Optimizing the date of an upgrading investment in a data network

Frédéric Morlot, Benoît Fourestié, Salah Eddine Elayoubi

Optimizing the date of an upgrading investment in a data network

Frédéric Morlot, Benoît Fourestié, Salah-Eddine Elayoubi
{frederic.morlot, benoit.fourestie, salaheddine.elayoubi}@orange-ftgroup.com

***Abstract*—Due to the introduction of new services, the volume of data transferred in mobile networks is rapidly growing and operators periodically face the necessity to upgrade their network. Such upgrades allow them to increase the capacity, and provide adequate Quality of Service (QoS). In this paper we propose a general framework for deriving the optimal date for a network upgrade. We show that this date is the result of a compromise betweenthe decrease of upgrade investment cost and the loss of profit generated by insufficient capacity. The upgrade should hence be performed when the loss of profit, derived using analytical capacity expressions, exceeds the expected discount. The model presented herein accounts for the randomness of the demand and upgrading cost functions, and results are given for a HSDPA network.**

***Index Terms*—data flows, brownian motion, quality of service, customer satisfaction.**

I. Introduction

As new high throughput services such as multimedia applications are introduced in the market, the traffic is significantly growing in mobile telecommunications networks. In particular, data flows are growing exponentially, as they have been doing since the middle 90' in the Internet. To face these soaring volumes of data to be transferred, mobile operators must periodically upgrade their equipments to offer higher throughputs lest they should undergo blocking problems. However as the demand does not increase steadily and must be considered as partly random, the upgrade dates are difficult to be forecast.

In this article we consider a HSDPA cellular network and assume for the sake of simplicity that there exists no priority between data flows. As the demand increases, the operator can choose between the following two strategies:

- *with admission control*: it sets a lower throughput threshold and thus has to prevent some users from entering the network to guarantee this minimum throughput. As the demand rises, more and more users are rejected: the *blocking rate* increases.
- *without admission control*: there is no throughput lower threshold. All users are admitted in the network, but the communication duration becomes longer and longer for each user until the network is saturated. The global throughput experienced in the network may result very small.

In the remainder of the article, we will assume that the operator has chosen the second option. Note that if it had chosen the admission control strategy, main results presented herein would still hold and generalization elements are given in section V. We further assume that the "fair power" sharing algorithm is used [1] since it can be shown that it has good stationary properties. As the demand rises, the operator increases its profit, since each user transfers more data and thus spends more money. When the network starts experiencing saturation problems, the global troughput and profit may fall. The operator must then upgrade its network and faces the following trade-off:

- The later the investment, the lower individual throughputs and customer satisfaction. Permanent insatisfaction will result into churn and additional loss of profit.
- The sooner the investment, the more expansive as upgraded element costs decrease with time.

This article aims at proposing a global framework for evaluating the optimal date for network upgrading, and its content is the following:

- First, operator's profit is calculated. Analytical values for capacity, individual throughput and satisfaction will be derived as a function of the demand. The randomness of the rising demand will be accounted for;
- Second, upgrading costs will be estimated assuming exponential decrease according to time;
- Third, elements are given in case that an admission control strategy is used;
- Fourth, results are given for a HSDPA network and show that the more random (volatile) the demand, the earlier the investment should be made.

To the best of our knowledge, this paper presents the first endeavour to analytically derive optimal dates of investments in mobile radio networks taking into account randomly growing data transfers, capacity and throughput evaluations and exponentially decreasing investment costs. Such an analysis will be of interest for future operator upgrading decisions.

