A genetikai algoritmusok alapjai

Absztrakt

A tanulmány a genetikai algoritmusok témájában végzett kutatások eredményét dolgozza fel, fókuszálva annak alapvető működésére és folyamataira. Jellemzésre kerülnek a genetikus algoritmusok alapfogalmai és bemutatásra kerülnek a folyamatban szereplő legfontosabb módosító lépéseket, az úgynevezett genetikai algoritmus operátoruait, amelyek főként a keresztezési és a mutációs műveletek. Szó esik a népesedési modellekről, a szülőválasztásról, a szelekciós folyamatokról és metódusokról, továbbá bemutatásra kerülnek az egyes leállítási feltételek. Megfogalmazásra kerülnek az egyes előnyök és hátrányok.

Kulcszavak: optimalizálás, genetikai algoritmusok, evolúció algoritmusok

1. Bevezetés

A gyártás és az anyagtudomány folyamatos fejlődése lehetővé tette a tervezőmérnökök számára az olyan komplex digitális modellek előállítását, amelyeket nehéz vagy lehetetlen a hagyományos gyártási módszerekkel és technológiákkal előállítani. Ez a példátlan gyártási rugalmasság a továbbiakban még több lehetőséget és kihívásokat kínál a digitális modellek számítógéppel segített tervezésében. Ezen lehetőségek mellett még a legtapasztaltabb tervezőmérnök intuíciója is alább maradhat. A tervezési kihívások érdekében számítógépes algoritmusok kerültek kifejlesztésre a probléma megoldására az adott tervezési célok és korlátok alapján. Ezt a fajta algoritmus-vezérelt tervezési folyamatot hívjuk ma generatív tervezésnek (GT), amelynek angol elnevezése a Generative Design (GD).

2. A Genetikai algoritmusok

Lényegét tekintve a Genetikai Algoritmus (GA) egy keresésen alapuló optimalizálási folyamat, amely a természetes szelekció alapelveit veszi alapul és használja fel [1]. A genetikai algoritmusok John Holland nevéhez fűzhető, de David Edward Goldberg is nagy sikereket ért el a kutatásaival ebben a témában [2], [3], [4]. Az eljárás széles körben alkalmazott optimális vagy közel optimális megoldások előállítására olyan problémák esetén, amelyeknek a megoldása túlságosan sok időt venne igénybe. Alapvetően az optimalizálás azt a munkafolyamatot jelenti, amely segítségével valami jobb eredmény érhető el, a megoldások keresésének fázisában. Minden folyamatban van egy bemeneti és egy kimeneti paraméter. Ezek alapján az optimalizálás arra koncentrálódik, hogy feltárásra kerüljön az a bemeneti érték, amely alapján a „legjobb" kimeneti érték adódik. A matematika nyelvén fogalmazva a „legjobb" definíciója egy vagy több célfüggvény minimalizálását vagy maximalizálását jelenti, amely a bemeneti paraméterek megfelelő változtatásának a függvénye. A folyamatban az összes lehetséges megoldás vagy értékek halmaza alkotja a keresési teret. Ebben a keresési térben létezik egy pont vagy azok halmaza, amely az optimális megoldást szolgáltatja. A genetikai algoritmusokban a meglévő probléma lehetséges megoldásainak sokasága helyezkedik el, majd ezek mutáción és rekombináción mennek keresztül, mint a természetben fellelhető genetikai folyamatokban és jelenségekben. Ezen mechanizmusok segítségével új egyedek jöhetnek létre és a folyamat több generáción keresztül megismétlődik. Minden megoldást jellemez egy úgynevezett fittségi érték, amely a célfüggvény értéke alapján kerül meghatározásra. A rátermettebb egyedek nagyobb esélyt szereznek arra, hogy párosodjanak. Ez egyezést mutat Darwin C. R. evolúciós elméletével [5]. Ebből az következhet, hogy folyamatosan fejlődnek ki jobb megoldások, amelynek csak a megállítási kritérium szab határt. A GA-k spontán folyamatok, de sokkal eredményesebbek, mint a véletlenszerű lokális keresés.

2.1. A genetikus algoritmusok alapfogalmai

A genetikai algoritmusok vizsgálata során meglehetősen fontos, hogy szó essen a módszer alapvető terminológiáiról.

