

INTRODUCTION

L'intérêt de la visualisation pour l'extraction de connaissances n'est plus à démontrer, et de nombreuses méthodes ont déjà été proposées. Des progrès sont néanmoins toujours possibles et nécessaires pour générer des visualisations interactives et adaptées aux besoins des utilisateurs. La question de l'évaluation des représentations générées et plus globalement des méthodes utilisées est également toujours d'actualité. Le problème de l'évaluation est plus complexe dans le cas des méthodes interactives que dans le cas des méthodes automatiques de fouille de données. En effet, en plus de la disponibilité des ensembles de données et de leurs tailles, se pose le problème de la reproduction à l'identique des expérimentations : quel public et quelles expérimentations choisir pour qu'une méthode soit validée ?

L'atelier Extraction et Visualisation des Connaissances, associé à la conférence EGC (Extraction et Gestion des Connaissances), est, depuis 2003, un lieu d'échange pour les chercheurs francophones de ce domaine.

Ce numéro spécial rassemble les versions étendues et actualisées des meilleurs papiers présentés au cours des éditions 2009 à 2011 de l'atelier Visualisation et Extraction des Connaissances d'EGC.

Les progrès réalisés par les outils de visualisation permettent à présent d'envisager leur utilisation pour observer de gros volumes de données, et les grands réseaux qui nous entourent sont pour cela des candidats très intéressants. Pour favoriser le passage à l'échelle, il est néanmoins nécessaire d'exploiter les propriétés de ces grands graphes, et en particulier leur haute densité locale, qui se traduit par l'existence d'une structure en communautés. Trois articles de ce numéro sont consacrés à la visualisation de communautés, reposant sur différents modes de construction de ces communautés (sur des critères purement topologiques, topologiques et structurels, conceptuels).

Le quatrième papier est également dédié à la visualisation de grands réseaux, appliquée à la sécurité, puisqu'il s'agit de détecter visuellement des événements anormaux dans la topologie d'un sous-ensemble de l'Internet.

Comme nous l'avons mentionné en introduction, la création et l'évaluation d'applications visuelles interactives est toujours un défi, auquel est consacré le dernier article de ce numéro.

Visualisation de grands réseaux et de leurs communautés

1. Visualisation interactive multi-échelle des grands graphes. Application à un réseau de blogs, Massoud Seifi (Communautés construites sur des critères topologiques)

De nombreux réseaux du monde réel peuvent être modélisés par des grands graphes. Réduire la complexité d'un graphe de manière à ce qu'il puisse être

facilement interprété par l'œil humain est une aide précieuse pour comprendre et analyser ce type de données. Nous proposons une méthodologie de visualisation interactive multi-échelle de grands graphes basée sur une classification hiérarchique des sommets qui nous permet de représenter ces graphes de manière lisible et interprétable. Nous appliquons notre méthodologie à un réseau de blogs francophones.

2. Détection et visualisation des communautés dans les réseaux sociaux, Juan David Cruz, Cécile Bothorel, François Poulet (Communautés construites sur des critères topologiques et sémantiques)

Les réseaux sociaux contiennent trois types de variables : structurelles, en représentant les interactions entre acteurs, de composition ou sémantique, qui décrivent les caractéristiques de chaque acteur, et d'affiliation, qui sont utilisées pour représenter l'appartenance des nœuds aux différents groupes. À l'heure actuelle la plupart des approches d'analyse des réseaux sociaux n'utilisent que l'information structurelle et d'affiliation, en perdant les descriptions sémantiques de chaque acteur. Cet article propose un modèle d'intégration de l'information provenant d'un réseau social qui permet d'identifier et de visualiser les communautés avec des critères sémantiques et structurels.

3. Extraction et visualisation de structures arborescentes à partir de treillis de concepts, Cassio Melo, Marie-Aude Afaure, Bénédicte Le Grand (Communautés construites sur des critères conceptuels)

Les logiciels classiques produits dans le contexte de l'analyse formelle de concepts font peu d'usage des techniques de visualisation, et produisent des représentations de treillis de concepts peu lisibles lorsque le nombre des concepts de ces treillis dépasse quelques dizaines. Ceci est problématique, dans la mesure où le nombre de concepts dans ces treillis croît significativement avec la taille des données et leur nombre de dimensions. Dans cet article, nous proposons plusieurs méthodes pour améliorer la lisibilité des treillis de Galois : tout d'abord par l'extraction et la visualisation d'arbres dérivés des structures conceptuelles en treillis, et deuxièmement par l'utilisation de techniques colorimétriques et de distorsion. Ces contributions constituent un pas important dans l'analyse visuelle de structures conceptuelles, puisqu'elles permettent à des experts du domaine d'explorer visuellement de plus grands ensembles de données, que les visualisations traditionnelles de treillis de concept ne peuvent pas représenter de manière efficace.

Visualisation de grands réseaux - Application à la sécurité

4. Visualisation de grands réseaux dynamiques - Application à la détection d'anomalies de routage dans l'Internet, Alice Albano, Bénédicte Le Grand, Matthieu Latapy

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 pour représenter l'évolution des liens d'un réseau au fil du temps. Les limites liées au facteur d'échelle nous ont amenés à proposer deux métaphores originales, restreintes

au suivi des nœuds du réseau. Les forces, les limites et la complémentarité de ces quatre représentations nous ont permis de dégager une ébauche de méthodologie de détection visuelle 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 illustrées sur un sous-ensemble du réseau internet.

Méthode de création et d'évaluation de visualisations

5. *Création et évaluation d'interfaces visuelles interactives à partir de services web*, Romain Vuillemot

Cet article présente une architecture basée sur les services web permettant la création et l'évaluation d'applications visuelles interactives. Les services web sont une standardisation de l'échange de données dans un système distribué, tel que le web. Ils servent principalement à la publication de données (site web, API), mais peuvent également servir au traitement de celles-ci. Nous montrons leur application à la création de représentations visuelles de données, par combinaison de services, afin de recomposer le modèle de référence de traitement des données. Les visualisations ainsi générées peuvent être rendues interactives, et leur combinaison permet de réaliser des tâches d'exploration et d'analyse visuelle de données. Ces services web produisant systématiquement des données d'usage (logs) nous donnons des éléments d'analyse de l'activité utilisateur à des fins de validation et d'évaluation de cette activité.

BÉNÉDICTE LE GRAND

Université Paris 1 Panthéon-Sorbonne

FRANÇOIS POULET

Université de Rennes 1, IRISA

Nous remercions vivement les relecteurs pour la qualité de leurs commentaires.

RELECTEURS

Michaël Aupetit – CEA LIST, Gif-sur-Yvette

Nadir Belkhiter – Université de Laval, Québec

Fanny Chevalier – OCAD University/University of Toronto, Canada

Thanh-Nghi Do – Université de Can Tho, Vietnam

Jean-Daniel Fekete – INRIA Saclay, Ile-de-France

Michel Soto – Université Paris 5, Paris

