

Grid technológia az allergiásokért

ÁDÁM RITA, BENCSIK ATTILA

Glia Számítástechnikai és Tanácsadó Kft.
{radam, abencsik}@glia.hu

Lektorált

Kulcsszavak: pollen, meteorológia, pollen-előrejelzés, meteorológiai adat, időjárás-előrejelzés, wise mind grid

A számítógépek kihasználatlanul hagyott kapacitását az időjárési viszonyok és a pollenkoncentráció kapcsolatának feltárására, az időjárás-előrejelzésen alapuló, minél pontosabb pollen-előrejelző rendszer kialakítására használjuk. A statisztikai összefüggések keresésekor nagy mennyiségű, több éves meteorológiai és pollen adatsorral dolgozunk, melyek közti bonyolult kapcsolatot adatbányászati módszerekkel kívánjuk felderíteni.

1. Bevezetés

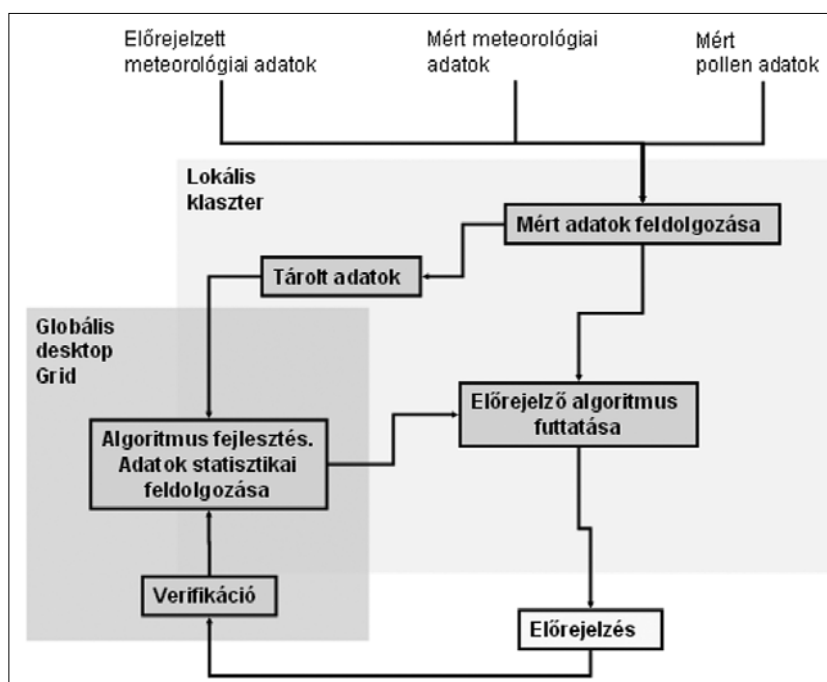
Számos olyan tudományos és műszaki probléma létezik, amelynek megoldására nem lehet a szokásos modellezési eljárásokkal választ találni a rendszer bonyolultsága miatt. A statisztikai analízis és a nagy számítási teljesítmény kombinációja viszont lecsökkentheti a kutatás időtartamát. A masszív kombinatorikának nevezett kutatási módszer használatával az elméletekből származó feltételezések kísérleti tesztelése nagyságrendekkel lerövidülhet. Az eljárás alkalmazása a szokásos számítógépes kapacitás mellett megoldhatatlan feladat. Ezt a problémakört igyekszik áthidalni a GRID technológia. A projektben az asztali számítógépek kihasználatlanul hagyott kapacitását az időjárési viszonyok és a pollenkoncentráció kapcsolatának feltárására, az időjárás-előrejelzésen alapuló, minél pontosabb pollen-előrejelző rendszer kialakítására használjuk. A statisztikai összefüggések keresésekor nagy mennyiségű, több éves meteorológiai és pollen adatsorral dolgozunk, melyek közti bonyolult kapcsolatot adatbányászati módszerekkel kívánjuk felderíteni. Az általunk kidolgozott módszer a GRID technológia új alkalmazását, az allergiás megbetegedések megelőzését tűzte ki célul. Ezt a procedúrát ismertetjük részletesen az alábbiakban.

