INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Project-Team Rap

Réseaux, Algorithmes et Probabilités

Rocquencourt

THEME COM

Activity Report

2004

Table of contents

1. Team

Research Director
Philippe Robert [DR]

Administrative Assistant
Virginie Collette [TR]

Staff members (INRIA)
Christine Fricker [CR]

External Members
Jacqueline Boyer [France Télécom R&D, Lannion]
Fabrice Guillemin [France Télécom R&D, Lannion]
Stéphanie Poisson [France Télécom R&D, Lannion]
Danielle Tibi [University of Paris VII]

Post-Doctoral Fellow
Nelson Antunes [University of Faro]

Ph. D. Students
Youssef Azzana [INRIA]
Nadia Benazzouna [France Telecom R&D Lannion]
Hanène Mohamed [INRIA]

2. Overall Objectives

The research team RAP (Networks, Algorithms and Communication Networks) created in 2004 is issued from a long standing collaboration between engineers at France Telecom R&D in Lannion and researchers from INRIA-Rocquencourt. The initial objective was to formalize and expand this fruitful collaboration.

At France-Telecom R&D in Lannion, the members of the team are experts in the analytical modeling of communication networks as well as on some of the operational aspects of networks management concerning traffic measurements on ADSL networks for example.

At INRIA-Rocquencourt, the members of RAP have a recognized expertise in modeling methodologies applied to stochastic models of communication networks.

From the very beginning, it has been decided that the efforts of RAP project will focus on few dedicated domains of application over a period of three or four years. The general goal of the collaboration is to develop, analyze and optimize algorithms for communication networks. For the moment, the current projects are:

1. Mathematical Models of Traffic Measurements of ADSL traffic.
2. Design of Algorithms to Sample TCP flows.

The RAP project also aims at developing new fundamental tools to investigate *probabilistic* models of complex communication networks. We believe that mathematical models of complex communication networks require a deep understanding of general results on stochastic processes. It could be argued that, since stochastic networks are "applied", general results concerning Markov processes (for example) are not of a real use in practice and therefore that ad-hoc results are more helpful. Recent developments in the study of communication networks have shown that this point of view is flawed. Technical tools such as scaling methods, large deviations and rare events, requiring a good understanding of some fundamental results concerning stochastic processes, are now used in the analysis of these stochastic models. Two domains are currently investigated

1. Design and Analysis of Algorithms for Communication Networks. See Section 3.2.
2. Analysis of scaling methods for Markov processes: fluid limits and functional limit theorems. See Section 3.3.

3. Scientific Foundations

3.1. Measurements and Mathematical Modeling

Keywords: *Passive measurements, TCP traces.*

3.1.1. Traffic Modeling

Characterization of Internet traffic has become over the past few years one of the major challenging issues in telecommunications networks. As a matter of fact, understanding the composition and the dynamics of Internet traffic is essential for network operators in order to offer quality of service and to supervise their networks. Since the celebrated paper by Leland *et al* on the self-similar nature of Ethernet traffic in local area networks, a huge amount of work has been devoted to the characterization of Internet traffic. In particular, different hypotheses and assumptions have been explored to explain the reasons why and how Internet traffic should be self-similar.

A common approach to describing traffic in a backbone network consists of observing the bit rate process evaluated over fixed length intervals, say a few hundreds of milliseconds. Long range dependence as well as self-similarity are two basic properties of the bit rate process, which have been observed through measurements in many different situations. Different characterizations of the fractal nature of traffic have been proposed in the literature (see for instance Norros on the monofractal characterization of traffic). An exhaustive account to fractal characterization of Internet traffic can be found in the book by Park and Willinger. Even though long range dependence and self similarity properties are very intriguing from a theoretical point of view, their significance in network design has recently been questioned.

While self-similar models introduced so far in the literature aims at describing the global traffic on a link, it is now usual to distinguish short transfers (referred to as mice) and long transfers (referred to as elephants) [28]. This dichotomy was not totally clear up to a recent past (see for instance network measurements from the MCI backbone network). Yet, the distinction between mice and elephants become more and more evident with the emergence of peer-to-peer (p2p) applications, which give rise to a large amount of traffic on a small number of TCP connections. The above observation leads us to analyze ADSL traffic by adopting a flow based approach and more precisely the mice/elephants dichotomy. The intuitive definition of a mouse is that such a flow comprises a small number of packets so that it does not leave or leaves slightly the slow start regime. Thus, a mouse is not very sensitive to the bandwidth sharing imposed by TCP. On the contrary, elephants are sufficiently large so that one can expect that they share the bandwidth of a bottleneck according to the flow control mechanism of TCP. As a consequence, mice and elephants have a totally different behavior from a modeling point of view.

