Marija Janković i Miljan Dašić

Simulacija kretanja indeksa berze optimizovanim fraktalnim funkcijama

Cilj rada je ispitivanje fraktalne prirode funkcije kretanja indeksa berze tokom vremena. Napisana je simulacija u programskom jeziku C++ koja generiše fraktalne funkcije koje kopiraju kretanje indeksa berze. Ra- čunarski generisane fraktalne funkcije su optimizovane genetičkim algoritmom, tako da vrednosti konačnih funkcija minimalno odstupaju od vrednosti indeksa berze. Zaključeno je da se optimizacijom fraktalnih funkcija može kopirati kretanje indeksa berze u određenim vremenskim intervalima.

Teorijski uvod

Berza u ekonomskom i pravnom smislu

Berza je javno, organizovano tržište kapitala. Definicija berze sadrži tri osnovna elementa: mesto na kome se trguje, predmet trgovanja i način trgovanja. Najčešće berza ima svoj objekat gde agenti trguju na licu mesta, mada se danas sve više berzanskih poslova obavlja elektronskim informacionim sistemima. Način trgovanja je zakonom strogo regulisan.

Ključna funkcija berze je da obezbedi kontinualno tržište hartija od vrednosti po cenama koje ne odstupaju bitno od onih po kojima su prethodno prodate. Kontinuitet tržišta smanjuje kolebanje cena hartija od vrednosti što dalje povećava njihovu sposobnost konverzije u gotovinu putem prodaje. Berza utvrđuje i publikuje cene hartija od vrednosti.

Postoji potreba da se kvantifikuju uspešnost rada berze i hartije od vrednosti kojima se na njoj trguje. U tu svrhu svaka berza ima parametar – indeks berze. Definicija i načina računanja indeksa berze ima mnogo. Uopšteno, indeks berze predstavlja analitičko oruđe za sve subjekte koje interesuje dinamika promena cena na datom tržištu kapitala koje berza pokriva. Konkretno, indeks berze je pozitivan racionalan broj koji se menja tokom vremena. U našem radu korišćeni su podaci o kretanju dnevnog indeksa nekih od najpoznatijih svetskih berzi. Podaci stari i do 50 godina dobavljeni su sa Interneta (http://finance.yahoo.com).

Marija Janković (1991), Beograd, Ratnih vojnih invalida 23/13, učenica 3. razreda Računarske gimnazije u Beogradu

Miljan Dašić (1990), Paraćin, Majora Marka 13, učenik 4. razreda Gimnazije u Paraćinu

MENTOR: dr Aleksandar Bogojević, Institut za fiziku, Beograd

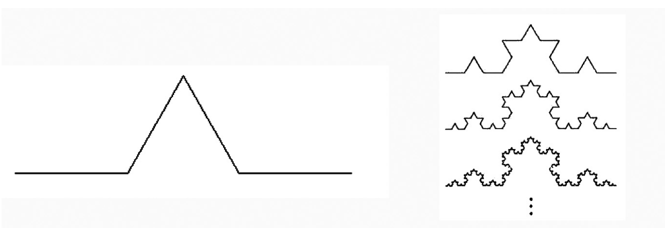
Berza u matematičkom smislu

Pokazuje se u praksi da berza nije stabilan i kontinuiran sistem kakav bi u idealnom slučaju trebalo da bude i kako je ekonomski i pravno definisana. Posmatrana u matematičkom smislu, berza je jako kompleksan sistem na koji utiče ogroman broj činilaca. Dešavanja na berzi uslovljena su velikim brojem faktora i jako su osetljiva na razne događaje koji eksplicitno nisu povezani sa berzom. Iz tog razloga čini se da je kretanje indeksa berze slučajno, haotično i nepredvidivo. Funkcija kretanja indeksa berze od vremena je diskretna i ne može se zapisati u analitičkom obliku.

