



Adam Mrozek

Uniwersytet Ekonomiczny w Katowicach
Wydział Informatyki i Komunikacji
Katedra Informatyki
adam.mrozek@ue.katowice.pl

Dariusz Badura

Wyższa Szkoła Biznesu w Dąbrowie Górniczej
Wydział Nauk Stosowanych
Katedra Informatyki
dbadura@wsb.edu.pl

PROBLEM STROJENIA ALGORYTMU OPTYMALIZACJI ROJEM CZĄSTEK W OPTYMALIZACJI CIĄGŁEJ

Streszczenie: Optymalizacja wybranego zagadnienia, polegająca na znalezieniu analitycznego rozwiązania wyznaczającego ekstremum opisującego to zagadnienie funkcji, jest bardzo często złożona. Analityczne rozwiązanie staje się czasem niemożliwe, szczególnie w przypadku, gdy funkcja jest sformułowana w sposób uwikłany. W wielu przypadkach nie istnieją też metody automatycznego rozwiązywania takich formuł. Do rozwiązywania wielu problemów optymalizacyjnych skutecznym narzędziem okazała się optymalizacja rojem cząstek (ang. *Particle Swarm Optimization*, PSO). Sam algorytm bywa także stosowany jako część innych niedeterministycznych algorytmów, tworząc konstrukcje hybrydowe. Biorąc pod uwagę skuteczność znajdowania rozwiązania, wśród innych podobnych metod optymalizacji algorytm PSO nie plasuje się na czołowym miejscu. Stąd liczne próby modyfikacji oraz ustalenia najbardziej optymalnych i uniwersalnych parametrów algorytmu PSO.

W artykule przedstawiono wyniki badań efektywności podstawowej wersji algorytmu optymalizacji rojem cząstek (PSO) dla problemów ciągłych o różnej skali złożoności. Autorzy podjęli także próbę oceny kosztu strojenia tego algorytmu dla małych problemów.

Słowa kluczowe: PSO, optymalizacja, strojenie.

JEL Classification: C60, C61, C68, C69.

Wprowadzenie

Optymalizacja procesów zachodzących w dziedzinie ciągłej jest zasadniczym problemem w wielu obszarach życia: w technice, zarządzaniu, ekonomii itp. Znanych jest wiele metod wyznaczania rozwiązań optymalnych dla prostych

zagadnień badanych zależności kilku zmiennych. Trudności pojawiają się, gdy zmiennych decyzyjnych jest więcej, przebieg funkcji celu nie jest gładki lub opisany złożonymi zależnościami, istnieją dwa lub więcej kryteria oceny lub gdy nałożone są nietrywialne ograniczenia. Dotychczas opracowane metody ograniczają się do rozwiązania każdego z powyższych przypadków z osobna. Kiedy znane są przebiegi funkcji celu w sposób analityczny, możliwe jest znalezienie wartości zmiennych decyzyjnych w taki sam sposób.

Znalezienie analitycznego rozwiązania wyznaczającego ekstremum dowolnej funkcji bywa jednak kłopotliwe lub czasem wręcz niemożliwe, gdy jest ona sformułowana w sposób uwikłany. Znane z literatury metody prowadzą do uzyskiwania złożonych postaci rozwiązań, które następnie wymagają redukcji. Dodatkowym zagadnieniem, poza dużym kosztem obliczeń, jest zapewnienie należytej dokładności i wiarygodności otrzymanych wyników. Często koszt obliczeń związany z weryfikacją dopuszczalności danego rozwiązania okazuje się znacznie mniejszy niż znalezienie rozwiązania optymalnego, trudnością natomiast jest znalezienie takiego rozwiązania, które spełniałoby warunek dopuszczalności. W licznych (także cytowanych) publikacjach autorzy dowodzą skuteczności regulacji parametrów opisanej metody zastosowanej dla rozwiązywania złożonych problemów (o wymiarze przestrzeni D_{max} większym niż 10 a nawet 30). Osiągnięto wyniki kosztem dużych nakładów obliczeniowych.

Wobec powyższego zostały wyciągnięte hipotezy badawcze:

1. Dla wielu średnio złożonych zadań dokładne ustalenie wartości wszystkich parametrów nie gwarantuje radykalnej poprawy skuteczności metody.
2. Zaproponowane różne odmiany algorytmu dla niezłożonych problemów możliwe są do uproszczenia do jednej wspólnej postaci, w której wartości stałych parametrów można ograniczyć do określonego zakresu.
3. Przeprowadzenie złożonych i kosztownych obliczeniowo procedur strojenia metody dla omawianych problemów nie jest adekwatne do uzyskiwanych rezultatów.

Wykonane eksperymenty obliczeniowe potwierdzają wysunięte hipotezy.

Niniejszy artykuł opisuje procedurę badawczą mającą na celu stwierdzenie lub zaprzeczenie potrzeby strojenia algorytmu optymalizacji rojem cząstek (PSO) w stosunku do zadań o średniej skali trudności (wymiar przestrzeni $D_{max}=8$) realizowanych w warunkach mocno ograniczonego możliwości sprzętowych.

W zadaniach optymalizacji wymiar w przestrzeni poszukiwań stanowi liczbę niezależnych zmiennych decyzyjnych. W badaniach zastosowano różne wersje wspomnianego algorytmu, poczynając od podstawowej kanonicznej wersji,

zaproponowanej w 1995 r. przez Kennedy'ego i Eberharta [1995], a kończąc na wersjach z wagą inercji (ω) oraz współczynnikiem ścisku (χ). Przystępując do planowania eksperymentów, autor niniejszego artykułu postawił hipotezę, iż w większości przypadków o średniej skali trudności przeprowadzenie dostrajania parametrów nie jest uzasadnione. Potwierdzenie tej tezy przemawiałoby za stwierdzeniem, że badana metoda wykazuje cechę samoadaptacyjnej, co w połączeniu z tym, że w stosunku do innych metod charakteryzuje się względną prostotą konstrukcji, czyni ją łatwą do realizowania w prostych układach sterowania.

Na podstawie analizy opisanych w literaturze przypadków można wysnuć hipotezę, iż wśród wielu parametrów metody możliwe jest wyselekcjonowanie tych, których zmiany mają istotny wpływ na zachowanie algorytmu. Właściwe ich ustawienie przyczynia się do uzyskania pożądanych cech metody, a ponadto ustawienie nie jest zazwyczaj stałe. Jednym z takich parametrów, które decydują o złożoności modelu, a co za tym idzie – koszcie obliczeń oraz jakości, jest zastosowana topologia. Topologia takiego układu ma ścisły związek z pojęciem sąsiedztwa, które może być definiowane w różny sposób. Badania ograniczono do metod z pojedynczym najlepszym wzorcem (w wersji globalnej – globalbest).

Szeroki przegląd dokonań w zakresie algorytmów PSO oraz porównania właściwości proponowanych rozwiązań można znaleźć w artykule Bonyadi i Michalewicza [2017].

