



Növénytermesztési tartamkísérletek vizsgálata földközeli és légi távérzékelési technológiával

Burai P., Lénárt Cs.

Debreceni Egyetem, Mezőgazdaságtudományi Kar, Víz- és Környezetgazdálkodási Tanszék, 4032 Debrecen, Böszörményi út 138.

ÖSSZEFOGLALÁS

A tanulmányunk célja kis táblaméretű tartamkísérletek távérzékeléssel történő vizsgálata. Két alföldi gazdaságban végeztünk vizsgálatokat, ahol kisparcellás tartamkísérleteket vannak beállítva. A távérzékelési felvételek elemzése során számba vettük a leggyakrabban használt távérzékelési adatforrásokat, ezek használhatóságát a kis parcellaméretű esetekben. Kutatásunk során TETRACAM ADC multispektrális (3 csatornás) kamerával földközeli (0,5-3m), illetve alacsony repülési (300-500m) magasságból készítettünk felvételeket. Megvizsgáltuk a kamerával készített felvételek geometriai pontosságát és elvégeztük a felvételek geometriai transzformációját. Az alacsony repülési magasságból (300-500m) készített felvételekkel nagy térbeli felbontású képeket készítettünk (0,4-0,6m), amelyek kalibrációját földi azonosító pontok használatával végeztük el. A terepi mintavételek során borítás és biomassza vizsgálatokat végeztünk, amelynek pontszerű adatait összehasonlítottuk a felvételekkel. A multispektrális elemzés alkalmával növényi vegetációs indexet (NDVI) számítottunk, amit a terepi mintavétel során mért levélfelület indexsel (LAI), szármagasság, szemtömeggel hasonlítottunk össze.

(Kulcsszavak: távérzékelés, GIS, fotogrammetria, multispektrális elemzés, NDVI)

ABSTRACT

Evaluate small plot-sized long period researches by airborne and near-field remote sensing

P. Burai, Cs. Lénárt

University of Debrecen, Centre of Agricultural Sciences, Department of Water and Environmental Management
H-4032 Debrecen, Böszörményi út 138.

The aim of the study was the evaluation of small plot-sized long-period researches by remote sensing. We examined two sites in the Hungarian Great Plain, where small parceled long-period researches are being set. By the analysis of remotely sensed images we defined the most often used remote sensed data sources, the usability of them in the case of the small parcel-sizes. During our research, we took near-field (0,5-3 m) photos in low flying height (300-500 m) by TETRACAM ADC multispectral (3 channels) camera. We analyzed the geometric accuracy of the images taken by the camera and made the geometric transformation of these images. We made high resolution images (0,4-0,6 m) from the photos, which were taken from low flying height (300-500 m). We calibrated them by using ground control points. We made overlay and biomass researches during the terrain sampling, and we compared the pointwise data of them with the images. We calculated NDVI during the multispectral analysis and we compared it with the Leaf Area Index (LAI), stem height and grain mass, which have been measured during the sampling.

(Keywords: remote sensing, GIS, photogrammetry, multispectral analysis, NDVI)

BEVEZETÉS

Növénytermesztési tartamkísérletek során különböző fajtákkal és kezelésekkel (tápanyag, növényvédelem, talajművelés, stb.) végeznek kis méretű parcellákban kísérleteket, ellenőrzött körülmények között. A vegetációs időszakban szükség lehet különböző mérésekre, amely a hagyományos mintavételi eszközökkel nagy munkaigényű, nehezen kivitelezhető. Eredményes vizsgálatokat végeztek különböző távérzékelő eszközökkel, a gyomosodottság (Tamás, 2005) és a növényi betegségek okozta elváltozások (Laudien *et al.*, 2003) precíziós mérésére. Napjainkban a távérzékelte felvételek felhasználásának erőteljes növekedése figyelhető meg. A távérzékelte felvételek elemzésével gyorsabban és olcsóbban nyerhetünk adatokat, mint a földi mintavétel során (Sabins, 1996). A távérzékelési eredmények megbízhatóságát elsősorban a térbeli és a spektrális felbontás határozza meg (Condit, 1970), azonban a különböző adatok felhasználhatóságát a felvételek nagysága, hozzáférhetősége, visszatérési idő nagyban befolyásolja. A távérzékelte adatok elemzésével a talajról és a növényzetről egyaránt nyerhetünk információt (Zilinyi, 2000), amelyek nagy előnye, hogy nem csak pontszerű adatot kapunk, hanem az állományon belüli eltérések térben is értékelhetőek (Seelan *et al.*, 2003). A tanulmányunkban kis táblaméretű tartamkísérletek távérzékeléssel történő vizsgálatát végezzük el TETRACAM ADC digitális multispektrális kamerával. A tanulmány első részében a kamera geometriai és radiometriai vizsgálatát mutatjuk be, majd a terepi mintavételek értékelését ismertetjük.

Ma már különböző térbeli felbontású és spektrális összetételű távérzékelte adatforrásokat használhatunk a mezőgazdasági és környezetvédelmi célú vizsgálatokhoz (Mucsi, 2004). Az 1. táblázat azoknak az adatforrásoknak az alapadatait tartalmazza, amelyeket a Debreceni Egyetem Víz- és Környezetgazdálkodási Tanszéke mezőgazdasági és környezetvédelmi célú elemzésre alkalmaz.

