Détection-Localisation Séquentielle d'Anomalies Volumiques dans un Réseau de Télécommunications

Lionel FILLATRE[1], Igor NIKIFOROV[2], Sandrine VATON[1]

[1]GET - ENST-Bretagne
Technopôle Brest Iroise - 29238 Brest cedex 3 - France

[2]ICD/LM2S - Université de Technologie de Troyes
12, rue Marie Curie - B.P. 2060 - 10010 Troyes cedex - France
lionel.fillatre@enst-bretagne.fr, igor.nikiforov@utt.fr,
sandrine.vaton@enst-bretagne.fr

Résumé – Ce papier s'intéresse à la détection et la localisation d'anomalies dans la distribution spatiale des demandes de trafic Origine-Destination (OD) d'un réseau de télécommunications à partir des volumes mesurés sur les liens du réseau. Puisque la matrice OD n'est pas estimable à partir des volumes des liens, la détection-localisation d'anomalies est un problème de décision statistique mal posé avec des paramètres de nuisances. Le trafic ambiant, assimilé à une matrice de trafic OD correspondante à l'état de fonctionnement nominal du réseau, est inconnu et il est donc considéré comme un paramètre de nuisance. La méthode présentée dans ce papier consiste à 1) proposer un modèle paramétrique parcimonieux du trafic ambiant basé sur des splines polynômiaux ; ce modèle permet de surmonter le caractère mal-posé du problème d'estimation de la matrice OD ; 2) trouver une fonction des volumes de trafic mesurés sur les liens qui est indépendante du trafic ambiant et qui est sensible aux anomalies et 3) détecter et localiser de manière séquentielle des anomalies en utilisant cette fonction des volumes de trafic. L'algorithme de détection-localisation proposé est optimal en ce sens que le délai de décision est minimisé pour une probabilité de fausse alarme et des probabilités de fausses localisations données.

Abstract – This paper proposes to detect and isolate from simple link load measurements the anomalies in the spatial Origin-Destination (OD) distribution of the traffic demand of a given autonomous system. Since the OD traffic matrix is not recoverable from link load measurements the anomaly detection-isolation in OD traffic is an ill-posed decision-making problem with nuisance parameters. The ambient traffic, i.e. the OD traffic matrix corresponding to the non-anomalous network state, is unknown and it is considered here as a nuisance parameter because it can mask the anomalies. The method proposed in this paper consists in 1) designing a parametric description of the ambient traffic by using a spline-based parsimonious model; this model permits us to overcome the obstacles of the ill-posed OD traffic matrix estimation problem; 2) finding a function of link measurements which is independent of the ambient traffic and sensitive to anomalies and 3) sequentially detecting and isolating anomalies by using this function. The proposed detection-isolation algorithm is optimal in the sense that the decision delay is minimized for a given mean false alarm rate and false isolation probabilities.

1 Motivation

Les anomalies de trafic (déni de service, balayage de ports, liens défectueux, etc.) ont des effets négatifs sur la qualité de services des réseaux de télécommunications. Elles peuvent notamment augmenter le temps de transit des données dans le réseau via la congestion des liens de communication. La détection-localisation de ces anomalies est donc un enjeu critique pour les opérateur réseaux qui doivent réagir rapidement et efficacement face à ces anomalies et activer des contre-mesures appropriées.

Un réseau est typiquement composé de plusieurs noeuds et chaque couple de noeuds distincts est susceptible de communiquer et de générer un flot de données dit flot Origine-Destination (OD). Puisqu'il n'existe qu'un nombre limité de liens physiques, chaque noeud du réseau n'est pas directement connecté avec tous les autres noeuds. Les administrateurs du réseau ont donc la charge de router (multiplexer et orienter) ces différents flots OD en respectant un certain niveau de qualité (débit, retard, etc). Malheureusement, puisque le nombre de liens est très nettement inférieur au nombre de flots OD, il est impossible de connaître précisément la quantité de trafic d'un flot OD, ce qui limite considérablement la rapidité de réaction des gestionnaires du réseau face à des événements imprévus. Au contraire, la volumétrie de trafic sur les liens physiques du réseau est mesurée de manière routinière via le protocole Simple Network Management Protocol (SNMP) [1] et il est donc souhaitable de diagnostiquer les anomalies éventuelles du réseau à partir des mesures SNMP.

