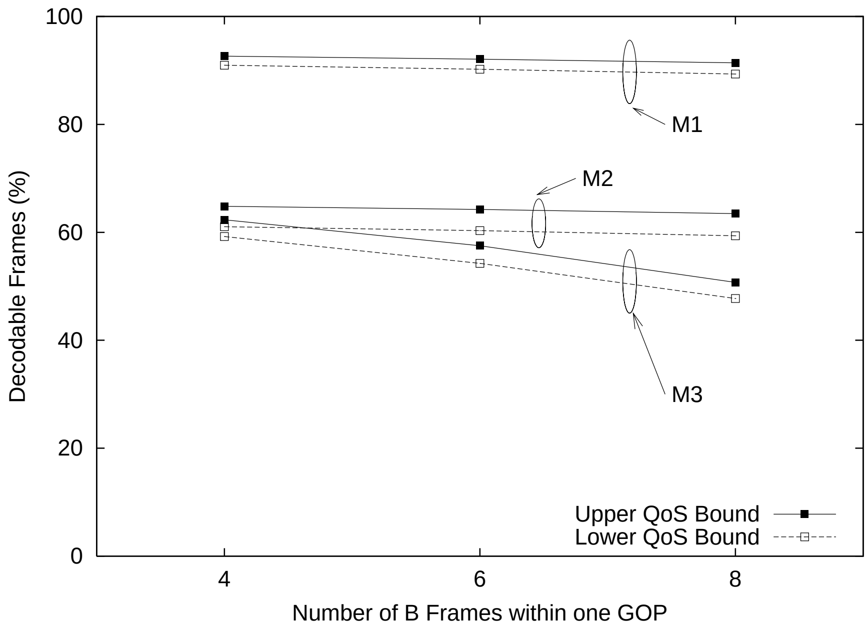
4.4.1 Effects of Compression

Experiment I compares the performance achieved by GOP patterns that have different compression factors. We consider all the GOP patterns that belong to either Class 1 or Class 2, as shown in Table 4.3. There are $k+1$ and $2(k+1)$ B frames at each GOP in Class 1 or Class 2, respectively. Figure 4.6 presents the fraction of decodable frames of patterns from each class handled by different mappings onto network drop precedences. The results presented here consider no loss tolerance ($dt = 1.0$).

The higher the compression is within each class (more B frames), the lower is the expected quality. For both GOP classes the decrease in quality due to an increasing compression is relatively small on either M1 or M2. However, a more significant degradation in quality is observed for M3. This degradation in quality because of an increasing compression is stronger in GOPs from Class 1 than in GOPs from Class 2. This points out the importance of the P frames. In Class 1, for the same number of B frames we have more



(a) Class 1



(b) Class 2

Figure 4.6: Effects of compression with $dt = 1.0$.

Table 4.3: GOP patterns for Experiments I and II.

GOP pattern	B frames
$IBPB$	$2 \times B$
$IBPBPB$	$3 \times B$
$\vdots$	
$IBPBPBPBPBPBPBPB$	$8 \times B$
$IBBPBB$	$4 \times B$
$IBBPBBPBB$	$6 \times B$
$IBBPBBPBBPBB$	$8 \times B$

P frames. Moreover, we observe the significant improvement on the expected quality of a video stream due to the protection of P frames.

4.4.2 Effects of Tolerance to Losses

In experiment II, we evaluate the QoS improvement if the decoder can tolerate a number of losses in a frame and thus allows to still consider such frames as being decodable ($dt < 1.0$). The adoption of a FEC scheme provides this feature at the expense of data redundancy. The intended FEC scheme [130] adds redundant information in a fashion that k video packets are protected by FEC in such a way that n packets ($n > k$) are transmitted, which are all of the same importance to the receiver. If at least k packets out of the n packets composing a frame are correctly received, the video information can be correctly decoded and the frame is considered decodable. This result is independent of which packets are lost within the frame.

We denote by c_0 the needed data rate to send the video stream when $dt = 1.0$. When $dt = 1.0$, the system is completely intolerant to losses and there is no redundant data in the video stream as no FEC scheme is in use. Given a certain decodable threshold (dt) value q ($0 \leq q \leq 1$), the needed capacity to send the redundant video stream is $c(q) = \frac{c_0}{q}$. Thus, considering $c(q) = c_0(1 + v(q))$, the overhead $v(q)$ imposed by the adoption of a certain level of tolerance to losses is given by

$$v(q) = \frac{c(q) - c_0}{c_0} = \frac{1 - q}{q}. \tag{4.7}$$

Similar relation between the FEC overhead and the level of loss tolerance has been found in [139] when modeling the usage of Priority Encoding Transmission (PET) [131] to transfer MPEG video streams. Therefore, to take into account the data redundancy due to the adoption of a FEC scheme, we use Equation 4.7 to consider an overhead of 33% for $dt = 0.75$ and of 100% for $dt = 0.50$ in the size of each frame. The GOP patterns are the same as for experiment I (Table 4.3). Under the M1 mapping variant, all packets from video frames are mapped onto P_h and thus the background traffic has a higher drop precedence when compared with video packets from any frame. As a consequence of not having many

Table 4.4: GOP patterns for Experiment III.

	GOP pattern	B Frames
1	$IBPB$	$2 \times B$
2	$IBPBPBPB$	$4 \times B$
3	$IBBPBB$	$4 \times B$
4	$IBPBPBPBPBPBPBPB$	$8 \times B$
5	$IBBPBBPBBPBB$	$8 \times B$

discards of video packets, the results for different values of dt almost overlap, diminishing the effects of some tolerance to losses. Under mapping variants where certain video packets and the background traffic are mapped onto the same drop precedence, the benefits of some tolerance to losses become clear as in Figures 4.7 and 4.8, which present different values of dt handled by the M2 and M3 mapping variants, respectively.

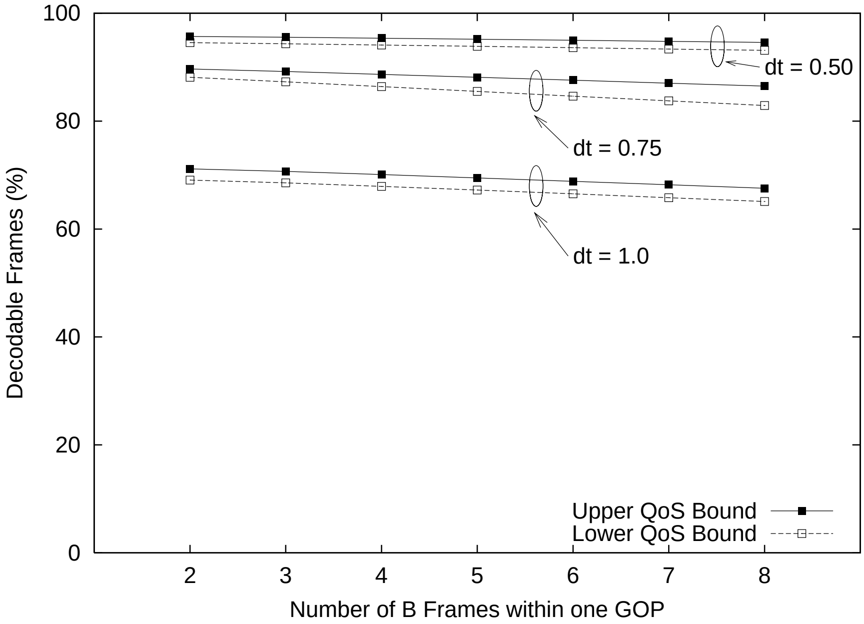
Considering the M2 mapping variant (Figure 4.7) and the same level of tolerance to losses, an increasing compression within each class of GOPs does not significantly influence the perceived quality of the delivered stream. Nevertheless, the application of a FEC scheme can strongly improve the delivery quality of the MPEG video stream (similar observation as in [132]). GOPs from Class 1 take a better advantage of a certain level of loss tolerance. This can be seen when comparing GOP patterns from the two classes with the same number of frames, for example, comparing $N_B = 6$ in Figure 4.7(a) and $N_B = 8$ in Figure 4.7(b).

For the M3 mapping variant (Figure 4.8), however, we observe that the quality offered by GOPs from Class 1 degrades more quickly with an increasing compression than the quality provided by GOPs from Class 2 (cf. Figure 4.8(a) in comparison with Figure 4.8(b)). This effect is due to the presence of more P frames in GOPs from Class 1. As P frames are not discriminated from B frames, contrasting Figures 4.7 and 4.8 shows, as expected, that the quality degrades faster under the M3 mapping variant than under M2, specially for GOPs belonging to Class 1. Once again, a larger number of P frames in GOPs from Class 1 contributes for such stronger degradation.

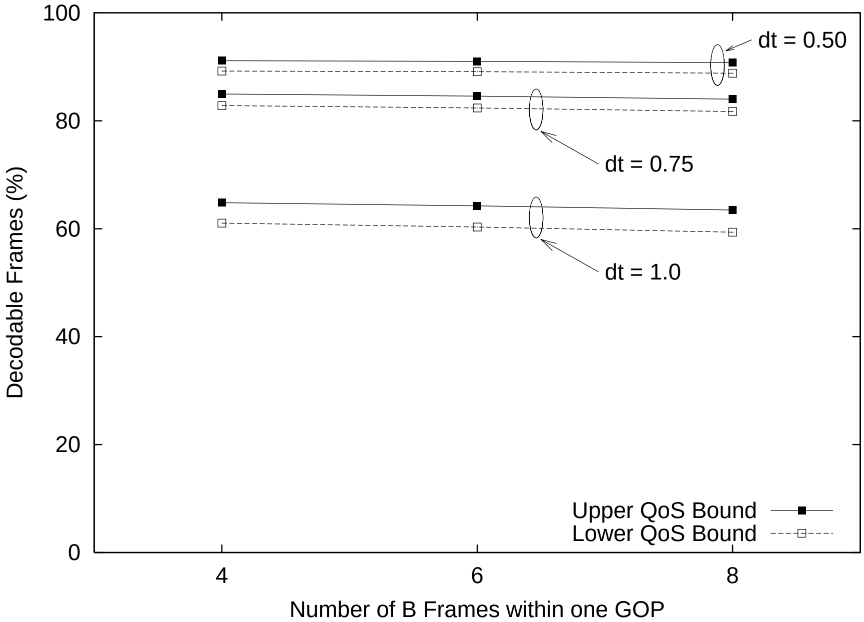
4.4.3 Effects of Drop Precedence Levels

Experiment III compares the appliance of different drop precedence levels or priority mappings (Table 4.2) to five representative GOP patterns presented in Table 4.4. Some of these GOP patterns present the same number of B frames, although distributed in a different manner depending on which class they belong to. Figure 4.9 shows the fraction of decodable frames for each pattern handled by a different drop precedence mapping and for different levels of loss tolerance.

For the same number of B frames, quality degrades more on GOPs from Class 1 than on GOPs from Class 2 as frames get a lower priority mapping. Note that for the same number of B frames, GOPs from Class 1 have more P frames. Again, this shows the importance of P frames in the hierarchical structure of MPEG. GOPs from Class 2 take a better advantage

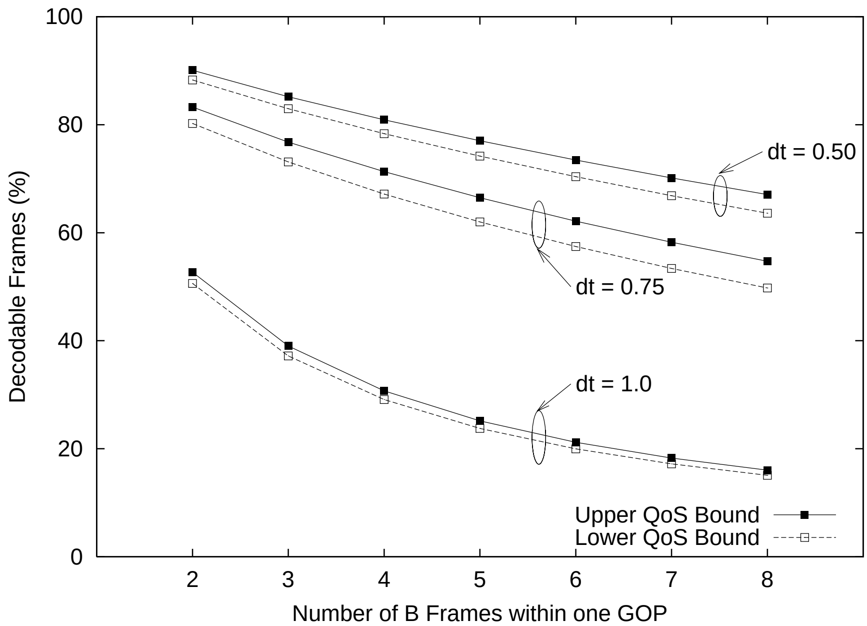


(a) Class 1

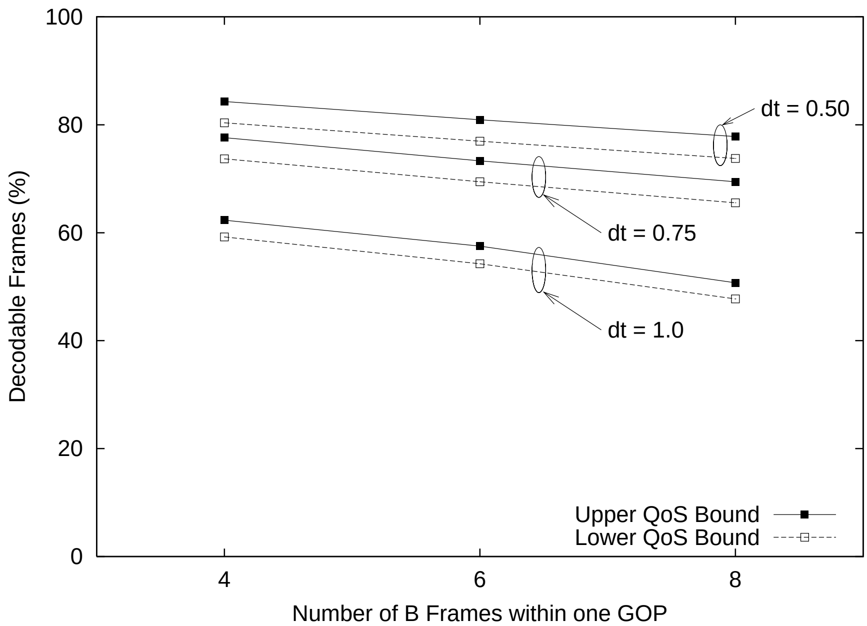


(b) Class 2

Figure 4.7: Effects of tolerance to losses for the M2 mapping variant.



(a) Class 1



(b) Class 2

Figure 4.8: Effects of tolerance to losses for the M3 mapping variant.

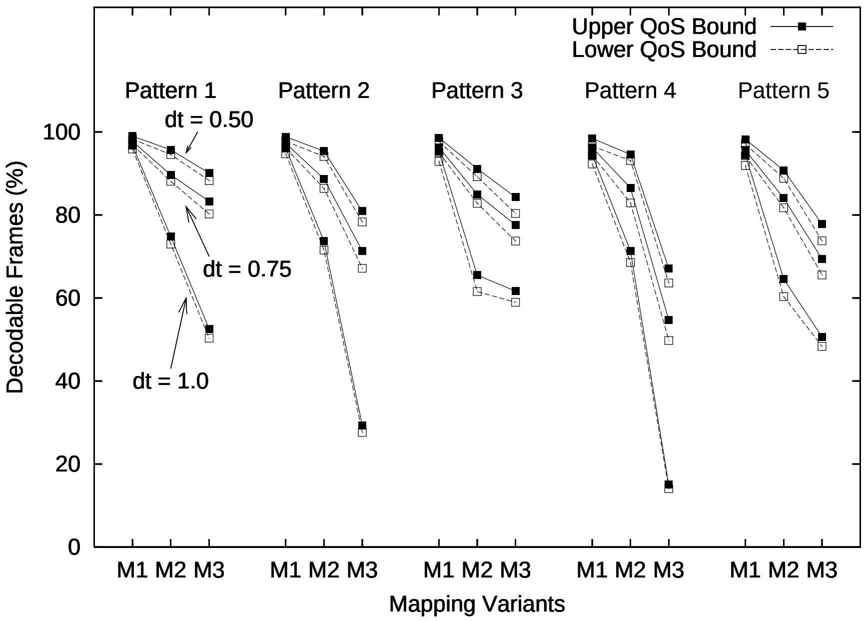


Figure 4.9: Drop precedence mappings for different GOP patterns.

of a certain tolerance to losses provided by a FEC scheme as frames get a lower priority mapping. As opposed to intuition, more compressed GOP patterns may have better quality than less compressed ones under specific circumstances (as an example, cf. Pattern 1 versus Pattern 3, for dt=1.0 and M3 mapping in Figure 4.9). This result may become more evident to the reader by considering that under a mapping as M3, a pattern like $IBBB$ is expected to get on average a better quality than a less compressed pattern like $IBPB$.

4.5 Conclusion

In this chapter, we have identified a variety of techniques that allow the improvement of the delivery quality of MPEG streams to end users. From an analytical model, we have evaluated the impact of frame losses in the quality of MPEG streams and in the waste of network resources. Our assessment has considered issues such as the use of redundancy either by applying a FEC scheme to tolerate losses or by changing the compression factor in a GOP pattern generation. We have also covered how to best use a differentiation scheme among different types of MPEG frames and how to adequately protect them from losses. Furthermore, we have evaluated the effects of simultaneously adopting different QoS strategies to improve the delivery quality of MPEG streams.

The obtained results pointed out the importance of P frames in the GOP structure. Besides protecting the most important I frames, any scheme that intends to differentially handle data from different types of frames must also strongly consider the protection of P frames to achieve a significant quality improvement. This is true, specially if the adopted

GOP pattern belongs to Class 1. The difference between the highest and the lowest quality as predicted by means of our analytical model, in order to get bounds for performance as observed on application layer, is always small and therefore negligible, which suggests the good predictability quality that may be achieved by the adopted model. Another key point is the need of adopting FEC mechanisms, as they can strongly improve the delivery quality over lossy networks. The combination of protecting more important frames (from the point of view of the hierarchical structure of MPEG encoding) with an additional tolerance to losses for the frames that are more susceptible to losses has been shown as an efficient strategy to improve the delivery quality of a video transmission. From our findings, GOP patterns belonging to Class 2 can typically achieve a stronger quality improvement by the different analyzed techniques. Hence, for most boundary conditions, we recommend the adoption of GOP patterns from Class 2 to take a better advantage of the analyzed QoS schemes.

4.6 Résumé du Chapitre

Vu son utilisation intense, le codage MPEG devient une solution attirante pour la distribution de flux audio et vidéo dans le réseau Internet. Le codage MPEG est recommandé comme le standard pour les futures applications de vidéo interactives [113]. L'Internet, étant basé sur le modèle "au mieux" (*best-effort*), celui-ci n'offre aucune différenciation parmi les paquets. Ainsi, des pertes sans différentiation de paquets dégradent fortement la qualité perçue par l'utilisateur d'une transmission vidéo MPEG. Cela arrive spécialement en raison de la structure hiérarchique du codage MPEG où une possible propagation d'erreurs peut générer d'importantes pertes de trames même en présence de taux de pertes de paquets plus faibles. Par exemple, un taux de pertes de paquets de 3% peut se traduire par une probabilité d'erreur de 30% au niveau des trames [114]. Des ressources du réseau peuvent être également gaspillées par la transmission d'informations qui deviennent inutile pour le récepteur [115]. Une fraction des données reçues peut devenir inutile pour le décodeur si les données reçues pour une trame ne sont pas suffisantes pour la décoder ou si une trame, dont la trame actuelle dépend, est undécodable. Dans ce chapitre, nous analysons de différents schémas de Qualité de Service (QdS) et les compromis résultants de l'adoption conjointe de ces schémas en visant une distribution efficace des flux vidéo MPEG.

D'après un modèle analytique, nous évaluons l'impact des pertes de trames sur la qualité des flux vidéo MPEG et sur le gaspillage de ressources du réseau. Notre évaluation vise des questions telles que l'utilisation de la redondance par l'application d'une technique de *Forward Error Correction* (FEC) pour la tolérance aux pertes, le changement du facteur de compression du codage MPEG ou la protection inégale de trames MPEG dans un environnement de Différentiation de Services (DS). De même, nous évaluons l'impact des pertes au niveau réseau dans la qualité perçue par l'utilisateur au niveau application.

Les résultats déterminent des bornes pour la qualité attendue par les utilisateurs finaux ainsi que des indications sur comment mieux profiter de l'utilisation conjointe des techniques de QdS évalués. Les résultats obtenus montrent l'importance des trames P dans la structure hiérarchique du codage MPEG. Tout schéma offrant un traitement différencié des données des différentes trames doit protéger les trames I, mais aussi accorder un importance particulière à la protection des trames P, pour obtenir un amélioration significative de la qualité. La différence entre les bornes inférieures et supérieures de qualité estimés par notre modèle analytique est faible, ce qui suggère une bonne précision atteinte par le modèle adopté. Un autre point clé est le besoin d'adoption des mécanismes de FEC à cause de leur capacité d'améliorer fortement la qualité de délivrance de la vidéo transmise à travers des réseaux présentant pertes de paquets. L'utilisation conjointe de la protection des trames les plus importantes du point de vue de la hiérarchie du codage MPEG, et de la tolérance aux pertes de paquets s'avère une stratégie efficace pour l'amélioration de la transmission avec qualité d'un flux vidéo MPEG.

Part II

Location-Awareness

Chapter 5

Location-Awareness

Knowing the physical location of a host from an identifier of that host, such as a name or IP address, enables a whole new class of location-aware applications. These applications take into account where the users are accessing from. Different location-aware applications may be conceived for different contexts like mobile ad hoc networks and the Internet. In the context of a mobile ad hoc network, examples of novel location-aware applications are: position-based filtering of content availability, navigation orientation, sending of a message to hosts within a certain region (geocasting), and geographic routing of packets. Examples of novel location-aware applications in the Internet context are: targeted advertising on web pages, automatic selection of a language to first display the content, accounting the incoming users based on their positions, restricted content delivery following regional policies, and authorization of transactions only when performed from pre-established locations.

In this chapter, we give background in providing location-awareness to networks. We first establish the difference between positioning systems and location services in Section 5.1. Motivation is illustrated by the description of some location-aware scenarios in Section 5.2. In Section 5.3, we formalize the location problem. We present in Section 5.4 an overview of the existing solutions to location-awareness in specific contexts such as cellular phone networks, mobile ad hoc networks, and the Internet.

5.1 Positioning Systems and Location Services

A positioning system allows a user or device to know its own position. The most known positioning system is the Global Positioning System (GPS) [140]. GPS satellites transmit signals that can be detected by anyone with a GPS receiver. Using the receiver, one can determine its location. The Galileo positioning system [141] is intended as an European alternative to United States GPS. Satellite-based systems such as GPS and Galileo, while very useful outdoors, are ineffective indoors because buildings block satellite transmissions. As mobile computing devices and local-area wireless networks receive increasing interest, alternative positioning systems, which are designed to work indoors, have been proposed. Examples of such indoors positioning systems are Active Badge [142] and RADAR [143].

A location service extends the notion of a positioning system. While a positioning system allows a user or device to only learn its own position, a location service allows a node to learn the position of any other participating node. For example, in a mobile ad hoc network that uses geographic forwarding of packets, a source must know the geographic position of any destination to which it wishes to send in order to label packets with that destination position. In this context, an example of location service is the Grid Location Service (GLS) [144] that is described in further detail in Section 5.4.2.4. A location service provides the capacity to indicate a node position from an identifier of that node and is fundamental to the concept of location-awareness. In this work, we focus on location services rather than positioning systems for both ad hoc networks and the Internet.

5.2 Location-Aware Scenarios

The concept of location-awareness allows the development of services that take advantage of knowing the physical location of hosts instead of only relying on identifiers of that hosts. In this section, we describe some example scenarios that are enabled by these services.

5.2.1 Geographic Routing of Packets

Suppose a location service in a mobile ad hoc network that provides to any node the physical position of any other node in the network. Nearby nodes have their position known by one-hop broadcast messages. Since a source node knows the position of all participating nodes, both nearby and remote, this source is then able to send a packet to any destination forwarding the packet to a nearby node that is physically closer to the destination node. Each intermediate node repeats this process, collaborating at every hop to get the packet closer to its destination. This kind of position-based routing is known as geographic routing [145, 146, 147, 148]. The advantage of this routing scheme concentrates on low state requirements to take a forwarding decision at each node. Geographic routing is thus commonly claimed as a promising alternative solution to the scalability problems found by other routing approaches in mobile ad hoc networks. The drawback of this approach relies on its strong dependence of an efficient location service.

5.2.2 Location-Awareness on Web Pages

Consider for instance a web site with touristic information about Paris. This web site may find interesting to sell targeted publicity to its sponsors. For example, an airline company is a potential sponsor for a touristic site since clients accessing this site are likely to be eventual customers in buying a ticket to flight to the touristic destination. Therefore, suppose airline company *XYZ* sponsors a touristic web site containing information about visiting Paris. If the touristic site is able to use a geographic location service to identify from where its clients are accessing it, the airline company *XYZ* can show up a targeted advertisement to the site visitors. The targeted advertisement might be a promotion for flights New York–Paris–New York shown to clients accessing from New York City. Similarly, the targeted advertisement would show a promotion on tickets Tokyo–Paris–Tokyo for clients accessing from Japan.

5.2.3 Restricted Content Delivery Following Regional Policies

Suppose that a great sports event is to be broadcasted live in the Internet. Nevertheless, following regional policies, the availability of such data should be restricted based on user location, like territorial rights management akin to TV broadcast rights. In other words, the great sports event should be broadcasted everywhere except to the region where it will take place. A geographic location service thus authorizes which client has access to the content based on the client's location. Likewise, content is only available to people accessing from certain regions or to users collocated with the content's server.

5.2.4 Location-Based Security Check

The success of electronic commerce is heavily dependent on customer's confidence to this system. Credit card numbers may be stolen or eavesdropped if appropriate countermeasures are not taken. Typically, such numbers may be used illicitly for electronic commerce purposes and who performs the illicit activity is not likely to be collocated with all victims. A supplemental service to help security in electronic commerce might be to perform a location-based security check before authorizing a transaction to be done. For example, a customer may authorize the use of its credit card to electronically buy something only when this transaction is requested from the customer's hometown. If this customer is about to travel, the administrator of the credit card may be previously warned and accordingly configure the location-based security checks following instructions provided by the customer. Since authorized locations are known, a transaction requested from elsewhere might generate warnings on untypical or unauthorized behavior of a customer.

5.3 Formalization of the Location Problem

Position information about nodes is the new element when networks become location-aware. The establishment of a relation between previous components of the network, such as addresses, and the newcomer location information will be called hereafter the location problem.

As a packet crosses a network from a source node to a destination node, the packet is successively forwarded at intermediate nodes. The forwarding decision at each intermediate node is taken by a routing function to select the best next hop node to send the packet to. This routing function bases its forwarding decision on data collected from the network and, typically, on an identifier of the destination. The maintenance of updated data available to the routing function and the criteria to select the best next hop node are a duty of the adopted routing protocol. From a high-level perspective, the main components here are the identifiers of nodes and the routing procedure. In a location-aware environment, however, a third component is added: the location information. Therefore, we define three main components in the general context of location-awareness:

- Identifier - a means of identifying a host without any ambiguity. A host may have more than one identifier. An identifier is represented by an element of any coherent structure of numbers, codes, or names. Different identifier structures may be simultaneously present to identify the same host. For example, in the Internet context, both the IP address and the name of the host are identifiers of the host.

- Routing - the process of finding a path from a source to every destination in a network. The routing procedure is carried out at each intermediate node to select the best next hop node to forward a packet to. The sequence of chosen nodes determines the path followed by the packet through the network.

- Location - the physical position of a host. The location is not necessarily fixed over time, as in the case of mobile hosts. Location information may be represented in several formats. For example, positions on an absolute or relative coordinate system, pairs of latitude and longitude, postal codes, names of streets, cities, provinces, or countries, are all considered valid location data.

Let I denote the set of all identifiers. Since the routing decision is based on a given identifier of the destination and on the topologically nearby nodes, we represent the routing decision at each node as a routing function $g_I : I \rightarrow I$. We represent by $N \subset I$ the subset containing all identifiers of nodes in the path of a packet from the source node to the destination node of the packet. Therefore, $N = \{n_1, n_2, \dots, n_m\}$ is the permutation, i.e. ordered subset, identifying the sequence of m nodes crossed by a packet in its path from the source node (n_1) to the destination node (n_m). The routing function g_I at each node $n_x \in N$, $x = 1, \dots, m-1$, selects from a set of possible next hop nodes the best option $n_{x+1} \in I$ to forward the packet. This routing function g_I is based on an identifier $i \in I$ of the destination n_m. Hence, at a node n_x, the next hop node is $n_{x+1} = g_I(i)$.

To enable location-aware applications, a relation between the identifier of a host and the current physical location of the host must be established. Let L denote the set of all locations. The function $f_L : I \rightarrow L$ maps an identifier $i \in I$ of a node to the current location $l \in L$ of the node. The availability of a relation between identifiers and locations of hosts allows, for example, the geographic routing of packets in a network. For a given destination identifier $i \in I$, the source of a transmission learns the location $l \in L$ of the destination by applying $l = f_L(i)$ and thereby using a location service. The location-based routing decision at each node is then represented by the routing function $g_L : L \rightarrow L$. Thus, the routing decision in geographic routing becomes a function of the location of the destination instead of a function of the identifier of the destination.

The function $f_I = L \rightarrow I$ answers the question of which nodes are at a given location or, alternatively, at a given area around this location. This function returns a set $W \subset I$, where the elements of W allow finding out who are the nearby nodes or the set of nearby nodes providing a service at a given range. Similarly, the discovery of who occupies a specific geographic area or the set of nodes providing a certain service at a specific geographic area is also made available. Geocasting [149] implements this function to identify the nodes belonging to the target area to where a message is being sent.

Figure 5.1 summarizes in a diagram the defined functions for the identifier (I) and the location (L) sets. The location relation $\mathcal{L} \subset I \times L$ is established. The relation $\mathcal{L}$ assigns locations to identifiers. Thus, the pair $(i, l) \in \mathcal{L}$ relates for a given identifier $i \in I$ of a

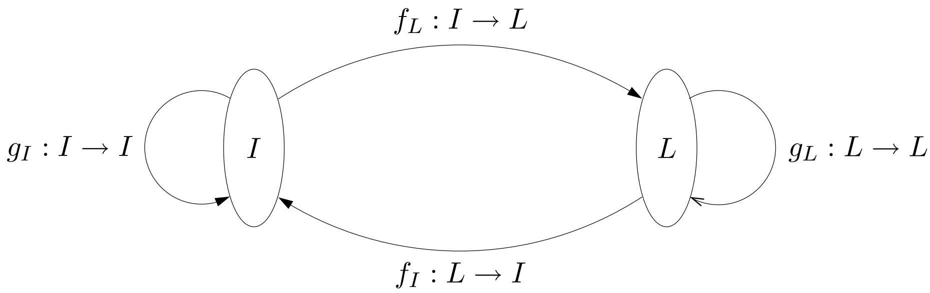


Figure 5.1: Diagram of functions for the identifier (I) and location (L) sets.

node n, the location $l \in L$ of the node n. Likewise, the identifier relation $\mathcal{I} \subset L \times I$ provides the identifiers of nodes occupying a given location.

5.4 Location-Awareness in Specific Contexts

The definition of the mapping functions f and the routing functions g that act on the identifier set I and on the location set L provides a high-level formalization of the location problem. When applying the model to specific contexts, with different and particular features, the implementation of each function changes. From this general perspective, we consider three specific contexts for applying location-awareness: GSM cellular systems, mobile ad hoc networks, and the Internet.

5.4.1 GSM

The Global System for Mobile Communication (GSM) [150] adopts a structure of databases to track user's mobility. In the GSM cellular system, each user has an unique identifier, the International Mobile Subscriber Identity (IMSI) number. Each wireless carrier keeps a database called Home Location Register (HLR) with mobile user information. Similarly, the Visitor Location Register (VLR) is a database that contains temporary information concerning the mobile users that are currently in the area, but have HLRs elsewhere. When a mobile user moves into a new location area, the new VLR updates the current location information in the correspondent HLR database of the mobile user. Calls directed to a given mobile user arrive to the carrier controlling the subscription of the user. The HLR of the subscriber carrier knows the current VLR in which the mobile user is registered. This operation maps a given IMSI number $i \in I$ onto a location $l \in L$ associated with the current VLR of the mobile user, implementing thus the location function $f_L : I \to L$.

