

Motifs Séquentiels pour la Sélection des Webviews à Matérialiser

Ali Ben Ammar*
Mouna Badis**, Abdelaziz Abdellatif***

*Institut Supérieur d'Informatique et de Gestion de Kairouan
3100, Kairouan. Tunisie
ali.benammar @ isd.rnu.tn
**Faculté des Sciences de Tunis
2092, Tunis. Tunisie
manoui4@yahoo.fr
***Faculté des sciences de Tunis
2092, Tunis. Tunisie
abdelaiziz.abdellatif@fst.rnu.tn

Résumé. Dans cet article nous proposons une approche pour la recherche d'indicateurs renforçant le choix des webviews à matérialiser. Un webview matérialisé est une instance statique, d'une page web dynamique, stockée au niveau du serveur web à un point de temps. Il sert à réduire le coût de génération répétitive des données sources des requêtes. Notre contribution consiste à analyser l'historique du site web, moyennant la technique de web usage mining, pour calculer un poids de matérialisation pour chaque webview. Ce poids servira par la suite pour l'algorithme de sélection des webviews à matérialiser afin de décider si un webview est recommandé pour la matérialisation ou non. Plus précisément, il sert pour estimer le profit (en termes de temps de réponse) pouvant être produit par la matérialisation d'un webview. Nos résultats d'expérimentation ont montré que notre approche permet de réduire le risque de matérialisation des webviews, dont le profit n'est certain au moment de leur sélection, de plus de 20%.

1 Introduction

La matérialisation des données est une technique permettant d'optimiser le temps de réponse des requêtes des usagers. Elle consiste à stocker les résultats de quelques requêtes pour réduire le coût de génération répétitive des données à partir des sources. Récemment, elle a été appliquée au niveau des Sites Web à Utilisation Intensive de Données (DIWS, Data Intensive Web Sites) où on parle de matérialisation de webviews. Un webview est une page web dynamique dont les données sont extraites à partir de bases de données structurées. Ainsi, la matérialisation des webviews consiste à transformer les données dynamiques en données statiques, c'est-à-dire la création d'une instance statique d'une page web dynamique à un instant donné.

Dans un environnement de DIWS, un webview peut être soit virtuel, c'est-à-dire toujours généré en ligne à partir des sources, ou matérialisé, c'est-à-dire stocké au niveau du serveur. La matérialisation des données au niveau des DIWS peut prendre deux formes : matérialisation des vues (les données structurées représentant les sources des webviews) au