1. Założenia i koncepcja badań

Wiele prac teoretycznych oraz przeprowadzonych eksperymentów [Ströyner, 2014; Bansal et al., 2011; Li-ping, Huan-jun, Shang-xu, 2005] miało na celu ustalenie optymalnej wielkości poszczególnych parametrów metody, a także wykazania przewagi zaproponowanych modyfikacji nad pozostałymi wersjami. Jako algorytmy porównawcze wybierano uproszczone wersje metody kanonicznej, w których występował tylko jeden ze składników (kognitywny lub społeczny) lub były one równe, tj. zostały ustalone na podstawie teoretycznych rozważań. Teoretyczne badania zbieżności podstawowego algorytmu PSO prowadzili M. Clerc [2012] oraz E. Ozcan i Ch.K. Mohan [2010]. Prace teoretyczne przeprowadzano na uproszczonym modelu. Uproszczenia polegały m.in. na wyeliminowaniu czynnika losowego, przyjęciu pewnych założeń co do postaci funkcji celu, ograniczeniu liczby badanych parametrów czy przyjęciu wartości skrajnych. Zabiegi te umożliwiły zastosowanie metod analitycznych i wyprowadzenie ogólnych warunków zbieżności metody. Ozcan i Mohan [2010], Clerc [2012], Kennedy i Engelbrecht [1995] oraz van den Bergh wspólnie z Engel-

brechtem [2006] przedstawili równanie modelu PSO w postaci stochastycznej, a następnie dokonując uproszczeń, wyprowadzili równanie na kolejną pozycję cząsteczki. Wprowadzili oni do modelu PSO tzw. czynnik zdolności eksploracyjnej stada – λ_{maxD} . Im jest on większy, tym szerszy jest zasięg przeszukiwania. Dobrym wyborem jest wartość bliska 0,9. Zakres zmian podstawowych parametrów i ich wzajemnej relacji ustalono następująco:

$$-1 < \omega < 1 \text{ i } c_1 + c_2 < 4(1 + \omega) \quad (1)$$

gdzie ω oznacza współczynnik inercji, c_1 i c_2 to stałe przyspieszenia (czynnik kognitywny i społeczny).

W innym uproszczeniu przedstawionym przez A. Adriansyaha i S.H.M. Amina [2006] przyjęto następujące założenia:

$$U_1 = U_2 = \frac{1}{2} \text{ oraz } c = \frac{c_1 + c_2}{2} \text{ i } p = \frac{c_1}{c_1 + c_2} pbest + \frac{c_2}{c_1 + c_2} qbest \quad (2)$$

gdzie:

U_1, U_2 – losowe zmienne o rozkładzie jednostajnym zastąpione wartością stałą.

Podstawiając powyższe do wzoru na ruch cząstek, otrzymano równanie kwadratowe, którego analiza pierwiastków dała następujące warunki konieczne, aby cząsteczka osiągnęła punkt równowagi:

$$\omega < 1, c > 0, 2\omega - c + 2 > 0 \quad (3)$$

Gdy pierwiastki równania są złożone z części rzeczywistej i urojonej, cząsteczki wykonują harmoniczne oscylacje wokół punktu równowagi. Jest to równoważne spełnieniu następującej nierówności:

$$\omega^2 + c^2 - 2\omega c - 2\omega - 2c + 1 < 0 \quad (4)$$

Bliższa analiza wykazała zależności, dla których trajektoria posiada odpowiednie zachowanie. Wydzielono 5 charakterystycznych regionów. Wynikiem badań teoretycznych jest przedstawienie zakresów stosowalności poszczególnych wartości wybranych parametrów oraz ich wzajemnych relacji: ω do $c_1 + c_2$ oraz c_1 do c_2 . Natomiast nie wytypowano żadnej uniwersalnej ich wartości. Dało to podstawę do próby standaryzowania metody PSO dla celów porównawczych oraz rozwoju prac nad modyfikacją postaci równania ogólnego PSO.

Poszczególne standardy SPSO 2006, 2007 i 2011 [Clerc, 2012] wprowadzały zarówno ograniczenia co do wielkości poszczególnych parametrów, jak i sposób ich zmian. W początkowych pracach polepszenie skuteczności algorytmu upatrywano w doborze odpowiednich warunków początkowych metody, gdyż w szczególnych przypadkach algorytm okazywał się rozbieżny. Regulację pomiędzy eksploracją a eksploatacją zapewnić miała odpowiednia strategia reakcji

na przekroczenie dopuszczalnych warunków brzegowych. Modyfikacje zaproponowane przez Shi i Eberharta [2012] wprowadzają losową zmienną o rozkładzie jednostajnym $r = U(0,1)$ oraz obie wartości c_1 i c_2 bliskie 2, a także maksymalną prędkość cząstki v_{max} . Przeprowadzane były także równoległe badania empiryczne metody z innymi rozkładami zmiennej losowej, przez którą mnożono czynniki kognitywne i społeczne (c_1 , c_2). Alternatywne postacie równania kanonicznego PSO zaproponowali Clerc oraz Kennedy, Carlisle, Jiang, Luo i Yang oraz Trelea czy Dozier [zob. Ströyner, 2014, s. 7]. Ich wzajemne korelacje prezentuje tabela 1.

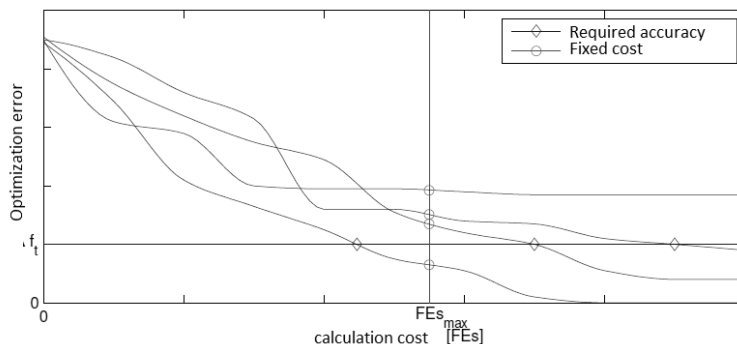
Tabela 1. Porównanie parametrów dla modelu ze stałą wagą inercji i współczynnikiem ściśku

Autor	Ω	c_1	c_2	λ_{maxD}	$\varphi_1=c_1/\omega$	$\varphi_2=c_2/\omega$	$\varphi=(\varphi_1+\varphi_2)$	χ	χ_1	k	k_1
Clerc oraz Kennedy	0.729	1.494	1.494	0.942	2.04938	2.04938	4.0987654	0.731257	0.66698	0.99691	1.09298
Trelea	0.6	1.7	1.7	0.889	2.83333	2.83333	5.6666667	0.296743	0.24960	2.02195	2.40384
Carlisle oraz Dozier	0.729	2.041	0.948	0.975	2.79972	1.30041	4.1001372	0.729688	0.66527	0.99906	1.095801
Jiang, Luo i Yang	0.715	1.7	1.7	0.995	2.37762	2.37762	4.7552448	0.430076	0.36796	1.6625	1.943123

Źródło: Opracowanie własne.

