

O rozproszonym algorytmie generowania wzorców temporalnych

Do analizy danych temporalnych, czyli danych zmieniających się w czasie i ułożonych w sposób chronologiczny wykorzystuje się algorytmy z zakresu szeroko rozumianej sztucznej inteligencji i technik drążenia danych (ang. *Data mining*) [2]. Jednym z podejść wspierających taką analizę jest teoria zbiorów przybliżonych [4] opierająca się na odkrywaniu wzorców temporalnych [5]. Odkrycie tego typu wzorców pozwala na predykcję zmian danych w czasie.

Algorytmy tego typu polegają na znajdowaniu wzorców o odpowiedniej jakości w tzw. oknie czasowym, podczas przesuwania go przez cały system informacyjny z ustalonym krokiem. Jeden z takich algorytmów został zaimplementowany m.in. we własnej aplikacji *Templates* [3], który do znajdowania wzorców w oknie wykorzystuje algorytm Apriori [1] (jeden z najszybszych sekwencyjnych algorytmów wyszukiwania wzorców), a do odkrywania wzorców temporalnych algorytm Synaka [5]. Dodatkowo w aplikacji *Templates* dokonano kilku modyfikacji obu wspomnianych algorytmów przyspieszających ich działanie. Mimo to mankamentem algorytmu odkrywania wzorców temporalnych jest duża złożoność obliczeniowa (szczególnie przy niskim progu jakości wzorców), co sprawia trudności w analizie danych o dużych rozmiarach, głównie ze względu na jego czasochłonność.

W niniejszym referacie zaprezentujemy rozproszony algorytm generowania wzorców temporalnych. Motywacją do jego opracowania, było zarówno znaczne skrócenie czasu obliczeń, stosunkowo powszechny dostęp do sieci lokalnych, jak również brak konieczności stosowania znacznie droższych komputerów o dużej mocy obliczeniowej wykorzystujących przetwarzanie równoległe. Opracowany algorytm rozproszony bazuje na obu ww. algorytmach.

Prowadzone badania są wspierane przez grant własny 3 T11C 005 28 finansowany przez Ministerstwo Nauki i Informatyzacji.

- [1] R. Agrawal, R. Srikant, *Fast Algorithms for Mining Association Rules*. In Proc. of the 20th Int'l Conference on Very Large Databases, Santiago, Chile, September 1994.
- [2] K. Cios, W. Pedrycz, R. W. Świniarski, *Data Mining Methods for Knowledge Discovery*, Kluwer Academic Publishers, Boston, MA, USA 1998.

- [3] A. Paluch, W. Rzaśa, Z. Suraj, *Presentation of the System for Temporal Data Analysis*. in: H.D. Burkhard, L. Czaja, G. Lindemann, A. Skowron, H. Schlingloff, Z. Suraj (Eds.), *Proceedings of the Concurrency, Specification and Programming (CS&P'2004)*. Workshop, Vol. 3: Multiagent Systems and Applications, Caputh, Germany, September 24-26, 2004, Informatik Bericht Nr. 170, Humboldt University, Berlin 2004, pp. 533-542.
- [4] Z. Pawlak, *Rough Sets - theoretical aspects of reasoning about data*. Kluwer Academic Publishers, Dordrecht, 1991.
- [5] P. Synak, *Temporal Templates and Analysis of Time Related Data*, in: W. Ziarko and Y. Yao (Eds.), *The Second International Conference, RSCTC'2000*, *Lecture Notes in Artificial Intelligence 2001*, Springer-Verlag, Berlin, 2001, pp. 420-427.