La détection et la localisation d'anomalies à partir des mesures SNMP est un problème difficile. Bien que largement étudié, ce problème reste toujours d'actualité car il présente plusieurs difficultés majeures. Premièrement, puisque le nombre de mesures SNMP est très nettement inférieur au nombre de flots OD, la connaissance des flots OD à partir des mesures SNMP est un problème inverse fortement mal posé très difficile à résoudre. Deuxième-

ment, une anomalie affectant un flot OD apparaît naturellement sur tous les liens physiques qui routent ce flot OD. Pour identifier le flot OD coupable via les mesures SNMP, il est donc nécessaire d'exploiter les corrélations entre les liens du réseau. Cette phase d'identification reste peu étudiée [7]. Enfin, il n'existe pas actuellement de méthode dont la fiabilité théorique, notamment en terme de fausse alarme et de bonne détection, soit clairement établie.

2 État de l'art et contribution

La détection d'anomalies dans les réseaux est un problème largement abordé dans la littérature [8, 9]. Le problème de localisation des anomalies, ou plus précisément des flots OD qui sont contaminés par ces anomalies, est nettement moins étudié [7]. Dans ces deux cas, la difficulté principale à surmonter est causée par le très grand nombre de flots OD à diagnostiquer par rapport au faible nombre de liens physiques. L'approche qui semble actuellement la plus efficace pour pallier cette difficulté tout en permettant la détection-localisation des flots anormaux s'appuie sur une décomposition PCA (Principal Component Analysis) de la matrice de trafic [7]. Concrètement, il s'agit d'une technique qui permet de réduire la dimensionalité du problème en projetant la matrice de trafic sur un sous-espace de dimension réduite. Ce sous-espace, appelé sous-espace normal, contient l'ensemble des matrices de trafic cohérentes avec la qualité de service souhaitée par l'opérateur (trafic normal). Toutes les autres matrices de trafic possibles sont alors considérées comme anormales puisqu'elle peuvent provoquées une dégradation de la qualité de service. La méthode PCA permet de mettre en évidence que la matrice de trafic peut être décrite par un faible nombre de variables. Cette méthode présente néanmoins plusieurs limites [6] : 1) elle est sensible au choix de la dimension du sous-espace normal ; 2) la création du sous-espace normal nécessite des données historiques exemptes d'anomalies, ce qui n'est pas toujours le cas en pratique. Le sous-espace normal peut alors ne pas être représentatif du trafic normal du réseau ; 3) les performances statistique des algorithmes de détection (probabilités de fausse alarme et de fausses localisations des flots OD anormaux) ne sont pas calculables *a priori* de manière fiable (mais estimées de manière numérique).

Ce papier propose une solution au problème de détection-localisation des flots OD anormaux en surmontant les limites de l'approche PCA. Dans un premier temps, nous introduisons un paramétrage linéaire parcimonieux des flots OD qui permet de limiter considérablement le nombre de paramètres inconnus décrivant le trafic. Ce paramétrage nécessite un faible nombre de données historiques et est peu sensible aux éventuelles anomalies contenues dans ces données historiques. Dans un second temps, nous proposons un algorithme séquentiel récursif asymptotiquement optimal pour détecter et localiser un flot OD anormal qui présente une augmentation, ou une diminution, brutale de trafic à partir d'un instant inconnu t_0 [2]. Quand une anomalie apparaît sur un flot OD, elle est "routée" par le réseau et le principe de détection-localisation exploite les corrélations entre les différents liens du réseau pour détecter et localiser ce flot OD anormal. Les propriétés d'optimalité statistique de cet algorithme sont clairement établies.

