الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne Démocratique et Populaire
وزارة التعليم العالي و البحث العلمي
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
المدرسة العليا للإعلام الآلي 08 ماي 1945 سيدي بلعباس
École Supérieure en Informatique
08 Mai 1945 Sidi Bel Abbès

THESIS

To obtain the diploma of **Ingénieur d'État**
Field: **Computer Science**
Specialty: **Ingénierie des Systèmes Informatiques (ISI)**

Theme

Anomaly Detection in Streaming Data through Closed Itemset and Rule-Based Approaches

---

Presented by:
SAKOUTI Yasmine

Submission Date: **Sept, 2024**
In front of the jury composed of:

Dr. BENSENANE Hamdane	ESI Sba	President
Pr RAHMOUN Abdelatif	ESI Sba	Co-Supervisor
Dr BENABDERRAHMANE Sid Ahmed	NYU Abu Dhabi	Supervisor
Dr. BENDAOUD Fayssal	ESI Sba	Examiner

Academic Year : 2023/2024

Abstract

Anomaly detection is crucial for maintaining data integrity across various domains, particularly in environments characterized by dynamic data flows.

A novel approach to anomaly detection utilizes the extraction of closed itemsets and association rules.

By leveraging a streaming-based algorithm, the proposed method efficiently identifies anomalies in real time, enhancing the detection of irregular itemsets in transactional datasets.

The effectiveness of this approach is evaluated through benchmark datasets, demonstrating its capability to achieve high accuracy.

Additionally, the integration of closed itemsets significantly reduces computational overhead while improving detection accuracy, making this approach suitable for applications with resource constraints.

Keywords : Anomaly Detection, Closed Itemset Mining, Streaming Data, Association Rules, Real-Time Data Processing.

Résumé

La détection des anomalies est cruciale pour maintenir l'intégrité des données dans divers domaines, en particulier dans les environnements caractérisés par des flux de données dynamiques.

Une nouvelle approche de la détection des anomalies utilise l'extraction d'ensembles d'items fermés et de règles d'association.

En s'appuyant sur un algorithme basé sur le streaming, la méthode proposée identifie efficacement les anomalies en temps réel, améliorant ainsi la détection des ensembles d'items irréguliers dans les ensembles de données transactionnelles.

L'efficacité de cette approche est évaluée à travers des ensembles de données de référence, démontrant sa capacité à atteindre une haute précision.

De plus, l'intégration des ensembles d'items fermés réduit considérablement la surcharge computationnelle tout en améliorant la précision de détection, rendant cette approche adaptée aux applications avec des contraintes de ressources.

Mots clés : Détection des Anomalies, Fouille d'Ensembles d'Éléments Fermés, Données en Flux, Règles d'Association, Traitement des Données en Temps Réel.

ملخص

يعتبر اكتشاف الشذوذ أمرًا مهما للحفاظ على سلامة البيانات عبر مجالات متنوعة، لا سيما في البيئات التي تتسم بتدفقات بيانات ديناميكية.

منهجية جديدة لاكتشاف الشذوذ تعتمد على استخراج مجموعات العناصر المغلقة وقواعد الارتباط.

من خلال الاستفادة من خوارزمية قائمة على تدفق البيانات، تحدد الطريقة المقترحة الشذوذ بكفاءة في الوقت الفعلي، مما يعزز اكتشاف مجموعات العناصر غير المنتظمة في مجموعات البيانات المعاملاتية.

يتم تقييم فعالية هذه الطريقة من خلال استخدام مجموعات بيانات مرجعية، مما يُظهر قدرتها على تحقيق دقة عالية.

بالإضافة إلى ذلك، يُقلل دمج مجموعات العناصر المغلقة بشكل كبير من الحمل الحاسوبي مع تحسين دقة الاكتشاف، مما يجعل هذه المنهجية مناسبة للتطبيقات ذات القيود على الموارد.

كلمات مفتاحية :

اكتشاف الشذوذ، التنقيب عن مجموعات العناصر المغلقة، تدفق البيانات، قواعد الارتباط، معالجة البيانات في الوقت الفعلي.