

# Détection visuelle d'événements dans des grands réseaux d'interaction dynamiques. Application à l'Internet

Bénédicte Le Grand, Matthieu Latapy

LIP6, Université Pierre et Marie Curie - Paris VI  
4 place Jussieu, 75252 Paris, France  
{benedicte.le-grand ;matthieu.latapy}@lip6.fr

**Résumé.** L'objectif des travaux présentés dans ce papier est de faciliter la détection visuelle d'événements dans des réseaux d'interaction dynamiques de grande taille.

Deux méthodes de visualisation classiques et «exhaustives» ont été étudiées, qui représentent l'évolution des liens du réseau au fil du temps. Les limites liées au facteur d'échelle nous ont conduits à proposer deux métaphores restreintes au suivi des nœuds du réseau. Les forces, les limites et la complémentarité de ces quatre métaphores nous ont permis de dégager une ébauche de méthodologie de détection d'événements dans la dynamique de grands réseaux d'interaction.

Les visualisations et la méthodologie présentées dans cet article sont génériques et applicables à tout type de nœuds et de liens ; elles sont ici appliquées pour illustration à un sous-ensemble du réseau Internet.

## 1 Contexte et objectifs

Alors que l'analyse et la visualisation de grands réseaux sont au cœur de nombreux travaux, l'étude de la dynamique de ces réseaux soulève encore de nombreux défis. L'objectif de nos travaux est de faciliter la détection visuelle d'événements (par exemple d'anomalies) liés à la dynamique de réseaux de grande taille. Bien que les métaphores et la méthodologie de visualisation présentées dans cet article soient génériques et applicables à tout type de nœuds et de liens, elles sont ici illustrées sur un cas particulier : un sous-ensemble du réseau Internet. Dans ce contexte, les nœuds du réseau sont des machines et les liens entre les nœuds correspondent aux liens physiques entre ces machines. Les détails de la construction du réseau sont présentés dans la Section suivante.

### 1.1 Réseau d'étude

Le réseau étudié ici est constitué des nœuds et des liens traversés au fil du temps par une machine donnée (désignée comme la *source* dans la suite de cet article) pour atteindre 3000

## Détection visuelle d'événements dans de grands réseaux d'interaction dynamiques

destinations de l'Internet<sup>1</sup>. Pour connaître, à un instant donné, le chemin suivi par les paquets pour aller de la source vers la première destination, le processus est le suivant :

- La source émet un message (plus spécifiquement un *paquet IP*) vers la première destination en limitant à 1 le nombre de routeurs qu'il est autorisé à traverser (grâce au champ *Time To Live* de l'en-tête IP). Le premier routeur traversé, puisqu'il n'est pas autorisé à transmettre le paquet à un autre routeur, envoie alors un message d'erreur à la source, pour lui signifier que le paquet n'a pas pu arriver à destination ; la source connaît ainsi l'adresse du premier routeur traversé (qui se trouve dans le champ *adresse source* de l'en-tête IP du paquet).
- On recommence l'opération en limitant successivement le champ *Time To Live* à 2, 3, ... N ; on obtient ainsi l'adresse des 2<sup>e</sup>, 3<sup>e</sup>, ... N-ième routeurs traversés par les paquets envoyés par la source vers la première destination.

Ce processus est ensuite répété pour chacune des destinations visées (3000 destinations distinctes), permettant d'obtenir une carte du réseau à partir de la source, comme l'illustre la Figure 1.

Cette carte a été recalculée périodiquement, toutes les 15 minutes environ pendant quasiment 3 semaines, 2000 fois en tout (voir (Magnien et al., 2009) pour plus de détails sur la collecte des données et sur l'optimisation du processus pour limiter le nombre de messages à envoyer). Les données étudiées dans ce papier représentent donc 2000 *itérations*, aussi appelées *passes*, qui constituent autant de vues successives de l'évolution de la topologie du réseau au fil du temps.