1. táblázat

A kutatás során alkalmazott távérzékelte adatforrások tulajdonságai

Távérzékelte felvétel (1)	Képméret (2)	Terepi felbontás (3)	Csatornák száma (4)	Csatornák (μm) (5)	Felvétel készítés időpontja (6)
LANDSAT 7 ETM+	170×183 km	MS: 30m TIR: 60m PAN: 15m	7	1.-5.: 0,45-1,75 7.: 2,08-2,35 6.: 10,40-12,50 PAN: 0,52-0,90	16 naponta (7)
SPOT 5	60×60 km	SWIR: 20m MS: 10m PAN: 5/2,5 m	4	MS: 1. 0,49-0,79; 2. 0,61-0,89; 3. 0,61-1,58; 4. 0,68-1,75; PAN: 0,49-0,69	-
DAIS - 7915	3990 m × 8735 m*	5m*	79	1.-32.: 0,40-1,00 (Si)** 33.-40.: 1,50-1,80 (InSb)** 41.-72.: 2,00 – 2,50 (InSb)** 73.: 3,00-5,00 (InSb)** 74.-79.: 8,00-12,60 (MCT)**	az időjárási viszonyoktól függ (8)
TETRACAM ADC	420×500 m***	0,5m	3	1.-4.: 0,45-0,89	az időjárási viszonyoktól függ

*2000m-es repülési magasságnál (2000m aircraft altitude); **DAIS szenzortípusok (DAIS sensors); ***300m-es repülési magasságnál (300m aircraft altitude);

Table 1: Technical information of the applied images

Image type(1), Image size(2), Pixel size(3), Bands number(4), Bandwith and sensors(5), Revisit time(6), 16days(7), Depending on the weather(8)

Az 1. táblázatban szereplő felvételek közül a Landsat és Spot adatok jól alkalmazhatók a nagyobb méretű területek idősoros vizsgálatában (Sabins, 1996), azonban a kis méretű táblák elemzése a nagy pixelméret (10-30m) miatt, még a szubpixeles elemzéssel sem ad megbízható eredményt (1. ábra). 2002-ben a Hysens program keretében alföldi mintaterületeken is készítettek DAIS 7915 hiperspektrális felvételeket, amelyek az 5m-es terepi felbontásnak és a nagy csatornaszámnak (79) köszönhetően jól alkalmazható földhasználat osztályozás és kvalitatív vizsgálatokra, például a talajdegradációs folyamatok vizsgálatában (Pechmann et al., 2003).

1. ábra

Különböző parcellaméret és távérzékelte felvétel pixelméreteinek összehasonlítása

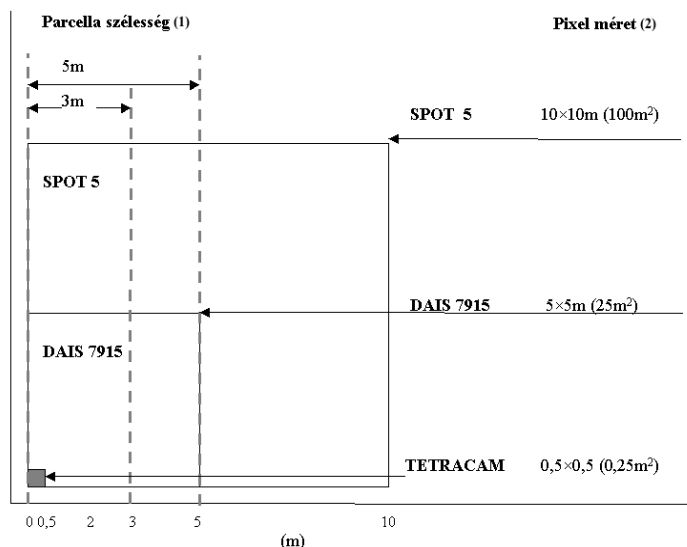


Figure 1: Comparing different parcell sizes and pixel sizes of remote sensing data

Parcell size(1), Pixel size(2)

ANYAG ÉS MÓDSZER

TETRACAM ADC digitális multispektrális kamera és az optika technikai paraméterei

- csatornák: zöld (520-600nm), vörös (620-750nm), közeli infra (750-950 nm)
- digitális felbontás: 1280×1020pixel = 1,3 Megapixel (Motorola CMOS)
- optika: 8,5mm; 1:1,3, 2/3'

A kamerában használt CMOS szenzor működését tekintve a széles körben elterjedt CCD (Charged Coupled Device) szenzorokhoz hasonlít, azonban alacsonyabb ára, kisebb energai igénye, alacsony zajszintje, a felvételek előnyös geometriai és radiometriai tulajdonságai miatt, egyre inkább teret nyer a digitális képalkotás technológiában (Samadzadegan et al., 2004). A kamera kalibrációját a helyszínen kell elvégezni. Egy fehér színű etalon lapról kell referencia felvételt készíteni, amelyet később a képek kalibrációhoz használunk.

2. ábra

TETRACAM ADC kamera



Figure 2: TETRACAM ADC digital camera

Adatok feldolgozása

A digitális kamerával készített felvételeket, a készülék szoftverével (Briv 32) TIFF formátumban tudjuk az adatokat letölteni. A felvételek feldolgozására Briv32, ENVI és ArcGIS térinformatikai szoftvereket használtunk

Mintaterület, mintavétel

A mintaterületeket a Hajdúböszörmény mellett fekvő Tedej Rt. tulajdonában lévő 1500 ha-os Pusztaság területen és a Debreceni Egyetem Látóképi Növénytermesztési Kísérleti telepén jelöltük ki.

