INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Project-Team rap

Réseaux, Algorithmes et Probabilités

Rocquencourt

THEME COM

Activity Report

2005

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 Moteau [France Télécom R&D, Lannion]
Danielle Tibi [University of Paris VII]

Post-Doctoral Fellow
Nelson Antunes [University of Faro (until February 2005)]
Stylianos Kouremenos [National Technical University of Athens (Starting from September 2005)]

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

2. Overall Objectives

2.1. 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) [24]. 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 [20], [23] 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 [23]. An alternative approach consists of performing a statistical analysis of flow as in [20], [21]. 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 situations 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.* [4], 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 [25] for example. In statistical physics, these methods are quite classical, see Comets [19].

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 [18] 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 [5].

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 [22].

4. New Results

4.1. Mathematical Models of Traffic Measurements

Participants: Nelson Antunes, Youssef Azzana, Christine Fricker, Fabrice Guillemin, Stéphanie Moteau, 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, due to the p2p algorithms. The analysis of traces leads to model the traffic by a $M/G/\infty$ queue 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.

A first approach gives that, in case of heavy traffic i.e. if the arrival rate λ tends to infinity and if the sampling step Δ tends to 0 while Δ/λ 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. If the duration of the flow is Weibull then the duration of the sampled flow, given that it is sampled more than twice, is also Weibull. It gives a way to estimate the parameters of the Weibull distribution. In practice, this is not satisfactory since the estimation of the tail distribution is not easy when the sampling step is large (one packet every thousand).

An alternative approach is to use quantities whose mean can be obtained as a function of the key parameters, typically the number W_k of flows sampled less (resp. more) than k times in a given time interval. There exists a scaling of Δ such that this mean tends to a constant. In this case, *Chen-Stein method* is used to prove the convergence in distribution of W_k to a Poisson distribution when the total number of flows is large. This method is powerful enough to give precise estimates of the distance of the distributions. When the mean tends to infinity, a normal approximation can be also obtained as a consequence. The system is reduced to dynamical urn model because the flows are not permanent in the time interval.

In practice, the ratio $\Delta\lambda$ is assumed to be a constant and the elephants can, in this case, be considered as permanent. It has been proved that a normal approximation holds for the number of flows sampled more than $k \geq 1$ times, when the ratio of the number of flows to the number of sampling times is small. Comparing experimental values obtained on traces and theoretical ones, we obtained a discrepancy which is probably be due to the bursty nature of the data elephants or the presence of mice. This point is currently under investigation.

4.1.2. On Line Algorithms For Traffic Measurements

We are interested here in detecting and estimating the number of flows traversing a router in the network. The characterization of the flow statistics is of interest for the detection of attacks or anomalies, it can be also used to charge the clients in function of the traffic generated, also in traffic engineering. Moreover, Internet providers can infer the clients application (Peer-to-Peer, voice over IP, web, ftp...) without looking at the packets contents.

We focus on big flows (those who exceed a certain number of packets T or occupy more than certain percentage of the total available bandwidth). Indeed, it is known that big flows represent the majority of the traffic volume, for example, we know that less than 9% of the flows exchanged between AS represent up to 70% of the total number of bytes exchanged between all the AS pairs. Also, for a lot of applications, the knowledge of those big flows is sufficient to characterize the traffic.

To answer this question, we proposed an algorithm based on the use of T parallel Bloom filters, each filter i has a counter C_i. Initially, all the T filters are empty and the different counters also initialized to 0. Upon the reception of a flow F, we look for the first parallel filter (determined by a hashing function) where flow F does not exist yet, then we increase the value of the counter of this filter by 1 and we fill the different bits corresponding to F by 1. When the size of the filters is well parametrized, all the flows of size bigger than i reach the filter i with a negligible proportion of flows of size smaller than i. Consequently, we use the value of the counter C_i as an estimator of the total number of flows of size larger than i. Since this algorithm must run in real time without interruption, all the filters become saturated after a while especially because of the contribution of mice and the estimation error becomes unacceptable. To deal with this problem, we have proposed an adaptive mechanism which cleans out the filters regularly and maintains the filling of the filters under a certain threshold (50% in our case). Indeed, as soon as this threshold is reached, we remove some packets by reinitializing the first parallel filters and moving them to the end of the T filters.