Populáció: az adott probléma összes kódolt megoldásának a részhalmaza,
Kromoszómák: egy lehetséges megoldás az adott problémára,
Gén: egy kromoszómának az egyik elemi pozíciója,
Allél: az az érték, amelyet egy gén egy adott kromoszómához vesz,
Genotípus: a számítási tér populációja,
Fenotípus: az a populáció a valós megoldási térben, amelyben a megoldások oly módon kerülnek ábrázolásra, ahogyan azok a valós világban jelennek meg.
Dekódolás és kódolás: egyszerű problémák esetén a fenotípus és a genotípus terei megegyeznek, azonban a legtöbb esetben a fenotípus és a genotípus terei eltérőek. A dekódolás egy megoldás transzformációja a genotípusból a fenotípus térbe, míg a kódolás a fenotípusból a genotípus térbe történő transzformáció folyamata.
Fitnesz függvény: az egyszerűen definiált fitneszfüggvény olyan függvény, amely a megoldást bemenetként veszi, és kimenetként a megoldás alkalmasságát állítja elő.
Genetikai operátorok: ezek változtatják meg az utódok genetikai összetételét. Ide tartozik a keresztezés, a mutáció és a szelekció [4], [7].

2.2. A megoldások kódolása

A genetikai algoritmus megvalósítása során az egyik legfontosabb szempont az, hogy milyen reprezentációt kívánunk használni a megoldásaink ábrázolásához [2], [4]. A bináris ábrázolás az egyik legegyszerűbb és széles körben használt ábrázolás a genetikus algoritmusokban. Egyes problémáknál, amikor a megoldási tér Boole-féle döntési változókból áll, azaz igen vagy nem, a bináris ábrázolás természetes.

0	1	1	0	1	0	1	0	1	1

1. Ábra. A genotípus bináris ábrázolása.

Azoknál a problémáknál, ahol a géneket folytonos, nem pedig diszkrét változókkal lehet definiálni, a valós értékű reprezentáció a legcélszerűbb. Ezeknek a valós értékű vagy lebegőpontos számoknak a pontossága azonban a számítógép függvénye lehet.

0,5	0,4	0,1	0,2	0,6	0,7	0,3	0,7	0,2	0,4

2. Ábra. A folytonos változókkal rendelkező genotípus ábrázolása.

A diszkrét értékű gének esetében nem korlátozható mindig a megoldási teret bináris értékekre. Ha négy szempont között kell különbséget tenni, akkor ilyen esetekben az egész számok ábrázolása lehet célszerű.

1	4	2	3	1	2	3	2	4	3

3. Ábra. A diszkrét változókkal rendelkező genotípus ábrázolása.

Sok problémában magát a megoldást az elemek sorrendje jelenti. Ilyen esetekben a permutációs ábrázolás a legkézenfekvőbb megoldás.

2	5	10	6	4	6	2	8	9	7

4. Ábra. A genotípus permutációs ábrázolása.

2.3. Népesedési modellek és a szülőválasztás

Az állandósult állapotú genetikus algoritmusok növekményes algoritmusoknak is nevezhetők. A generációs modellben „n" utód generálódik, ahol „n" a populáció mérete, és az iteráció végén a teljes populáció helyettesítődik. A fitnesz érték kiszámítása ismételten történik [4]. A legtöbb esetben az alkalmassági függvény és a célfüggvény megegyezik azzal, hogy a cél az adott célfüggvény maximalizálása vagy minimalizálása. A szülők kiválasztása nagyon fontos a genetikus algoritmusokban a konvergencia aránya szempontjából, mivel a megfelelő szülők jobb megoldásokat eredményeznek.

A fitnesz arányos kiválasztás a szülőválasztás egyik legkedveltebb módja. Ebben minden egyén fitneszértékével arányos valószínűséggel válhat szülővé. Az alkalmassággal arányos szelekciónak több esete lehetséges:

Rulettkerék kiválasztás: egy kördiagram „n" darab egységre kerül felosztásra és az egyes körcikkek területe arányos a szülők fitneszértékével. Egy fixpont segítségével, a diagramot megforgatva megtörténik a kiválasztás,
Sztochasztikus mintavétel (SUS): hasonlít a rulettkerék kiválasztásához, azonban több fixpont létezik, így az összes szülő egy időben kerülhet kiválasztásra.
Versenyválasztás: véletlenszerűen kerül kiválasztásra „k" egyed a populációból majd közülük kiválasztásra kerül a legjobb, amely szülővé válhat. A „Tournament Selection" az irodalomban is rendkívül népszerű, mivel akár negatív fitnesz értékekkel is működhet [10].
Rangsor kiválasztás: a populáció minden egyedét fitneszértéke szerint rangsorolják. A szülők kiválasztása az egyén rangjától függ és nem az alkalmasságtól. Akkor használatos, ha a populáció egyedeinek fitneszértékei nagyon közel állnak egymáshoz.
Véletlenszerű kiválasztás: a szülők véletlenszerűen kerülnek kiválasztásra a meglévő populációból. Nincs szelekciós nyomás a rátermettebb egyénekre, ezért ezt a stratégia nem népszerű [4], [7].

2.4. A „crossover" keresztezési operátorok

A keresztezési operátor esetén egynél több szülő kerül kiválasztásra és azok genetikai adottságait felhasználva egy vagy több utód kerül előállításra.
A „One Point Crossover" egypontos keresztezés egy véletlenszerű keresztezési pontot választ ki, majd a két szülő genotípusának egy-egy részét cseréli fel, hogy ezáltal új utódokhoz juthasson [8].

0	1	2	3	4	5	6	7	8	9	→	0	1	2	3	4	2	4	6	4	7
4	7	8	3	1	2	4	6	4	7		4	7	8	3	1	5	6	7	8	9

5. Ábra. Az egypontos keresztezés.

A „Multi Point Crossover" többpontos keresztezés az egypontos keresztezés továbbfejlesztése, ahol a váltakozó szegmenseket felcserélik, hogy új utódokat kapjanak.

0	1	2	3	4	5	6	7	8	9	→	0	1	2	4	6	3	6	7	8	9
5	6	9	4	6	3	5	6	5	9		5	6	9	3	4	5	5	6	5	9

6. Ábra. A többpontos keresztezés.

Az egységes keresztezésben nem kerül felosztásra a kromoszóma különböző szegmensekre, hanem minden gén külön kezelendő. Lényegében minden kromoszómához igen-nem alapon eldöntésre kerül, hogy bekerül-e az utódba vagy sem.

0	1	2	3	4	5	6	7	8	9	→	0	7	2	3	1	2	6	6	8	7
4	7	8	3	1	3	4	6	4	7		4	1	8	3	4	5	5	7	4	9

7. Ábra. Az egységes keresztezés.

A teljes aritmetikai rekombináció, általában egész számok ábrázolása esetén használatos, amely úgy működik, hogy a két szülő súlyozott átlagát veszik alapul a következő képletekkel:

$$\text{utód}=\alpha\cdot x+(1-\alpha)\cdot y \qquad (1)$$

Nyilvánvaló, hogy α=0,5 esetén az utódok meg fognak egyezni egymással.

0,1	0,1	0,2	0,2	0,3	0,3	0,4	→	0,2	0,15	0,2	0,2	0,3	0,25	0,3
0,3	0,2	0,2	0,2	0,3	0,2	0,2		0,2	0,15	0,2	0,2	0,3	0,25	0,3

8. Ábra. A teljes aritmetikai rekombináció.

A „Davis Order Crossover" vagy rövidítve „OX1" egy permutáció alapú keresztezéshez használatos. Működésében a szülőben két véletlenszerű keresztezési pont által kimetszett szakasz átmásolódik az utódra. A második szülő második keresztezési pontjától kezdve átmásolódik a fennmaradó fel nem használt számsor a második szülőről az első utódra. Ez megismétlődik a második utódnál is a szülői szerep megfordításával.

0	1	2	3	4	5	6	7	8	9	→				3	4	5	6			
9	8	0	2	9	1	2	3	5	6											
0	1	2	3	4	5	6	7	8	9	→	3	5	6	3	4	5	6	9	8	0
9	8	0	2	9	1	2	3	5	6											

9. Ábra. Az OX1 keresztezés.

Számos keresztezési művelet létezik, amelyet számos publikáció és szakirodalom taglal, mint például a „Partial Mapped Crossover", amelynek rövidítése „PMX", az „Order Based Crossover", amely rövidítése „OX2", a „Shuffle Crossover" és a „Ring Crossover" [6], [8], [9], [10].