II. Operator's Profit

The operator profit depends on the amount of data flowed by the network, i.e. on the sum of individual throughputs. However individual throughputs depend on a large number of parameters and are difficult to calculate. The profit thus depends on the following parameters:

- the mean daily traffic demand per cell (in Mbits/sec/cell). We denote it by X_t, where the time $t = 0, 1...T$ varies *discretely*, for example day by day. $[0, T]$ stands for the considered time period.
- the customer satisfaction

  - before the upgrading: $S_t^1 = f^1(X_t)$
  - after the upgrading: $S_t^2 = f^2(X_t)$
- the daily profit per surface unit generated by X_t, possibly balanced by S_t^i

$$\pi_t^i = g(X_t, S_t^i) = g(X_t, f^i(X_t))$$

If we denote by t_0 the investment date ($0 < t_0 < T$), the total profit $\Pi_T(t_0)$ actualized at the date $t = T$ is calculated as follows:

$$\Pi_T(t_0) = \sum_{t=0}^{t_0-1} e^{\zeta(T-t)}\pi_t^1 + \sum_{t=t_0}^{T} e^{\zeta(T-t)}\pi_t^2, \qquad (1)$$

where ζ is the actualization rate. For simplicity, we assume that ζ is constant during the period $[0, T]$.

A. Traffic Demand

We assume that the network is formed by circular cells of radius R and that the demand is uniformly distributed in the cell. We have:

$$X_t = \lambda_t \times E[\xi],$$

where λ_t is the mean arrival rate per cell at the date t and $E[\xi]$ is the mean size of a typical data flow. In the following we assume that $E[\xi]$ remains stable during $[0, T]$, so that X_t is proportional to λ_t, whereas λ_t is growing. This model is strictly equivalent to the case where the number of active users is constant, but they initiate connections more often. Furthermore a model where the rate of calls λ_t would be constant with $E[\xi]$ increasing would lead to similar results.

To model the evolution of $(X_t)_{t\in\mathbb{N}}$, let us consider it as the daily *sampling* of a continuous stochastic process $(\tilde{X}_t)_{t\in\mathbb{R}^+}$. Usually, one monitors the 24 mean demands over one hour each, keeps the second or third highest, and multiply it by a given factor. As many random phenomena related to a social behaviour (e.g. [2][5]), we assume that $\tilde{X}_t$ is growing according to the following stochastic differential equation:

$$\begin{cases} \frac{d\tilde{X}_t}{\tilde{X}_t} = \alpha dt + \sigma dW_t \\ \tilde{X}_0 = x_0 \qquad \text{almost surely,} \end{cases} \qquad (2)$$

where W_t is a standard brownian motion. In other words, $\tilde{X}_t$ is a geometric brownian motion, whose solution is given by (see [4]):

$$\tilde{X}_t = x_0 e^{(\alpha - \sigma^2/2)t + \sigma W_t}, \quad t \in \mathbb{R}^+$$

Note that we have here $\mathbb{E}[X_t] = x_0 e^{\alpha t}$. Hence, α is the trend of the demand, whereas σ is its volatility. Figure 1 shows the daily evolution of the traffic according to time obtained using the described model.

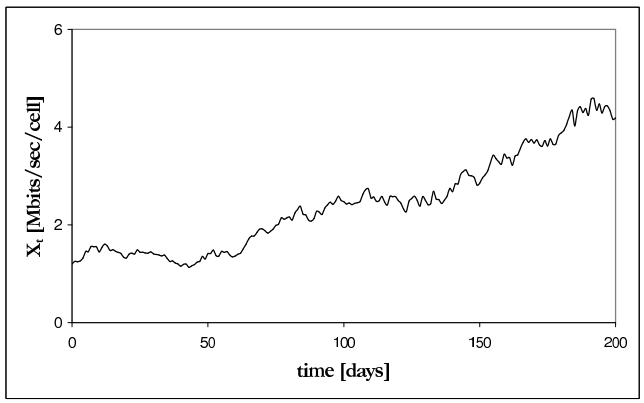

Fig. 1. Daily traffic evolution. α = -log(0.85)/30 day^{-1} (which corresponds to an average increasing of 15% per month), σ = 0.04, x_0 = 1.2 Mbit/sec/cell.

