

Agnieszka Jóźwiak

Akademia Ekonomiczna we Wrocławiu

WYKORZYSTANIE PROGRAMOWANIA GENETYCZNEGO DO GENEROWANIA SYGNAŁÓW KUPNA – SPRZEDAŻY INSTRUMENTÓW FINANSOWYCH

1. Wstęp

Powszechnie znaną wśród inwestorów na rynkach finansowych strategią inwestycyjną jest długoterminowa strategia „kup i trzymaj”. Alternatywne wobec niej metody analizy technicznej są bardziej dynamiczne i generują sygnały kupna i sprzedaży oparte na krótkoterminowych fluktuacjach cen. Celem artykułu jest przedstawienie sposobu wykorzystania programowania genetycznego do automatycznego generowania sygnałów kupna i sprzedaży papierów wartościowych.

Pierwsze badania w zakresie analizy technicznej przeprowadzone zostały w latach 1960-1970. Autorzy odwoływali się do hipotezy rynku efektywnego i wykluczali stosowanie analizy technicznej na rynkach finansowych [Aleksander 1964; Jensen, Bennington 1970; Fama 1966]. Późniejsze badania wykazały jednak, że wyniki otrzymane dzięki regułom technicznym przeważają nad wynikami strategii „kup i trzymaj”, a co za tym idzie, można poszukiwać związków pomiędzy bieżącą ceną instrumentu finansowego a cenami historycznymi [Pruit, White 1988; Brock, Lakonishok, LeBaron 1992]. Analiza techniczna jest również z powodzeniem wykorzystywana do prognozowania kierunku zmian kursu na rynkach walutowych [Sweeney 1986].

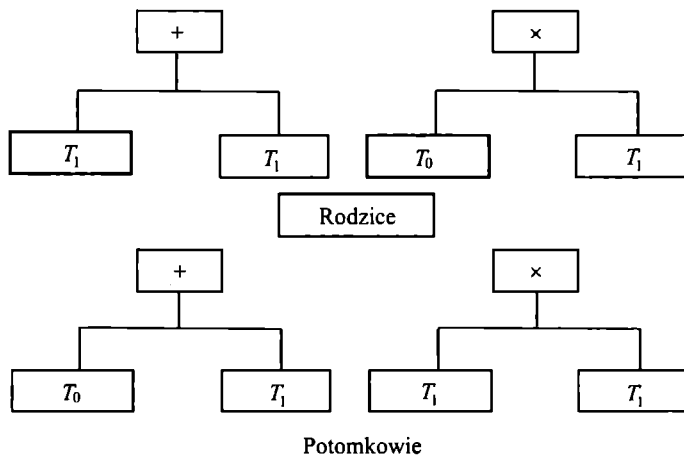
Do automatycznego generowania sygnałów kupna i sprzedaży coraz częściej używa się narzędzi z zakresu sztucznej inteligencji. Przykłady zastosowania do tego celu algorytmów genetycznych można znaleźć w pracy [Holland 1975]. Jednak postać reguł bardziej adekwatną do problemu zapewnia programowanie genetyczne [Koza 1992]. Reguły przedstawione są tu jako programy o strukturze drzew. Pierwsze próby wykorzystania programowania genetycznego nie dały satysfakcjonujących rezultatów [Chen, Yeh 1996], jednakże kolejne wyniki są obiecujące [Neely, Weller, Dittmar 1997; Neely, Weller 1999; Potvin, Soriano, Vallee 2004].

Artykuł jest próbą zastosowania programowania genetycznego do wygenerowania reguł technicznych dla wybranych spółek notowanych na warszawskiej Giełdzie Papierów Wartościowych. W analizie nie dopuszczono krótkiej sprzedaży i nie ujęto kosztów transakcyjnych.

2. Programowanie genetyczne

Programowanie genetyczne należy do klasy algorytmów ewolucyjnych. Algorytm ewolucyjny to probabilistyczny algorytm działający na populacji osobników zakodowanych w ściśle określony sposób. Populacja ewoluuje w czasie, imitując procesy znane z teorii ewolucji Darwina, takie jak selekcja, krzyżowanie i mutacja osobników. „Genetyczność” wynika z wyboru krzyżowania jako dominującego mechanizmu ewolucji, co sprawia, że programowanie genetyczne jest bliższe algorytmom genetycznym niż np. strategiom ewolucyjnym, w których dominująca jest mutacja.