A mintaterületeken a vizsgált parcelláknak a szélessége: 1,5-3 - 5-30 m, azonban a 3 és az 5 méter széles parcellák a legjellemzőbbek. A háttéradatok mérését a felvételek készítésének időpontjába végeztük el. A földközeli mérések esetében a 1×1m-es mintavételi keretet használtunk. A légi felvételek értékeléséhez a háttéradatok meghatározásához pontszerű mintavételt alkalmaztunk. A mintavételi helyek pozícióját precíziós DGPS méréssel határoztuk meg. A teszt felvételeket cukorrépa, kukorica és búza állományokról készítettük.

EREDMÉNY ÉS ÉRTÉKELÉS

A felvételek geometriai értékelése

A felvételek lineáris torzítási viszonyainak vizsgálatára 11×13 pontból álló (5cm x,y távolságú) alaphálót használtunk. Az alapháló szomszédos pontjai között távolságmérést alkalmaztunk. Megmértük az x, y és átlós irányú (c) távolságokat, majd összehasonlítottuk a valódi x, y, c (standard) távolságokkal (2. táblázat)

A legnagyobb eltéréseket a standard távolságtól ($x_{\max}-x_s$), ($y_{\max}-y_s$), ($c_{\max}-c_s$) a képszeleken mértük (3. táblázat)

Jelentős eltéréseket tapasztaltunk a kép szélső területein, ezért a felvételek értékelésénél feltétlenül szükséges a geometriai korrekció. A légi felvételek elemzése esetében földi azonosító pontokat alkalmaztunk (Ground Control Point, GCP) a geometriai korrekció során. Referencia értéknek földi illesztőpontok EOY koordinátái adtuk meg, amelyek pozícióját precíziós (cm-pontos) GPS méréssel mértünk a terepen.

Mivel a vizsgált parcellák sík területen találhatók (a képen maximális szintkülönbség 3m), jelentős domborzati hatással nem kellett számolni, a transzformáció során elsőrendű polinominális korrekciót használtunk. A polinominális transzformáció könnyen alkalmazható, kis számítási igényű és a sík mintaterület esetében hatékonyan alkalmazható (Jensen, 1996).

2. táblázat

A vizsgált felvétel torzításának statisztikai adatai

Távolság (1)	Standard érték (2)	Szórás (3)	CV % (4)
X	1,13	0,064	5,67
Y	1,13	0,045	4,05
C	1,59	0,076	6,75

Table 2: The distortion statistic of the analysed image

Directon(1), Standard value(2), Standard deviation(3), Variation coefficient(4)

3. táblázat

A vizsgált felvétel maximális torzítása x,y,z irányba

Eltérés (1)	$\Delta t_{\max}$	$\Delta t_{\max\%}$
$(x_{\max}-x_s)$	0,13	11,50
$(y_{\max}-y_s)$	0,09	7,96
$(c_{\max}-c_s)$	0,49	30,81

Table 3: The maximal distortion of x,y,z directon

Deviation(1)

A földi illesztőponok és a felvétel közötti geometriai különbséget az X és Y értékekből számított rmse (root mean square error) értékkel adhatjuk meg, ahol:

$$\text{rmse} = [(X' - X)^2 + (Y' - Y)^2]^{1/2} \quad (1)$$

ahol: X,Y: földi azonosító pont (GCP) EOVS koordináta

X',Y': transzformált kép EOVS koordináta

A geometriai transzformáció után a légifelvételek rmse = 1,1-2 között mozgott. A hiba javítását a földi referenciapontok növelésével érték el. A földközeli mintavételeknél a szabályos mintavételi keret sarokpontjainak a koordinátáit kell megadni, illetve a kép szabályos kivágásával lehet a térbeli torzítást javítani. A felvételeket lehetőség szerint állványról, a felszínre merőlegesen kell elkészíteni.

Radiometriai és statisztikai vizsgálatok

Radiometriai és statisztikai elemzésre a tedeji 200×150m-es mintaterületről 300m-es repülési magasságból 2004. május 17.-én készített légi felvételt alkalmaztunk. A felvétel egy olyan búzatábláról készült, ahol homogén állományt (azonos fajta, sortáv, ezermagtömeg, kezelés) termesztettek különböző sőtartalmú területen. A sőtartalom a mintaterület É-D tengelyében csökken, amelyet a növényállomány biomasszaváltozása jelez (3. ábra).

3. ábra

**A közeli infra(1), zöld(2) vörös(3) és a csatornák reflektancia értékei
a mintaterület É-D tengelyében**

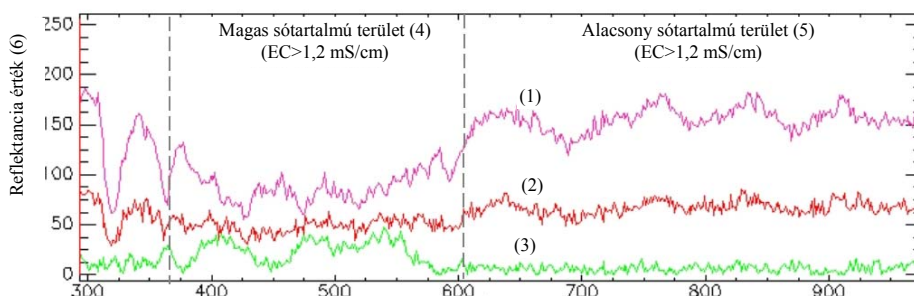


Figure3. The reflectance value of the investigated area of the vertical profile

Near infrared(1), Green (2), Red band(3), High EC area(4), Low EC area(5), The values of the reflectance(6)