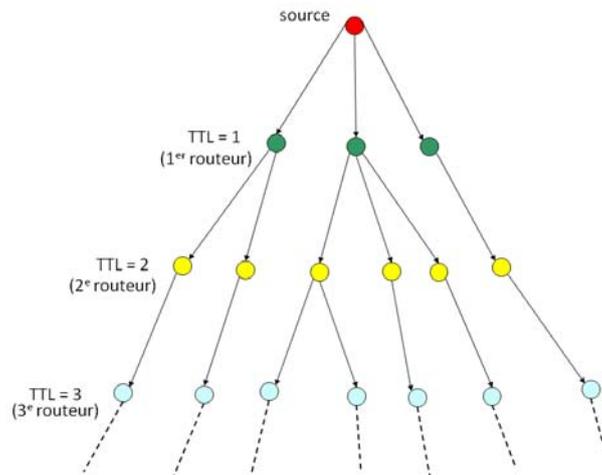


FIG. 1 – Construction du réseau

---

<sup>1</sup> Ces machines destinations appartiennent à la plate-forme *PlanetLab* et sont réparties partout dans l'Internet.

## 1.2 Objectif

L'objectif des travaux présentés dans cet article est de détecter visuellement des évolutions inattendues du réseau : de tels événements sont intrinsèquement liés à la *dynamique* du réseau, dans la mesure où la présence d'un nœud ou d'un lien sur une carte donnée n'est jamais anormale en soi ; c'est l'apparition ou la disparition de nœuds ou de liens qui peut (éventuellement) constituer des événements anormaux dans l'évolution d'un réseau IP. Des modifications soudaines et ponctuelles des routes suivies par les messages peuvent refléter des anomalies, alors que l'alternance entre plusieurs chemins est parfois normale dans le cadre d'un équilibrage de la charge sur le réseau (appelé *load balancing*).

Une première difficulté consiste donc à définir la notion d'événement ; les travaux de (Hamzaoui et al., 2010), appliqués au même type de données, visent à détecter automatiquement les événements « statistiquement significatifs ». Une interprétation est ensuite nécessaire pour identifier ceux qui constituent effectivement des anomalies. Pour cela, les instants particuliers détectés sont étudiés et comparés à des notifications d'erreurs dans le réseau (tickets d'Abilène). Les statistiques peuvent s'avérer peu intuitives pour l'utilisateur final ; par ailleurs, l'intérêt du data mining visuel ayant été prouvé dans de nombreux contextes, l'objectif de ce papier est d'étudier comment la visualisation pourrait faciliter la détection d'événements dans la dynamique de réseaux de grande taille.

Les auteurs de (Teoh et al., 2003) utilisent des métaphores de représentation classiques pour détecter des anomalies de routage du protocole BGP. L'interactivité permet aux utilisateurs d'étudier en détail certains systèmes autonomes (AS) pour interpréter des anomalies qui se sont produites à des instants connus (ce qui n'est pas applicable à nos travaux, puisque l'on ignore quand des événements se sont produits). Ces travaux ont été étendus dans (Teoh et al., 2006), où les visualisations sont couplées à du data mining et permettent d'obtenir des résultats intéressants dans ce contexte spécifique. Dans nos travaux, nous tentons de conserver une approche aussi générique que possible, afin de pouvoir l'appliquer à tout type de réseaux d'interaction (par exemple des réseaux sociaux).

De nombreuses visualisations de réseau (en particulier des tentatives de représentation du réseau Internet) ont été proposées – souvent par des opérateurs ou des fournisseurs d'accès (Akamai), mais leur intérêt est le plus souvent de déceler des tendances globales, car les événements exceptionnels sont masqués par la très grande quantité de données (voir (Claffy et al.) qui décrit *Mapnet*, un outil de visualisation de l'infrastructure des backbones de fournisseurs internationaux). Par ailleurs, il s'agit le plus souvent de représentations statiques qui n'ont pas vocation à être mises à jour fréquemment.

D'autres types de réseaux d'interaction ont récemment suscité un grand intérêt en terme de visualisation ; c'est en particulier le cas de la blogosphère (Discovery, 2007).

