
La sédimentation visuelle de données

Outil et technique pour visualiser les flux de données à destination du grand public.

DOMAINE : Visualisation de données

TYPE : Logiciel, technique

MOTS-CLES : Sédimentation visuelle, Visualisation dynamique, Flux de données, Données dynamiques, Temps réel, Utilisateurs non experts, Conception, Métaphore.

KEYWORDS: Design, Information visualization, Dynamic visualization, Dynamic data, Data stream, Real time, Design Metaphor, JavaScript, Token.

DOI:10.3199/ISI.18.1.1-n © 2013 Lavoisier

1. Objectif du projet

En 2013¹, cinq cent millions sont envoyés par jour. Cette quantité ne cesse de croître, et explose pour les événements à forte popularité où 150 000 tweets peuvent être envoyés par secondes. Il n'est pas possible pour un humain de traiter toutes ces informations de part leur volume, vitesse, et hétérogénéité. Ces critères correspondent à la définition du 'Big data', et de nombreux modèles statistiques aident à structurer, analyser et prendre des décisions afin de répondre à une question précise sur ces données (comme les prédire). Cependant la visualisation de données reste nécessaire dans la compréhension de ces données brutes, la visualisation de la décision et enfin leur communication, en particulier vers un large public.

Visualiser des flux de données, même à un rythme de mise à jour de quelques éléments par secondes est déjà un challenge : l'humain a une capacité d'attention limitée et des études ont montré qu'il ne peut pas suivre plus de deux ou trois éléments à la fois. Dans un environnement de travail, l'humain peut être distrait et même interrompu pendant une longue période de temps, durant laquelle les données ont été mises à jour (comme de nouveaux Tweets). Comprendre la mise à jour, quand elle a eu lieu, et comment, doit être supporté par l'interface graphique. Ce besoin est prédominant dans des domaines où la réponse de l'utilisateur doit être très rapide juste après la réception de données, par exemple en gestion d'urgences, de presse, en bourse, ou dans le *crowdsourcing* en temps réel (Huron & Fekete, 2012),

¹ Estimation en moyenne, <https://blog.twitter.com/2013/new-tweets-per-second-record-and-how>

etc. Pour répondre à ces exigences, nous devons créer de nouvelles techniques qui permettent de visualiser les données en transits en temps réel, tout en gardant le bénéfice de visualiser la tendance générale. C'est pour répondre à ces besoins que nous avons créé « Visual Sedimentation » (Huron, Vuillemot, & Fekete, 2013, 2012).

2. Résultats

La sédimentation visuelle est une technique inspirée du processus physique de la sédimentation. Ce processus est le résultat de roches qui chutent, puis s'entassent en couches et se sédimentent au fur et à mesure du temps. Dans *Visual Sedimentation* nous appliquons cette analogie à des flux de données numériques. Les flux de données continus sont d'abord découpés en éléments discrets. Puis chaque élément est associé à un jeton (éléments graphiques). Chaque jeton est alors affiché à l'écran, il est soumis à une simulation des forces physiques comme celles de la gravité et du magnétisme. A la fin de leur chute, les jetons s'empilent en formant un tas, dans lequel ils s'agrègent en décroissant visuellement au fur et à mesure du temps. Au final, le tas de jeton se transforme en un diagramme classique comme un camembert ou un diagramme en bâton, dont la taille est mise à jour en fonction du nombre de jetons accumulés.

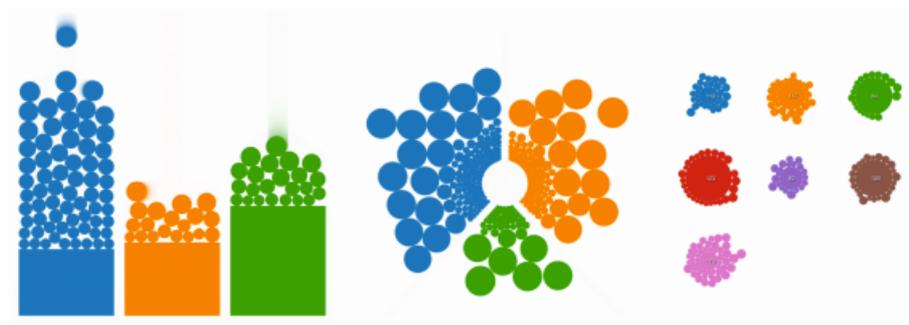


Figure 1. La sédimentation visuelle appliquée au diagramme en bâton à gauche, à un diagramme en camembert au centre et à un « bubble chart » à droite.

En utilisant cette analogie pour concevoir des visualisations nous permettons à un utilisateur non expert d'utiliser une connaissance préexistante à la visualisation pour mieux la comprendre. L'évolution des jetons dans l'espace est prédictible et la représentation des mises à jour est lisible. La visualisation peut être lue par un utilisateur sans nécessiter un grand apprentissage a priori. La sédimentation visuelle est donc une technique qui permet ainsi de visualiser la mise à jour sur des diagrammes classiques (Figure 1.) comme les graphiques en bâtons, les camemberts, les « bubbles charts », etc. Basée sur les contraintes définies plus haut, cette technique est aussi générative et permet de créer une multitude de nouvelles visualisations.