2. A kutatás célja

A jelen cikkben ismertetett algoritmus-keresés során a meteorológiai változók (hőmérséklet, csapadék, relatív nedveség, légnyomás, szélesebesség) és a levegőben jelenlévő pollenkoncentráció közötti kapcsolatot keressük. A minél pontosabb összefüggések feltárásának célja az időjárás-előrejelzésen alapuló pollen-előrejelző rendszer kialakítása. Az algoritmus-keresés során külön vizsgál-

juk az egyes meteorológiai változók hatását a pollenkoncentráció alakulására, valamint ezeket súlyozzuk különböző időtávokkal. A meteorológiai viszonyok a pollen kibocsátó növény fejlődését befolyásolják, így a későbbi pollenkibocsátó képességére is hatással vannak. Természetesen más-más módon befolyásolhatja például a parlagfűvet a fejlődési fázisban kapott napfénytartam, az akkor fújó szél vagy a lehulló csapadék. Ezért az egyes változókról már kiinduláskor feltételezzük, hogy a múltra vonatkozóan különböző súlyt kapnak. A célunk megtalálni azokat az algoritmusokat, amelyek a meteorológiai körülmények figyelembe vételével (legyen szó a mai napról, vagy akár az elmúlt fél évről), a meteorológiai előrejelzésen alapulva minél pontosabban megbecslik a várható pollenkoncentrációt. Az eredményeinket a kutatás lezárultával egy új pollen-előrejelző rendszer formájában tervezzük hasznosítani.

1. ábra Az előrejelző rendszer összetevői



3. Econet Wise Mind Grid

A feladat megoldására az econet által kifejlesztett Wise Mind Desktop Grid rendszert választottuk. A Wise Mind egy ipari felhasználásra kifejlesztett magas biztonsági paraméterekkel rendelkező Desktop Grid rendszer. A WM Desktop Grid rendszert kialakításából fakadóan lehet lokális és globális szinten is alkalmazni.

A lokális szint azt jelenti, hogy a teljes Desktop Grid „házon” belül ugyanazon tűzfal mögött van, azaz lokális hálózattal vannak a Desktop Grid erőforrásai összekötve. A globális szint azt jelenti, hogy az erőforrások interneten keresztül kapcsolódnak egymáshoz. A kutatási programban az algoritmuskeresésre a rendszer globális kiépítését használjuk, a napi pollen előrejelzést pedig az econet irodai hálózatára telepített lokális rendszer segítségével végezzük.

4. A GRID szerepe

A statisztikai összefüggések keresésekor nagy mennyiségű, tíz évre vonatkozó meteorológiai és pollen mérési adatbázist dolgozunk fel. Az algoritmus keresésekor a számítási hatékonyság növelésére alkalmazzuk a GRID technológiát. A GRID rendszerek alkalmasak hatalmas mennyiségű adat tárolására, illetve feldolgozására. Mi ebből a feldolgozás részt használjuk ki.

A projekt megvalósításához egy weboldalt hozunk létre, ahol a leendő résztvevők regisztrálják magukat, ezután letölthetik a kapcsolódáshoz szükséges programot és a feldolgozandó adatbázisokat. Ezután az internetre csatlakozott számítógépek felhasználatlan kapacitását hasznosítjuk és algoritmusokat tesztelünk rajtuk. Ezek az algoritmusok keresik a meteorológiai elemek és a levegőben mérhető pollenkoncentráció közötti kapcsolatot.

Első közelítésben lineáris kapcsolatot feltételeztünk a változók között:

$$\sum a_i \cdot t + b_i \cdot u + c_i \cdot f \dots = \text{pollen koncentráció,}$$

ahol

t – a hőmérséklet,

u – a szélesség,

így vettük figyelembe a többi változót is

(relatív nedvesség, légnyomás, csapadék)

a, b, c, \dots stb. – a keresett együtthatók

(lehetnek függvények is),

i – az időlépcső

(ez a különböző változóknál lehet más és más).

Az általunk előállított becslés és a valódi, mért pollenkoncentráció közti eltérés mérésére, azaz a módszer hibájára az átlagos négyzetes hibát választottuk több szakirodalmi hivatkozás alapján:

$$\text{RMS} = 1/n(\sum (m_i - e_i)^2)^{1/2}$$

ahol m_i – az i -edik mérés,

e_i – az i -edik előrejelzés,

n – az adatok száma.

Ezzel a hibafüggvénnyel mértük a későbbi algoritmusaink hibáját is.

Az első algoritmus keresés során kiderült, hogy lineáris kapcsolatnál bonyolultabb összefüggést kell keresnünk a koncentráció és a meteorológiai elemek között. Második közelítésben olyan függvényeket tartalmazó nemlineáris kapcsolatokat vizsgáltunk, amelyek a változók közötti bonyolult kapcsolatot feltételezéseink szerint reprezentálhatják.