Przedstawione sugestie nie dają jednoznacznej odpowiedzi na pytanie, jaka jest najlepsza postać dla algorytmu PSO. Nie ma jednoznacznych miar jakości rankingu ważności cech. Ponadto oprócz zastosowań stacjonarnej optymalizacji pożądaną cechą staje się zdolność do śledzenia zmieniającego się stanu [Carlisle, 2002]. Konkluzją wyciągniętą z obserwacji cytowanych badań jest wniosek, że jedyny sposób na efektywne działanie algorytmu to przeprowadzenie ustalenia parametrów dla każdego z problemów indywidualnie. Kolejne modyfikacje dały metody dynamicznie zmienne według zadanego schematu lub metody adaptacyjne [Bansal et al., 2011].

W ramach wielu badań zaproponowano zakres i postać głównych parametrów metody oraz wskaźniki oceny ich efektywności. Umożliwiło to porównanie ich między sobą, jak również między innymi metodami optymalizacyjnymi.

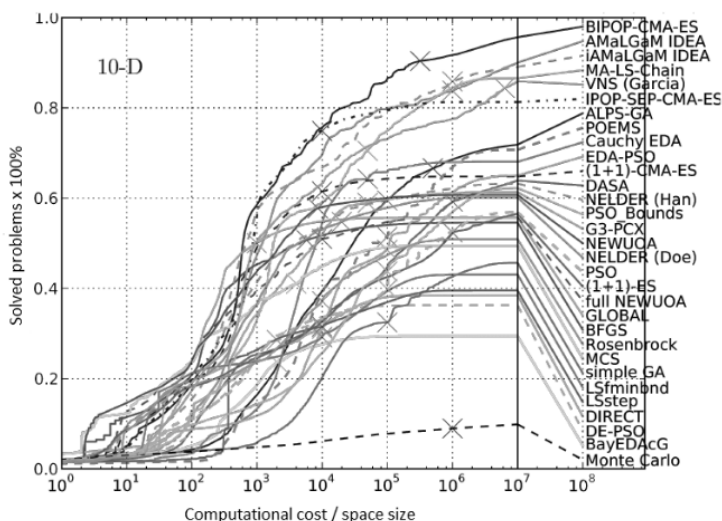


Rys. 1. Błąd obliczeń vs koszt

Źródło: Opara, Arabas [2012].

Porównania głównie dotyczą liczby uruchomień funkcji celu (FE) przy założonym poziomie akceptowalnych wyników błędu optymalizacji lub wysokości jego poziomu osiągniętego po wykonaniu założonej liczby przebiegów [Opara, Arabas, 2012].

Innym kryterium porównawczym był udział uruchomień algorytmu zakończonych sukcesem w ogólnej liczbie uruchomień. W tym przypadku – przyjmując górny limit maksymalnej liczby wyliczeń funkcji ewaluacji $FE_{max} = 10^7$. Uśrednione wyniki dla różnych testowych funkcji i metod opracowali K. Opara i J. Arabas [2012].



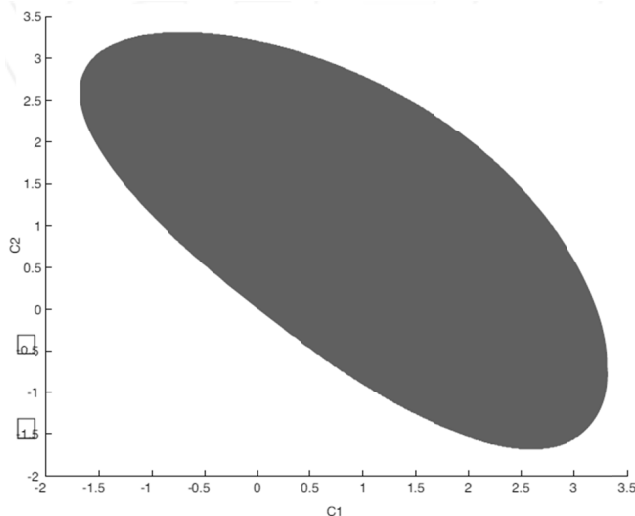
Rys. 2. Efektywność algorytmów według K. Oparę i J. Arabasa

Źródło: Opara, Arabas [2012].

Z obliczeń Opary i Arabasa wynika, że badane wersje algorytmu PSO, w porównaniu z innymi metodami, nie wypadają najlepiej. Średni procent rozwiązanych problemów to odpowiednio ok. 40% DE-PSO i 55% dla PSO. Można tu zauważyć charakterystyczne punkty załamania, po których kontynuacja obliczeń nie przynosi znacznej poprawy. Dla większości metod kształtują się między 10^3 a 10^5 , dla PSO są to odpowiednio: 10^4 i 10^6 . Wyniki dotyczą 10-wymiarowej przestrzeni. Takie wyniki uzasadniają dalsze prace nad polepszeniem skuteczności. Zaobserwowano jednak, że wyraźne polepszenie widoczne jest dla dużych wymiarów przeszukiwanej przestrzeni. Przeprowadzone badania miały na celu potwierdzić lub obalić ustalone hipotezy badawcze:

1. Dla wielu średnio złożonych zadań dokładne ustalenie wartości wszystkich parametrów nie gwarantuje radykalnej poprawy skuteczności metody.
2. Zaproponowane różne odmiany algorytmu dla niezłożonych problemów możliwe są do uproszczenia do jednej wspólnej postaci, w której wartości stałych parametrów można ograniczyć do określonego zakresu.
3. Przeprowadzenie złożonych i kosztownych obliczeniowo procedur strojenia metody dla omawianych problemów nie jest adekwatne do uzyskiwanych rezultatów.

Postawione na wstępie pytanie – czy można pominąć strojenie wszystkich parametrów – autorzy niniejszego artykułu, nawiązując do poprzedników, opracowali, wykonując wielokrotnie obliczeń eksperymentalnych. Przyjęto następujące zakresy zmian parametrów: $\omega \in [-1, 1]$ z krokiem 0,25, $c_1 \in [-1, 5, 3, 5]$ i c_2 z tego samego zakresu dla obszaru wewnątrz figury wyznaczonej na wykresie (rys. 3).



Rys. 3. Współzależności gwarantujące zbieżność procesu iteracyjnego $\{DX_t\}$

Źródło: Jiang, Luo, Yang [2007].

Zakres zmian ustalono na podstawie analizy szczególnych przypadków. Stwierdzenie wyraźnej tendencji poprawy wskaźnika jakości wraz z monotonicznie zmieniającą się wartością parametru sugerowałoby, iż metoda jest czuła na jego zmianę. Natomiast występowanie ekstremum dla różnych wartości w zależności od badanej funkcji przywodziłoby na myśl, że regulacja tego parametru jest wskazana, gdy ogólna postać (charakterystyka) zależności funkcyjnych nie jest znana.

2. Opis eksperymentów

W celu przeprowadzenia eksperymentów wykonano program obliczeniowy napisany w języku C++ z użyciem biblioteki OpenMP. Zaimplementowano ogólną postać algorytmu PSO w wersji asynchronicznej. Stan bieżący obliczeń monitorowano z określonym interwałem liczby wykonanych funkcji ewaluacji. W każdym punkcie kontrolnym obliczane były wartości średnie i skrajne uzyskane w danej populacji. Obliczenia wykonywane były na laptopie Toshiba Satellite A6630 11M o parametrach podanych na rysunku 4.