Fraktali

Fraktali su geometrijski objekti čije je osnovno svojstvo samosličnost. Dobijaju se rekurzivno ponovljenim geometrijskim ili računskim postupkom. Mnogi objekti u prirodi imaju fraktalnu prirodu, poput oblaka, brokolija i nekih drugih biljaka, morskih obala, snežnih pahulja i planinskih venaca. U računarskoj grafici se upravo generisanjem fraktala dobijaju svi ovi prirodni objekti. Definiše se polazni oblik i postupak, odnosno geometrijska transformacija koja se rekurzivno ponavlja, teoretski do u beskonačnost, a praktično do željene rezolucije.

Pokazano je da funkcija kretanja indeksa berze izgleda samoslično na svim nivoima vremenskih skala – dan, nedelju dana, meces dana, godina (Mandelbrot 1999). Dakle, moguće je prikazati ovo kretanje fraktalnom funkcijom. Fraktalna funkcija je fraktalna kriva koja zadovoljava definiciju funkcije, tj. za svaku vrednost nezavisne promenljive iz definisanog domena daje tačno jednu vrednost. Fraktalne funkcije su česte u prirodi i važne u mnogim fizičkim procesima, tako da se uopšteno ovaj metod može primeniti i za ispitivanje drugih funkcija fraktalne prirode.



Slika 1. Generisanje poznatog fraktala – Kohove krive

Figure 1. Generating a well known fractal – Koch curve

Genetički algoritam

Heuristički algoritmi su grupa algoritama koji imaju izuzetnu primenu u skupu problema nerešivih egzaktno uopšte ili nerešivih egzaktno u realnom vremenu. Njihova prednost je relativno kratko vreme izvršavanja algoritma, a rezultat samo približno tačno, ali dovoljno dobro rešenje.

Za optimizaciju fraktalnih funkcija kojima želimo da simuliramo kretanje indeksa berze izabran je genetički algoritam iz grupe heurističkih algoritama. Genetički algoritam (GA) radi po uzoru na molekul DNK u prirodi. Zadatak mu je da poveća funkciju prilagođenosti jedinke okolini, gde je u našem slučaju jedinka generator fraktalne funkcije, a prilagođenost je korelacija sa kretanjem indeksa berze.

GA u koracima:

1. Generisanje početne populacije jedinki.
2. Evaluacija jedinki – računanje funkcije prilagođenosti.
3. Selekcija jedinki – jedinke se sortiraju prema funkciji prilagođenosti i određen procenat ukupnog broja jedinki (one koje su bolje prilagođene) koristi se za stvaranje nove populacije.
4. Stvaranje nove populacije – reprodukcija selektovanih (bolje prilagođenih) jedinki korišćenjem rekombinacija (crossing-over) i mutacija.
5. Ako je dostignut zadat broj generacija, algoritam se završava. Ako ne, vraća se na korak 2 sa novostvorenom populacijom.

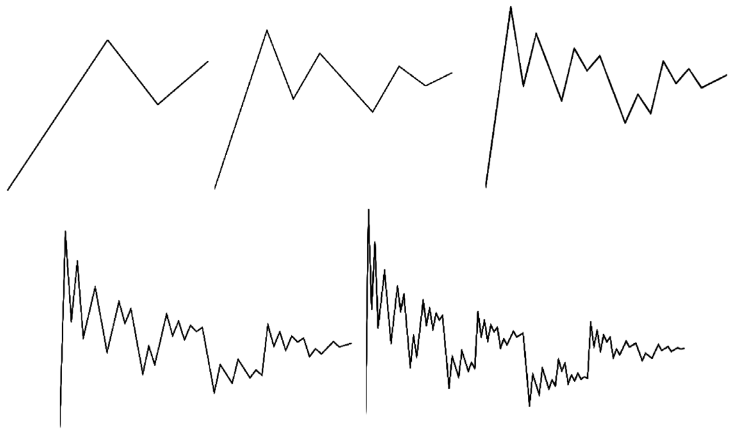
Simulacija

Za osnovnu ideju o realizaciji simulacije poslužila je ideja o fraktalnim generatorima izložena u radu Benoit B. Mandelbrota (Mandelbrot 1999), jednog od začetnika teorije fraktala. Kao ulaz simulacije služi funkcija zavisnosti indeksa berze od vremena za neku odabranu berzu, odnosno vrednosti indeksa berze za neki vremenski period – određeni broj dana, sati, nedelja ili meseci. Izlaz simulacije predstavlja fraktalna funkcija koja datu funkciju kretanja indeksa kopira u zadovoljavajućoj meri.