2.5. A mutációs operátorok

A mutáció röviden úgy jellemezhető, mint egy kis véletlenszerű módosítás a kromoszómalánchan, hogy további új megoldás jöhessen létre. Ezeket az operátorokat a genetikai populáció sokféleségének fenntartására és bevezetésére használják. Megfigyelések alapján a mutáció elengedhetetlen a genetikus algoritmusok konvergenciájához, ellentétben a keresztezési műveletekkel.

Az egyik leg alapvetőbb mutációs operátor a „Bit Flip Mutáció", amely esetén kiválasztásra kerül egy vagy több véletlenszerű gén, amelyek megfordításra kerülnek. Ez az operátor a binárisan kódolt genetikus algoritmusok esetében használatos [8], [9], [10].

0	1	1	0	1	1	1	1	0	0	→	0	1	1	0	1	0	1	1	0	0

10. Ábra. A „Bit Flip" mutáció.

A véletlenszerű átállítás „Random Reseting" a „Bit Flip" operátor adoptálása az egész számokkal való kódolás esetére, amelyben a megengedett értékek halmazából kiválasztásra kerül egy véletlenszerű érték, ami felcseréli a véletlenszerűen kiválasztott gént értékét.

0	1	2	6	8	2	4	7	5	8	→	0	1	2	6	8	2	4	2	5	8

11. Ábra. A „Random Reseting" mutáció.

A „Swap" mutáció esetén véletlenszerűen kiválasztásra kerül két gén a kromoszómaláncon belül, amelyek utána felcserélődnek. Ez a metódus a permutáció alapú kódolásoknál gyakori.

8	1	7	6	9	5	4	7	2	8	→	8	1	7	6	9	2	4	7	5	8

12. Ábra. A „Swap" mutáció.

A „Scramble" mutáció esetén a teljes kromoszómából kiválasztásra kerül a géneknek egy részhalmaza, majd az érték véletlenszerű keverésen esnek át. A „Scramble" mutáció a permutációs alapú kódolásnál lett népszerű eljárás.

0	8	3	6	9	2	4	7	2	8	→	0	8	3	9	2	4	6	7	2	8

13. Ábra. A „Scramble" mutáció.

Az inverziós mutáció esetén kiválasztásra kerül a gének egy részhalmaza, mint például a „Scramble" mutációnál, de a részhalmaz megkeverése helyett csupán a teljes karakterláncot invertálódik az alhalmazban.

0	4	3	6	9	2	4	7	2	8	→	0	4	3	4	2	9	6	7	2	8

14. Ábra. Az inverziós mutáció.

A túlélők kiválasztására vonatkozó feltétel meghatározza, hogy mely egyedeket szükséges törölni és melyikeket kell megtartani a következő generációban. Néhány genetikus algoritmus úgynevezett elitizmust alkalmaz, amely azt jelenti, hogy a népesség aktuális legrátermettebb tagja mindig a következő generációba kerül. Ezért a legrátermettebb egyed semmilyen körülmények között nem módosítható meg.

Az életkor alapú kiválasztás az a módszer, ahol a fitneszérték nem ismert. Arra a feltevésre alapul, hogy minden egyed egy véges nemzedék erejéig kerülhet bele a populációba. Így, ha már a feltétel nem teljesül, akármilyen jó az egyed fitneszértéke, az kiszorul az adott populációból [8], [9], [10].

Életkor	Fitneszérték	Kromoszóma
1	6	P1
4	5	P2
6	2	P3
9	2	P4
3	8	P5
7	9	P6
10	9	P7
3	5	P8
5	4	P9

Meglévő populáció

+

Fitneszérték	Kromoszóma
2	C1
9	C2

→

Életkor	Fitneszérték	Kromoszóma
3	8	P1
2	6	P2
5	3	P3
0	2	C1
4	4	P5
6	7	P6
0	9	C2
3	2	P8
2	1	P9

Új populáció

15. Ábra. Az életkor alapú szelektálás.

Az ábrán látható példában az életkor az a szám, amelyet egy adott egyed egy populációban megél. A populáció legidősebb tagjai kikerülnek a populációból, a többi egyed életkora pedig egyel növekszik. A fitneszérték alapú szelekcióban az utódok általában a populáció legrosszabb fittségi értékkel rendelkező egyedeit helyettesítik. A legkevésbé alkalmas egyedek kiválasztása történhet versenykiválasztás vagy fitneszértékkel arányos kiválasztás.