In our approach, we think that describing statistical properties of the Internet traffic at the packet level is not appropriate, mainly because of the strong dependence properties noticed above. It seems to us that, at this time scale, only signal processing techniques (wavelets, fractal analysis, ...) can lead to a better understanding of Internet traffic. It is widely believed that at the level of users, independence properties (like for telephone networks) can be assumed, just because users behave quite independently. Unfortunately, there is not, for the moment, a stochastic model of a typical user activity. Some models have been proposed, but their number of parameters is too large and most of them cannot be easily inferred from real measurements. We have chose to look at the traffic of elephants and mice which is an intermediate time scale. Some independence properties seem to hold at that level and therefore the possibility of Markovian analysis. Note that despite they are sometimes criticized, Markovian techniques are, basically, the *only* tools that can give a sufficiently precise description of the evolution of various stochastic models (average behavior, distribution of the time to overflow buffers,...).

3.1.2. Sampling the Internet Traffic

Traffic measurement is an issue of prime interest for network operators and networking researchers in order to know the nature and the characteristics of traffic supported by IP networks. The exhaustive capture of traffic traces on high speed backbone links, with rates larger than 1 Gigabit/s, however, leads to the storage and the analysis of huge amounts of data, typically several TeraBytes per day. A method of overcoming this problem is to reduce the volume of data by sampling traffic. Several sampling techniques have been proposed in the literature (see for instance [24][27] and references therein). In this paper, we consider the deterministic $1/N$ sampling, which consists of capturing one packet every other N packets. This sampling method has notably been implemented in CISCO routers under the name of NetFlow which is widely deployed nowadays in commercial IP networks.

The major issue with $1/N$ sampling is that the correlation structure of flows is severely degraded and then any digital signal processing technique turns out very delicate to apply in order to recover the characteristics of original flows [27]. An alternative approach consists of performing a statistical analysis of flow as in [24][25]. The accuracy of such an analysis, however, greatly depends on the number of samples for each type of flows, and may lead to quite inaccurate results. In fact, this approach proves efficient only in the derivation of mean values of some characteristics of interest, for instance the mean number of packets or bytes in a flow.

3.1.3. Algorithms of Sampling

Deriving the general characteristics of the TCP traffic circulating at some edge router has potential applications at the level of an ISP. It can be to charge customers propotionaly to their use of the network for example. It can be also to detect what is now called "heavy users".

Another important application is to detect the propagation of worms, attacks by denial of service (DoS). And, once the attack is detected, to counter it with an appropriate algorithmic approach. Due to the natural variation of the Internet traffic, such a detection (through sampling !) is not obvious. Robust algorithms have to be designed to achieve such an ambitious goal. An ultimate (and ambitious !) goal would be of having an automatic procedure to counter this kind of attacks.

3.1.4. Goals

- Propose a fairly *simple and accurate* estimation of the traffic circulating in an ADSL network. A limited number of parameters should characterize the traffic at the first order. Note that ADSL traffic is significantly different from the usual academic traffic analyzed up to now (more than 80% of the ADSL traffic is from Peer to Peer networks).
- *Infer* through sampling the parameters of the model proposed to describe the ADSL traffic.
- *Design and analyze algorithms* to detect in sampled traffic attacks by worms or DoS and more generally unusual events.

3.2. Design and Analysis of Algorithms

Keywords: *Data Structures, Stochastic Algorithms*.

The stochastic models of a class of generic algorithms with an underlying tree structure, the splitting algorithms, have a wide range of applications. To classify the massive data sets generated by traffic measurements, these algorithms turn out to be fundamental. Hashing mechanisms such as Bloom filters are currently investigated at the light of these new applications. These algorithms have also been used for now more than 30 years in various areas, among which

- Data structures. Fundamental algorithms on data structures are used to sort and search. They are sometimes referred to as divide and conquer algorithms.
- Access Protocols. These algorithms are used to give a distributed access to a common communication channel.
- Distributed systems. Recently, algorithms to select a subset of a group of identical communicating components like Ad-hoc networks, sensor networks and more generally mobile networks are using a related approach.

This class of algorithms is fundamental, their range of applications is very large and, moreover, they have a nice underlying mathematical structure. Trees are the main mathematical objects to describe them. The associated stochastic processes can be seen as a discrete version of fragmentation processes which have been recently thoroughly investigated by Bertoin, Pitman and others. They are also related to random recursive decompositions of intervals introduced by Mauldin and Williams and their developments in fractal geometry by Falconer, Lapidus, etc...