Wykorzystanie	Szybkość	Maksymalna szybkość:	1.60 GHz
99%	1.58 GHz	Gniazda:	1
		Rdzenie:	4
Procesy	Wątki	Dojścia	Procesory logiczne:
84	1163	32272	8
			Wirtualizacja:
			Włączone
			Pamięć podręczna poziomu 1:
			256 KB
			Pamięć podręczna poziomu 2:
			1.0 MB
			Pamięć podręczna poziomu 3:
			6.0 MB
Czas pracy			
4:03:03:39			

Rys. 4. Parametry techniczne środowiska obliczeniowego

Źródło: Opracowanie własne.

Dla każdego z przeprowadzonych eksperymentów ustalono warunki początkowe symulacji: typy wariantów badanych algorytmów obejmujących standardową postać kanoniczną z wagą inercji (ω), jak i ze współczynnikiem ścisku (χ), przyjmując proponowane w literaturze standardowe wielkości dla czynnika kognitywnego oraz społecznego. Jako wariant zastosowano PSO globalbest synchronicznie uaktualniane. Określone zostały: zbiór testowanych funkcji, liczba powtórzeń eksperymentów, ograniczenie dziedziny oraz liczba niezależnych zmiennych (wymiarów). Jako zbiór testowych problemów wykorzystano cztery funkcje: De Jonga (sferyczna), równoległy do osi hiperelipsoid, De Jonga (hiperelipsoid osi równoległej), Schwefela (obrócony hiperelipsoid).

W obliczeniach przyjęto, że w każdym z wymiarów dozwolone wartości dziedziny mieszczą się w przedziale $x \in [-x_{max}; x_{max}]$ dla x_{max} równego 20. Obliczenia wykonano dla 4, 6 i 8 wymiarów. Eksperymenty były wykonywane 100-krotnie dla każdej instancji, przy przyjęciu trzech przebiegów dla każdej z wartości współczynnika odbicia $\gamma \in [0,1]$ z krokiem 0,03.

Kryteria zakończenia poszczególnych faz obliczeń uwzględniały założenie różnej dynamiki postępów realizacji algorytmu w trakcie jego realizacji. Wprowadzono 7 faz w zależności od konsumpcji zasobów, tzn. czasu procesora mierzonego liczbą wywołań (B_i) funkcji ewaluacji FEs . Nakład obliczeniowy dla poszczególnych faz dobrano proporcjonalnie do liczby wymiarów przypadających na jedną pracującą cząstkę i określono dla kolejno dobranych budżetów [Yuan et al., 2010]. Pierwszy z nich wyliczony został według wzoru:

$$B_1 = 40 \cdot (2 \cdot k + 2) \quad (5)$$

gdzie k jest wymiarem problemu. Budżety kolejnych poziomów (i) to zgodnie ze wzorem:

$$B_i = 2^{i-1} \cdot B_1 \quad (i = 2, 3, 4) \quad (6)$$

Ustalony został maksymalny koszt czasowego wykonania wszystkich 30 uruchomień w ramach jednej instancji (3 godz.):

- liczby wywołań funkcji ewaluacji przypadającej na jeden wymiar: 10 000,
- liczby iteracji zależnej od wielkości populacji ($N = 1$ do 200, z krokiem 2, 5 i 10),
- zainicjowania zmiennych początkowych liczbami losowymi za pomocą generatora liczb losowych Marsen Twister z zadanego przedziału.

Dla położenia przyjęto: $\pm x_{max}$, a dla prędkości: $\pm x_{max} / 2$. W eksperymentach nie zastosowano kryterium przerwania obliczeń wynikającego z braku postępów. Ostatnim etapem badań było wykonanie zestawu symulacji oraz zebranie wyników.

2.1. Eksperyment 1

W pierwszym eksperymencie wykorzystano standardową postać równania SPSO 2007 z wagą inercji ω . Przyjęto w nim następujące założenia:

- topologia – *gbest*: dla roju o liczności S ,
- warunki brzegowe – odbijanie ze współczynnikiem odbicia γ :

$$\begin{cases} \text{gdy } (x_{i,d}(t+1) < \min_d) & \text{to } \begin{cases} x_{i,d}(t+1) = \min_d \\ v_{i,d}(t+1) = \gamma \cdot v_{i,d}(t) \end{cases} \\ \text{gdy } (x_{i,d}(t+1) > \max_d) & \text{to } \begin{cases} x_{i,d}(t+1) = \max_d \\ v_{i,d}(t+1) = -\gamma \cdot v_{i,d}(t) \end{cases} \end{cases} \quad (7)$$

- uaktualnianie – szeregowo,
- rozmiar roju: $S = 10 + \lceil 2\sqrt{d} \rceil$,
- maksymalna ilość iteracji $t_{max} = 10\,000$,
- funkcjami dopasowania są wybrane z zestawu funkcji testowych w dziedzinie liczb rzeczywistych ograniczone do przedziałów podanych dla każdej funkcji symetrycznie ± 20 ,
- uaktualnienie prędkości i położenia zgodnie ze standardem SPSO 2007 następuje według wzoru:

$$\begin{cases} v_{i,d}(t+1) = v_{i,d}(t) + U(0, c_1) (p_{i,d}(t) - x_{i,d}(t)) + U(0, c_2) (l_{i,d}(t) - x_{i,d}(t)) \\ x_{i,d}(t+1) = v_{i,d}(t) + x_{i,d}(t) \end{cases} \quad (8)$$

gdzie:

t – numer aktualnego kroku czasowego,

$x_i(t)$ – pozycja i -tej cząstki w chwili t ,

$v_i(t)$ – prędkość i -tej cząstki w chwili t ,

$p_i(t)$ – ostatnia najlepsza pozycja i -tej cząsteczki,

$l_i(t)$ – najlepsza pozycja znaleziona przez sąsiednie cząsteczki,

$U(0, c_1)$ i $U(0, c_2)$ są realizacjami zmiennych losowych o rozkładzie jednostajnym.

Indeksy i oraz d oznaczają odpowiednio numer cząstki (i) oraz wymiar jej położenia (d). Wartości c_1 i c_2 są stałe dla całego algorytmu. Dla większej precyzji w miejsce U wprowadzono następujące oznaczenia: $r_{1,j} \cdot c_1$ i $r_{2,j} \cdot c_2$: $r_1, r_2 = U(0,1)$, wskazując, że są one stałe dla danej iteracji, lecz różne dla każdego z wymiarów:

$$v_{i,d}(t+1) = \omega v_{i,d}(t) + c_1 r_{1,d}(t) \cdot (p_{i,d}(t) - x_{i,d}(t)) + c_2 r_{2,d}(t) \cdot (l_{i,d}(t) - x_{i,d}(t)) \quad (9)$$

$$\omega = \frac{1}{2 \ln(2)} \approx 0,721 \text{ i } c_1 = c_2 = c = \frac{1}{2} + \ln(2) \approx 1,193 \quad (10)$$

W kolejnych eksperymentach ustalono współczynnik przenikania γ (według tabeli 2).