Generisanje početne populacije

Na početku simulacije generiše se na slučajan način početna populacija od 20 jedinki i broj jedinki u populaciji je stalan tokom cele simulacije. Jedna jedinka jeste generator fraktalne funkcije. Geni te jedinke su koordinate tačaka generatora (od 4 do 10 tačaka). Geni su, dakle, realni brojevi u intervalu $[0, 1]$. X-koordinata prve tačke u generatoru uvek je 0, a poslednje 1. Fraktalna funkcija konstruiše se reskaliranjem generatora na prostor između dve zadate tačke (prva i poslednja uzeta vrednost indeksa berze). Postupak se rekurzivno ponavlja reskaliranjem generatora na susedne tačke fraktala dobijenog u prethodnoj iteraciji i to sve do trenutka kada je dobijen određen broj tačaka. Na slici 2 prikazan je u koracima opisani postupak.

U toku izvršavanja genetičkog algoritma fraktalni generator ponaša se analogno naslednoj materiji kod živih bića. Kroz iteracije GA, koje



Slika 2.
Postupak konstrukcije fraktala u našoj simulaciji – od generatora sa 4 tačke do fraktala od 100 tačaka

Figure 2.
Construction of a fractal in our simulation from a generator with 4 points to a fractal with 100 points

nazivamo i generacijama, generatori, odnosno jedinke, menjaju se i „uče" osobinama neophodnim za preživljavanje. U ovom slučaju, da bi jedinka preživela, fraktalna funkcija koja nastaje od date jedinke mora postići što manje odstupanje od funkcije promene indeksa berze koja čini ulazni podatak simulacije.

Važno je istaći razliku između ovde primenjenog GA i evolucije koja se odvija u prirodi. Nasledne materije živih bića na Zemlji jesu RNK i DNK, molekuli koji se sastoje od nizova nukleotida. U oba molekula razlikujemo 4 vrste nukleotida. Uobičajeno je da se prilikom primene GA ovaj model uprosti na dve vrste nukleotida – 0 ili 1. Geni su, dakle, uobičajeno strogo diskretni. Međutim, pokazuje se da se GA lako može modifikovati tako da radi sa genima predstavljenim realnim brojevima.

Evaluacija i selekcija

Neka je N broj učitanih vrednosti indeksa berze, odnosno broj posmatranih dana (dan je najčešće korišćena rezolucija u ovoj simulaciji). Fraktalna funkcija konstruiše se za svaku jedinku, odnosno generator, na N tačaka, koristeći prvu i poslednju tačku učitanu iz podataka o berzi. Za svaku od tih fraktalnih funkcija izračunava se σ, koeficijent prilagođenosti jedinke, odnosno mera sličnosti konstruisane fraktalne funkcije i funkcije kretanja indeksa berze. Dobijeni broj predstavlja funkciju prilagođenosti za svaku od jedinki. U datoj formuli Δy je razlika prave vrednosti indeksa i vrednosti fraktalne funkcije u i-toj tački, a y je prava vrednost indeksa u toj tački:

$$\sigma = 1 - \frac{\sum_{i=1}^{n} \frac{\Delta y_i}{y_i}}{N}.$$

Uspešnije jedinke populacije tekuće generacije opstaju, dok se manje uspešne odbacuju. Od dvadeset jedinki tekuće populacije, pet najuspešnijih ući će u po dva reproduktivna čina, narednih 10 po uspešnosti ući će u po jedan reproduktivni čin, a poslednjih pet jedinki biće zanemareno u potpunosti. Reproduktivni čin predstavlja dobijanje dve nove jedinke na osnovu dve slučajno sparene jedinke iz tekuće populacije. Čin se sastoji iz genskih rekombinacija i mutacija.