A very large subset of the literature has been devoted to the static case analysis, mainly because of its applications in theoretical computer science. In the dynamic case, i.e. when the shape of the tree changes according to some random events, little work has been done for this class of algorithms. Their analysis has been, for the moment, mainly achieved by using analytical methods with functional transforms, complex analysis techniques and inversions theorems. Curiously, despite of the intuitive underlying stochastic structures, probabilistic analyses of these algorithms are quite scarce (see Devroye for example).

3.2.1. Goals

- *Static case*. Generalize and simplify the results currently proved by using analytic tools. Prove limit theorems for *distributions* instead of averages as it is currently the case.
- *Dynamic case*. Study renormalization techniques to analyze tree algorithms under heavy traffic. The understanding of the fundamental features of these algorithms with a traffic of requests is a major issue in this domain. Because of the quite complex technical framework, the partial results obtained up to now with analytical tools hide, in some way, the general phenomena.

3.3. Scaling of Markov Processes

Keywords: *Fluid Limits, Functional Limit Theorems, Statistical Physics*.

As the complexity of communication networks increases (and, consequently, the algorithms regulating them), the classical mathematical methods used to estimate the stationary behavior, the transient behavior show more and more their limitations. For a one/two-dimensional Markov process describing the evolution of some network, it is sometimes possible to write down the equilibrium equations and to solve them. When the number of nodes is more than 3, this kind of approach is not, in general, possible. The key idea to overcome these difficulties is to consider limiting procedures for the system:

- by considering the asymptotic behavior of the probability of some events like it is done for large deviations at a logarithmic scale or for heavy tailed distributions, or looking at Poisson approximations to describe a sequence of events associated to them.
- by taking some parameter η of the model and look at the behavior of the system when it approaches some critical value η_c. In some cases, even if the model is complicated, its behavior simplifies as $\eta \to \eta_c$: some of the nodes grow according to some classical limit theorem and the rest of the nodes reach some equilibrium which can be described.
- by changing the time scale and the space scale with a common normalizing factor N and let N goes to infinity. This leads to functional limit theorems, see below.

The list of possible renormalization procedures is, of course, not exhaustive. But for the last ten years, this methodology has become more and more developed. Its advantages lies in its flexibility to various situation and also to the interesting theoretical problems it has raised since then.

3.3.1. An Example of Scaling Methods: TCP

In our past work the Congestion Avoidance Algorithm of the TCP protocol has been analyzed by using such a technique. The equilibrium of the *one*-dimensional Markov chain associated to this algorithm is not known for the moment. A large number of papers have been written on this famous AIMD Algorithm. But either it was, in some way, idealized or approximations were used without justifications. In a series of papers Dumas *et al.* [2], Guillemin *et al.* [11], a conveniently rescaled (time and space) Markov process has been analyzed in the limit when the loss rate of packets of some long connection was converging to 0. It provided a *rigorous* analysis to the scaling properties of this important algorithm of TCP.

3.3.2. Fluid Limits

A fluid limit scaling is a particular important way of scaling a Markov process. It is related to the first order behavior of the process, roughly speaking it amounts to a functional law of large numbers for the system considered.

It is in general quite difficult to have a satisfactory description of an ergodic Markov process describing a stochastic network. When the dimension of the state space d is greater than 1, the geometry complicates a lot any investigation : Analytical tools such as Wiener-Hopf techniques for dimension 1 cannot be easily generalized to higher dimensions. It is possible nevertheless to get some insight on the behavior of these processes through some limit theorems. The limiting procedure investigated consists in speeding up time and scaling appropriately the process itself with some parameter. The behavior of such rescaled stochastic processes is analyzed when the scaling parameter goes to infinity. In the limit one gets a sort of caricature of the initial stochastic process which is defined as a *fluid limit*.

A fluid limit keeps the main characteristics of the initial stochastic process while some stochastic fluctuations of second order vanish with this procedure. In "good cases", a fluid limit is a deterministic function, solution of some ordinary differential equation. As it can be expected, the general situation is somewhat more complicated. These ideas of rescaling stochastic processes have emerged recently in the analysis of stochastic networks, to study their ergodicity properties in particular. See Rybko and Stolyar [29] for example. In statistical physics, these methods are quite classical, see Comets [23].

Multi-Class Networks. The state space of the Markov processes encountered up to now were embedded into some finite dimensional vector space. For $J \in \mathbb{N}$, $J \geq 2$ and $j = 1,...J$, λ_j and μ_j are positive real numbers. It is assumed that J Poissonnian arrivals flows arrive at a single server queue with rate λ_j for $j = 1,..., J$ and customers from the jth flow require an exponentially distributed service with parameter μ_j. All the arrival flows are assumed to be independent. The service discipline is FIFO.

