

KLÍMAMODELLEZÉS ÉS ELOSZTOTT RENDSZEREK

Salga Péter, salga@thor.agr.unideb.hu

Debreceni Egyetem, AVK, Gazdasági- és Agrárinformatikai Tanszék

Hufnagel Levente, levente.hufnagel@uni-corvinus.hu

Budapesti Corvinus Egyetem, KTK Matematika és Informatika Tanszék

Bevezetés

A klímaváltozás mint lehetséges globális környezeti probléma, ma a tudomány és a világpolitika egyik központi kérdésévé vált. A világ kutatóinak döntő többsége egyetért abban, hogy megkezdődött a globális felmelegedés időszaka. Nemzetközi adatok szerint 1950 és 2003 között a Föld felszínének átlaghőmérséklete 13,87 °C-ról 14,52 °C-ra növekedett. A fosszilis energiaforrásokból 1950-ben 1612 millió tonna szén jutott a légkörbe, 2003-ban pedig 6999 millió tonna. A viták ma már arra koncentrálnak, hogy a globális felmelegedés csupán természeti jelenség-e, vagy pedig antropogén hatások következménye is. Valószínű, hogy mindkét tényező együttesen érvényesül, és kölcsönösen felerősítik egymást. A globális felmelegedés következménye a szélsőséges meteorológiai események számának és intenzitásának növekedése, ami a globális vagy regionális klíma megváltozását eredményezheti (VAHAVA, 2006).

A kutatások az elmúlt két-három évtizedben több területen folytak:

- a klímaváltozás jellemzése, klímaszcenáriók kidolgozása,
- a klímaváltozás várható természeti, gazdasági és társadalmi hatásainak feltárása,
- a klímaváltozáshoz való alkalmazkodási stratégiák kidolgozása.

A klímaváltozás és hatásai, mint sztochasztikus rendszerek, belső törvényszerűségekkel rendelkeznek, még ha ezek nem is definiálhatóak matematikai egzaktsággal. Ez a felismerés valószínűsíti az analóg technikák alkalmazásának eredményességét a klímaváltozás hatásainak vizsgálatánál is.

Feldolgozva az elmúlt évek időjárási, mezőgazdasági és talajtani adatait, mintaelemzésen alapuló becslés valószínűleg meg neurális hálózatok segítségével viszonylag nagy számú paraméter figyelembe vételével. A rendszer működéséhez nincs szükség matematikai szabályok és algoritmusok kidolgozására, a hálózat a hasonlóság elve alapján osztályozza az időbeli paraméter-értékeket, és nagy megbízhatóságú becslést tesz a jövőbeli értékekre.

Mivel a bemeneti paraméterek hézagossága vagy bizonytalansága miatt fuzzy bemeneti-értékek használandóak, valamint az előrejelzés nagy időintervallumokra történik, a kimeneti értékek inkább tendenciákat jeleznek a rendszerben. A modell fontos alkalmazási területe lehet a különböző kiindulási klíma-paraméterek mezőgazdasági hatásainak összehasonlítására, elemzésére is.

Nagy számú paraméter együttes vizsgálata a hagyományos klíma modellek segítségével igen nagy számítási kapacitást igényel, amely csak elosztott rendszerek segítségével oldható meg. A párhuzamos feldolgozás segíti az ún. paraméter-tanulmányok kidolgozását is, ahol ugyanaz a modell kerül kiértékelésre akár több milliószor, a kezdeti paraméterek kis változtatásaival.

Szintén a nagy számítási kapacitás teszi lehetővé, a hagyományos szcenáriók elemzését olyan szélsőséges időjárási szituációkban, amelyeket éppen a klímaváltozás modelljei jósolnak, de eddig nem kerültek be a modellek vizsgált paramétereire közé.

A modern információtechnológia segítséget nyújthat az előrejelzések pontosításában. Az elosztott rendszerek (klaszter, grid rendszerek) a nagy számításigényű, szcenáriók kidolgozásában jelenthetnek segítséget. A neurális hálózat a mintaelemzéses becsléseknél, míg a fuzzy logika a bizonytalan mértékű paraméterek kezelésénél alkalmazható. A különféle

optimalizáló eljárások (pl. genetikus algoritmus) pedig a scenáriókra adható legjobb technológiai válaszok elemzését teszik lehetővé.

A klímaváltozás mezőgazdasági hatásmodelljei

A klímaváltozás a meteorológiai elemek átlagos értékeinek olyan jelentős megváltozását jelenti, amelynek környezeti, gazdasági, társadalmi hatása van. Az átlagos értékek változása hosszabb periódusra vonatkozik. (Harnos N., 2003)

A klímaváltozás mezőgazdasági hatásmodelljeinek vizsgálata egyértelműen arra enged következtetni, hogy

- az agroökológiai zónák eltolódnak a hőmérsékletemelkedés és a vízhasznosítási hatékonyság javulása következtében,
- a növénytermesztés hatékonysága a középső és magas szélességi övezetekben (alapvetően a fejlett országokban) javulni fog - a növekvő fotoszintézis, a hosszabb tenyészidőszak és fagymentes periódus miatt,
- a legtöbb fejlődő országban a termőképesség csökkenni fog (a cereáliák esetében mintegy 10 %-kal), ami jelentős ellátási gondokat jelent az adott térségben,
- a trópusi és szubtrópusi régiókban, Afrika Száhel-övezetében válik legkritikusabbá a helyzet. (Fisher, G. et al. 2001)

A klímaváltozás modellezése többszálú kutatás, ahol az ún. scenáriók eredményei jelentősen eltérhetnek egymástól (bár tendenciában megegyeznek), ezért a mezőgazdaságra gyakorolt hatásokat is csak az egyes modell-scenáriók függvényében tudjuk vizsgálni.