Za svaku generaciju jedinki, za svaku populaciju, računa se srednji koeficijent prilagođenosti koji se koristi kao merilo napretka jedinki kroz generacije. Srednji koefiijent prilagođenosti obeležavaćemo u daljem tekstu sa β i on je jednak aritmetičkoj sredini koeficijenata σ svake jedinke u tekućoj populaciji. I σ i β jesu realni brojevi, ne nužno, ali gotovo uvek u intervalu [0, 1].

Aritmetičke genske rekombinacije

Po uzoru na rekombinacije hromozoma u prirodi, dešavaju se i rekombinacije u računarskom svetu. Za razliku od klasične rekombinacije (jednostavne razmene gena između dva niza gena) koja se primenjuje kada su geni dati iz određenog skupa diskretnih vrednosti (npr. bitovski – nule i jedinice), aritmetička rekombinacija se koristi kada su geni dati kao realni brojevi. Parametri aritmetičkih rekombinacija – α i k, određuju se na slučajan način. Parametar α predstavlja slučajan realan broj između nula i jedan, dok parametar k predstavlja ceo broj koji određuje mesto na kome se seku nizovi tačaka fraktalnih generatora dve jedinke; uzima vrednost od 0 do dužine kraće jedinke (dužina je broj tačaka njenog fraktalnog generatora, 4 do 10).

Koordinate tačaka generatora novonastalih jedinki (dece) računaju se kao linearna kombinacija vrednosti starih jedinki (roditelja) koristeći parametar α. Ako su c_1 i c_2 uređeni parovi koordinata po jedne tačke iz generatora prvog i drugog deteta, a p_1 i p_2 uređeni parovi koordinata po jedne tačke roditelja, odvija se sledeća linearna kombinacija:

$$c_1 = \alpha \cdot p_1 + (1-\alpha) \cdot p_2$$

$$c_2 = \alpha \cdot p_2 + (1-\alpha) \cdot p_1$$

Na ovaj način menjaju se vrednosti gena (koordinata tačaka) do mesta presecanja određenog parametrom k. Od tog mesta, prvo dete ostale gene dobija direktno od prvog roditelja, a drugo od drugog. Važno je uočiti da nije neophodno da jedinke imaju isti broj gena. Primer aritmetičke rekombinacije (k = 2, α = 0.3):

0.0 0.2 | 0.6 0.3 0.2 0.5

$\overline{\mathbf{0.6}}$ $\overline{\mathbf{0.9}}$ | 0.4 0.3 0.5 0.5

=

0.42 0.69 0.6 0.3 0.2 0.5

$\overline{\mathbf{0.18}}$ $\overline{\mathbf{0.41}}$ 0.4 0.3 0.5 0.5

Korišćenjem aritmetičkih rekombinacija vrlo brzo bi jedinke cele populacije imale gotovo identične prve gene, tj. onoliko gena koliko ima generator najmanje dužine. Na prvi pogled, ovo je dobro jer sistem brzo konvergira nekom rešenju. Međutim, kako se brže dolazi do zasićenja, sistem ne dobija šansu da evoulira do nekog potencijalno boljeg rešenja. Zbog toga je korišćen i jednostavniji, obični princip rekombinacija:

0.0 0.2 | 0.6 0.3 0.2 0.5

$\overline{\mathbf{0.6}}$ $\overline{\mathbf{0.9}}$ | 0.4 0.3 0.5 0.5

=

$\overline{\mathbf{0.6}}$ $\overline{\mathbf{0.9}}$ 0.6 0.3 0.2 0.5

0.0 0.2 0.4 0.3 0.5 0.5

Mutacije

Po uzoru na prirodu primenjuju se slučajne promene gena – mutacije. Svrha mutacija uglavnom jeste da unesu novi genetički materijal. Rekombinacije, kao što im i naziv kaže, reorganizuju postojeći materijal. Za evoluciju je potrebna pojava novih osobina. Međutim, u našem GA, novi genetički materijal unosi se i mutacijama, ali u slučaju korišćenja aritmetičkih rekombinacija i tim rekombinacijama. Bez obzira što se u rekombinacijama novi geni dobijaju direktno iz starih, dobijaju se novi brojevi, te samim tim novi genetički materijal. Kao i u prirodi, nivo mutacija je razumno nizak, pošto bi visok nivo mutacija stalno slučajno menjao jedinke, čime se gubi kontrola i usporava evolucija.