Ilyen függvényekre példa:

$$\text{const}_1 \cdot (\ln(x))^{\text{const}_2}$$

$$\ln(\text{const}_1 + \text{const}_2 \cdot x + \text{const}_3 \cdot x^{\text{const}_3})$$

$$(\text{const}_1 \cdot \exp(x))$$

$$\text{const} \cdot \exp(x), \text{const}_1 \cdot \exp(\text{const}_2 \cdot x)$$

ahol

x – bármely változó, a konstans, pozitív egész szám.

A fenti függvények csak példák, hiszen külön algoritmusokat próbálunk ki mind az egyes meteorológiai elemekre, mind az év adott pollenterhelési időszakaira vonatkozóan.

Az algoritmusok kidolgozására egy vezérlő szerver osztja ki a feladatokat és értelmezi az egyes számítógépek eredményeit. A tesztek eredményét a résztvevők a weboldalon keresztül folyamatosan nyomon követhetik. A weboldal arra is szolgál, hogy a kutatás eredményeit folyamatosan publikáljuk, így az bárki számára hozzáférhető lesz.

A programban abban az esetben is részt lehet venni, ha nem rendelkezünk folyamatos internet kapcsolattal (ami manapság már viszonylag ritka). Ekkor az internetre való fellépéskor a GRID-hez szükséges kliens program feltölti az elvégzett számítások eredményeit a vezérlő szerverre. Állandó internetkapcsolat esetén az adatáramlás a vezérlő szerverre valódi időben megvalósul.

Miután a kapcsolatot feltártuk, a meteorológiai előrejelzésen alapuló pollen-előrejelzést tudunk készíteni oly módon, hogy az összefüggést alkalmazzuk a mért adatokra, emellett az időjárás-előrejelzést is figyelembe vesszük.

A fent leírt algoritmus-keresést először a parlagfűre, később a fennmaradó 34, hazánkban előforduló pollenre szeretnénk végrehajtani. Az így létrejövő rendszer a teljes országot, a teljes allergén időszakot és az összes hazai pollent lefedő előrejelzést szolgáltatna.

A projekt egy nagyobb lélegzetű hazai együttműködés, a 2005-ben elkezdődött HaGrid projekt [1] egyik részfeladatáént fut. A HaGrid az econet.hu Informatikai Zrt, az Országos Meteorológiai Szolgálat, a MTA Számítástechnikai és Automatizálási Kutató Intézet, a Glia Számítástechnikai és Tanácsadó Kft., a DDC Automatika Kereskedelmi és Szolgáltató Kft. és az Env-in-Cent Környezetvédelmi Tanácsadó Iroda együttműködésében valósul meg. A program célja, hogy a Desktop GRID technológiát olyan formára dolgozza át, hogy az alkalmas legyen mind zárt, vállalati környezeten belüli, mind azon kívüli feladatok megoldására.

A megvalósításhoz a résztvevők három központot (SZTAKI, OMSZ, econet) hoznak létre, amely a rendszer

alapját alkotja. A GRID rendszert ezután különböző célú alkalmazásokkal tesztelik, ezek egyike a pollen-előrejelző rendszer. Ezzel a projekt egy jövőbeni, hazai szolgáltató rendszer alapjait igyekszik megteremteni, amelyben a PC tulajdonosok becsatlakozhatnak a (vállalati és egyetemi) kutatás-fejlesztésbe.

A kutatás témájához kapcsolódik még a SZTAKI, a BME, az ELTE és a Compaq együttműködésével létrejött SzuperGrid [2] projekt (2002-2003), amelynek keretében a SZTAKI adaptálta a P-GRADE programfejlesztő rendszert a gridhez [3]. Ennek segítségével intézményen belül a különböző, heterogén informatikai erőforrásokat összekapcsolták. A P-GRADE egyik jelentős alkalmazási területe az Országos Meteorológiai Szolgálatnál futtatott ultrarövid távú előrejelzés informatikai hátterének biztosítása volt.

A grid technológiát a meteorológia több területén alkalmazzák, ezek közül napjainkban kiemelkedően fontos a klímakutatás. A különböző meteorológiai-éghajlati modellek a legnagyobb kapacitású számítógépeket igénylik szerte a világon. A klímakutatásban világviszonylatban élen járó Hadley Centre [4], a brit meteorológiai szolgálat, klímakutató intézete több egyetemmel karöltve grid rendszerre alapozva a században várható éghajlatváltozásokat próbálja minél pontosabban előrejelezni [5].

