



EXTENSION DU MODÈLE MULTIDIMENSIONNEL AUX FAITS QUALITATIFS. APPLICATION À L'ANALYSE EN LIGNE DES GÊNES DES CHANTIERS URBAINS PAR MME FATIHA AMANZOUGARENE

Présentée par : Fatiha Amanzougarene Discipline : informatique Laboratoire : PRISM

Résumé :

Les entrepôts de données et les systèmes OLAP constituent les principaux éléments d'un système d'information décisionnel. Ces dernières années, plusieurs travaux ont été menés, afin d'étendre les concepts des entrepôts de données classiques pour traiter de nouveaux types de données, dites complexes (texte, multimédia, géographique, etc...). Dans cette thèse, nous nous sommes intéressés à la problématique de l'intégration de l'information qualitative dans l'analyse multidimensionnelle. Nos travaux sont guidés par une étude de cas sur les gênes des chantiers urbains. Après avoir défini la notion de gêne et déterminer les différents facteurs qui interviennent dans son évaluation, nous avons mis en évidence le besoin d'un modèle de représentation qualitative s'appuyant sur des règles et des connaissances expertes. Or, l'analyse multidimensionnelle dans les entrepôts de données traditionnelles ne considère que des mesures quantitatives. Notre principale contribution est donc d'étendre le modèle multidimensionnel aux mesures

qualitatives exprimées sous forme de termes linguistiques. Partant du constat que les connaissances expertes sont parfois incomplètes, notre deuxième contribution est de proposer une méthode originale de reconstruction de données manquantes dans le contexte des entrepôts de données. Ce modèle consiste à combiner la technique des k-plus proches voisins (KPPV) avec la programmation par contraintes pour assurer une meilleure prédiction des valeurs manquantes dans une table de faits d'un entrepôt de données classique ou qualitatif.

Abstract :

Data warehouses and OLAP systems constitute the main elements of decision support systems. In recent years, several studies have been conducted in order to extend the capabilities of conventional data warehouse to handle complex data types (e.g., text, multimedia, geographic, etc.)

In this thesis, we focused on the integration problem of qualitative information in multidimensional analysis. Our work are guided by a case study on urban building sites annoyances. After defining the notion of annoyance and determined the factors involved in its evaluation, we highlighted the need for a qualitative representation model based on rules and expert knowledge. However, conventional multidimensional data models only consider quantitative measures. Therefore, our main contribution is to extend the multidimensional model to treat qualitative measures expressed as linguistic terms. Considering that expert knowledge are sometimes incomplete, our second contribution is to propose an original model for missing data reconstruction in the context of data warehouses. This model consists in combining the constraint programming and a technique of machine learning, namely the k-nearest neighbor algorithm. In addition to its application in classical data warehouses, our model adapts to qualitative data warehouses, as in the annoyances analysis of urban building sites.

INFORMATIONS COMPLÉMENTAIRES

Mohand-Saïd HACID, Professeur des Universités, à l'Université Claude Bernard Lyon 1/Laboratoire d'InfoRmatique, Image et Systèmes d'Information (LIRIS) - CNRS UMR 5205 - Villeurbanne - Rapporteur

Mohamed QUAFAROU, Professeur des Universités, à l'Université d'Aix-Marseille/Laboratoire des Sciences de l'Information et des Sciences (LSIS) - UMR CNRS 7296 - Marseille - Rapporteur

Karine ZEITOUNI, Professeur des Universités, à l'Université de Versailles Saint-Quentin-en-Yvelines/Laboratoire Parallélisme, Réseaux, Système, Modélisation (PRISM) - Versailles - Directeur de thèse

Mohamed CHACHOUA, Enseignant-Chercheur, à l'Ecole des Ingénieurs de la Ville de Paris - Paris - Co-encadrant de thèse

Anne LAURENT, Professeur des Universités, à l'Université Montpellier 2/Laboratoire d'Informatique, de Robotique et de Microélectronique de Montpellier (LIRMM) - CNRS UMR 5506 - Montpellier - Examineur

Patrick MARCEL, Maître de Conférences, à l'Université François Rabelais /Laboratoire d'Informatique (LI) - Blois - Examineur

Contact : dredval service FED : theses@uvsq.fr