



HAL
open science

Arabesque, application d'exploration et de géovisualisation de données de flux et de réseaux

Etienne Côme, Thomas Bapaume, Laurent Jégou, Françoise Bahoken,
Grégoire Le Campion, Nicolas Roelandt

► To cite this version:

Etienne Côme, Thomas Bapaume, Laurent Jégou, Françoise Bahoken, Grégoire Le Campion, et al.. Arabesque, application d'exploration et de géovisualisation de données de flux et de réseaux. SAGEO 2019, Nov 2019, Clermont-Ferrand, France. hal-02299349

HAL Id: hal-02299349

<https://hal.science/hal-02299349>

Submitted on 23 Jan 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Proposition pour une démonstration de *Arabesques*

**Côme E.¹, Bapaume T.¹, Jégou L.², Bahoken F.^{3,4}, Maisonobe M.⁴,
Roelandt N.⁵, Le Campion G.⁶**

1. COSYS-GRETTIA, IFSTTAR / Université Paris-Est
Cité Descartes, 14-20 Boulevard Newton, F-77420 Champs-sur-Marne, France.
etienne.come@ifsttar.fr
2. UMR 5193 LISST / Université de Toulouse 2 / CNRS
Maison de la Recherche, 5, allées Antonio-Machado, F-31058 Toulouse Cedex 9,
jegou@univ-tlse2.fr
3. AME-SPLOTT, IFSTTAR / Université Paris-Est
Cité Descartes, 14-20 Boulevard Newton, F-77420 Champs-sur-Marne, France
francoise.bahoken@ifsttar.fr
4. UMR 8504 Géographie-Cités / Équipe PARIS / CNRS
13, rue du Four, F-75006 Paris, France
marion.MAISONOBE@cnrs.fr
5. AME, IFSTTAR / Université Paris-Est
Cité Descartes, 14-20 Boulevard Newton, F-77420 Champs-sur-Marne, France
nicolas.roelandt@ifsttar.fr
6. UMR 5319 Passages, Équipe ARD / CNRS
Maison des Suds, 12 Espl. des Antilles, F-33600 Pessac, France
Gregoire.LECAMPION@cnrs.fr

RÉSUMÉ. Le projet gflowiz (geographic flow visualisation) est une collaboration interdisciplinaire réalisée dans l'objectif de concevoir et développer Arabesques, une nouvelle application web dédiée à la géo visualisation de données de flux et de réseaux spatiaux. Cette proposition présente différentes possibilités d'actions sur les informations géographiques, statistiques et cartographiques qui concourent à la réalisation de cartographies correspondantes.

ABSTRACT. The gflowiz (geographic flow visualisation) project is an interdisciplinary collaboration carried out with the objective of designing and developing Arabesques, a new web application dedicated to the geo visualization of flow datasets and spatial networks. This proposal presents different possibilities for action on geographical, statistical and cartographic information that contribute to dedicated maps.

MOTS-CLÉS : matrice OD, hyper matrice, graphe multiplexe, geoweb, cartographie.

KEYWORDS: origine-destination matrix, hyper matrix, multiplex graph, geoweb, flow mapping.

1. Introduction

Le projet gflowiz (*geographic flow visualisation*) est une collaboration interdisciplinaire financée par l'Institut français des sciences et technologies des transports, de l'aménagement et des réseaux (IFSTTAR). Réalisée en collaboration avec le Centre national de la recherche scientifique (CNRS), elle a pour objectif de concevoir et développer une nouvelle application web dédiée à la géo visualisation de données de flux et de réseaux spatiaux, *Arabesques*.

Cette application, inscrite dans le cadre de la Loi Lemaire pour une république numérique, relève de l'objectif général visant à accroître la compréhension des déterminants géographiques de la mobilité spatiale de biens et de personnes. Elle souhaite répondre à l'impératif de visualisation des résultats de recherches fondamentales ou appliquées relevant de cadres de développements théoriques et méthodologiques pouvant être considérés comme transversaux, à différents thèmes (population, habitat, environnement, transport, ...), interdisciplinaires par essence (géographie, démographie, environnement, géomatique, sciences pour l'ingénieur, sciences humaines et sociales, ...) en raison des défis sociétaux liés à l'essor de la société numérique qu'ils permettent de relever.