Koristimo dve vrste mutacija. Najpre sa određenom verovatnoćom (korišćena vrednost 1%, odnosno 5% za obične rekombinacije) za određen procenat (1%, odnosno 5%) menjamo vrednost neke od koordinata tačaka generatora novonastalih jedinki, a zatim, takođe sa definisanom verovatnoćom (10%), u generator neke jedinke umećemo novu slučajno generisanu tačku ili izbacujemo neku od tačaka koje se već nalaze u tom generatoru. Za navedene vrednosti verovatnoća mutacija utvrđeno je da daju dobre rezultate.

Završetak rada algoritma

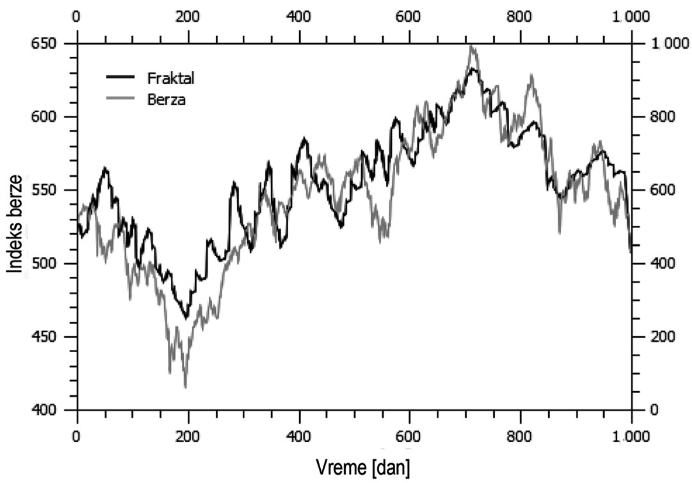
Algoritam radi zadati broj generacija, stvarajući novu populaciju. Potreban je veliki broj generacija kako bi se dobile dobro prilagođene jedinke.

Rezultati i diskusija

Simulacija je puštana za 15 različitih berzi. Korišćeni su najuticajniji svetski, kontinentalni i regionalni indeksi (www.finance.yahoo.com). Podaci su iz različitih vremenskih perioda, a sami indeksi računati su na nekoliko različitih načina i imaju različite brojeve komponenti, te su korišćeni podaci raznovrsni u svakom pogledu. GA je puštan za različit broj generacija. Fraktalna funkcija koja u velikoj meri kopira funkciju kretanja indeksa berze dobijena je za svih 15 berzi.

Na slici 3 prikazane su funkcija kretanja indeksa Njujorške berze (NYA index, novi metod računanja indeksa) i fraktalna funkcija dobijena upotrebom GA. Nulta tačka na grafiku odgovara 31. decembru 1965. godine, a svaka naredna tačka odgovara narednom radnom danu. Dakle, rastojanje po x-osi između dve tačke jeste jedan dan (odnosno tri dana između tačaka koje odgovaraju petku i ponedeljku). Dobijena je fraktalna funkcija koja oslikava funkciju kretanja indeksa. Prikazana fraktalna funkcija je jedna od jedinki poslednje generacije, a u poslednjoj generaciji sve jedinke su izuzetno slične, većina je čak i identična. Ovo je tipičan primer rada simulacije i rezultati slične uspešnosti postignuti su i za sve ostale korišćene berze, za oba tipa genskih rekombinacija.

Konkretno, primer sa grafika na slici 3 dobijen je korišćenjem aritmetičkih genskih rekombinacija kroz 5000 generacija na 1000 tačaka. Srednji koeficijent prilagođenosti za populaciju poslednje generacije iznosio je približno 0.97.