5. Összefoglalás

A fent részletezett módszer a GRID rendszerek újfajta alkalmazását mutatja be. A programhoz csatlakozó résztvevők számítógépein olyan algoritmusokat kívánunk futtatni, amelyek nagy méretű, több évre vonatkozó meteorológiai és pollen-adatbázisok feldolgozásával megkeresik a köztük fellelhető összefüggéseket és segítenek a pollenterhelés előrejelzésének megvalósításában.

Irodalom

- [1] Hungarian Advanced Grid
<http://hagrid.hu>
- [2] Magyar Szuperszámítógép Grid
<http://mszgrid.iif.hu/>
- [3] Kacsuk, P.:
A magyar grid rendszerek és fejlesztési irányai. VIII., Országos (Centenárium) Neumann Kongresszus, Budapest, 2003.
- [4] Hadley Centre for Climate Prediction and Research
<http://www.metoffice.gov.uk/research/hadleycentre/>
- [5] Climate prediction project
<http://www.climateprediction.net/>

Hírek

A Cisco újgenerációs Unified Wireless Network egységes, vezeték nélküli hálózati megoldásának része a Cisco Aironet 1250 sorozatú hozzáférési pont, a világ első vállalati szintű 11n szabványnak megfelelő készüléke, valamint a 48 Gbit sebességű, kiválóan méretezhető Cisco Catalyst 6500 alapú WLAN-vezérlőrendszer, illetve a Unified Wireless Network 4.2-es változatában debütáló vezetékes és vezeték nélküli szolgáltatások. A Catalyst kapcsolócsalád egyetlen Ethernet-portról is megoldja a két rádiósávú Aironet 1250 sorozatú hozzáférési pontok áramellátását.

A Cisco új Aironet 1250-je a világ első Wi-Fi tanúsítvánnyal rendelkező **802.11n 2.0**-s szabványtervezetét követő hozzáférési pontja, és az egyetlen olyan kereskedelmi forgalomban kapható termék, amely része volt a Wi-Fi Alliance 802.11n 2.0 tervezet tesztelési környezetének, amellyel a későbbiekben az összes egyéb terméket hitelesíteni fogják a Wi-Fi-együttműködés szempontjából. A Cisco a rádiófrekvenciás kommunikációban szerzett tapasztalatát felhasználva ötszörös teljesítménynövekedést ért el a többcsatornás bemenettel és kimenettel rendelkező, úgynevezett MIMO-technológia terén. Ezzel sokkal megbízhatóbbá tehető a vezeték nélküli lefedettség, ami különösen a rádiófrekvenciás felhasználás szempontjából problémás területeken jelent nagy előnyt, mint például az egészségügy, a felsőoktatás, a raktározás, a logisztika és a gyáripár.

A Cisco 802.11n technológiával felvértezett, újgenerációs vezeték nélküli termékei a következőket nyújtják:

- A Catalyst 6500-hoz tartozó, a 802.11n technológia fokozatos, vagy azonnali, nagy volumenű bevezetését egyaránt támogató Cisco WiSM szolgáltatásmodul révén a felhasználók további vezérlők hozzáadásával bővíthetik a kapacitást igényeik szerint. Rugalmas méretezhetőségével a teljes szervezetre kiterjedő WLAN-bevezetések is gyorsan megoldhatók.

- A Cisco az új Unified Wireless Network Software nevű programjának 4.2-es verziója kibővített mobil szolgáltatásokat biztosít. Idetartozik a vállalati "mesh" alapú WLAN-hálózat, a hozzáférési pontok figyelése és a migrációs segédletek, valamint a vendégfelhasználók számára biztosított egységes vezetékes és vezeték nélküli hozzáférés. Emellett ezek a szolgáltatások fejlett vezeték nélküli hangátviteli funkciókkal, a távoli elérési pontok hibátűrő, átfedő elhelyezésével és integrált spektrumanalízissel rendelkeznek.

- A továbbfejlesztett kliensprogramok, így a Cisco Secure Services Client 5.0 által is kínált új lehetőségek az egyszerűbb vállalati szintű, egyetlen kliensbiztonsági és felügyeleti keretrendszer kialakítását szolgálják.