A képfeldolgozás során megvizsgáltuk a légi felvétel általános statisztikai paramétereit

4. ábra

A vörös(a) zöld(b) és közeli infra(c) csatornák hisztogramjai

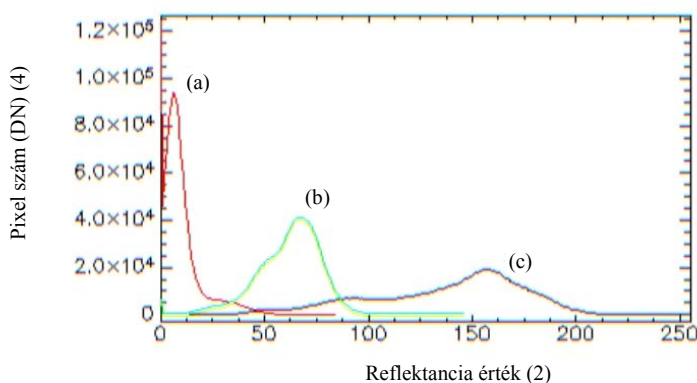


Figure 4. Histogram of the values of the red (a), green (b) and near infrared (c) bands

Pixel number(1), Reflectance value(2)

A csatornák hisztogramjainak elemzése során a vörös csatornánál erőteljes asszimetria figyelhető meg, a hisztogram lefutása a minimumértékek irányába nem egyenletes (4. ábra). A közeli infra tartomány erőteljes szórást mutat, amit a heterogén növényállomány reflektancia-értékeinek változása okoz. A korrelációs mátrix az egyes csatornák közötti kapcsolatot mutatja meg. Szoros kapcsolat van a felvétel reflektanciáértékeinek vörös és közeli infra csatornáik között (4. táblázat).

4. táblázat

A vizsgált felvétel csatornái között számított kovariancia és korrelációs mátrix

Csatornák (1)	Min	Max	Átlag (2)	Szórás (3)
1	0	255	137,209	37,912
2	0	84	9,872	9,649
3	0	146	61,878	15,018

Kovariancia mátrix (3)				Korrelációs mátrix (4)			
Csatornák	1	2	3	Csatornák	1	2	3
1	1437,325	-177,479	499,971	1	1	-0,48514	0,878083
2	-177,479	93,114	-36,781	2	-0,48514	1	-0,2538
3	499,971	-36,781	225,561	3	0,878083	-0,2538	1

Table 4. Covariation and correlation matrix between the bands of the analysed image

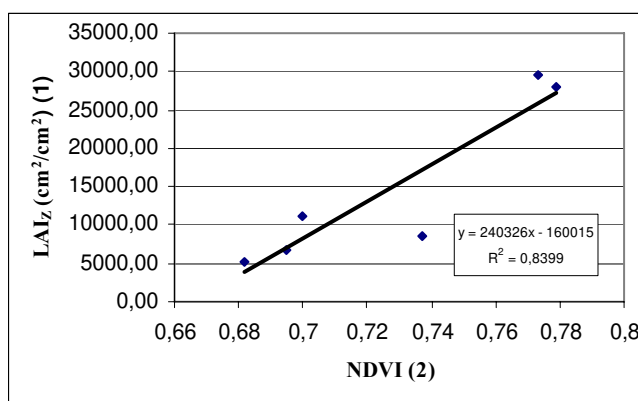
Bands(1), Average(2), Standard deviation(3), Covariation matrix(4), Correlation matrix(5)

A TETRACAM ADC kamera földközeli alkalmazásai:

A biomassza vizsgálatra a közeli infra és a vörös csatornák reflektancia értékeiből számított Normalizált Vegetációs Indexet (NDVI) használtuk. Az vegetációs index NDVI használata széles körben elterjedt a vegetáció elemzésben, földhasználat értékelésben (David *et al.*, 1999; Running *et al.*, 1995). Az NDVI index értékeiből a levélfelület index (LAI) is nagy biztonsággal meghatározható (Jordan *et al.*, 1969).

A Látóképi Tangazdaság területén kukorica hibridek levélfelület indexét (LAI) mértük és hasonlítottuk össze a kb. 3 m magasról készített felvételekből számított NDVI-vel. A levélfelület mérésénél a zöld, fotoszintetikusan aktív levélfelületet mértük (LAI_Z). Az NDVI és LAI között szoros korrelációt számítottunk, majd elvégeztük a lineáris függvény illesztését (5. ábra, 5. táblázat).

5. ábra

NDVI – LAI_Z korreláció kukorica hibridek vizsgálatánál (n=6)Figure 5. Correlation between NDVI and LAI_Z (n=6)

Green Leaf Area Index(1), Normalized Difference Vegetation Index(2)

5. táblázat

A vizsgált kukorica hibridek LAI és NDVI adatai

Hibridek (1)	LAI (cm ² /cm ²) (2)	NDVI (3)
Káma	29441.66	0.773
Vilma	27998.14	0.779
Madona	11199.01	0.7
Debreceni 377SC	8668.423	0.737
Uranus	6806.499	0.695
Szegedi SC352	5056.934	0.682