B. Customer Satisfaction

To evaluate the customer satisfaction, let us first compute the flow throughput $\gamma_t(r)$ a user can expect at distance r from the center of the cell as carried out in [1]. The resource of a single downlink channel is time-shared between active users. Denote by ϕ_u the fraction of time the BS transmits to user u, with $\sum_u \phi_u$ = 1. The data rate of user u is then

$$C_u(r) = C(r) \times \phi_u,$$

where $C(r)$ is the peak data rate, obtained in the absence of any other user in the cell, i.e., for ϕ_u = 1. When there are x users, the "fair power" sharing is defined by $\phi_u = \frac{1}{x}$. We have:

$$C(r) = min\left(C_0, \frac{Z}{b} \times \frac{\Gamma(r)}{\eta + I(r)}\right),$$

where:

- C_0 is the maximum peak rate (which depends on channel bandwidth and coding efficiency);
- Z is the cell chip rate;
- b is a lower bound for energy-per-bit to noise density ratio (E_b/N_0), such that the error probability will not be higher than a target threshold;
- $\Gamma(r)$ is the path loss between the BS and user u;
- η is the thermal noise to received power ratio;
- $I(r)$ is the interference to received power ratio

.

Figure 2 shows the peak data rate in a hexagonal cell. In the interference term, we only take immediate adjacent cells in account. Furthermore, we assume that the path loss $\Gamma(r)$ is decreasing according to $1/r^4$.

Note that we considered the ideal case where a continuous set of peak rates is achievable. But in practice, because of coding constraints, the set of achievable peak rates is discrete, and locally constant.

We define the cell load as (see [1]):

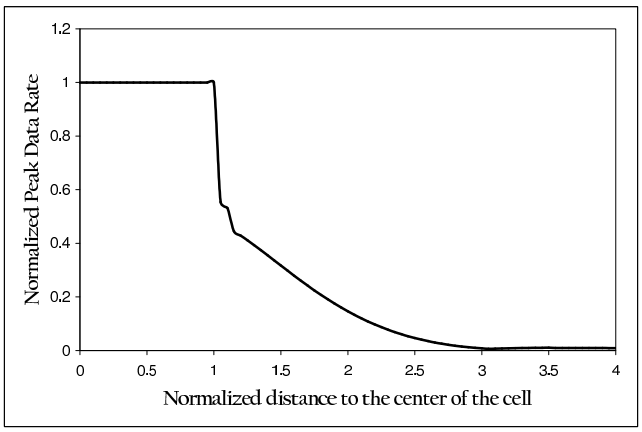

Fig. 2. Peak data rate against distance to the center of the cell. The distance and the rate are normalized with respect to the cell's radius and to C_0. Z/b = 3 Mbits/sec, η = 10%.

$$\rho_t = \frac{X_t}{\pi R^2} \int_0^R \frac{2\pi r dr}{C(r)},$$

where R is the radius of the cell.

- if $\rho_t > 1$, the cell is overloaded and one can show ([1]) that the number of active users grows indefinitely; any individual data rate tends to zero, and the cell is saturated.
- if $\rho_t < 1$, the cell is underloaded and the number of active users tends to a finite stationary regime.

We are naturally led to introduce:

$$X_{max} = \pi R^2 \left(\int_0^R \frac{2\pi r dr}{C(r)} \right)^{-1}, \tag{3}$$

such that $\rho_t = X_t / X_{max}$. Hence $\rho_t < 1 \iff X_t < X_{max}$.