The simulations we made show that the least we erase the packets the best is the estimation (it is better to remove one packet than T packets). Indeed, by totally cleaning out the T filters, we remove the contribution of all the mice but we make a lot of errors in the detection and the statistics of the elephants in the contrary of erasing only one packet. The simulations show also that the relative error of the total number of elephants is maintained low around 3 to 4% and is stabilized over a long period of time. The first moments of the elephants size (average, and variance) show also a satisfying concordance with real statistics of elephants.

These algorithms have been successfully tested on ADSL traces corresponding to two hours of traffic.

4.2. 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 is 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 is 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.

The results obtained last year have be extended to the case where the tree structure of the topology is not anymore regular and deterministic but is a Galton-Watson tree. If F is the number of leaves of a node, several cases have been separately investigated when $\mathbb{E}(F^2) < +\infty$ or when $\mathbb{E}(F^2) = +\infty$ and $\mathbb{E}(F) < +\infty$, or finally when $\mathbb{E}(F) = +\infty$.

4.3. Stability Properties of Loss Networks

Participants: Nelson Antunes, Christine Fricker, Philippe Robert, Danielle Tibi.

A new class of stochastic networks has been introduced and analyzed. Their dynamics combine the key characteristics of the two main classes of queueing networks : loss networks and Jackson type networks.

1. Each node of the network has finite capacity so that a request entering a saturated node is rejected as in a loss network.
2. Requests visit a subset of nodes along some (possibly) random route as in Jackson or Kelly's networks.

This class of networks is motivated by the mathematical representation of cellular wireless networks. Such a network is a group of base stations covering some geographical area. The area where *mobile users* communicate with *a base station* is referred to as *a cell*. A base station is responsible for the bandwidth management concerning mobiles in its cell. New calls are initiated in cells and calls are handed over (transfered) to the corresponding neighboring cell when mobiles move through the network. A new or a handoff call is accepted if there is available bandwidth in the cell, otherwise, it is rejected.

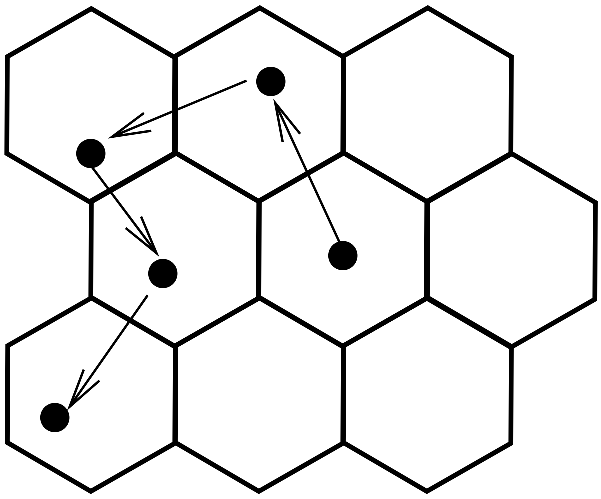

Figure 1.

The time evolution of these networks has been analyzed by considering two limiting regimes

- Heavy traffic limits.
  The arrival rates and capacities at nodes are proportional to some factor N which gets large.
- Thermodynamic limits.
  The number of nodes of the network goes to infinity.

The time evolution of the network can be (roughly) described as follows. A stochastic process $(\overline{X}_N(t))$ associated with the state of the network for the parameter N is introduced : $\overline{X}_N(t)$ is the vector describing the number of requests of different classes at the nodes of the network. As N goes to infinity, it is proved that $(\overline{X}_N(t))$ converges to some function $(x(t))$, satisfying the deterministic equation

$$\frac{d}{dt}x(t) = F\left(x(t)\right), \qquad t \geq 0. \tag{1}$$

The equilibrium points of the limiting process are contained in the set of solutions x of the equation $F(x) = 0$.

It is shown in Antunes *et al.* [16] that for the heavy traffic limit, there is a unique equilibrium point. The proof uses a dual method approach to study the fixed point equations together with some convenient inequalities.

For the thermodynamic limit, it is shown in Antunes *et al.* [17] that there are situations where *several equilibrium points* coexist. This result has practical important implications for communication networks : It implies that, in some cases, the network will stay a long time in a set of states where a class of calls will be rejected and after this long time, it will switch to a set of states where this class of calls has a higher acceptance rate and, again after a long time, it switches back to the first set of states and so on. At the mathematical level, this is the situation where the function F has at least two stable points and a saddle point. The proof uses an interesting correspondence between two energy functions defined in different state spaces.

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.