5.4.2 Mobile Ad Hoc Networks

The high rates of change in the topology of an mobile ad hoc network, due to node mobility, raise important problems to conventional topology-based routing protocols. Such conventional routing protocols are implementations of the routing function $g_I : I \to I$ in our notation. In mobile ad hoc networks, nodes that are physically close are likely to be close in the network topology. Geographic routing [145] takes advantage of this similarity to base their forwarding decision on location information instead of identifier information. Hence, the geographic routing protocols are implementations of the routing function $g_L : L \rightarrow L$. Although the geographic routing of packets is the main motivation behind the location services in mobile ad hoc networks, other location-aware applications may be developed if the position of nodes is available, such as geocasting and navigation orientation [151].

Providing a location service in the context of mobile ad hoc networks is a difficult problem. When a source node requests the position of a destination node, the source only knows an identifier of the destination node. There is no static relation between the node location and the node identifier as a consequence of node mobility. The location service itself must do its task using geographic routing. The problem is to establish a dynamic relation between the identifier of the destination node and the physical location of the destination in order to allow the geographic routing of packets towards that destination. When a source needs to learn the location of a given node a, the source contacts a location server of the node a for the desired information. The position obtained by the source is the position registered within the location service by the node a. A location service in a mobile ad hoc network is therefore an implementation of the function $f_L : I \rightarrow L$.

To allow the sources to learn the physical position of a specific destination, a location service is needed. As mobile nodes move, they must update their positions in the service. When a node wants to obtain the position of a certain node, the node requests that information to the service. In classic cellular networks, like GSM, centralized position servers maintain position informations about the nodes in the network. In mobile ad hoc networks, if this server were part of the ad hoc network itself there would be a problem to learn the position of the server. A query would become a "chicken-and-egg" problem: without a position server, it is not possible to get position information; meanwhile without position information, the server can not be reached. Furthermore, due to the dynamics of ad hoc networks, there is no guarantee that at least one fixed location server will be reachable in a given ad hoc network. In the following, the main proposals for decentralized location services in ad hoc networks are presented.

5.4.2.1 Distance Routing Effect Algorithm for Mobility (DREAM)

Within the Distance Routing Effect Algorithm for Mobility (DREAM) framework [152], each node keeps a database that stores position information about each other node in the network. The frequency that each node sends an update of its position to each other node in the network is determined by the distance separating the nodes and the mobility rate. This distance considers what is called distance effect: the greater the distance separating two nodes, the slower they appear to be moving with respect to each other. Therefore, nodes provide more accurate position information to nodes that are closer to them than to nodes that are far away. This is achieved by sending update messages with an "age" indication that corresponds to how far from the sender that message travels before it is discarded (spatial resolution). The mobility rate considers that the faster a node moves, the more often it communicates its location (temporal resolution). Using spatial and temporal resolution, a

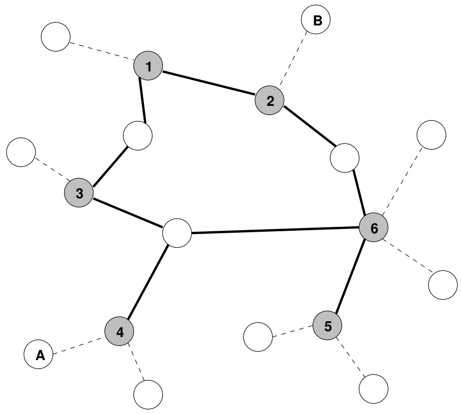


Figure 5.2: Quorum system.

node controls how far and how often it must update its location.

The DREAM approach requires that all nodes maintain position information about every other node. Thus, each mobile node floods its own position information at different update rates for distinct parts of the entire network. Since the process uses global flooding, it is not suited for large networks.

5.4.2.2 Quorum-based System

A quorum-based location management is proposed by Haas *et al.* [153]. A quorum system is constituted by organizing objects, position databases in this case, into sets (called quorums), where every two quorums intersect and no quorum contains another quorum. A subset of all mobile nodes composing the network is chosen to host the position databases. A virtual backbone is built between this subset of nodes using a non-position-based ad hoc routing mechanism. The nodes belonging to the virtual backbone compose the subsets (quorums) that host the position databases.

Figure 5.2 illustrates the backbone structure in an ad hoc network. In this example, nodes one through six are chosen to host position databases. The virtual backbone built between these nodes is represented by the solid lines. The location databases are then arranged into quorums. From figure 5.2, an example quorum system is: {{1,2,3}, {1,4,5}, {2,4,6}, {3,5,6}}. Note that every two quorums intersect. To update its location, a mobile node, say A for instance, contacts the closest node that belongs to the backbone set: the one with location database 4, in this case. The backbone node selects one of the quorums, {2,4,6}, to receive the update associated with a sequence number. When another mobile node, say B, wants to obtain the position information about A, B sends a message to the closest node that belongs to the backbone set, in this case, node 2. The backbone node 2 queries a quorum, e.g. {1,4,5}, for the location of A. In this case, location database 4 sends back to B the location information of A. Two consecutive query or update operations performed by the same mobile host may not access the same quorum either. Thus, it is possible that B receives an older copy of the location information of A. Then, it is up to B to verify the sequence numbers in all received messages and pick up the latest one. The position of the destination node is provided to the source node by the databases at the intersection between the querying quorum and the updated quorum used by the destination host in one of its last update operations.

The quorum-based system uses location servers that form a virtual backbone distributed among the network nodes. A major drawback of this approach is its dependency on a non-geographic routing protocol to build the virtual backbone, which greatly increases the implementation complexity and may compromise the scalability of this approach.

5.4.2.3 Home-Agent Based Approach

The location update in the Home-Agent based approach [154] for ad hoc networks is inspired in the one used in cellular phone networks and mobile IP. When a phone user moves away from his home server (agent) to a new location, it periodically sends its current location to its fixed home agent. When a phone call is made to that user, the call is first sent toward the user's home agent. The home agent then directs the call toward the user's current position.

In the Home-Agent based approach for ad hoc networks, a node uses its identifier to feed a hash function and obtain a position in the deployment region. Nodes around the selected position act as location servers for the node. Likewise, nodes wanting to learn the position of a node from the node identifier use the same hash function to know where to search for the information. The region selected by the hash function may be viewed as a rendezvous point between the location updates of a node w and the nodes looking for the position of node w. The SLURP proposal [155] adopts a similar strategy to deal with the location management in ad hoc networks. The Tribe protocol [156] uses a similar scheme to implement an indirect routing strategy for wireless self-organizing networks.

5.4.2.4 Grid Location Service (GLS)

The Grid Location Service (GLS) [144] divides the area containing the ad hoc network into a hierarchy of squares. The smallest square is referred to as an order-1 square. A group of four order-1 squares composes an order-2 square, and so on. To avoid an overlap, a particular order-n square is part of only one order-$(n + 1)$ square. Thus, there is an invariant: a node is located in exactly one square of each size. This system of increasing square sizes provides a context in which a node selects fewer and fewer location servers at greater distance. Each node knows the positions of the neighboring nodes that share with it the same order-1 square. Each node replicates its position information in a subset of the network nodes that act as the node location servers.

The strategy to select the nodes that act as location servers for a given node adopts a notion of near node IDs, defined as the least ID greater than a node's own ID. The ID space is considered to be circular, i.e. the ID numbers wrap around after the highest possible ID. To demonstrate how the mechanism works, an example of the GLS scheme is provided in Figure 5.3. When node 14 wants to distribute its position information, it sends position updates to the nearest ID in each of the three surrounding order-1 squares. Thus, the

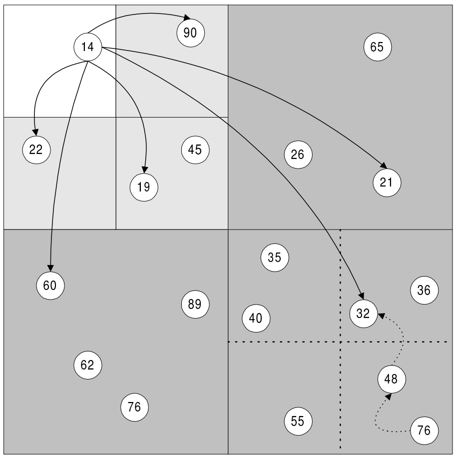


Figure 5.3: Example of the Grid Location Service (GLS).

position information concerning node 14 is available at nodes 19, 22, 90, and at all nodes that share the order-1 square with node 14. In the surrounding order-2 squares, nodes 21, 32, and 60 are chosen because they have the nearest ID with respect to 14 in their respective order-2 squares. This process is repeated at increasing order squares until all the area of the ad hoc network is covered. Therefore, the density of location servers for a given node decreases logarithmically with the distance from that node. Actually, a node routes update packets to its location servers without knowing their identities. Suppose that a node N wants to recruit a location server in some order-n square. Then, node N sends a packet, using geographic forwarding, to that square. The first node belonging to the square that receives the packet from N progressively forwards the position information to nodes with closer IDs in a process that resembles a query for node N location. When the information reaches the correct node, the position of N is stored.

The query process for a node location is performed as follows. Suppose that node 76 in Figure 5.3 wants to learn the position of node 14. It begins to search a nearby node that has the information about node 14. In the example, the closest location server to node 76 that hosts position information about node 14 is node 32. Nevertheless, node 76 ignores that node 32 holds the required information, but it is able to discover this information. Node 76 knows all the nodes in the order-1 square where it is located. Therefore, if any node in this square knows the position of node 14 it would be the node with the closest ID to 14, node 48 in this case. Node 76 queries node 48 for the position information about node 14. Since node 48 does not host the requested position, it must query another node for the information. Note that node 48 is a location server in its order-1 square for all nodes belonging to all surrounding order-2 squares that have IDs between 14 and 47, in this case nodes 32, 35, 36, and 40. As a consequence, node 48 is aware of the closest node to node 14 in its order-2 square. Node 48 thus queries node 32 for information about node 14 and, in fact, node 32 hosts the information. Thus, there is a “path” of descending node IDs from each of the squares of all orders to a valid location server.

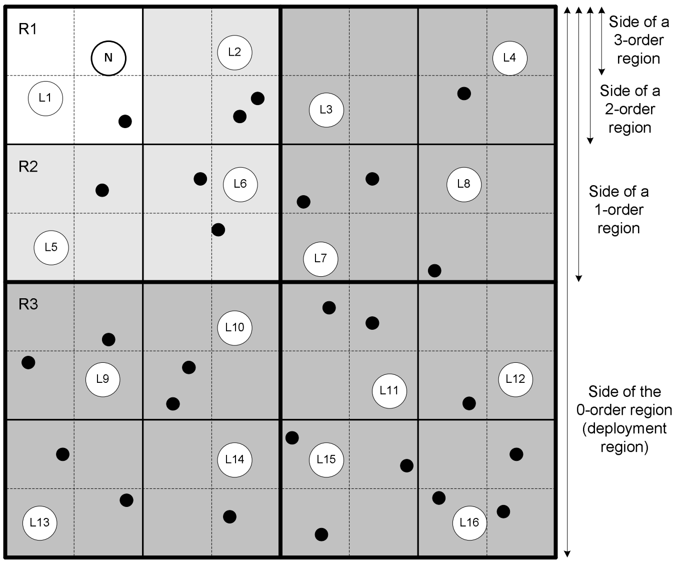


Figure 5.4: Example of the Distributed Location Management (DLM).

5.4.2.5 Distributed Location Management (DLM)

Similarly to the Grid Location Service (GLS), the Distributed Location Management (DLM) scheme [157] proposes to partition the entire network into a hierarchical grid. For a given node N in the network, a subset of nodes is selected to act as location servers for node N. Each of the chosen location servers is responsible for a region of the grid and is periodically updated with node N location. Nevertheless, instead of selecting the location server in a certain partition as the one that has the closest ID to node N ID, DLM uses a hash function to map node N ID directly into a set of partitions in the network. Then, any node present in a selected partition may act as a location server for node N. The DLM proposition specifies the notion of addressing the location of a given node by a hierarchy of different order regions. As shown in Figure 5.4, the 0-order region covers the entire deployment region. Each n-order region is partitioned into $a_n \times b_n$ $(n+1)$-order regions. In the example, $a_0 = 2$ and $b_0 = 2$. DLM also introduces the notion of the smallest common region of two nodes as the smallest region that is able to contain both nodes. For example, in Figure 5.4, the smallest common region of node N and L_1 is the 2-order region R_1, and the smallest common region of N and L_{15} is the entire deployment region (0-order region).

The partitioning scheme of DLM allows to identify the location of a node by pointing out in which region the node resides. Thus, the addressing approach in DLM labels regions in accordance with their orders and locations. The relative address of a $(n+1)$-order region with respect to its n-order region is a string (i, j), where $i \in \{0, 1, \ldots, a-1\}$, $j \in \{0, 1, \ldots, b-1\}$. Therefore, (i, j) represents the relative position of this $(n+1)$-order region in the n-order region with the upper left corner as the origin point. For example, in Figure 5.4, the relative address of the 3-order region where node N is located with respect to the 2-order region R_1 is (0,1). The relative address of the 1-order region R_2 with respect to the 0-order region R_3 is (0,0). The complete address of a region is the concatenation of the relative addresses from the regions of different orders wherein the referred region is located. For example, the complete address (1,0).(0,1).(0,1) refers to the 3-order region where node L_{10} is located.

The density of location servers is one server for each m-order region. Assuming $m = 2$, node N in Figure 5.4 has as location servers nodes $L_1, \ldots, L_{16}$. Node N complete address is (0,0).(0,0).(0,1). If the smallest common region of node N and its location server L_i is an n-order region, then L_i only stores the $(n + 1)$ partial address of node N. For example, as the smallest common region between node N and L_8 is the 0-order region, then L_8 only stores N's 1-level partial address which is (0,0). The same happens to all location servers outside the region R_2. The location servers L_2, L_5, and L_6 store the 2-level address of node N, that is, (0,0).(0,0). Just L_1 stores the complete address (0,0).(0,0).(0,1). Each time node N moves, it only needs to update a subset of its location servers. The position information is kept at decreasing accuracy levels as the distance from the node's region to the location server's region increases.

Suppose that a given node A wants to obtain the position of node N. Node A is capable of obtaining the location of the subset of N's location servers by carrying out the same hash function used by node N to select its location servers. Then node A sends its request to the nearest location server of node N. The position information stored in the location server indicates at a possible range of accuracies the region where node N resides. Following these indications, the query message is able to search for location servers that store information of increasing accuracy until it reaches node N and learns the accurate location information about node N.

5.4.2.6 Considerations on Location Services for Ad Hoc Networks

Currently, there are two main approaches to deal with the location management and scalability issues in mobile ad hoc networks [158]: the home zone approach and the hierarchical approach. The home-agent based approach and the SLURP protocol are examples of the former, where a zone in the deployment region is chosen to be a rendezvous point between location updates and location queries. The latter hierarchically divides the deployment region. Nodes located at different portions of the hierarchy act as location servers for other nodes. Thus, for each node, some other nodes act as location servers distributed in different density levels in the hierarchical structure throughout the deployment region. Examples of such an approach for location management in mobile ad hoc networks are GLS and DLM. Cheng *et al.* [159] propose a hybrid approach to combine the strengths of SLURP and GLS. Multiple home zones are assigned to each node in a hierarchical fashion throughout the deployment region aiming to reduce the volume of update messages traveling long distances in the network.

Under both the home zone and the hierarchical approaches, a source wishing to learn the position of a destination node sends a dedicated location query packet to one or more location servers. The location server then forwards the request according to its most updated location information toward the destination to obtain the most accurate location information for the source from the destination itself. After receiving a query packet forwarded by a location server, the destination node sends its current position to the requesting source node. Upon receiving the position of the destination from the destination itself, the source may then label the packets and start transmitting the flow through geographic routing. In Chapter 6, we propose a TCP-tailored approach that reduces location management costs for TCP flows in mobile ad hoc networks.

5.4.3 Internet

Another location problem is to discover the geographic position of a fixed node on the Internet. In the current Internet, however, there is no direct relation between a host identifier, such as a name or IP address, and the host geographic location. The novel location-aware applications then require the deployment of a geographic location service for Internet hosts. A DNS-based approach was proposed in RFC 1876 [160]. Nevertheless, the DNS-based approach is not widely adopted since it requires changes in the record structure of the DNS records and administrators have no motivation to register location records. The adoption of geographic distance between Internet hosts as a network distance metric is evaluated in [161]. Lakhina *et al.* [162] investigate the geographic location of Internet components to create a base for the development of geographically-driven topology generation methods.

Here we review some techniques to estimate the geographic location of Internet hosts. The location estimation offers an approximate location with an associated error. Since an estimation of the host's location is provided, the analyzed techniques are implementations of an estimator $\widehat{f_L}$ to the location function $f_L : I \rightarrow L$. For a given identifier $i \in I$ of a host, the estimator $\widehat{f_L}$ provides an approximate location $\widehat{L} \subset L$ for the host, where the correct location $f_L(i) \in \widehat{L}$.

5.4.3.1 Queries on Whois Databases

Tools such as IP2LL [163] and NetGeo [164] query Whois databases in order to obtain the location information recorded therein to infer the geographic location of a host. Such an information, however, may be inaccurate and stale. Moreover, if a large and geographically dispersed block of IP addresses is allocated to a single entity, the Whois databases may contain just a single entry for the entire block. As a consequence, a query onto the Whois databases provides the registered location of the entity that controls the block of IP addresses, although the concerned hosts may be geographically dispersed.

5.4.3.2 Location Inference from DNS names

This technique infers the geographic location of a host based on the DNS name of the host or another nearby node, being the base of GeoTrack [165], VisualRoute [166], and GTrace [167]. Quite often network operators assign names to routers that have some geographic meaning, presumably for administrative convenience. For example, the name `bcr1-so-2-0-0.Paris.cw.net` indicates a router located in Paris, France. Likewise, a router name may have an indication of location, such as an abbreviation for a city, state, or country name. Some ISPs even use the airport code of the city they are located in.

Airport codes follow a worldwide standard and thus such a naming convention eases the task of determining a router's location. A router is said to be recognizable if its geographic location can be extracted from its DNS name. Using the `traceroute` tool, the network path between a probe machine and the target host to be located is determined. The technique then extracts the location information from the DNS names of the intermediate routers along the path whenever it is possible, thus tracing a geographic path to the target host. The location estimation of the target host is the location of the last recognizable router in the path, i.e. the closest router to the target, if not the target itself.

The location inference from DNS names have some important drawbacks. First, not all names contain an indication of location, limiting the number of recognizable routers. Second, since there is no standard, operators commonly develop their own rules for naming their routers even if the names are geographically meaningful. Therefore, the parsing rules to recognize a location from a node name must be specific to each operator. The creation and management of these rules is a challenging task as there is no standard to follow. Third, since the position of the last recognizable router in the path toward the host to be located is used to estimate the position of such a host, a lack of accuracy is also expected.

5.4.3.3 Location Inference from Clustering the IP Address Space

In this technique, the IP address space is splitted into clusters such that all hosts with an IP address within a cluster are likely to be co-located. Knowing the location of some hosts in the cluster and assuming they are in agreement, the technique infers the location of the entire cluster. An example of this technique is GeoCluster [165]. The basic approach adopted by GeoCluster to determine a geographic cluster is to associate topological clusters to geographic clusters as follows. Routing and address allocation in the Internet is hierarchical. An autonomous system (AS) aggregates routing information of hosts that are under its administration. Then the routes for hosts within an AS are typically advertised to the rest of the Internet as a single aggregate, identified by an address prefix (AP). It should be noted that a single AS may advertise more than one AP due to some specific policy or performance considerations like load balancing. Information about the APs are derived from the Border Gateway Protocol (BGP) used for inter-AS routing in the Internet. The technique assumes that an AS and its associated APs often correspond to a geographic cluster. Even if an AS covers a large geographic area, the associated APs that are advertised by BGP may provide more specific clusters. In summary, the base GeoCluster algorithm, discovers APs from BGP data and assumes that these APs are geographic clusters. If the location of some hosts in the same geographic cluster is known, this location is used as a location estimation of the host to be located. Padmanabhan and Subramanian [165] also propose a subclustering algorithm to deal with a large AS that may spam a large geographic area.

Clustering the IP address space to geographically locate an Internet host, however, relies on information that tends to be partial and possibly inaccurate. The information is partial because it comprises location information for a relatively small subset of the IP address space. Moreover, such an information may be inaccurate because the databases rely on data provided by users, which may be unreliable to provide correct location information.

5.4.3.4 Measurement-Based Location Inference

The measurement-based location inference is based on delay measurements and the exploitation of a possible correlation between geographic distance and network delay. This technique is the base of GeoPing [165]. The correlation between geographic distance and network delay if considered worldwide is weak to moderate [168], although we show in Section 7.2 that it tends to be stronger in specific regions with richer connectivity. The location estimation of a host is based on the assumption that hosts with similar network delays to some fixed probe machines tend to be located near each other. This assumption is similar to the one exploited by wireless positioning systems such as RADAR [143] concerning the relationship between signal strength and distance. Therefore, given a set of landmarks with a well known geographic location, the location estimation for a target host to be located is the location of the landmark presenting the most similar delay pattern to the one observed for the target host.

5.4.3.5 Considerations on Location Techniques for the Internet

Proxies impose a fundamental limitation on all location techniques that depend on the client's IP address [165]. Many web clients access the network from behind proxies and firewalls. Therefore, the IP address that a client presents to the external network may actually correspond to a proxy. A client and a proxy may be in close proximity, as in the case of a caching proxy in a university campus. In some cases, however, they may be apart, as in the case of some large ISPs that concentrate a cluster of proxies for their clients in a centralized location no matter where the clients are accessing from. Since the IP address seen by the external network is the address of the proxy, the location techniques infer the geographic location of the proxy, which may be inaccurate in the case the client and the proxy are not in close proximity.

We focus on the measurement-based inference of the geographic location of Internet hosts. An alternative to performing direct delay measurements to compose the delay patterns is to estimate the delay between the concerned hosts. Several frameworks for estimating the delay between Internet hosts have been recently proposed, such as IDMaps [169], Global Network Positioning (GNP) [170], King [171], Lighthouses [172], Big-Bang Simulation (BBS) [173], Virtual Landmarks [174], and Internet Coordinate System (ICS) [175]. These frameworks can be used as an underlying infrastructure to provide estimations of the delay patterns from the fixed probe machines to the landmarks and to the host to be located. In Chapter 7, we present the measurement-based location technique in further detail, formalizing the measurement-based geographic location problem. We also address issues such as the placement of landmarks and probe machines as well as how to measure the similarity between the measured delay patterns.

5.5 Résumé du Chapitre

La possibilité de connaître la localisation physique d'un hôte en fournissant uniquement son identifiant ouvre la voie pour une classe toute entière d'applications conscientes de la localisation des hôtes. De différentes applications conscientes de la localisation peuvent être conçues dans de différents contextes, tels que les réseaux ad hoc mobiles et l'Internet. Dans le cas des réseaux ad hoc mobiles, des exemples de nouvelles applications conscientes de la localisation sont : le filtrage du contenu basé sur la localisation, l'envoi d'un message aux nœuds présents dans une certaine région (*geocasting*) et le routage géographique de paquets. Dans le contexte de l'Internet, des exemples de nouvelles applications conscientes de la localisation sont : la publicité ciblée sur des sites web, la sélection automatique de la langue, le contrôle d'accès au contenu selon des politiques régionales ainsi que l'autorisation des transactions commerciales électroniques seulement lorsqu'elles sont réalisées à partir des endroits spécifiques.

Dans ce chapitre, nous formalisons le problème de la localisation d'un point de vue général. Nous définissons les relations entre la nouvelle information de localisation et les composants traditionnels comme les identifiants d'hôtes et le routage entre eux. Lorsqu'on applique le modèle général à un contexte spécifique, ayant des caractéristiques particulières, l'implémentation spécifique d'un service de localisation change. Nous présentons alors l'état de l'art des services de localisation comprenant trois domaines différents : les réseaux téléphoniques GSM, les réseaux ad hoc mobiles et l'Internet. Le GSM adopte une structure de bases de données pour la gestion de la localisation des utilisateurs. Chaque opérateur garde deux bases de données : la première enregistre où se trouve ses utilisateur et la deuxième enregistre les utilisateurs visitant sa zone de couverture. Lorsqu'un utilisateur se déplace, il enregistre auprès de la base de données de son opérateur d'origine sa position actuelle. Un appel destiné à un utilisateur est d'abord acheminé vers son opérateur qui à son tour l'achemine vers l'opérateur visité.

La structure fixe utilisée par le GSM ne peut pas s'appliquer aux réseaux ad hoc mobiles vu leur manque d'infrastructure fixe. Les principales propositions à la gestion de la localisation dans les réseaux ad hoc mobiles sont l'approche *home zone* et l'approche hiérarchique. La première établit par une fonction de hachage à partir de l'identifiant d'un nœud A, une région où les nœuds appartenant à cette région seront les serveurs de localisation d'A. La proposition *Home-Agent* [154] est un exemple d'une telle approche. L'approche hiérarchique partage la région de déploiement sur une hiérarchie des serveurs de localisations. Les serveurs de localisation d'un certain nœud sont distribués à différents niveaux de densité à travers la structure hiérarchique. Le schéma GLS (*Grid Location Service*) [144] est un exemple d'approche hiérarchique à la gestion de localisation dans les réseaux ad hoc mobiles.

Dans le contexte de l'Internet, de différentes propositions existent pour estimer la localisation géographique d'un hôte. Une approche basée sur le DNS a été proposée par le RFC 1876 [160]. Cependant, cette approche n'est par largement adoptée car elle demande des modifications aux registres DNS et les administrateurs n'ont pas de motivation pour le faire. Padmanabhan et Subramanian [165] proposent trois approches pour inférer la localisation géographique d'un hôte à partir de son adresse IP : l'inférence à partir du nom DNS, l'inférence à partir d'un groupement des adresse IP, et l'inférence basée sur des mesures. Dans cette thèse, nous nous concentrons sur les problèmes liés à l'inférence basée sur des mesures.

Chapter 6

Location Management in Mobile Ad Hoc Networks

In mobile ad hoc networks, autonomous mobile nodes cooperate with each other to implement network functionalities, such as packet routing. Since nodes move arbitrarily without previous notification, the design of scalable packet routing protocols for mobile ad hoc networks becomes a challenging task [176, 177, 178]. The high rates of change in the topology due to node mobility impose scalability problems on conventional topology-based routing protocols. Under geographic routing [145, 146, 147, 148], each node only needs to know its own position, the position of the destination, and the positions of its one-hop neighbors to take a forwarding decision. The state requirements at each node depend on the density of nodes instead of the total number of destinations. Therefore, geographic routing claims to be nearly "stateless", as it is based only on few local states. The routing decision at each node concerns the selection of the next-hop among the node neighbors in order to get the packet closer to its destination. Commonly, each node determines its own position through the use of the Global Positioning System (GPS) [140] or some other type of positioning system [179, 180]. However, inexpensive and simple handheld devices may be expected to compose an ad hoc network. Such equipments are not likely to have a GPS receiver or any other positioning system by power, cost, and size constraints. A possible solution suggests that these limited nodes rely on a nearby location aware node to act as a location proxy for them [181]. The positions of nearby nodes are typically learned through one-hop broadcasts.

Location management concerns the deployment of a scalable location service able to provide a given source with the position of any eventual destination. Camp *et al.* [182] evaluate proactive and reactive alternatives for geographic routing in mobile ad hoc networks. Under geographic routing, a source must learn the geographic position of any eventual destination to be able to label packets with the destination position. Therefore, geographic routing protocols are dependent on the existence of scalable location services [158, 183] to provide a mechanism for sources to learn the positions of the destinations. Providing a scalable location service in the context of mobile ad hoc networks is a difficult problem because there is no fixed infrastructure. When a source requests the position information of a destination, the source only knows the destination identifier. As a consequence of node mobility, there is no static relationship between a node identifier and the node location. Moreover, the location service itself uses geographic routing. In order to allow the geographic routing of packets toward the destination, the location management establishes a dynamic relationship between the destination identifier and the destination physical location. Previous efforts on scalable location management, like the home-agent based approach [154], SLURP [155], GLS [144], and DLM [157] have been revised in Chapter 5. These proposals for location management in ad hoc networks, however, disregard specific features of TCP flows.

In this chapter, we propose a TCP-tailored approach [184] to the location management in mobile ad hoc networks. The main proposals in location management operate in two sequential and distinct phases: the location query of the position of a destination node and the transmission of a flow toward the destination node. We propose to use the initial packet of the flow to learn the location of the destination instead of adopting dedicated location query packets. Such an approach specially benefits TCP flows since it exploits particular features of the TCP connection mechanism. For TCP flows, our approach improves the performance of the location management whenever a source begins the procedure of learning the location of a destination node. The TCP-tailored approach can be applied to previous works with minor changes on their particular features, thus enhancing their performance for TCP flows. Results show that the proposed approach reduces the transmission costs in location management for the TCP flows with respect to the conventional schemes, specially favoring short-lived TCP flows. Furthermore, a reduction on location management costs also implies a reduction in the battery consumption of nodes.

6.1 TCP-Tailored Approach

Geographic routing selects the next node to forward a packet using the positions of the current node, of the nearby nodes, and of the destination node. The packet path from the source toward its destination is composed by the sequence of nodes that get the packet closer to its destination at each hop. An exception is the presence of void regions. In such a case, the transmission range of the current node does not reach the destination although the current node is the closest node to the destination among all its neighbors. To solve this problem, geographic routing proposals such as GPSR [145] are designed to route packets around these void regions. A packet may be even forwarded, in an intermediate step, to a node farther from the destination than the current one. Despite such a fact, the final goal of geographic routing remains to get the packet closer to the destination as long as the positioning of nodes allows it. Routing around a void region is actually making the packet move toward the destination, even if not at each step.

Let $G = (V, E)$ denote the graph G consisting of the set V of all nodes and the set of edges $E = \{(x, y) \mid x, y \in V \text{ and } \mathcal{D}(x, y) \leq R\}$. The value R is the transmission range of nodes and $\mathcal{D}(x, y)$ is the euclidean distance between the positions of nodes x and y. We define $E_{x,y} \subset E$ as the ordered subset of edges that composes the path between the nodes x and y. For example, in Figure 6.1, $E_{a,d} = \{(a, b), (b, c), (c, d)\}$.

The scenario of adopting geographic routing and a location service to allow a source to learn the position of the requested destination has two distinct phases:

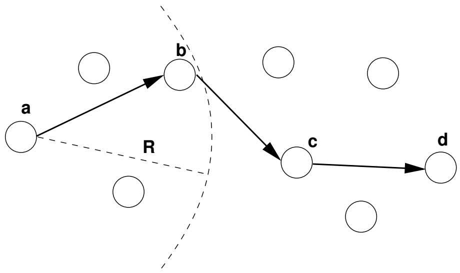


Figure 6.1: Geographic routing in a mobile ad hoc network.

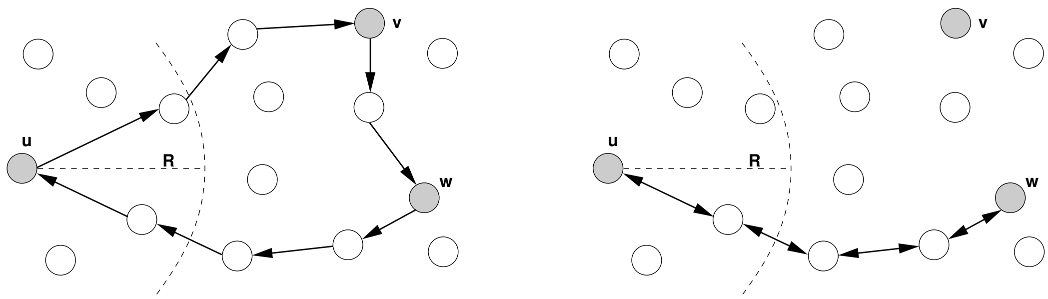
- Location query – a source must query a location service (composed of one or more location servers) for position information about the destination. When the query arrives at a location server, it is forwarded toward the destination to obtain the most possible accurate position information. The query packet path may actually include more than one location server with increasing accuracy about the actual location of the destination. Upon arriving at the destination, the location query induces the destination to update the source with its current location. Figure 6.2(a) presents the path of a location query from the source node u requesting the position of the destination node w to the node v that acts as a location server for node w.