2.2. Podsumowanie pierwszego eksperymentu

Wyniki pierwszej próby ustalenia współczynnika przenikania wykazały, że liczba sukcesów rośnie proporcjonalnie wraz z jego wzrostem. Badania poprzestano przy wartości tego parametru równej 0,5, czyli wytraceniu połowy końcowej prędkości. Wynik ten nadal okazał się o połowę gorszy od wyników przedstawionych przez Oparę i Arabasa [2012]. Pozwoliło to przypuszczać, że wariant z całkowitym wytłumieniem prędkości jest najgorszy.

Tabela 2. Efektywność dla średniego kosztu

γ	Średnia liczba sukcesów	Odchylenie standardowe	Średni koszt	Odchylenie standardowe	Wartość procentowa sukcesów
0,05	114,818	16,142	35 828,182	2 448,950	13,67%
0,1	129,364	13,684	35 180,545	859,323	15,40%
0,2	131,818	12,392	35 437,182	1 795,856	15,69%
0,25	135,909	14,781	34 838,545	1 137,183	16,18%
0,3	149,818	20,029	34 199,818	1 568,310	17,84%
0,4	143,091	14,335	34 564,818	871,516	17,03%
0,5	149,636	14,583	33 750,000	1 271,951	17,81%

Źródło: Opracowanie własne.

2.3. Eksperyment drugi

W drugim eksperymencie zostały zmienione warunki początkowe, w których uzależniono początkową prędkość od odległości od krańców przedziału:

$$\begin{cases} x_i(0) = U(\min_d, \max_d) \\ v_i(0) = \frac{U(\min_d, \max_d) - x_i(0)}{2} \\ p_i(0) = x_i(0) \\ l_i(0) = \arg \min_{j \in N_i(x(0))} (f(p_j(0))) \end{cases} \quad (11)$$

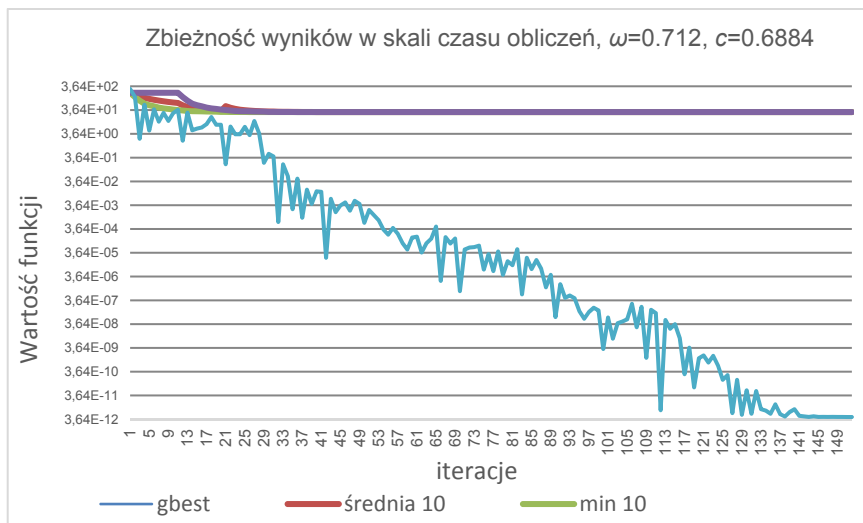
Przystąpiono do zbadania zależności jakości otrzymanych wyników od wielkości stada. W tym celu zbudowano następujące dwa modele:

1. Stado liczące 30 indywidualnych osobników: wielkość sąsiedztwa równa 0 (*lokalbest*, $c_1 = c_2 = c$).
2. Stado liczące 30 osobników z czynnikiem socjalnym równym 0 (*globalbest*, $c_1 = c$, $c_2 = 0$).

W badaniach przyjęto umowną wartość inercji ω , szacując ją w następujący sposób. Założono, że czynnikiem decyzyjnym jest różnica pomiędzy wartościami funkcji ewaluacji w kolejnych krokach. W takim przypadku zmniejszanie się

jej świadczyłoby o zbliżaniu się do ekstremum. Pozostawienie wartości parametrów bez zmian miałyby zapewnić kontynuowanie trendu. W przypadku wzrostu tej różnicy, a więc oddaleniu się od ekstremum, ingerencja polega na zawróceniu cząsteczki, zmianie kierunku i redukcji prędkości o pewien współczynnik. Wielkość tego współczynnika dobierana jest deterministycznie bądź heurystycznie. W przypadku połowy ostatniej prędkości następny krok przenosi cząstkę w połowę drogi. Optymalna wielkość współczynnika redukcji zależy od rozmiarów przestrzeni $x_{max} - x_{min}$, oczekiwanej liczby kroków do zatrzymania algorytmu t_{max} oraz minimalnej prędkości, dla której uznaje się jakikolwiek postęp: v_{min} .

Obliczenia wykazały, że przy zakładanym koszcie na kolejnych progach wyliczona wartość ω waha się pomiędzy 0,92 a 0,98, natomiast c pomiędzy 1,43 a 1,48. Dotychczasowe badania miały na celu zbadanie, jakie mogą być najlepsze parametry i rezultaty dla pojedynczych cząstek (działających w odosobnieniu). Obserwacjom poddano zarówno wynik offline, jak i online. Okazało się bowiem, że niejednokrotnie cząsteczki nie utrzymują najlepszych pozycji – bezwładność ich jest zbyt duża. Zadawalające wyniki osiąga 1 lub 2 na 30 badanych cząstek, a pozostałe średnio ustalają duże gorsze wyniki. Ponadto wraz z przebiegiem obliczeń wyniki te się nie polepszają. Najlepsze otrzymane rezultaty przedstawiają poniższe wykresy zbieżności.



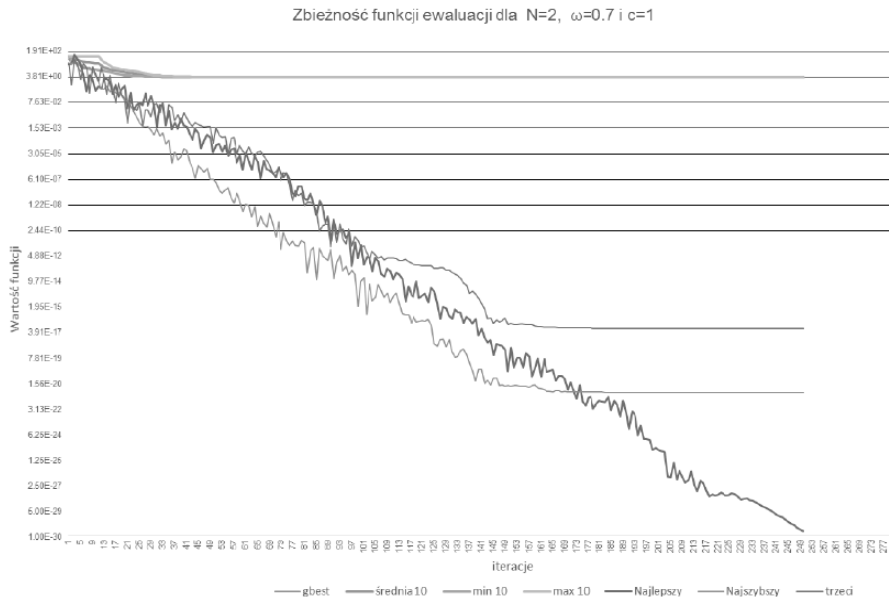
Rys. 5. Zbieżność funkcji ewaluacji dla $N = 1$, $\omega = 0,712$ i $c = 0,6884$

Źródło: Opracowanie własne.