3 Position statistique du problème

Soit un vecteur de mesures Y_t obtenu en sortie d'un système $\mathcal{F}$ statique :

$$Y_t = \mathcal{F}(X_t, \boldsymbol{\theta}_t, \boldsymbol{\xi}_t, t)$$

où X_t est un vecteur d'état inconnu (paramètre de nuisance), $\boldsymbol{\xi}_t$ est un bruit aléatoire et $\boldsymbol{\theta}_t$ est le vecteur des paramètres informatifs. Le paramètre $\boldsymbol{\theta}_t$ peut changer à un instant inconnu t_0, dit instant de changement, ce qui conduit naturellement aux deux types d'hypothèses possibles :

$$\text{hypothèse simple :} \quad \boldsymbol{\theta}_t = \begin{cases} \boldsymbol{\theta}_0 & \text{si} \quad t < t_0 \\ \boldsymbol{\theta}_l & \text{si} \quad t \geq t_0 \end{cases},$$

$$\text{hypothèse composite :} \quad \boldsymbol{\theta}_t \in \begin{cases} \Theta_0 & \text{si} \quad t < t_0 \\ \Theta_l & \text{si} \quad t \geq t_0 \end{cases},$$

où $l = 1, \ldots, m$. Les vecteurs $\boldsymbol{\theta}_0, \ldots, \boldsymbol{\theta}_m$ et les ensembles $\Theta_0, \Theta_1, \ldots, \Theta_m$ sont connus. Par exemple, lorsque le système fonctionne dans son mode nominal, le paramètre vectoriel est $\boldsymbol{\theta}_t = \boldsymbol{\theta}_0$. Si une défaillance de type l se produit à l'instant inconnu t_0, le paramètre vectoriel devient $\boldsymbol{\theta}_t = \boldsymbol{\theta}_l$, $l = 1, \ldots, m$, pour $t \geq t_0$. Bien évidemment, les hypothèses composites sont naturellement plus difficiles à traiter que les hypothèses simples. L'algorithme de détection-localisation doit calculer un couple (N, ν) basé sur les observations $(Y_t)_{t \geq 1}$ où N est le temps d'arrêt (alarme) où le changement de type $\nu \in \{1, \ldots, m\}$ est détecté et identifié.

Détaillons le critère d'optimalité retenu pour évaluer la qualité d'un algorithme séquentiel de détection-localisation. Nous considérons uniquement le cas d'hypothèses simples car nous montrerons dans la suite du papier qu'il est possible de ramener le problème de détection-localisation d'anomalies dans un réseau à un problème de choix entre hypothèses simples. Soit une famille de distributions $\mathcal{P} = \{P_l,\ l = 0, \cdots, m\}$ avec les densités $\{f_l,\ l = 0, \cdots, m\}$. Nous observons une suite de variables aléatoires indépendantes $(Y_t)_{t \geq 1}$ de manière séquentielle telles que $Y_1, Y_2, \cdots, Y_{t_0 - 1}$ suivent la distribution P_0 et $Y_{t_0}, Y_{t_0+1}, \cdots$ suivent la distribution P_l. L'instant de changement t_0 et le numéro l sont inconnus (mais non aléatoire). Soit $P^l_{t_0}$ la distribution des observations $Y_1, Y_2, \cdots, Y_{t_0-1}, Y_{t_0}, \cdots$ quand $t_0 = 1, 2, \cdots$ et Y_{t_0} est la première observation qui suit la distribution P_l et $\mathbb{E}^l_{t_0}$ désigne l'espérance mathématique sous $P^l_{t_0}$. Pour des applications exigeant une grande fiabilité, il est nécessaire de garantir des niveaux de fausse alarme et fausse localisation strictes. Un critère convenable et assez aisément manipulable consiste à minimiser le délai moyen maximum de détection-localisation (voir [4, 5]) :

$$\overline{\mathbb{E}}(N) \stackrel{\text{def.}}{=} \sup_{t_0 \geq 1, 1 \leq l \leq m} \mathbb{E}^l_{t_0}(N - t_0 + 1 \mid N \geq t_0) \qquad (1)$$

sous les contraintes :

$$\mathbb{E}_0(N) \geq \gamma,\ \sup_{t_0 \geq 1} \mathbb{P}^l_{t_0}(\nu = j \neq l | N \geq t_0) \leq \beta \qquad (2)$$

