

Vers une décision Skyline intelligente

Sana Nadouri^{*,**}, Allel Hadjali^{*}, Zaidi Sahnoun^{**}

^{*}ENSMA - Laboratoire LIAS (Poitiers, 86360) sana.nadouri@ensma.fr,
allel.hadjali@ensma.fr

^{**}UC2 - Laboratoire LIRE (Constantine, 25000)
sana.nadouri@univ-constantine2.dz, zaidi.sahnoun@univ-constantine2.dz

Dans cet article, nous proposons une approche B2D (DataBase To Decision tables) une adaptation du Skyline pour optimiser les Systèmes Interactifs d'Aide à la Décision (SIAD). Actuellement, il n'y a pas d'application du modèle Skyline dans les SIAD. L'opérateur Skyline, qui a été introduit pour la première fois par Borzsony et al. (2001), renvoie les meilleurs tuples en faisant un compromis optimal de plusieurs dimensions définies par le décideur Hose (2016), il n'est pas nécessaire de définir une fonction de score pour classer les éléments car il se base sur la relation de dominance de Pareto : soit p et q deux tuples avec le même nombre de dimensions d , p domine q , si p est au moins aussi bon que q sur toutes les dimensions et meilleur que q sur au moins une dimension. **Formellement** : (en supposant que plus la valeur est petite, meilleure elle est) $p \prec q \implies \forall i \in [1,d] : p_i \leq_i q_i$ and $\exists j \in [1,d] : p_j <_j q_j$, si ni $p \prec q$ ni $q \prec p$, ils sont incomparables.

L'analyse Skyline dans les Bases de Données (BD) a été largement étudiée, ce type d'analyse ne s'applique pas seulement aux BD relationnelles, mais elle peut aider dans tous les domaines dans lesquels une décision est requise Borzsony et al. (2001). Le paradigme du Skyline n'a jamais été utilisé auparavant sur les tables de décision. L'utilisation de cet opérateur a plusieurs avantages : l'amélioration des performances (en réduisant la taille de la table), la qualité de la décision et la persistance de la réponse. Nous avons d'abord procédé à une analyse critique des algorithmes du Skyline existants et choisi l'algorithme Salsa de Bartolini et al. (2006, 2008).

L'approche dépend de l'extraction de critères pour appliquer le processus Skyline. On utilise le modèle HDMSM (Hybrid Decision-Making Support Model) Wang et al. (2015) pour définir les critères et Salsa pour l'intégration du Skyline en modifiant la dernière phase de HDMSM pour obtenir rapidement des décisions fiables. Salsa s'efforce d'éviter l'analyse de l'intégralité des données triées, contrairement aux propositions précédentes. C'est le premier algorithme qui exploite les valeurs d'une fonction de notation monotone pour trier les données à comparer, nous avons appliqué ce mécanisme aux tables de décision pour obtenir les meilleures décisions à l'aide des critères extraites après utilisation de HDMSM. Les tables de décision sont similaires aux tables des BD relationnelles. Au lieu d'avoir des valeurs, nous avons des conditions et chaque tuple représente une décision qui peut être sélectionnée si toutes les conditions sont satisfaites, la différence reste dans les types de valeur comme les conditions composées. Dans les SIAD, la décision dépend de plusieurs critères souvent contradictoires. C'est la raison pour laquelle nous avons choisi le Skyline pour sélectionner les décisions en tenant compte de manière équitable de tous les critères choisis. B2D a deux parties : (i) HDMSM

pour formuler le problème avec un ensemble de critères ; et (ii) la génération des décisions à partir des tables de décision.

Notre étude de cas porte sur la proposition de traitements adéquats à un patient, nous avons les analyses du patient et ses antécédents. Le médecin propose des traitements au SIAD. La tâche consiste à trouver les traitements les plus appropriés en fonction du patient. Le SIAD contient des tables de décision avec symptômes et un ensemble de traitements associés à ces symptômes. Le processus consiste à appliquer Salsa sur ces tables après avoir déterminé les critères de décision, le système propose ou corrige les décisions du médecin. Nous avons effectué deux tests. **Test 1** : nous avons testé l'évolution du temps d'exécution qui dépend de la taille de la table de décision. **Test 2** : les mêmes variations avec une différence dans les choix de décision, les patients et la taille des analyses. Les premiers résultats sont intéressants et encourageants (Fig. 1).

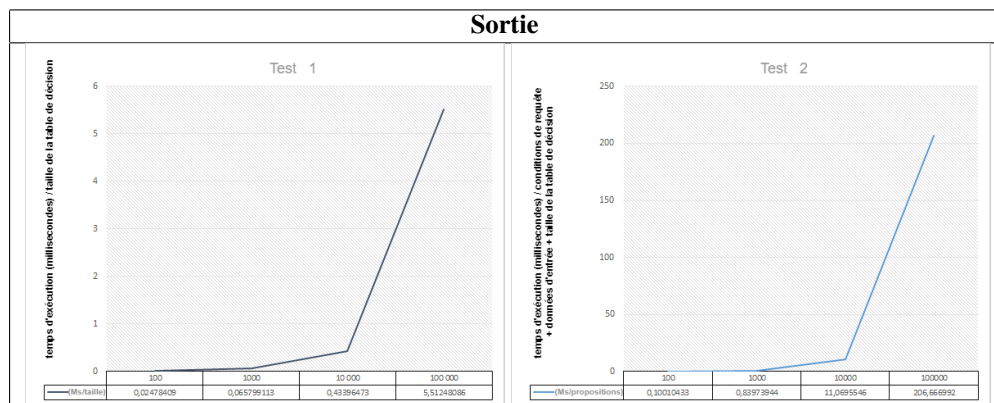


FIG. 1 – Evaluation de l'approche B2D

Pour répondre à l'une des faiblesses des SIAD, nous avons proposé l'approche B2D qui exploite le paradigme du Skyline. Ensuite, nous avons mis en œuvre et vérifié notre proposition dans différentes situations (taille/diversité/cas), les premiers résultats sont très prometteurs. Pour les travaux futurs, nous comptons (i) mettre en œuvre l'ensemble du SIAD et le comparer avec le processus normal ; (ii) étendre cette étude aux décisions réparties en exploitant l'idée du Skyline de groupe.

Références

- Bartolini, I., P. Ciaccia, et M. Patella (2006). Salsa : Computing the skyline without scanning the whole sky. In *15th ACM Int. Conf. on Inf. and Knowledge Management, CIKM '06*, New York, NY, USA, pp. 405–414. ACM.
- Bartolini, I., P. Ciaccia, et M. Patella (2008). Efficient sort-based skyline evaluation. *ACM Trans. Database Syst.* 33, 31 :1–31 :49.
- Borzsony, S., D. Kossmann, et K. Stocker (2001). The skyline operator. In *17th Int. Conf. on Data Eng.*, pp. 421–430.
- Hose, K. (2016). Skyline queries. *Datenbank-Spektrum* 16(3), 247–251.
- Wang, C.-S., H.-L. Yang, et S.-L. Lin (2015). To make good decision : A group dss for multiple criteria alternative rank and selection. *2015*, 1–15.