

COMPARISON OF SELECTED TRADE CLASSIFICATION ALGORITHMS ON THE WARSAW STOCK EXCHANGE

Joanna Olbryś¹, Michał Mursztyn²

¹ Faculty of Computer Science, Białystok University of Technology, Białystok, Poland

² Student of Faculty of Computer Science, Białystok University of Technology, Białystok, Poland

Abstract: Empirical equity market microstructure research often requires knowledge about trade sides. It is important to recognize the side that initiates the transaction and to distinguish between the so called buyer- and seller-initiated trades. Unfortunately, most data sets do not identify the trade direction. Therefore, researchers rely on indirect trade classification rules to infer trade sides. The aim of this paper is to implement and compare four trade classification algorithms on the Warsaw Stock Exchange (WSE). The high-frequency data for the selected WSE-listed companies covers the period from January 3, 2005 to June 30, 2015. To check the robustness of empirical results to the choice of sample, three adjacent sub-periods of equal size: the pre-crisis, crisis, and post-crisis period are analysed. The Global Financial Crisis on the WSE is formally established as the period June 2007-February 2009.

Keywords: trade classification algorithms, intraday data, Polish stock market

PORÓWNANIE WYBRANYCH ALGORYTMÓW KLASYFIKACJI TRANSAKCJI NA GIEŁDZIE PAPIERÓW WARTOŚCIOWYCH W WARSZAWIE S.A.

Streszczenie Badania empiryczne w zakresie mikrostruktury rynku często wymagają wiedzy na temat stron transakcji. Szczególnie ważna jest identyfikacja strony inicjującej transakcję oraz podział na transakcje inicjowane przez nabywcę i sprzedającego. Niestety, większość baz danych intraday nie zawiera takich informacji. Dlatego badacze korzystają z pośrednich metod klasyfikacji. Celem pracy była implementacja i porównanie czterech algorytmów klasyfikacji transakcji na Giełdzie Papierów Wartościowych w Warszawie S.A. Badanie przeprowadzono z wykorzystaniem śróddziennych danych transakcyjnych akcji trzech

reprezentatywnych firm, należących do grup spółek dużych (grupa BIG), średnich (grupa MEDIUM) i małych (grupa SMALL). Były to odpowiednio spółki: KGH, MCI oraz ENP. Wszystkie charakteryzują się stosunkowo wysoką płynnością na tle grup, do których należą. Analizy objęły okres od 3 stycznia 2005 do 30 czerwca 2015, z wyróżnieniem trzech jednako licznych podokresów: przed kryzysem, kryzys, po kryzysie. Okres globalnego kryzysu finansowego na giełdzie warszawskiej został ustalony w sposób formalny jako przedział czasowy czerwiec 2007-luty 2009.

Słowa kluczowe: algorytmy klasyfikacji transakcji, dane intraday, polski rynek kapitałowy

Artykuł zrealizowano w ramach pracy badawczej S/WI/1/2014.