pour $1 \leq l, j \neq l \leq m$, où $\mathbb{E}_0(N)$ est le temps moyen avant une fausse alarme et $\sup_{t_0 \geq 1} \mathbb{P}^l_{t_0}(\nu = j \neq l | N \geq t_0)$ correspond à la probabilité de fausse localisation la plus élevée (pire cas). Une borne inférieure asymptotique $n(\gamma, \beta)$ pour le délai moyen maximum de détection-localisation (1)-(2) dans la classe

$$\mathcal{K}_{\gamma,\beta} = \{(N,\nu) : \mathbb{E}_0(N) \geq \gamma, \max_{1\leq l\leq m} \max_{1\leq j\neq l\leq m} \sup_{t_0\geq 1} \mathbb{P}^l_{t_0}(\nu = j | N \geq t_0) \leq \beta\}$$

est donnée par (voir [5]) :

$$n(\gamma,\beta) \stackrel{\text{def.}}{=} \inf_{(N,\nu)\in\mathcal{K}_{\gamma,\beta}} \overline{\mathbb{E}}(N) \gtrsim \max\left\{\log\frac{\gamma}{\rho^*_{\mathrm{d}}}, \log\frac{\beta^{-1}}{\rho^*_{\mathrm{i}}}\right\}$$

lorsque $\min\{\gamma, \beta^{-1}\} \to \infty$, où $\rho^*_{\mathrm{d}} = \min_{1\leq j\leq m} \rho_{j,0}$, $\rho^*_{\mathrm{i}} = \min_{1\leq l\leq m} \min_{1\leq j\neq l\leq m} \rho_{l,j}$ et $\rho_{l,j} = \mathbb{E}^l_1\left(\log\frac{f_l(Y_t)}{f_j(Y_t)}\right)$. Un algorithme optimal récursif (N^*, ν^*), qui atteint asymptotiquement la borne $n(\gamma,\beta)$, est donné par (voir [4, 5]) :

$$N^* = \min_{1\leq l\leq m}\{N(l)\},\ \nu^* = \arg\min_{1\leq l\leq m}\{N(l)\} \qquad (3)$$

$$N(l) = \inf\left\{t \geq 1 : \min_{0\leq j\neq l\leq m} [g_t(l) - g_t(j) - h_{l,j}] \geq 0\right\} \qquad (4)$$

$$g_t(l) = (g_{t-1}(l) + z_t(l))^+ \text{ avec } z_t(l) = \log\frac{f_l(Y_t)}{f_0(Y_t)}, \qquad (5)$$

$g_0(l) = 0$ pour tout $1 \leq l \leq m$, $g_t(0) = 0$ pour tout t et $(\cdot)^+ = \max\{\cdot, 0\}$. Les seuils $h_{l,j}$ sont définis par :

$$h_{l,j} = \begin{cases} h_{\mathrm{d}} & \text{if } 1 \leq l \leq m \quad \text{et } j = 0 \\ h_{\mathrm{i}} & \text{if } 1 \leq l, j \leq m \quad \text{et } j \neq l \end{cases}, \qquad (6)$$

où h_{d} est le seuil de détection et h_{i} est le seuil de localisation. Le choix des seuils h_{d} and h_{i} est discuté (avec des commentaires pratiques et des simulations) dans [4].

