

## **XEWGraph : Outil de Visualisation et Analyse des Hypergraphes pour un Système d'Intelligence Economique**

Zakaria Boulouard\*, Amine El Haddadi\*\*, Anass El Haddadi\*\*\*,  
Lahcen Koutti\*, Abdelhadi Fennan\*\*

\*Faculté des Sciences Agadir

zboulouard@gmail.com

lkoutti@yahoo.fr

\*\*FST Tanger

am.elhaddadi, afennan@gmail.com

\*\*\*ENSA Al Hoceima

anass.elhaddadi@gmail.com

L'Intelligence Economique (ou Competitive Intelligence) est une démarche et un processus organisationnel qui permet à l'entreprise d'être plus compétitive, d'une part en surveillant l'environnement et ses changements, et d'autre part, en surveillant les changements internes.

Les graphes sont parmi les outils les plus puissants de visualisation. Ils permettent d'afficher, tout type d'information utile et peuvent, donc, bien répondre aux besoins d'Intelligence Economique. Contrairement aux graphes de la Business Intelligence basés sur des données structurées, les graphes de l'IE sont basés sur des données non-structurées (voire massives), donc à la fois difficiles à analyser et à visualiser. Dans ce cas, la clé qui va nous permettre de tirer des informations utiles et de les afficher sous forme d'un graphe est "la catégorisation". Par exemple, en veille scientifique, afin d'afficher les réseaux sociaux liant les auteurs, cela serait beaucoup plus simple de les catégoriser selon les articles qu'ils ont publiés en commun, ou encore, selon leurs pays.

L'idée est de proposer un outil qui va gérer et organiser les données sous forme relationnelle appelée « Préconnaissance » et de là, extraire les connaissances implicites qui seront mises en forme et présentées aux décideurs sous forme de graphes décisionnels. Contrairement aux outils de visualisation de graphes déjà présents (Gephi, VisuGraph, etc...), notre outil Xplor EveryWhere (XEW) présente un nouveau module qui donne un nouveau cadre pour l'analyse des données massives et l'affichage des résultats sous forme de graphes. En effet, XEWGraph permet d'afficher des larges graphes sur des plateformes web et mobiles en tirant profit, d'un côté de la puissance de l'architecture client-serveur de XEW qui offre un partage intuitif de l'information décisionnelle, et d'un autre côté de l'approche que nous avons adoptée pour la visualisation des graphes, et qui nous a permis d'obtenir à la fois une plus grande vitesse dans la génération de la préconnaissance et une représentation beaucoup plus significative des graphes visualisés.

L'approche adoptée est de dessiner les larges graphes à visualiser en se basant sur les spécificités des hypergraphes.

**Définition.** Un Hypergraphe  $H = (X, E)$  est un couple dans lequel  $X$  est un ensemble de noeuds et  $E$  est un ensemble d'hyperedges  $e_k$  contenant un nombre arbitraire de noeuds de  $X$ .

Dans notre approche, nous avons modifié cette définition en proposant un lemme stipulant que l'hyperedge  $e_k$  est un ensemble contenant un nombre arbitraire de noeuds de  $X$  ayant une caractéristique en commun.

**Lemme.** Soit  $X = \{x_i / i \in 1, \dots, n\}$  un ensemble de noeuds  $x_i$ .

Soit  $y_j$  avec  $j \in 1, \dots, m$  une caractéristique du noeud  $x_i$ , on pourrait écrire  $x_i$  tel que

$$x_i = \{y_j / j \in 1, \dots, m\}$$

L'hyperedge  $e_k \in E$  avec  $k \in 1, \dots, p$  est redéfini donc comme étant :

$$e_k = \{x_i \in X / \exists j \in 1, \dots, m ; x_i(y_j) = cte\}$$

**Application.** Cette écriture de l'hyperedge nous a permis d'isoler une caractéristique commune à l'ensemble du corpus de données et de l'utiliser pour représenter les hyperedges contenant l'ensemble des noeuds ayant cette même caractéristique.

Comme précisé dans la description de l'approche adoptée, la définition d'une caractéristique et son utilisation pour la visualisation d'hyperedges représentant des points ayant la même valeur de cette caractéristique nous a permis de gagner des avantages considérables en le comparant au prototype web de VisuGraph que nous avons précédemment essayé d'intégrer à Xplor Everywhere :

- Beaucoup moins de temps lors de la phase de génération de la préconnaissance. En effet, nous avons pu exécuter cette phase en 1h contre 2h30 pour un corpus de 65000 sources de données traitées par un PC standard.
- Côté visualisation des données. En effet, nous avons pu visualiser trois fois plus de données significatives avec le même moteur Javascript.

Nous avons présenté sur ce papier un nouveau module de visualisation de données massives faisant partie du Système d'Intelligence Economique Xplor Everywhere qui donne la possibilité d'avoir plus de flexibilité et de partager l'information décisionnelle à grande échelle grâce à son architecture web (client-serveur). Ceci a été possible grâce à l'adoption d'une approche basée sur les hypergraphes.

Nous comptons améliorer encore plus les performances de XEWGraph en adoptant, d'un côté, une architecture clustérisée basée sur Hadoop, et d'un autre côté en mettant en place un algorithme d'optimisation des graphes pour faciliter la navigation dans les différentes typologies des graphes (réseaux sociaux, sémantiques, de collaborations, d'alliances, etc...), en se basant sur les principes de permutation et d'hybridation des noeuds.

## Summary

The Competitive Intelligence System Xplor Everywhere helps searching, visualizing, and sharing useful data. In this paper, we will introduce Xplor Everywhere and its newest feature called XEWGraph, which is dedicated to the analysis of massive data and visualization of hypergraphs.