A une échelle plus réduite que l'Internet, ce qui est également le cas du réseau étudié ici, des outils de visualisation de graphes tels que *GUESS* (Guess), *Gephi* (Gephi), *Tulip* (Tulip), *Pajek* (Batageli, 2009) ou *Commetrix* (Commetrix), sont extrêmement intéressants. Ces outils permettent en effet de générer des représentations visuelles de qualité (la représentation

## Détection visuelle d'événements dans de grands réseaux d'interaction dynamiques

proposée en section 2.1 utilise d'ailleurs l'outil *GUESS*). Notre objectif n'est pas de développer un nouvel outil, mais une méthodologie pouvant s'appuyer sur des outils existants.

Notre approche consiste donc à étudier des métaphores « naturelles » de visualisation de graphes et à proposer d'autres représentations du réseau. Certaines représentations sont intrinsèquement statiques et leurs différentes instances doivent être comparées les unes aux autres pour retrouver une dynamique (on peut par exemple visionner des cartes successives sous forme de vidéo), alors que d'autres reflètent la dynamique –partielle ou globale– sur une image unique.

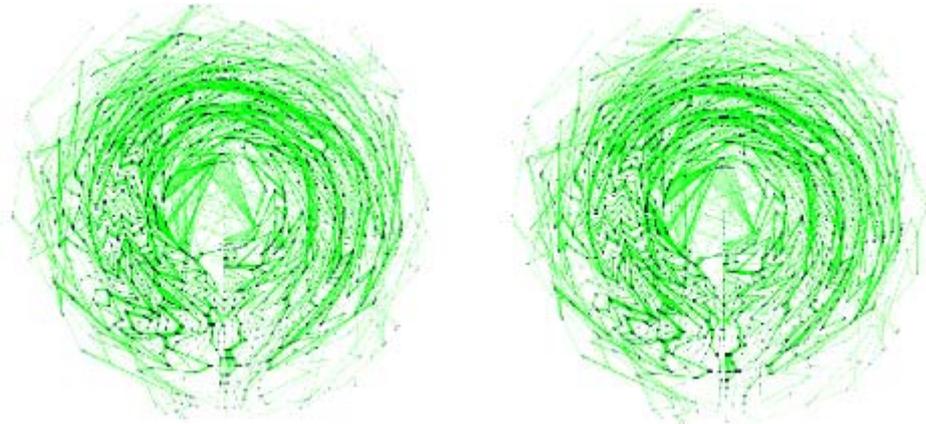
La suite de l'article est organisée comme suit : nous présentons tout d'abord (Section 2) deux méthodes de visualisation classiques et «exhaustives», représentant les liens du réseau (et, pour l'une d'entre elle, les liens et les nœuds). Les limites liées au facteur d'échelle nous conduisent à proposer dans la Section 3 deux métaphores originales et restreintes aux seuls nœuds du réseau. Les forces, les limites et la complémentarité de ces quatre métaphores nous permettent de conclure en proposant une ébauche de méthodologie de détection d'événements dans la dynamique de grands réseaux d'interaction ; nous concluons ce papier en présentant les perspectives d'avenir de ces travaux.

## 2 Visualisation des liens du réseau

### 2.1 Représentation sous forme de graphe (d'arbre)

La visualisation d'un réseau d'interaction sous forme de graphe, constitué des machines du réseau et des liens physiques entre elles, est très naturelle ; de plus, elle constitue une représentation très complète, puisqu'elle permet de visualiser aussi bien les liens que les nœuds du réseau. L'étude des liens permet de détecter un plus grand nombre d'événements que le seul examen des nœuds : en effet, les liens sont forcément modifiés en cas d'apparition ou de disparition de nœuds, mais ils peuvent également évoluer au sein du même ensemble de nœuds (ce qui est indétectable si l'on se limite à l'étude des apparitions et disparitions des nœuds au fil du temps).

Dans le cas du réseau étudié ici, le graphe construit à chacune des 2000 itérations est en réalité un arbre (ce qui est lié à la manière dont le réseau est construit, comme expliqué dans la Section 1.1). Cette propriété particulière est mise à profit pour notre visualisation : la source est placée au centre et ses voisins immédiats sont disposés sur un cercle autour d'elle. Les nœuds suivants sont placés sur un cercle de rayon plus important etc. Le rayon du cercle sur lequel se trouve un point indique la distance entre ce nœud et la source (en nombre de sauts).