- Flow transmission – after receiving the position information from the destination itself, the source node starts the transmission of the flow toward the destination node. During a transmission, both the source and the destination nodes must update their positions to each other if they move. This update position information is piggybacked onto either data or acknowledgment packets. Regular update messages may also be adopted in the absence of acknowledgments. After the first position discovery, the messages are exchanged directly between the source and the destination nodes without any participation of location servers, as shown in Figure 6.2(b).

The path leading from source node u to destination node w passing through node v is referred to as the *indirect path* in contrast with the *direct path* leading directly from node u to node w (see Figure 6.2). We consider the cost in consumed energy of transmitting one packet over an edge as being dependent on the edge length and on the packet size. We also consider that edges have symmetric costs. The function $\mathcal{S}(p)$ is defined returning the size of packet p. The total cost of sending one packet p from node u to node v is defined as

$$\mathcal{C}(u, v, p) = \mathcal{S}(p) \sum_{(x,y)\in E_{u,v}} \mathcal{D}(x, y). \tag{6.1}$$

A node u wishing to learn about the position of a node w queries a node v that acts as a location server for node w. The total cost of sending a sequence of n flow packets (f) from node u to node w is defined as follows. This cost includes the cost of a query packet (q) going from node u to node w through node v, i.e. taking the *indirect path*, and back to



(a) Location query (b) Flow transmission

Figure 6.2: Two phases in conventional location services for geographic routing.

node u through the *direct path*. The total cost also comprises the transmission of the n flow packets and their respective m ($1 \leq m \leq n$) acknowledgments packets (a) directly through the *direct path* between nodes u and w. In the absence of acknowledgment packets, we consider a to represent the position update messages. Hence, the total cost $\mathcal{C}_1$ of the conventional approaches considering one query packet is

$$\mathcal{C}_1 = \mathcal{C}(u, v, q) + \mathcal{C}(v, w, q) + \mathcal{C}(w, u, q) + n\ \mathcal{C}(u, w, f) + m\ \mathcal{C}(w, u, a). \tag{6.2}$$

We propose an improvement to the process of determining the location information and the sending of the packets of a flow through geographic routing. The initial packets of a flow are sent to query the location servers and then those initial packets are forwarded to their destination instead of using a dedicated query packet to perform the task. In this case, the initial k packets ($1 \leq k \leq n$) of a flow follow the *indirect path* until the source receives the first acknowledgment or update packet from the destination. From this point on, the source can then label the $n - k$ remaining packets with the destination accurate location allowing them to use the *direct path*. This proposal does not eliminate the query for locations, but it instead avoids the usage of a dedicated query packet and the waiting time for a query answer imposed on a source for starting a transmission. The total cost $\mathcal{C}_2$ of the proposed approach is then given by

$$\mathcal{C}_2 = k\ \mathcal{C}(u, v, f) + k\ \mathcal{C}(v, w, f) + (n - k)\ \mathcal{C}(u, w, f) + m\ \mathcal{C}(w, u, a). \tag{6.3}$$

The proposed scheme is only worth adopting when $\mathcal{C}_2 \leq \mathcal{C}_1$. Allowing for the possible different sizes of the different packets, it may not be worth sending a large packet through a long, and thus expensive, *indirect path* in comparison with the sending of a smaller extra query packet. After the initial packets traverse the longer *indirect path* and the source receives the destination location, the remaining packets take the *direct path* between the source and destination nodes. Results from the performance analysis presented in Section 6.2 show that the proposed scheme always reduces the cost if the first packet of a flow to be transmitted has a size equivalent to the size of the dedicated query packet.

Such conditions match the conditions observed in the presence of TCP flows. TCP uses a three-way handshake mechanism to perform the TCP connection establishment [185]. The source node first sends a `SYN` packet, which only contains the TCP header and no payload, and waits for an acknowledgment from the destination before sending an acknowledgment packet to establish the connection.

6.2 Performance Analysis

The assessment of the proposed scheme allows for: the different sizes of the different packets, the size of the flow to be sent, and the relative positioning of the concerned nodes (the source, the destination, and the location server). To analyze the proposed scheme, we consider the minimum bound costs of each approach. Combining the triangle inequality, $\mathcal{D}(u,w) \leq \mathcal{D}(u,v) + \mathcal{D}(v,w)$ for all nodes $u, v, w \in V$ [186], with Equation (6.1), we get the lower bound of the total cost of sending one packet p from node u to node v as

$$\begin{aligned} \mathcal{C}_{\min}(u,v,p) &= \min_{u,v \,\in\, V} \left(\mathcal{S}(p) \sum_{(x,y) \in E_{u,v}} \mathcal{D}(x,y) \right) \\ &= \mathcal{S}(p)\ \mathcal{D}(u,v). \end{aligned} \tag{6.4}$$

Applying the minimum bound costs as stated in Equation (6.4) to Equations (6.2) and (6.3), we get the lower bound costs of $\mathcal{C}_1$ and $\mathcal{C}_2$, respectively. It is expected that a geographic routing protocol would make its best to achieve a cost as close as possible to the considered lower bounds. This happens since a packet is forwarded by each intermediate node to the reachable node that makes the packet get closer to the destination. The lower bound costs are achieved in the case all intermediate nodes are perfectly aligned with the source and destination nodes.

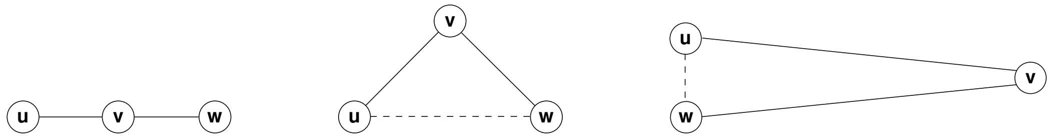
6.2.1 Dispersion of Nodes and Cost Improvement

The degree of dispersion between the *direct path* leading from the source node u to the destination node w and the *indirect path* containing node v, which stores the location information of node w, is defined as

$$\Delta = \frac{\mathcal{D}(u,w)}{\mathcal{D}(u,v) + \mathcal{D}(v,w)}, \qquad 0 < \Delta \leq 1. \tag{6.5}$$

Figure 6.3 illustrates the possible values of Δ as a function of the relative positioning of nodes u, v, and w. Values of Δ close to one indicate a small degree of dispersion between the *indirect* and the *direct paths*. On the other hand, small values of Δ represent a high degree of dispersion.

We also define the cost improvement Ψ provided by the proposed scheme in comparison with the cost of the conventional approach as



(a) $\Delta = 1$ (b) $\Delta < 1$ (c) $\Delta \to 0$

Figure 6.3: The relative positioning of nodes and the values of Δ.

$$\Psi = 1 - \frac{\mathcal{C}_2}{\mathcal{C}_1}. \tag{6.6}$$

For the lower bound costs of $\mathcal{C}_1$ and $\mathcal{C}_2$, the cost improvement Ψ is given by

$$\Psi = \frac{\mathcal{S}(q)\ (1+\Delta) - k\ \mathcal{S}(f)\ (1-\Delta)}{\mathcal{S}(q)\ (1+\Delta) + n\ \mathcal{S}(f)\ \Delta + m\ \mathcal{S}(a)\ \Delta}. \tag{6.7}$$

6.2.2 Effects of the Relative Size of Packets

We evaluate the achieved cost improvement of the proposed scheme considering the relative sizes of query and flow packets as well as different values of Δ. Figure 6.4 shows the cost improvement Ψ for different ratios between the sizes of query and flow packets. The size of the query packet is 40 bytes. We consider query and acknowledgment packets as having the same size. We also consider in this analysis a data flow consisting of one flow packet, i.e. $n = k = m = 1$. Figure 6.4 shows the cost improvement Ψ when the proposed scheme is worth using ($\mathcal{C}_2 \leq \mathcal{C}_1$).

When the initial packet of the flow has the same size of the query packet, the proposed scheme is always worth using. As the flow packet gets larger than the query packet, sending it through a longer *indirected path* becomes less advantageous. Depending on the relative positioning of the concerned nodes (Δ lower than a certain value), sending an initial packet larger than a query packet through the longer *indirect path* is not interesting. Therefore, the proposed scheme specially benefits flows with small initial packets, i.e. packets with size equivalent to the size of a query packet. TCP flows have such a characteristic because of the initial `SYN` packet sent by the connection establishment mechanism. When $\Delta = 1$, the cost improvement Ψ for each scenario reaches its upper bound value given by

$$\Psi_{\max} = \frac{2\ \mathcal{S}(q)}{2\ \mathcal{S}(q) + n\ \mathcal{S}(f) + m\ \mathcal{S}(a)}. \tag{6.8}$$

Given that the upper bound $\Psi_{\max}$ is a function of the relative sizes of the concerned packets, each case presents its own upper bound $\Psi_{\max}$.

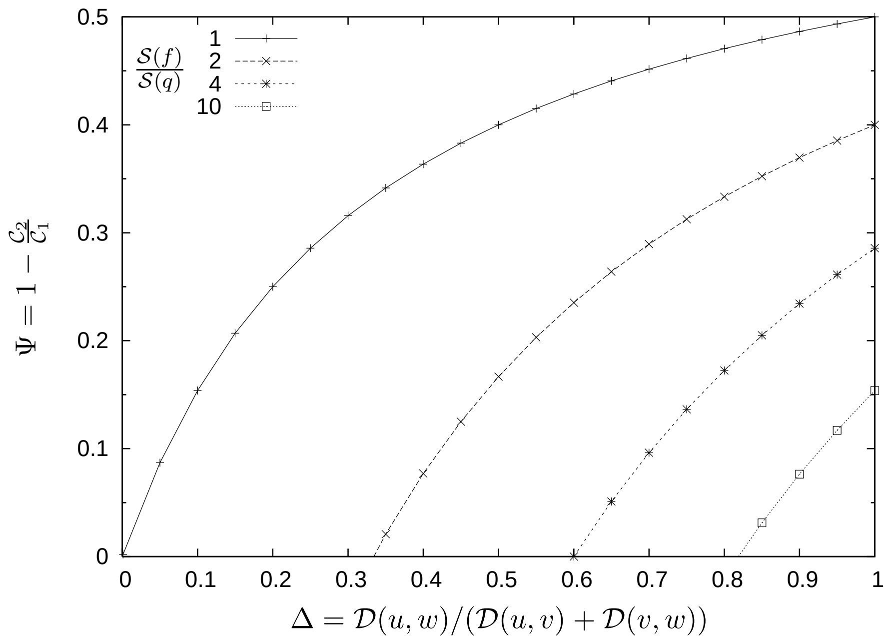
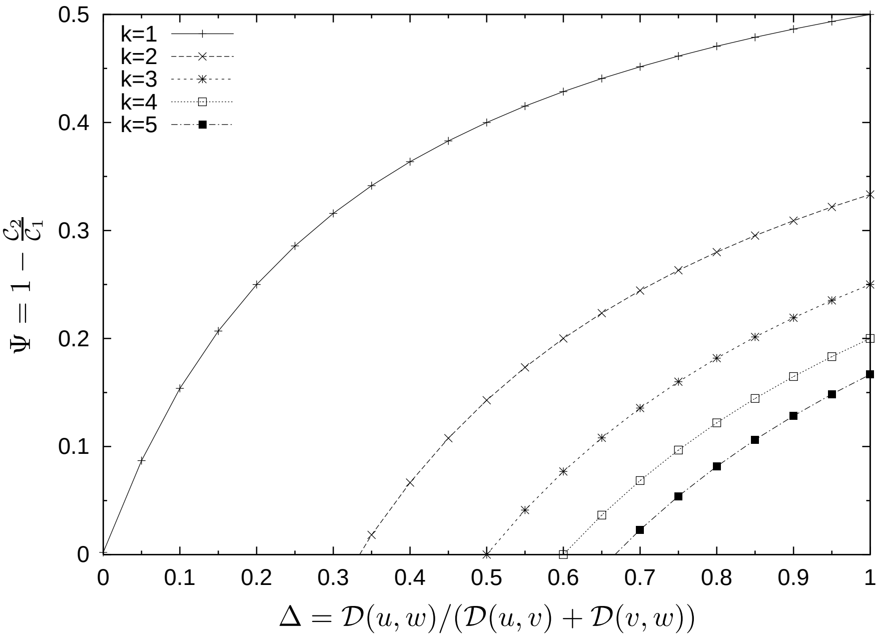


Figure 6.4: Effects of the size of packets and the relative positioning of nodes.

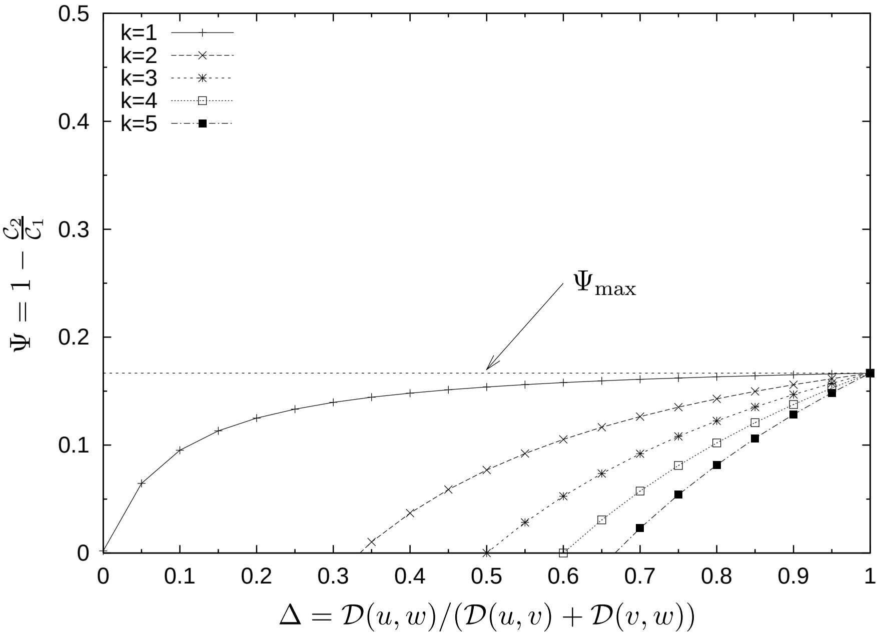
6.2.3 Effects of the Number of Initial Packets Sent

We now evaluate the case where more than one initial flow packet ($1 \leq k \leq n$) from the data flow is sent through the *indirect path*. Two options apply: either all flow packets ($n = k$) or just the k first packets out of the n packets composing the flow are sent through the *indirect path*. Figure 6.5 presents the results for both cases with flow and query packets having the same size ($\mathcal{S}(f)/\mathcal{S}(q) = 1$). We consider one acknowledgment packet for each flow packet, i.e. $m = n$. Figure 6.5(a) shows the results for $n = k = 1, \ldots, 5$, meaning that all packets from the flow of n packets are sent through the *indirect path*. We observe that the shorter the flows are, the greater are the benefits, indicating the proposed scheme favors short-lived flows.

If we consider a flow containing $n = 5$ packets that sends its $k = 1, \ldots, 5$ initial packets through the *indirect path*, we get the results presented in Figure 6.5(b). We observe that is better to adopt the first packet of a flow to learn the position of the destination and send the remaining packets through the *direct path* afterward, as shown for $k = 1$. In Figure 6.5(b), all cases converge to the same value of the upper bound of the cost improvement given by $\Psi_{\max}$ because the upper bound is independent of the value of k (see Equation (6.8)). It should be noted that the previous presented results are completely consistent with the upper bound $\Psi_{\max}$. Since the upper bound is a function of the quantity n of packets composing the flow and of the size of the concerned packets, each analyzed scenario has its own upper bound $\Psi_{\max}$.



(a) $n = k$



(b) $n = 5$

Figure 6.5: Effects of the number of initial packets sent through the *indirect path*.

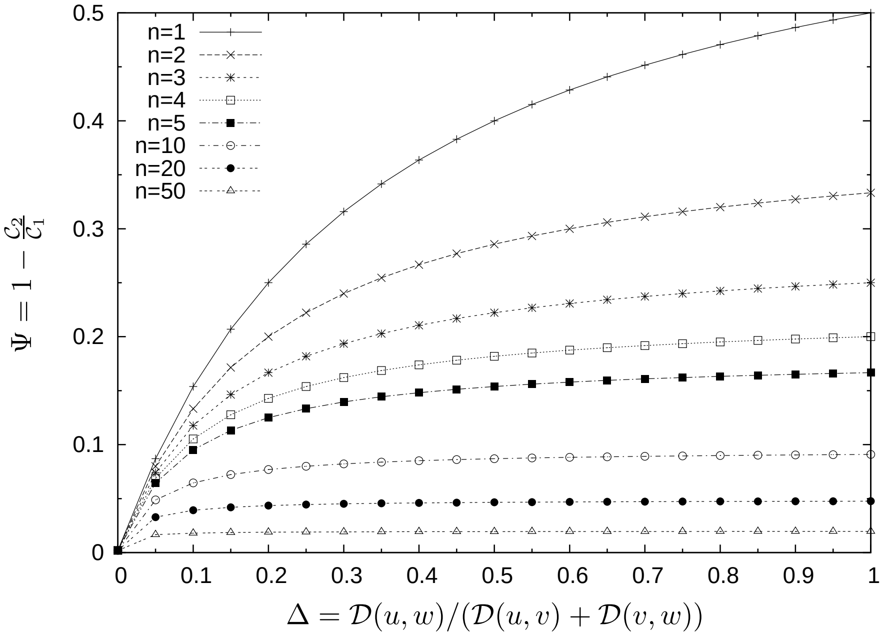


Figure 6.6: Effects of the flow size and relative positioning of nodes.

6.2.4 Effects of the Size of the Flow

As the size of the data flow to be sent increases, the cost of discovering the destination location is diminished when facing the whole cost of the flow transmission. Figure 6.6 presents the cost improvement Ψ for different flow sizes as a function of different values of the dispersion Δ due to the relative positioning of the concerned nodes. The flow and query packets have the same size ($\mathcal{S}(f)/\mathcal{S}(q) = 1$). In the cost of the proposed scheme, we consider the sending of one initial packet ($k = 1$) out of the n packets composing the flow through the *indirect path*. In the case of a TCP flow, $k = 1$ means the sending of the first `SYN` packet through the *indirect path* and all the remaining packets through the *direct path*.

Short-lived flows with ten packets or less ($n \leq 10$) present the most significant cost improvement. This cost reduction obtained by the short-lived flows ranges from 5% to 50% for dispersion of nodes $\Delta \geq 0.05$ depending on the size of the short-lived flow. Note that if $\mathcal{D}(u,w) \leq R$, the source node u and the destination node w are neighbors and thus no location query toward any remote location server is needed. From the triangle inequality and considering $\mathcal{D}(u,w) > R$, we observe that the dispersion Δ of the concerned nodes is actually bounded by $\frac{R}{2\ \mathcal{D}(u,v)+R} < \Delta \leq 1$, if a location service is about to be used. Thus, depending on the size of the short-lived flow, a cost improvement of 5% to 50% is achieved if $\mathcal{D}(u,v) \leq 9.5\ R$. This condition assures the dispersion Δ of the concerned nodes to be $0.05 \leq \Delta \leq 1$.

6.2.5 Effects on TCP and UDP Flows

In the case of a TCP flow, a source node first sends a `SYN` packet because of the three-way handshake mechanism that performs the TCP connection establishment [185]. We propose to exploit this feature by sending the first `SYN` packet to learn the location of the destination through the *indirect path* and the remaining packets through the *direct path*. This procedure always enhances the performance of the location management as shown in Figure 6.5 for $k = 1$. As shown in Figures 6.4 and 6.5, this proposed TCP-tailored scheme reduces the cost of a TCP transmission because the first packet of the TCP flow (the `SYN` packet) has the same size of the query packet and is the only packet sent through the *indirect path*.

Short-lived TCP flows, as most web requests for instance, represent a large portion (up to 80% [107, 187]) of the traffic in the current Internet. We consider short-lived TCP flows as having five to ten packets. The minimum of five packets takes into account the minimum number of packets sent by the source to have a complete TCP connection: two packets for connection establishment, one data packet, and two packets for connection termination. Supposing a similar traffic profile between the Internet and future mobile ad hoc networks, a cost reduction of up to 16% is achieved for the most part of flows (see Figure 6.6). The minority of flows, either flows of larger size ($n > 10$) or short-lived flows being transmitted with a dispersion $\Delta < 0.05$, has also a cost reduction, but limited to 5%.

UDP flows may not take advantage of the proposed scheme. The sending of a first flow packet larger than a query packet through the *indirect path* may become more costly than the adoption of a small dedicated query packet. In this case, the adoption of the proposed scheme may be advantageous, or not, depending on the relative size of the concerned packets and the relative positioning of the concerned nodes (see Figure 6.4). Based on the transport protocol adopted by the flow to be transmitted, a source may perform the location management using either the TCP-tailored approach for the TCP flows or the conventional approach for the remaining flows. Such a conservative policy assures to a location management scheme added with the TCP-tailored approach at least the performance of the same scheme with no addition for non-TCP flows and an enhanced performance for TCP flows.

6.2.6 Estimating the Dispersion of Nodes and the Cost Improvement

From Equations (6.2) and (6.3), the TCP-tailored approach is worth using ($\mathcal{C}_2 \leq \mathcal{C}_1$) if the positioning of the concerned nodes satisfies the condition

$$\Delta \geq \frac{k\ \mathcal{S}(f) - \mathcal{S}(q)}{k\ \mathcal{S}(f) + \mathcal{S}(q)}. \tag{6.9}$$

This condition is consistent with the results presented in Figures 6.4 and 6.5. From Equation (6.9), the source knows the minimum value of Δ that assures the proposed scheme as being worth to adopt considering some characteristics of the flow to be transmitted. Nevertheless, it should be noted that Δ measures the dispersion of the relative positioning between the three nodes u, v, and w. Clearly, the position of the destination node w is unknown as this is exactly why the source node u has to make use of the location service. Therefore, there are no means to *a priori* compute the exact value of the dispersion Δ.

The greater the value of the dispersion Δ of the concerned nodes, the greater the cost improvement Ψ is. For a given flow, Equation (6.7) associates a cost improvement $\Psi = \psi$ with a dispersion $\Delta = \delta$. Therefore, if we estimate the probability that the dispersion Δ is greater than δ, we estimate the probability that the cost improvement Ψ due to the TCP-tailored approach is greater than ψ. At the moment of querying a location service, the source node u only knows the distance $\mathcal{D}(u, v)$ to the location server v of node w and some characteristics of the flow. Thus, our estimation of the dispersion Δ must be carried out based on this limited knowledge.

We propose an alternative method to estimate the probability of having a value of dispersion Δ greater than a given value δ. As a consequence, we estimate the probability of having a cost improvement Ψ due to the TCP-tailored approach greater than a given value ψ. Suppose node u located at the position $(0, 0)$ of a coordinate system covering the deployment region of the ad hoc network. The node u has to contact the location server v to obtain the location information about node w. Suppose now node v located at position $(1, 0)$ on the same coordinate system of node u. Node u wishes to know the location of node w that may be located anywhere else in the deployment region, but not in the cover range R around node u. By definition, if the location service is about to be used, $\mathcal{D}(u, w) > R$. Figure 6.7 depicts the dispersion Δ between the positioning of the related nodes if the position of node w on the coordinate system is (x, y). Without loss of generality, we adopt the values on the x and y-axis as units of $\mathcal{D}(u, v)$, which is the known information at node u. The closer node w is to node u in comparison with the distance $\mathcal{D}(u, v)$, the lower is the value of the dispersion Δ (see also Figure 6.3(c)).

The projected areas in the xy plane (Figure 6.7) are the contour plots of the dispersion Δ, for $\Delta = 0.4, 0.6, 0.8$. The contour plots represent the cross sections of the surface with the planes having such values of Δ. If node w is located inside the area $A_{\Delta=\delta}$, then the relative positioning of nodes u, v, and w has a dispersion Δ bounded by δ. Note that $A_{\Delta=i} \supseteq A_{\Delta=j}$, for $i \geq j$. If $\Delta > \delta$, then knowing some characteristics of the flow to be transmitted and using Equation (6.7) allows to evaluate that the TCP-tailored approach has a cost improvement $\Psi > \psi$. For example, the area $A_{\Delta=0.8}$ indicated in Figure 6.7 represents the contour for $\Delta = 0.8$. Therefore, if node w is located inside the area $A_{\Delta=0.8}$, then $\Delta \leq 0.8$. Similarly, if node w is located outside $A_{\Delta=0.8}$, anywhere else in the deployment region, then $\Delta > 0.8$. Considering the scenario of Figure 6.6 and applying Equation (6.7), if $\Delta > 0.8$ and $n = 10$, then the cost improvement is $\Psi > 9\%$.

We compute the probability that $\Delta > \delta$ and this probability is the same that the cost improvement Ψ is greater than a value ψ obtained applying Equation (6.7) accordingly. The computation of such a probability is performed as follows. Let $\mathcal{A}$ denote the area of the complete deployment region and A_R denote the circular area around node u with range R. We also denote by $|A|$ the surface area of the region delimited by A. Again, if $\mathcal{D}(u, w) < R$, node w would be a nearby node with respect to node u and no location query would be needed. If a location query is being considered, node w is certainly outside the region around node u whose surface area is $|A_R| = \pi R^2$. The probability that $\Delta > \delta$ is the same that node w is located outside the area $A_{\Delta=\delta}$. Considering $(A_{\Delta=\delta} \cup A_R) \subset \mathcal{A}$, the probability that the cost improvement Ψ is greater than a given value ψ, equivalent to a dispersion $\Delta = \delta$ using Equation (6.7), is

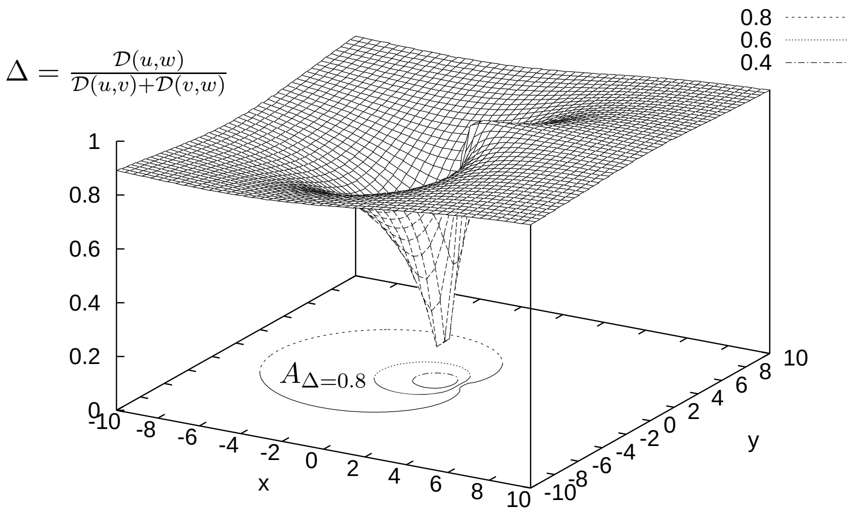


Figure 6.7: Dispersion Δ for different positions of node w to be located.

$$P(\Psi > \psi) = P(\Delta > \delta) = 1 - \frac{|A_{\Delta=\delta} - A_R|}{|\mathcal{A}|}. \tag{6.10}$$

The computation of the probability presented in Equation (6.10) considers the deployment region $\mathcal{A}$ known to all nodes. This would be the case if nodes are confined within a well known region. If individual nodes are not aware of the dimensions of the deployment region they are located in, we propose an adaptive mechanism to estimate the size of the deployment region $\mathcal{A}$. In this case, at the beginning, nodes have no knowledge of the total size of the deployment region. As nodes contact other remote nodes, they learn the distance from themselves to these other nodes. For example, to carry out a transmission, the source node u learns the distance $\mathcal{D}(u, w)$ to the destination node w and the distance $\mathcal{D}(u, v)$ to the location server v of node w. If nodes keep the maximum distance $d_{\max}$ they have used, $d_{\max}$ can be used to perform an estimation $\widehat{\mathcal{A}}$ of the unknown deployment region $\mathcal{A}$. The estimated size of the deployment region is then $|\widehat{\mathcal{A}}| = \pi d_{\max}^2$. Considering no initial knowledge of the total deployment region, the probability that the cost improvement Ψ is greater than a given value ψ, equivalent to a dispersion $\Delta = \delta$ using Equation (6.7), is

$$P(\Psi > \psi) = P(\Delta > \delta) = 1 - \frac{|A_{\Delta=\delta} - A_R|}{|A_{\Delta=\delta} \cup A_R \cup \widehat{\mathcal{A}}|}. \tag{6.11}$$

At the very first transmission of node u that requires a location query, the only knowledge node u has of the remaining ad hoc network is the distance $\mathcal{D}(u, v)$ to the location server v of node w to be located. Therefore, $d_{\max}$ is initialized to $\mathcal{D}(u, v)$. The estimated size of the deployment region is thus $|\widehat{\mathcal{A}}| = \pi d^2(u,v)$ in this first step. As node u communicates with other farther nodes, the knowledge of the deployment region increases since $d_{\max}$ is updated with information of the farthest reachable known node. Alternatively, a node could also extract indications about how large the deployment region is from the packets the node forwards. Packets under geographic routing carry the position of their sources and destinations. As a node forwards a packet toward a destination, such a node can use the position of the source and the destination present on the packet to faster improve its estimation of the size of the deployment region the node is located in.

6.3 Conclusion

We have proposed a TCP-tailored approach to the location management in mobile ad hoc networks. The proposed approach sends the first packet of a TCP flow (inherently a `SYN` packet) to learn the position of the flow's destination instead of adopting dedicated location query packets. This approach is shown to be specially advantageous to the TCP flows reducing their transmission costs in the location management procedure and thereby reducing the energy consumption of nodes. The achieved cost reduction is a function of the spatial dispersion of the concerned nodes, the size of packets, and the size of flows. The TCP-tailored approach gives a better advantage to short-lived TCP flows (up to 10 packets) providing higher cost reductions for the most part of flows. The short-lived TCP flows represent about 80% of the flows observed in the current Internet [107, 187]. We consider the number of flows because the needed query of a location service only concerns the initialization of a flow transmission. Supposing a similar traffic profile between the Internet and the future mobile ad hoc networks, the TCP-tailored approach benefits the majority of flows in the location management procedure.

The TCP-tailored approach is applicable to the main previous proposals for location management in mobile ad hoc networks. Applying the proposed TCP-tailored approach, a source sends the first packet of TCP flows to learn the destination's position using the particular features of each previous proposal (either home zone or hierarchical approaches for instance). Therefore, the main previous proposals reduce the transmission costs in the location management procedure for the TCP flows with minor changes, independent of the particular features of such previous proposals. Based on the transport protocol adopted by the flows to be transmitted, a source selects to apply either the TCP-tailored approach (for the majority of flows composed by the TCP flows) or the conventional approach (for the remaining flows). As each cooperative node in the ad hoc network seeks to adopt the method having the lowest cost at each transmission, the adoption of the TCP-tailored approach provides a performance improvement of the location management as well as a reduction in the battery consumption of nodes.

6.4 Résumé du Chapitre

Dans les réseaux ad hoc mobiles, des nœuds indépendants coopèrent les uns avec les autres pour assurer des fonctionnalités du réseau, tel que le routage de paquets. Étant donné que les nœuds peuvent se déplacer indépendamment en présentant plusieurs profils de mouvement, la conception de protocoles de routage résistants au facteur d'échelle devient une tâche complexe à élaborer [176]. Des changements fréquents de la topologie, conséquence de la mobilité des nœuds, implique de sérieux problèmes de passage à l'échelle pour les protocoles de routage traditionnels basés sur la topologie. Le routage géographique [145] traite le problème de résistance au facteur d'échelle. Chaque nœud n'a besoin de connaître que la localisation de ses voisins et du nœud destinataire pour prendre une décision d'acheminement d'un paquet. La décision de routage à chaque nœud correspond au choix du prochain saut parmi les voisins du nœud de façon à rapprocher le paquet du destinataire final. Chaque nœud peut déterminer sa propre position par GPS ou par n'importe quel autre service de positionnement [179]. Les positions des nœuds voisins sont découvertes par la diffusion de messages autour du nœud.