Table 5. LAI and NDVI data of the analysed corn hybrids

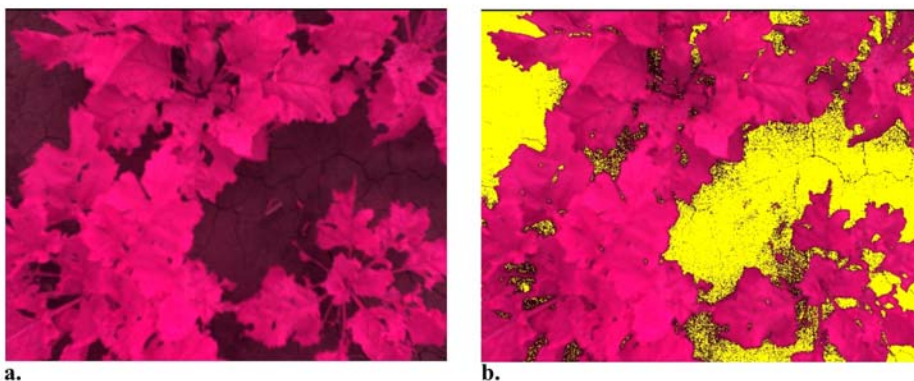
Hybrids(1), Green Leaf Area Index(2), Normalized Difference Vegetation Index(3)

Cukorrépa állományban a tényleges borítást vizsgáltuk (canopy %), amely számítását a Briv 32 szoftver a talaj színeképének megadása után automatikusan kiszámít. (6. ábra)

Az eljárás előnye a gyors és pontos borítás vizsgálat (a pontos terület kiszámítása), továbbá lehetőség van az árnyékkompenzáció alkalmazására.

6. ábra

Cukorrépa földközeli hamisszínes felvétele (a) és a kiszámított felszínborítás (b)



r: közeli infra, g: vörös, b: zöld csatorna (r: near-infrared g: red b: green) (canopy=74,8%)

Figure 6. False color (a) and canopy (b) ground closed images of sugar beet are taken by TETRACAM ADC

A TETRACAM ADC kamera alkalmazása légi felvételek készítésére:

Búza állomány vizsgálata tedeji mintaterületen:

A légifelvételtől számított NDVI értékeket összehasonlítottuk a 2004. júliusi terepi mintavétel (n=11) szemtermés (g/m²) értékével, ahol szoros korrelációt mutattunk ki (R²=0,90) (7. ábra, 6. táblázat).

7. ábra

A különböző búza állományról készített normalizált vegetációs index (NDVI) és szemtermés (g/m^2) közötti összefüggés ($n=11$)

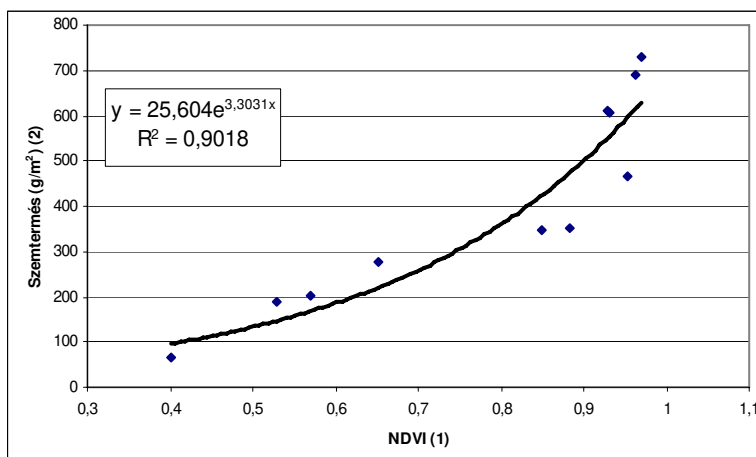


Figure 7: Correlation between NDVI and yield (g/m^2) in winther wheat ($n=11$)

Normalized Difference Vegetation Index(1), Yield (2)

6. táblázat

A minták szemtermés és NDVI adatai

Minták (1)	Szemtermés (g/m^2) (2)	NDVI
1	353.5	0.883
2	466.55	0.952
3	275.1	0.651
4	201.25	0.571
5	606.9	0.932
6	689.15	0.962
7	188.65	0.53
8	67.2	0.401
9	347.55	0.85
10	610.75	0.93
11	728.35	0.97

Table 6: The data of the samples of the yields and NDVI

Samples(1), Yield(2), Normalized Difference Vegetation Index(3),

Az $25,604e^{3,3031x}$ függvény x tagja helyére behelyettesítettük a mintaterületből számított NDVI képet, amelynek pixel-értékéből kiszámítható a mintaterület termésterképe.

KÖVETKEZTETÉSEK

A növénytermesztési tartamkísérletek kis parcellamérete (3-5m széles parcellák) a nagy felbontású távérzékelt felvételek használatát indokolják. Tanulmányunkban TETRACAM ADC digitális kamerával készített felvételeket vizsgáltunk föld közeli és légi felvételezésre. Alacsony repülési magasságból (300-500m) a felvételek 0,4-0,6m geometriai felbontásúak. A földközeli és légi felvételeket vizsgálva megállapítható, hogy a kamerába beépített nagy látószögű optika jelentős torzítást okoz, amelyet geometriai transzformációval tudunk csökkenteni. A légi felvételek esetében a geometriai korrekció során célszerű növelni a földi azonosító-pontok számát. A geometriai torzítás miatt a földközeli felvételeknél is célszerű elvégezni a transzformációt. A felvételek radiometriai tulajdonságait vizsgálva a pixel értékek minden csatorna esetében dinamikusan változtak, a vegetáció reflexiók tulajdonságainak megfelelően.