**FIG. 2** – *Arbre à 2 itérations successives*

L'étude visuelle de la dynamique du graphe consiste alors en la visualisation (et la comparaison) des graphes correspondant aux cartes successives du réseau. Pour obtenir ces graphes (ici des arbres comme nous l'avons expliqué plus haut), les coordonnées de la totalité des nœuds rencontrés au cours des 2000 itérations ont été fixées<sup>2</sup>, et seuls les nœuds et liens présents à une itération donnée sont affichés sur l'arbre correspondant. L'outil *GUESS* a été utilisé pour le dessin du graphe (mais les coordonnées ont été pré-calculées en amont). La Figure 2 illustre l'évolution entre deux itérations ; l'objectif est par exemple de détecter des portions d'arbre qui apparaissent ou disparaissent d'une itération à une autre, traduisant des modifications de routage.

Une première difficulté liée à cette méthode est le choix de la position de chaque nœud, qui a un impact important sur la lisibilité de la visualisation (par exemple à cause des arcs qui se coupent).

Par ailleurs, à moins que les changements de route ne se produisent tout près de la machine source, il est très difficile de distinguer des changements significatifs d'un arbre à un autre, comme le montre la Figure 2. Cette métaphore de visualisation, pourtant naturelle pour représenter une structure arborescente, n'est pas satisfaisante car elle introduit un biais important : les modifications proches de la source sont beaucoup plus facilement observables puisqu'elles modifient une grande portion de l'arbre, à l'inverse des changements concernant les nœuds périphériques.

Cette méthode de visualisation n'est donc envisageable que pour une approche locale ; elle ne pourra être utilisée qu'une fois qu'un sous-ensemble de nœuds et un sous-ensemble d'itérations auront été sélectionnés, c'est-à-dire lorsque l'on saura « où » (quels nœuds) et « quand » (quelles itérations) regarder.

---

<sup>2</sup> Chaque carte instantanée comprend en moyenne 12000 nœuds. Le nombre total de nœuds rencontrés pendant la période étudiée (2000 itérations) est de l'ordre de 20000.

### Approche différentielle

Dans la visualisation étudiée précédemment, chaque carte (arbre) correspond à la vision statique du réseau à un instant donné. Il est possible de refléter un peu de dynamique sur une carte unique en adoptant une représentation que nous qualifions de *différentielle*. Pour une itération donnée, on représente les nœuds qui sont « apparus », c'est-à-dire les nœuds qui étaient absents pendant les  $n$  itérations précédentes<sup>3</sup>, et qui sont présents à l'itération courante. Il est possible d'ajouter une contrainte supplémentaire en imposant que ces nœuds soient présents pendant  $m$  itérations à partir de l'itération courante (cela permet de distinguer des nœuds apparus de façon très ponctuelle de ceux qui restent présents à partir du moment où ils sont apparus). Réciproquement, on peut s'intéresser aux nœuds qui ont « disparu » à un instant donné (en utilisant également les paramètres  $n$  et  $m$ ).

Cette représentation différentielle permet de mettre en avant les apparitions et disparitions de nœuds et de liens, et donc de capturer une certaine dynamique sur une représentation unique. Les paramètres  $n$  et  $m$  sont respectivement appelés *marge arrière* et *marge avant*.

L'approche différentielle est adaptée pour mettre en évidence une dynamique régulière telle que l'alternance entre deux routes pour l'équilibrage de charge sur le réseau : il suffit pour cela de fixer la marge arrière à 1 (on regarde les nœuds absents à l'itération précédente) et la marge avant à 0 (nœuds présents à l'itération courante) : seuls les nœuds apparus à l'itération courante et absents à l'itération précédente seront affichés. Par contre, les événements « imprévisibles » sont difficilement repérables puisque les valeurs à choisir pour ces deux paramètres ne sont pas connues.