Osobniki w programowaniu genetycznym kodowane są w postaci drzew, w których gałęziami (węzłami) są funkcje należące do skończonego zbioru funkcji $F = \{f_1, f_2, f_3, \dots, f_{N_{func}}\}$, gdzie N_{func} jest liczbą dostępnych funkcji. Natomiast liśćmi są zmienne, stałe i parametry należące do zbioru symboli terminalnych $T = \{a_1, a_2, a_3, \dots, a_{N_{term}}\}$, gdzie N_{term} jest liczbą dostępnych symboli terminalnych.



Rys. 1. Operacja krzyżowania. W obu drzewach zostały zaznaczone węzły wylosowane do krzyżowania

Źródło: opracowanie własne.

Procedura ewolucji osobników w ramach programowania genetycznego jest następująca:

1. Generowanie populacji początkowej.

Stworzenie P losowych programów i obliczenie dla nich dopasowania.

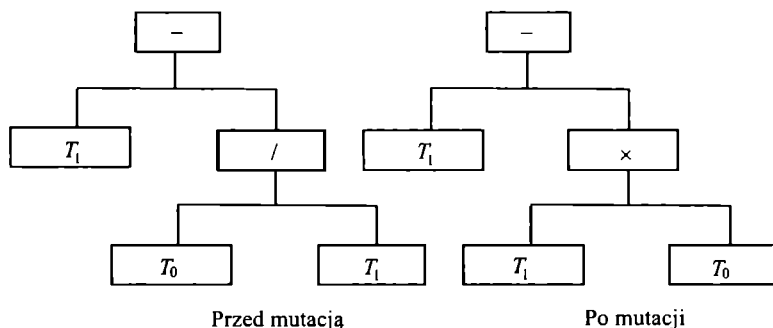
2. Selekcja: wybór osobników, które w następnych krokach zostaną poddane operacjom krzyżowania i mutacji. Prawdopodobieństwo wyboru osobnika jest zależne od jego dopasowania.

3. Krzyżowanie: polega na wymianie poddrzew pomiędzy dwoma drzewami (rys. 1).

4. Mutacja: polega na wygenerowaniu nowego poddrzewa w miejsce losowo wybranego (rys. 2).

5. Obliczenie dopasowania dla nowo powstałych osobników.

6. Powtórzenie kroków 2-5 dla ustalonej liczby generacji.



Rys. 2. Operacja mutacji

Źródło: opracowanie własne.

Końcowym rezultatem jest najlepszy program wygenerowany podczas obliczeń.

Niech $F = \{+, *, -, /\}$ i $T = \{T_0, T_1\}$. Operacje krzyżowania i mutacji osobników przedstawiono odpowiednio na rys. 1 i 2.

3. Opis metody

Programowanie genetyczne generuje reguły w postaci programów o strukturze drzew. Reguły są binarnymi funkcjami, które zwracają zero (sygnał kupna) lub 1 (sygnał sprzedaży).

Kodowanie. Programy te są konstruowane przy użyciu zbioru funkcji i symboli terminalnych. Zbiór funkcji składa się z operatorów:

- arytmetycznych: $+$, $-$, $/$, $\times$;
- logicznych: i , lub , nie ;
- relacyjnych: $<$, $>$;
- opartych na danych historycznych:
 - $\text{średnia}(n)$ – oblicza średnią z ostatnich n notowań;
 - $\text{min}(n)$ – minimalna cena z ostatnich n notowań;
 - $\text{max}(n)$ – maksymalna cena z ostatnich n notowań;
 - $\text{lag}(n)$ – cena sprzed n notowań.

Terminalami są liczby całkowite z przedziału $<0; 100>$.