En utilisant le routage géographique, une source a besoin d'avoir un moyen d'apprendre la position géographique d'une destination pour être capable d'étiqueter les paquets avec la position de cette destination. Les protocoles de routage géographique pour les réseaux ad hoc mobiles sont alors dépendants de l'existence de services de localisation capables de fournir aux sources les positions des destinataires [158]. Fournir un service de localisation efficace dans le contexte de réseaux ad hoc mobiles constitue un problème difficile vue l'absence d'infrastructure. Lorsqu'un nœud source demande la position d'un nœud destinataire, il a uniquement connaissance de l'identifiant du nœud destinataire. Le problème consiste à établir un rapport dynamique entre l'identifiant d'un nœud destinataire et sa localisation pour permettre le routage géographique de paquets vers lui.

Dans ce chapitre, nous proposons une approche TCP pour la gestion de la localisation dans les réseaux ad hoc mobiles. Nous proposons une approche qui bénéficie en particulier aux flux TCP lorsqu'une source démarre la procédure de découverte de la localisation du nœud destinataire. Les efforts précédents en gestion de la localisation pour les réseaux ad hoc mobiles opèrent en deux étapes successives et distinctes : la recherche de la localisation et la transmission d'un flux. Nous proposons d'utiliser les premiers paquets d'un flux à transmettre afin de rechercher la localisation de la destination au lieu d'utiliser des paquets de recherche dédiés. On montre que l'approche proposée réduit les coûts pour les flux TCP par rapport à l'approche traditionnelle.

Chapter 7

Geographic Location of Internet Hosts

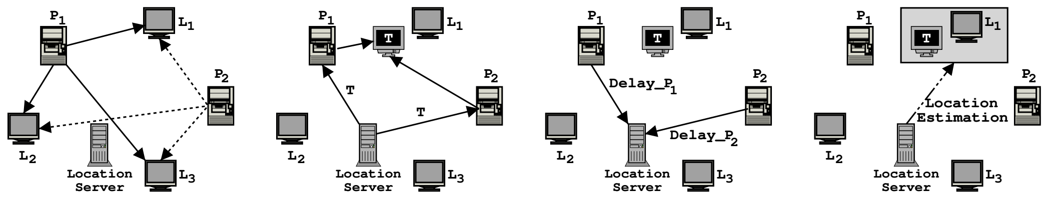
Location-aware applications take into account where the users are accessing from and thereby can offer novel functionalities in the Internet. In this chapter, we focus on improving the accuracy of the geographic location estimation of Internet hosts inferred from delay measurements. Host locations are inferred by comparing delay patterns of geographically distributed landmarks, which are hosts with a known geographic location, with the delay pattern of the target host to be located. Probe machines measure the delay patterns toward landmarks and the target host. We investigate the correlation between geographic distance and network delay. This correlation is weak to moderate if considered worldwide, whereas we show it is stronger in regions with richer connectivity. Moreover, we identify two key points that influence the accuracy of the Internet host location from delay measurements. The accuracy basically depends on the placement of landmarks and probe machines as well as on how efficiently the similarity between delay patterns is evaluated. Therefore, we aim at improving the accuracy of the host location estimation by: (i) strategically placing landmarks and probe machines; and (ii) selecting models to best measure the similarity between the delay pattern of each landmark and the one of the target host.

Landmarks are expected to reflect where most users and hosts are. The number and position of the landmarks are key points for the accuracy of the location estimation and for the impact on network load due to measurements. The problem of finding the best location to place Internet resources, to reduce both the traffic load on networks and the delay perceived by end users, was addressed by Al-Fawzan *et al.* [188] and Giladi *et al.* [189] for network servers, Krishnan *et al.* [190] for caches, Qiu *et al.* [191], Radoslavov *et al.* [192], Cronin *et al.* [193], and Bassali *et al.* [194] for mirrors, and Shah *et al.* [195] for filters in content-based multicast. We use, in this chapter, similar concepts to find the best placement of landmarks and probe machines to provide accurate geographic location estimations for the majority of Internet hosts. We propose and evaluate a demographic placement approach [196, 197] that considers the geographic distribution of users, and consequently of hosts to be located, to place landmarks and probe machines. Our results show that the demographic placement provides a relatively small number of landmarks able to represent a large portion of users (hosts) within a limited coverage distance. Adopting fewer landmarks to locate a host implies a lower amount of measurement traffic. We also verify that for a limited number of landmarks, the demographic placement improves the representativeness of each landmark. This improvement results in closer landmarks and more accurate location estimations for the most part of hosts to be located. We also apply our demographic approach to place the probe machines. Probe machines are placed on sites likely to have enough network infrastructure to make their deployment feasible and in a fashion to avoid gathering redundant measurement data. Another key issue concerning the location estimation accuracy is the evaluation of how similar a landmark and the target host delay patterns are. Measuring the similarity between different items is fundamental for accuracy of recommender systems [198] and pattern analysis [199]. In our case, a similarity model compares the delay patterns and determines the landmark with the most similar delay pattern with respect to the one of the target host. Hence, we investigate different similarity models and evaluate their accuracy for the measurement-based geographic location of an Internet host. Padmanabhan and Subramanian [165] adopt the Euclidean distance as a way to assess the similarity of the observed delay patterns. One of our similarity models considers the previously adopted Euclidean distance and we use this distance as a reference for evaluation. Our results [200] show that Euclidean distance is outperformed by other similarity models that provide more accurate location estimations of the target host.

7.1 Inferring Host Geographic Locations from Delay Measurements

GeoPing [165] infers a host geographic location from delay measurements. In general, a moderate correlation between distance and delay prevents the capture of this relationship under a mathematical model. Therefore, GeoPing adopts an empirical approach based on the observation that hosts sharing similar delays to other fixed hosts tend to be near each other geographically.

We formalize the problem of inferring a host location from delay measurements as follows. Consider a set $\mathcal{L} = \{L_1, L_2, \ldots, L_K\}$ of K landmarks. Landmarks are reference hosts with a well known geographic location. Consider a set $\mathcal{P} = \{P_1, P_2, \ldots, P_N\}$ of N probe machines. Fig. 7.1 illustrates the steps in inferring a host location from delay measurements, which are detailed along this section. The probe machines periodically determine the network delay, which is actually the minimum delay of several measurements, to each landmark (Fig. 7.1(a)). Therefore, each probe machine P_x, $1 \leq x \leq N$, keeps a delay vector $\mathbf{d}_x = (d_{1x}, d_{2x}, \ldots, d_{Kx})$, where d_{ix} is the delay between the probe machine P_x and the landmark $L_i \in \mathcal{L}$. Suppose one wants to determine the location of a given target host T. A location server that knows the landmark set $\mathcal{L}$ and the probe machine set $\mathcal{P}$ is then contacted. The location server asks the N probe machines to measure the delay to host T (Fig. 7.1(b)). Each probe machine P_x, $1 \leq x \leq N$, returns a delay vector $\mathbf{d}'_x = (d_{1x}, d_{2x}, \ldots, d_{Kx}, d_{Tx})$, i.e. the delay vector $\mathbf{d}_x$ plus the just measured delay to host T (Fig. 7.1(c)). After receiving the delay vectors from the N probe machines, the location server is able to construct the delay matrix $\mathbf{D}$ with dimensions $(K + 1) \times N$:



(a) (b) (c) (d)

Figure 7.1: Inferring a host location from delay measurements.

$$\mathbf{D} = \begin{bmatrix} d_{11} & d_{12} & \dots & d_{1N} \\ d_{21} & d_{22} & \dots & d_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ d_{K1} & d_{K2} & \dots & d_{KN} \\ d_{T1} & d_{T2} & \dots & d_{TN} \end{bmatrix} \tag{7.1}$$

The delay vectors gathered by the demanding location server from the probe machines correspond to the columns of the delay matrix $\mathbf{D}$. The location server then compares the lines of the delay matrix $\mathbf{D}$ to estimate the location of host T. To infer the location of host T, the landmark L presenting the most similar delay pattern with respect to the delay pattern of host T is determined. The corresponding location of the landmark L is the location estimation of host T (Fig. 7.1(d)). The delay matrix $\mathbf{D}$ combined with the knowledge of the location of the landmarks of the set $\mathcal{L}$ compose a delay map recording the relationship between network delay and geographic location.

The placement problem we deal with involves the number of needed landmarks and where to place these landmarks and the probe machines. We are thus interested in where to place a finite number of landmarks to maximize the representativeness of each placed landmark. Fewer landmarks imply a lower amount of measurement traffic injected in the network. The amount of measurements may be evaluated as follows. Let Δ denote the time interval adopted by the probe machines to periodically gather the delay from the landmarks of the set $\mathcal{L}$. The total number of measurements $\mathcal{M}$ to estimate the location of h hosts in a time interval τ is

$$\mathcal{M}(h,\tau) = 2N\left(\left\lceil \frac{\tau}{\Delta} \right\rceil K + h\right). \tag{7.2}$$

It should be noted that each measurement may consist of one to several delay samples, but only the minimum value is considered to not take into account delays due to network congestion. In the case we send p `ping` packets to estimate the minimum RTT between a probe machine and a landmark, the amount of measurement traffic injected in the network is actually given by $p\mathcal{M}$.

One alternative to periodically gathering the delay seen by the set of probe machines to the landmarks is to have representative delay matrixes. A representative delay matrix might characterize the delay patterns between the set of probe machines to the set of landmarks for a specific weekday or period of the day. For example, a given delay matrix representing the delay patterns for the daytime period of a normal work day is likely to be different of a delay matrix representing the delay patterns for the nighttime period of a weekend day. Thus, the location estimation of a target host may be performed using a set of representative delay matrixes, avoiding a periodical gather of delay measurements. Hybrid solutions may also be envisaged using periodical measurements to update the adopted representative delay matrixes.

The inference from delay measurements relies on a possible correlation between geographic distance and network delay. We study this correlation in further detail in Section 7.2. We also indicate how the technique of host location inference from delay measurements handles factors that may weaken the correlation between distance and delay. Since the accuracy basically depends on the placement of the concerned components and on how efficiently the similarity between delay patterns is evaluated, these are the main issues we address in this chapter to improve the location estimation accuracy. We propose and evaluate a demographic placement approach to deal with the placement of a limited number of landmarks and probe machines in Section 7.3. We also investigate different similarity models in Section 7.4. These models measure the similarity between delay patterns and determine the landmark to be used as a location estimation for the target host.

7.2 Correlation between Geographic Distance and Network Delay

In order to study the correlation between geographic distance and network delay, we adopt two data sets:

- LibWeb – delay measurements performed from the LIP6 laboratory located in Paris, France, to 135 target hosts with well-known locations all over the world in June 2002. The set of target hosts is mainly composed of university sites extracted from library web (LibWeb) servers around the world [201]. The geographic distribution of the target hosts is as follows: 56 in North America, 44 in Western Europe, 11 in Asia, seven in Eastern Europe, 7 in Latin America, 4 in the Middle East, 3 in Africa, and 3 in Oceania. From the 135 original target hosts, 109 hosts have answered the `ping` requests. The considered delay is the minimum of several measurements to not take into account delays due to congestion in intermediate routers.

- RIPE – data collected from the Test Traffic Measurements (TTM) project of the RIPE network [202]. The data set we consider is composed by the 2.5 percentile of the delay observed from each RIPE host to each other host in the set during a period of 10 weeks from early December 2002 until February 2003. All 55 hosts on the RIPE network are equipped with a GPS card, thus allowing their exact geographic position to be known. The hosts in the RIPE network are geographically distributed as follows: 42 in Western Europe, 5 in the US, 3 in Eastern Europe, 2 in the Middle East, 2 in Oceania, and 1 in Asia.

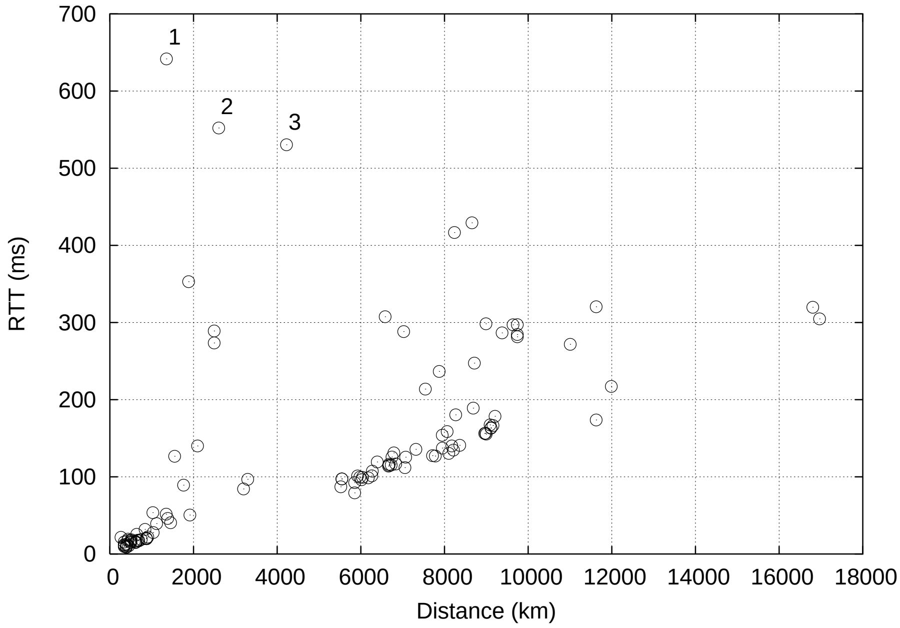


Figure 7.2: Scatter plot of geographic distance and network delay (LibWeb).

7.2.1 Experimental Results for the LibWeb Data Set

Fig. 7.2 presents the scatter plot of the geographic distance and the minimum delay between our probe machine and each target host for the LibWeb data set. A weak correlation is observed between geographic distance and network delay worldwide, resulting in a coefficient of correlation $R = 0.2971$. Some points are significantly away from the others. For about the same geographic distance, the observed delay may be greater by one order of magnitude. For example, points 1, 2, and 3 correspond to hosts located in Algeria, Turkey, and Iran, respectively. The `ping` packets from Paris to Algeria (point 1), for instance, actually make their way through routers located in New York. Geographic properties of Internet routing are studied in further detail in [203]. Routes toward some locations are actually far from the straightforward physical direction. This inconvenient is strongly reinforced by the poor connectivity in certain regions. Poor connectivity weakens the correlation between geographic distance and network delay. From the set of 109 answering hosts, the 80 hosts located in North America and Western Europe have been identified. These regions have the richest connectivity linking their hosts. Fig. 7.3 shows the correlation between geographic distance and network delay considering only the hosts located in North America and Western Europe. The regions that have a richer connectivity have a much stronger correlation between geographic distance and network delay. Even if there are a few isolated points away from the other points, two main clusters are observed in Fig. 7.3. These clusters correspond to hosts located in Western Europe and North America. Excluding the points that clearly remain out of the pack for both clusters (one outlier for the Western Europe cluster and two

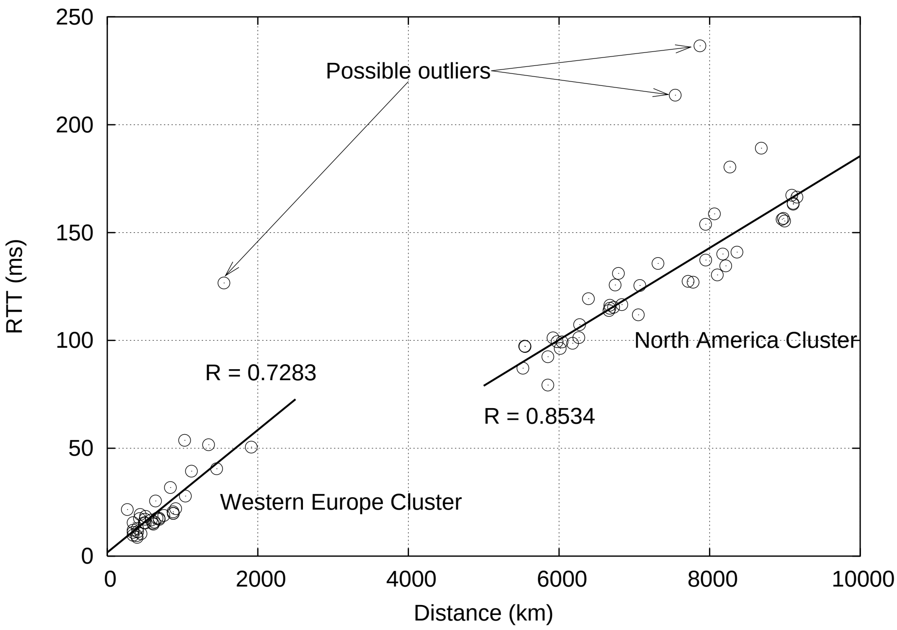


Figure 7.3: Correlation between distance and delay (richer connectivity in LibWeb).

outliers for the North America cluster) results in the following coefficients of correlation: $R = 0.7283$ for the Western Europe cluster and $R = 0.8534$ for the North America cluster.

7.2.2 Experimental Results for the RIPE Data Set

In contrast with the LibWeb data set used in the experiment of Section 7.2.1, where one probe machine gathers delay measurements from 135 landmarks, the RIPE data set gives us the possibility of considering multiple measurement points. As we have the measured delay from each single host to any other host in the RIPE network, any host in the set may be viewed as a probe machine, as a landmark, or both. The delays observed between each pair of RIPE hosts are asymmetric, presumably due to asymmetries in routing and in concurrent traffic load. As a consequence, we consider individually the viewpoint of each host toward the other hosts in the set. The correlation between geographic distance and network delay within the whole RIPE data set is shown in Fig. 7.4. For the whole RIPE data set, there is a moderate correlation between distance and delay ($R = 0.6272$). Nevertheless, we have observed that one single host in the RIPE network significantly contributes to weaken such a correlation. Fig. 7.5 presents the correlation for the RIPE data set without the outlier host. A significantly stronger correlation between distance and delay ($R = 0.8983$) is observed disregarding the outlier host. Performing a `traceroute` toward this outlier host, one single link on the path strongly contributes to increase the end-to-end delay. In accordance with the experiments presented in Section 7.2.1, we observe once more that a richer connectivity strengths the correlation between geographic distance and network delay.

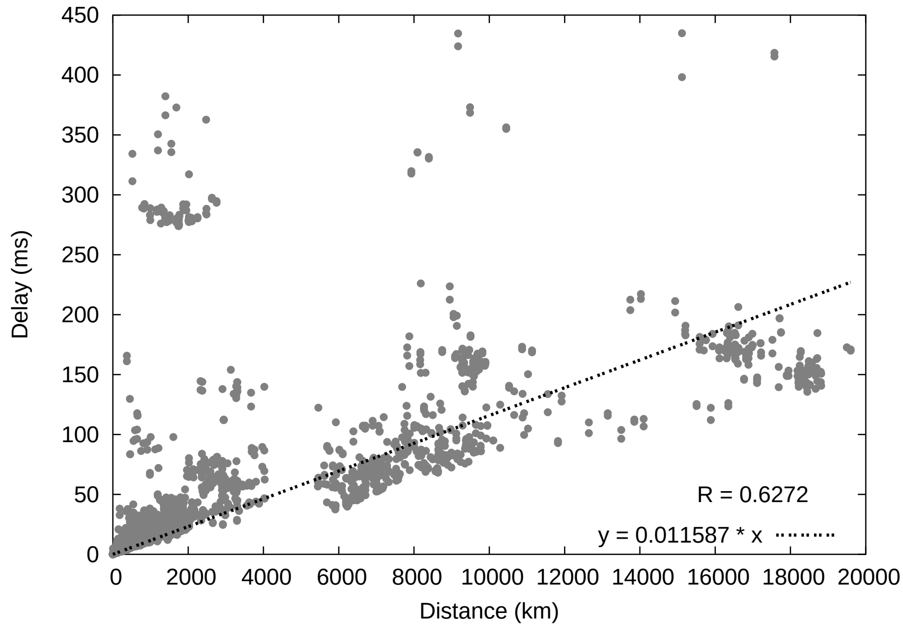


Figure 7.4: Correlation between distance and delay for the RIPE data set.

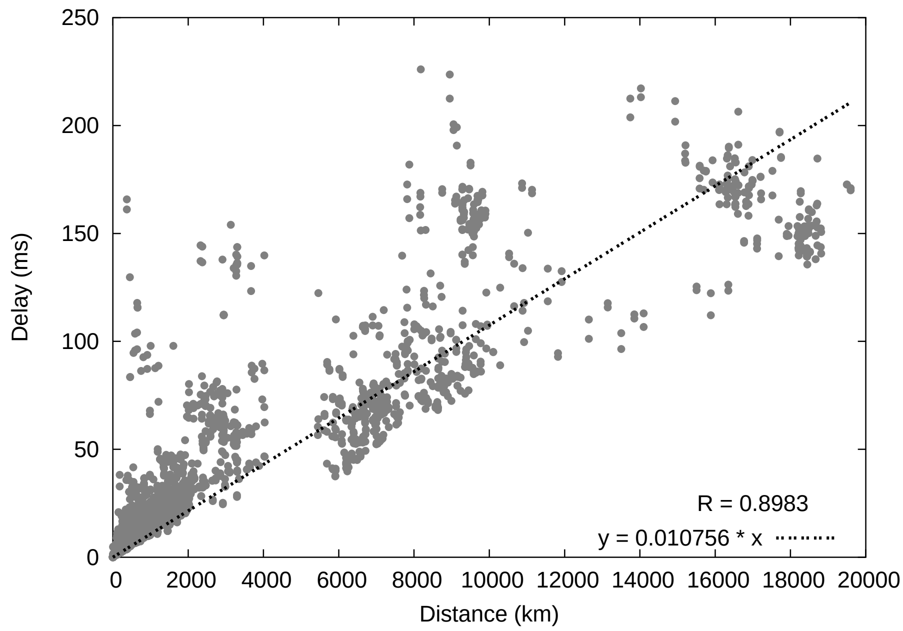


Figure 7.5: Correlation disregarding the outlier host in the RIPE data set.

7.2.3 Considerations on the Host Location from Delay Measurements

From our study we observe a weak to moderate correlation between geographic distance and network delay worldwide. Similar results are presented in [168]. Nevertheless, we observe a much stronger correlation on regions with richer connectivity, indicating that the correlation becomes stronger as connectivity becomes richer. Furthermore, within regions with richer connectivity, we observe that hosts located in nearby areas present similar delays to fixed measurement points. Therefore, we verify the assumption that hosts with similar network delays to some fixed probe machines tend to be located close to each other. Recent findings [162, 204] indicate a strong correlation between population and router density in economically developed countries. Moreover, most users, and consequently most hosts to be located, are likely to be in regions presenting richer connectivity, whereby a stronger correlation between distance and delay holds.

Two main aspects contribute to the robustness of the host inference from delay measurements against factors that may weaken the correlation between distance and delay. First, delay is measured from multiple geographically distributed probe machines rather than from one single location. Second, the minimum delay, among several delay samples, is considered rather than an individual delay sample.

7.3 Demographic Placement of Landmarks and Probe Machines

A landmark is a reference to be used as a location estimation of a certain quantity of hosts supposed to be nearly located. As a landmark is a reference for the geographic position of a set of nearby hosts, it should ideally represent the position of as many hosts as possible. Areas with high host density with a co-located landmark provide location estimations that reflect more accurately the positions of a large number of hosts. Therefore, landmarks are expected to indicate where most users (and hosts) are. A strong correlation between population and router density is found in economically developed countries and in urban areas characterized by population density peaks [162, 204]. Most users, and consequently most hosts to be located, are likely to be in regions with rich connectivity. We propose a demographic placement approach to address the issue of placing landmarks and probe machines. In this approach, we place landmarks and probe machines according to the user (host) population distribution.

The main urban agglomerations spread worldwide are considered since they offer the highest concentration of users (hosts to be located). We consider all urban agglomerations with more than one million inhabitants [205], totalizing 407 locations. For illustration, Table 7.1 shows the top 10 urban agglomerations worldwide. It is known that the Internet infrastructure varies dramatically across different regions throughout the world. Therefore, we weight the populations of the different agglomerations with the number of Internet users in the country the agglomeration belongs to over the total population of the country. In applying such a weight, we estimate the main user agglomerations worldwide to be covered by the demographic placement. Table 7.2 presents the top 10 user agglomerations worldwide

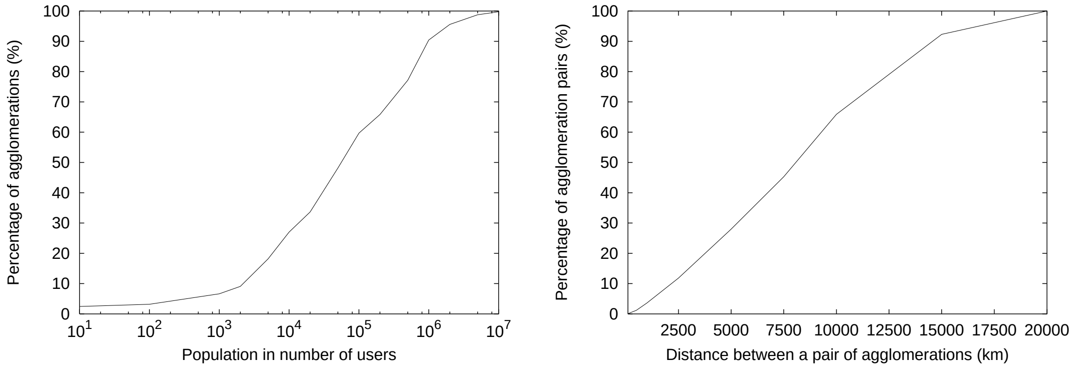
Table 7.1: Top 10 urban agglomerations worldwide.

Rank	Urban Agglomeration	Country	Total Population
1	Tokyo	Japan	34,900,000
2	New York	USA	21,600,000
3	Seoul	South Korea	21,150,000
4	Mexico City	Mexico	20,750,000
5	São Paulo	Brazil	20,250,000
6	Bombay	India	18,150,000
7	Osaka	Japan	18,000,000
8	Delhi	India	17,150,000
9	Los Angeles	USA	16,800,000
10	Jakarta	Indonesia	15,850,000

Table 7.2: Top 10 user agglomerations worldwide.

Rank	User Agglomeration	Country	User Population
1	New York	USA	11,496,837
2	Los Angeles	USA	8,941,984
3	Tokyo	Japan	7,449,579
4	Seoul	South Korea	6,755,013
5	Chicago	USA	5,003,253
6	Washington	USA	4,178,248
7	London	UK	3,868,029
8	San Francisco	USA	3,858,892
9	Osaka	Japan	3,842,190
10	Philadelphia	USA	3,353,244

out of the total 407 considered agglomerations. Data on estimations of the Internet users and total population of each country are available in [206]. Denoting the set of user agglomerations as $\mathcal{A}$, the landmark placement problem can now be formally defined. The set $\mathcal{A}$ represents the 407 user agglomerations, totalizing 173,696,253 estimated users [205, 206]. The cumulative user distribution over $\mathcal{A}$ is presented in Fig. 7.6(a). The mean distance between each pair of elements in $\mathcal{A}$ is 8167 km. The cumulative distribution of this distance is shown in Fig. 7.6(b). The landmark placement problem is to find a set of landmarks $\mathcal{L} \subseteq \mathcal{A}$ with K landmarks subject to an optimization condition $f(\mathcal{L})$. The possible locations of landmarks are limited to the locations of the agglomerations ($\mathcal{L} \subseteq \mathcal{A}$). From the viewpoint of the optimization condition $f(\mathcal{L})$, the solution to the placement problem under this restriction may be inferior to a solution that allows landmarks to be placed anywhere, i.e. not necessarily on the locations of the agglomerations. Allowing landmarks to be placed



(a) User distribution

(b) Distance distribution

Figure 7.6: Cumulative user and distance distributions over the set $\mathcal{A}$.

anywhere might provide geographically closer sites to place landmarks with respect to the fixed set of agglomerations $\mathcal{A}$. The resulting locations, however, might include sites with no network infrastructure such as rural areas, forests, oceans, and alike. Limiting the placement of landmarks to the locations of the agglomerations is thus justified as it is suitable to place landmarks where is likely to exist a previous network infrastructure. Furthermore, agglomerations with a high level of network infrastructure provide a large set of hosts able to be used as landmarks.

Eligible landmarks are any host able to echo `ping` messages and known to be located within the indicated agglomerations, thus resulting in a potentially large number of eligible landmarks in large user agglomerations. Such landmarks may even be unsuspecting participants in the procedure. As a consequence, for each considered urban agglomeration, a set of oblivious hosts may be selected to constitute a pool of eligible landmarks. Since the probe machines need to use one landmark from a certain urban agglomeration, the probe machines pick one of the eligible landmarks located in the urban agglomeration. The association between a geographic location and the pool of eligible landmarks located therein may be carried out using DNS, performing load balancing among co-located landmarks. Each pool of landmarks might be named `new-york.netlandmark.info`, `london.netlandmark.info`, and so forth rather than adding IP addresses into the code of probe machines, thus enhancing robustness against individual landmark failures.

We consider two complementary approaches to determine the placement of the set of landmarks $\mathcal{L}$ of size K taking into account the concentration of users in $\mathcal{A}$. Such approaches have been used to determine the placement of different kinds of facilities like fire stations, hospitals, or police departments [207]. In the first approach, given the maximum distance from an agglomeration to a landmark, we want to determine the minimum number of landmarks needed to cover all agglomerations. If the condition of covering all agglomerations is relaxed, a smaller number of landmarks may cover a large portion of the considered space

Table 7.3: Adopted notation for placing landmarks and probe machines.

g_{ij}	geographic distance between agglomerations i and j
h_i	number of users at agglomeration i
G	geographic coverage distance
W	maximum distance from any agglomeration to the nearest landmark
M	minimum distance between any pair of placed probe machines
a_{ij}	$\begin{cases} 1 & \text{if agglomeration } i \text{ can cover demands at agglomeration } j, \\ 0 & \text{if not.} \end{cases}$
X_i	$\begin{cases} 1 & \text{if there is a landmark on agglomeration } i, \\ 0 & \text{if not.} \end{cases}$
Z_i	$\begin{cases} 1 & \text{if agglomeration } i \text{ is covered,} \\ 0 & \text{if not.} \end{cases}$
Y_{ij}	$\begin{cases} 1 & \text{if agglomeration } i \text{ is assigned to a landmark at site } j, \\ 0 & \text{if not.} \end{cases}$
Q_i	$\begin{cases} 1 & \text{if there is a probe machine on agglomeration } i, \\ 0 & \text{if not.} \end{cases}$

of users as the user distribution is unequal throughout the agglomerations. Thus, given the maximum distance between an agglomeration and a landmark, we want to know how many landmarks are needed to cover at least a certain portion of the considered users, if not all. This model is known as the maximum covering location model [208]. In the second approach, after fixing the number of landmarks (K) to be located, we minimize the maximum distance from any agglomeration to its nearest landmark. This problem is known as the K-center problem [208].

The placement of probe machines is also investigated. We apply the demographic approach to place probe machines on sites with enough network infrastructure to make their deployment feasible and in a fashion to avoid gathering redundant measurement data. Table 7.3 presents the adopted notation to model the problem of placing landmarks and probe machines.

7.3.1 Maximum Covering Location Model

The maximum covering location model is obtained when the number of covered hosts to be located is maximized, considering a limited number of landmarks and a fixed coverage distance of each landmark. A landmark covering a certain region is a location estimation for the hosts within that region. The demographic placement of landmarks takes into account the concentration of users within the agglomerations. Using the notation from Table 7.3, the maximum covering location model is expressed by the objective function

$$f_1(\mathcal{L}) = \max \sum_i h_i Z_i \tag{7.3}$$

with the following constraints:

$$Z_i \leq \sum_j a_{ij} X_j \qquad \forall i \tag{7.4}$$

$$\sum_j X_j \leq K \tag{7.5}$$

$$X_j = 0, 1 \qquad \forall j \tag{7.6}$$

$$Z_i = 0, 1 \qquad \forall i \tag{7.7}$$

The constraint (7.4) states that users at the agglomeration i are covered if at least one site that covers agglomeration i is selected to host a landmark. The constraint (7.5) stipulates that we locate no more than K landmarks. The constraints (7.6) and (7.7) are the constraints of integrality for the decision variables X and Z.