## 2.2 Matrice d'adjacence

Une première simplification de la métaphore précédente consiste à représenter uniquement les liens (et non plus les liens et les nœuds). La représentation choisie pour cela est la matrice d'adjacence : il s'agit d'une matrice carrée dont les lignes et les colonnes correspondent aux machines rencontrées lors d'au moins une itération sur la période étudiée. Pour une itération donnée, un point est affiché aux coordonnées  $[i, j]$  de la matrice s'il existe un lien entre le nœud  $i$  et le nœud  $j$  à cette itération. Cette représentation est illustrée sur la Figure 3. Tous les nœuds rencontrés au cours des 2000 itérations sont représentés dans la matrice, afin de pouvoir repérer des apparitions ou des disparitions de liens.

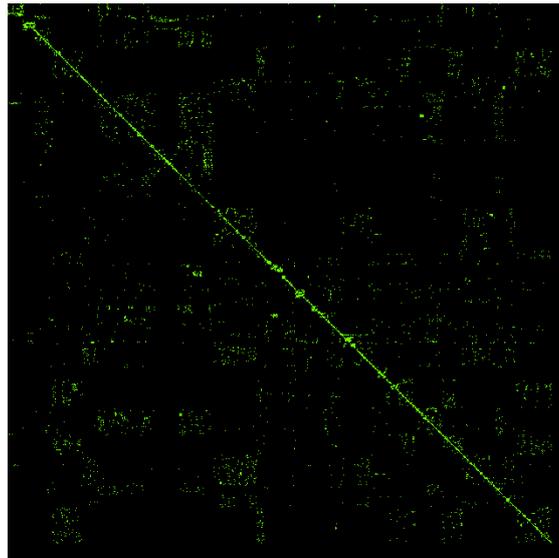
Une anomalie de routage peut être supposée lorsqu'une forte modification du nombre de nœuds et/ou de liens est observée entre deux itérations successives (visible par exemple lorsque l'on fait défiler les images les unes après les autres dans une vidéo<sup>4</sup>). Des tels changements se traduisent par l'apparition et la disparition soudaines de points dans la matrice. Ils sont plus aisément détectables si des « blocs » de liens apparaissent ou disparaissent, ce qui n'est visible que si ces liens sont regroupés au sein de la matrice d'adjacence.

---

<sup>3</sup>  $n$  est un paramètre à ajuster selon le type d'événement que l'on cherche à mettre en évidence.

<sup>4</sup> Voir la vidéo située à l'url [http://www.complexnetworks.fr/videos.php?video\\_id=24](http://www.complexnetworks.fr/videos.php?video_id=24).

L'ordre des nœuds dans la matrice est donc important car c'est lui qui rend possible l'apparition de tels blocs. L'objectif est donc de regrouper dans une même zone les machines qui sont proches en terme de topologie de réseau, afin de voir des blocs de liens qui apparaissent ou disparaissent d'une itération à une autre. Le classement des nœuds dans l'ordre croissant de leur adresse IP semble adapté, dans la mesure où les machines d'un même sous-réseau se trouvent ainsi groupées dans la matrice d'adjacence.



**FIG. 3** – *Matrice d'adjacence*

Comme la représentation sous forme de graphe, cette méthode a l'avantage de s'intéresser aux liens entre les machines et pas seulement aux machines elles-mêmes.

Cette représentation permet de faire la distinction entre l'équilibrage de charge, qui se traduit par des clignotements, et les changements brutaux de route, visibles par l'apparition de blocs blancs ou noirs (traduisant les apparitions et les disparitions de liens). Comme nous l'avons expliqué précédemment, ces blocs ne sont visibles que si les liens qui apparaissent ou disparaissent se situent dans une même zone de la matrice, et donc si l'ordre des nœuds est adapté.

Le problème de la taille de la matrice d'adjacence se pose, puisque le nombre total de nœuds distincts présents au cours d'au moins une itération sur les 2000 est d'environ 20000. Afin de pouvoir représenter cette matrice sous une forme plus synthétique, une solution consiste à agréger sur un unique pixel plusieurs valeurs à l'intérieur d'un intervalle d'adresses IP: dans ce cas, un point présent dans la matrice indique qu'au moins un lien existe entre deux machines quelconques de cet intervalle. Cette représentation synthétique implique néanmoins une perte d'information et dépend de la manière dont les coupures sont effectuées entre les différents intervalles d'adresses IP.