Notre méthode *Visual Sedimentation* a été appliquée dans différents contextes comme dans l'engagement du public d'une émission de télévision (Huron Samuel, Vuillemot Romain, & Fekete Jean Daniel, 2013), lors de tâches d'annotation collective en temps réel (Huron Samuel, Isenberg Petra, & Fekete Jean Daniel, 2013), pour visualiser les tweets des candidats présidentiels durant la campagne de 2012 (figure 2. A), ou encore pour voir l'évolution du travail dans un logiciel de gestion de versions (figure 2. G).

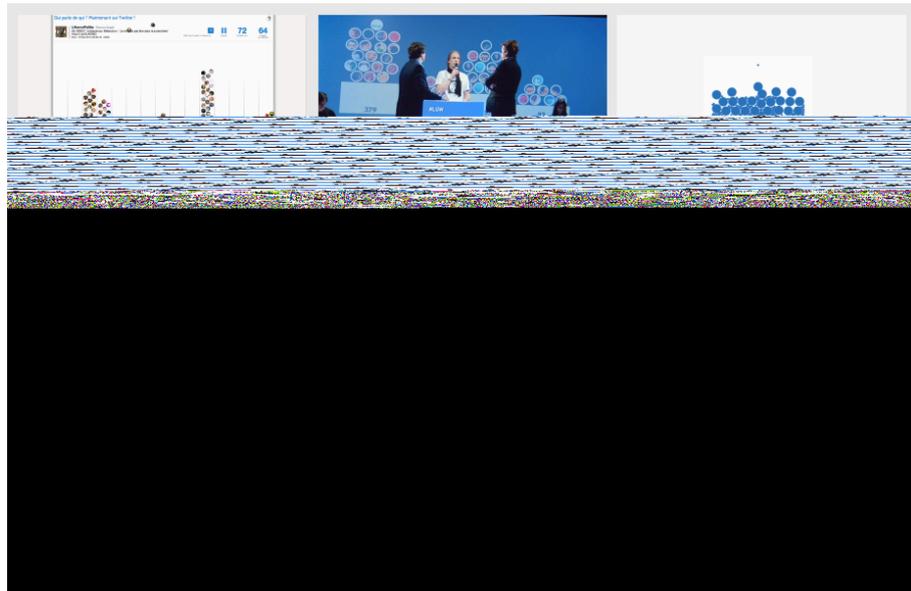


Figure 2. Différentes applications de la sédimentation visuelle sur différents flux de données, (A) *BubbleT* : tweets citant un candidat à l'élection présidentielle (B) *BubbleTV* : tweet utilisant différents hashtags dans une application de Social TV (C) *Sediclock* (D) flux RSS de Stackoverflow (E) Statistiques de l'évolution de la population (F) Mises à jour de wikipedia (G) Commit dans un logiciel de gestion de versions (H) Tweets citant la couleur d'un MMS (I) Détails.

3. Distribution et contact

Visual Sedimentation est une librairie JavaScript Open Source distribuée sur un site dédié² et sur GitHub³. *Visual Sedimentation* a été conçu et développé à l'Institut

² <http://www.visualseedimentation.org/>

³ <https://github.com/INRIA/VisualSedimentation>

de recherche et d'innovation du Centre Pompidou⁴ et à l'INRIA, par Samuel Huron et Romain Vuillemot, sous la direction de Jean Daniel Fekete.

Bibliographie

- Huron, S., & Fekete, J.-D. (2012). User-Supplied Sentiments in Tweets. *IEEE Visweek : 2nd Workshop on Interactive Visual Text Analysis*. Oct 2012, Seattle. Retrieved from <http://hal.inria.fr/hal-00734407>
- Huron, S., Vuillemot, R., & Fekete, J.-D. (2012). Towards Visual Sedimentation. *IEEE VisWeek 2012 Electronic Conference Proceedings*. Seattle, Washington, USA, États-Unis. Retrieved from <http://hal.inria.fr/hal-00734084>
- Huron, S., Vuillemot, R., & Fekete, J.-D. (2013). Visual Sedimentation. *IEEE, 2013, IEEE Transactions on Visualization and Computer Graphics, 19 (12), pp. 2446-2455*. Retrieved from <http://hal.inria.fr/hal-00846260>
- Huron S., Isenberg P., & Fekete J.-D. (2013). PolemicTweet: Video Annotation and Analysis through Tagged Tweets. *Proceedings of the IFIP TC13 Conference on Human-Computer Interaction (INTERACT), Sep 2013, Heidelberg, Germany. Springer, 14th IFIP TC 13 International Conference, Cape Town, South Africa, September 2-6, 2013, Proceedings, Part II, 8118, pp. 135-152*. Retrieved from <http://hal.inria.fr/hal-00817591>
- Huron S., Vuillemot R., & Fekete J.-D (2013). Bubble-TV: Live Visual Feedback for Social TV Broadcast. *ACM CHI 2013 Workshop : Exploring and enhancing the user experience for television (2013)*. Retrieved from <http://hal.inria.fr/hal-00796354>

⁴ <http://www.iri.centrepompidou.fr/>