A vázolt általános érvényű megállapítások ismeretében felmerül a kérdés, hogyan tudunk következtetni a fontosabb gazdasági növények hozamára, az egyes régiók, illetve a világ élelmiszertermelési kapacitására, biztonságára. A legáltalánosabban használt eljárás a növényi növekedési modellek alkalmazása (szimulációs modellezés).

Bár a klíma-növény kapcsolat vizsgálatánál figyelembe szokás venni a

- csapadék,
- hőmérséklet,
- CO₂-koncentráció,
- termőhely, földhasználat

paramétereket, a modellek vizsgálati alapvetően két paramétercsoportra (illetve az azokból levezethető tényezőkre) korlátozódnak:

- hőmérsékletnövekedésre,
- a légköri CO₂-koncentrációra. (Harnos Zs., 2005)

Nem veszik számításba pl. a technológia változását, a növények genetikai változását, az extrém időjárási szituációk kialakulását és a maradék paramétereket is általában kisebb csoportokra elkülönülten vizsgálják.

A növényi növekedési modellek általánosan használtak ma már a kísérlettervezésben, a termelésstervezésben, a mezőgazdasági területek, régiók termőképességének jellemzésére. A legismertebb az IBSNAT modellesalád, amelyet a legfontosabb szárazföldi növényekre (őszi búza, kukorica, szója stb.) dolgoztak ki és használnak világszerte, többek között a klímaváltozás várható hatásainak az elemzésére is. Európában legelterjedtebb modellek az AFRCWHEAT és a SIRIUS, illetve ezek adaptált, továbbfejlesztett változatai. Magyarországon az AFRCWHEAT adaptált változatával készültek számítások (Harnos N., 2003.), amelyek az eredményei összefoglalva láthatók az alábbi táblázatban.

	Tényleges termés (t/ha)		Szimulált termés (t/ha) AFRCWHEAT	
	Várható érték	Szórás	Várható érték	Szórás
Győr				
Tényleges időjárás	4,88	0,63	4,9	0,5
HADCM 2 klímaszcenárió				
CO ₂ -koncentr.: 360 ppm			3,7	0,4
CO ₂ -koncentr.: 500 ppm			4,2	0,5
Debrecen				
Tényleges időjárás	4,69	0,69	4,8	0,6
HADCM 2 klímaszcenárió				
CO ₂ -koncentr.: 360 ppm			3,8	0,4
CO ₂ -koncentr.: 500 ppm			4,4	0,5

A búza szimulált termésátlagai (Harnos, N., 2003)

A klímaváltozás természetesen az állatvilágra, így az egyes mezőgazdasági termények kártevőire is hatást gyakorol. Sokszor éppen az egyes rovarok szezonális dinamikájának megváltozásából következhetnek a klímaváltozás hatásaira (Hufnagel, L., Gaál, M., 2005).

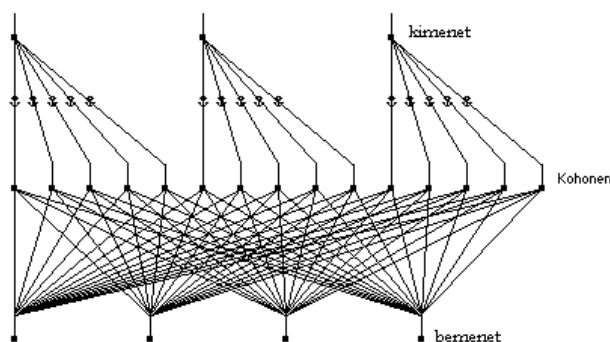
Neurális hálózatok és fuzzy rendszerek

Neurális hálózatok

1943-ban W. S. McCulloch és W. Pitts az idegrendszer működését bináris matematikai modell segítségével írták le. Absztrakciójuk a biológiai kutatásból az idegsejtmodellt számítási eszközzé tette. Felfedezésüktől azonban sok időnek kellett eltelnie ahhoz, hogy "igazi" feladatokra alkalmassá váljék. 1986-ban publikálta D. E. Rumelhart, G. E. Hinton és R. J. Williams az általános célú számításokra alkalmas neurális hálózatok tanítási algoritmusát, a backpropagation-t.

A neurális hálózatok alapelemei a neuronok. A neuron súlyozza a bemeneteire érkező jeleket, ezeket összeadja, majd a válaszfüggvény segítségével számítja a kimenet értékét. Leggyakrabban az egyszerűen deriválható logisztikus szigmoid függvényt használják válaszfüggvényként.

A neuronokból szerveződik a hálózat oly módon, hogy az egyes számító elemek rétegeket képeznek. Valamennyi réteg összes neuronja az előtte és az utána álló réteg összes neuronjával kapcsolatban van, a saját rétegének neuronjaival azonban nem.



Neurális hálózat

Az ábra egy neurális hálózat felépítését mutatja. Az egyes rétegek neuronjainak paramétereit együttesen súlymátrixba és küszöbvektorba foglalhatjuk össze. Az egész neurális hálózat kimenete így alakul:

$$y = f(\mathbf{W}_2 \cdot f(\mathbf{W}_1 \cdot \mathbf{p} + \mathbf{b}_1) + \mathbf{b}_2),$$

ahol f a logisztikus szigmoid-függvény, \mathbf{W}_1 és \mathbf{W}_2 a két réteg súlymátrixa, \mathbf{b}_1 és \mathbf{b}_2 küszöbvektorok, \mathbf{p} a bemeneti és \mathbf{y} a kimeneti vektor.

A hálózat használatához a súly- és küszöbértékeket kell meghatároznunk. A meghatározást tanulásnak nevezik. A tanulás során ismert bemenet-kimenet párok (az ún. tananyag) segítségével úgy határozzuk meg a hálózati paramétereket, hogy a hálózat hibája minimális legyen. Ennél az optimalizálási feladatnál fontos, hogy a válaszfüggvény egyszerűen deriválható legyen.

Matematikai szempontból a neuronhálózatok a való világ egyfajta statisztikai modelljei, amelyeket a paraméterek hangolásával teszünk alkalmassá a feladatok megoldására.