## Détection visuelle d'événements dans de grands réseaux d'interaction dynamiques

Par conséquent, la matrice d'adjacence est elle aussi préconisée pour une représentation locale de la topologie du réseau, pour laquelle on peut se limiter à l'observation d'une portion de la matrice. Comme pour la représentation sous forme de graphe (d'arbre), cela implique en particulier de savoir « où » regarder dans la matrice (i.e. quelles machines).

On peut noter que l'approche différentielle est également possible avec cette métaphore : dans ce cas, un point présent dans la matrice à une itération donnée indique l'apparition d'un lien qui n'existait pas lors des précédentes itérations (ou l'inverse si l'on s'intéresse aux liens qui ont disparu). Les avantages et les inconvénients de cette visualisation différentielle sont les mêmes que pour la représentation d'arbre différentiel proposé dans la Section 2.1.

Bien que très complètes, les visualisations qui représentent les liens entre les machines posent des problèmes en terme d'échelle. Le nombre de liens potentiels entre les 20000 machines observées au cours d'au moins une itération est trop important pour qu'ils puissent être visualisés intégralement, comme nous l'avons vu précédemment. Dans la Section suivante, nous proposons deux métaphores originales, limitées à la représentation des nœuds du réseau.

### 3 Visualisation des nœuds du réseau

#### 3.1 Courbe de Hilbert

L'objectif de cette représentation est de montrer sous une forme synthétique et affichable sur un écran<sup>5</sup>, les machines présentes dans le réseau à une itération donnée. Pour cette visualisation, tous les nœuds rencontrés au cours de la mesure sont ordonnés sur une droite par ordre croissant de leur adresse IP. Afin de passer à une représentation en 2D tout en garantissant que des nœuds proches sur la droite restent proches sur la visualisation, les coordonnées des points sont fixées à l'aide d'une courbe de Hilbert (Figure 4).

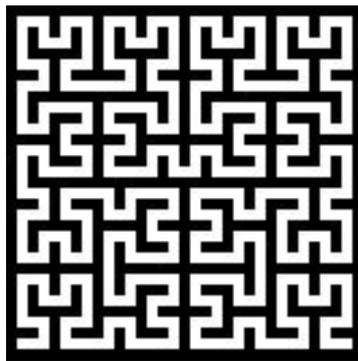


FIG. 4 – Principe des courbes de Hilbert

---

<sup>5</sup> Il s'agit en effet de la principale limite des matrices d'adjacence.

Cette représentation est beaucoup plus synthétique que la matrice d'adjacence, puisque l'on représente uniquement –par des points blancs- les nœuds présents à une itération donnée, au lieu de représenter les liens (Figure 5<sup>6</sup>). Cette métaphore est donc adaptée à une représentation de tous les nœuds. Il s'agit cependant, comme pour les autres visualisations étudiées jusqu'ici, de vues instantanées ; il faut, là encore, savoir « où » regarder (i.e. quelle portion de l'image) et « quand » (quelles itérations). Le repérage d'anomalies dans la dynamique du réseau peut s'effectuer en visionnant les images successivement et en essayant de voir apparaître des blocs d'apparitions ou de disparitions de nœuds. Cependant, l'ordre des nœuds a, là encore, un fort impact sur l'existence de blocs blancs ou noirs et donc sur la détection de modifications du routage.

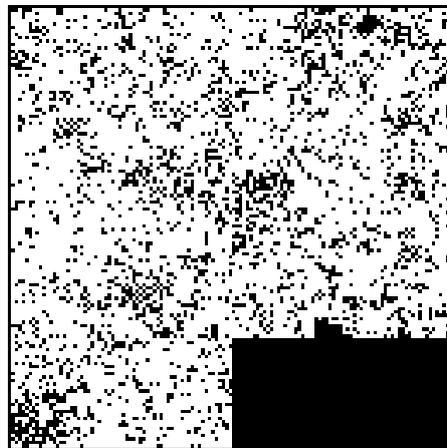


FIG. 5 – Visualisation « Hilbert »*t*

L'approche différentielle est aussi possible avec ce type de représentation : un point sur la courbe de Hilbert est représenté si le nœud correspondant est présent à cette itération alors qu'il était absent aux précédentes (ou inversement).