Let us denote the flow throughput of users at distance r by $\gamma_t(r)$ (it is the ratio of the mean flow size to the mean flow duration). Then, Bonald and Proutière [1] show that, if $\rho_t < 1$:

$$\gamma_t(r) = C(r)(1 - \rho_t). \tag{4}$$

By assumption, the mean flow size (the numerator in $\gamma_t(r)$) does not vary significantly, whereas the mean flow duration (the denominator) does as the load increases. Thus, if we want to compute the mean flow throughput, the significant number to calculate is the *harmonic* mean of $\gamma_t(r)$ over the cell (balanced by the proportion of active users between r and $r + dr$; remember we suppose users are uniformly distributed in the cell. In other words, we have to calculate the arithmetic mean of $1/\gamma_t(r)$ over the cell. From (4) we deduce:

$$\overline{\gamma_t} = R^2 \left(\int_0^R \frac{2 r dr}{\gamma_t(r)} \right)^{-1} = X_{max} - X_t,$$

and of course, in overload, $\overline{\gamma_t}$ = 0.

Finally, we can summarize the whole calculation by:

$$\overline{\gamma_t} = (X_{max} - X_t)^+ \tag{5}$$

where, for a given real x, $x^+ = max(x, 0)$. Note that $\overline{\gamma_t}$ = 0 if and only if $X_t \geq X_{max}$, i.e., if and only if the cell is saturated.

Now we can compute the customer satisfaction, which can reasonably be supposed to depend on $\overline{\gamma_t}$. Since subjective satisfactions have been shown to be more sensitive to small variations at low throughputs than at high throughputs, Enderlé and Lagrange propose in [6] to model the customer satisfaction as a negative exponential function of the throughput:

$$S_t = e^{-\beta/(X_{max} - X_t)^+}.$$

For example, β can be chosen as: $\beta = log(2) \cdot \gamma_{1/2}$, where $\gamma_{1/2}$ is the throughput value ensuring a satisfaction of 50%. Once again, note that S_t = 0 if and only if $X_t \geq X_{max}$, i.e., if and only if the cell is saturated.

C. Daily Profit

Let us recall that X_t is sampled day by day, for example during the second or the third highest peak hour. Within 24 hours, the operator transmits $\mu min(X_t, X_{max})$ to active users, where μ is a multiplicative factor between the peak hour and the whole day. Typically, $\mu \approx 4 \cdot 3600 \approx 14000$ sec. Since taxation is applied to the volume of transfers and not to the duration, the gross daily profit per cell is given by:

$$\pi^{gross} = \delta \mu min(X_t, X_{max}) \tag{6}$$

where δ is the transfer price (say in \$/Mbit). However the gross profit should be weighed by the customer satisfaction to account for the quality of the communications. The *net profit* is thus calculated as the product of π^{gross} by S_t

$$\pi^{net} = \delta \mu min(X_t, X_{max}) e^{-\beta/(X_{max} - X_t)^+}.$$

If S_t = 0, i.e. if the cell is saturated, the net profit is null. If the satisfaction is maximal, i.e., S_t = 1, the net profit is equal to the gross profit (6). To sum up we have:

$\pi_t = \delta \mu X_t e^{-\beta/(X_{max} - X_t)}$	*if* $X_t < X_{max}$
$\pi_t = 0$	*otherwise*

Intuitively, as the demand rises, X_t will increase as will the profit (Fig. 3). Then, the profit will decrease because the unsatisfaction effect becomes dominant. In case no upgrading action is taken, the profit will progressively tend to zero. Once the operator decides to upgrade the network, two solutions are possible:

- make the network denser, i.e. install new base stations in the network
- install additional transmitters operating on different frequency bands with the view to increase the throughputs. In such case we obtain a higher value of X_{max}, so that:

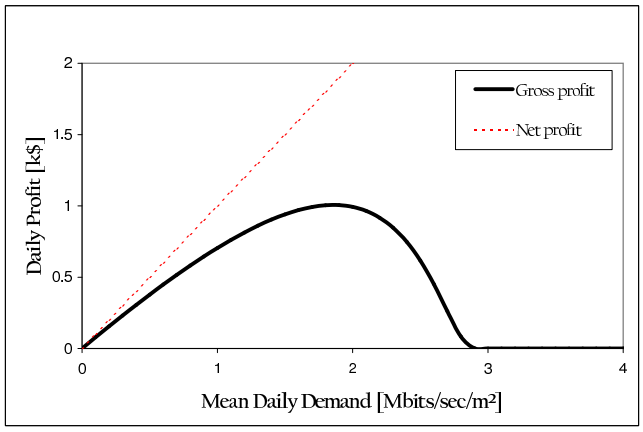

Fig. 3. Daily Profit generated by demand. X_{max} = 3 Mbits/sec/cell, δ = 0.1 $/Mbit, μ = 14000 sec, β = 0.7 Mbit/sec (which corresponds to $\gamma_{1/2} \approx 1$ Mbit/sec).

$$\pi_t = \delta\mu min(X_t, X'_{max})e^{-\beta/(X'_{max}-X_t)^+}$$

In this article we restrict our analysis to the latter case, i.e., X_{max} changes but other variables do not.

D. Total Profit

As we have now derived π_t analytically, we can compute numerically the expectation of $\Pi_T(t_0)$ given t_0 using the Monte Carlo method. Results for the calculation of the daily profit calculated using Scilab are shown in Fig. 4.

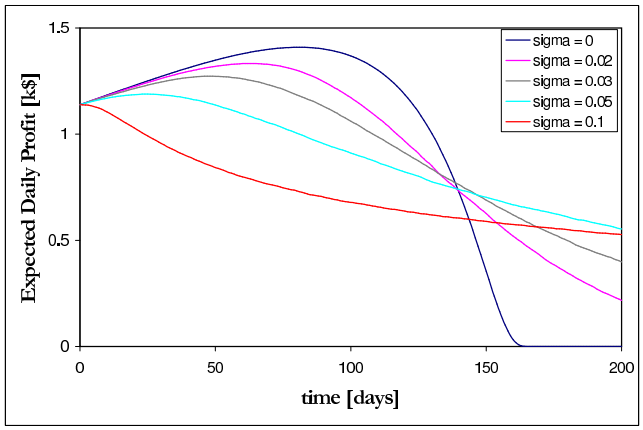

Fig. 4. Expected Daily Profit during the period. Parameters are identical to those used in former figures.

We observe that the higher sigmais , the more important the risk is and the sooner the maximal value of the expected profit is reached. Additionally, the higher sigma is, the slower the curve tends to zero. When σ becomes paramount, the curve flattens out.

From the previous simulation, we can deduce the total profit shown in Fig. 5. We observe that the expected profit diminishes as the volatility increases, because of the risk.

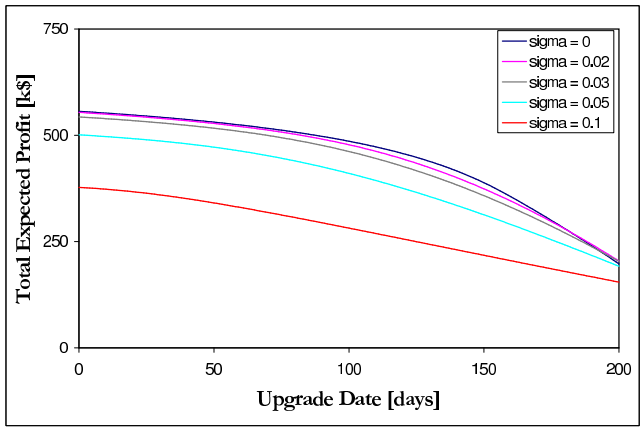

Fig. 5. Total Expected Profit. X_{max} = 8 Mbits/sec/cell after the upgrade. ζ = -log(0.95)/365 day^{-1}, which corresponds to a 5% / year actualization.

To summarize the results:

- for a given value of σ, the earlier the investment, the happier the consumers, and the higher the profit;
- at a given date, the higher σ, the lower the profit, since the system can be saturated earlier.

However, we must now take the investment cost into account. This is the subject of next section.