4.4.1. Dynamic tree algorithm

This is a dynamic version of a class of algorithms analyzed by Mohamed and Robert [8]. The splitting procedure is the same, but a phenomenon of immigration has to be considered : on every leaf of the associated tree, every time unit, new messages arrive following a Poisson process of parameter λ. Contrary to the « static case », the boundary conditions turn out to complicate a lot the resolution of the problem. A new probabilistic tool has to be used ; an auto-regressive process whose invariant density plays an important role to determine the asymptotic behavior of the cost of the algorithm.

4.4.2. Leader election algorithm

A related algorithm, a leader election algorithm, has been analyzed. It has been previously investigated by Janson and Szpankowski with analytic methods. This algorithm is used in the context of a distributed system of n stations sharing a common channel of communication that can transmit only one message by unit of time. We assume that every station which sends a message to network can listen at the same time to the channel and so discern one of three possible information on the state of this one ; a *collision* when it there at least two tries of transmission, a *silence* when none of the stations tried to send its message or a *success* when exactly one station tries transmission. Question is then how these stations can, by using the same protocol, identify one of them as a *leader* to coordinate the whole system ?

Such algorithm based on a process of random elimination has varied applications in distributed systems field such as cell telephones and networks of wireless communications. The problem of leader election in networks computer science is fundamental to assure communications and synchronization of the different components of the system. This problem was also studied in the context of radio networks.

Formally, the algorithm of leader election divides an initial group of n items into two subgroups, eliminates one of two and continues the same process until finding one *leader*. If at a given level k all items are eliminated, algorithm starts again from the previous level $k-1$.

Our study, based on probabilistic techniques, allows to simplify the analysis of such algorithm and especially to eliminate the implicit dependency of its asymptotic behavior as it is the case in the expression established by Janson and Szpankowski. Besides, an explicit representation of the associated oscillation phenomenon has also been obtained. These results are obtained via a careful analysis of the following probabilistic functional equation

$$h(x) = h(px) + h(qx)\, e^{-px} + f(x) = \mathbb{E}\left(h(Ax)\, e^{-p 1_{\{A=q\}}}\right) + f(x),$$

where A is a random variable of distribution $\mathcal{W} = p\,\delta_p q\,\delta_q$ and f a given function. The use of a simple iterative scheme gives an explicit expression of the average cost of the algorithm $\mathbb{E}(H_n)$. It is proved that the centered average cost of the algorithm $H_n - \lfloor -\log_p(n) \rfloor$ is asymptotically identical to a periodical function F, whose explicit expression is known, of $-\log_p(n)$

$$\mathbb{E}(H_n) - \lfloor -\log_p(n) \rfloor = F(-\log_p(n)) + O(1/n).$$

4.4.3. *Extensions to Stationary Sequences*

The results of Mohamed and Robert [8] have been extended to the case where the branching procedure are not independent but are driven by a dynamical system. These results are known (See Vallée and its co-workers) to hold for some dynamical systems generated by the iterations of some function on $[0, 1]$. Our approach gives a further extension to general dynamical systems. It uses a general version of the renewal theorem for stationary sequences together with a representation of the cost function as counting functional.

5. Contracts and Grants with Industry

5.1. Contracts

Participation to the CRE with France Telecom « Mathematics of Internet Measurements ». Two years contract starting from 2005.

Participation to the RNRT project « OSCAR » on the measurements in the Internet. Three years contract starting from 2005.

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 starting from 2004.

Participation to the ANR Projet Blanc « SADA » on the Discrete Random Structures, three year contract starting from 2005. Participants : University of Bordeaux, University of Caen, Computer science department of Ecole Polytechnique, INRIA Algo and Rap projects, University of Versailles.

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 :
Nelson Antunes (University of Algarve), Nelly Litvak (University of Twente, Kavita Ramanan (Carnegie Mellon University) and Bert Zwart (Eurandom).

7. Dissemination

7.1. Leadership within scientific community

Philippe Robert is the Chairman of the Project Committee of INRIA-Rocquencourt.

Philippe Robert has been the referee for the PhD thesis by P. Brown from France-Telecom R&D Sophia-Antipolis.

Philippe Robert is « Professeur Chargé de Cours » at the 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

Philippe Robert was invited as lecturer at the ALEA'2005 Conference at Luminy.

Christine Fricker and *Philippe Robert* were at the Large Deviations Workshop from July 3rd to 5th in Ottawa, Canada.

Christine Fricker, *Hanène Mohamed*, *Philippe Robert* and *Danielle Tibi* gave talks at the Informs Conference from July 6th to 8th in Ottawa, Canada.