A TETRACAM ADC kamera jól alkalmazható a pontos növényi borítás, biomassza vizsgálatban, elsősorban a közeli infra és a vörös csatornák reflektancia értékeiből számított NDVI index használatával. A szántóföldi kultúrákban végzett terepi mérések eredményeit összehasonlítva a számított NDVI értékekkel szoros korreláció számítottunk a biomassza, a levélfelület index (LAI), szármagasság, és a szentömeg között. A nagy felbontású felvételek lehetőséget adnak az állományon belüli kis kiterjedésű folyamatok térbel vizsgálatára. A kamera alkalmazásának nagy előnye a terepi alkalmazhatóság, könnyű kezelhetőség, gyors adatnyerés. A nagy felbontású légi felvételek esetében a kis képméret az optika kedvezőtlen geometriai paraméterei miatt az eszköz csak kisebb területek felvételezésére alkalmas.

IRODALOM

- Condit, L. (1970). The spectral reflectance of American soils. In: Photogrammetric Engineering and Remote Sensing, 36. 955-966.
- David P.T., Warren B.C., Robert E. K., Karin S.F., John M.B. (1999). Relationships between Leaf Area Index and Landsat TM Spectral Vegetation Indices across Three Temperate Zones Sites. In: Remote Sensing of Environment, 70. 52-68.
- Janseen, L.L.F., Jaarsma, M.N., van der Linden, E.T.M. (1990). Integrating Topographic data with Remote Sensing for Land-Cover Classification. In: Photogrammetric Engineering & Remote Sensing, 56. 11. 1503-1506.
- Jensen, J.R. (1996). Introductory digital image processing: a remote sensing perspective. Second edition. Prentice-Hall, Upper Sadle River, New Jersey
- Jordan, C.F. (1969). Deviation of leaf-area index from quality of light on the forest floor. In: Ecology 50. 663-666.
- Laudien, R., Bareth, G., Doluschitz, R. (2003). Analysis of hyperspectral field data for detection of sugar beet diseases. In: Proceedings of EFITA 2003 Conference, Debrecen.
- Mucsi, L. (2004). Műholdas távérzékelés. Libellus kiadó : Szeged
- Pechmann I., Tóth T., Tamás J., Kardeván P., Róth L., Burai P., Katona Zs. (2003). Eltérő talajszótartalmú növényzeti foltok elkülönítése hiperspektrális technológiával (Földminősítés és földhasználati információ, Keszthely)
- Running, S.W., Loveland, T.R., Pierce, L.L., Nemani, R.R., Hunt Jr., E.R., (1995). A remote sensing classification logic for global land cover analysis. In: Remote Sensing of Environment, 51. 39-48.

- Sabins, F.F. (1996). Remote Sensing. Principles and Interpretation. W.H. Freeman and Co. : Los Angeles
- Samadzadegan, F., Hahn, M., Sarpulaki, M., Mostofi, N. (2004). Geometric and radiometric evaluation of the potential of a high resolution CMOS – camera. In: Proceedings of ISPRS XXth Congress, Istanbul
- Seelan, S.K., Laguet, S., Casady, G.M., Seielstad, G.A. (2003). Remote sensing application for precision agriculture: A learning community approach. In: Remote Sensing of Environment, 88. 157-169.
- Tamás, J. (2005). Széles spektrumú kézi kamera alkalmazhatósága a terepi gyomfelvelelésekre. In: Növényvédelem 41. 2. 53-59.
- Turner, D.P., Cohen, W.B., Kennedy, R.E., Fassnacht, K.S., Briggs, J.M., (1999). Relationships between Leaf Area Index and Landsat TM Spectral Vegetation Indices across three temperate zone sites. In: Remote Sensing of Environment, 70. 52-68.
- Zilinyi V. (1995). Természetes felszínek spektrális reflexiós tulajdonságai, és hasznosításuk az optikai távérzékelés interpretációjában. Egyetemi Doktori Értekezés, DATE, MTK, Debrecen

Levelezési cím (*Corresponding author*):

Burai Péter

Debreceni Egyetem, Mezőgazdaságtudományi Kar

Víz- és Környezetgazdálkodási Tanszék

4032 Debrecen, Böszörményi út 138.

University of Debrecen, Centre of Agricultural Sciences

Department of Water and Environmental Management

H-4032 Debrecen Böszörményi út 138.

Tel.: 36-52-508-444

e-mail: pburai@gissserver1.date.hu



Cukorgyári termeltetési rendszer térinformatikai értékelése

Nagy I., Takács P.

Debreceni Egyetem Agrártudományi Centrum, Víz- és Környezetgazdálkodási Tanszék, 4032 Debrecen, Böszörményi út 138.

ÖSSZEFOGLALÁS

A technológiák fejlődésével (GIS, GPS, precíziós gazdálkodás) egyre több nagyobb méretű mezőgazdasági vállalkozás tárolja és elemzi termelési adatait számítógépes rendszerek segítségével. A jellemzően térinformatikai (GIS) rendszerekben tárolt adatok/adatbázisok előnyei sokrétűek: a földrajzi koordinátákhoz köthető adatok megjelenítése és kezelése; sokrétű adatformátumok (digitális talaj és kataszteri térképek, távérzékelte felvételek, mérési adatok, stb.) egyidejű kezelése; korszerű geostatistikai elemzések elvégezhetősége. Cikkünkben az Eastern Sugar Rt. kabai cukorgyárának termeltetési rendszeréhez kapcsolódó komplex térinformatikai adatbázis kialakításának első lépéseit ismertetjük. (Kulcsszavak: cukorgyári termesztési rendszer, térinformatikai adatbázis, adatintegritás)

ABSTRACT

GIS based evaluation of the production system of a sugar factory

I. Nagy, P. Takács

University of Debrecen, Centre of Agricultural Sciences, Faculty of Agriculture, Department of Water- and Environmental Management, Debrecen, H-4032, Böszörményi street 138.