We adopt a greedy approach [208] to obtain the maximum coverage with time complexity $O(|\mathcal{A}|^2 K)$. After placing the first landmark to cover the most uncovered demand, the algorithm greedily looks for the best location for the next landmark until K landmarks are placed. The greedy algorithm is as follows:

$\mathcal{L} \leftarrow \emptyset$; $\mathcal{A}' \leftarrow \mathcal{A}$
while $(|\mathcal{L}| < K)$ **do**
    Find agglomeration $A \in \mathcal{A}'$ that covers the most uncovered demand.
    $\mathcal{L} \leftarrow \mathcal{L} \cup A$
    Set $C \subseteq \mathcal{A}'$ to be the set of all agglomerations covered by A.
    $\mathcal{A}' \leftarrow \mathcal{A}' - C$
end while
$\mathcal{L}$ is the set of K landmarks

Fig. 7.7 presents the number of landmarks needed to cover a certain percentage of the considered number of users for different values of coverage distance. The incremental user coverage decreases as additional landmarks are assigned. There is a tradeoff between the number of landmarks and the accuracy of the location information represented by the coverage distance. For larger coverage distances, one landmark acts as a location estimation for a larger amount of users, but at a lower accuracy. As a consequence, depending on the requirements on location accuracy of the demanding application, a limited number of landmarks is enough to provide the requested information.

Fig. 7.8 shows the percentage of covered users achieved for three policies of landmark placement and coverage distances of 10 and 500 km. The used policies were: random placement, geographic placement, and demographic placement. Under random placement, the landmarks are randomly selected, disregarding both the concentration of users and the proximity between agglomerations. Geographic placement disregards the user concentration ($h_i = 1, \forall i$) within the candidate agglomerations to locate landmarks. The number of

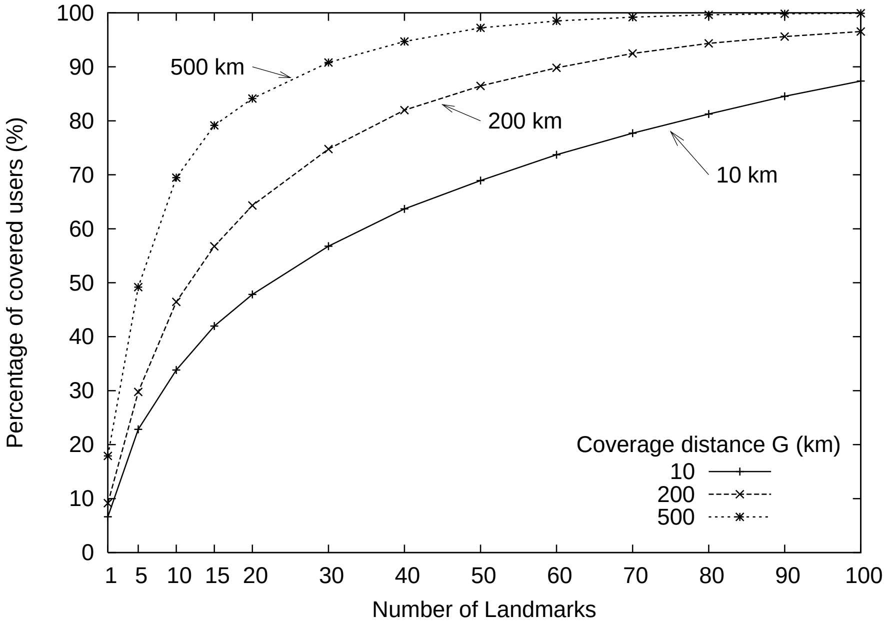
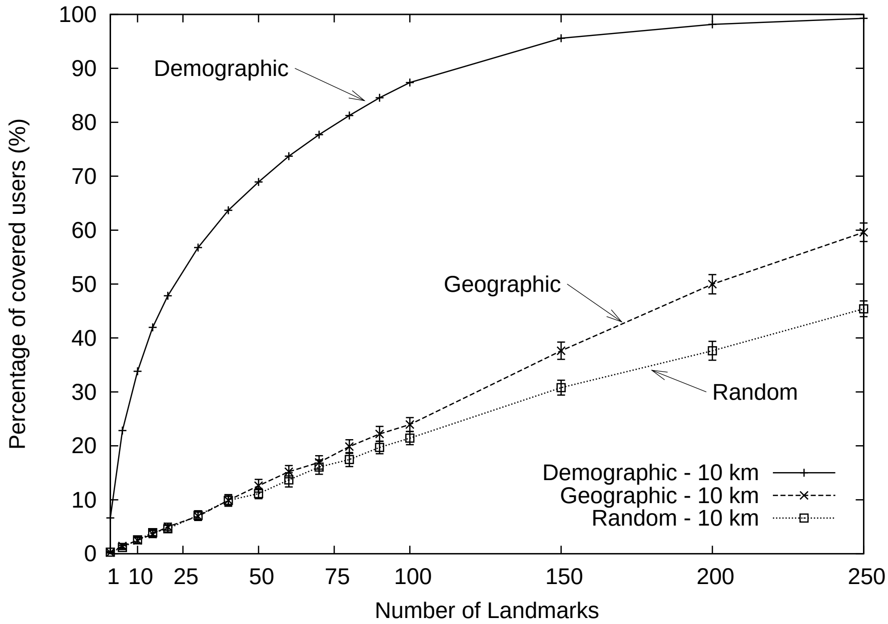


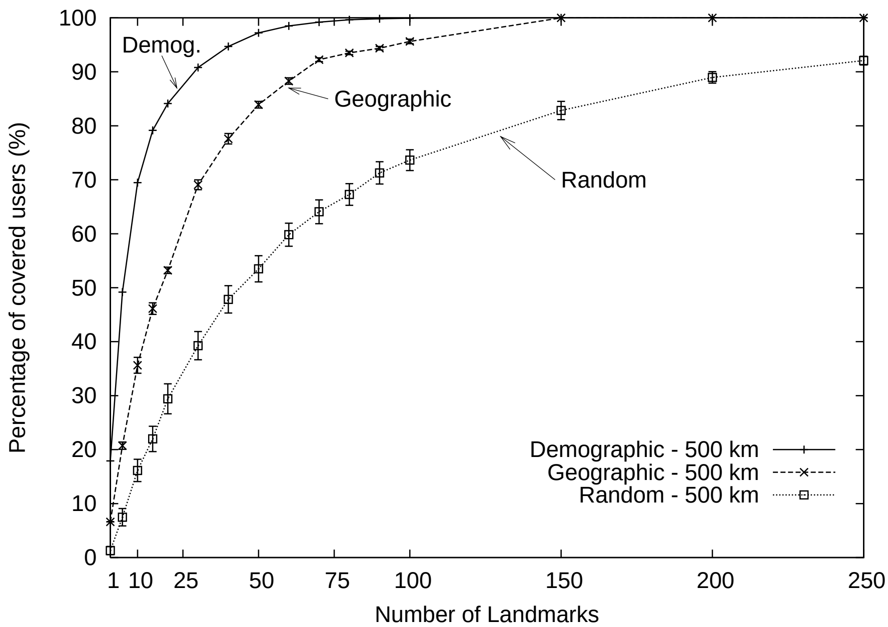
Figure 7.7: Covered users in the maximum coverage location model.

agglomerations within the range given by the coverage distance G around a agglomeration determines the weight of the agglomeration. Among agglomerations with equal weight, the elected agglomeration to locate the landmark is randomly chosen. Error bars in the results from the random and geographic placement represent the 99% confidence interval. The demographic placement of landmarks significantly improves the representativeness of the chosen landmarks in terms of hosts to be eventually located when compared with the random and the geographic placement policies.

The proposed methodology provides not only the number of needed landmarks to cover a certain percentage of users, but also the specific geographic locations of such landmarks. The geographic distribution of the landmarks indicated by the demographic placement approach is visualized using the GeoPlot tool [209]. Fig. 7.9 illustrates the demographic placement of 50 landmarks to cover almost 90% of the considered number of users for a coverage distance of 250 km. In Fig. 7.9, landmarks are concentrated in three main regions: USA, Western Europe, and Japan. Such results are consistent with recent findings [162, 204] that show a high density of Internet infrastructure on such regions. The high density of placed landmarks on these three regions reflects the concentration of users therein. There are some areas of relatively high user concentration elsewhere, as in Brazil and Australia. Results indicate where to place landmarks to improve the representativeness of each landmark given the constraints on the accuracy and the number of landmarks to be placed.



(a) $G = 10$ km



(b) $G = 500$ km

Figure 7.8: Covered users in demographic, geographic, and random placement.

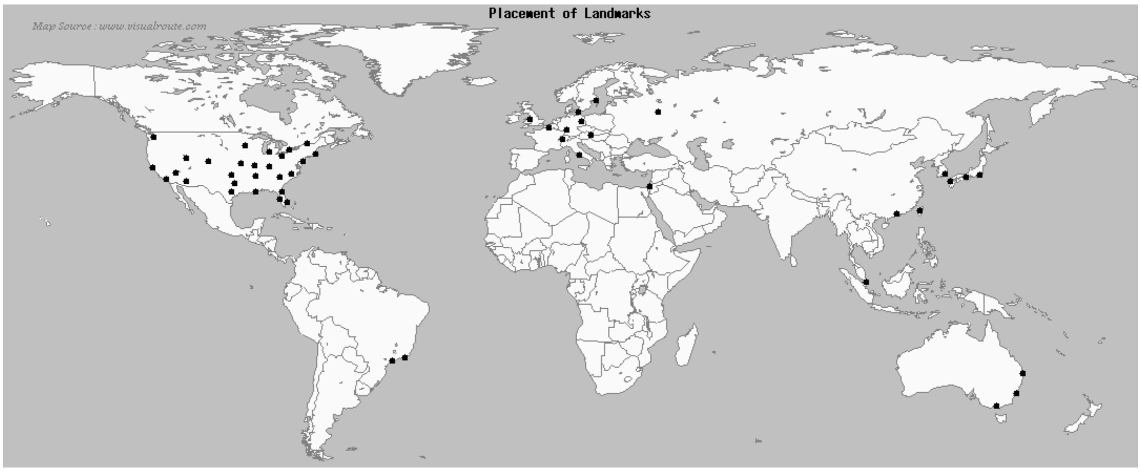


Figure 7.9: Placement of 50 landmarks for a coverage distance of 250 km.

7.3.2 K-center Problem

We now tackle the problem of minimizing the maximum distance between an agglomeration and the nearest landmark while considering a fixed number of landmarks. Even if an agglomeration i is within the coverage distance of a landmark j ($a_{ij} = 1$), it may be assigned to another closer landmark k ($Y_{ij} = 0$ and $Y_{ik} = 1$). Therefore, $a_{ij} \geq Y_{ij}$ as there is no reason to assign an agglomeration to a landmark other than the closest one. This minimization problem is known as the K-center problem [208] and may be formulated using the notation defined in Table 7.3 by the objective function

$$f_2(\mathcal{L}) = \min W \tag{7.8}$$

and it is subject to the following constraints:

$$\sum_j X_j = K \tag{7.9}$$

$$Y_{ij} \leq X_j \qquad \forall i, j \tag{7.10}$$

$$W \geq \sum_j g_{ij} Y_{ij} \qquad \forall i \tag{7.11}$$

$$X_j = 0, 1 \qquad \forall j \tag{7.12}$$

$$Y_{ij} = 0, 1 \qquad \forall i, j \tag{7.13}$$

The constraint (7.9) stipulates that K landmarks are to be located. The constraint (7.10) states that agglomeration i can only be covered by site j if site j has a landmark. The constraint (7.11) defines the lower bound on the maximum distance, which is being minimized. In other words, the maximum distance (W) must be greater or equal than the distance between any agglomeration i and the landmark j to which the agglomeration i is assigned. The constraint (7.12) is the integrality constraint on the decision variable X. The constraint (7.13) requires the agglomeration i to be assigned to only one landmark j.

In the demographic approach, in order to consider the concentration of users in each agglomeration, constraint (7.11) should be replaced by $W \geq h_i \sum_j g_{ij} Y_{ij}, \forall i$. Otherwise, the minimization problem leads to a geographic placement of landmarks. One solution to the problem is to enumerate each possible subset of size K out of the $|\mathcal{A}|$ candidate locations and then verify which one provides the minimum value of W. Nevertheless, even for moderate values of $|\mathcal{A}|$ and K such an enumeration is not realistic, as the K-center problem is known to be NP-Complete [210]. An alternative binary search algorithm [208] to approximately solve the K-center problem is outlined as follows:

1: $G_{\max} \leftarrow \max_{i,j}\{g_{ij}\}$
2: $G_L \leftarrow 0$; $G_H \leftarrow G_{\max}$
3: **while** $(G_H \neq G_L)$ **do**
4:     $G \leftarrow \lfloor (G_H + G_L)/2 \rfloor$
5:     Compute the smallest set of landmarks $\mathcal{L}(G)$ that covers the entire set of agglomerations $\mathcal{A}$ for a coverage distance of G.
6:     **if** $(|\mathcal{L}(G)| \leq K)$ **then**
7:         $G_H \leftarrow G$
8:     **else**
9:         $G_L \leftarrow G + 1$
10:     **end if**
11: **end while**
12: G_L is the solution to the objective function and $\mathcal{L}(G_L)$ provides the locations of the K landmarks for the solution.

Using the binary search approach, we are able to solve the K-center problem with time complexity $O(|\mathcal{A}|K^2 \log G_{\max})$. This approach defines a lower bound G_L and an upper bound G_H on the maximum distance between an agglomeration and the nearest landmark. The approach then successively narrows the range between such bounds until they converge into the smallest coverage distance that allows a set of K landmarks to cover all agglomerations. This coverage distance is the smallest maximum distance between an agglomeration and the nearest landmark for K placed landmarks, thus being the approximate solution of the objective function (7.8). The set of K landmarks that covers all agglomerations for such a coverage distance is the solution to the K-center problem. The binary search algorithm considers unweighted distances, resulting in a geographic placement of landmarks. In the demographic strategy, to consider demand-weighted distances, step 1 in the binary search algorithm should state $G_{\max} \leftarrow \left[\max_{i,j}\{g_{ij}\}\right]\left[\max_i\{h_i\}\right]$. Furthermore, in computing the smallest set of landmarks that covers the entire set of agglomerations (step 5 in the binary search algorithm), an agglomeration j is able to cover an agglomeration i if $g_{ij}h_i \leq G$.

We compare the results from the random, geographic, and demographic placement policies in Fig. 7.10. The geographic placement performs the best, providing the lowest maximum geographic distances between agglomerations and the nearest landmark. The demographic approach provides higher maximum and average distances, but such results mask the concentration of users within the agglomerations. Adding more landmarks may even keep the maximum distance observed under the demographic placement policy unchanged. A remote agglomeration with low user concentration may be kept far from the nearest landmark as additional landmarks are used to decrease the distance between more user populated agglomerations to the nearest landmark. This is illustrated in Fig. 7.10(a) between 50 and 150 placed landmarks as the maximum distance keeps leveled off. In spite of that, as shown in Fig. 7.10(b), the average distance to the nearest landmark keeps decreasing in this same range as the density of landmarks in denser user areas increases.

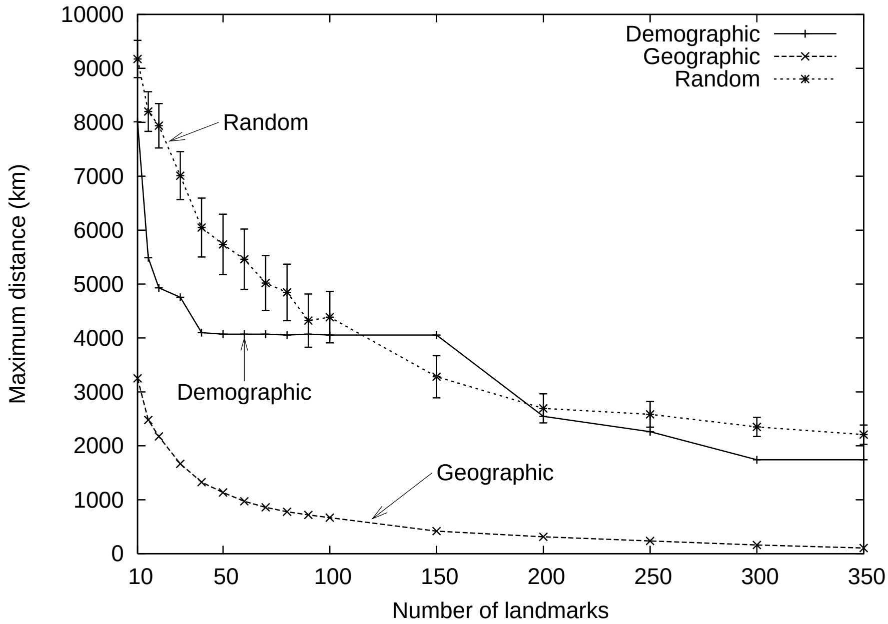
The demographic placement considers the user concentrations within the different agglomerations. As a consequence, the demographic placement pushes the worst-case distances toward the farthest and least user concentrated agglomerations. High user concentrated agglomerations are assigned to closer landmarks and the agglomerations with the highest user concentrations host the landmarks. These results are presented in Fig. 7.11 using 50 and 100 placed landmarks. The demographic strategy provides smaller distances between the agglomerations and their nearest landmark for the most part of users at the expense of leaving farther landmarks for the remote and least user concentrated agglomerations.

Fig. 7.12 illustrates the difference in adopting either unweighted (geographic placement) or weighted (demographic placement) distances to be minimized in the K-center problem. Fig. 7.12(a) presents the placement of 50 landmarks considering the geographic placement, thus disregarding the concentration of users. Not considering the concentration of users leads to a more uniform geographic distribution of the 50 landmarks worldwide and to a lower maximum unweighted or purely geographic distance W between any agglomeration and its nearest landmark. In the other hand, Fig. 7.12(b) presents the placement of 50 landmarks considering weighted distances under the demographic placement. The maximum (unweighted or purely geographic) distance W between an agglomeration and the nearest landmark is higher, but the worst cases are pushed to remote agglomerations with the lowest concentrations of users. The result is a dense placement of landmarks in the regions presenting the highest concentrations of users like the USA, Western Europe, and Japan.

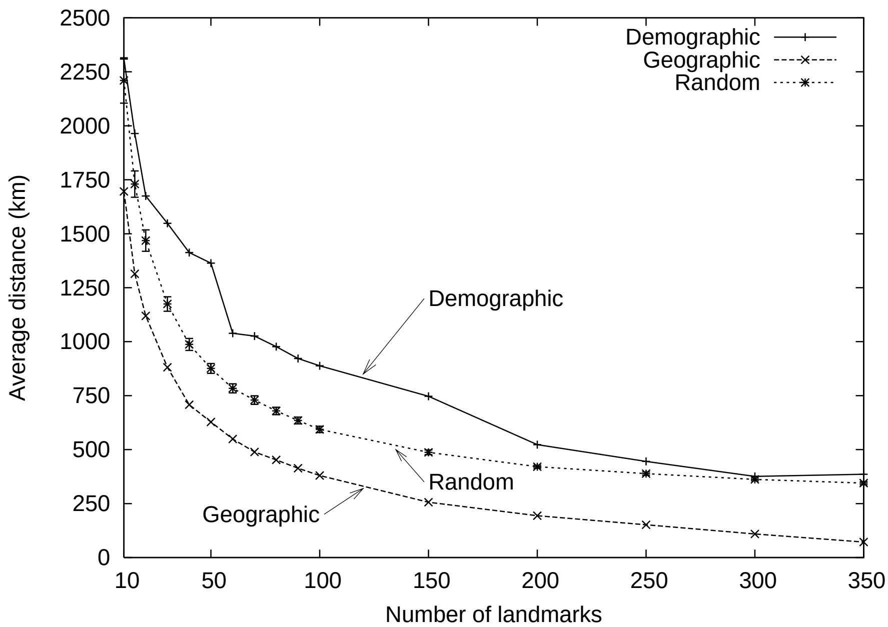
7.3.3 Placement of Probe Machines

In this subsection, we address the issue of where to place the probe machines. Probe machines measure the delay to a target host and regularly gather the delays to landmarks. Probe machines that are near each other may share common paths to some remote landmarks or target hosts, thus providing redundant information to the location estimation decision. Therefore, our first goal is to make a geographically sparse distribution of probe machines to avoid shared paths. This sparse placement of probe machines avoids a possible correlation between the delay patterns gathered from the landmarks. In order to make the deployment of probe machines feasible, they are placed on agglomerations with better network infrastructure. The second goal is thus to maximize the number of users on the agglomerations selected to host probe machines. We adopt the number of users (hosts) as a means to reflect the level of network infrastructure on a given agglomeration.

The problem of placing probe machines is to find a set $\mathcal{P} \subseteq \mathcal{A}$ with N probe machines that maximizes the minimum weighted distance between any pair of placed probe machines.

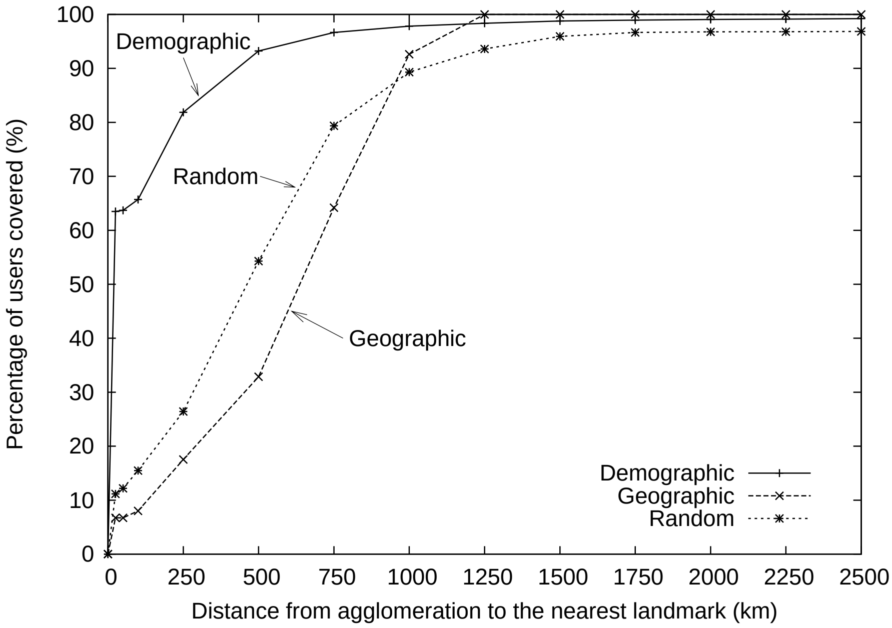


(a) Maximum geographic distance

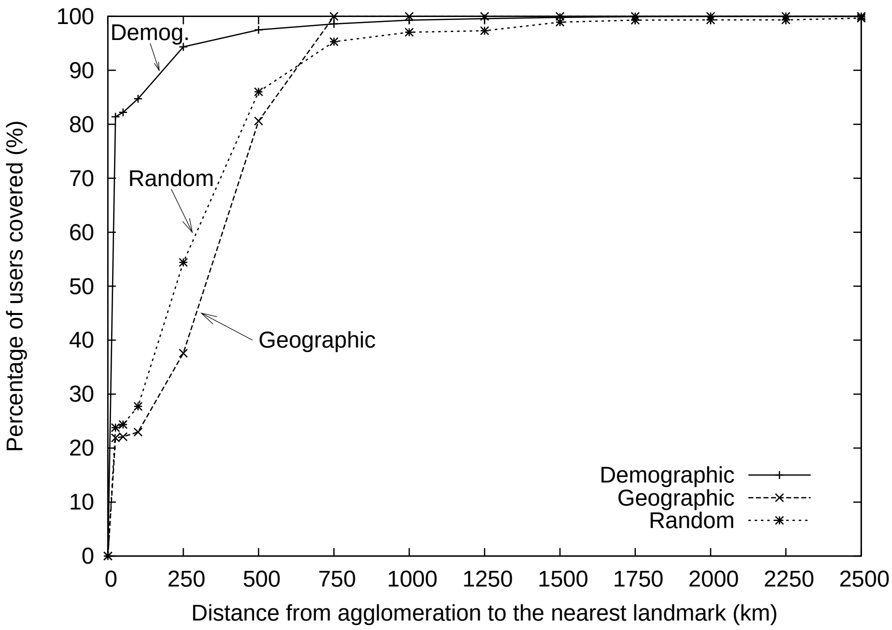


(b) Average geographic distance

Figure 7.10: Distance from an agglomeration to the nearest landmark (km).

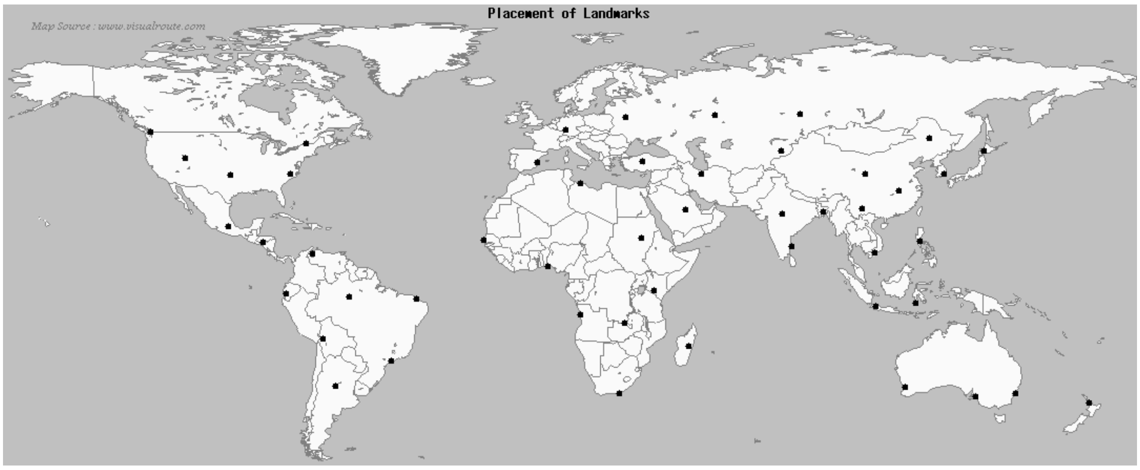


(a) $K = 50$

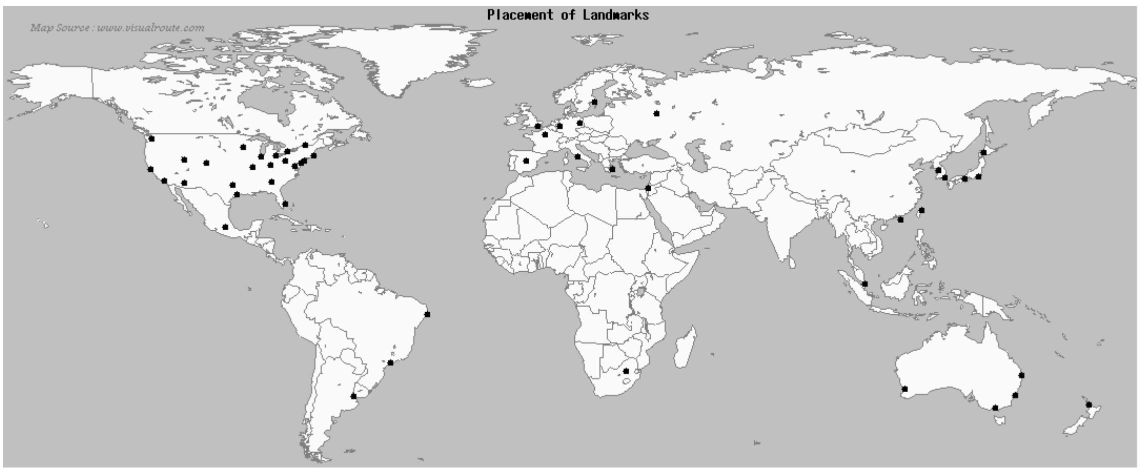


(b) $K = 100$

Figure 7.11: Covered users as a function of the distance to the nearest landmark.



(a) Geographic placement (disregards the concentration of users)



(b) Demographic placement (considers the concentration of users)

Figure 7.12: K-center solution for placing 50 landmarks.

We adopt the weighted distance between agglomerations to consider the user concentration at each agglomeration and thus place probe machines on agglomerations that have better network infrastructure. Using the notation defined in Table 7.3, the problem of placing probe machines is given by the objective function

$$f_3(\mathcal{P}) = \max M \tag{7.14}$$

and is subject to the following constraints:

$$\sum_i Q_i = N \tag{7.15}$$

$$M \leq h_i \sum_j g_{ij} Q_j \qquad \text{if } Q_i = 1, \forall i \tag{7.16}$$

$$Q_i = 0, 1 \qquad \forall i \tag{7.17}$$

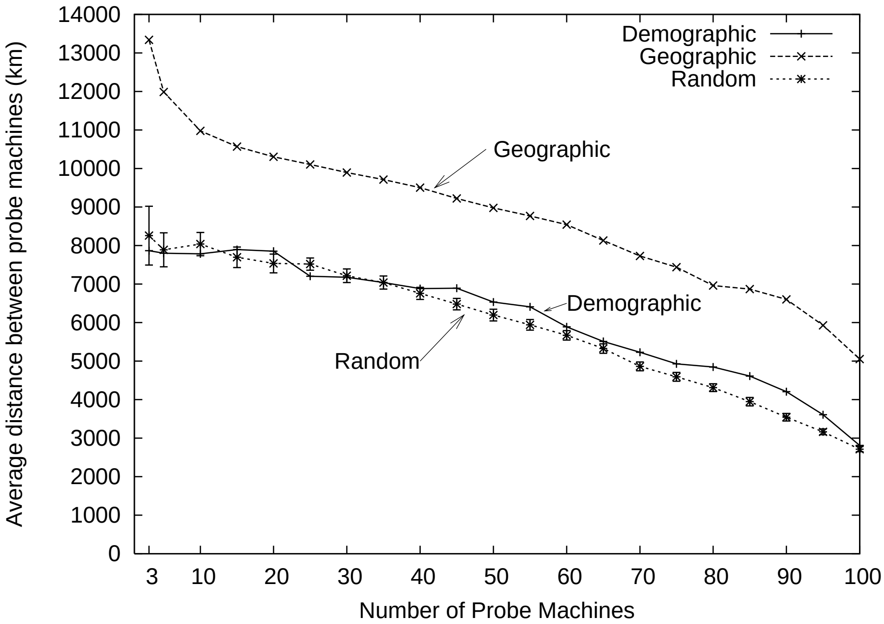
The constraint (7.15) states that N probe machines are to be located. The constraint (7.16) defines the upper bound on the minimum distance between two placed probe machines, which is being maximized. The minimum weighted distance M must be less than or equal to the distance between any pair of probe machines i and j. This constraint makes no sense if $Q_i = 0$ (no probe machine at agglomeration i) as M is the weighted distance between a pair of probe machines i and j. The constraint (7.17) is the integrality constraint on the decision variable Q.

We propose a greedy approach to solve the problem of placing probe machines with time complexity $O(|\mathcal{A}|^2 N)$. This greedy approach is as follows:

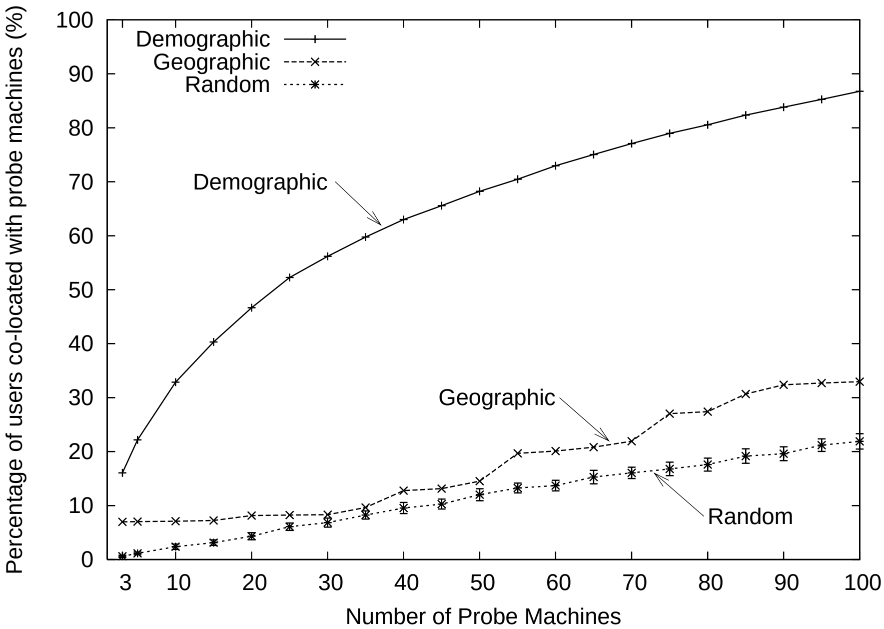
$\mathcal{A}' \leftarrow \mathcal{A}$
$\mathcal{P} \leftarrow A \in \mathcal{A}$ with $\max_i\{h_i\}$
while $(|\mathcal{P}| < N)$ **do**
    Find agglomeration $A \in \mathcal{A}'$ that has the maximum (weighted) distances toward the previously placed probe machines.
    $\mathcal{P} \leftarrow \mathcal{P} \cup A$
    $\mathcal{A}' \leftarrow \mathcal{A}' - A$
end while
$\mathcal{P}$ is the set of N probe machines

The set of probe machines is initialized with the agglomeration that presents the highest user concentration. Afterwards the algorithm greedily places a new probe machine to maximize the weighted distance between the new probe machine and the previously placed probe machines until N probe machines are placed. The first placed probe machine influences the provided solution. Nevertheless, the decision to place the first probe machine on the agglomeration with the highest user concentration is consistent with the argumentation of placing probe machines where exists enough network infrastructure.