Youssef Azzana gave a talk at the ITC'18 in Beijing.

Philippe Robert gave a talk at the Dynamical Systems Workshop in Dijon from September 14 to 16 and at the seminar of ergodic theory in Rennes, at the Algo seminar in Rocquencourt and at École Polytechnique.

Nelson Antunes, *Christine Fricker* and *Philippe Robert* were at Performance'05 from October 5th to 7th in Juan-les-Pins, France.

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° 3, 2002, p. 860–889.

[2] V. DUMAS, F. GUILLEMIN, P. ROBERT. *A Markovian analysis of Additive-Increase Multiplicative-Decrease (AIMD) algorithms*, in "Advances in Applied Probability", vol. 34, n° 1, 2002, p. 85–111.

[3] C. FRICKER, P. ROBERT, D. TIBI. *A degenerate central limit theorem for single resource loss systems*, in "Annals of Applied Probability", vol. 13, n° 2, 2003, p. 561–575.

[4] F. GUILLEMIN, P. ROBERT, B. ZWART. *AIMD algorithms and exponential functionals*, in "Annals of Applied Probability", vol. 14, n° 1, 2004, p. 90–117.

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

Articles in refereed journals and book chapters

[6] N. ANTUNES, C. FRICKER, F. GUILLEMIN, P. ROBERT. *Perturbation Analysis of a Variable $M/M/1$ Queue : A probabilistic Approach*, in "Advances in Applied Probability", To Appear.

[7] N. ANTUNES, C. FRICKER, F. GUILLEMIN, P. ROBERT. *Integration of streaming services and TCP data transmission in the Internet*, in "Performance Evaluation", vol. 62, n° 1-4, October 2005, p. 263–277.

[8] H. MOHAMED, P. ROBERT. *A probabilistic analysis of some tree algorithms*, in "Annals of Applied Probability", vol. 15, n° 4, November 2005, p. 2445–2471.

[9] F. PIERA, R. MAZUMDAR, F. GUILLEMIN. *On product-form stationary distributions for reflected diffusions with jumps in the positive orhtant*, in "Advances in Applied Probability", vol. 37, n° 1, March 2005, p. 212–228.

[10] P. ROBERT. *On the asymptotic behavior of some Algorithms*, in "Random Structures and Algorithms", vol. 27, n° 2, September 2005, p. 235–250.

[11] B. SERICOLA, F. GUILLEMIN, J. BOYER. *Sojourn times in the M/PH/1 processor sharing queue*, in "Queueing Systems, Theory and Applications", vol. 50, n° 1, 2005, p. 109–130.

Publications in Conferences and Workshops

[12] N. ANTUNES, C. FRICKER, F. GUILLEMIN, P. ROBERT. *Perturbation analysis of the area swept under the the queue length process of a variable $M/M/1$ queue*, in "Performance'05, Juan les Pins", IFP WG 7.3, 2005, http://www-rocq.inria.fr/~robert/src/papers/2005-1full.pdf.

[13] Y. AZZANA, F. GUILLEMIN, P. ROBERT. *A Stochastic Model for Topology Discovery of Tree Networks*, in "Proceedings of ITC'19, Beijing", 2005, http://www-rocq.inria.fr/~robert/src/papers/2004-6.pdf.

[14] N. BEN AZZOUNA, F. GUILLEMIN, S. POISSON, P. ROBERT, C. FRICKER, N. ANTUNES. *Inverting sampled ADSL traffic*, in "Proceedings of ICC'05, Seoul", 2005, http://www-rocq.inria.fr/~robert/src/papers/2004-7.pdf.

Miscellaneous

[15] N. ANTUNES, J. BOYER, C. FRICKER, F. GUILLEMIN, S. MOTEAU, P. ROBERT. *A probabilistic approach to packet sampling in the Internet*, Submitted, October 2005.

[16] N. ANTUNES, C. FRICKER, P. ROBERT, D. TIBI. *Analysis of Loss Networks with Routing*, Preprint, October 2005, http://hal.ccsd.cnrs.fr/docs/00/05/26/56/PDF/Full-Heavy.pdf.

[17] N. ANTUNES, C. FRICKER, P. ROBERT, D. TIBI. *On Stochastic Networks with Multiple Stable Points*, Preprint, November 2005.

Bibliography in notes

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

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

[20] 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.

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

[22] 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.

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

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

[25] 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, n° 3, 1992, p. 3–26.