Köznapi értelemben – és nem teljesen precízen – minden olyan feladatra alkalmazható neurális háló, amit az ember könnyen meg tud oldani, de a digitális számítógép számára igen nehéz vagy megoldhatatlan feladatot jelent (pl. arcfelismerés, kézírás-olvasás, repülőgép vezetés).

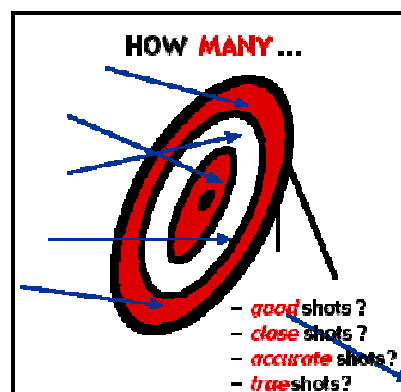
Fuzzy logika

A hagyományos logikában a változók két értéket vehetnek fel: 0-t és 1-et. Ez az érték felfogható egy halmaz elemeinek hozzátartozási értékeként is. A fuzzy logika ezt a hozzátartozást általánosítja: e két szám között végtelen sok hozzátartozás vehető fel.

Az általánosításhoz azonban át kell definiálni a logikai műveleteket. A fuzzy logika megteremtője, *L. A. Zadeh* 1975-ben az alábbi módon határozta meg a három alapműveletet:

$$\begin{aligned} \text{ÉS} &- \min(m_A, m_B) \\ \text{VAGY} &- \max(m_A, m_B) \\ \text{NEM} &- 1 - m_A \end{aligned}$$

ahol m_A az A halmaz, m_B a B halmaz hozzátartozási értéke. A fenti műveletek igazak a hagyományos logikai értékekre is.



Egy lövést jellemezhet az, hogy „vagy eltaláltuk a céltáblát, vagy nem”, de pontosabb jellemzés, ha szavakkal megadjuk azt is, hogy mennyire találtunk a céltábla közepébe

A fuzzy logika két legfontosabb alkalmazási területe a (folyamat)szabályozás és a döntéshozás, amelyek igen nagy segítséget nyújtanak pontosan nem meghatározható értékű adatok feldolgozásánál.

Neuro-fuzzy feldolgozás

A neurális hálózat bemeneti és kimeneti értékei is lehetnek bizonytalanok, még gyakrabban hiányosak, ahol a meglévő adatokból következtetünk a hiányzókra. Sokszor előfordul, hogy nem is kívánjuk pontosan megadni egy paraméter értékét, csak egy viszonyítást a többi értékhez (pl.: kisebb csapadékmennyiség, nagyon nagy CO₂ koncentráció).

A módszer használatának két rendkívüli előnye a pontosságnövekedés és a sebesség. A döntéshozatal a fuzzy logikában használatos műveleteken alapszik, ezek pedig gyorsan számíthatók.

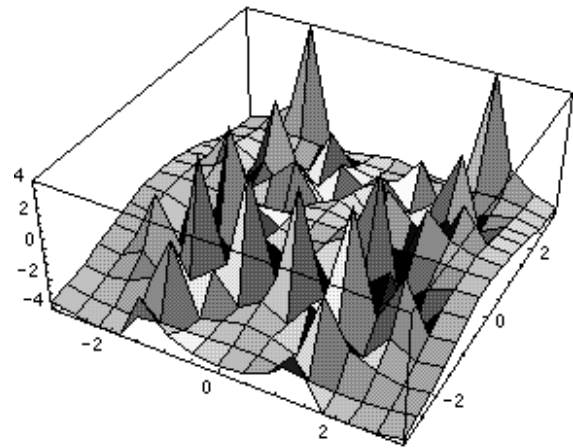
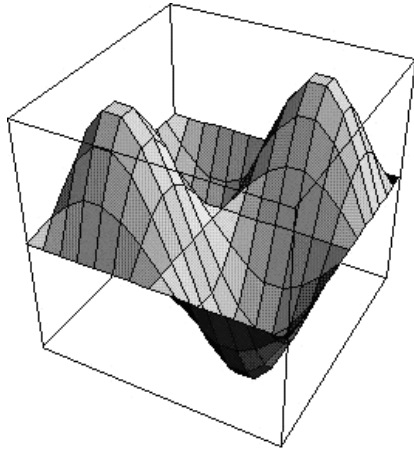
A fuzzy logikán alapuló döntés további erőssége abban áll, hogy különböző típusú adatok kezelésére nyújt jó megoldást. A szabályrendszer a fuzzy logika nyelvezetén átalakítható táblázatokká, amelyben szereplő elemek kódolják a szabályokat. A táblázatok módosítása nem jelent akkora programozói munkát, mint az "igazi" szakértői rendszereknél, így az egyre újabb és újabb adatok kezelhetők.

Genetikus algoritmus

A genetikus algoritmus az optimalizálási problémák megoldására alkalmazott módszer, mely az élőlények öröklődési folyamatait másolja. A módszer egyik fő előnye, hogy a számítástechnikában előforduló problémák egy nagyon széles osztályára alkalmazható, ugyanakkor általában nem használ területfüggő tudást, és akkor is működik, ha a feladat struktúrája kevésbé ismert.

Másképp megfogalmazva a genetikus algoritmus egy szélsőérték-keresési (optimalizáló) eljárás, mely viselkedésében a természetes evolúció folyamatát utánozza. A megoldások egy adott lépésben érvényes halmazát populációnak nevezik, és ezen az evolúciós módszertanból vett operátorokat használ a populáció egy lépésben való megváltoztatására, amely változás visszavonhatatlan.