Sur le plan scientifique, cette application ambitionne d'innover dans la manipulation de données numériques décrivant des flux et réseaux localisés. C'est pourquoi elle s'inscrit, d'une part, dans le nouveau paradigme de la « cartographie de visualisation » (Mac Eachren, 2004) consistant à combiner dans un même environnement des dispositifs de géo visualisation et de traitement des informations géographiques et statistiques – à visée descriptive dans un premier temps. En autorisant, d'autre part, la prise en compte de jeux de données relationnelles complexes, pouvant être à la fois volumineux et présenter différentes dimensions des mobilités spatiales (plusieurs catégories thématiques et/ou temporalités). Une attention particulière est en outre apportée au rendu en lui-même, que ce soit pour le dessin ou pour la sémiologie cartographique des figurés. La volonté d'une amélioration qualitative des images produites devrait conduire à développer une approche sensée de la géo visualisation des mobilités et des interactions spatiales.

L'objet de cette démonstration est de présenter le fonctionnement de la toute première version de *Arabesques*, en appui de l'atelier *Le Geoweb des cartes de flux et réseaux* réalisé en marge de la conférence SAGEO'2019 pour le compte de l'Action Prospective *Théoriser, Observer, Analyser, Modéliser le Geoweb* (GEOWEB) du GdR MAGIS.

Le développement présenté fait suite à une première phase qui a donné lieu à un état de la question réalisé sur les possibilités actuellement offertes de manipulation de données relationnelles à des fins d'analyse cartographique (Bahoken et al., 2019). L'application relève – à ce stade de développement – de la formation du premier des trois *clusters* identifiés par une typologie de l'existant. Il est composé d'applications

dont la représentation principale est la *carte de flux*, réalisée pour symboliser principalement des *déplacements immatériels*, c'est-à-dire origine-destination selon une logique de raisonnement cartographique classique, mobilisant essentiellement la variable rétinienne *Taille*.

Cette famille d'applications se caractérise également par la présence au sein du même environnement, de *plusieurs représentations secondaires* présentées selon *plusieurs modalités graphiques*. La réalisation d'une enquête en ligne a ensuite permis d'identifier les besoins spécifiques d'une communauté pour ce type d'application, ouverte dans son code source et libre d'utilisation. Elle a enfin abouti à la définition du problème à résoudre : proposer un ensemble d'outils d'analyse et de représentation de matrices d'interactions géographiques ou territoriales complexes, c'est-à-dire denses (concernant de nombreux acteurs) et riches, car catégorielles et temporelles.

Cette démonstration va focaliser l'attention sur les principales possibilités d'actions sur les composantes géographiques, statistiques et cartographiques de cette toute nouvelle application appelée *Arabesques*.

2. L'information géographique

La première composante relève de l'information géographique qui sert de support (de fond de carte) ou d'habillage de la cartographie.

2.1. Jeux d'échelles et résolution

L'éventail des possibilités de représentations graphiques offertes aux usagers de cette nouvelle application est indépendant de l'échelle géographique des données à analyser. En effet, les possibilités de dessin qui sont proposées sont les mêmes quel que soit le niveau géographique.

Afin d'éviter d'éventuelles divergences entre les tailles des plus petits et des plus gros figurés, nous avons partiellement suivi une suggestion générale de Waldo Tobler d'ajuster le dessin des figurés linéaires à la résolution spatiale. Notre solution a pour objectif de garantir une perception locale du plus petit élément et globale de l'ensemble des motifs à peu près équilibrée, y compris en cas de zoom / dé zoom.

Cet ajustement des tracés est rendu possible grâce à un ratio défini empiriquement, fixé par défaut à 0,02% de la plus grande dimension du rectangle d'encombrement maximal du jeu de données de flux - de l'information statistique et non géographique (voir *infra*).