4 Modèle de mesures

Considérons un réseau constitué de r noeuds et n liens unidirectionnels $\gamma(\ell)$ où $y_t(\ell)$ désigne le volume de trafic (typiquement en octet par seconde) sur le lien $\gamma(\ell)$ à l'instant t. Soit $x_t(i,j)$ le flot de trafic OD du noeud i au noeud $j \neq i$. La matrice de trafic $X_t = \{x_t(i,j)\}$ est ordonnée dans le sens lexicographique pour obtenir un vecteur colonne $X_t = (x_t(1), \ldots, x_t(m))^T$ où X^T désigne la matrice transposée de X et $m = r(r-1)$. Définissons la matrice de routage A comme la matrice $A = (a(\ell,k))$ de taille $n \times m$ où $0 \leq a(\ell,k) \leq 1$ représente la fraction du volume de trafic du flot OD k qui est routée sur le lien ℓ. Ceci conduit au modèle linéaire : $Y_t = A\,X_t$. Sans perte de généralité, il est admis que la matrice A est de rang n. Le problème fondamental des données SNMP réside dans le fait que $n \ll m$. Par conséquent, l'estimation de la matrice de trafic X_t à partir des mesures SNMP nécessite de résoudre un problème inverse mal-posé. Pour surmonter cette difficulté, le trafic normal moyen du réseau (trafic ambiant) est modélisé comme une combinaison linéaire de fonctions de base connues (des splines polynomiaux) dont les coefficients de pondération sont inconnus, ce qui s'écrit $\mathbb{E}[X_t] = B\boldsymbol{\mu}_t$, où la matrice B de taille $m \times q$ est connue et $\boldsymbol{\mu}_t \in \mathbb{R}^q$ est un vecteur de coefficients inconnus. Les variations temporelles du trafic ambiant moyen sont alors expliquées par les variations temporelles du vecteur $\boldsymbol{\mu}_t$. La matrice résultante $H = A\,B$ est supposée de rang plein colonne. Ainsi, puisque $q < n \ll m$ (modèle parcimonieux), le nombre de paramètres caractérisant le trafic ambiant devient suffisamment petit pour que le problème de détection-localisation admette une solution. Quand le trafic circulant sur le réseau est dépourvu d'anomalie, nous obtenons finalement le modèle linéaire :

$$Y_t = H\boldsymbol{\mu}_t + A\boldsymbol{\xi}_t,$$

où $Y_t = (y_t(1), \ldots, y_t(n))^T$ et $\boldsymbol{\xi}_t$ est un processus AutoRégressif (AR) d'ordre 1 et de paramètre $0 \leq \phi < 1$ connu :

$$\boldsymbol{\xi}_t = \phi\,\boldsymbol{\xi}_{t-1} + \boldsymbol{\varepsilon}_t$$

où $\boldsymbol{\varepsilon}_t \sim \mathcal{N}(0, \Sigma)$ est un bruit blanc Gaussien de matrice de covariance Σ connue, diagonale et de taille $m \times m$.

5 Détecteur-localisateur

Nous souhaitons détecter et localiser une anomalie $\boldsymbol{\theta}$ additive affectant seulement un flot OD. Si, à l'instant t, le niveau du flot l est anormalement élevé (ou bas) par rapport à son niveau normal, alors le vecteur des mesures Y_t est nécessairement contaminé par une anomalie $\boldsymbol{\theta} = \theta\mathbf{a}(l)$ où $\mathbf{a}(l)$ désigne la l-ème colonne normalisée de A et θ désigne l'amplitude inconnue du changement apparu sur le flot OD l. Notre objectif consiste alors à détecter l'anomalie $\boldsymbol{\theta}$ et à déterminer l'indice l du flot OD contaminé le plus rapidement possible. Au moyen d'un pré-filtrage, la composante AR du bruit est aisément éliminée, ce qui conduit au modèle :

$$\begin{aligned} \widetilde{Y}_t &= Y_t - \phi\,Y_{t-1} \\ &= H\tilde{\boldsymbol{\mu}}_t + \boldsymbol{\zeta}_t\ [+\boldsymbol{\theta} = \theta\mathbf{a}(l)], \end{aligned}$$

où $1 \leq l \leq m$, $\tilde{\boldsymbol{\mu}}_t = \boldsymbol{\mu}_t - \phi\,\boldsymbol{\mu}_{t-1}$, $\boldsymbol{\zeta}_t \sim \mathcal{N}(0, A\Sigma A^T)$. Puisque $\tilde{\boldsymbol{\mu}}_t$ est inconnu ($\tilde{\boldsymbol{\mu}}_t \in \mathbb{R}^q$), il est naturel d'assimiler ce paramètre à un paramètre de nuisance qui peut éventuellement masquer certaines anomalies. Le problème est invariant sous l'action du groupe des translations $G = \{g : g(Y) = Y + HC,\ C \in \mathbb{R}^q\}$ (voir une introduction au principe d'invariance dans [3]). La statistique invariante maximale (également appelé "vecteur de parité") $Z_t = W\widetilde{Y}_t$ correspond à la transformation des données mesurées $\widetilde{Y}_t$ en un ensemble de $n - q$ variables linéairement indépendantes par projection sur l'espace noyau de la matrice H. La matrice $W^T = (w_1, \ldots, w_{n-q})$ de taille $n \times (n-q)$ est composée des vecteurs propres $w_1, \ldots, w_{n-q}$ de la matrice de projection $P_H = I_n - H(H^TH)^{-1}H^T$ associés à la valeur propre 1, où A^{-1} est l'inverse de A. La matrice W satisfait les conditions suivantes : $WH = 0$, $W^TW = P_H$, $WW^T = I_{n-q}$. L'utilisation du principe d'invariance associée au fait que la matrice $\widetilde{\Sigma} = WA\Sigma A^TW^T$ est connue nous permet de considérer uniquement la statistique $\widetilde{Z}_t = \widetilde{\Sigma}^{-\frac{1}{2}}W\widetilde{Y}_t$.