Jednoczesne występowanie operatorów logicznych i arytmetycznych w ramach jednego drzewa nakłada pewne ograniczenia na jego strukturę. Operatory arytmetyczne będą umiejscowione w dolnej części drzewa, a logiczne w górnej (włączając również korzeń drzewa). Łącznikiem pomiędzy tymi dwoma rodzajami operatorów są operatory relacyjne. Operator relacyjny może się także znajdować w korzeniu drzewa.

Obliczenie dopasowania. Dla każdego osobnika obliczane jest dopasowanie zdefiniowane jako średnia różnica pomiędzy kwotą, którą dysponuje inwestor na koniec kwartału, jeżeli sygnały kupna lub sprzedaży generowane są co tydzień, a kwotą, którą miałby, gdyby zastosował strategię „kup i trzymaj”. Kapitał początkowy wynosi 1000 PLN.

Dane. Obliczenia dotyczyły następujących spółek notowanych na warszawskiej Giełdzie Papierów Wartościowych: BRE Bank SA, Dębica SA, Wólczanka SA oraz KGHM Polska Miedź SA.

Zbiór danych treningowych obejmował 1400 notowań z okresu od 10.07.1997 do 17.02.2003.

Zbiór danych testowych obejmował 558 notowań z okresu od 18.02.2003 do 05.05.2005 roku.

Parametry programowania genetycznego. Aby uzyskać optymalne parametry dla programowania genetycznego, zostały przeprowadzone wstępne eksperymenty. Na ich podstawie wybrano optymalne wartości parametrów algorytmu. Przedstawiono je w tab. 1. Po przeprowadzeniu testów dla 10 000, 20 000 oraz 30 000 osobników wielkość populacji została ustalona na 30 000, pomimo znaczącego wydłużenia czasu obliczeń. Liczbę generacji ustalono na 50 ze względu na brak znacznej poprawy jakości rozwiązania w późniejszych generacjach. Pozostałe parametry w większości zostały przyjęte arbitralnie jak w pracy [Koza 1992].

Tabela 1. Parametry programowania genetycznego

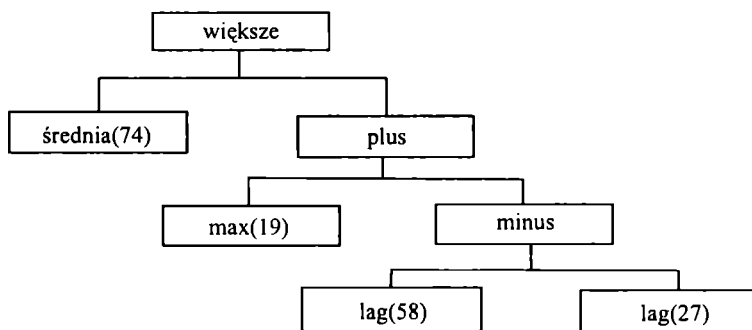
Rozmiar populacji	30 000
Liczba generacji	50
Prawdopodobieństwo mutacji	0,05
Maksymalna liczba węzłów (gałęzi) w drzewie w populacji początkowej	14
Maksymalna liczba węzłów (gałęzi) w drzewie w kolejnych generacjach	20

Źródło: opracowanie własne.

3. Wyniki i wnioski

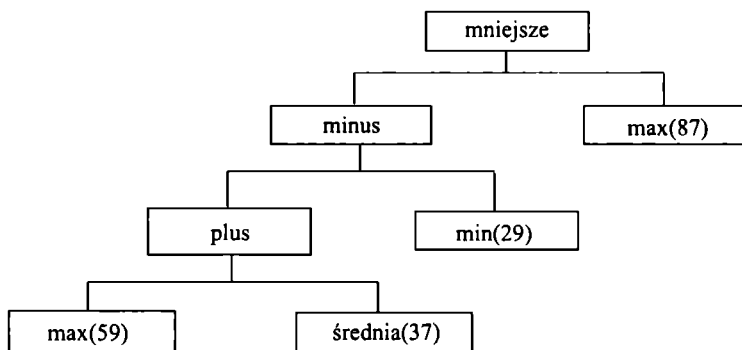
Na kolejnych rysunkach (3-6) przedstawiono schematy blokowe otrzymanych reguł kupna sprzedaży dla wybranych spółek giełdowych.