Les approches différentielles permettent de mettre en évidence la dynamique des nœuds sur une représentation statique (pour une itération donnée). Le paramétrage des marges avant et arrière est cependant difficile. L'idéal serait de représenter la dynamique de toutes les itérations sur une seule représentation statique ; ceci est l'objet de la métaphore présentée dans la Section suivante, dite en « code barre ».

---

<sup>6</sup> Le nombre total de machines rencontrées pendant la période de mesure n'étant pas une puissance de 2, la Figure 5 comporte une partie « vide » qui correspond au rectangle situé en bas à droite.

### 3.2 Code barre

La visualisation que nous qualifions de *code barre* (voir les Figures 6 et 7 pour comprendre l'origine de ce nom !) fournit une représentation globale dans le temps, c'est-à-dire qu'elle permet de représenter toutes les itérations sur une même vue.

Sur cette visualisation, la présence d'un point blanc aux coordonnées  $(i, j)$  signifie que le nœud  $j$  est présent à l'itération  $i$ . Comme c'était le cas pour la matrice d'adjacence, cette représentation de 20000 (nœuds) sur 2000 (itérations) est trop volumineuse pour tenir sur un écran ; néanmoins, il est inutile de regarder des images successives pour observer des apparitions ou des disparitions de nœuds (ou idéalement de blocs de nœuds, qui sont plus facilement repérables que des nœuds isolés), ce qui constitue un avantage par rapport à la matrice d'adjacence.



FIG. 6 – Problème durable

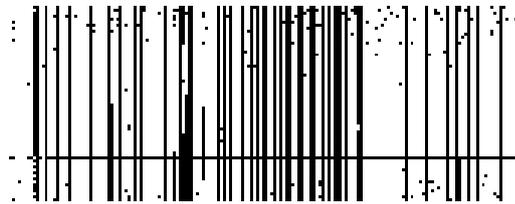


FIG. 7 – Problème ponctuel

Cette représentation présente donc un gros avantage sur les métaphores étudiées précédemment : elle capture sur une seule image la dynamique du réseau sur toute la période étudiée. Il s'agit par conséquent de la représentation la mieux adaptée pour identifier les itérations et les machines concernées par des modifications de routage : elle peut permettre de répondre aux questions « où » et « quand » regarder.

Bien que très prometteuse, cette visualisation présente malgré tout quelques faiblesses. Une première limite est liée à sa dépendance à l'ordre des nœuds, comme c'était le cas pour les matrices d'adjacence et pour les courbes de Hilbert.

Une autre faiblesse est liée à la taille importante de l'image générée (de l'ordre de 20000 \* 2000), qui ne permet pas de l'afficher dans son intégralité à une résolution suffisante ; il faut par conséquent s'appuyer sur des techniques automatiques pour repérer des blocs dans les données (qui correspondent à un problème relativement durable, comme l'illustre la Figure 6) ou des lignes horizontales (traduisant un problème ponctuel comme le montre la Figure 7).

## 4 Conclusion et perspectives

Dans cet article, nous avons analysé et comparé plusieurs métaphores de visualisation – classiques ou originales- visant à faciliter la détection d'événements (par exemple d'anomalies) dans des réseaux dynamiques de grande taille. Le réseau particulier étudié ici est un sous-ensemble de l'Internet, constitué de machines et de routeurs, dont l'évolution a été suivie pendant 3 semaines.

L'analyse des forces et faiblesses de ces représentations nous ont conduits à la conclusion suivante quant à l'ébauche d'une méthodologie de détection d'événements dans la dynamique des grands réseaux d'interaction :

- La visualisation en « code barre » est la plus complète du point de vue de la dynamique ; il s'agit donc de la métaphore à utiliser en première intention pour repérer les moments et/ou les machines affectées par des événements inhabituels.
- Les représentations matricielle et sous forme de courbe de Hilbert sont également susceptibles de mettre en évidence des événements, en particulier si un sous-ensemble d'itérations a été identifié par la méthode précédente.
- La visualisation sous forme d'arbre (spécifique au réseau étudié) est la moins adaptée car l'impact d'une modification du routage est extrêmement variable en fonction de l'endroit où elle se produit –en terme de distance à la source.

