Dimensionnement de réseau avec prévisions de demandes incertaines et contrainte de monoroutage

Olivier Klopfenstein

Dimensionnement de réseau avec prévisions de demandes incertaines et contrainte de monoroutage

Olivier Klopfenstein

Orange Labs, 38-40 rue du gl Leclerc, 92130 Issy-les-Moulineaux, France

On considère le problème du dimensionnement de réseau avec demandes incertaines. On suppose que chaque demande en trafic est gaussienne, et doit être routée selon un seul chemin dans le réseau. Sur le plan théorique, on montre que le routage aux plus faibles coûts marginaux est un bon algorithme d'approximation pour ce problème. Une comparaison numérique avec une résolution optimale est réalisée.

Keywords: Dimensionnement de réseau, demandes incertaines, contraintes probabilistes, mono-routage

1 Introduction

La gestion opérationnelle des réseaux doit faire face aux incertitudes dues au trafic. Lorsqu'un réseau existe et doit être exploité, il faut prendre en compte la variabilité du trafic des clients connectés. On parle dans ce cas d'*ingénierie de trafic*, dont les prises de décision concernent des échelles de temps relativement courtes (typiquement, du mois au semestre). Pour un réseau de données (internet par exemple), les grandeurs permettant d'évaluer la qualité de service sont notamment la bande passante offerte, le délai de routage et la probabilité de perte de paquets. Ces grandeurs sont bien sûr liées entre elles, et sont concernées directement par les phénomènes de congestion.

Mais il est nécessaire aussi d'anticiper l'évolution du trafic à plus long terme (à une année ou plus), voire de concevoir des réseaux pour l'introduction de nouveaux services. On parle alors de *planification* du réseau et de son dimensionnement. L'incertitude première à gérer concerne la prévision de l'évolution globale du trafic, par exemple l'évolution du nombre de clients par service, ainsi que celle de leurs usages.

Il se trouve que pour chacune des deux échelles de temps décrites ci-dessus, les approximations gaussiennes sont généralement justifiées pour les réseaux de données multi-services. On propose un modèle d'optimisation du dimensionnement d'un réseau maillé (coeur de réseau) qui repose sur une description gaussienne des demandes en trafic. Un de nos objectifs est de maîtriser le risque associé à un dimensionnement, ce qui conduit naturellement à un problème d'optimisation sous contraintes probabilistes. Enfin, on prend en compte une contrainte souvent requise en pratique : chaque demande ne doit être acheminée que par un seul chemin dans le réseau.

Il existe aujourd'hui dans la littérature un très grand nombre de publications consacrées à la prise en compte de l'incertitude sur la prévision de la demande pour la planification de réseaux. On renonce ici à une bibliographie représentative de ce domaine pour simplement en donner les principaux axes de développement. Une grande partie de ces travaux se concentre sur les approches d'optimisation robuste, où le réseau est dimensionné de manière à satisfaire un ensemble de scénarios de trafic donnés (cf par exemple [1, 10]). Cependant, en pratique, la définition de ces ensembles de scénarios n'est pas toujours aisée. Une autre famille de papiers utilise des programmes stochastiques à plusieurs étapes avec recours [7]. Au niveau méthodologique, le problème de la définition des scénarios à considérer demeure. Les publications relatives au dimensionnement de réseau avec contraintes probabilistes sont très peu nombreuses. Dans [8], les auteurs traitent du dimensionnement de réseaux ATM avec contrôle de la probabilité de blocage des connexions. Récemment, [9] a étudié un modèle proche de celui présenté ici.

Compte tenu des restrictions sur la taille du papier, les preuves des résultats sont omises.

2 Modèles mathématiques

On note $G = (N,A)$ le graphe orienté du réseau. À l'échelle de l'ingénierie de trafic, concentrons-nous sur un lien $a \in A$ du réseau. On note c_a sa capacité et $\ell(a)$ sa charge. L'ensemble des flux qui le traversent peuvent être vus comme des sources de trafic stochastiquement indépendantes, dont l'agrégation $\ell(a)$ a un comportement gaussien (sur des périodes de temps où le trafic est *stationnaire*, par exemple aux heures les plus chargées). Cette intuition a été vérifiée par de nombreuses études pratiques sur des réseaux de données (voir par exemple [6]). Alors la quantité $\mathbb{P}(\ell(a) > c_a)$ est la probabilité de congestion du lien.