We compare the results provided to the problem of placing probe machines by the demographic, geographic, and random placements in Fig. 7.13. Under the geographic placement, to consider unweighted distances, the constraint (7.16) is replaced by $M \leq \sum_j g_{ij} Q_j$, if $Q_i = 1$, for all i. In Fig. 7.13(a), we compare the average distance observed between all pairs of the N placed probe machines for the three placement policies. The geographic placement performs the best in sparsely distributing the probe machines as it presents the largest average distances. Nevertheless, such results place probe machines on remote user agglomerations that not necessarily have a good network infrastructure since the geographic placement disregards user concentration throughout the agglomerations.



(a) Distance between probe machines



(b) Users co-located with probe machines

Figure 7.13: Placing probe machines.

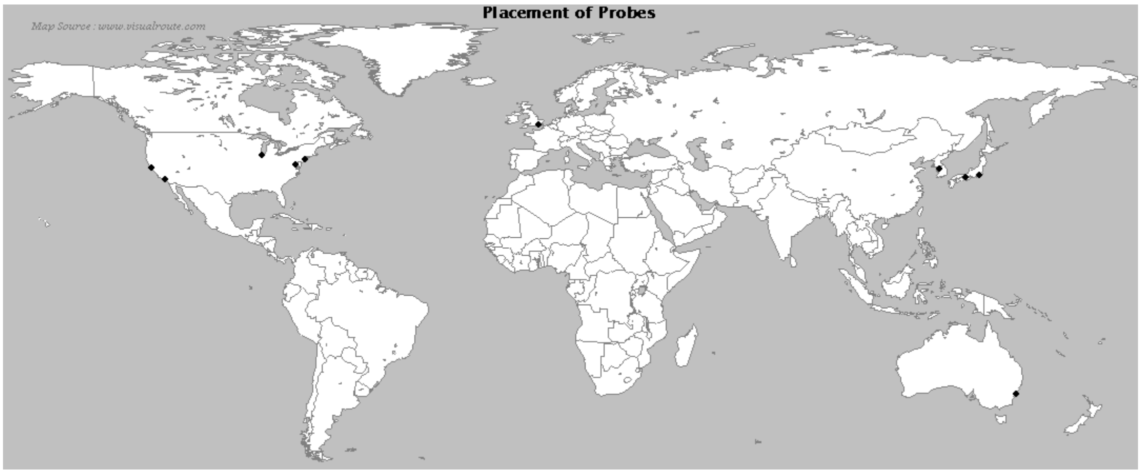


Figure 7.14: Demographic placement of 10 probe machines.

Fig. 7.13(b) shows that the demographic placement locates the probe machines on agglomerations with high density of users, which is our adopted criterion to measure the level of network infrastructure. The demographic placement largely outperforms the geographic and random placements. In the demographic placement, probe machines are placed on agglomerations that have a sufficient network infrastructure to make the deployment of probe machines feasible. Meanwhile, the demographic placement is still able to provide a relatively sparse distribution of probe machines to avoid the shared paths toward landmarks and hosts to be located as shown in Fig. 7.13(a).

Fig. 7.14 illustrates the placement of 10 probe machines adopting the demographic placement. Since the proposed methodology considers the user concentration, the probe machines are placed on the regions with the highest user concentrations. Meanwhile, such probe machines are also geographically spread out on far regions that present a relatively good network infrastructure, like Australia for instance. For 10 placed probe machines, there is already some probe machines placed in relatively nearby areas, indicating that a larger number of probe machines may provide a certain level of redundant information.

7.4 Measuring the Similarity between Delay Patterns

In this section, we investigate how to best measure the similarity between the delay pattern of each landmark and the one observed for the target host. The delay patterns result from the partial viewpoints gathered by the distributed probe machines. The landmark that presents the most similar delay pattern with respect to the one of the target host provides the location estimation of that host. Measuring the similarity of the concerned delay patterns is thus a key point for the accuracy of the host location from delay measurements.

7.4.1 Similarity Models

We define the similarity function $\mathcal{S}(\mathbf{x}, \mathbf{y}) : \mathbb{R}^N \to [0, 1]$ to measure the degree of similarity between two delay patterns $\mathbf{x}$ and $\mathbf{y}$ where N is the number of adopted probe machines. We adopt such a function to evaluate the degree of similarity between the delay patterns gathered by the probe machines from each landmark and from the target host to be located. In this subsection, we introduce three similarity models [198]. Each similarity model provides its own manner of implementing the similarity function $\mathcal{S}(\mathbf{x}, \mathbf{y})$. The resulting similarity level between the delay patterns falls in the interval $0 \leq \mathcal{S}(\mathbf{x}, \mathbf{y}) \leq 1$. The closer the similarity level is to 1, more similar the delay patterns are.

7.4.1.1 Cosine-based Similarity

In the first similarity model, the two delay patterns are thought of as two vectors in a N-dimensional delay space. The similarity between them is measured by computing the cosine of the angle θ between these two vectors. The cosine of the angle θ between the delay vectors $\mathbf{x}$ and $\mathbf{y}$ is computed by

$$\cos\theta = \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{x}\|\|\mathbf{y}\|} \tag{7.18}$$

where "$\cdot$" denotes the dot-product of the two vectors and $\|\mathbf{x}\|$ is the Euclidean size of vector $\mathbf{x} \in \mathbb{R}^N$, i.e. $\|\mathbf{x}\| = \sqrt{\sum_{i=1}^{N} x_i^2}$.

In order to be scaled to the interval $[0, 1]$, the cosine-based similarity between the delay patterns $\mathbf{x}$ and $\mathbf{y}$, denoted by $\mathcal{S}_{\cos}(\mathbf{x}, \mathbf{y})$, is given by

$$\mathcal{S}_{\cos}(\mathbf{x}, \mathbf{y}) = \frac{\cos\theta + 1}{2}. \tag{7.19}$$

7.4.1.2 Correlation-based Similarity

In this model, similarity is measured by computing the coefficient of correlation between the two delay patterns $\mathbf{x}$ and $\mathbf{y}$. The coefficient of correlation is defined as

$$\mathrm{corr}(\mathbf{x}, \mathbf{y}) = \frac{\sigma_{\mathbf{xy}}^2}{\sigma_{\mathbf{x}}\sigma_{\mathbf{y}}}, \tag{7.20}$$

where $\sigma_{\mathbf{xy}}^2$ denotes the covariance between delay patterns $\mathbf{x}$ and $\mathbf{y}$, and $\sigma_{\mathbf{x}}$ is the standard deviation of $\mathbf{x}$. The correlation-based similarity, denoted by $\mathcal{S}_{\mathrm{cor}}(\mathbf{x}, \mathbf{y})$, is scaled to the interval $[0, 1]$ and given by

$$\mathcal{S}_{\mathrm{cor}}(\mathbf{x}, \mathbf{y}) = \frac{\mathrm{corr}(\mathbf{x}, \mathbf{y}) + 1}{2}. \tag{7.21}$$

7.4.1.3 Distance-based Similarity

We represent the distance $u(\mathbf{x}, \mathbf{y})$ between two delay patterns $\mathbf{x}$ and $\mathbf{y}$ by

$$u(\mathbf{x}, \mathbf{y}) = L_p = \left[\sum_{i=1}^{N} |x_i - y_i|^p\right]^{\frac{1}{p}}, \quad p > 0. \tag{7.22}$$

The distance function $u(\mathbf{x}, \mathbf{y})$ belongs to the L_p family of functions. When $p = 1$, we have the Manhattan or city-block distance. In contrast, for $p = 2$, we have the Euclidean distance. Furthermore, $0 < p < 1$ results in a non-metric distance function adequate to be used if distances do not satisfy the triangle inequality [211]. Shepard [212] argues in favor of an exponential decay function for a distance-based similarity model. A flexible distance-based similarity model, denoted by $\mathcal{S}_{\text{dis}}(\mathbf{x}, \mathbf{y})$, which includes the exponential decay function is given by

$$\mathcal{S}_{\text{dis}}(\mathbf{x}, \mathbf{y}) = \mathrm{e}^{-(u(\mathbf{x},\mathbf{y})/\beta)^{\alpha}}, \quad \beta > 1,\ \alpha > 0, \tag{7.23}$$

where β is a scaling factor defined as

$$\beta = \max_{i,j} \left(1, \quad \frac{1}{2} \left(\frac{\sum_i \mathbf{x}_i}{\|\mathbf{x}\|} + \frac{\sum_j \mathbf{y}_j}{\|\mathbf{y}\|} \right) \right). \tag{7.24}$$

Since $\mathcal{S}_{\text{dis}}(\mathbf{x}, \mathbf{y})$ should decrease as $u(\mathbf{x}, \mathbf{y})$ increases, then $\alpha > 0$. The same kind of function is explored for measuring similarity in different contexts like pattern analysis and similarity theory [199, 211, 213].

7.4.2 Experimental Results for the Similarity Models

In this subsection, we analyze the performance of the different similarity models to provide a location estimation of a target host to be located. We adopt the RIPE data set to evaluate the three similarity models. The hosts in the RIPE data set are considered one at a time the target host to be located. The remaining hosts in the set are then used as probe machines and landmarks to perform the location estimation of the target host. We repeat this procedure to evaluate the resulting location estimation of each host in the data set.

7.4.2.1 Ranking Landmarks

As we know the geographic position of each host in the RIPE data set due to the GPS cards, we are able to determine the *ideal* landmark to be chosen by the similarity models being evaluated. The ideal landmark is the geographically closest landmark to the target host. We analyze the performance of the similarity models by ranking the ideal landmark for each target host. Each model provides a list of landmarks in descending order of similarity presented by their delay patterns with respect to the delay pattern of the target host. We then observe the rank of the ideal landmark on the ordered list resulting from each evaluated similarity model. This rank is obtained for each host in the RIPE data set being considered as a target at a time. Fig. 7.15 presents the average rank of the ideal landmark for each similarity model.

The best performance in ranking the landmarks is obtained by the distance-based similarity model when $p = 1$, equivalent to the city-block distance. The city-block distance also has the lowest sensitivity to the considered values of α since it presents a similar performance independent of the value of this parameter. Moreover, the city-block distance outperforms the Euclidean distance previously adopted in [165] as well as the cosine-based and correlation-based similarity models. Furthermore, the previously adopted Euclidean

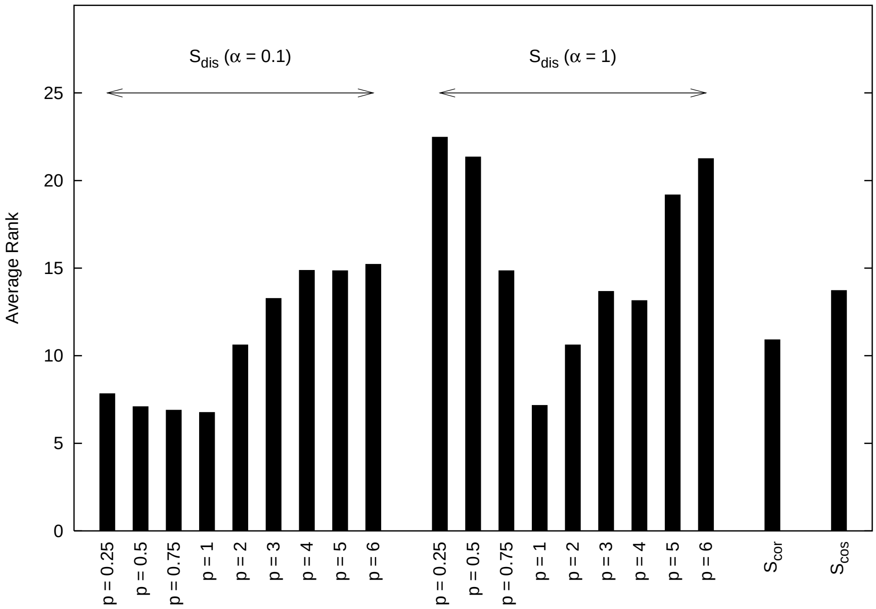


Figure 7.15: Average rank of the ideal landmark.

distance performs similarly to the correlation-based similarity model. For data following a standard normal distribution, i.e. the mean equal to zero and variance equal to one, the correlation is indeed a linear transformation of the squared Euclidean distance. This relationship is given by $\mathrm{corr}(\mathbf{x},\mathbf{y}) = 1 - L_2^2(\mathbf{x},\mathbf{y})/2N$, where $L_2(\mathbf{x},\mathbf{y})$ is the Euclidean distance between $\mathbf{x}$ and $\mathbf{y}$. Even if in our experiments the delay patterns $\mathbf{x}$ and $\mathbf{y}$ clearly do not follow the standard normal distribution, our findings indicate a certain level of correlation between the correlation-based and the distance-based (with Euclidean distance) similarity models.

We note that the distance-based similarity model is more sensitive to the variation of the distance parameter p for larger values of α. The adopted values of p and α influences on the performance of the distance-based similarity model. The results obtained by the correlation-based and cosine-based similarity models are even comprised within the performance range of the distance-based model. For $\alpha = 0.1$, relatively good performances are achieved by the distance-based model with non-metric distances ($0 < p < 1$). Some experiments show that the triangle inequality does not hold over all parts of the Internet [169, 214]. These experiments indicate that the adoption of non-metric distances may be more adequate than metric distances on specific parts of the Internet.

The cumulative probability of ranking the ideal landmark for different similarity models is depicted in Fig. 7.16. We compare the distance-based similarity model (presented in the form $\mathcal{S}_{\mathrm{dis}}$–$p$–$\alpha$) for city-block and Euclidean distances with the correlation-based and cosine-based similarity models. As the Euclidean distance has been previously adopted, we use it as a performance reference for comparison. We observe that the distance-based

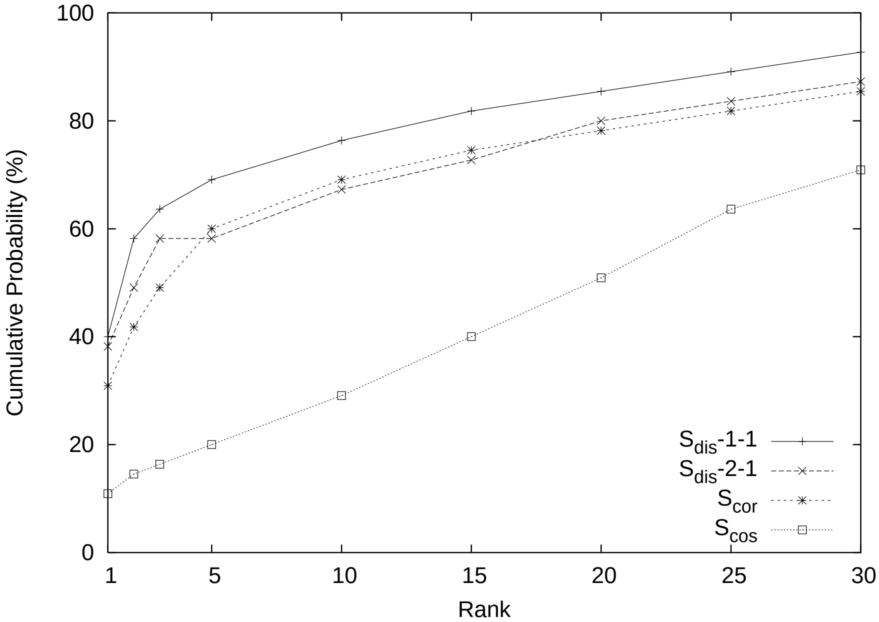


Figure 7.16: Cumulative probability of the rank of the ideal landmark.

similarity model with city-block distances outperforms the remaining ones. The correlation-based similarity model provides an equivalent performance to the distance-based similarity model with Euclidean distance.

7.4.2.2 Distance Accuracy

We now evaluate the distance accuracy of the location estimation of each considered similarity model. The location estimation corresponds to the location of the landmark chosen by each similarity model, i.e. the first ranked landmark in the resulting ordered list of each similarity model. Therefore, we evaluate the performance of the different similarity models by comparing the error distance from the selected landmark and the target host to be located.

Fig. 7.17 presents the average distance between the target host and the elected closest landmark. These results are the average distance of each RIPE host, when being considered one at a time a target host, to the nearest landmark indicated by each similarity model. For each RIPE host used as a target host, the remaining hosts are then used as landmarks and probe machines. As in the ranking landmarks evaluation, the best performance is achieved by the distance-based similarity model for the city-block distance ($p = 1$) and $\alpha = 1$. Nevertheless, the insensitivity to α is not the same for the distance accuracy as in the evaluation of ranking the ideal landmark. The correlation-based similarity model presents a similar performance to the non-metric distances with $\alpha = 0.1$, outperforming the distance-based similarity model with the reference Euclidean distance.

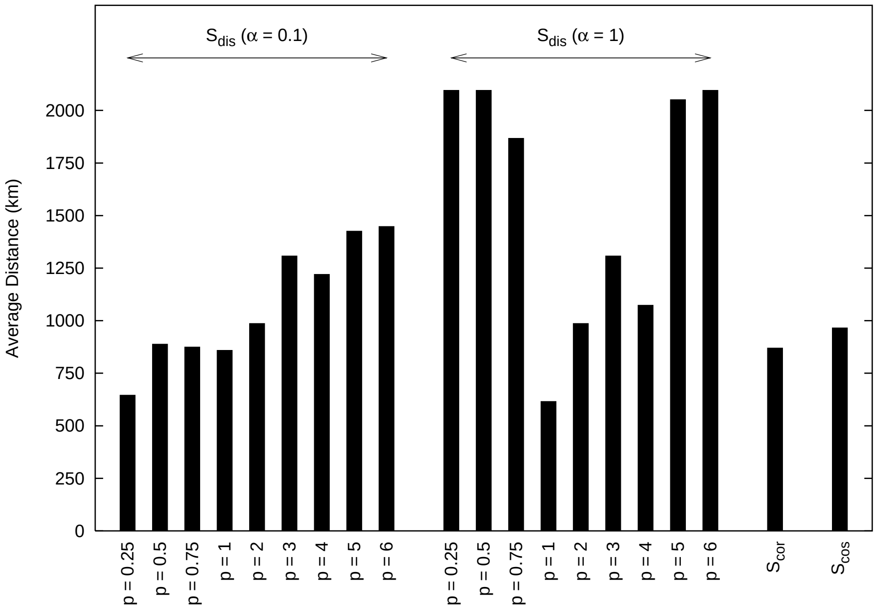


Figure 7.17: Average distance between the target host and the location estimation.

The cumulative probability of the distance between the target host and the selected closest landmark is shown in Fig. 7.18. Again, we compare the correlation-based and the cosine-based similarity models with the distance-based similarity model for city-block and Euclidean distances. The “Upper Bound” is the best possible performance for the RIPE data set. This upper bound consists of the distances from the target hosts to their respective ideal landmarks. The city-block distance ($\mathcal{S}_{\text{dis}}$–1–1) outperforms the remaining models, presenting the closest performance to the upper bound.

It should be noted that the RIPE data set is relatively small. Using one host as the target host results in 54 landmarks to choose from to infer a location estimation. Even if the elected landmark is the geographically closest landmark to the target host, not necessarily it is nearby that target host. For example, the RIPE host in Tokyo is the only one in Asia and its ideal landmark is located in Finland, 7819 kilometers away. As a consequence of the limited number of elements in the RIPE data set, the average distance from the ideal landmarks to their respective target hosts is 405 kilometers. This distance is the best average distance the similarity models can reach for the RIPE data set.

7.5 Conclusion

This chapter focused on a geographic location service of Internet hosts based on a technique that infers host locations using delay measurements from geographically distributed landmarks. We aimed at improving the accuracy of the geographic location estimation of the target host to be located.

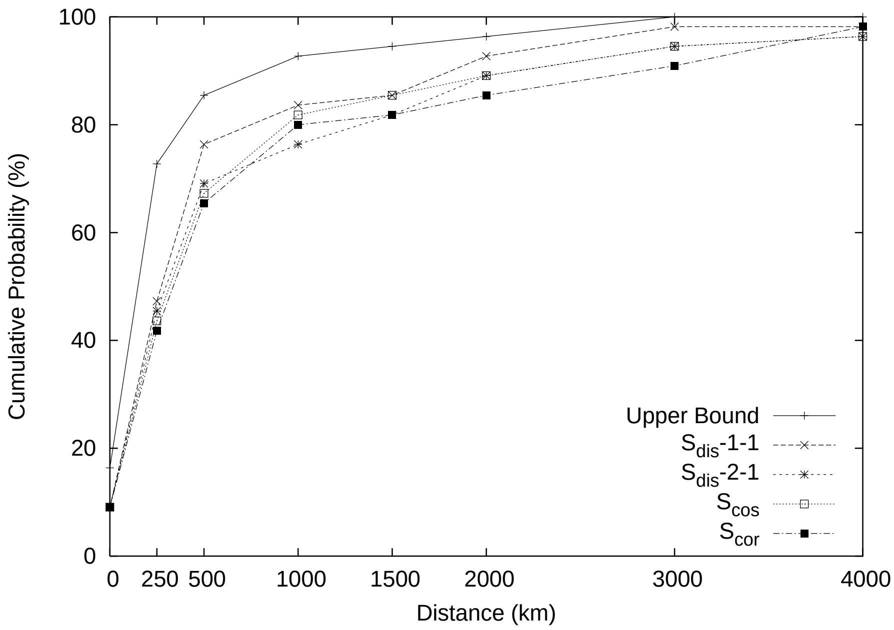


Figure 7.18: Cumulative probability of the distance (km).

The contributions of this chapter are based on three key points. First, we studied the correlation between geographic distance and network delay. Poor connectivity weakens such a correlation. We observed for two different data sets that the correlation between distance and delay becomes stronger as connectivity becomes richer. Thereafter, we identified two key issues that influence the accuracy of the resulting geographic location estimation of a target host: the placement of landmarks and probe machines, and the similarity model that compares the observed delay patterns. Thereby we addressed the problem of strategically placing landmarks and probe machines to improve the accuracy of the geographic location estimation. We proposed and evaluated a demographic placement approach that considers the geographic distribution of users (hosts) to place landmarks and probe machines. Results show that the proposed demographic placement allows a relatively small number of landmarks to represent a large portion of users within a limited coverage distance. Fewer landmarks also imply a lower amount of measurement traffic in the network. These landmarks are placed on areas of high user density, thus providing closer landmarks and more accurate location estimations for most hosts to be located. Probe machines are placed on locations that have enough network infrastructure and sparsely distributed to avoid gathering redundant data. Furthermore, in order to address the second key issue on accuracy, we investigated three different similarity models and how accurate they are for Internet host location from delay measurements. This investigation was carried out using experimental data from the RIPE measurement infrastructure. The similarity models select the landmark that provides the location estimation of the target host to be located, thus being a key component for accuracy. Despite such a fact, Euclidean distance was the only similarity model that has been previously adopted. Our findings show that the distance-based similarity model with $\alpha = 1$ and city-block distance outperforms in accuracy the remaining similarity models. This result includes the Euclidean distance that has been used as a reference for comparison. Therefore, strategically placing the landmarks and probe machines as well as using a more adequate similarity model contribute to improve the accuracy of the measurement-based geographic location of Internet hosts.

As an ongoing work at the time of the writing, we are deploying a location service based on measurements for the geographic location of Internet hosts. We use the NIMI (National Internet Measurement Infrastructure) infrastructure [215] as our probe machines[1]. We are using about 2300 landmarks that are sparsely distributed worldwide. These landmarks are extracted from the LibWeb [201] data set. We expect, by using this infrastructure, to deploy a measurement-based geographic location of Internet hosts and use it as a basis for new location-aware applications.

7.6 Résumé du Chapitre

Les applications conscientes de la localisation prennent en compte le lieu à partir du quel les clients accèdent au réseau et peuvent de cette manière introduire de nouvelles fonctionnalités dans l'Internet. Dans ce chapitre, nous nous concentrons sur l'amélioration de la précision de l'estimation de la localisation géographique d'un hôte basée sur des mesures. La localisation d'hôtes Internet est inférée en comparant des profils de délai d'hôtes de référence (*landmarks*) qui ont leurs localisations géographiques bien connues avec le profil de l'hôte à localiser. Des machines sondes réalisent les mesures ayant par but la détermination de ces profils de délai. Nous identifions deux points clés qui jouent sur la précision de l'estimation de la localisation. Cette précision dépend du placement des hôtes de référence et des machines sondes ainsi que de l'évaluation de la similarité observée parmi les profils de délai mesurés.

Les hôtes de référence doivent représenter la région où la plupart des utilisateurs se trouvent. Le problème de trouver comment placer au mieux les ressources Internet a été étudié pour différents cas tels que le placement des miroirs [194] ou des serveurs de réseau [189]. Dans ce chapitre, nous adoptons des concepts similaires pour placer des hôtes de références et machines sondes en visant à améliorer la précision de l'estimation pour la plupart des hôtes Internet. Nous proposons une politique de placement démographique. Les résultats obtenus montrent que le placement démographique fournit un nombre relativement petit d'hôtes de référence capable de représenter une portion importante des utilisateurs avec une limite de distance fixe. L'utilisation d'un nombre plus faible d'hôtes de référence réduit aussi le volume de mesures engendrées dans le réseau. Le fait de placer de manière démographique les machines sondes permet le placement distribué de ces machines pour éviter la redondance des données de mesures.

L'évaluation de la similarité d'un profil de délai d'hôtes de référence et de celui de l'hôte cible à localiser est aussi un point clé pour la précision de l'estimation de la localisation. Dans

[1] We are grateful to Andrew Adams (Pittsburgh Supercomputing Center – PSC) and Vern Paxson (ICIR) for the access to the NIMI infrastructure.

notre cas, un modèle de similarité compare les profils de délai et détermine l'hôte de référence ayant le profil de délai le plus similaire à celui de d'hôte cible. La localisation de cet hôte de référence sera l'estimation de localisation de l'hôte cible. Nous étudions par conséquent différents modèles de similarité et nous évaluons leur précision pour la localisation d'hôtes Internet par mesures. Padmanabhan et Subramanian [165] utilisent la distance euclidienne comme modèle de similarité. Un de nos modèles inclue cette distance et nous l'utilisons comme référence pour la comparaison avec les autres modèles de similarité étudiés. Les résultats montrent que d'autres modèles de similarité ont une meilleure performance que la distance euclidienne ce qui contribue à une estimation de localisation géographique plus précise des hôtes Internet.

Chapter 8

Conclusion

The main goals of this thesis were to study two possible evolutions in network functionality that can influence the design and development of new applications. These possible evolutions are the support of Quality of Service (QoS) and the location-awareness on networks. Applications that simultaneously exploit the resulting new functionalities from the support of QoS and location-awareness may be envisaged. An example is a content delivery network (CDN) that delivers multimedia content and allows a location-based adaptation of the available content. Nevertheless, the needed mechanisms to implement these new functionalities are mutually independent. Implementations are actually dependent on the particular characteristics of the context they target, such as either on ad hoc networks or the Internet. Therefore, each part of this work was dedicated to one of these possible evolutions that can leverage the functionality of networks as seen by the end users.

In the first part of this work, we investigated the problem of providing QoS support for voice and MPEG video traffic. Such QoS-sensitive streams require guarantees from the network that are seldom achieved in the current Internet. Therefore, the study of QoS mechanisms to support multimedia traffic such as voice and video is of utmost importance given the high demand for the quality delivery of multimedia content. From the experience obtained with the IntServ/RSVP architecture, it has been shown that keeping one state per flow is unfeasible at the core routers since this approach does not scale. The DiffServ architecture emerged as being a scalable alternative for QoS support. This scalability comes at the expense of providing QoS guarantees to aggregates of flows instead of individual flows.

We evaluated the support of voice traffic offered by the Expedited Forwarding (EF) scheme of the DiffServ architecture. The voice traffic was modeled by both CBR and On-Off sources. Our evaluation also considered the implementation of the EF service by either a priority queue or a WRR scheduler. Since the EF service is intended to provide strong guarantees to the differentiated aggregate, it is expected to constitute an expensive service because of the amount of needed resources. Therefore, we studied the trade-off between allowing an additional but tolerable shaping delay and contracting a less costly amount of resources. Allocating more resources to the EF service does not significantly influence the performance of the concurrent best-effort traffic. Although the best-effort traffic receives a smaller amount of resources, it is allowed to use the bandwidth left unused by the EF aggregate in idle periods. Resources are actually allocated in preference to the EF aggregate and, if unused, they are made available to the best-effort traffic, thus enhancing network utilization.

Our assessment of the multimedia traffic support also comprises MPEG video streams. We investigated a variety of techniques that improve the delivery quality of MPEG video streams to end users. These techniques provide unequal protection of the different frame types that compose an MPEG video stream against indiscriminate packet losses within the network. We considered techniques such as adding redundancy either by applying a FEC scheme to tolerate losses or by changing the compression factor in a GOP pattern generation. We have also covered how to best use a differentiation scheme among different types of MPEG frames and how to adequately protect them from losses. Applying these different QoS strategies together constitutes a complex task. Therefore, we also analyzed the effects of the joint adoption of these techniques to better support the delivery quality of MPEG video streams. The developed analytical model allows the determination of bounds for performance as observed at the application layer. Results show that the combined adoption of a protection of the more important frames (from the viewpoint of the hierarchical MPEG structure), and an additional tolerance to losses through a FEC mechanism is an efficient strategy to improve delivery quality in video transmissions.

In the second part of this work, we addressed issues related to introducing location-awareness in networks. Networks conventionally rely on a structure composed by identifiers of the network components and routing procedures that are able to route information between these components. The addition of location-awareness introduces a new element that is the physical location of such network components. Novel location-aware applications and services may be envisaged if the physical locations of network components and users can be known. Nevertheless, to deploy these applications, a location service must be available to map the identifiers of the components onto the locations of the components. In this work, we investigated the location-awareness issue and the specific solutions to location services in different contexts such as mobile ad hoc networks and the Internet.

Considering mobile ad hoc networks, we proposed a TCP-Tailored approach to the location management. We first identified that the main proposals for location management in mobile ad hoc networks share two steps as a common characteristic: the location query for the position of the destination node and the flow transmission toward this node. Our proposed approach sends the first packet of a TCP flow (inherently a `SYN` packet) to learn the position of the flow's destination instead of adopting dedicated location query packets. The analysis we performed show that short-lived TCP flows specially benefit from the proposed scheme when compared with the adoption of a dedicated query packet. Short-lived TCP flows, such as web requests for instance, constitute a large portion of traffic on the current Internet. Supposing a similar traffic profile between the Internet and the future mobile ad hoc networks, the TCP-tailored approach benefits to the majority of flows in the location management procedure.

In the context of geographically locating Internet hosts, we concentrated in improving the accuracy of the location estimation inferred from delay measurements. The measurement-based geographic location of Internet hosts relies on a possible correlation between network delay and geographic distance. From our experiments, we showed that although this correlation is weak to moderate worldwide, it is much stronger in regions with rich connectivity. The robustness of this technique against factors that weaken the correlation between delay and distance comes from using multiple geographically distributed probe machines to perform the delay measurements as well as considering the minimum delay among several delay samples rather than an individual delay sample. We proposed a demographic placement policy to take into account the distribution of users to place landmarks and probe machines. Using our proposed demographic placement policy, a relatively small number of landmarks is able to represent a large portion of users within a limited coverage distance. Moreover, adopting fewer landmarks also imply a lower amount of measurement traffic in the network. These landmarks are placed on areas of high user density, thus providing closer landmarks and more accurate location estimations for most hosts to be located. Probe machines are placed on locations that have enough network infrastructure and sparsely distributed to avoid gathering redundant data. Different similarity models were also evaluated. The similarity model gives the landmark whose location should be used as location estimation for the target host to be located. We show that the choice of the similarity model is fundamental for the accuracy of the location estimation. Our findings show that strategically placing landmarks and probe machines as well as using an adequate similarity model are key points for the accuracy of the location estimation in the measurement-based geographic location of Internet hosts.