Technological development enables for more and more middle and large sized agricultural firms to store their data about their production on computers. Most database are stored in GIS-based systems, which have several advantages: handling and visualizing data with geographic coordinates; capability to use various data formats (digital soil and cadastre maps, remotely sensed images, field measurements, etc.); geostatistical analyses can be performed. In our article we intrude the first steps of building the Eastern Sugar Ltd.' complex GIS database in the area of the sugar factory at Kaba.

(Keywords: sugar factory crop production system, GIS data base, data integration)

BEVEZETÉS

A mezőgazdasági termelőnek folyamatos törekvése, hogy a termőhely tulajdonságait egyre nagyobb részletességgel ismerje meg és ezeket az ismereteket fokozott hatékonysággal tudja felhasználni. A termőhelyi környezetet napjainkig, a termesztési gyakorlat a maga bonyolultságában azonban főleg a meghatározó folyamatokra alapozva és nagyobb művelési egységekben kezelte (Tamás, 1999). Folyamatos technikai fejlesztés után az 1980-1990 közötti évtizedben következett be az a technikai-technológiai és szemléletbeli áttörés, melynek hatására bekövetkezett a térinformatika magyarországi megjelenése. A mezőgazdasági gyakorlatban a szántóföldi növénytermesztés keretei között a helymeghatározáson alapuló technika alkalmazása

lehetőséget nyújt ahhoz, hogy a korábbi, a termesztési egységekre (táblákra) vonatkozó információk (pl. hozamtérkép, talajtérkép, növényi kórokozók, kártevők és gyomok által károsított terület és a károsítás foka) egységes rendszerben legyenek kezelhetők és a beavatkozások – trágyázás, növényvédelmi munkák, stb. – adott helyhez kötötten történhessenek (Németh és Jolánkai, 2002).

A térinformációs rendszerek helyhez kötött információk gyűjtésére, kezelésére, elemzésére és megjelenítésére szolgálnak. Az információk elemzésében fontos szerepet játszik a térbeliség, a megjelenítésben pedig a képi jelleg (Detrekői és Szabó, 1995).

A helyhez kapcsolható információk következtében a hely szinte önként adódik a különböző információk összekapcsolásának eszközeként (Detrekői és Szabó, 1995). Az információk közel 80-90%-a a térhez kapcsolódik. Az információs társadalom kialakulása idején ez a szakterület alkalmazását kikerülhetetlenné teszi (Tamás, 2001). Az adatok térinformatikai elemzése egyszerűbb, könnyebb kezelhetőséget biztosít, az egyes adatok átláthatóbbak, de nem hanyagolható el a megjeleníthetőség jelentősége sem, hiszen a hagyományos térképi ábrázolásokkal szemben a térinformatikai rendszerek lényegesen gazdagabb megjelenítési lehetőségekkel rendelkeznek (képernyőn, plotteren, nyomtatón való adatközlés, raszter- és vektorábrázolás, színek, vonalvastagság, textúra változtathatósága stb.) (Buttenfield and Mackaness, 1991). Mindemellett megfelelő minőség esetén a térinformációs rendszerekben tárolt, illetve előállított új információkra teljesülhetnek a Kesselyák (2001) által előírt, illetve megkívánt tulajdonságok, vagyis a sértetlenség, a biztonság, az adatvédelem/adatbiztonság, a titkosság, a hitelesség, a rendelkezésre állás, a könnyű kezelhetőség, a könnyű visszakereshetőség, valamint az archiválhatóság.

A cukorrépa – összehasonlítva más növényeinkkel – egyik legfiatalabb kultúrnövényünk. Magyarországra a répát Tessedik Sámuel 1767-ben hozta be Németországból. A hazai ipar alig 200 éves múltra tekint vissza. Az első répacukorgyártó üzem munkába állítására 1808-ban, Ercsiben került sor (Kettinger és Ördög, 2004). A nagyobb cukorgyárak építése az 1830-as években kezdődött, mely az 1880-as évekre befejeződött, majd a magyar cukoripar virágzásnak indult.

Az I. világháború kitöréséig az országban 31 cukorgyár működött, a háború után már csupán 12. Magyarországon az 1930-as évekig főleg németek által nemesített cukorrépát termesztettek. A II. világháború után a cukorgyárakat államosították, 1971-től a cukorgyárak önálló vállalatok lettek, trösztí irányítással, mely 1980-ban szűnt meg.

A rendszerváltozáskor Magyarországon 12 cukorgyár működött. A privatizáció 1991-ben kezdődött meg, a jelenlegi tulajdonosi szerkezet 1997-re alakult ki. 3 nagy európai cukorgyártó társaság: az osztrák Agrana, a francia Eridania Béghin-Say és a brit és francia Eastern Sugar szerzett többségi tulajdont. Ők jelentős fejlesztéseket hajtottak végre a cukorgyárakban, növelték a feldolgozó kapacitást, korszerűsítették a technológiát.