A natural way to describe this process is to take the state space of the finite strings with values in the set $\{1, ..., J\}$, i.e. $\mathcal{S} = \cup_{n\geq 0}\{1, ..., J\}^n$, with the convention that $\{1, ..., J\}^0$ is the set of the null string. If $n \geq 1$ and $x = (x_1, ..., x_n) \in \mathcal{S}$ is the state of the queue at some moment, the customer at the kth position of the queue comes from the flow with index x_k, for $k = 1, ..., n$. The length of a string $x \in \mathcal{S}$ is defined by $\|x\|$. Note that $\|\cdot\|$ is not, strictly speaking, a norm. For $n \geq 1$, there are J^n vectors of length n; the state space has therefore an exponential growth with respect to that function. Hence, if the string valued Markov process $(X(t))$ describing the queue is transient then certainly the length $\|X(t)\|$ converges to infinity as t gets large. Because of the large number of strings with a fixed length, the process $(X(t))$ itself has, a priori, infinitely many ways to go to infinity. Bramson [22] has shown that complicated phenomena could indeed occur. It turns out that the "classical" fluid limits methods of the finite dimensional case cannot be used in such a setting. This is probably one of the most challenging question in the domain to be able to propose new methods to tackle the problems due to the infinite dimension of the state space. Dantzer and Robert [1] derives results in this direction. See also the corresponding Chapter of Robert [4].

3.3.3. Goals

The general goals are, in some way, contained in the previous sections. They will consist in developing scaling techniques in the various cases encountered in sampling problems or tree algorithms where the traffic will be supposed to be close to saturation. The following fundamental questions will be analyzed

- Study the impact of randomness in fluid limit processes. This has been already partially investigated in Dantzer and Robert [1].
- Develop techniques to investigate metastability phenomena observed in some models of networks in the scaling limit due to mean field approach. See Kelly [26].

4. New Results

4.1. Mathematical Models of Traffic Measurements

Participants: Nelson Antunes, Nadia Benazzouna, Christine Fricker, Fabrice Guillemin, Philippe Robert.

4.1.1. Sampling ADSL traffic

The exhaustive capture of traces on high speed backbone link leads to the storage and the analysis of huge amount of data. In order to limit the consumption of memory in routers, passive traffic measurements employ sampling at the packet level. Indeed sampling techniques are implemented on CISCO routers (under the name of NetFlow). Flow statistics are formed by routers from the sampled substream of packets. Sampling entails a loss of information. The first question is whether sampling succeed in estimating the characteristics of the original traffic.

The aim of the study is to estimate the parameters of the real ADSL traffic from the sampled traffic. We use an a priori knowledge of the traffic, through the model developed in our previous work from the analysis of ADSL traces. Here the model is simplified a lot because mice are not seen by sampling and p2p traffic is predominant. Roughly speaking, traffic is mainly composed by p2p elephants. More precisely, the flows are chunks of elephants, probably due to the p2p algorithms. The analysis of traces leads to model the traffic by a $M/G/\infty$ where the customers are flows and their duration has a Weibull distribution. Sampling consists in choosing a customer at random every time step Δ. The traffic is characterized by a few parameters which have to be estimated: The arrival rate and the two parameters of the Weibull distribution of the flow duration.

The following results hold. First, in case of heavy traffic i.e. the arrival rate λ tends to infinity, if the sampling step Δ tends to 0 while $\Delta\lambda$ tends to a constant c then the sampling times of a permanent flow are the instants of a Poisson process with intensity $1/c$. This property is used to determine the arrival rate λ.

The number of sampling times of an elephant is shown to have the same distribution as the number of points of a Poisson process during a time interval whose length has a Weibull distribution. The exact asymptotics of the queue of the distribution is expressed. It is related to a result of Asmussen, Kluppelberg and Sigman. As a consequence, the duration of the sampled flow, given that it is sampled more than twice, is shown to be Weibull if the flows have such a distribution. It gives a way to estimate the parameters η and μ of the Weibull distribution which characterize the tail distribution of an elephant. Nevertheless, this is not satisfactory since the estimation of this tail distribution is not easy when the sampling step is large (one packet every thousand). The idea we want to use is to show that the number of flows which are seen k times converges in heavy traffic to a variable which has a Poisson distribution depending on the key parameters.

4.2. Integration of traffics in IP Networks

Participants: Nelson Antunes, Nadia Benazzouna, Christine Fricker, Fabrice Guillemin, Philippe Robert.

For traffic analysis, we adopt in this study a flow based approach and the popular mice and elephants dichotomy. A TCP flow is characterized by a sequence of packets characterized by four integers: source and destination addresses, source and destination ports. A mouse is a TCP flow with less than 20 packets, only the slow start phase of TCP protocol is used. On the contrary, due to their lengths, elephants share the remaining bandwidth because of the flow control mechanism of TCP. These two types of flows have therefore a completely different behavior from a modeling point of view.