A kritériumfüggvény felületét több ponton vizsgálja, míg a többi eljárás (strukturális algoritmusok, lineáris programozási eszközök) a felületen csak egy pontot mozgat, minek következtében – a többi eljárástól különbözően – nem csak a lokális szélsőérték helyeket találja meg, hanem globális optimummal szolgál. Az 1. ábrán a bal oldali ábra egy egyszerű feltételrendszer két szélsőérték-hellyel rendelkező megoldásterét mutatja, míg a jobb oldali ábrán egy bonyolult feltételrendszerű optimalizáció lokális maximumai láthatóak. A klasszikus eszközök csak a bal oldali ábrán képesek a globális maximumot megtalálni, a jobboldali ábrán valamelyik lokális maximumhelynél megállnak. Ezzel szemben a genetikus algoritmus a jobb oldali megoldásterben is nagy biztonsággal a legmagasabb csúcsot találja meg.



Mindenhol alkalmazható, ahol az a feladat, hogy a sok lehetséges megoldás közül a legjobbat találjuk meg, ahol az értéket egy értékelőfüggvény, másnéven rátermettségi függvény (*fitness function*) adja meg. A genetikus algoritmus megoldások egy populációját tartja fenn, azaz egyszerre több megoldással dolgozik. Az aktuális populációból minden lépésben egy új populációt állít elő úgy, hogy a szelektív operátor által kiválasztott legrátermettebb elemeken (szülőknön) alkalmazza a rekombinációs és mutációs operátorokat. Az alapgondolat az, hogy mivel általában minden populáció ez előzőnél rátermettebb elemeket tartalmaz, a keresés folyamán egyre jobb megoldások állnak rendelkezésre.

Mivel egyszerre több pontos keresést végez a problématerben, amely egyrészt kellő robusztusságot biztosít - hiszen ha egy keresési út zsákutcának bizonyul, az még nem jelenti az egész algoritmus kudarcát. Másrészt a problémának nem egy, hanem több különböző, közel optimális megoldását nyújtja, amelyből akár a felhasználó is kiválaszthatja a neki tetsző - eddig még nem deklarált szubjektív feltételeknek megfelelő - példányt.

Operátorok

Az operátorokat két csoportba szokás sorolni. Az elsőbe tartoznak azok, amelyek egy megoldásból állítanak elő egy újabbat, ezek a mutációk. A másik csoportba tartoznak azok, amelyek több kiindulási megoldásból (szülőből) állítanak elő újabb megoldást a szülők kódjainak valamilyen kombinációja segítségével. Ezeket rekombinációknak nevezik. Az itt tárgyalt kódoláshoz is megadható rengetegféle mutáció és rekombináció, itt azokat ismertetjük, amelyeket a legszélesebb körben alkalmaznak.

- Keresztezés (crossover)

Az egyik alapvető genetikus operátor. Feladata a determinisztikus keresés megvalósítása, vagyis a problématerben a meglévő tudás alapján próbál előrelépni. Segítségével két kiinduló egyedből egy újat hozhatunk létre úgy, hogy ennek genetikai anyaga egyaránt tartalmaz elemeket mindkét kiinduló egyedből.

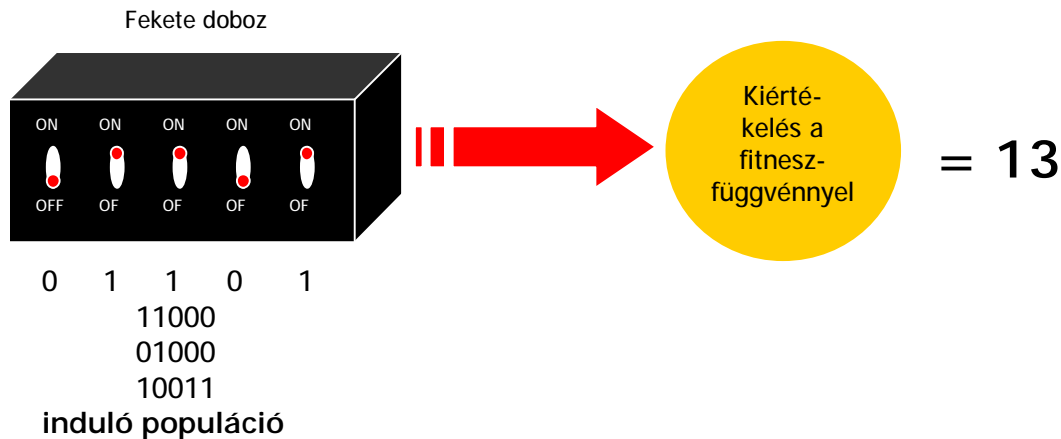
A keresztezés alkalmazásával olyan új egyedeket szeretnénk előállítani, amelyek szüleik előnyös tulajdonságait öröklik. Azt azonban nem tudjuk garantálni, hogy ez valóban így is lesz, ugyanis a genetikus operátorok nem vizsgálják meg az egyedek belső szerkezetét, nem válogatják ki belőlük a kedvező tulajdonságokért felelős részeket. A genetikus anyag egy adott szakaszáról sokszor nem is lehet eldönteni, hogy az kedvező tulajdonságokat kódol vagy sem.

- Mutáció

A másik alapvető genetikus operátor. Feladata a heurisztikus keresés megvalósítása, a problémater eddig még fel nem fedezett részében lévő lehetséges egyedekhez próbál eljutni. A mutáció során az algoritmus az egyed egy véletlenszerűen kiválasztott helyén lévő elemét az alaphalmaz egy másik elemére cseréli.