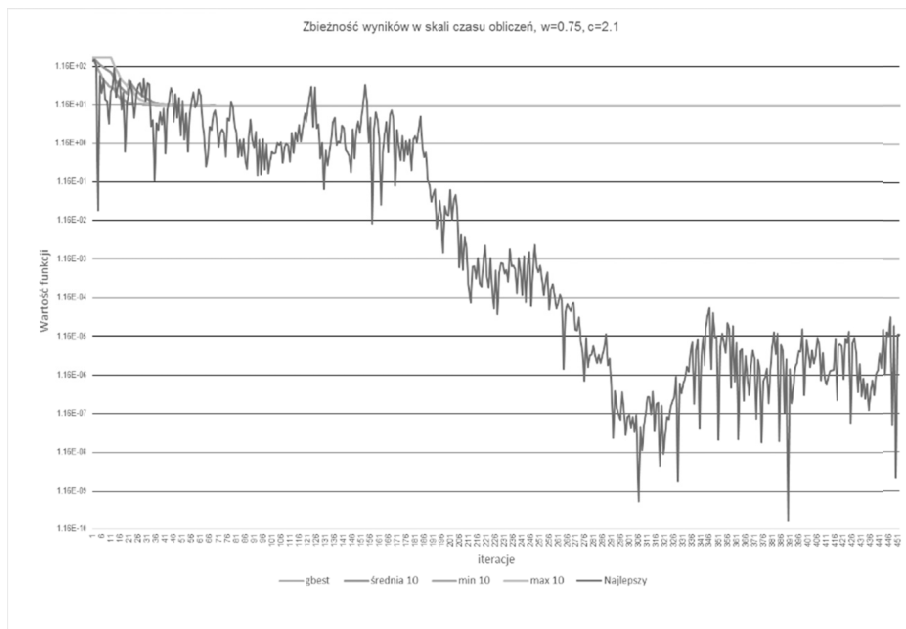
Rys. 6. Zbieżność funkcji ewaluacji dla $N = 1$, $\omega = 0,721$ i $c = 1,3768$

Źródło: Opracowanie własne.



Rys. 7. Zbieżność funkcji ewaluacji dla $N = 2$, $\omega = 0,7$, $c = 1$

Źródło: Opracowanie własne.



Rys. 8. Zbieżność funkcji ewaluacji dla $N = 2$, $\omega = 0,75$, $c = 2,1$

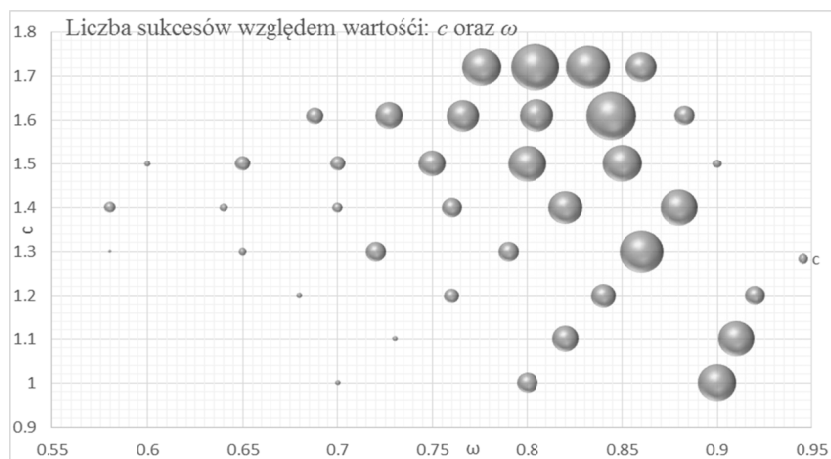
Źródło: Opracowanie własne.

Wartości ω z zakresu od 0,712 do 0,721 oraz c z zakresu od 0,6 do 1,4 wydają się równie dobre. W kolejnych eksperymentach wprowadzono drugiego osobnika (stado o liczebności $N = 2$), przyjęto także zaokrągloną wartość $\omega = 0,7$ (rys. 7) oraz $\omega = 0,75$ (rys. 8) dla średniej wartości $c = 2,1$.

Eksperyment ten pokazuje, jaki wpływ na szybkość uzyskania oczekiwanych rezultatów ma wiedza o przybliżonym wyniku obliczeń. W eksperymencie tym nie zastosowano jako pomocniczej informacji o położeniu najlepszej cząsteczki, jednak para potrafiła wykorzystać chwilową najlepszą pozycję każdej z dwójga cząsteczek. Wartość jej szybko przekroczyła próg 10^{-4} . Skutkiem tego była poprawa średniej poniżej 10^{-5} oraz wartości najlepszych nawet poniżej 10^{-30} , i to już po 250 iteracjach.

Dwa kolejno najlepsze rezultaty ustaliły się zaledwie po 160 iteracjach na poziomie 10^{-20} oraz 10^{-17} . Jest to bardzo dobry wynik, gdyż daje rozwiązanie z prawdopodobieństwem ok 10%. Jako model do dalszych testów wybrano wariant PSO globalbest z liczebnością stada $N = 30$. Dla tego modelu badano wpływ parametrów ω oraz c na szybkość zbieżności. Zbieżność tę stwierdzono w sytuacji przekroczenia progu dokładności 10^{-3} po 10 000 kroków. Brak osiągnięcia tego warunku uznano za porażkę, osiągnięcie – za sukces. Obliczenia

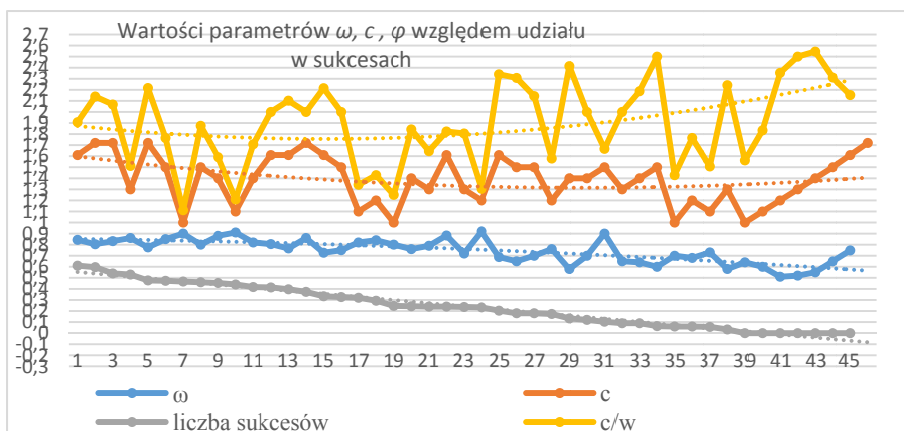
powtarzano 10 razy dla każdej z badanych funkcji testowych. Procentowy udział sukcesów do wszystkich prób obrazuje poniższy wykres (rys. 9). Na rysunku wielkość kuli odpowiada pozycji w rankingu.