III. Upgrading Investment

The upgrading cost is a decreasing function of time. The decrease of the cost is due to many factors, for example the R&D progress, and also the serialization in the manufacturing chain. Moore's law states that electronic devices' capacity doubles every 18 months. So in this paper we assume it decreases exponentially:

$$K(t) = K_0 e^{-\epsilon t},$$

where ϵ is the depreciation rate.

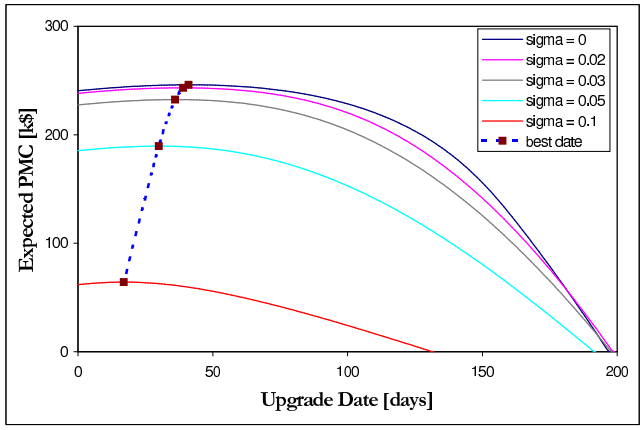

Fig. 6. Expected PMC during the period. K_0 = 300 k$ and ϵ = -log(0.5)/365, which corresponds to a 50% / year cost decrease.

The number to maximize is what we call the total *Profit-minus-Cost* (PMC), actualized at date T:

$$\begin{aligned} PMC(t_0) &= \Pi_T(t_0) - K(t_0)e^{\zeta(T-t_0)} \\ &= \Pi_T(t_0) - K_0 e^{\zeta T-(\epsilon+\zeta)t_0}. \end{aligned}$$

Note that generally $\epsilon \gg \zeta$, so that $\epsilon + \zeta \approx \epsilon$. Results when taking the $K(t)$ component in our simulations and results are shown in Fig. 6.

Again we show that the higher the volatility is, the lower the expected profit is. This result was expectable because of the increased risk. Furthermore, we see that the optimal investment date decreases according to the volatility, as shown in Fig. 6. However, works on investment under uncertainty obtain the opposite result [5]. In fact, as will be shown in next section, the significant value to observe is the optimal *actualized* investment date, and we will observe the opposite trend.

IV. Strategy

As time goes by, we have realizations of the daily demand. These values can thus be used to derive a more precise optimal date than the one we found at t = 0 assuming demand randomness. Instead of calculating expectations, we can now take in account what *really* happens during the period. We can implement the following algorithm: each day $t \in [0, T]$, we perform deterministic calculations over $[0, t]$ and expectations over $[t, T]$, with an initial value x_t of traffic demand which is now a deterministic parameter. Each computation thus provides us with a new investment date $t_0 \in [t, T]$. In other words, t_0 is daily *actualized*. When t_0 becomes close to t, it is time to invest. Note that such a situation ($t_0 \to t$) will inevitably happen by continuity, since at the end of the process we have $t = t_0 = T$. Interestingly we can take the upgrading delay into account: if for example we need 1 month of delay for the deployment, we will plan it as soon as $t_0 - t \leq 30$ days.

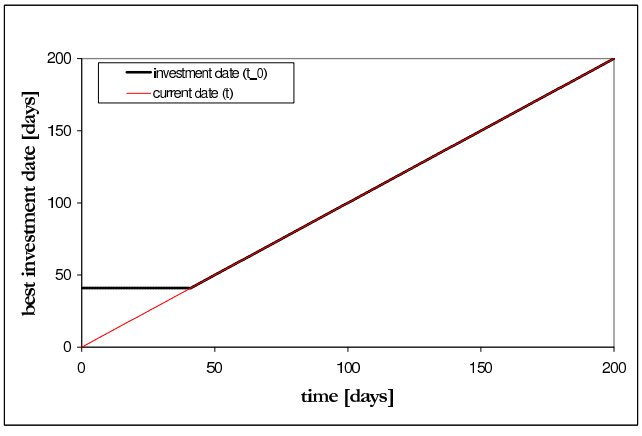

Fig. 7. Best investment date during actualization algorithm in the deterministic case. Investment will take place at t = 41 days.