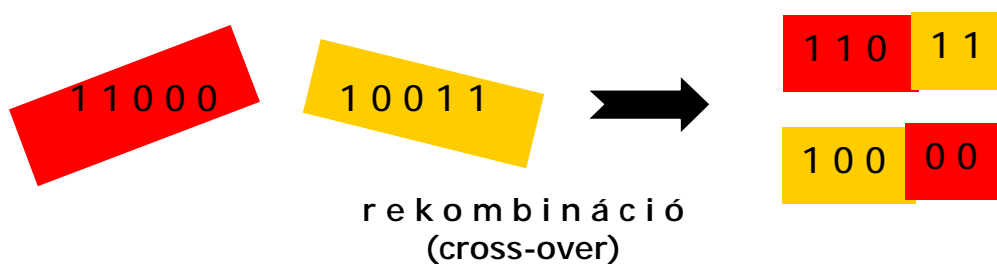
Egy egyszerű genetikus algoritmus példa

Az ábrán látható fekete doboz kapcsolóinak állása meghatározza azt a decimális számértéket, amit a doboz digitális kijelzőjén látunk. Tegyük fel, hogy nem tudjuk, hogy a kapcsoló állapotainak megfelelően bináris számot számolja át tízes számrendszerbe a fekete doboz. A feladatunk ekkor az, hogy úgy állítsuk be a kapcsolókat, hogy a lehető legnagyobb értéket lássuk a doboz kijelzőjén.



A fekete doboz feladat

A megoldás első lépéseként a kapcsolók lehetséges állapotát reprezentáló tetszőleges 5-jegyű bináris számokat adunk meg; ez lesz az induló populációnk. A populáció egyes egyedei kisselektálódnak minden szaporodási ciklus végén, úgy hogy csak a legrátermettebbek maradhatnak a populációban. A legrátermettebbek meghatározása a fitness-függvény segítségével történik, ami minden problémánál speciálisan kerül meghatározásra. A fekete doboznál a bináris számot decimálissá átalakító függvény a fitness-függvény, és a szaporodási ciklusok után csak a legnagyobb függvényértékű (legnagyobb fitnessű, legrátermettebb) egyedeket hagyjuk benn a populációban.



A rekombinációs operátor működése a fekete doboz feladatban

1 1 0 0 0



1 1 0 **1** 0

mutáció

A mutációs operátor működése a fekete doboz feladatban

A szaporodási ciklusokban az evolúciós operátorok segítségével változik meg a populáció. Az egyes egyedeken előre beállított valószínűséggel rekombináció vagy mutáció végezhető. A genetikus algoritmus eredménye igen nagymértékben függ az evolúciós operátorok beállított rátájától.

Elosztott rendszerek

A számítástechnika fejlődésének lendületét az magyarázhatja, hogy képes egyre több embert meggyőzni arról, hogy szüksége van rá, használata megkönnyíti az életét, és egyre több felhasználási területet talál magának. Az idők során a számítógépekre rábízott feladatok mennyisége mindig gyorsabban növekedett, mint az egy processzorba sűrítendő számítási teljesítmény. Kezdetben csak egyszerű számítások elvégzése volt a feladat, de mostanra eljutottunk hatalmas adatbázisok komplett feldolgozásáig, hogy azokból statisztikai adatokat nyerjünk ki. Mindeközben a meglévő számítási teljesítmény növelésének igénye mellett megjelentek egyéb, addig még nem látott kívánalmak is, mint megbízhatóság, rendelkezésre állás, szervizelhetőség. Ezen új feladatok némelyikének már csak nehezen lehet megfelelni egyetlen processzorral (ha egyáltalán meglehet).

Eljutottunk tehát a többprocesszoros rendszerek nélkülözhetetlenségéig. Ezek olyan előnyös plusz tulajdonságokkal rendelkeznek, melyek miatt megéri a többlet energia befektetést létrehozásukra áldozni. Többféle többprocesszoros rendszert különböztetünk meg. Mindegyiknek van olyan egyedi jellemzője, ami miatt mindnek van létjogosultsága.

Több processzor egy rendszerben (SMP)

A csúcscategóriás asztali PC-kben és munkaállomásokban elterjedt megoldás az SMP (Symmetrical Multiprocessing) azaz a szimmetrikus multiprocesszor. Az SMP párhuzamosan kapcsolt processzorokból áll, melyek egy közös rendszerbuszra csatlakoznak. Mint a definíció is mutatja: a párhuzamosan kapcsolt processzorok egyenrangúak, nincs semmilyen különbség közöttük, bármelyikük végre tud hajtani bármilyen feladatot.

Ezt azért fontos kihangsúlyoznunk, mert a kezdetben megjelent többprocesszoros rendszerek némelyike nem volt szimmetrikus. Ezekben a rendszerekben egy vagy több processzort különböző feladatokra dedikáltak, jellemzően az operációs rendszer vagy egyéb alrendszer futtatására. Ezek a processzorok nem futtattak felhasználói kódot. Ezeket az aszimmetrikus rendszerek hamar kiszorították a szimmetrikus modellek, mert a dedikált processzorok rendszerint rendkívül terheltek voltak, míg a felhasználó kódot futtató processzorok sokat álltak, várva arra, hogy dedikált társaik befejezzék a munkát.

Az SMP rendszerekben már minden processzoron futtatott az operációs rendszer és a felhasználói kód is, ezzel jobban eloszlott a terhelés. Sajnos, a rendszerben lévő összes erőforrás közös, azaz mindegyik processzor képes hozzáférni és vezérelni mindegyik erőforrást. Az egész rendszernek egy közös I/O egysége és memória egysége van, amelyen a processzoroknak osztozniuk kell. Ebben az esetben a több processzor csak a számítási kapacitást növeli meg, hiszen a memória és az I/O sávszélességét el kell osztaniuk egymás között, mert közösen használják őket. Így az egy processzorra jutó keresztmetszet, sávszélesség kisebb lesz, mint egy egyprocesszoros gép esetén. Ezért maximum 16 processzoros rendszert érdemes ezzel az architektúrával készíteni. Mert ekkora már annyira

leszükül a sávszélesség, hogy a processzorok mindig csak adatokra várnának. A sok processzoron egyetlen közös operációs rendszer fut.