Trace captures have been done by France Telecom R&D. TCP traffic has been collected on an Internet backbone link connecting different ADSL areas. A significant part of it are p2p applications and hence large elephants.

4.2.1. Interaction of TCP flows

The integration of two types of flows sharing a channel is an important issue in the design of communication networks. It applies in the context of IP networks to the case of best effort traffics (TCP) and streaming traffics (UDP). It is also relevant to describe the interaction of short TCP flows and large TCP flows. (See the above section).

The basic model analyzed here consists in a bottleneck link with variable capacity receiving TCP connections. The varying capacity is due to the UDP traffic or the mice traffic depending on the model considered. The general idea being that, contrary to "greedy" traffic like UDP traffic or mice traffic, the large TCP flows adapt their throughputs to the state of the link.

The varying capacity can be seen as driven by a stationary Gaussian process, an Ornstein-Uhlenbeck process. This assumption is natural based on the observations of traffic measurements and also, from a mathematical point of view, since it is known that the superposition of many small connections converges to such processes.

Up to now, the problem of expressing the invariant distribution of these systems is largely unsolved and known to be very difficult. Some earlier work by Nunez-Queija solves the problem in the context of a varying capacity driven by a Markov Modulated Point Process (MMPP). The result obtained are expressed in terms of complicated matrix expressions involving the (numerous) parameters of the MMPP. They are not easy to use in practice. This is also the reason why an Ornstein Uhlenbeck process has been taken, it is quite simple since it has only two parameters, the mean and the variance.

It has been chosen to study the case where the varying capacity oscillates around a fixed value μ. Perturbation methods are used to describe the stationary behavior of such a system. To tackle this problem two approaches have been used. The first one which is analytic consists in expanding the solution of the PDE associated to the dynamic of the system with respect to the perturbation parameter. This is the approach in Fricker *et al.* [7]. As a result a reduced load approximation has been proved when the capacity varies linearly with respect to the Ornstein-Uhlenbeck process.

4.2.2. Probabilistic expansion of a time varying $M/M/1$ queue

The other approach is probabilistic, it consists in analyzing the effect of the perturbation on one cycle of the Markov process. Under quite general assumptions, careful calculations lead to the expansion up to the second order of some of the characteristics of the queue.

In Antunes [20], we consider that the number of elastic flows in a link is represented by an $M/M/1$ queue whose server rate depends on the state of a stationary process $(X(t))$. Specifically, $(X(t))$ is the environment created by the superposition of unresponsive traffic so that the server rate at time t is $\mu + \varepsilon p(X(t))$, where p is some given function. Under the assumption that the server rate is weakly perturbed ($\varepsilon \ll 1$) by the environment, the system can be studied via perturbation analysis.

Two important measures of performance and throughput of a queue are the length and the area swept under the occupation process during a busy period. If the busy period is the time taken by the server to clear the queue, the are under the busy period is the total time spent by all customers waiting for the queue to be empty. The main results of this work are the explicit expressions of the two first terms of the power series expansion in ε of the mean value of the length and the area under the busy period. The first order term is consistent with the Reduced Service Rate approximation, which is equivalent that the service rate reduces to the average capacity of the variable queue. The second order term stresses the importance of the evolution of the varying capacity, through its correlation function.

Several applications of the expansion are studied. The case of non-negative and non-positive perturbation functions are considered. When the perturbation function is non-positive, the environment uses a part of the capacity of the $M/M/1$ queue with constant service rate μ. This application is motivated by the coexistence of elastic and streaming traffic where priority is given to stream traffic in a buffer of a router. The bandwidth available for non priority traffic is the transmission link reduced by the bit rate of streaming traffic. It is shown that for non-positive perturbation function, the variation of the service rate has a negative impact on the performance of the system when compared with a queue with constant service rate $\mu + \mathbb{E}(p(X(0))$.

Considering the complexity of the the second order term for a general perturbation function, we evaluate the system in two limit regimes, termed slow and fast environment. The environment is scaled by a factor $\alpha > 0$, such that at time t the environment is supposed to be $X(\alpha t)$. When the parameter α tends to infinity, the variations on the environment process completely vanish and the service rate reduces to the average capacity of the variable queue. On the other hand, as α goes to zero, the environment is frozen and is given by the initial state of the process.

The fast and slow environment provide an explicit estimates of the second order term, where the slow environment yields a worst performance of the perturbation queue. It is not clear that these limit regimes give a lower and upper bound of the performance of the queue. If the intuition leads to such conclusion, it remains an open problem to establish it rigorously.

Current research involves the expansion of the mean sojourn time of a customer in an $M/M/1$ processor sharing queue with perturbation. Functionals of a branching processes are used to get the first order expansion.