2.2. Jeux de données géographiques

L'application autorise l'import de ses propres jeux de données, aux formats .JSON et geoJSON, leur projection dans différents systèmes et leur personnalisation pour la représentation (voir §4.).

Des fonds géonumériques sont proposés à titre complémentaire, à des fins d'habillage ; il s'agit à ce stade de couches issues de *OpenStreetMap* et de *NaturalEarthData*.

En cas d'absence d'informations géographiques associées à un jeu de données statistique, il est possible de géocoder un fichier .CSV sous réserve que ce dernier comporte *a minima* un identifiant d'unités spatiales valide. À défaut de cela, les nœuds correspondants seront supprimés lors d'une vérification de l'intégrité du fichier.

Plusieurs possibilités de géocodage sont proposées pour différents niveaux territoriaux généraux, tels que les codes ISO3 et UNEP pour les pays, régions ou sous-régions du monde ; les identifiants du découpage européen NUTS 1,2,3 ; les codes postaux ou INSEE pour les différents découpages français ; les identifiants de territoires spéciaux, comme le Grand Paris, le *greater London* notamment.

3. L'information statistique

La seconde composante consiste à interagir avec une information statistique selon différentes modalités : exploration, calcul d'indicateurs, filtres interactifs numériques et/ou graphiques en soutien à l'exploration visuelle.

3.1. Jeux de données statistiques

Ces possibilités sont offertes pour toutes les dimensions d'une ou de plusieurs matrices empilées formant un même jeu de données, que l'information soit évaluée ou non, disponible pour son cœur (des liens) et/ou en ses marges (des nœuds).

Les matrices peuvent en effet être simples ou complexes :

– Matrice simple (ordinaire) : il n'existe qu'une seule relation, éventuellement bilatérale ou orientée, pour un même couple de lieux ;

– Matrice complexe, est une hyper matrice (ou matrice empilée) équivalente à un graphe multiple qui comporte plusieurs liens bilatéraux ou non, pour un même couple de nœuds.

Quel que soit leur type, les matrices doivent être fournies dans un format long présentant *a minima* trois colonnes (identifiants des lieux d'origine, de destination et valeur (mesure ou modalité) du lien).

3.2. Calcul d'indicateurs basiques

Outre les paramètres classiques résumant les distributions statistiques (min, max, ...), le calcul de quelques indicateurs de base est réalisé par défaut lors du chargement des données ; il s'applique aux liens et aux nœuds. Fondés sur des approches mixtes, de réseaux et gravitaires, ces calculs proposent par exemple au nombre de degrés (valué ou non), à des mesures de centralité, au calcul d'un volume ou d'une balance (soldes) des échanges, à la mesure de leurs espacements. En

présence d'hyper matrices ou de graphe multiplexe, des opérations supplémentaires sont réalisées (voir § 3.3).

3.3. Possibilités de filtrage

3.3.1. Filtrage numérique

Le filtrage numérique s'applique aux données quantitatives de stock (absolues) ou de taux (continues ou pseudo-continue). Il est réalisé soit visuellement à l'aide d'une fenêtre de sélection sur un histogramme interactif à fenêtre coulissante (*slider*), soit de manière numérique (en indiquant un seuil). La part d'information sélectionnée, les valeurs minimales et maximales sont indiquées automatiquement sur la figure. Pour les liens, la possibilité d'un filtrage selon la distance géodésique parcourue est en outre proposée.

3.3.2. Filtrage temporel

Les possibilités de filtrage des matrices temporelles ne sont actuellement disponibles que pour les liens. Elles portent sur le choix d'une date ou d'une période, sous une forme numérique ou graphique (*slider*). Trois formats de date sont disponibles (chaîne, horaire, horodatage).