Finalement, le problème de détection-localisation se résume à tester les hypothèses

$$\widetilde{Z}_t \sim \begin{cases} \mathcal{N}(0, I_{n-q}) & \text{si } t < t_0 \\ \mathcal{N}(\theta_l \mathbf{v}(l), I_{n-q}), \quad \theta_{l,1} \leq |\theta_l| \leq \theta_{l,2} & \text{si } t \geq t_0 \end{cases} \quad (7)$$

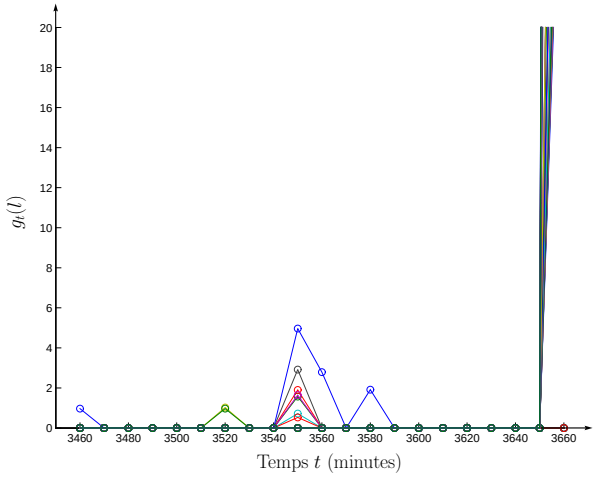
où $\mathbf{v}(l) = \widetilde{\Sigma}^{-\frac{1}{2}} W \mathbf{a}(l)$ est un vecteur connu définissant la direction d'une anomalie de type l dans l'espace des statistiques invariantes. Pour simplifier le problème, puisque l'amplitude θ_l du flot anormal est complètement inconnue, il est admis que cette amplitude est uniformément distribuée sur $[-\theta_{l,2}; -\theta_{l,1}] \cup [\theta_{l,1}; \theta_{l,2}]$. Pour éviter de fausses localisations dans le cas de larges variations de la norme euclidienne des signatures $l \mapsto \|\mathbf{v}(l)\|_2$ sur $[1; m]$, les valeurs $\theta_{l,1}$ et $\theta_{l,2}$ sont choisies telles que $\theta_{l,1}\|\mathbf{v}(l)\|_2 = \bar{\theta}_1$ et $\theta_{l,2}\|\mathbf{v}(l)\|_2 = \bar{\theta}_2$ pour tout $l \in [1; m]$ où $\bar{\theta}_1, \bar{\theta}_2$ sont deux amplitudes fixées en fonction de la sensibilité souhaitée du détecteur-localisateur. Le problème de détection-localisation appliqué à la statistique invariante maximale (7) est alors résolu en utilisant l'algorithme (3)-(6).

