

Schéma multidimensionnel dédié pour l'OLAP des Tweets

Maha Ben Kraiem^{*,**}, Kais Khrouf^{*}, Jamel Feki^{*}, Franck Ravat^{**}, Olivier Teste^{**}

^{*}Laboratoire Mir@cl, Université de Sfax
Route de l'Aérodrome km 4, B.P. 1088, 3018 Sfax, Tunisie
Maha.BenKraiem@yahoo.com, Khrouf.Kais@isecs.rnu.tn, Jamel.Feki@fsegs.rnu.tn
^{**} IRIT, Université Paul Sabatier Toulouse III
118, route de narbonne, F-31062 Toulouse Cedex 9, France
{Franck.Ravat, Olivier.Teste}@irit.fr

Résumé. Les tweets permettent l'échange de faits et d'opinions entre les utilisateurs du réseau social "Twitter". Le nombre de tweets échangés ne cesse d'augmenter et constitue ainsi une nouvelle source importante d'informations. L'application des techniques OLAP "On-Line analytical Processing" sur ces gros volumes de tweets permet d'extraire de nouvelles informations et/ou connaissances concernant par exemple le comportement des usagers ou les sujets émergents. Cet article propose un schéma multidimensionnel générique dédié à l'OLAP des données dynamiques (tweets).

1 Introduction

Dans le monde interconnecté actuel, la vitesse de diffusion des informations amène à une formation des données tendant vers de plus en plus d'immédiateté. En effet, les grands réseaux sociaux permettent le partage et la diffusion de l'information de manière quasi-instantanée. Depuis leur apparition, les blogs et les réseaux sociaux ont rendu les internautes plus actifs au sein des réseaux participatifs. Ces derniers s'inscrivent dans une tendance technologique plus large, appelée Web 2.0, qui met le Web en valeur avec davantage d'interaction et de collaboration.

L'exploitation de ces nouvelles formes de publications suscite des applications décisionnelles exceptionnelles pour des décideurs non conventionnels. En effet, ces décideurs utilisent ces gros volumes de données comme nouvelles ressources documentaires pour y puiser de l'information qui n'était pas accessible par le biais des moyens classiques. Depuis son apparition en 2006, le site Web Twitter¹ s'est développé de telle manière qu'il est actuellement le dixième site le plus visité au monde². Twitter est une plate-forme de *microblogging*. Il s'agit d'un système de partage d'informations via lequel un utilisateur a la possibilité, soit de suivre d'autres utilisateurs qui postent des messages courts, soit de diffuser ses propres messages à ses propres suiveurs ("followers").

1. <http://twitter.com>.

2. <http://www.alexa.com/siteinfo/twitter.com>.