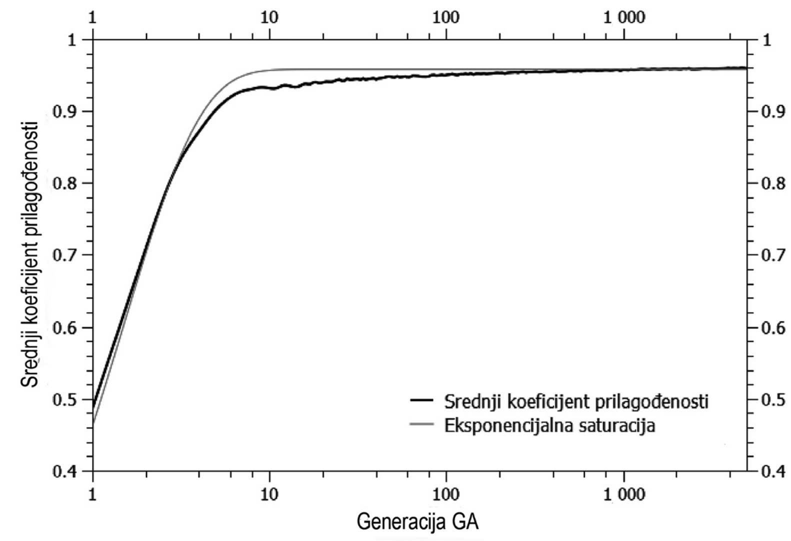


Slika 3.
Kretanje indeksa berze

Figure 3.
Stock market index changes by time (day): black line – fractal; gray line – stock market

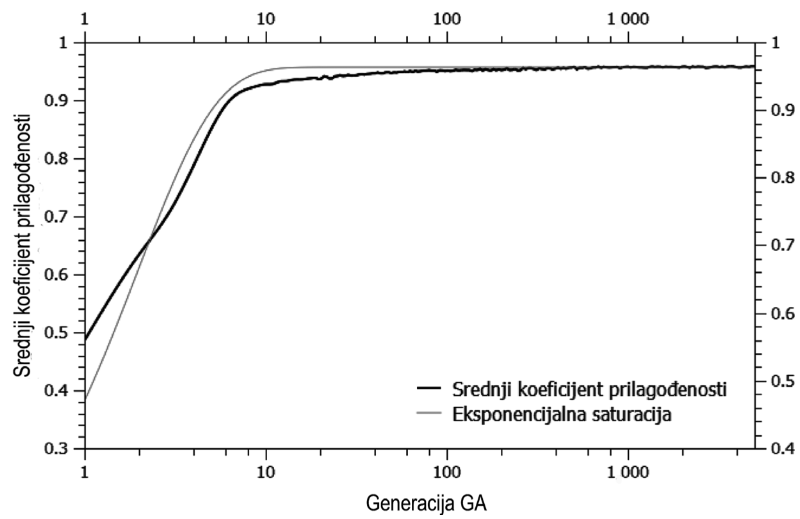
Uočeno je da pri broju generacija višestruko manjem od 5000 jedinke ne uspevaju da evoluiraju toliko dobro, a pri višestruko većem broju generacija dobijaju se rezultati gotovo iste uspešnosti.

Simulacija je za podatke o NYA indeksu puštana 100 puta sa gore navedenim vrednostima parametara i upotrebom aritmetičkih i običnih rekombinacija. Na slikama 4 i 5 prikazano je usrednjeno kretanje srednjeg koeficijenta prilagođenosti kroz generacije za tih 100 izvođenja, koje nam kazuje o napretku jedinki tokom evolucije. Oba grafika fitovana su



Slika 4.
Optimizacija genetičkim algoritmom – aritmetičke genske rekombinacije

Figure 4.
Optimization with a genetic algorithm – arithmetic genetic recombinations, mean suitability coefficient by generation:
black line – mean suitability coefficient
gray line – exponential saturation