Future Work

The wide adoption of the IEEE 802.11 standard for wireless local area networks and the emerging of ad hoc networks lead to an increasing demand for QoS support on these networks. Scarce bandwidth and mobility add complexity to the problem of providing QoS in wireless networks as opposed to wired networks. The varying performance of the radio channel and the access method that equally shares the channel among all hosts also contribute to this complexity [216]. Changing some parameters of the medium access control (MAC) layer as a function of the required QoS has been the focus of some research efforts. Another approach is to couple mobility and QoS management. However, the support of QoS to deliver multimedia content in wireless networks remains an open research area.

Developing and enhancing techniques to geographically locate an entity in the Internet are the focus of some recent research. Nevertheless the implications of doing it, especially in regards to privacy and security, are not entirely investigated. Recently, the IETF created a new Working Group, GeoPriv (Geographic Location/Privacy) [217]. The objective of GeoPriv is to assess the authorization, integrity, and privacy requirements that must be met in order to transfer location information, or authorize the release or representation of such an information through an agent. Standardization on the format to express location information is also envisaged. The combination of these elements should provide a service capable of transferring geographic location information in a private and secure fashion.

Topologically-aware construction of overlay networks and server selection [218] adopts landmarks as references of proximity to group nodes that are relatively close to each another in terms of network delay. This binning strategy around some landmarks avoids unnecessary high latency hops, thus improving routing performance. Demographically placing the reference landmarks closer to where most users are, as proposed in this work, may then improve the performance of the binning strategy as more users would be closer to the reference landmarks.

The deployment of a location service based on measurements for the geographic location of Internet hosts may be envisaged using the NIMI (National Internet Measurement Infrastructure) infrastructure [215]. The NIMI boxes can act as our probe machines. In our initial experiments, we are using about 2300 landmarks that are sparsely distributed worldwide. These landmarks are extracted from the LibWeb [201] data set. We expect to use this infrastructure as a first step to provide a basis for new location-aware applications. The location-aware construction of overlay networks is an example of possible application served by the location service we intend.

8.1 Résumé du Chapitre

Dans la première partie de cette thèse, nous avons étudié les problèmes liés au support de Qualité de Service (QdS) pour le trafic de voix et de vidéo MPEG. Ces flux demandent des garanties de QdS qui sont rarement atteintes dans l'Internet actuel. Ainsi, l'étude des mécanismes capables de transporter du trafic multimédia devient importante.

Nous avons évalué le service offert au trafic de voix lorsqu'il est traité par le schéma *expedited forwarding* (EF) de l'architecture DiffServ. Les résultats montrent que l'augmentation de la fraction de service attribuée au agrégat EF n'affecte pas fortement le trafic concurrent appartenant au service "au mieux" (*best-effort*). Cette situation se prolonge tandis que le trafic "au mieux" utilise la bande passante qui a été laissée sans utilisation par le trafic EF pendant des périodes d'inactivité. Ensuite, nous évaluons d'après un modèle analytique l'impacte des pertes de trames sur la qualité des flux vidéo MPEG et sur le gaspillage de ressources du réseau. Les résultats offrent des bornes sur la qualité espérée par les utilisateurs finaux ainsi que des indices sur comment mieux profiter de l'adoption conjointe des schémas de QdS évalués. Un point clé est le besoin d'adoption des mécanismes de FEC à cause de leur capacité d'améliorer fortement la qualité de délivrance de la vidéo transmise à travers des réseaux présentant pertes de paquets.

Dans la deuxième partie de la thèse, nous formalisons le problème de localisation d'un point de vue général. Un service de localisation est conçu pour fournir la localisation d'un composant du réseau à partir de son identifiant. Après avoir présenté un état de l'art sur les services de localisation, nous avons étudié l'application de ces services à différents contextes tels que les réseaux ad hoc mobiles et l'Internet.

Pour la gestion de la localisation dans les réseaux ad hoc mobiles, nous proposons une approche qui bénéficie en particulier aux flux TCP. On démontre que l'approche proposée réduit les coûts pour les flux TCP par rapport à l'approche traditionnelle qui ne prend pas en compte des particularités des ces flux. Dans le contexte de l'Internet, nous nous concentrons sur l'amélioration de la précision de l'estimation de la localisation géographique d'un hôte basée sur des mesures. Nous proposons alors une politique de placement démographique pour placer les hôtes de référence et les machines sondes en prenant en compte la distribution des utilisateurs. Nous étudions aussi des modèles de similarité pour comparer les profils de délai et déterminer l'hôte de référence ayant le profil de délai le plus similaire à celui de l'hôte cible. Les résultats montrent que le placement stratégique des hôtes de référence et des machines sondes ainsi que le choix adéquat du modèle de similarité sont des points clés pour la précision de l'estimation de localisation. La construction des réseaux en surcouches (*overlay*) basés sur la localisation géographique des hôtes ainsi que les aspects liés à la confidentialité des informations de localisation restent des domaines ouverts de recherche.

Appendix A

Decodable Frames in a GOP

The expected number of correctly decodable frames in a GOP is calculated based on the probability for a frame (I, P, or B) to be directly lost and the impact of such a loss in the subsequent frames due to the hierarchical structure of MPEG encoding. The model of Equation 4.1 takes into account all the consequences of direct frame losses which impose indirect frame losses within a GOP, though the direct frame losses themselves are assumed to be mutually independent. We denote by $S(F)$ the probability that a frame F is considered decodable. Each term of Equation 4.1 refers to the effects of the direct loss of one particular type of frame within a given GOP_x based on the probability of correctly decoding the frame and by considering its dependencies:

1. The probability $S(I)$ that the I frame in GOP_x is decodable is

   $S(I) = 1 - \xi_I$.

   Thus, the expected number of correctly decodable I frames in GOP_x is

   $N_{\text{dec}}^{I} = 1 - \xi_I$.

2. The probabilities $S(P_1)$, $S(P_2)$, …, $S(P_{N_P})$ that each of the N_P frames of type P in GOP_x is decodable are

   $S(P_1) = (1 - \xi_I)(1 - \xi_P)$,

   $S(P_2) = (1 - \xi_I)(1 - \xi_P)^2$,

   $\vdots$

   $S(P_{N_P}) = (1 - \xi_I)(1 - \xi_P)^{N_P}$.

   The expected number of correctly decodable P frames in GOP_x is thus

   $N_{\text{dec}}^{P} = \sum_{j=1}^{N_P} S(P_j) = (1 - \xi_I) \sum_{j=1}^{N_P} (1 - \xi_P)^j$.

3. As consecutive B frames have the same dependency throughout the hierarchical structure, we consider consecutive B frames as composing a B group. Each GOP pattern has $(M - 1)$ B frames in each B group. We represent by B_n each of the B frames

that compose the n^{th} B group. The probabilities $S(B_1)$, $S(B_2)$, ..., $S(B_{\frac{N}{M}-1})$ that each B frame in each of the $\frac{N}{M}-1$ initial B groups is considered decodable are
$S(B_1) = (1-\xi_I)(1-\xi_P)(1-\xi_B)$,
$S(B_2) = (1-\xi_I)(1-\xi_P)^2(1-\xi_B)$,
$\vdots$
$S(B_{\frac{N}{M}-1}) = (1-\xi_I)(1-\xi_P)^{\frac{N}{M}-1}(1-\xi_B)$.
The last B group is also affected by the I frame of GOP_{x+1}. Thus,
$S(B_{\frac{N}{M}}) = (1-\xi_I)^2(1-\xi_P)^{\frac{N}{M}-1}(1-\xi_B)$.
As each B group has $(M-1)$ B frames and $N_P = \frac{N}{M}-1$, the expected number of correctly decodable B frames in GOP_x is

$$\begin{aligned}
N^B_{\text{dec}} &= (M-1)\sum_{j=1}^{\frac{N}{M}} S(B_j) \\
&= (M-1)(1-\xi_I)\sum_{j=1}^{N_P}(1-\xi_P)^j(1-\xi_B) \\
&+ (M-1)(1-\xi_I)^2(1-\xi_P)^{N_P}(1-\xi_B) \\
&= (M-1)(1-\xi_I)(1-\xi_B) \\
&\times \left((1-\xi_I)(1-\xi_P)^{N_P} + \sum_{j=1}^{N_P}(1-\xi_P)^j\right).
\end{aligned}$$

Therefore, the expected number of correctly decodable frames in a GOP is $N_{\text{dec}} = N^I_{\text{dec}} + N^P_{\text{dec}} + N^B_{\text{dec}}$, that is equivalent to Equation 4.1.

Publications

International Journals

[1] A. Ziviani, B. E. Wolfinger, J. F. de Rezende, O. C. M. B. Duarte, and S. Fdida, "Joint adoption of QoS schemes for MPEG streams," *Multimedia Tools and Applications, Kluwer Academic Publishers, ISSN: 1380-7501*, 2003. Accepted for publication.

[2] A. Ziviani, J. F. de Rezende, and O. C. M. B. Duarte, "Evaluating the expedited forwarding of voice traffic in a differentiated services network," *International Journal of Communication Systems, John Wiley and Sons, ISSN: 1074-5351*, vol. 15, no. 9, pp. 799–813, Nov. 2002.

International Conferences

[1] A. Ziviani, S. Fdida, J. F. de Rezende, and O. C. M. B. Duarte, "Demographic placement for Internet host location," in *Proc. of the IEEE GLOBECOM'2003*, (San Francisco, CA, USA), Dec. 2003.

[2] A. Ziviani, S. Fdida, J. F. de Rezende, and O. C. M. B. Duarte, "Similarity models for Internet host location," in *Proc. of the IEEE International Conference on Networks - ICON'2003*, (Sydney, Australia), pp. 81–86, Sept. 2003.

[3] A. Ziviani, S. Fdida, J. F. de Rezende, and O. C. M. B. Duarte, "Placement issues in measurement-based host location," in *Proc. of the 9th European Summer School and IFIP Workshop on Next Generation Networks - EUNICE 2003*, (Budapest - Balatonfüred, Hungary), pp. 173–180, Sept. 2003.

[4] A. Ziviani, S. Fdida, J. F. de Rezende, and O. C. M. B. Duarte, "Enhancing location management in mobile ad hoc networks," in *Proc. of the IFIP Mediterranean Ad Hoc Networking Workshop - Med-Hoc-Net 2002*, (Sardinia, Italy), Sept. 2002.

[5] A. Ziviani, B. E. Wolfinger, J. F. de Rezende, O. C. M. B. Duarte, and S. Fdida, "On the combined adoption of QoS schemes to improve the delivery quality of MPEG video streams," in *Proc. of the Intern. Symp. on Performance Evaluation of Computer and Telecommunications Systems - SPECTS'2002*, (San Diego, CA, USA), pp. 25–32, July 2002.

[6] A. Ziviani, J. F. de Rezende, O. C. M. B. Duarte, and S. Fdida, "Improving the delivery quality of MPEG video streams by using differentiated services," in *Proc. of the European Conference on Multiservice Networks – ECUMN'2002*, (Colmar, France), pp. 107–115, Apr. 2002.

[7] A. Ziviani, J. F. de Rezende, O. C. M. B. Duarte, and S. Fdida, "Evaluating voice traffic in a differentiated services environment," in *Proc. of the 17th International Teletraffic Congress – ITC17*, (Salvador, Brazil), pp. 907–918, Dec. 2001.

[8] A. Ziviani, J. F. de Rezende, and O. C. M. B. Duarte, "Towards a differentiated services support for voice traffic," in *Proc. of the IEEE GLOBECOM'99*, vol. 1, (Rio de Janeiro, Brazil), pp. 59–63, Dec. 1999.

French-spoken Conferences

[1] A. Ziviani, S. Fdida, J. F. de Rezende, and O. C. M. B. Duarte, "Une approche TCP pour la gestion de la localisation dans les réseaux ad hoc mobiles," in *Proc. of the Colloque Francophone sur la Gestion de Réseaux et de Services - GRES'2003*, (Fortaleza, Brazil), Feb. 2003.

[2] A. Ziviani, J. F. de Rezende, O. C. M. B. Duarte, and S. Fdida, "Services différenciés pour les flux MPEG," in *Proc. of the Journées Doctorales Informatique et Réseaux – JDIR'2002*, (Toulouse, France), pp. 51–60, Mar. 2002.

References

[1] Christopher Lefelhocz, Bryan Lyles, Scott Shenker, and Lixia Zhang, "Congestion control for best-effort service: Why we need a new paradigm," *IEEE Network*, vol. 10, no. 1, pp. 10–19, Jan. 1996.

[2] Xipeng Xiao and Lionel Ni, "Internet QoS: the big picture," *IEEE Network*, vol. 13, no. 2, pp. 1–13, Mar. 1999.

[3] Janusz Gozdecki, Andrzej Jajszczyk, and Rafal Stankiewicz, "Quality of service terminology in IP networks," *IEEE Communications Magazine*, vol. 41, no. 3, pp. 153–159, Mar. 2003.

[4] David D. Clark, "Adding service discrimination to the Internet," Tech. Rep., MIT, Sept. 1995.

[5] Scott Shenker, "Fundamental design issues for the future Internet," *IEEE Journal on Selected Areas in Communications*, vol. 13, no. 7, pp. 1176–1188, Sept. 1995.

[6] Eric C. Rosen, Arun Viswanathan, and Ross Callon, "Multiprotocol label switching architecture," *RFC 3031*, Jan. 2001.

[7] E. Crawley, R. Nair, B. Rajagopalan, and H. Sandick, "A framework for QoS-based routing in the Internet," *RFC 2386*, Aug. 1998.

[8] Robert Braden, David D. Clark, and Scott Shenker, "Integrated services in the Internet architecture: an overview," *RFC 1633*, June 1994.

[9] Kathleen Nichols, Steven Blake, Fred Baker, and David L. Black, "Definition of the differentiated services field (DS field) in the IPv4 and IPv6 headers," *RFC 2474*, Dec. 1998.

[10] Steven Blake, David L. Black, Mark Carlson, Elwyn Davies, Zheng Wang, and Walter Weiss, "An architecture for differentiated services," *RFC 2475*, Dec. 1998.

[11] Jeremy Lawrence, "Designing multiprotocol label switching networks," *IEEE Communications Magazine*, vol. 39, no. 7, pp. 134–142, July 2001.

[12] Lixia Zhang, Stephen E. Deering, Deborah Estrin, Scott Shenker, and Daniel Zappala, "RSVP: A new resource reservation protocol," *IEEE Network*, vol. 7, no. 5, pp. 8–18, Sept. 1993.

[13] Robert Braden, Lixia Zhang, Steve Berson, Shai Herzog, and Sugih Jamin, "Resource reservation protocol (RSVP) – version 1 functional specification," *RFC 2205*, Sept. 1997.

[14] John Wroclawski, "The use of RSVP with IETF integrated services," *RFC 2210*, Sept. 1997.

[15] Paul P. White, "RSVP and integrated services in the Internet: A tutorial," *IEEE Communications Magazine*, vol. 35, no. 5, pp. 100–106, May 1997.

[16] Loa Andersson, Paul Doolan, Nancy Feldman, Andre Fredette, and Bob Thomas, "LDP specification," *RFC 3036*, Jan. 2001.

[17] Jim Boyle, Vijay Gill, Alan Hannan, Dave Cooper, Daniel O. Awduche, Blaine Christian, and Wai Sum Lai, "Applicability statement for traffic engineering with MPLS," *RFC 3346*, Aug. 2002.

[18] Daniel O. Awduche, "MPLS and traffic engineering in IP networks," *IEEE Communications Magazine*, vol. 37, no. 12, pp. 42–47, Dec. 1999.

[19] George Swallow, "MPLS advantages for traffic engineering," *IEEE Communications Magazine*, vol. 37, no. 12, pp. 54–57, Dec. 1999.

[20] Eric C. Rosen and Yakov Rekhter, "BGP/MPLS VPNs," *RFC 2547*, Mar. 1999.

[21] Karthik Muthukrishnan and Andrew Malis, "A core MPLS IP VPN architecture," *RFC 2547*, Mar. 1999.

[22] George Apostolopoulos, Roch Guérin, Sanjay Kamat, Ariel Orda, Tony Przygienda, and Doug Williams, "QoS routing mechanisms and OSPF extensions," *RFC 2676*, Aug. 1999.

[23] Claudio E. Casetti, Renato Lo Cigno, Marco Mellia, Maurizio M. Munafò, and Zsóka Zoltán, "A new class of QoS routing strategies based on network graph reduction," in *Proc. of the IEEE INFOCOM'2002*, New York, NY, USA, June 2002.

[24] Luís Henrique Maciel Kosmalski Costa, Serge Fdida, and Otto Carlos Muniz Bandeira Duarte, "Developing scalable protocols for three-metric QoS routing," *Computer Networks*, vol. 39, no. 6, pp. 713–727, Aug. 2002.

[25] Fernando Kuipers, Piet Van Mieghem, Turgay Korkmaz, and Marwan Krunz, "An overview of constraint-based path selection algorithms for QoS routing," *IEEE Communications Magazine*, vol. 40, no. 12, pp. 50–55, Dec. 2002.

[26] Saurabh Srivastava, Balaji Krithikaivasan, V. Venkatachalam, Cory Beard, Deep Medhi, W. Alanqar, and A. Nagarajan, "A case study on evaluating the benefits of MPLS traffic engineering through constraint-based routing and network controls," in *Proc. of the IEEE International Conference on Communications - ICC'2002*, New York, NY, USA, Apr. 2002.

[27] Paul P. White and Jon Crowcroft, “The integrated services in the Internet: State of the art,” *Proceedings of the IEEE*, vol. 85, no. 12, pp. 1934–1946, Dec. 1997.

[28] John Wroclawski, “Specification of the controlled-load network element service,” *RFC 2211*, Sept. 1997.

[29] Scott Shenker, Craig Partridge, and Roch Guérin, “Specification of guaranteed quality of service,” *RFC 2212*, Sept. 1997.

[30] Abhay K. Parekh and Robert G. Gallager, “A generalized processor sharing approach to flow control in integrated services networks: The single node case,” *IEEE/ACM Transactions on Networking*, vol. 1, no. 3, pp. 344–357, June 1993.

[31] Abhay K. Parekh and Robert G. Gallager, “A generalized processor sharing approach to flow control in integrated services networks: The multiple node case,” *IEEE/ACM Transactions on Networking*, vol. 2, no. 2, pp. 137–150, Apr. 1994.

[32] Jean-Yves Le Boudec and Patrick Thiran, *Network Calculus*, Springer Verlag LNCS 2050, June 2001.

[33] Yoram Bernet, “The complementary roles of RSVP and differentiated services in the full-service QoS network,” *IEEE Communications Magazine*, vol. 38, no. 2, pp. 154–162, Feb. 2000.

[34] Fred Baker, Carol Iturralde, Francois Le Faucheur, and Bruce Davie, “Aggregation of RSVP for IPv4 and IPv6 reservations,” *RFC 3175*, Sept. 2001.

[35] Martin Karsten, “Experimental extensions to RSVP – remote client and one-pass signalling,” in *Proc. of the International Workshop on Quality of Service – IWQoS’2001*, Karlsruhe, Germany, June 2001.

[36] Huirong Fu and Edward W. Knightly, “Aggregation and scalable QoS: A performance study,” in *Proc. of the International Workshop on Quality of Service – IWQoS’2001*, Karlsruhe, Germany, June 2001.

[37] Jens Schmitt, Martin Karsten, and Ralf Steinmetz, “On the aggregation of deterministic service flows,” *Computer Communications*, vol. 24, no. 1, pp. 2–18, Jan. 2001.

[38] A. Mankin, Fred Baker, Robert Braden, S. Bradner, M. O’Dell, A. Romanow, A. Weinrib, and Lixia Zhang, “Resource reservation protocol (RSVP) – version 1 applicability statement,” *RFC 2208*, Sept. 1997.

[39] Brian Carpenter and Kathleen Nichols, “Differentiated services in the Internet,” *Proceedings of the IEEE*, vol. 90, no. 9, pp. 1479–1494, Sept. 2002.

[40] R. Guérin, L. Li, S. Nadas, P. Pan, and V. Peris, “The cost of QoS support in edge devices: An experimental study,” in *Proc. of the IEEE INFOCOM’99*, New York, NY, USA, Mar. 1999.

[41] Van Jacobson, Kathleen Nichols, and Kedarnath Poduri, "An expedited forwarding PHB," *RFC 2598*, June 1999.

[42] Bruce Davie, Anna Charny, Jon Bennett, Kent Benson, Jean-Yves Le Boudec, Bill Courtney, Shahram Davari, Victor Firoiu, and Dimitrios Stiliadis, "An expedited forwarding PHB (per-hop behavior)," *RFC 3246*, Mar. 2002.

[43] Juha Heinanen, Fred Baker, Walter Weiss, and John Wroclawski, "Assured forwarding PHB group," *RFC 2597*, June 1999.

[44] Anna Charny, Jon Bennett, Kent Benson, Jean-Yves Le Boudec, Angela Chiu, Bill Courtney, Shahram Davari, Victor Firoiu, Charles Kalmanek, and K. K. Ramakrishnan, "Supplemental information for the new definition of the EF PHB (expedited forwarding per-hop behavior)," *RFC 3247*, Mar. 2002.

[45] Kathleen Nichols, Van Jacobson, and Lixia Zhang, "A two-bit differentiated services architecture for the Internet," *RFC 2638*, July 1999.

[46] M. Katevenis, S. Sidiropoulos, and C. Courcoubetis, "Weighted round-robin cell multiplexing in a general-purpose atm switch chip," *IEEE Journal on Selected Areas in Communications*, vol. 9, no. 8, pp. 1265–1279, Oct. 1991.

[47] Anna Charny and J. Y. Le Boudec, "Delay bounds in a network with aggregate scheduling," in *Proc. of the International Workshop on Quality of future Internet Services - QofIS'00*, Berlin, Germany, Sept. 2000.

[48] Zhi-Li Zhang, Zhenhai Duan, Lixin Gao, and Yiwei Thomas Hou, "Decoupling QoS control from core routers: A novel bandwidth broker architecture for scalable support of guaranteed services," in *Proc. of the ACM SIGCOMM'2000*, Stockholm, Sweden, Aug. 2000.

[49] Jon C. R. Bennett, Kent Benson, Anna Charny, William F. Courtney, and J. Y. Le Boudec, "Delay jitter bounds and packet scale rate guarantee for expedited forwarding," in *Proc. of the IEEE INFOCOM'2001*, Anchorage, AK, USA, Apr. 2001.

[50] S. Floyd and V. Jacobson, "Random early detection gateways for congestion avoidance," *IEEE/ACM Transactions on Networking*, vol. 1, no. 4, pp. 397–413, Aug. 1993.

[51] Bob Braden, David D. Clark, Jon Crowcroft, Bruce Davie, Steve Deering, Deborah Estrin, Sally Floyd, Van Jacobson, Greg Minshall, Craig Partridge, Larry Peterson, K. K. Ramakrishnan, Scott Shenker, John Wroclawski, and Lixia Zhang, "Recommendations on queue management and congestion avoidance in the Internet," *RFC 2309*, Apr. 1998.

[52] David D. Clark and Wenjia Fang, "Explicit allocation of best effort packet delivery service," *IEEE/ACM Transactions on Networking*, vol. 6, no. 4, pp. 362–373, Aug. 1998.

[53] Juha Heinanen and Roch Guérin, "A single rate three color marker," *RFC 2697*, Sept. 1999.

[54] Juha Heinanen and Roch Guérin, "A two rate three color marker," *RFC 2698*, Sept. 1999.

[55] Mohamed A. El-Gendy and Kang G. Shin, "Equation-based packet marking for assured forwarding services," in *Proc. of the IEEE INFOCOM'2002*, New York, NY, USA, June 2002.

[56] Ikjun Yeom and Narasimha Reddy, "Impact of marking strategy on aggregated flows in differentiated services network," in *Proc. of the International Workshop on Quality of Service - IWQoS'99*, London, England, June 1999.

[57] Ikjun Yeom and A. L. Narasimha Reddy, "Adaptive marking for assured forwarding service," *IEEE Communications Letters*, vol. 6, no. 10, pp. 461–463, Oct. 2002.

[58] Biswajit Nandy, Jeremy Ethridge, Abderrahmane Lakas, and Alan Chapman, "Aggregate flow control: Improving assurances for differentiated services network," in *Proc. of the IEEE INFOCOM'2001*, Anchorage, AK, USA, Apr. 2001.

[59] Nicolas Christin, Jorg Liebeherr, and Tarek Abdelzaher, "A quantitative assured forwarding service," in *Proc. of the IEEE INFOCOM'2002*, New York, NY, USA, June 2002.

[60] Aimin Sang, Haifeng Zhu, and San-qi Li, "Weighted fairness guarantee for scalable differentiated services (DiffServ)," *Computer Communications*, vol. 24, no. 1, pp. 51–63, Jan. 2001.

[61] Andrew Odlyzko, "Paris metro pricing: The minimalist differentiated services solution," in *Proc. of the International Workshop on Quality of Service - IWQoS'99*, London, England, June 1999.

[62] Paul Hurley, Jaen-Yves Le Boudec, Patrick Thiran, and Mourad Kara, "ABE: Providing a low-delay service within best effort," *IEEE Network*, vol. 15, no. 3, pp. 60–69, May 2001.

[63] Y. Moret, S. Fdida, C. Galand, and A. Fichou, *Evaluation of the ER Algorithm ER-AQLES in Different ABR Environments*, in ATM Networks: Performance Modelling and Analysis, Chapman and Hall, 1998.

[64] Constantinos Dovrolis and Parameswaran Ramanathan, "A case for relative differentiation services and the proportional differentiation model," *IEEE Network*, vol. 13, no. 5, Sept. 1999.

[65] Constantinos Dovrolis, Dimitrios Stiliadis, and Parameswaran Ramanathan, "Proportional differentiated services: Delay differentiation and packet scheduling," *IEEE/ACM Transactions on Networking*, vol. 10, no. 1, pp. 12–26, Feb. 2002.

[66] Constantinos Dovrolis and Parameswaran Ramanathan, "Dynamic class selection and class provisioning in proportional differentiated services," *Computer Communications*, vol. 26, no. 3, pp. 204–221, Feb. 2003.

[67] Constantinos Dovrolis, Dimitrios Stiliadis, and Parameswaran Ramanathan, "Proportional differentiated services: Delay differentiation and packet scheduling," in *Proc. of the ACM SIGCOMM'99*, Cambridge, MA, USA, Sept. 1999.

[68] Constantinos Dovrolis and Parameswaran Ramanathan, "Proportional differentiated services, part II: Loss rate differentiation and packet dropping," in *Proc. of the International Workshop on Quality of Service - IWQoS'2000*, Pittsburgh, PA, USA, June 2000.

[69] Jong-Tae Park, Jong-Wook Baek, and James Won-Ki Hong, "Management of service level agreements for multimedia Internet service using a utility model," *IEEE Communications Magazine*, vol. 39, no. 5, pp. 100–106, May 2001.

[70] Jim Martin and Arne Nilsson, "On service level agreements for IP networks," in *Proc. of the IEEE INFOCOM'2002*, New York, NY, USA, June 2002.

[71] Clarence Filsfils and John Evans, "Engineering a multiservice IP backbone to support tight SLAs," *Computer Networks*, vol. 40, no. 1, pp. 131–148, Sept. 2002.

[72] Dan Grossman, "New terminology and clarifications for DiffServ," *RFC 3260*, Apr. 2002.

[73] Andrea Westerinen, John Schnizlein, John Strassner, Mark Scherling, Bob Quinn, Jay Perry, Shai Herzog, An-Ni Huynh, Mark Carlson, and Steve Waldbusser, "Terminology for policy-based management," *RFC 3198*, Nov. 2001.

[74] Rares Serban, Chadi Barakat, and Walid Dabbous, "Dynamic resource allocation in core routers of a DiffServ network," in *Proc. of the Asian Computing Science Conference - ASIAN 2002*, Hanoi, Vietnam, Dec. 2002.

[75] Laurent Mathy, Christopher Edwards, and David Hutchison, "The Internet: A global telecommunications solution?," *IEEE Network*, vol. 14, no. 4, pp. 46–57, July 2000.

[76] Bernd E. Wolfinger and Martin Zaddach, "Techniques to improve quality-of-service in video communications via best-effort networks," in *Proc. of the International Conference on Networking - ICN'2001*, Colmar, France, July 2001.

[77] Liren Zhang, Li Zheng, and Koh Soo Ngee, "Effect of delay and jitter on voice/video over IP," *Computer Communications*, vol. 25, no. 9, pp. 863–873, June 2002.

[78] Henning Schulzrinne, "Converging on Internet telephony," *IEEE Internet Computing*, vol. 3, no. 3, pp. 40–43, May 1999.

[79] Pawan Goyal, Albert Greenberg, Charles R. Kalmanek, Willian T. Marshall, Partho Mishra, Doug Nortz, and K. K. Ramakrishnan, "Integration of call signaling and resource management for IP telephony," *IEEE Internet Computing*, vol. 3, no. 3, pp. 44–52, May 1999.

[80] Bo Li, Mounir Hamdi, Dongyi Jiang, Xi-Ren Cao, and Yiwei Thomas Hou, "QoS-enabled voice support in the next-generation Internet: Issues, existing approaches and challenges," *IEEE Communications Magazine*, vol. 38, no. 4, pp. 54–61, Apr. 2000.

[81] Mahbub Hassan, Alfandika Nayandoro, and Mohammed Atiquzzaman, "Internet telephony: Services, technical challenges, and products," *IEEE Communications Magazine*, vol. 38, no. 4, pp. 96–103, Apr. 2000.

[82] Maarten J.C. Buchli, Danny De Vleeschauwer, Jan Janssen, Annelies Van Moffaert, and Guido Petit, "Resource allocation and management in DiffServ networks for IP telephony," in *Proc. of the International Workshop on Network and Operating Systems Support for Digital Audio and Video – NOSSDAV'2001*, Port Jefferson, NY, USA, June 2001.

[83] Artur Ziviani, José Ferreira de Rezende, and Otto Carlos Muniz Bandeira Duarte, "Evaluating the expedited forwarding of voice traffic in a differentiated services network," *International Journal of Communication Systems, John Wiley and Sons, ISSN: 1074-5351*, vol. 15, no. 9, pp. 799–813, Nov. 2002.

[84] N. Semret, R. R.-F. Liao, A. T. Campbell, and A. A. Lazar, "Peering and provisioning of differentiated Internet services," in *Proc. of the IEEE INFOCOM'2000*, Tel-Aviv, Israel, Mar. 2000.

[85] Ulrich Fiedler, Polly Huang, and Bernhard Plattner, "Towards provisioning Diff-Serv intra-nets," in *Proc. of the International Workshop on Quality of Service – IWQoS'2001*, Karlsruhe, Germany, June 2001.

[86] Thyagarajan Nandagopal, Narayanan Venkitaraman, Raghupathy Sivakumar, and Vaduvur Bharghavan, "Delay differentiation and adaptation in core stateless networks," in *Proc. of the IEEE INFOCOM'2000*, Tel-Aviv, Israel, Mar. 2000.

[87] Aaron Striegel and Govindarasu Manimaran, "Packet scheduling with delay and loss differentiation," *Computer Communications*, vol. 25, no. 1, pp. 21–31, Jan. 2002.

[88] Fang Hao, Ellen Zegura, and Mostafa Ammar, "Supporting server selection in differentiated service networks," in *Proc. of the IEEE INFOCOM'2001*, Anchorage, AK, USA, Apr. 2001.

[89] Giuseppe Bianchi, Nicola Blefari-Melazzi, and Mauro Femminella, "Per-flow QoS support over a stateless differentiated services IP domain," *Computer Networks*, vol. 40, no. 1, pp. 73–87, Sept. 2002.

[90] Aaron Striegel and Govindarasu Manimaran, "A scalable approach for DiffServ multicasting," in *Proc. of the IEEE International Conference on Communications - ICC'2001*, Helsinki, Finland, June 2001.

[91] Aaron Striegel and Govindarasu Manimaran, "A survey of QoS multicasting issues," *IEEE Communications Magazine*, vol. 40, no. 6, pp. 82–87, June 2002.

[92] Xin Wang and Henning Schulzrinne, "Pricing network resources for adaptive applications in a differentiated services network," in *Proc. of the IEEE INFOCOM'2001*, Anchorage, AK, USA, Apr. 2001.