Masszívan párhuzamos rendszerek (MPP)

A masszívan párhuzamos rendszerek (MPP – Massively Parallel Processing) esetén a nagyobb számítási kapacitás létrehozását úgy oldják meg, hogy sok azonos funkciójú építőelemet tesznek egymás mellé és összekötik őket. Így sokan hajtják végre egyszerre (egymással párhuzamosan) ugyan azt a műveletet különböző változókkal. Természetesen nem szükséges, hogy az összes építőelem teljesen azonos legyen. Lehetnek elemek, melyeknek speciális funkciójuk van a rendszerben, de az egy időben azonos feladatot végrehajtó elemek mind azonosak, vagy legalábbis azonosnak tűnnek a feladat végrehajtása szempontjából. Egy építőkocka egy processzort és egy memóriát tartalmaz. Az építőkockák (az egyes processzor – memória párosítások) egymással párhuzamosan tudnak feladatot végrehajtani, ők alkotják a rendszer csomópontjait (node). Ezeket a csomópontokat a nagy darabszámuk miatt igen sokféle módon lehet összekapcsolni. Az egyes összekapcsolási módok a szomszédok számát, a gyűrűalkotás lehetőségét is definiálják. Igen bonyolult, egyedi célokra specializált számítógépet tudnak létrehozni az összeköttetések megfelelő megtervezésével. Miután egyedileg kell megtervezni, minden egyes speciális esetre a rendszert, és az igen sok elemből áll, ezért igen drága az előállítás. Speciális feladat lehet például nagyméretű mátrixokon való gyors műveletvégzés.

Az egyes csomópontokat speciális hálózat köti össze. A kész rendszer általában nem bővíthető, csak akkor, ha összeállításakor nem töltötték fel benne az összes csomópontot processzor és memória elemmel. Az egyes építőelemek darabonként nem fejleszthetőek, esetleg egyszerre az összes egyidejű cseréjével. Idővel, a technológia fejlődésével az egészet, az alapjaitól kell újratervezni. Az új processzorok új összeköttetési módokat követelhetnek maguknak. A rendszer processzorainak száma elérheti a több százat is. Ennek a megvalósításnak hatalmas előnye, hogy nincs előre definiált rendszerbusz, ami korlátozhatná a sávszélességet. A rendszer tulajdonságai nagymértékben igazíthatóak a teljesítendő igényekhez, hiszen egyedileg kell tervezni. Persze ez sok pluszmunkával és borsos árral jár. De bizonyos esetekben elkerülhetetlen. Miután a rendszert egyedi feladatok megoldására élezi ki, ezért nem lesz univerzális. Azaz amire tervezték, azt nagyon gyorsan fogja végrehajtani, de ettől még bizonyos feladatokat lehet, hogy egy asztali PC-nél is lassabban fog elvégezni.

Az egyes csomópontok felhasználásának két fő lehetősége van, vagy mindegyik csomópont azonos feladatot lát el, vagy soros feldolgozást végeznek, azaz egy csomópont a végrehajtás egy fázisáért felel csak, annak végrehajtása után továbbadja a következő csomópontnak további végrehajtásra.

Klaszter (Cluster)

A klaszter egy csoport önálló, egyedül is működőképes számítógép. A rendszer minden csomópontja (node) egy-egy önállóan is működőképes számítógép. A csomópontok szabványos kommunikációs hálózattal vannak összekötve. Ez a szabványos hálózat akár egy *switch* is lehet. Mivel a rendszer szabványos elemekből építhető, ezért az egységnyi számítási kapacitásért itt kell legkevesebbet fizetni, ezért nevezik a klasztert a „szegényember szuperszámítógépének”. A szabványos elemekből való építkezés másik előnye a könnyű építhetőség és a rendszer méretének jó skálázhatósága. A rendszer az alkalmazások felé egy nagy erőforrás képét mutatja. A programoknak nem kell feltétlenül támogatniuk a klaszteren való futtatást. Ha más architektúrára lettek optimalizálva, akkor is futni fognak, igaz nem olyan sebességgel, mint az eredeti célrendszeren. Ha egy csomópont meghibásodik, akkor annak feladatait, - a hiba detektálása után - a rendszerbe kapcsolt másik gép átveheti. A

rendszerben nem egyetlen operációs rendszer van, hanem annyi, ahány különálló számítógép. A klaszterek létrehozásának okai között szerepelhet az alábbi három fontos szempont:

- Magas rendelkezésre–állású klaszterek: két vagy több csomópontot tartalmaz, melyek egymás tükrei, azaz tartalmuk megegyezik. Ha az elsődleges rendszer meghibásodik például hardverhiba miatt, akkor a megfelelő hibakezelő szoftver a második rendszert teszi meg elsődleges rendszerré, és így megszakítás nélkül lehet dolgozni a számítógépekkel.
- Terhelés-elosztó klaszterek: nagyvállalatok előszeretettel alkalmaznak klasztereket a web szolgáltatásokhoz, vagy az adatbázis tranzakciók kezeléséhez. A klaszter a bejövő kérést ahhoz a csomóponthoz irányítja, amelyik a leginkább képes ezt lekezelni.
- Tudományos számítások: nagy műveletigényű tudományos számítások elvégzéséhez is remekül használható. Az MPP architektúrával ellentétben sokkal egyszerűbben és olcsóbban bővíthető, gyakorlatilag csak egy újabb számítógép hálózatba kötését, és minimális újrakonfigurálást igényel. A legtöbb szoftver, amely speciálisan klasztereken történő futtatáshoz készült, tetszőlegesen skálázható, nem igényel újrafordítást, ha megváltozik a klaszter mérete.

Előnyök és hátrányok – összehasonlítás

A fent ismertetett architektúrák mindegyikének ismeretes előnye és hátránya is. Az SMP teljesítménye csak a processzorok számával növelhető (egy ideig lineárisan), azonban hamar elérjük azt a felső korlátot, amely fölött már a buszrendszer keresztmetszete szab határt a teljesítménynek. A speciális, több processzort kezelő hardver ráadásul drága. A szuperszámítógép nagyon gyors, amikor arra a feladatra használjuk, amire megépítették. Ezért cserébe csillagászati összegeket fizethetünk, és speciálisan kell megírni programjainkat, hogyha futtatni akarjuk szuperszámítógépünkön. A klaszterek könnyen bővíthető, jó ár/teljesítmény megoldást jelentenek a fentiekkel szemben. Sajnos azonban itt is mindenhez speciális szoftver szükséges. Figyelemre méltó azonban, hogy az 500 leggyorsabb számítógép listájának legdinamikusabban növekvő részét pontosan a klaszterek teszik ki.