Slika 5.
Optimizacija genetičkim algoritmom – obične rekombinacije

Figure 5.
Optimization with a genetic algorithm – usual recombinations, mean suitability coefficient by generation:
black line – mean suitability coefficient
gray line – exponential saturation

funkcijama eksponencijalne saturacije (zasićenja). Eksponencijalne krive prikazane na tim graficima odgovaraju jednačinama:

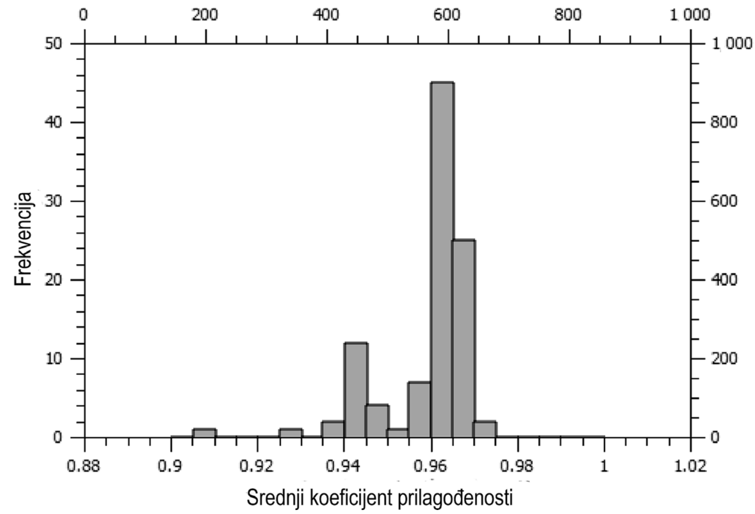
$$f(x) = 0.96(1 - e^{-0.66x})$$

za simulaciju sa aritmetičkim rekombinacijama, i:

$$f(x) = 0.96(1 - e^{-0.5x})$$

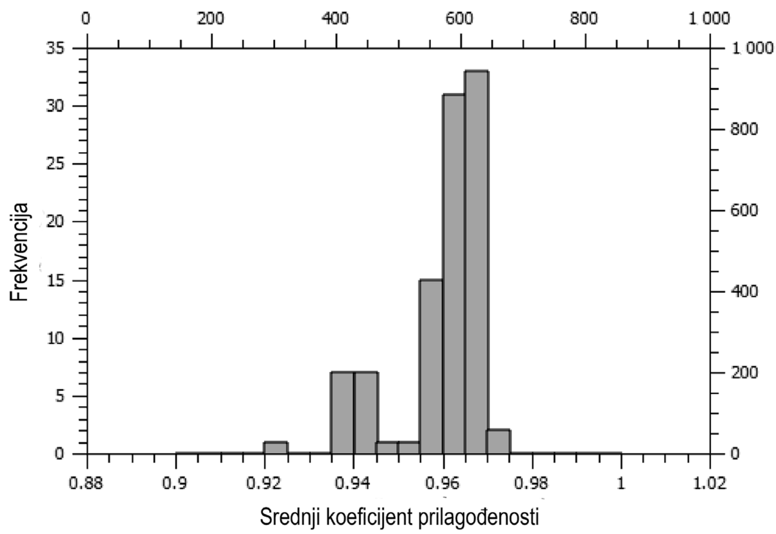
za krivu na grafiku koji odgovara simulaciji sa običnim genskim rekombinacijama.

Na slikama 6 i 7 prikazana je histogramom raspodela vrednosti dobijenih kao srednji koeficijent prilagođenosti za poslednju populaciju



Slika 6.
Raspodela srednjeg koeficijenta prilagođenosti – aritmetičke genske rekombinacije

Figure 6.
Distribution of the mean suitability coefficient – arithmetic genetic recombinations, frequency by mean suitability coefficient



Slika 7.
Raspodela srednjeg koeficijenta prilagođenosti – obične genske rekombinacije