4.3. Algorithms to Infer Topologies

Participants: Youssef Azzana, Fabrice Guillemin, Philippe Robert.

The inference of the Internet topology is highly relevant in studying the spread of attacks and malicious programs such as worms and DOS through the network. It helps also to change the routing in order to balance the load and troubleshoot operational problems and also for network management. Recently, many protocols like multicast applications, traffic matrix estimation rely on the knowledge of the network topology to optimize the service provision and to increase the quality of service perceived by end users.

One popular approach to discover the network topology consists in using the theory of random graphs (Erdos and Renyi graphs, small world). It permits the construction of a random graph based on some local properties. Indeed, it has been observed that the degree distribution obeys to a power law. However, it's worth noting that a small error for example in the estimation of the power law parameter due to incomplete data may lead to erroneous interpretations. Another method exploits the BGP messages exchanged by different AS (Autonomous Systems). Thus, it's possible to construct the AS graph simply by listening to BGP messages. Then, one can refine the graph by looking for the IGP messages also. The most widely used method is the traceroute probing. In this approach the network is considered like a black box which is gradually explored. Traceroutes between two different hosts allows the discovery of the whole routers along the path between them. Indeed, the source transmitting the traceroute message gradually increment the TTL field of the packets sent to the destination (the number of hopes traversed) which make it possible to obtain the list of intermediate routers. Practically, a certain number of machines considered as sources proceeds by executing traceroutes to a list of destinations and the results are merged to construct the global map of the Internet.

In Azzana *et al.* [21], the network is represented as a tree, this might be unrealistic,but the different maps obtained by projects like CAIDA, ... suggest a tree like topology of the Internet, besides the network has a hierarchy structure. Moreover, if we consider the view collected by one source, the network is exactly a tree routed at this source.

The aim of our model is the study of the traceroute efficiency. This can be measured by the proportion of discovered nodes. More precisely, we randomly probe a number of destinations, we obtain a close formula of the average and variance of the proportion of discovered nodes. We derive also some expansions when the network size becomes large or when the number of destinations is small. Ultimately, a comparison of several tree models to see when traceroute discover easily the network is discussed. These results are extended to the case of Galton-Watson trees.

4.4. Analysis of Splitting Algorithms

Participants: Hanène Mohamed, Philippe Robert.

Algorithms with an underlying tree structure are quite common in computer science and communication networks. Splitting algorithms are examples of such algorithms.

A splitting algorithm is a procedure that divides recursively into subgroups an initial group of n items until each of the subgroups obtained has a cardinality strictly less than some fixed number D. A common problem is, given an initial number n of requests, to estimate the time it takes to complete the algorithm. In the language of trees, it amounts to give an asymptotic expression of the number R_n of nodes of the corresponding tree.

Classical methods in such a framework use complex analysis techniques (via Mellin transform) and various functional inversion theorems. A typical result in this domain is to establish that the sequence $\mathbb{E}(R_n)/n$ is equivalent to $F(\log_2(n))$, where F is some periodic function whose Fourier transform is known. Up to know, precise asymptotic results have been obtained only in this way.

A new, probabilistic approach to tackle this class of problems has been proposed. It consists in deriving a probabilistic representation of the quantities of interest to evaluate the asymptotic behavior of these algorithms. The basic ingredients are:

- A probabilistic de-poissonnization method which avoids to establish the, sometimes, painful functional inversion theorems.
- A convenient use of Fubini's Theorem.
- The use of the key renewal theorem to establish the convergence of the sequences. This is the critical point of the analysis. This is where the advantage of the probabilistic approach is notable.

The paper Robert [16] presents the method described above. Several classical results are re-derived through this method (with more precise error terms in fact). Moreover, purely analytical results, i.e. without any probabilistic background in the formulation of the problem, are also tackled with this technique. The asymptotic behavior of harmonic sequences for example.

The paper Mohamed and Robert [14] considers a very general splitting mechanism which consists in splitting a set of n items into a random number of groups and for each group the probability of assigning an item to it is given by some *random* variable. This model has been formulated by Devroye to analyze the depth of the associated tree. In this paper we analyze the total number R_n of nodes of the tree. It is shown that under some condition the sequence $\mathbb{E}(R_n)/n$ is converging to some constant whose expression is derived. Otherwise, the sequence is oscillating on a logarithmic scale according to a periodic function which has an explicit *integral* expression.

When the splitting mechanism is symmetrical, a new probabilistic representation of $\mathbb{R}_n$ in terms of Poisson processes gives quite directly a central limit theorem (obtained by Jacquet and Regnier) which is considered as a very technical result when proved via analytical methods.

5. Contracts and Grants with Industry

5.1. Contracts

Participation to the CRE with France Telecom “Bandwidth Allocation in the Internet”. This contract is for two years.