In Fig. 7, we show the results in the deterministic case, i.e., σ = 0. For t = 0, we find t_0 = 41 days as expected by Fig. **??**. Since there is no randomness, t_0 is fixed until $t = t_0$. Then, t_0 remains "stuck" to t. In the random case, we can observe some fluctuations before t_0 becomes stuck to t (see Fig. 8). In fact, the optimal date we find (60 days) seems to be higher than the one found at the beginning (30 days). Is it a hazard, or is the mean final best result always greater ? To examine that assumption, we simulate several possible scenarios, in other words, several possible trajectories of the brownian demand. For each trajectory, we apply the actualization algorithm and we compute a final result, averaged over all the simulations (see Fig. 9): now, we are in accordance with classic investment studies.

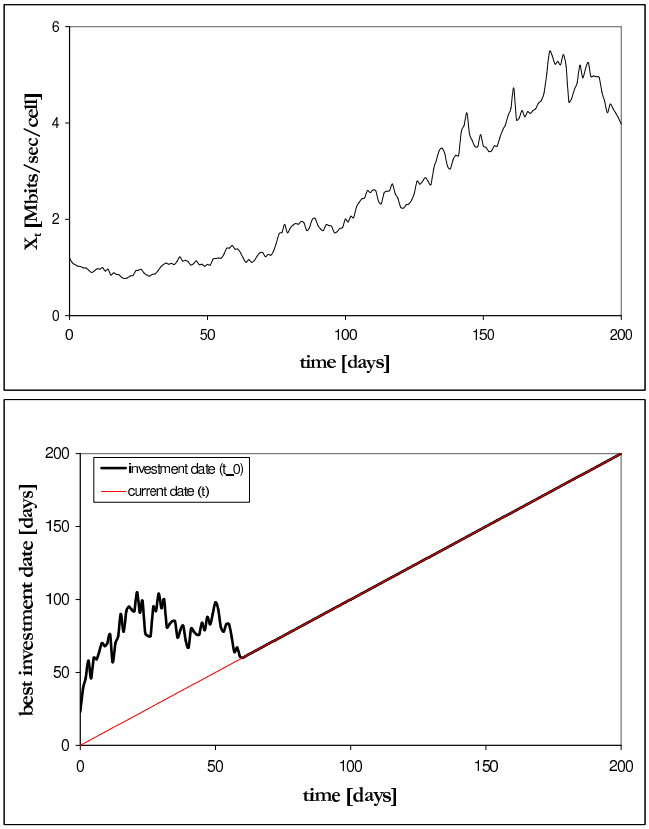

Fig. 8. Best investment date during actualization algorithm in the random case (σ = 0.05), for one given trajectory. Here, investment will take place at around t = 60 days. When X_t does not grow as fast as expected, the actualized investment date is postponed. But as soon as X_t has sufficiently grown, the actualized date is stabilized.