À l'échelle de la planification de réseau, l'incertitude touche la valeur du trafic prise pour référence pour le dimensionnement. Cette valeur de référence, pour un type de trafic donné, peut être par exemple la valeur moyenne de l'intensité du trafic, ou encore sa valeur-pic (ce sont deux exemples extrêmes). Toujours est-il que cette valeur de référence est impossible à prévoir parfaitement. Elle dépend notamment de l'évolution du nombre d'utilisateurs du réseau, ainsi que de leur comportement. Entre chaque paire de points du réseau (origine-destination), chaque service se trouve ainsi caractérisé par une distribution de sa valeur de référence. Si l'on pense au théorème de la limite centrale, l'agrégation de ces services entre l'origine et la destination considérées justifie une approximation gaussienne du trafic global. Cette approximation est d'autant plus justifiée au niveau de chaque lien, où l'on additionne les trafics issus de demandes ayant des couples origine-destination distincts. La quantité $\mathbb{P}(\ell(a) > c_a)$ est alors un niveau de risque de congestion associé au dimensionnement du lien a.

On note ρ_a le coût d'une unité de trafic écoulée sur le lien $a \in A$ du réseau. On considère un ensemble I de demandes en trafic, chaque demande $i \in I$ ayant pour source s_i, pour destination d_i, et pour trafic b_i. On introduit le problème de dimensionnement de réseau sous contraintes probabilistes, avec contrainte de monoroutage des demandes (appelé par la suite *CCNDP*, pour *Chance-Constrained Network Design Problem*) :

$$\begin{array}{lll}
\min & \sum_{a\in A} \rho_a . c_a & \quad (1)\\
\text{s.c.} & \sum_{a\in A^+(v)} x_{ia} - \sum_{a\in A^-(v)} x_{ia} = \begin{cases} 1 & \text{si } v = s_i, \\ -1 & \text{si } v = d_i, \\ 0 & \text{sinon} \end{cases} \quad \forall i \in I, v \in V, & \quad (2)\\
& \mathbb{P}\left(\sum_{i\in I} b_i x_{ia} \leq c_a\right) \geq 1-\varepsilon, \quad \forall a \in A, & \quad (3)\\
& x_{ia} \in \{0,1\}, \quad \forall i \in I, a \in A, & \quad (4)\\
& c_a \geq 0, \quad \forall a \in A. & \quad (5)
\end{array}$$

Les variables de décision sont x et c. La variable x_{ia} indique si la demande $i \in I$ est routée sur l'arc a. La variable c_a représente la quantité de bande passante consommée sur le lien a. Les égalités (2) sont les contraintes classiques de conservation du flot. Le dimensionnement de chaque arc a se fait selon un niveau de risque dont la contrainte (3) impose qu'il soit inférieur à $\varepsilon \in [0; 0,5]$.

On suppose que les demandes sont stochastiquement indépendantes, et suivent chacune une loi normale : $\forall i \in I, b_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$. Il est bien connu que dans ce cas, l'équation (3) pour $a \in A$ devient (cf e.g. [5]) :

$$\sum_{i\in I} \mu_i x_{ia} + \Phi^{-1}(1-\varepsilon). \sqrt{\sum_{i\in I} \sigma_i^2 . x_{ia}^2} \leq c_a.$$

Sous ces hypothèses, le problème *CCNDP* consiste à trouver le multiflot mono-routé qui minimise le coût : $\sum_{a\in A} \rho_a . \left(\sum_{i\in I} \mu_i x_{ia} + \Phi^{-1}(1-\varepsilon). \sqrt{\sum_{i\in I} \sigma_i^2 . x_{ia}^2}\right)$. On a supposé ici, pour éviter toute complication inutile, que les volumes des demandes étaient stochastiquement indépendants. Mais l'intégration de corrélations se fait sans difficulté. Noter que le problème obtenu est NP-difficile [3].