6 Évaluation sur données réelles

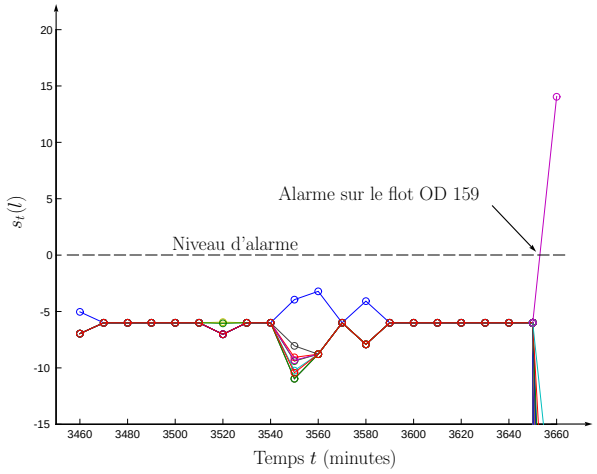
Cette méthode a été évaluée avec des mesures SNMP recueillies sur un grand réseau commercial international composé de 50 noeuds, $n = 168$ liens surveillés et $m = 2450$ flots OD non nuls. Le comportement typique des fonctions de décision $g_t(l)$ et $s_t(l) = \min_{0 \leq j \neq l \leq m}[g_t(l) - g_t(j) - h_{l,j}]$ est illustré sur la figure 1 pour $\phi = 0.95$, $h_{\mathrm{d}} = 6$, $h_{\mathrm{i}} = 6$, $\bar{\theta}_1 = 8.85$ et $\bar{\theta}_2 = 20$. Dans cet exemple, il est admis (du point de vue d'un opérateur humain examinant manuellement les données recueillies) que l'anomalie débute à l'instant 3660. Dès cet instant, plusieurs fonctions $g_t(i, 0)$ croissent rapidement, ce qui traduit l'apparition d'une anomalie. Les fonctions $s_t(i)$ servent alors à "diagnostiquer" les flots OD : lorsque la fonction $s_t(i)$ dépasse 0, le flot OD i est déclaré anormal. Sur la figure 1.(b), seule la fonction $s_t(159)$ associée au flot OD 159 est croissante et dépasse le seuil. Dès l'instant 3660, l'algorithme détecte et localise le flot OD 159 anormal.

7 Conclusion et perspectives

Le problème de détection-localisation rapide d'anomalies volumiques dans un réseau de télécommunications est traité comme un problème de détection-localisation séquentiel en présence de paramètres de nuisances (trafic ambiant). Puisque le nombre de mesures est nettement inférieur au nombre de flots OD, un paramétrage linéaire à base de fonctions splines est proposé pour décrire le trafic normal du réseau. S'appuyant sur le principe d'invariance, le trafic ambiant est éliminé et les anomalies sont détectées.



(a)



(b)

FIG. 1 – Réalisations typiques des fonctions de décision (a) $g_t(l)$ et (b) $s_t(l) = \min_{0 \leq j \neq l \leq m}[g_t(l) - g_t(j) - h_{l,j}]$ en fonction du temps t (en minutes) pour du trafic réel sur un réseau commercial international.

Références

[1] A. Feldmann et al. Deriving traffic demands for operational IP networks : methodology and experience. *IEEE/ACM Trans. Networking*, 9(3) :265–279, 2001.

[2] A. Lakhina et al. Diagnosing network-wide traffic anomalies. In *ACM SIGCOMM*, 2004.

[3] E.L. Lehman. *Testing Statistical Hypotheses, Second Edition*. Chapman & Hall, 1986.

[4] I. Nikiforov. A simple recursive algorithm for diagnosis of abrupt changes in random signals. *IEEE Trans. Inform. Theory*, 46(7) :2740–2746, November 2000.

[5] I. Nikiforov. A lower bound for the detection/isolation delay in a class of sequential tests. *IEEE Trans. Inform. Theory*, 49(11) :3037–3046, 2003.

[6] Haakon Ringberg et al. Sensitivity of PCA for traffic anomaly detection. In *ACM SIGMETRICS*, 2007.

[7] Augustin Soule et al. Detectability of traffic anomalies in two adjacent networks. In *Passive and active measurement conference*, 2007.

[8] Marina Thottan and Chuanyi Ji. Anomaly detection in IP networks. *IEEE Trans. Signal Processing*, 51(8) :2191–2204, 2003.

[9] Y. Zhang, Z. Ge, A. Greenberg, and M. Roughan. Network anomography. In *IMC'05*, 2005.