3.3.3. Filtrage catégoriel

Le filtrage catégoriel s'applique aux données qualitatives nominales disponibles dans les jeux de données de nœuds et/ou de liens. Trois possibilités sont proposées graphiquement et numériquement :

- Sélection de variables multiples : l'utilisateur choisit une ou plusieurs des variables qualitatives disponibles (par exemple, les préfectures et sous-préfectures s'il dispose d'une variable spécifiant le profil administratif des villes françaises) ou le type de liens (par exemple, import ou export pour des données d'échanges commerciaux) ;
- Sélection d'une variable unique : l'utilisateur choisit une seule variable (pour les nœuds : par exemple, les flux au départ et à destination de Clermont-Ferrand pour des données de mobilité ; les flux de type import pour des données commerciales) ;
- Sélection par suppression d'une variable unique : l'utilisateur choisit une ou plusieurs variables portant sur les nœuds et les liens à supprimer. Si la suppression sélection porte sur des nœuds, tous les liens correspondants (par exemple, la région Ile-de-France pour une analyse à l'échelle nationale) sont supprimés.

3.4. Possibilités d'agrégation

Des fonctions d'agrégation attributaires, relatives à la gestion des hyper matrices ou graphes multiplexes, sont proposées pour réduire le nombre de figurés à représenter - ne pas représenter, par exemple, les n liens qui correspondent aux n dates disponibles.

Par défaut, c'est la fonction somme qui est appliquée au chargement des données, en l'absence de précisions. L'utilisateur peut toutefois choisir d'appliquer une fonction moyenne, minimum, maximum ou médiane calculée sur l'ensemble des matrices ou des graphes fournis. Il est également possible de choisir une seule date ou bien d'agréger les données, selon une fonction donnée, sur une période ou pour des catégories.

Cette fonction est importante, car elle définit par défaut l'intensité des teintes et l'opacité des figurés correspondants (voir *infra*).

4. L'information cartographique

La troisième composante de *Arabesques* est dédiée au rendu cartographique, par une action sur la symbolisation (ponctuelle ou linéaire) des données ou sur les fonds servant de support ou d'habillage.

Le dessin des différentes couches et de leurs objets est finement paramétrable, afin de tenir compte de l'éventuelle complexité de l'information, en termes de densité, nécessitant une gestion particulière des superpositions et des dispositions des liens et nœuds (voir Fig. 1). Les liens les plus gros sont placés au premier plan par défaut à l'inverse des cercles les plus larges



FIGURE 1. *Dispositions et gestion des superpositions*

La représentation cartographique relève notamment de la richesse (type de caractères, catégories, temporalités) des informations symbolisées par les différents objets géographiques représentés : des liens, des nœuds. Il est possible de contextualiser leur représentation par des informations relatives à leur structure *via* la réalisation de cartes choroplèthes (par exemple de statistiques socio-économiques, environnementales ou autres, si les données correspondantes sont fournies), ou bien l'affichage d'informations complémentaires d'habillage disponibles par défaut. Les différents éléments de paramétrage sont ensuite repris dans une légende placée sur la vue principale.

Les éléments classiques de la conception cartographique thématique (échelle graphique, légende), de même que la possibilité d'ajouts graphiques complémentaires (titre, etc.) sont disponibles par défaut.

4.1. Symbologie des nœuds

Le dessin des nœuds est enrichi grâce à une fenêtre interactive permettant de paramétrer leur apparence : leur *Taille*, leur *Forme* (cercle uniquement), leur *Teinte* et leur *opacité* ; ainsi que les labels correspondants (voir Fig. 2).