3 Routage selon les plus faibles coûts marginaux

Dans le cas où toutes les demandes sont bien connues (i.e. toutes les variances sont nulles), le problème est résolu simplement en routant chaque demande selon le plus court chemin au sens des poids $\{\rho_a\}_{a\in A}$ (routage selon les plus faibles coûts marginaux). On peut construire des exemples qui montrent cette démarche n'est pas toujours optimale. On va cependant mettre en évidence qu'elle reste le plus souvent très pertinente.

On note n le nombre de noeuds du réseau. On suppose dans cette partie que les coûts ρ_a sont entiers. Les résultats suivants montrent que si la variance des demandes est suffisamment petite, le routage optimal s'effectue selon les plus faibles coûts marginaux :

Proposition 1 *Si toutes les demandes ont le même écart-type* $\sigma > 0$, *et si :*

$$\sigma \leq \frac{\min_{i\in I}\mu_i}{\Phi^{-1}(1-\varepsilon).(n-1).\max_{a\in A}\rho_a},$$

alors il existe un routage optimal tel que chaque demande est routée selon un de ses plus courts chemins au sens des poids $\{\rho_a\}_{a\in A}$.

Proposition 2 *Supposons que tous les coûts* $\{\rho_a\}_{a\in A}$ *prennent la même valeur* $\bar{\rho}$, *et que le réseau est k-connexe. Si toutes les demandes ont le même écart-type* $\sigma > 0$, *et si :*

$$\sigma \leq \frac{k.\min_{i\in I}\mu_i}{\Phi^{-1}(1-\varepsilon).(n+k-2)},$$

alors il existe un routage optimal tel que chaque demande est routée selon un de ses plus courts chemins en nombre de bonds.

Les hypothèses ci-dessus sont relativement restrictives en pratique. Par exemple, pour un réseau biconnexe de $n = 20$ noeuds, avec $\varepsilon = 0,1$, on obtient : $\sigma \leq 0,078.\min_{i\in I}\mu_i$. D'autre part, il est important de noter que, même sous ces hypothèses restrictives, n'importe quel routage aux plus courts chemins n'est pas nécessairement optimal.

Proposition 3 *Supposons que les hypothèses de la proposition 1, ou de la proposition 2, sont satisfaites. Alors le routage optimal est le routage* $\hat{x}$ *aux plus courts chemins selon les poids* $\{\rho_a\}_{a\in A}$ *qui minimise la quantité :* $\sum_{a\in A}\rho_a\sqrt{\sum_{i\in I}\hat{x}_{ia}}$.

Cependant, si le routage selon les chemins de plus faibles coûts marginaux n'est pas toujours optimal, il est toujours une bonne approximation de l'optimum. C'est ce que montrent les résultats suivants.

Proposition 4 *Soit* $\hat{x}$ *un routage des demandes aux plus courts chemins selon les poids* $\{\rho_a\}_{a\in A}$. *On suppose qu'il existe* $\alpha \geq 0$ *tel que, pour toute demande i :* $\sigma_i \leq \alpha.\mu_i$. *Soit* x^* *une solution optimale, on a :*

$$f(\hat{x}) \leq \frac{1+\alpha\Phi^{-1}(1-\varepsilon)}{1+\alpha\Phi^{-1}(1-\varepsilon)/\sqrt{|I|}}.f(x^*).$$

On notera que dans le cas où il n'y a pas d'incertitude ($\alpha = 0$), on retrouve bien le coefficient 1. Idem lorsque l'on considère une seule demande ($|I| = 1$).

Corollaire 1 *Supposons qu'il existe* $\alpha \geq 0$ *tel que, pour toute demande i :* $\sigma_i \leq \alpha.\mu_i$. *Alors router les demandes selon les plus courts chemins au sens des poids* $\{\rho_a\}_{a\in A}$ *est une* $\left(1+\alpha\Phi^{-1}(1-\varepsilon)\right)$ *- approximation de l'optimum.*

Ce résultat indique que router selon les plus faibles coûts marginaux sera une bonne approximation de l'optimum lorsque les incertitudes (i.e. les variances) sont faibles et que la garantie probabiliste recherchée n'est pas trop forte. On donne ci-dessous un tableau de quelques valeurs de ce coefficient d'approximation, selon ε et α :