Fitneszérték	Kromoszóma
4	P1
6	P2
3	P3
1	P4
8	P5
9	P6
2	P7
7	P8
8	P9

Meglévő populáció

+

Fitneszérték	Kromoszóma
1	C1
2	C2

→

Fitneszérték	Kromoszóma
2	P1
4	P2
3	P3
1	C1
4	P5
5	P6
2	C2
2	P8
1	P9

Új populáció

16. Ábra. A Fitneszérték alapú szelektálás.

Az ábrán megfigyelhető, hogy a legkisebb fittneszértékkel P4 és P7 egyedek rendelkeznek, ennek megfelelően ők kerülnek lecserélésre.

2.6. Leállítási feltételek

A genetikai algoritmus befejezési vagy leállítási feltétele dönt arról, hogy mikor ér véget egy genetikus algoritmus futása. Megfigyelések alapján, kezdetben az algoritmus nagyon gyorsan fejlődik, de a későbbi szakaszokban, ahol a differencia már nagyon kicsi, ott jelentősen lelíttődik. Jó megoldás lehet működtetni egy számlálót, amely nyomon követi azokat a generációkat, amelyeknél nem történt javulás a populációban. Az algoritmus akkor fejeződik be, amikor a számláló elér egy előre meghatározott értéket [7], [8], [9], [10].

2.7. A genetikai algoritmusok előnyei és hátrányai

A genetikai algoritmusok gyorsabbak és hatékonyabbak a hagyományos módszerekhez képest. Folyamatos és diszkrét funkciókat, valamint többcélú problémákat egyaránt optimalizál. A „jó" megoldások listáját tartalmazza, nem csak egyetlen megoldást. Mindig választ kap a problémára, ami idővel javul. Azonban a genetikus algoritmusok nem alkalmasak minden típusú problémára, különösen azokra nem, amelyek egyszerűek és származtatott információk állnak rendelkezésre. Az alkalmassági érték kiszámítása ismételten történik, ami bizonyos problémák esetén költséges lehet a számítási szempontok alapján. Sztochasztikus folyamatot feltételezve nincs garancia arra, hogy a megoldás optimális és a legjobb minőségű.

3. Összefoglalás

Az alábbi tanulmány összefoglalta a genetikai algoritmusok legfontosabb fogalmait és működésének alapjait. Bemutatta a folyamatban szereplő legfontosabb lépéseket, az úgynevezett genetikai algoritmus operátorokat, amelyek főként a keresztezési és a mutációs műveletek. Szó esett a népesedési modellekről, a szülőválasztás, szelekciós folyamatokról és azok metódusairól, továbbá bemutatásra kerültek az egyes leállítási feltételek is.

Irodalomjegyzék

1. Erdősné, S. Cs.; Gyurecz, Gy.; Janik, J.; Körtélyesi, G.: *Mérnöki Optimalizáció*, egyetemi tananyag, (2012) ISBN 978-963-279-538-6
2. David, E. G.; Manohar, P. Samtani.: *Engineering optimization via genetic algorithm*, ASCE, pp. 471-482., (1986)
3. Holland, J.: *Adaptation in Natural and Artificial Systems*, MIT Press, (1992), ISBN 978-0262581110.
4. Mitchell, Melanie.: *An Introduction to Genetic Algorithms*, MIT Press, (1996), ISBN 9780585030944
5. Darwin, C. R.: 1872. *The origin of species by means of natural selection, or the preservation of favoured races in the struggle for life*, (1872), London
6. David, E. G., C.H. Kuo.: *Genetic algorithms in pipeline optimization*, Journal of Computing in Civil Engineering, (1987)
7. Z. Michalewicz.: *Algorithms + data structures = evolution programs*, Springer, (1994)
8. R. L. Haupt, S. E. Haupt.: *Practical genetic algorithms*, (2004), ISBN 0-471-45565-2
9. Kalyanmoy D.: *Multi-Objective Optimization Using Evolutionary Algorithms: An Introduction, (2011)*
10. J. Yang, C. K. Soh.: *Structural optimization by genetic algorithms with tournament selection*, Journal of computing in civil engineering, (1997)

SZABÓ KRISTÓF
egyetemi tanársegéd és PhD hallgató
Miskolci Egyetem, GÉIK - Szerszámgépészeti és Mechatronikai Intézet, Szerszámgépek Intézeti Tanszéke

2023.