Rys. 9. Względna liczba sukcesów w zależności od wartości parametrów ω oraz c dla 30-elementowej populacji w modelu globalbest dla wybranych funkcji sześcioparametrowych

Źródło: Opracowanie własne.

Na rysunku 10 pokazano zależność wartości poszczególnych parametrów (oś odciętych) od pozycji w rankingu (oś rzędnych), kropkami zaznaczono trend.



Rys. 10. Wartości parametrów względem udziału w sukcesach

Źródło: Opracowanie własne.

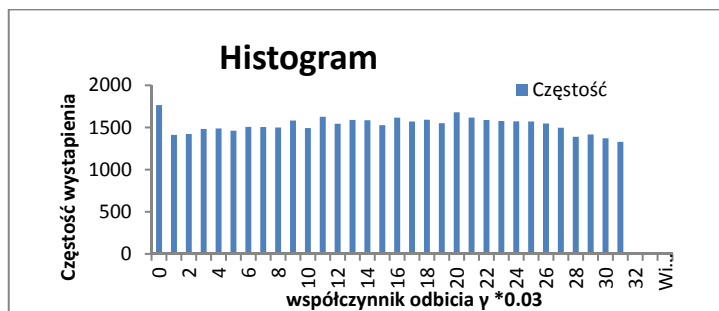
Dla porównania zestawiono poniżej zakres i względny błąd parametrów uzyskanych dla wyliczonego błędu, dla poszczególnych progów dokładności. Wyselekcjonowano zestawy, dla których liczba poprawnie otrzymanych wyników jest większa niż mediana, tzn. dla których najszybciej osiągnięto sukces. Wartości te podano w tabeli 3. Puste pola lub wartość 0 oznaczają, że nie osiągnięto wskazanej dokładności.

Tabela 3. Rozbieżności w otrzymanych parametrach dla wybranych funkcji sześcioparametrowych

Najszybsze: osiągnięty sukces przy 1., 2. i 3. ograniczeniu kosztu									
	Dla progu 1			Dla progu 2			Dla progu 5		
Dla progu 1	bł 1 >50	bł 2	bł 3	bł 1 >100	bł 2 >0	bł 3	bł 1 >1300	bł 2 >800	bł 3 >100
min ω	0,645	0	0	0,645	0,645				0,645
max ω	0,69075	0	0	0,72125	0,6755				0,75175
zmiana [%]	7,09%			11,82%	4,73%				16,55%
min c_1	1,35	0	0	0,72	1,53				0,81
max c_1	2,88	0	0	2,88	2,07				2,52
zmiana [%]	113,33%			300,00%	35,29%				211,11%
min c_2	0,51	0	0	0,51	1,32				0,87
max c_2	2,04	0	0	2,67	1,86				2,58
zmiana [%]	300,00%			423,53%	40,91%				196,55%
min φ_1	2,093,023	0	0	11,162,791	2,317,304				112,305
max φ_1	4,361,984	0	0	43,619,841	3,064,397				3,493,934
zmiana [%]	108,41%			290,76%	32,24%				211,11%
min φ_2	0,738328	0	0	0,7223796	1,954,108				1,206,239
max φ_2	3,162,791	0	0	41,395,349	2,817,115				3,860,465
zmiana [%]	328,37%			473,04%	44,16%				220,04%
min c_1/c_2	0,661765	0	0	0,2696629	0,822581				0,313953
max c_1/c_2	5,647,059	0	0	56,470,588	1,568,182				2,896,552
zmiana [%]	753,33%			1994,12%	90,64%				822,61%
min f	4,15E-07	0	0	2,19E-09	2,19E-09				9,40E-25
max f	9,92E-05	0	0	8,70E-05	9,10E-09				9,22E-17
min hit	1	0	0	1	0				0
max hit	98	0	0	262	32				1260

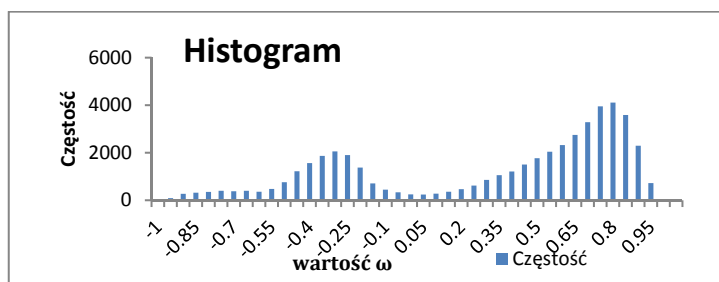
Źródło: Opracowanie własne.

Bardziej obrazowo wpływ poszczególnych badanych wartości parametrów na liczbę zadawalających wyników metody pokazują poniższe histogramy.



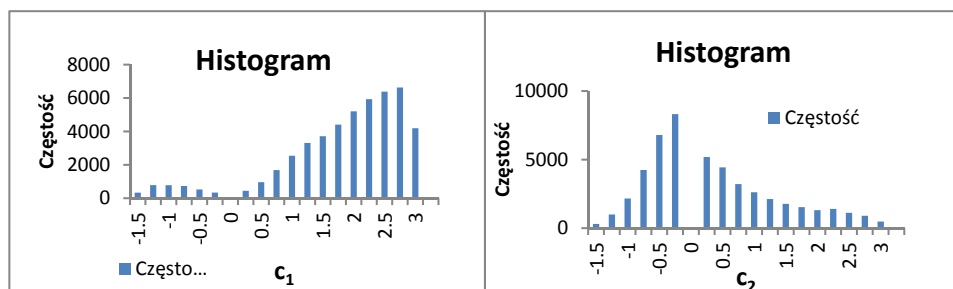
Rys. 11. Wartości współczynnika odbicia w zakończonych sukcesem obliczeniach badanych funkcji dla $N = 4$ i $d = 4$

Źródło: Opracowanie własne.



Rys. 12. Wartości współczynnika inercji ω a liczba sukcesów obliczeń badanych funkcji

Źródło: Opracowanie własne.

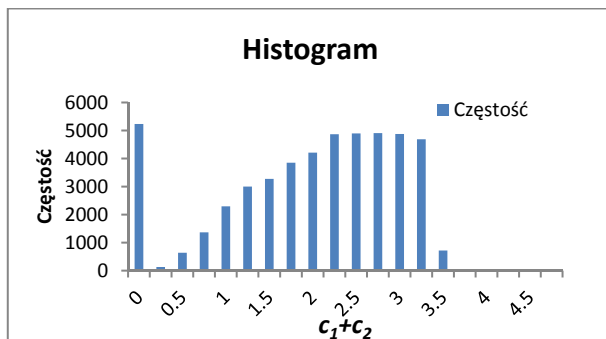


Rys. 13. Wartości współczynników c_1 i c_2 oraz ich udział w obliczeniach zakończonych sukcesem dla badanych funkcji ($N = 4$, $d = 4$)

Źródło: Opracowanie własne.