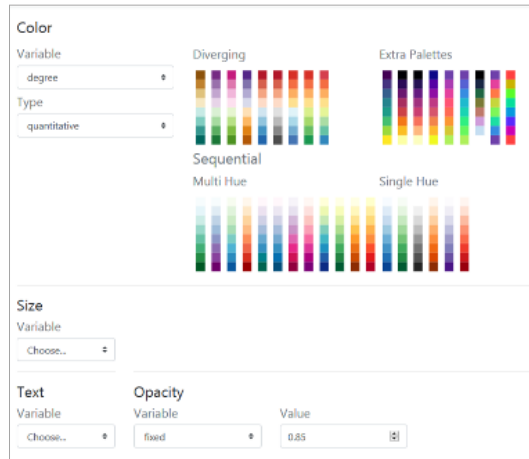


FIGURE 2. Fenêtre de paramétrage des nœuds

4.1. Symbologie des liens

Le dessin des liens est plus riche que celui des sommets. Notre application propose donc une possibilité d'action sur leur géométrie qui enrichit leur sémiologie, par l'ajout de la *Forme* et l'*Orientation*, qui définit leur style (voir Fig. 3).

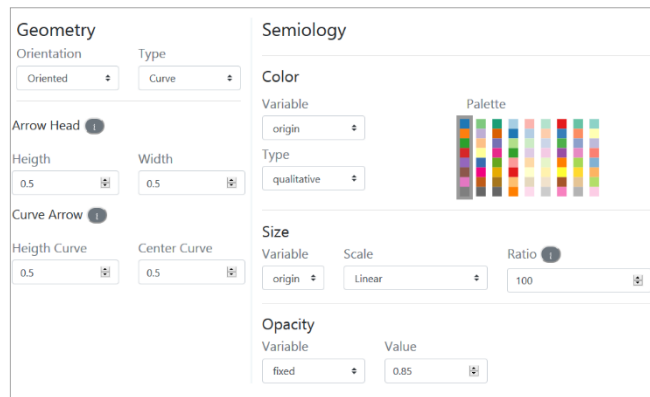


FIGURE 3. Fenêtre de paramétrage des liens

Il est en effet possible d'agir sur toutes les dimensions géométriques du lien :

- la *Taille* permet le paramétrage de l'épaisseur du lien en fonction d'une valeur ;
- la *Valeur*, associée à la *Couleur*, permet le paramétrage de la teinte du lien, éventuellement de son intensité en fonction de variables respectivement catégorielles ou numériques (de taux) ;
- l'*Opacité* des liens, qui est fondée sur l'intensité de la teinte calibrée pour les matrices complexes est fixée par défaut à 0,85, mais peut être modifiée ;
- l'*Orientation*, lorsqu'elle est présente, permet le paramétrage des dimensions de la tête de la demi-flèche (voir Fig. 1) ;
- la *Forme* qui autorise le dessin
 - d'une droite ;
 - d'une courbe, avec paramétrage de sa hauteur (point maximal atteint) et de sa pente à l'aide d'un point d'inflexion ;
 - d'un triangle droit ;
 - d'un triangle courbe, sous la forme de goutte d'eau.
- le ratio, déjà évoqué, dont la valeur par défaut peut être modifiée afin d'optimiser le dessin en fonction de la résolution spatiale.

Conclusion

Le projet gflowiz (*geographic flow visualisation*) est un projet innovant visant à créer *Arabesques*, une application web dédiée à la géo visualisation de données de flux et de réseaux spatiaux. Il tient compte des informations collectées grâce à un inventaire des applications existantes dans le domaine et d'une enquête réalisée auprès de professionnels pour collecter leurs besoins en termes d'application web.

Une attention particulière a été apportée aux composantes géographiques, statistiques et cartographiques de l'application, afin d'obtenir un ensemble homogène et cohérent avec la littérature actuelle.

Arabesques offre de nombreuses possibilités tant dans les données en entrées qu'elle permet de charger et manipuler (matrices simples ou complexes), que dans les possibilités de filtrage, d'agrégation, de représentation et de symbologie.

Remerciements :

L'équipe de gflowiz souhaite remercier les porteur.e.s de l'Action Prospective Théoriser, Observer, Analyser, Modéliser le Geoweb du Groupement de recherche 2340 du CNRS : Méthodes et Applications pour la Géomatique et l'Information spatiale (GDR MAGIS), pour leur accueil chaleureux, le support et le soutien pour l'organisation de la session de la conférence SAGEO'2019, dédiée à la première présentation de cette application.