Elosztott rendszerek

„Az elosztott rendszer önálló számítógépek olyan összessége, amely kezelői számára egyetlen koherens rendszernek tűnik.” Ez a definíció két feltételt fogalmaz meg. Az első szerint az elosztott rendszer több számítógépből áll, amelyek lehetnek ugyanolyanok vagy különbözőek. Az előbbi esetén homogén, míg az utóbbi esetén heterogén a rendszerről beszélünk. A másik feltétel szerint a rendszernek el kell rejtenie az első tulajdonságot, azaz több szempontból is egységes képet kell mutatnia a felhasználók felé. Ez utóbbi tulajdonságot nevezzük átlátszóságnak.

A hardver szempontjából valójában minden elosztott rendszer pusztán több CPU együttese. A rendszerek lehetnek a CPU-k összekötésének módja szerint busz alapúak vagy kapcsolóhálózat alapúak, de csoportosíthatóak a CPU-hoz tartozó memória elhelyezkedése szerint is. Ha minden CPU ugyanazt a memóriát látja, akkor közös memóriásnak nevezzük, míg ha minden CPU-hoz külön memória tartozik, akkor osztott memóriásnak nevezzük a rendszert.

Elosztott rendszerek építésének egyik legegyszerűbb lehetősége, ha több különálló számítógépet hálózattal kapcsolunk össze. Az elosztott programok futtatásának céljából összekötött számítógépek egy úgynevezett klasztert (*cluster*) alkotnak. A klaszterek állhatnak olyan munkaállomásokból, amelyekre az a jellemző, hogy a tulajdonosaik időnként használják őket, de az idő nagy részében csak szabadon állnak. Ezeknél a szabad számítási kapacitás kihasználása a cél.

Egy másik lehetőség, hogy kizárólag elosztott programok futtatásának céljából kötnek össze számítógépeket. Ezeket a klasztereket nevezzük dedikált klasztereknek. A dedikált klaszterek manapság felveszik a versenyt a szuperszámítógépekkel, azok árának töredékéért.

Grid

Az adatintegráció és központosított adatelérés leghatékonyabb és legígéretesebb eszközei kétségkívül a grid rendszerek (Buyya, 1999). Az osztott számítógépes rendszerek teljesen új megközelítést hozta a grid technológia, melyen a hálózatba kapcsolt számítógépek kommunikációs protokolljait, az általuk kínált szolgáltatásokat, és az erőforrásaik elosztását meghatározó infrastruktúrát értünk. A grid úgy fogja össze a benne lévő számítógépeket, hogy azok megmaradnak önálló rendszereknek lenni, miközben felesleges erőforrásaikkal dolgoznak a közös feladaton (Buyya, 2002). A grid rendszerek segítségével hatékony adatintegrációra van lehetőség, és a megsokszorozott processzorok szuperszámítógép méretű számítási kapacitást adnak a nagy erőforrás-igényű programok futásához. Ezen kívül a grid technológia lehetőséget nyújt a mérőeszközök, monitorozó rendszerek kollaboratív használatára, a felhasználónak csak a gridbe kell belépnie, a többit az infrastruktúra elrejt a grid felhasználója előtt.

Klíma-változás hatásainak vizsgálata elosztott neuro-fuzzy rendszerek segítségével

A klíma-változás és hatásai, mint sztochasztikus rendszerek, belső törvényszerűségekkel rendelkeznek, még ha ezek nem is definiálhatóak matematikai egzaktsággal. Ez a felismerés valószínűsíti az analóg technikák alkalmazásának eredményességét a klíma-változás hatásainak vizsgálatánál is.

Feldolgozva az elmúlt évek időjárási, mezőgazdasági és talajtani adatait, mintaelemzésen alapuló becslés valószínűleg meg neurális hálózatok segítségével viszonylag nagy számú paraméter figyelembe vételével. A rendszer működéséhez nincs szükség matematikai szabályok és algoritmusok kidolgozására, a hálózat a hasonlóság elve alapján osztályozza az időbeli paraméter-értékeket, és nagy megbízhatóságú becslést tesz a jövőbeli értékekre. Hasonló rendszereket sikeresen alkalmaznak tőzsdei előrejelzésben (Alyuda), meteorológiában (Hsieh, W et al., 1998) vagy folyók vízszintjének előrejelzésében (Zhang, B. et al, 1998), de a klíma-modellezésben is (Krasnopolsky, V. M. et al, 2005).

Mivel a bemeneti paraméterek hézagossága vagy bizonytalansága miatt fuzzy bemeneti-értékek használandóak, valamint az előrejelzés nagy időintervallumokra történik, a kimeneti értékek inkább tendenciákat jeleznek a rendszerben. A modell fontos alkalmazási területe lehet a különböző kiindulási klíma-paraméterek mezőgazdasági hatásainak összehasonlítására, elemzésére is. A neuro-fuzzy rendszer nagy előnye, hogy nagyszámú paraméter vizsgálata esetén is elegendően gyors lehet ahhoz, hogy egy PC-n is valós időben szolgáltatssa az eredményeket.

Nagy számú paraméter együttes vizsgálata a hagyományos klíma modellek segítségével igen nagy számítási kapacitást igényel, amely csak elosztott rendszerek segítségével oldható meg. A párhuzamos feldolgozás segíti az ún. paraméter-tanulmányok kidolgozását is, ahol ugyanaz a modell kerül kiértékelésre akár több milliószor, a kezdeti paraméterek kis változtatásaival. Szintén a nagy számítási kapacitás teszi lehetővé, a hagyományos scenáriók elemzését olyan szélsőséges időjárási szituációkban, amelyeket éppen a klíma-változás modelljei jósolnak, de eddig nem kerültek be a modellek vizsgált paraméterei közé.