Z pierwszego histogramu (rys. 11) wynika, że współczynnik odbicia, a więc początkowa prędkość, kierunek i zwrot, nie ma zbytniego znaczenia. Znaczenie mają natomiast współczynniki c_1 i c_2 . Histogram na rysunku 14 przedstawia taką zależność wyprowadzoną empirycznie. Jak widać, suma mieszcząca się w prze-

dziale od 2 do 3,4 daje równie dobre wyniki już dla stada liczącego 4 osobniki (i problemu o wymiarze równym 4).



Rys. 14. Wartości sumy współczynników c_1 i c_2 oraz ich udział w obliczeniach zakończonych sukcesem dla badanych funkcji ($N = 4$, $d = 4$)

Źródło: Opracowanie własne.

Podsumowanie

Przeprowadzone badania miały na celu zweryfikowanie poprawności tezy, że przy prostych, tj. z małą ilością zmiennych parametrów, modelach optymalizowanych funkcji łączny koszt dostrajania parametrów oraz wykonywania optymalizacji ciągłej metodą PSO znacznie przekracza koszt wykonania podstawowej metody dla standardowych wartości parametrów przy takiej samej skuteczności. Dla porównania wybrane zostały 4 podstawowe funkcje sześciu zmiennych. Wykonano po 30 testów dla wariantów globalbest, odnotowując postęp w trakcie obliczeń. Wyniki zbierano w różnych kategoriach: osiągnięcie sukcesu dla maksimum ilości iteracji oraz dla 1/5 jej wielkości, przyjmując za sukces różne progi.

Głównym kryterium jakości było ustalenie wysokiej powtarzalności i tendencji wyników w funkcji wartości mierzonych parametrów. Stwierdzenie takowego uzasadniałoby potrzebę ustalenia wielkości tychże parametrów na drodze wstępnego dostrajania. Brak stałej tendencji uznawany jest za argument znoszący potrzebę wstępnego dostrajania, gdyż jakość metod nie zależy w takim wypadku od wartości parametrów. By zilustrować statystycznie, jaki wpływ na osiągnięcie sukcesu, tj. osiągnięcie punkty ekstremalnego lub jemu bliskiego, zostały wykonane histogramy. Analizowano wpływ zarówno liczności stada, jak i zależność między nią a ilością wymiarów na zdolność osiągnięcia i utrzymania minimalnych wartości badanych funkcji. Sytuacje, w których incydentalnie zo-

stała osiągnięta wartość minimalna, charakteryzuje niska liczba udanych ewaluacji, natomiast stała tendencja i wysoka zbieżność do optymalnego wyniku powoduje wysoki wskaźnik sukcesów w tej samej ilości prób, a ponadto małą odległość między najodleglejszymi pozycjami cząstek po zakończeniu obliczeń. Uzyskane wyniki potwierdziły wysunięte na wstępie hipotezy.

Literatura

- Adriansyah A., Amin S.H.M. (2006), *Analytical and Empirical Study of Particle Swarm Optimization with a Sigmoid Decreasing Inertia Weight* [w:] Regional Postgraduate Conference on Engineering and Science, 2006, Johore.
- Bansal J.C., Singh P.K., Saraswat M., A. Verma, S.S. Jadon, A. Abraham (2011), *Inertia Weight Strategies in Particle Swarm*, http://www.softcomputing.net/nabic11_7.pdf (dostęp: 15.05.2016).
- Bergh van den F., Engelbrecht A.P. (2006), *A Study of Particle Swarm Optimization Particle Trajectories*, www.elsevier.com/locate/ins (dostęp: 28.05.2016).
- Bonyadi M.R., Michalewicz Z. (2017), *Particle Swarm Optimization for Single Objective Continuous Space Problems: A Review*, "Evolutionary Computation", Vol. 25(1), s. 1-54.
- Carlisle A.J. (2002), *Applying the Particle Swarm Optimizer to Non-stationary Environments*, 16.12.2002, https://static.aminer.org/pdf/PDF/000/261/986/particle_swarm_optimization_in_non_stationary_environments.pdf (dostęp: 28.05.2016).
- Clerc M. (2012), *Standard Particle Swarm Optimisation*, <https://hal.archives-ouvertes.fr/hal-00764996/document>, (dostęp: 28.05.2016).
- Kennedy J., Eberhart R. (1995), *Particle Swarm Optimization*, Proceedings of IEEE International Conference on Neural Networks, Vol. IV, s. 1942-1948.
- Li-ping Z., Huan-jun Y., Shang-xu H. (2005), *Optimal Choice of Parameters for Particle Swarm Optimization*, 24.10.2005, <http://www.zju.edu.cn/jzus> (dostęp: 28.05.2016).
- Opara K., Arabas J. (2012), *Procedury oceny algorytmów optymalizacji ciągłej*, 03.10.2012, https://www.researchgate.net/publication/272506132_Procedury_oceny_algorytmow_optymalizacji_cigalej (dostęp: 28.05.2016).
- Ozcan E., Mohan C.K. (2010), *Analysis of a Simple Particle Swarm Optimization System*, Syracuse, New York.
- Shi Y., Eberhart R.C. (2012), *Empirical Study of Particle Swarm Optimization* [w:] Proceedings of the IEEE Congress on Evolutionary Computation, IEEE Press, s. 1945-1950.
- Ströyner C. (2014), *Particle Swarm Optimization & Parameter Selection*, 17.10.2014, http://www.mayr.informatik.tu-muenchen.de/konferenzen/Ferienakademie14/slides_papers/paper_Christoph_Stroessner.pdf (dostęp: 28.05.2016).

Yuan Z., Montes de Oca A., Birattari M., Stützle T. (2010), *Modern Continuous Optimization Algorithms for Tuning Real and Integer Algorithm Parameters*, 02.10.2010, <http://iridia.ulb.ac.be/~mbiro/paperi/YuaMonBirStu2010ants.pdf> (do-step: 28.05.2016).

THE PROBLEM OF TUNING THE PARTICLE SWARM OPTIMIZATION ALGORITHM IN CONTINUOUS OPTIMIZATION

Summary: The optimization of the chosen issue of finding an analytical solution to determine the extreme describing this function is very often complex. An analytical solution is sometimes impossible, especially when the function is formulated in an entangled way. In many cases, there are no methods for automatically solving such formulas. Optimization of the particle swarm (Particle Swarm Optimization, PSO) has proved to be an effective tool for solving many optimization problems. Hence numerous attempts to modify and determine the most optimal and universal parameters of the PSO algorithm. The algorithm itself is also used as a part of other non-deterministic algorithms to create hybrid constructions.

The article presents the results of research on the effectiveness of the basic version of the particle swarm optimization algorithm (PSO) for continuous problems of varying complexity scale. The authors also attempted to evaluate the cost of tuning this algorithm for small problems. The carried out computational experiments confirm the hypotheses advanced.

Keywords: PSO, optimization, tuning.