Średnie wyniki osiągnięte w okresie kwartalnym za pomocą reguł wygenerowanych w programowaniu genetycznym są wysokie zarówno dla danych treningo-



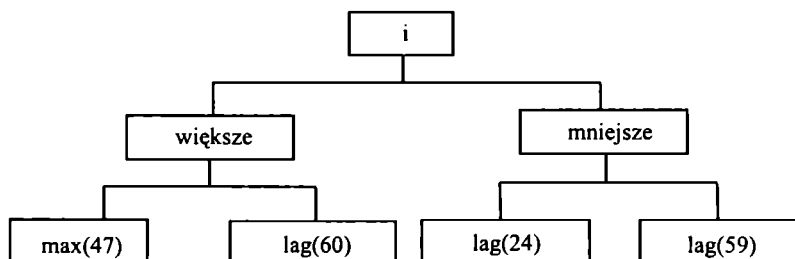
Rys. 3. Najlepsza reguła techniczna wygenerowana dzięki programowaniu genetycznemu dla spółki BRE Bank SA

Źródło: opracowanie własne.



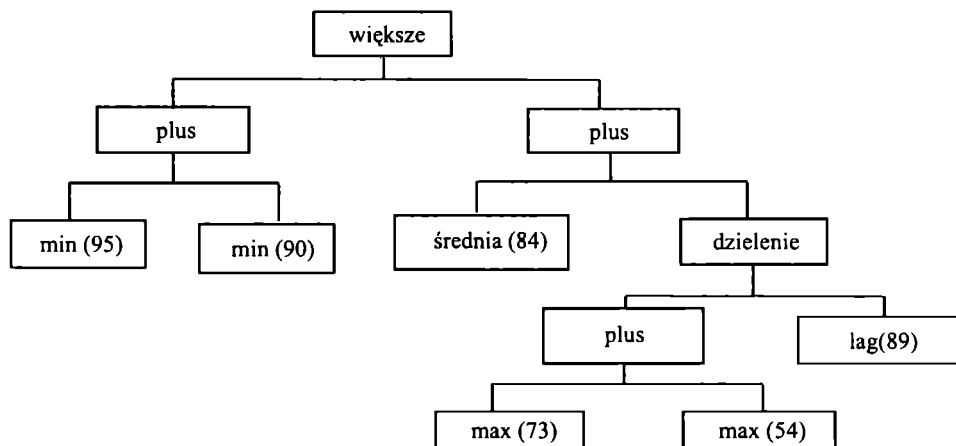
Rys. 4. Najlepsza reguła techniczna wygenerowana dzięki programowaniu genetycznemu dla spółki Dębica SA

Źródło: opracowanie własne.



Rys. 5. Najlepsza reguła techniczna wygenerowana dzięki programowaniu genetycznemu dla spółki KGHM Polska Miedź SA

Źródło: opracowanie własne.



Rys. 6. Najlepsza reguła techniczna wygenerowana dzięki programowaniu genetycznemu dla spółki Wólczanka SA

Źródło: opracowanie własne.

wych, jak i testowych (tab. 2). Jednak minimalne wartości kapitału na koniec kwartału, jeżeli sygnały kupna i sprzedaży generowane są co tydzień, są znacznie niższe od kapitału początkowego wynoszącego 1000 PLN. Tak duże odchylenie od wartości średniej sprawia, że stosowanie reguł przedstawionych na rys. 3-6 jest obciążone dużym ryzykiem.

W porównaniu ze strategią „kup i trzymaj” reguły otrzymane w wyniku programowania genetycznego dają średnio dla danych treningowych i testowych wyższe stopy zwrotu (tab. 3).

Tabela 2. Średnia, minimalna i maksymalna kwota otrzymana po kwartalnym okresie inwestycji, jeżeli sygnały kupna-sprzedaży były generowane co tydzień, a kwota zainwestowana na początku kwartału to 1000 PLN [zł]

	Dane treningowe			Dane testowe		
	min	średnia	max	min	średnia	max
BRE	702,5	1161,33	1984,4	852	1089,69	1542,4
Dębica	572,5	1064,26	1926,8	846	1065,33	1576
KGHM	706,9	1042,84	2069,5	746,9	1131,28	1553,4
Wólczanka	862,2	1170,6	8191,9	762,7	1167	2237,1

Źródło: opracowanie własne.