V. Generalization

So far, the actualization algorithm has been illustrated in a simple situation. However it could be applied to a wide range of different situations. For example, we could introduce:

- admission control: the network then guarantees to each user a minimal data throughput, but some users are blocked. In that case, the satisfaction S_t depends on the blocking rate, not on the throughput. [1] gives an analytic expression of the blocking rate:

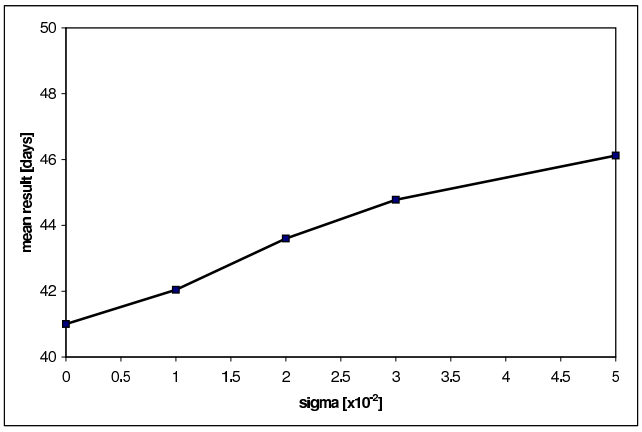

Fig. 9. Mean final date returned by actualization algorithm regarding volatility. As expected, this date becomes later while volatility is growing.

$$B = \frac{\rho_t^m}{1 + \rho_t + ... + \rho_t^m},$$

where ρ_t is the cell load (refer to section II-B) and m is the maximum allowed number of users in one cell. Then we assume that the customer satisfaction decreases exponentially according to blocking rate, and the actualization algorithm still holds.

- different power allocations. In that case, we have to compute a mean satisfaction over the cell, balanced by the probability density function of the user position.
- coding constraints, which entail discontinuities in the peak data rate curve [1].
- discontinuities and randomness in the investment cost curve to modelize R&D's breakthroughs. This random component is added to the expected PMC's computation.

VI. Conclusion

In this paper, we presented an analytical approach for determining the optimal date of an upgrade investment in mobile radio networks. The strategy we adopted is:

- first, analytically model the operator profit and the customer satisfaction;
- second, compute the optimal upgrading date with an actualization algorithm.

The first step is of utmost importance because it has to reflect as accurately as possible the real situation of the network: Radio Ressource Management (RRM) algorithms (such as admission control, power allocation) must be taken into account to correctly model the capacity of the network. To illustrate this modelling, we consider a HSDPA data network and show how to calculate the cell capacity, the user's throughput and the satisfaction, and the operator's profit. Note that the demand parameters, such as the volatility, the trend and the mean size of data flows, have to be evaluated carefully because results are rather sensitive to input variations, especially to volatility (Fig. **??**). Models should therefore be calibrated using measurements taken in the past to avoid too big discrepancies between forecasts and reality.

In the second step, we introduced the cost of the upgrading investment and developed an actualization algorithm that calculates the optimal investment time using the above described analytical modeling.

The method presented herein should be of interest for operators, allowing them to make optimal decisions as to when upgrade their radio networks.

Acknowledgment. We are especially grateful to Thomas Bonald for the useful pieces of advice he gave to us.

References

[1] T. Bonald, A. Proutière (France Telecom R&D), *Wireless Downlink Data Channels: User Performance and Cell Dimensioning*, The 57th IEEE Semiannual, April 2003.

[2] Y. d'Halluin, Peter A. Forsyth, and Kenneth R. Vetzal, *Managing Capacity for Telecommunications Networks Under Uncertainty*, IEEE/ACM Transactions on Networking, vol. 10, N4, August 2002.

[3] L. Salahaldin and T. Granger, *Investing in Sustainable Transport Infrastructure under Investment Cost and Population Growth Uncertainty*, Université Paris Dauphine, 2006.

[4] N. El Karoui, E. Gobet, *Introduction au Calcul stochastique, première partie*, Imprimerie de l'École Polytechnique, p. 88, 2003.

[5] L. Salahaldin and T. Granger, *Investing in Sustainable Transport to Relieve Air Pollution under Population-Growth Uncertainty* , 9th Annual International Conference on Real options, Paris, June 2005.

[6] N. Enderlé, X. Lagrange, *User Satisfaction Models and Scheduling Algorithms for Packet-Switched Services in UMTS*, Proceedings of IEEE VTC spring, April 2003.