Participation to the RNRT project “Métropolis” on the measurements in the Internet. Four years contract ending in 2004.

Participation to the ACI GAP “Graphs, Algorithms and Probability” on the algorithms of peer to peer networks. Participants: INRIA, LIAFA and LRI. Three years contract.

Participation to the ACI Xtrems on the rare events in the modeling of the Internet. Participants: ENST, INRIA and University of Evry. Three years contract.

Participation to E-Next, a network of excellence of EC.

Participation to the ACI Masse de données “FLUX” on the probabilistic counting methods of large data sets occurring in traffic measurements, biological sequences, dictionaries. Participants: INRIA (Algo project), INRIA (Rap project) and University of Montpellier. Three years contract.

6. Other Grants and Activities

6.1. National initiatives

Philippe Robert et *Fabrice Guillemin* are participating to the “Action Spécifique Métrologie”. The other members are Pascal Abry (ENS-Lyon), Daniel Kofman (ENST), Philippe Owezarski (LAAS) and Kavé Salamatian (Paris VI).

6.2. European initiatives

RAP is participating to the E-next network of excellence of EC. This network involves many research teams throughout Europe. In France, participants include LIP6, INRIA-Sophia, LAAS,...This network is a continuation of the efforts of RAP team in the domain of traffic measurement.

6.3. Visiting scientists

RAP team has received the following people:
Christian Gromoll (Eurandom), Raouf Jaïbi (University of Tunis), Nelly Litvak (University of Twente, Boris Miller (IPPI, Moscow), Kavita Ramanan (Carnegie Mellon University), Ahmed Kharroubi (University of Casablanca) and Bert Zwart (Eurandom).

7. Dissemination

7.1. Leadership within scientific community

Philippe Robert is, from September 2004, Chairman of the Project Committee of INRIA-Rocquencourt.

Philippe Robert has been the referee for the PhD thesis by P. Moyal from ENST-Paris. He participated to the CR2 jury at INRIA-Rocquencourt.

Philippe Robert is “Professeur Chargé de Cours à l’École Polytechnique” in the department of applied mathematics. He is in charge of lectures on mathematical modeling of networks.

7.2. Teaching

Christine Fricker gives Master2 lectures “Stochastic Processes” at the University of Versailles St-Quentin.

Philippe Robert gives Master2 lectures “Stochastic Networks” in the laboratory of the Probability of the University of Paris VI. He is also giving lectures in the “Majeure de Mathématiques Appliquées et d’Informatique” on Networks and Algorithms at the École Polytechnique.

7.3. Conference and workshop committees, invited conferences

Nelson Antunes, *Youssef Azzana* and *Christine Fricker* were at the ARC TCP meeting from October 12 to October 14 at INRIA-Sophia .

Christine Fricker was from April 9th to April 11th and from July 15th to July 18th, at France Telecom R&D in Lannion, France.

Fabrice Guillemin was at the Globecom 2004 conference from December 6 to December 8 in Dallas, USA.

Philippe Robert gave the opening conference at ENS-Cachan “Algorithms, Data Structures and Communication Networks” on September 17.

8. Bibliography

Major publications by the team in recent years

[1] J.-F. DANTZER, P. ROBERT. *Fluid limits of string valued Markov processes*, in "Annals of Applied Probability", vol. 12, n^{o} 3, 2002, p. 860–889, http://www-rocq.inria.fr/~robert/src/papers/2002-2.pdf.

[2] V. DUMAS, F. GUILLEMIN, P. ROBERT. *A Markovian analysis of Additive-Increase Multiplicative-Decrease (AIMD) algorithms*, in "Advances in Applied Probability", vol. 34, nº 1, 2002, p. 85–111, http://www-rocq.inria.fr/~robert/src/papers/2002-1.pdf.

[3] C. FRICKER, P. ROBERT, D. TIBI. *A degenerate central limit theorem for single resource loss systems*, in "Annals of Applied Probability", vol. 13, nº 2, 2003, p. 561–575, http://www-rocq.inria.fr/~robert/src/papers/2003-1.pdf.

[4] P. ROBERT. *Stochastic Networks and Queues*, Stochastic Modelling and Applied Probability Series, vol. 52, Springer, New-York, June 2003.

Articles in referred journals and book chapters

[5] N. BEN AZZOUNA, F. CLÉROT, C. FRICKER, F. GUILLEMIN. *Modeling ADSL traffic on an IP backbone link*, in "Special Issue of Annals of Telecommunications on Traffic Engineering and Routing", vol. 59, nº 11-12, nov-dec 2004.