Figure 7.
Distribution of the mean suitability coefficient – usual recombinations, frequency by mean suitability coefficient

tokom tih 100 izvođenja, koji nam govori o verovatnoći dobijanja najboljih mogućih vrednosti konačno dobijenog koeficijenta prilagođenosti u GA. Za podatke prikazane na oba ova histograma izračunato je da srednja vrednost iznosi 0.96, a standardna devijacija 0.01. Histogrami dobijeni za ostale berze veoma su slični prikazanim, uključujući pad frekvencije srednjeg koeficijenta prilagođenosti koji se u ovom slučaju javlja na vrednosti od 0.95. U histogramima za različite berze razlike čine pre svega najverovatnija vrednost koeficijenta prilagođenosti i interval na kome se računa frekvencija tih vrednosti, a pri čijoj se različitoj veličini bolje ili lošije uočava pomenuti pad frekvencije.

Zaključak

Napisana je simulacija u programskom jeziku C++ kojom je minimizirano odstupanje fraktalnih funkcija od funkcija kretanja indeksa berze optimizacijom genetičkim algoritmom. Pošto koeficijent prilagođenosti brzo konvergira ka dovoljno dobrom rešenju, zaključeno je da je naš GA efikasan i upotrebljiv za kopiranje berzanskih funkcija. Dobijene su fraktalne funkcije koje kvalitetno oslikavaju petnaest najznačajnijih svetskih berzi različitih karakteristika.

Došlo se do zaključka da usrednjeni srednji koeficijent prilagođenosti kroz vreme u zavisnosti od proteklih generacija tokom izvršavanja GA raste po zakonu eksponencijalne saturacije, nezavisno od tipa korišćenih genskih rekombinacija. U proseku, konvergira ka vrednosti od 0.96. Ispitivana je raspodela srednjeg koeficijenta prilagođenosti populacije poslednje generacije tokom većeg broja puštanja simulacija.

Privlačna tema za dalje istraživanje svakako je i ekstrapolacija fraktalne funkcije koja kvalitetno predstavlja berzu, odnosno predviđanje daljeg ponašanja berze. Pouzdana metoda predviđanja ponašanja kretanja indeksa i cena na berzi neminovno bi svom korisniku finansijski obezbedila budućnost.

Zahvalnost. Zahvaljujemo se našem mentoru, dr Aleksandru Bogojeviću sa Instituta za fiziku u Zemunu, na idejama, sugestijama i uvodu u ovu zanimljivu i modernu oblast računarske fizike. Takođe, velika zahvalnost ide našem petničkom mentoru, Andriji Jovanoviću, koji nam je uvek brzo i efikasno priskakao u pomoć oko programerskih problema i čiji saveti i ideje su znatno doprineli bogatstvu istraživanja.

Literatura

Mandelbrot B. B. 1999. How fractals can explain what's wrong with Wall Street. *Scientific American*, **280** (2): 70. Dostupno na: http://www.scientificamerican.com/article.cfm?id=multifractals-explain-wall-street

Chaos and Fractals. Dostupno na: http://www.pha.jhu.edu/čldb/seminar/ifs.html

Fractal. Dostupno na: http://www.mathworld.wolfram.com/Fractal.html

Metodologija za izračunavanje indeksa BELEXline. Dostupno na: http://www.belex.rs/data/2007/03/00004500.pdf

Random fractals and the stock market. Dostupno na: http://classes.yale.edu/fractals/RandFrac/welcome.html

Yahoo! Finance. Dostupno na: http://finance.yahoo.com/

Genetički algoritam. Dostupno na: http://www.obitko.com/tutorials/genetic-algorithms/ga-basic-description.php

Marija Janković and Miljan Dašić

Simulation of Stock Exchange Index Using Optimized Fractal Functions

The purpose of this project is to examine the fractal nature of the stock exchange index as a function of time. We have written a computer simulation in the C++ programming language that generates fractal functions which approximate the stock exchange index as a function of time. Fractal functions generated by a computer were optimized by a genetic algorithm, so that the values of final functions tend to be minimally different from the values of the stock exchange index. It was concluded that by using optimized fractal functions, it is possible to approximate changes of stock exchange index in given time intervals. Fractal functions are frequent in nature and important for many physical processes, so this method can be used more generally for simulating other fractal functions.