[93] J. Antonio García-Macías, Franck Rousseau, Gilles Berger-Sabbatel, Leyla Toumi, and Andrzej Duda, "Quality of service and mobility for the wireless Internet," *Wireless Networks*, vol. 9, no. 4, pp. 341–352, July 2003.

[94] Van Jacobson, Kathleen Nichols, and Kedar Poduri, "The virtual wire behavior aggregate," *Internet Draft*, Mar. 2000, `draft-ietf-diffserv-ba-vw-00`.

[95] Srinivasan Keshav, *An Engineering Approach to Computer Networking*, Addison-Wesley, 1997.

[96] Mansour J. Karam and Fouad A. Tobagi, "Analysis of the delay and jitter of voice traffic over the Internet," in *Proc. of the IEEE INFOCOM'2001*, Anchorage, AK, USA, Apr. 2001.

[97] Artur Ziviani, "Voz sobre serviços diferenciados na Internet (Voice over differentiated services in the Internet," M.S. thesis, Programa de Engenharia Elétrica - COPPE/UFRJ, Sept. 1999.

[98] Artur Ziviani, José Ferreira de Rezende, and Otto Carlos Muniz Bandeira Duarte, "Towards a differentiated services support for voice traffic," in *Proc. of the IEEE Global Communications Conference - GLOBECOM'99*, Rio de Janeiro, Brazil, Dec. 1999, vol. 1, pp. 59–63.

[99] Artur Ziviani, José Ferreira de Rezende, Otto Carlos Muniz Bandeira Duarte, and Serge Fdida, "Evaluating voice traffic in a differentiated services environment," in *Proc. of the 17th International Teletraffic Congress – ITC17*, Salvador, Brazil, Dec. 2001, pp. 907–918.

[100] Difu Su, Jaideep Srivastava, and Jey-Hsin Yao, "Investigating factors influencing QoS of Internet phone," in *Proc. of the IEEE International Conference on Multimedia Computing and Systems – ICMCS'99*, Florence, Italy, June 1999.

[101] Catherine Boutremans and Jean-Yves Le Boudec, "Adaptive joint playout buffer and FEC adjustment for Internet telephony," in *Proc. of the IEEE INFOCOM'2003*, San Francisco, CA, USA, Apr. 2003.

[102] ITU-T, "One-way transmission time," Recommendation G.114, Mar. 1993.

[103] Guang-Liang Li, Bo Li, Jun-Hong Cui, and Fang-Ming Li, "Transient QoS guarantee for packet voice," in *Proc. of the IEEE Global Communications Conference - GLOBECOM'98*, Sydney, Australia, Nov. 1998.

[104] Kevin Fall and Kannan Varadhan, *The ns Manual*, UC Berkeley, LBL, USC/ISI, and Xerox PARC, June 2003, Available at `http://www.isi.edu/nsnam/ns/`.

[105] Mischa Schwartz, *Broadband Integrated Networks*, Prentice Hall, 1996.

[106] Hassan Naser, Alberto Leon-Garcia, and Osama Aboul-Magd, "Voice over differentiated services," *Internet Draft*, Dec. 1998, `draft-naser-voice-diffserv-eval-00`.

[107] Kimberly Claffy, Greg Miller, and Kevin Thompson, "The nature of the beast: Recent traffic measurements from an Internet backbone," in *Proc. of the Internet Global Summit - INET'98*, Geneva, Switzerland, July 1998.

[108] Henning Schulzrinne, Stephen L. Casner, Ron Frederick, and Van Jacobson, "RTP: A transport protocol for real-time applications," *RFC 1889*, Jan. 1996.

[109] Nortel Networks, "Packet policing on the passport 8600," *White Paper*, 2000.

[110] Cisco Systems, "AutoQoS for voice over IP (VoIP)," *White Paper*, 2003.

[111] Olivier Verscheure, Xavier Garcia, Gunnar Karlsson, and Jean-Pierre Hubaux, "User-oriented QoS in packet video delivery," *IEEE Network*, vol. 12, no. 6, pp. 12–21, Nov. 1998.

[112] Rod Green, S. I. Woolley, N. W. Garnham, and K. P. Jones, "Experimental testbed results for broadband residential video service QoS management," in *Proc. of the IEEE International Conference on Communications - ICC'2002*, New York, NY, USA, Apr. 2002.

[113] Rob Koenen, "MPEG-4: Multimedia for our time," *IEEE Spectrum*, vol. 36, no. 2, pp. 26–33, Feb. 1999.

[114] Jill M. Boyce and Robert D. Gaglianello, "Packet loss effects on MPEG video sent over the public Internet," in *Proc. of the ACM Multimedia 98*, Bristol, England, Sept. 1998.

[115] Jim Wu and Mahbub Hassan, "The issue of useless packet transmission for multimedia over the Internet," *Computer Communications*, vol. 26, no. 12, pp. 1240–1254, July 2003.

[116] Marwan Krunz, "Bandwidth allocation strategies for transporting variable-bit-rate video traffic," *IEEE Communications Magazine*, vol. 37, no. 1, pp. 40–46, Jan. 1999.

[117] Artur Ziviani, Bernd E. Wolfinger, José Ferreira de Rezende, Otto Carlos Muniz Bandeira Duarte, and Serge Fdida, "On the combined adoption of QoS schemes to improve

the delivery quality of MPEG video streams," in *Proc. of the Intern. Symp. on Performance Evaluation of Computer and Telecommunications Systems - SPECTS'2002*, San Diego, CA, USA, July 2002, pp. 25–32.

[118] Artur Ziviani, Bernd E. Wolfinger, José Ferreira de Rezende, Otto Carlos Muniz Bandeira Duarte, and Serge Fdida, "Joint adoption of QoS schemes for MPEG streams," *Multimedia Tools and Applications, Kluwer Academic Publishers, ISSN: 1380-7501*, 2003, Accepted for publication.

[119] Dirk Farin, Niels Mache, and Peter H. N. de With, "SAMPEG: a scene adaptive parallel MPEG-2 software encoder," in *Proc. of the SPIE Visual Communications and Image Processing - VCIP'01*, San Jose, CA, USA, Jan. 2001.

[120] Roger Crooks, "An analysis of MPEG encoding techniques on picture quality," *White Paper – Tektronix*, 1998.

[121] Deepak S. Turaga and Tsuhan Chen, "Hierarchical modeling of variable bit rate video sources," in *Proc. of the 11th Packet Video Workshop - PV'2001*, Kyongju, Korea, Apr. 2001.

[122] Marcelo Dias de Amorim, Otto Carlos Muniz Bandeira Duarte, and Guy Pujolle, "An improved MPEG behavioral analysis with autonomous parameter-extracting algorithm for strict video applications," in *Proc. of the IEEE International Conference on Communications - ICC'2000*, New Orleans, LA, USA, June 2000.

[123] M. F. Alam, Mohammed Atiquzzaman, and M. A. Karim, "Traffic shaping for MPEG video transmission over the next generation Internet," *Computer Communications*, vol. 23, no. 14–15, pp. 1336–1348, Aug. 2000.

[124] Michael Hemy, Urs Hengartner, Peter Steenkiste, and Thomas Gross, "MPEG systems streams in best-effort networks," in *Proc. of the 9th Packet Video Workshop - PV'1999*, New York, NY, USA, Apr. 1999.

[125] Jitae Shin, Jong-Won Kim, and C.-C. Jay Kuo, "Content-based packet video forwarding mechanism in differentiated services networks," in *Proc. of the 10th Packet Video Workshop - PV'2000*, Cagliari, Italy, May 2000.

[126] Yevgeni Koucheryavy, Dmitri Moltchanov, and J. Harju, "A top-down approach to VoD traffic transmission over DiffServ domain using the AF PHB class," in *Proc. of the IEEE International Conference on Communications - ICC'2003*, Anchorage, AK, USA, May 2003.

[127] Jitae Shin, Jong-Won Kim, Daniel C. Lee, and C.-C. Jay Kuo, "Adaptive packet forwarding for relative differentiated service and categorized packet video," in *Proc. of the IEEE International Conference on Communications - ICC'2001*, Helsinki, Finland, June 2001.

[128] Jitae Shin, J. Kim, and C.-C. J.Kuo, "Quality of service mapping mechanism for packet video in differentiated services network," *IEEE Transactions on Multimedia*, vol. 3, no. 2, pp. 219–231, June 2001.

[129] Fabio D'Agostino, Enrico Masala, Laura Farinetti, and Juan Carlos De Martin, "A simulative study of analysis-by-synthesis perceptual video classification and transmission over DiffServ IP networks," in *Proc. of the IEEE International Conference on Communications - ICC'2003*, Anchorage, AK, USA, May 2003.

[130] Pascal Frossard and Olivier Verscheure, "Joint source/FEC rate selection for quality-optimal MPEG-2 video delivery," *IEEE Transactions on Image Processing*, vol. 10, no. 12, pp. 1815–1825, Dec. 2001.

[131] Andres Albanese, Johannes Blömer, Jeff Edmonds, Michael Luby, and Madhu Sudan, "Priority encoding transmission," *IEEE Transactions on Information Theory*, vol. 42, no. 6, pp. 1737–1744, Nov. 1996.

[132] Bernd E. Wolfinger, "On the potential of FEC algorithms in building fault-tolerant distributed applications to support high QoS video communications," in *Proc. of the ACM Symp. on Principles of Distributed Computing - PODC'97*, Santa Barbara, CA, USA, Aug. 1997.

[133] Kihyuno Park and W. Wang, "QoS-sensitive transport of real-time MPEG video using adaptive redundancy control," *Computer Communications*, vol. 24, no. 1, pp. 78–92, Jan. 2001.

[134] Artur Ziviani, José Ferreira de Rezende, Otto Carlos Muniz Bandeira Duarte, and Serge Fdida, "Improving the delivery quality of MPEG video streams by using differentiated services," in *Proc. of the European Conference on Multiservice Networks – ECUMN'2002*, Colmar, France, Apr. 2002, pp. 107–115.

[135] Artur Ziviani, José Ferreira de Rezende, Otto Carlos Muniz Bandeira Duarte, and Serge Fdida, "Services différenciés pour les flux MPEG," in *Proc. of the Journées Doctorales Informatique et Réseaux – JDIR'2002*, Toulouse, France, Mar. 2002, pp. 51–60.

[136] Marwan Krunz and Satish K. Tripathi, "On the characterization of VBR MPEG streams," in *Proc. of the ACM SIGMETRICS'97*, Seattle, WA, USA, May 1997.

[137] K. Heidtmann, J. Kerse, T. Suchanek, Bernd E. Wolfinger, and Martin Zaddach, "Fehlertolerante Videokommunikation über verlustbehaftete Paketvermittlungsnetze," in *Proc. of the KiVS'2001*, Springer-Verlag, Ed., Feb. 2001.

[138] Frank H. P. Fitzek and Martin Reisslein, "MPEG-4 and H.263 video traces for network performance evaluation," *IEEE Network*, vol. 15, no. 6, pp. 40–54, Nov. 2001.

[139] Rainer Storn, "Modeling and optimization of PET-redundancy assignment for MPEG sequences," Tech. Rep. TR-95-018, International Computer Science Institute, Berkeley, May 1995.

[140] Per Enge and Pratap Misra, "Special issue on global positioning system," *Proceedings of the IEEE*, vol. 87, no. 1, pp. 3–15, Jan. 1999.

[141] The European Union, *GALILEO – European Satellite Navigation System*, `http://europa.eu.int/comm/dgs/energy_transport/galileo/`.

[142] Andy Harter and Andy Hopper, "A distributed location system for the active office," *IEEE Network*, vol. 8, no. 1, pp. 62–70, Jan. 1994.

[143] Paramvir Bahl and Venkata N. Padmanabhan, "RADAR: An in-building RF-based user location and tracking system," in *Proc. of the IEEE INFOCOM'2000*, Tel-Aviv, Israel, Mar. 2000.

[144] Jinyang Li, John Jannotti, Douglas S. J. de Couto, David R. Karger, and Robert Morris, "A scalable location service for geographic ad hoc routing," in *Proc. of the IEEE/ACM Mobicom'00*, Boston, MA, USA, Aug. 2000, pp. 120–130.

[145] Brad Karp and H. T. Kung, "GPSR: Greedy perimeter stateless routing for wireless networks," in *Proc. of the IEEE/ACM Mobicom'00*, Boston, MA, USA, Aug. 2000.

[146] Rahul Jain, Anuj Puri, and Raja Sengupta, "Geographical routing using partial information for wireless ad hoc networks," *IEEE Personal Communications*, vol. 8, no. 1, pp. 48–57, Feb. 2001.

[147] Tracy Camp, Jeff Boleng, Brad Williams, Lucas Wilcox, and William Navidi, "Performance comparison of two location based routing protocols for ad hoc networks," in *Proc. of the IEEE INFOCOM'2002*, New York, NY, USA, June 2002.

[148] Ivan Stojmenovic, "Position-based routing in ad-hoc networks," *IEEE Communications Magazine*, vol. 40, no. 7, pp. 128–134, July 2002.

[149] Young Bae Ko and Nitin H. Vaidya, "Flooding-based geocasting protocols for mobile ad hoc networks," *Mobile Networks and Applications*, vol. 7, no. 6, pp. 471–480, Dec. 2002.

[150] Moe Rahnema, "Overview of the GSM system and protocol architecture," *IEEE Communications Magazine*, vol. 31, no. 4, pp. 92–100, Apr. 1993.

[151] Yu-Chee Tseng, Shih-Lin Wu, Wen-Hwa Liao, and Chih-Min Chao, "Location awareness in ad hoc wireless mobile networks," *IEEE Computer*, vol. 34, no. 6, pp. 46–52, June 2001.

[152] Stefano Basagni, Imrich Chlamtac, Violet R. Syrotiuk, and Barry A. Woodward, "A distance routing effet algorithm for mobility (DREAM)," in *Proc. of the IEEE/ACM Mobicom'98*, Dallas, TX, USA, Oct. 1998, pp. 76–84.

[153] Zygmunt Haas and Ben Liang, "Ad-hoc mobility management with uniform quorum systems," *IEEE/ACM Transactions on Networking*, vol. 7, no. 2, pp. 228–240, Apr. 1999.

[154] Ivan Stojmenovic, *Location updates for efficient routing in ad hoc wireless networks*, in Handbook of Wireless Networks and Mobile Computing, pp. 451–471, John Wiley & Sons, Feb. 2002.

[155] Seung-Chul M. Woo and Suresh Singh, "Scalable routing protocol for ad hoc networks," *Wireless Networks*, vol. 7, no. 5, pp. 513–529, Sept. 2001.

[156] Aline Carneiro Viana, Marcelo Dias de Amorim, Serge Fdida, and José Ferreira de Rezende, "Indirect routing using distributed location information," in *Proc. of the IEEE International Conference on Pervasive Computing and Communications - PerCom 2003*, Dallas-Fort Worth, TX, USA, Mar. 2003.

[157] Yuan Xue, Baochun Li, and Klara Nahrstedt, "A scalable location management scheme in mobile ad-hoc networks," in *Proc. of the IEEE Conference on Local Computer Networks - LCN'2001*, Tampa, FL, USA, Nov. 2001.

[158] Martin Mauve, Jörg Widmer, and Hannes Harstenstein, "A survey on position-based routing in mobile ad-hoc networks," *IEEE Network*, vol. 15, no. 6, pp. 30–39, Nov. 2001.

[159] Christine T. Cheng, Howard L. Lemberg, Sumesh J. Philip, Eric van der Berg, and Tao Zhang, "SLALoM: A scalable location management scheme for large mobile ad-hoc networks," in *Proc. of the IEEE WCNC'2002*, Orlando, FL, USA, Mar. 2002.

[160] Christopher Davis, Paul Vixie, Tim Goowin, and Ian Dickinson, "A means for expressing location information in the domain name system," *Internet RFC 1876*, Jan. 1996.

[161] Bradley Huffaker, Marina Fomenkov, Daniel J. Plummer, David Moore, and k claffy, "Distance metrics in the Internet," in *Proc. of the IEEE International Telecommunications Symposium - ITS'2002*, Natal, Brazil, Sept. 2002.

[162] Anukool Lakhina, John W. Byers, Mark Crovella, and Ibrahim Matta, "On the geographic location of Internet resources," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 6, pp. 934–948, Aug. 2003.

[163] University of Illinois at Urbana-Champaign, *IP Address to Latitude/Longitude*, `http://cello.cs.uiuc.edu/cgi-bin/slamm/ip2ll/`.

[164] David Moore, Ram Periakaruppan, Jim Donohoe, and Kimberly Claffy, "Where in the world is netgeo.caida.org?," in *Proc. of the INET'2000*, Yokohama, Japan, July 2000.

[165] Venkata N. Padmanabhan and Lakshminarayanan Subramanian, "An investigation of geographic mapping techniques for Internet hosts," in *Proc. of the ACM SIGCOMM'2001*, San Diego, CA, USA, Aug. 2001.

[166] Visualware Inc., *VisualRoute*, `http://www.visualware.com/visualroute/`.

[167] CAIDA, *GTrace*, `http://www.caida.org/tools/visualization/gtrace/`.

[168] Gerco Ballintijn, Maarten van Steen, and Andrew S. Tanenbaum, "Characterizing Internet performance to support wide-area application development," *Operating Systems Review*, vol. 34, no. 4, pp. 41–47, Oct. 2000.

[169] Paul Francis, Sugih Jamin, Cheng Jin, Yixin Jin, Danny Raz, Yuval Shavitt, and Lixia Zhang, "IDMaps: A global Internet host distance estimation service," *IEEE/ACM Transactions on Networking*, vol. 9, no. 5, pp. 525–540, Oct. 2001.

[170] T. S. Eugene Ng and Hui Zhang, "Predicting Internet network distance with coordinates-based approaches," in *Proc. of the IEEE INFOCOM'2002*, New York, NY, USA, June 2002.

[171] Krishna P. Gummadi, Stefan Saroiu, and Steven D. Gribble, "King: Estimating latency between arbitrary Internet end hosts," in *ACM Internet Measurement Workshop 2002*, Marseille, France, Nov. 2002.

[172] Marcelo Pias, Jon Crowcroft, Steve Wilbur, Tim Harris, and Saleem Bhatti, "Lighthouses for scalable distributed location," in *Proc. of the Second International Workshop on Peer-to-Peer Systems - IPTPS'03*, Berkeley, CA, USA, Feb. 2003.

[173] Yuval Shavitt and Tomer Tankel, "Big-bang simulation for embedding network distances in Euclidean space," in *Proc. of the IEEE INFOCOM'2003*, San Francisco, CA, USA, Mar. 2003.

[174] Lying Tang and Mark Crovella, "Virtual landmarks for the Internet," in *ACM Internet Measurement Workshop 2003*, Miami, FL, USA, Oct. 2003.

[175] Hyuk Lim, Jennifer C. Hou, and Chong-Ho Choi, "Constructing Internet coordinate system based on delay measurement," in *ACM Internet Measurement Workshop 2003*, Miami, FL, USA, Oct. 2003.

[176] Young-Bae Ko and Nitin H. Vaidya, "Location-aided routing (LAR) in mobile ad hoc networks," in *Proc. of the IEEE/ACM Mobicom'98*, Dallas, TX, USA, Oct. 1998.

[177] Zygmunt Haas, Joseph Y. Halpern, and Li Li, "Gossip-based ad hoc routing," in *Proc. of the IEEE INFOCOM'2002*, New York, NY, USA, June 2002.

[178] Robert Castañeda, Samir R. Das, and Mahesh K. Marina, "Query localization techniques for on-demand routing protocols in ad hoc networks," *Wireless Networks*, vol. 8, no. 2/3, pp. 137–151, Mar./May 2002.

[179] Jeffrey Hightower and Gaetano Borriello, "Location systems for ubiquitous computing," *IEEE Computer*, vol. 38, no. 8, pp. 57–66, Aug. 2001.

[180] Srdan Capkun, Maher Hamdi, and Jean-Pierre Hubaux, "GPS-free positioning in mobile ad-hoc networks," *Journal of Cluster Computing*, vol. 5, no. 2, Apr. 2002.

[181] Douglas S. J. de Couto and Robert Morris, "Location proxies and intermediate node forwarding for practical geographic forwarding," Tech. Rep. MIT-LCS-TR-824, MIT Laboratory for Computer Science, June 2001.

[182] Tracy Camp, Jeff Boleng, and Lucas Wilcox, "Location information services in mobile ad hoc networks," in *Proc. of the IEEE International Conference on Communications - ICC'2002*, New York, NY, USA, Apr. 2002.

[183] Artur Ziviani, Serge Fdida, José Ferreira de Rezende, and Otto Carlos Muniz Bandeira Duarte, "Enhancing location management in mobile ad hoc networks," in *Proc. of the IFIP Mediterranean Ad Hoc Networking Workshop - Med-Hoc-Net 2002*, Sardinia, Italy, Sept. 2002.

[184] Artur Ziviani, Serge Fdida, José Ferreira de Rezende, and Otto Carlos Muniz Bandeira Duarte, "Une approche TCP pour la gestion de la localisation dans les réseaux ad hoc mobiles," in *Proc. of the Colloque Francophone sur la Gestion de Réseaux et de Services - GRES'2003*, Fortaleza, Brazil, Feb. 2003.

[185] W. Richard Stevens, *TCP/IP Illustrated*, vol. 1, Addison-Wesley, Reading, MA, 1994.

[186] Thomas H. Cormen, Charles E. Leiserson, and Ronald L. Rivest, *Introduction to Algorithms*, MIT Press, Cambridge, MA, 1990.

[187] Sean McCreary and K. Claffy, "Trends in wide area IP traffic patterns - a view from Ames Internet Exchange," in *Proc. of the ITC Specialist Seminar*, Monterey, CA, USA, Sept. 2000.

[188] Mohammad A. Al-Fawzan and Fahad A. Hoymany, "Placement of network servers in a wide-area network," *Computer Networks*, vol. 34, no. 3, pp. 355–361, Sept. 2000.

[189] Ran Giladi, Ephraim Korach, and Rony Ohayon, "Placement of network resources in communication networks," *Computer Networks*, vol. 43, no. 2, pp. 195–209, Oct. 2003.

[190] P. Krishnan, Danny Raz, and Yuval Shavitt, "The cache location problem," *IEEE/ACM Transactions on Networking*, vol. 8, no. 5, pp. 568–582, Oct. 2000.

[191] Lili Qiu, Venkata N. Padmanabhan, and Geoffrey M. Voelker, "On the placement of web server replicas," in *Proc. of the IEEE INFOCOM'2001*, Anchorage, AK, USA, Apr. 2001.

[192] Pavlin Radoslavov, Ramesh Govindan, and Deborah Estrin, "Topology-informed Internet replica placement," *Computer Communications*, vol. 25, no. 4, pp. 384–392, Mar. 2002.

[193] Eric Cronin, Sugih Jamin, Cheng Jin, Anthony R. Kurc, Danny Raz, and Yuval Shavitt, "Constrained mirror placement on the Internet," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 7, pp. 1369–1382, Sept. 2002.

[194] Harpal Singh Bassali, Krishnanand M. Kamath, Rajendraprasad B. Hosamani, and Lixin Gao, "Hierarchy-aware algorithms for CDN proxy placement in the Internet," *Computer Communications*, vol. 26, no. 3, pp. 251–263, Feb. 2003.

[195] Rahul Shah, Ravi Jain, and Farooq Anjum, "Efficient dissemination of personalized information using content-based multicast," in *Proc. of the IEEE INFOCOM'2002*, New York, NY, USA, June 2002.

[196] Artur Ziviani, Serge Fdida, José Ferreira de Rezende, and Otto Carlos Muniz Bandeira Duarte, "Placement issues in measurement-based host location," in *Proc. of the 9th European Summer School and IFIP Workshop on Next Generation Networks - EUNICE 2003*, Budapest - Balatonfüred, Hungary, Sept. 2003, pp. 173–180.

[197] Artur Ziviani, Serge Fdida, José Ferreira de Rezende, and Otto Carlos Muniz Bandeira Duarte, "Demographic placement for Internet host location," in *Proc. of the IEEE Global Communications Conference - GLOBECOM'2003*, San Francisco, CA, USA, Dec. 2003.

[198] Badrul Sarwar, George Karypis, Joseph Konstan, and John Riedl, "Item-based collaborative filtering recommendation algorithms," in *Proc. of the International World Wide Web Conference - WWW10*, Hong Kong, May 2001.

[199] Simone Santini and Ramesh Jain, "Similarity measures," *IEEE Transactions on Pattern Analysis and Machine Inteligence*, vol. 21, no. 9, pp. 871–883, Sept. 1999.

[200] Artur Ziviani, Serge Fdida, José Ferreira de Rezende, and Otto Carlos Muniz Bandeira Duarte, "Similarity models for Internet host location," in *Proc. of the IEEE International Conference on Networks - ICON'2003*, Sydney, Australia, Sept. 2003.

[201] *LibWeb*, `http://sunsite.berkeley.edu/Libweb`.

[202] *RIPE Test Traffic Measurements*, `http://www.ripe.net/ttm/`.

[203] Lakshminarayanan Subramanian, Venkata N. Padmanabhan, and Randy Katz, "Geographic properties of Internet routing: Analysis and implications," in *Proc. of USENIX 2002*, Monterey, CA, USA, June 2002.

[204] Soon-Hyung Yook, Hawoong Jeong, and Albert-László Barabási, "Modeling the Internet's large-scale topology," *Proc. of the National Academy of Sciences (PNAS)*, vol. 99, pp. 13382–13386, Oct. 2002.

[205] Thomas Brinkhoff, *City Population*, `http://www.citypopulation.de`.

[206] Central Intelligence Agency (CIA), *The World Factbook 2002*, Jan. 2002, `http://www.cia.gov/cia/publications/factbook`.

[207] Constantine Toregas, Ralph Swain, Charles Revelle, and Lawrence Bergman, "The location of emergency service facilities," *Operations Research*, vol. 19, no. 6, pp. 1363–1373, 1971.

[208] Mark S. Daskin, *Network and Discrete Location*, John Wiley & Sons, New York, NY, 1995.

[209] CAIDA, *GeoPlot*, `http://www.caida.org/tools/visualization/geoplot/`.

[210] Michael R. Garey and David S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*, W. H. Freeman & Co., New York, NY, 1979.

[211] David W. Jacobs, Daphna Weinshall, and Yoram Gdalyahu, "Classification with non-metric distances: Image retrieval and class representation," *IEEE Transactions on Pattern Analysis and Machine Inteligence*, vol. 22, no. 6, pp. 583–600, June 2000.

[212] R. N. Shepard, "Toward a universal law of generalization for psychological science," *Science*, vol. 237, pp. 1317–1323, 1987.

[213] Daniel M. Ennis, Joseph J. Palen, and Kenneth Mullen, "A multidimensional stochastic theory of similarity," *Journal of Mathematical Psychology*, vol. 32, no. 4, pp. 449–465, Dec. 1988.

[214] Stefan Savage, Andy Collins, Eric Hoffman, John Snell, and Tom Anderson, "The end-to-end effects of Internet path selection," in *Proc. of the ACM SIGCOMM'99*, Cambridge, MA, USA, Sept. 1999.

[215] Vern Paxson, Jamshid Mahdavi, Andrew Adams, and Matt Mathis, "An architecture for large-scale Internet measurement," *IEEE Communications Magazine*, vol. 36, no. 8, pp. 48–54, Aug. 1998.

[216] Martin Heusse, Franck Rousseau, Gilles Berger-Sabbatel, and Andrzej Duda, "Performance anomaly of 802.11b," in *Proc. of the IEEE INFOCOM'2003*, San Francisco, CA, USA, Apr. 2003.

[217] Geographic Location/Privacy (geopriv) IETF Working Group, ," `http://www.ietf.org/html.charters/geopriv-charter.html`.

[218] Sylvia Ratnasamy, Mark Handley, Richard Karp, and Scott Shenker, "Topologically-aware overlay construction and server selection," in *Proc. of the IEEE INFOCOM'2002*, New York, NY, USA, June 2002.

List of acronyms

ABE	*Alternative Best Effort*
AP	*Address Prefix*
AQM	*Active Queue Management*
AS	*Autonomous System*
BGP	*Border Gateway Protocol*
CBR	*Constant Bit Rate*
CDN	*Content Delivery Network*
DiffServ	*Differentiated Services*
DS	*Differentiated Services*
EBM	*Equation-Based Marking*
FEC	*Forward Error Correction*
FTP	*File Transfer Protocol*
GOP	*Group of Pictures*
GSM	*Global System for Mobile Communication*
HLR	*Home Location Register*
IETF	*Internet Engineering Task Force*
IMSI	*International Mobile Subscriber Identity*
IntServ	*Integrated Services*
IP	*Internet Protocol*
ISP	*Internet Service Provider*
LIP6	*Laboratoire d'Informatique de Paris VI*
LDP	*Label Distribution Protocol*
LSP	*Label Switched Path*
LSR	*Label Switching Router*
MAC	*Medium Access Control*
MOS	*Mean Opinion Score*
MPEG	*Moving Picture Experts Group*
MPLS	*MultiProtocol Label Switching*
NIMI	*National Internet Measurement Infrastructure*
NS	*Network Simulator*
PET	*Priority Encoding Transmission*
PHB	*Per-Hop Behavior*
PMP	*Paris Metro Pricing*
PRR	*Priority Round-Robin*

QoS	*Quality of Service*
RED	*Random Early Detection*
RFC	*Request For Comments*
RIO	*RED with In and Out*
RSVP	*ReSource reserVation Protocol*
RTT	*Round Trip Time*
SLA	*Service Level Agreement*
SLS	*Service Level Specification*
TCA	*Traffic Conditioning Agreement*
TCM	*Three Color Marking*
TCP	*Transmission Control Protocol*
TCS	*Traffic Conditioning Specification*
UDP	*User Datagram Protocol*
VBR	*Variable Bit Rate*
VLR	*Visitor Location Register*
VPN	*Virtual Private Network*
WDM	*Wavelength-Division Multiplexing*
WRR	*Weighted Round-Robin*

List of figures

List of tables

Qualité de Service et Localisation d'Hôtes

Résumé Cette thèse se concentre sur deux problèmes : (i) le support de Qualité de Service (QdS) aux trafics de voix et de vidéo ; et (ii) la localisation d'hôtes dans les réseaux ad hoc mobiles et dans l'Internet.

L'Internet, basé sur le service “au mieux” (*best-effort*), n'offre aucune différenciation parmi les paquets. Ainsi, la qualité perçue par l'utilisateur d'une transmission audio ou vidéo peut être fortement dégradée. Dans ce travail, on évalue le support fourni par l'architecture de différenciation de services aux transmissions de voix et vidéo dont la besoin ne cesse de croître. Les résultats montrent que le mécanisme de différenciation de services adopté améliore la qualité de la distribution des flux de voix et de vidéo. Nous évaluons l'efficacité démontrée par différents modèles de trafic de voix en utilisant une allocation additionnelle de bande passante et l'impacte du *shaping* du trafic de voix. Nous démontrons aussi que l'utilisation conjointe d'une protection inégale des trames du flux vidéo et de la tolérance aux pertes de paquets s'avère une stratégie efficace pour la transmission de qualité de la vidéo MPEG.

La connaissance de la localisation physique d'un hôte basée sur son identifiant ouvre la voie pour une classe toute entière d'applications exploitant cette information. Nous proposons, pour les réseaux ad hoc mobiles, d'utiliser les premiers paquets d'un flux pour apprendre la position de son destinataire au lieu d'utiliser un paquet de recherche dédié. Nous démontrons que cette approche réduit les coûts de transmission comparés à ceux des approches traditionnelles pour les flux TCP. Nous avons aussi évalué la performance, dans le contexte de l'Internet, des procédures de localisation géographique d'un hôte basées sur des mesures. Ces mesures sont réalisées par des machines sondes vers l'hôte à localiser ainsi que vers des hôtes de références qui ont une localisation géographique connue. Nous proposons et évaluons une politique démographique de placement des hôtes de référence et des machines sondes. Nous évaluons aussi les modèles de similarité qui comparent les profils de délai pour estimer la localisation. Nos résultats démontrent que le placement stratégique des hôtes de références et des machines sondes ainsi que le choix adéquat du modèle de similarité à adopter contribuent à une estimation de localisation géographique plus précise des hôtes Internet.

Mots clés Réseaux, qualité de service, trafic multimédia, connaissance de la localisation.

Quality of Service and Location-Awareness

Abstract is on page ix

Keywords Computer networks, quality of service, multimedia traffic, location-awareness.