[6] P. DUBE, F. GUILLEMIN, R. MAZUMDAR. *Scale functions of Lévy processes and busy periods of infinite capacity M/GI/1 queues.*, in "Journal of Applied Probability", vol. 41, nº 4, 2004.

[7] C. FRICKER, F. GUILLEMIN, P. ROBERT. *Perturbation analysis of an $M/M/1$ queue in a diffusion random environment*, in "Advances in Applied Probability", Submitted, January 2004.

[8] F. GUILLEMIN. *Spectral analysis of birth and death processes.*, in "Journal of Applied Probability", Submitted, 2004.

[9] F. GUILLEMIN, R. MAZUMDAR. *Rate conservation laws for multidimensional processes of bounded variation with applications to priority queueing systems*, in "Methodology and Computing in Applied Probability", To appear, 2004.

[10] F. GUILLEMIN, D. PINCHON. *Analysis of the weighted fair queueing system with two classes of customers with exponential service times.*, in "Journal of Applied Probability", To appear, 2004.

[11] F. GUILLEMIN, P. ROBERT, B. ZWART. *AIMD algorithms and exponential functionals*, in "Annals of Applied Probability", vol. 14, nº 1, 2004, p. 90–117.

[12] F. GUILLEMIN, P. ROBERT, B. ZWART. *Tail asymptotics for processor sharing queues*, in "Advances in Applied Probability", vol. 36, nº 2, 2004, p. 525–543.

[13] I. MITRANI, P. ROBERT. *On the ASTA Property in a Feedback Processor-Sharing Queue*, in "Performance Evaluation", vol. 58, nº 1, October 2004, p. 81-85.

[14] H. MOHAMED, P. ROBERT. *A probabilistic analysis of some tree algorithms*, in "Annals of Applied Probability", Submitted, 2004.

[15] F. PIERA, R. MAZUMDAR, F. GUILLEMIN. *On product-form stationary distributions for reflected diffusions with jumps in the positive orhtant*, in "Journal of Applied Probability", Submitted, 2004.

[16] P. ROBERT. *On the Asymptotic Behavior of some Algorithms*, in "Random Structures and Algorithms", To Appear, vol. XX, 2004.

[17] B. SERICOLA, F. GUILLEMIN, J. BOYER. *Sojourn times in the M/PH/1 processor sharing queue*, in "Queueing Systems, Theory and Applications", Submitted, 2004.

[18] N. BEN AZZOUNA, F. GUILLEMIN. *Impact of Peer-to-Peer Applications on Wide Area Network Traffic: An Experimental Approach*, in "Globecom, Dallas", IEEE, 2004.

[19] N. BEN AZZOUNA, F. GUILLEMIN, S. POISSON, P. ROBERT, C. FRICKER, N. ANTUNES. *Inverting sampled ADSL traffic*, in "ICC 2005", Submitted, X, 2004.

Miscellaneous

[20] N. ANTUNES, C. FRICKER, F. GUILLEMIN, P. ROBERT. *Analysis of an $M/M/1$ queue with a varying capacity*, September 2004.

[21] Y. AZZANA, F. GUILLEMIN, P. ROBERT. *A Stochastic Model for Topology Discovery of Tree Networks*, June 2004.

Bibliography in notes

[22] M. BRAMSON. *Instability of FIFO queueing networks*, in "Annals of Applied Probability", vol. 4, n° 2, 1994, p. 414–431.

[23] F. COMETS. *Limites hydrodynamiques*, in "Astérisque", Séminaire Bourbaki, Vol. 1990/91, n° 201-203, 1991, p. Exp. No. 735, 167–192 (1992).

[24] N. DUFFIELD, C. LUND, M. THORUP. *Properties and Prediction of Flow Statistics Properties and Prediction of Flow Statistics*, in "ACM SIGCOMM Internet Measurement Workshop", November 2002, p. 6–8.

[25] N. DUFFIELD, C. LUND, M. THORUP. *Estimating Flow Distributions from sampled Flow statistics*, in "SIGCOMM", 2003, p. 25–29.

[26] R. J. GIBBENS, P. J. HUNT, F. P. KELLY. *Bistability in communication networks*, in "Disorder in physical systems, New York", Oxford Sci. Publ., Oxford Univ. Press, 1990, p. 113–127.

[27] N. HOHN, D. VEITCH. *Inverting sampled traffic*, in "IMC", October 2003, p. 27–29.

[28] V. PAXSON, S. FLOYD. *Wide area traffic: The failure of the Poisson assumption*, in "IEEE/ACM Trans. on Networking", 1995, p. 226-244.

[29] A. N. RYBKO, A. L. STOLYAR. *On the ergodicity of random processes that describe the functioning of open queueing networks*, in "Problems on Information Transmission", vol. 28, nº 3, 1992, p. 3–26.