	$\varepsilon = 0,3$	$\varepsilon = 0,2$	$\varepsilon = 0,1$	$\varepsilon = 0,05$
$\alpha = 0,05$	1,026	1,042	1,064	1,082
$\alpha = 0,10$	1,052	1,084	1,128	1,164
$\alpha = 0,15$	1,079	1,126	1,192	1,247

4 Tests numériques

On utilise les plans coupants de Kelley au sein d'un algorithme de *branch-and-bound* afin de résoudre le problème *CCNDP* à l'optimum. Cet algorithme a été utilisé du fait de sa grande simplicité d'implémentation. Il est évidemment extrêmement rudimentaire par rapport à l'état de l'art de la littérature (on renvoie par exemple à [4, 11, 2]), et peut être amélioré. Le tableau 1 donne quelques résultats numériques préliminaires, obtenus sur un graphe de 50 noeuds et 123 liens bi-directionnels (246 arcs). Dans tous les tests numériques, on considère $\alpha = 5\%$. Cette valeur signifie qu'il n'y a quasiment aucune chance que le trafic dévie de sa valeur moyenne de plus de 15% (3 fois l'écart-type). On observe que l'écart entre l'optimum et un routage aux plus faibles coûts marginaux est effectivement très faible (inférieur à 1%).

nb de demandes	proba. $1-\varepsilon$	sol. heuristique	optimum val.	écart
30	0,8	34388	34227	0,5%
	0,9	38348	38192	0,4%
	0,95	35489	35173	0,9%
40	0,8	45437	45220	0,5%
	0,9	46194	45864	0,7%
	0,95	46819	46395	0,9%

TAB. 1: Résultats numériques

Ce travail préliminaire ouvre de nombreuses perspectives, par exemple pour des analyses qualitatives de la bande passante nécessaire dans les réseaux pour des niveaux de qualité de service donnés. Sur le plan algorithmique, les observations faites ici peuvent servir de base à la résolution de problèmes plus proches de la réalité, par exemple avec des capacités modulaires de dimensionnement.

Références

[1] W. Ben-Ameur et H. Kerivin, Routing of uncertain demands, *Optimization and Engineering*, vol. 3 (2005), pp. 283-313.

[2] Pierre Bonami, Andreas Waechter, Lorenz Biegler, Andrew Conn, Gerard Cornuejols, Ignacio Grossmann, Carl Laird, Jon Lee, Andrea Lodi, Francois Margot and Nicolas Sawaya, An Algorithmic Framework for Convex Mixed Integer Nonlinear Programs, *Discrete Optimization*, vol. 5 (2008), pp. 186-204.

[3] G.M. Guisewite and P.M. Pardalos, Algorithms for the single-source uncapacitated minimum concave-cost network flow problem, *Journal of Global Optimization*, Volume 19, Issue 2 (1991), pp.121-139.

[4] F.S. Hillier, Chance-constrained programming with 0-1 or bounded continuous decision variables, *Management Science*, 14-1 (1967), pp. 34-57.

[5] P. Kall and J. Mayer, *Stochastic Linear Programming*, Springer, 2005.

[6] J. Kilpi and I. Norros, Testing the Gaussian approximation of aggregate traffic, *Proc. of the 2nd ACM SIGCOMM Workshop on Internet measurement* (2002), pp. 49-61.

[7] A. Lisser, A. Ouorou, J.-Ph. Vial et J. Gondzio, Capacity planning under uncertain demand in telecommunication networks, Technical Report 99.13, Logilab, Department of Management Studies, University of Geneva, Switzerland (1999).

[8] E.A. Medova, Chance-Constrained Stochastic Programming for Integrated Services Network Management, *Annals of Operations Research*, Vol. 81 (1998), pp.213-229.

[9] K. Meesublak, Network design under demand uncertainty, *Proc. of the Asia-Pacific Advanced Network Meeting* (2008).

[10] A. Ouorou, Robust capacity assignment in telecommunications, Comput. Manage. Sci. 3 (2006).

[11] A. Weintraub and J. Vera, A cutting plane approach for chance constrained linear programs, *Operations Research*, 39-5 (1991), pp. 776-785.