De nombreuses perspectives s'ouvrent à la visualisation pour la recherche d'événements dans la dynamique des réseaux :

- Il apparaît dans quasiment toutes les métaphores étudiées ici que la notion d'ordre sur les nœuds du réseau a un fort impact sur la mise en évidence de changements dans la dynamique d'un réseau. Un ordre naturel existe pour les adresses IP puisqu'il s'agit de valeurs numériques –même si cet ordre n'est pas forcément significatif lorsque l'on dépasse l'échelle du sous-réseau- mais ça n'est pas le cas pour tous les réseaux (par exemple les réseaux sociaux). Une première perspective consiste donc à généraliser cette notion d'ordre au sein des nœuds d'un réseau d'interaction.
- L'utilisation conjointe de méthodes statistiques et de visualisation représente la deuxième piste que nous souhaitons suivre pour approfondir les travaux présentés dans cet article.

## Références

- Akamai, Visualizing the Internet, [http://www.akamai.com/html/technology/visualizing\\_akamai.html](http://www.akamai.com/html/technology/visualizing_akamai.html).
- Assia Hamzaoui, Matthieu Latapy and Clémence Magnien, Detecting Events in the Dynamics of Ego-centered Measurements of the Internet Topology, *Proceedings of International Workshop on Dynamic Networks (WDN)*, in conjunction with WiOpt 2010.
- V. Batagelj, Visualization of Large Networks, Chapter in *the Encyclopedia of Complexity and System Science* (editor in chief Bob Meyers), in the Complex Networks section (section editor Geoffrey Canright), Springer Verlag, 2009.
- K. Claffy, B. Huffaker, Macroscopic Internet visualization and measurement, Caida Mapnet tool, <http://www.caida.org/tools/visualization/mapnet/summary.html>.
- Commetrix, Dynamic Network Visualization & Analysis, <http://www.commetrix.de/>
- Welcome To The Blogosphere, Discovery Magazine, May 2007, <http://discovermagazine.com/2007/may/map-welcome-to-the-blogosphere>
- Gephi, the Open Graph Viz Platform, <http://gephi.org/>
- GUESS, the Graph Exploration System, <http://graphexploration.cond.org>
- Clémence Magnien, Frédéric Ouedraogo, Guillaume Valadon, Matthieu Latapy, Fast dynamics in Internet topology: preliminary observations and explanations, *Fourth International Conference on Internet Monitoring and Protection (ICIMP 2009)*, May 24-28, 2009, Venice, Italy.
- Soon-Tee Teoh, Kwan-Liu Ma, S. Felix Wu, Dan Massey, Xiaoliang Zhao, Dan Pei, Lan Wang, Lixia Zhang, and Randy Bush, Visual-based Anomaly Detection for BGP Origin AS Change Events, Proc. of the 14th IFIP/IEEE Workshop on Distributed Systems: Operations and Management, 2003.
- Soon Tee Teoh, Supranamaya Ranjan, Antonio Nucci, and Chen-Nee Chuah, BGP Eye: A New Visualization Tool for Real-time Detection and Analysis of BGP Anomalies, Proc. of the ACM Conf. on Computer and Communications Security Workshop on Visualization for Computer Security, 2006.
- TULIP, Graph visualization Software, <http://tulip.labri.fr/>

## Summary

This paper aims at enabling the visual detection of unexpected « events » in large dynamic networks. Two exhaustive visualization metaphors are studied, which represent the evolution of links over time. Due to scalability issues, we propose two original metaphors restricted to the nodes of the network. The strengths, weaknesses and complementarity of these 4 approaches allow us to propose the first steps of a methodology for the visual detection of anomalies in large dynamic interaction networks.

The methodology and visualizations presented in this article are generic and may be applied to any type of nodes and links; they are illustrated here on a subset of the Internet.