Az infrastrukturális alapot a Debreceni Egyetem Gazdasági- és Agrárinformatikai tanszékén kiépült PBS alapú klaszter képezné, amely a bekapcsolódási pontot szolgáltatja a NorduGrid nemzetközi grid rendszer felé. Az elosztott alkalmazások java nyelven íródnak, a párhuzamos feldolgozásukról a klaszterben egy JBoss alkalmazás szerver gondoskodik.

A genetikus algoritmus eszközei segítségével a lokális adaptációs stratégiák részegységeikre és elemeikre bonthatóak, majd meghatározható ezen elemek olyan kombinációja, amely optimális választ jelent a klímaváltozás hatásaira. A „populáció” egy egyede lehet például megadott N, P, K tartalommal rendelkező talaj (az ehhez szükséges műtrágyázás kiszámítható), talajtípus, öntözés mennyisége, stb.) és növekedési modell és szakértői rendszer segítségével állapítja meg a program, hogy ehhez az egyedhez mekkora termésátlag valószínűsíthető az adott scenárió esetén. Az elemek és részegységek kapcsolatát elemző szakértői rendszerek segítségével tehát kiértékelhetőek az egyes stratégiák „rátermettségi értékei”, sőt azok ráfordítási értékben is összehasonlíthatóak.

A kutatás egy alternatív szempontot adhat a klímaváltozás mezőgazdasági hatásainak vizsgálatához, valamint egy kiindulópontja lehet azoknak a modelleknek, amelyek nagyszámú klíma-paraméter figyelembevételével készülnek.

Irodalomjegyzék

- 1) Alyuda Market Forecasting System - <http://www.alyuda.com/>
- 2) Barsi Árpád (1998): Felszínborítottság térképezése neuro-fuzzy módszerrel GIS-környezetben (forrás: http://www.fomi.hu/internet/magyar/szaklap/1998/06/1998_06_2.htm)
- 3) Downing, Thomas E. et al. (2000): Climate Change Climatic Variability and Agriculture in Europe, An Integrated Assessment, Research Report No. 21., University of Oxford, Environmental Change Unit
- 4) Fischer, Günther et al. (2001): Global Agro-ecological Assessment for Agriculture in the 21st Century. IIASA and FAO
- 5) L. J. Fogel, A. J. Owens, and M. J. Whals (1966) *Artificial Intelligence through Simulated Evolution*. Wiley
- 6) Goldberg, David E.(1989): Genetic algorithms in search, optimization and machine learning. Addison-Wesley
- 7) Harnos Noémi (2003): A klímaváltozás hatásának szimulációs vizsgálata őszi búza termelésére. AGRO-21 Füzetek. 31. szám
- 8) Harnos Zsolt (2005): A klímaváltozás és lehetséges hatásai a világ mezőgazdaságára, Magyar Tudomány, 2005/7.
- 9) Herdon Miklós, Salga Péter (2005): AgriGrid – agrárinformatikai kutatás-fejlesztési igények és lehetőségek, Debreceni Egyetem, Gazdasági- és Agrárinformatikai Tanszék
- 10) Holland, J. H. (1973): Genetic algorithms and the optimal allocation of trials. SIAM Journal of Computing. 2(2)
- 11) William W. Hsieh and Benyang Tang (1998): Applying Neural Network Models to Prediction and Data Analysis in Meteorology and Oceanography. *Bulletin of the American Meteorological Society*: Vol. 79, No. 9
- 12) Jelasity Márk (1999): Genetikus algoritmusok. In Iván Futó, editor, *Mesterséges Intelligencia*, Aula Kiadó
- 13) Kevin Swingler (1996): Applying Neural Networks, Academic Press, San Diego
- 14) John R. Koza (1992): Genetic Programming: On the Programming of Computers by Means of Natural Selection. MIT Press
- 15) Vladimir M. Krasnopolsky, M. S. Fox-Rabinovitz and D. Chalikov (2005) : Decadal Climate Simulations Using Accurate and Fast Neural Network Emulations for the NCAR Community Atmospheric Model Radiation, 16th Conference on Climate Variability and Change
- 16) NorduGrid – <http://www.nordugrid.org>
- 17) I. Rechenberg (1973): Evolutionsstrategie: Optimierung Technischer Systemenach Prinzipien der Biologische Evolution. Fromman-Holzboog, Stuttgart

- 18) P. Salga, R. Szilágyi, T. Toth, A. Pokol (2004): Application of CNN-GRID technology, Information Systems for Agriculture, Forestry and Rural Areas - enlargement of EU 10th International Conference 2004, Seč u Chrudimi, Czech Republic
- 19) Salga Péter, Szilágyi Róbert, Herdon Miklós (2004): Genetikus algoritmus alkalmazása a mezőgazdasági termelés optimalizálásában, Agrárinformatikai kutatók-fejlesztők és felhasználók nyári egyetemi szekció, Gödöllő
- 20) P. Salga, M. Herdon, B. Kónya, R. Szilágyi (2005): Thinking in Grid - Initiatives for Hungarian AgriGrid System, EFITA/WCCA2005, Vila Real, Portugal
- 21) Zhang B. and R. S. Govindaraju (1998): "Using Modular Neural Networks to Predict Watershed Runoff," *ASCE Water Resources Engineering Conference '98*
- 22) A globális klímaváltozás: hazai hatások és válaszok (2006), KvVM-MTA "VAHAVA" projekt összefoglalása, Budapest
- 23) L. Hufnagel, M. Gaál (2005): Seasonal dynamic pattern analysis in service of climate change research (A methodical case-study, monitoring and simulation based on an aquatic insect community), *APPLIED ECOLOGY AND ENVIRONMENTAL RESEARCH* 3(1): 79-132.