Niekorzystną cechą z punktu widzenia jakości rozwiązania jest zmniejszanie się różnorodności populacji wraz z postępem obliczeń. Znaczna część końcowej populacji to osobniki podobne. W dalszych badaniach planowane jest wprowadzenie mechanizmów immunologicznych w celu zapewnienia zróżnicowania populacji.

Tabela 3. Średnia kwartalna stopa zwrotu dla reguł otrzymanych za pomocą programowania genetycznego i dla strategii „kup i trzymaj” [%]

	Dane treningowe	Dane treningowe (strategia „kup i trzymaj”)	Dane testowe	Dane testowe (strategia „kup i trzymaj”)
BRE	16,13	13,71	8,96	4,61
Dębica	6,42	4,58	6,53	5,47
KGHM	4,28	0,50	13,12	12,06
Wólczanka	17,05	13,12	16,69	18,50

Źródło: opracowanie własne.

Ponieważ jedynym kryterium wyboru najlepszego osobnika jest średni zysk, metoda nie uwzględnia preferencji inwestora. Planuje się wprowadzenie dodatkowych kryteriów wpływających na wartość dopasowania (np. miar ryzyka).

Ograniczeniem jest także możliwość zastosowania metody do jednego waloru. Planowana jest próba zastosowania programowania genetycznego do stworzenia portfela akcji o strukturze dynamicznej.

Literatura

- Aleksander S.S., *Price Movements in Speculative Markets: Trends of Random Walk*, [w:] *The Random Character of Stock Market Prices*, red. Cootner P.H, MIT Press, Cambridge 1964.
- Brock W., Lakonishok J., LeBaron B., *Simple Technical Trading Rules and the Stochastic Properties of Stock Returns*, „Journal of Finance” 1992 nr 47.
- Chen S-H., Yeh C-H., *Toward a Computable Approach to the Efficient Market Hypothesis: an Application of Genetic Programming*, „Journal of Economic Dynamic and Control” 1996 nr 21.
- Fama E., *Efficient Capital Markets: a Review of Theory and Empirical Work*, „Journal of Finance” 1970 nr 39.
- Holland J.H., *Adaptation in Natural and Artificial Systems*, The University of Michigan Press, Ann Arbor 1975.
- Jensen M., Bennington G., *Random Walks and Technical Theories: Some Additional Evidences*, „Journal of Finance” 1970 nr 25.
- Koza J.R., *Genetic Programming – On programming Computers by Means of Natural Selection*, MIT Press, Cambridge, 1992.
- Neely C., Weller P., *Technical Trading Rules in the European Monetary System*, „Journal of International Money and Finance” 1999 nr 18.
- Neely C., Weller P., Dittmar R., *Is Technical Analysis in the Foreign Exchange Market Profitable? A Genetic Programming Approach*, „Journal of Financial and Quantitative Analysis” 1997 nr 32.
- Potvin J., Soriano P., Vallee M., *Generating Trading Rules on the Stock Markets with Genetic Programming*, „Computers and Operations Research” 2004 vol. 31 nr 7.
- Pruit S.W., White R.E., *The CRISMA Trading System: Who Says Technical Analysis Can't Beat the Market?* „Journal of Portfolio Management” 1988 nr 14.
- Sweeney R.L., *Beating the Foreign Exchange Market*, „Journal of Finance” 1986 nr 14.

GENERATING FINANCIAL TRADING RULES WITH GENETIC PROGRAMMING

Summary

Technical analysis is aimed at devising trading rules capable of exploiting short-term fluctuations on the financial markets. This market timing approach may be a profitable alternative to the buy-and-hold approach. This paper presents generating short-term trading rules on the stock market with genetic programming. Computational results, based on historical pricing data, for 4 Polish companies listed on the Warsaw stock exchange market are included.