Az európai cukorpiac átfogó reformjának köszönhetően a hazai cukorgyárak is szerkezetváltásra kényszerülnek. A kabai Eastern Sugar Rt. fennmaradása érdekében sürgető feladat a többszempontú optimalizálás végrehajtása, hiszen a cukorágazatban bekövetkező változások és a támogatások átcsoportosulása következtében a termelők egy része a cukorrépa termesztés feladására kényszerül. A versenyhelyzet a cukorgyáraktól is megköveteli a költséghatékonysági szempontú gazdálkodás ártérértékelését. A hatékony működéshez átfogó ismeretekre van szükség valamennyi termelési alkörzetről. Munkánk során a kabai cukorgyár termelési rendszerének adatain nyugvó integrált térinformatikai adatbázist hozunk létre, melynek segítségével kiválasztható a gyár működéséhez legoptimálisabb termeltetés, ezáltal javítva annak cukorpiaci versenyképességét.

ANYAG ÉS MÓDSZER

Az Eastern Sugar cukorrépa feldolgozással foglalkozó vállalatcsoport. A cég elsődleges üzleti tevékenysége EU minőségű cukor előállítása a helyi fogyasztói szükségletek kielégítésére. Az Eastern Sugar Csoport Közép-Európában az egyik legnagyobb cukorgyártó, mely az egykori kabai Hajdúsági Cukorgyár felvásárlásával alakult 1991-ben. Ma az Eastern Sugar Csoport 5 gyárat működtet 3 országban (egyet Magyarországon (Kabán), egyet Szlovákiában (Dunaszerdahelyen) és hármat Csehországban (Nemcice nad Hanou-n, Kojetin-ben és Hrochuv Tynec-ben) (1. ábra). Évente több mint 300.000 tonna cukrot állít elő. Méretének köszönhetően az Eastern Sugar Csoport az egyik vezető cukorrépa feldolgozó vállalat a régióban.

1. ábra

Az Eastern Sugar Csoport gyárainak elhelyezkedése



Forrás (Source): www.easternsugar.hu

Figure 1 Location of the Eastern Sugar Group's sugar factories

A térinformatika rohamos fejlődésének köszönhetően az Eastern Sugar Rt. is belépett azon felhasználók sorába, akik termeltetési adataikat térinformatikai eszközök segítségével akarják feldolgozni. Ennek egyik kiemelkedő oka a megnövekedett információ mennyiség, valamint az adatok helyhez kapcsolhatósága. A széleskörű alapadatokból kiépített adatbázis mindemellett alapul szolgál a termelési alkörzetekben végrehajtandó optimalizálás elvégzéséhez. Vizsgálataink kiterjednek többek között az egyes termelési alkörzetekben gazdálkodók terméshozam adatainak értékelésére, a gépellátottság felmérésre, továbbá a szántóföldről a cukorgyárba kerülésből adódó szállítási költséget is vizsgáljuk. A termelési alkörzeteket 1:10.000 méretarányú kataszteri térképek digitalizálását követően válogattuk le. Az adatszolgáltatásban az Eastern Sugar Rt. volt segítségünkre. A cukorgyár munkatársai kérdőíves felmérést végeztek termelőik körében a 2004-es évre vonatkozóan. Ezen kérdőívek is feldolgozásra kerültek munkánk során.

A cukorgyár 2 fő termeltetési körzetre bontható: a kabai és a békéscsabai körzetre. Vizsgálatainkat a kabai körzetre terjesztettük ki, ezt a termeltetési körzetet 45 térképszelvény fedi le. Ezen a termelési körzeten belül a cukorgyárnak beszállító cégek

területei eloszlásuk szerint 6 településre terjednek ki (Hajdúböszörmény, Nagyhegyes, Debrecen, Eszék, Hajdúszoboszló, Balmazújváros). A 2. ábrán Nagyhegyes egyik kataszteri térképét láthatjuk. A kataszteri térképeket ArcView 3.3 program segítségével dolgoztuk fel.

EREDMÉNY ÉS ÉRTÉKELÉS

A munka első részében a meglévő termesztéstechnológiai adatokból komplett adatbázist alakítottunk ki. Az adatok beszerzése folyamatosan történik, jelenleg az Eastern Sugar Rt. munkatársai által, a termelők körében végzett kérdőíves felmérésből származó adatok állnak rendelkezésünkre 4 termelési alkörzetre. A kérdőívben szereplő kérdések az általános termelési színvonal vizsgálatára irányultak.

Az adatbázis alapjául 1:10.000 méretarányú kataszteri térképek szolgálnak. Ezek rektifikálást követő digitalizálása után a helyrajzi számok megadásával leválogattuk az egyes termelők szántóterületeinek nagyságát minden egyes termelési alkörzetre. Ezt követően meghatároztuk az ily módon felosztott területek nagyságát az ArcView 3.3 program segítségével. A m²-ben megkapott adatokat átváltottuk hektárba, majd összesítettük először cégenként, majd termelési alkörzetenként a szántóterületek nagyságát (3., 4. ábra).

2. ábra

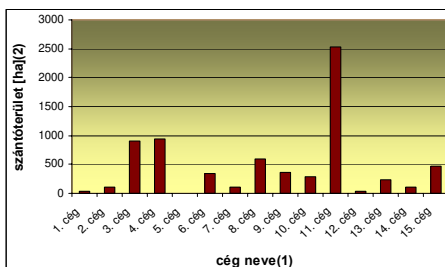
Nagyhegyes egyik kataszteri térképének részlete,
rajta a gazdálkodók területeinek elhelyezkedésével



Figure 2 Part of Nagyhegyes village's cadasteric map, indicating the lands of the agricultural firms

3. ábra

Az Eastern Sugar Rt. termelőinek szántóterület nagysága



4. ábra

Szántóterületek nagysága termelési alkörzetenként

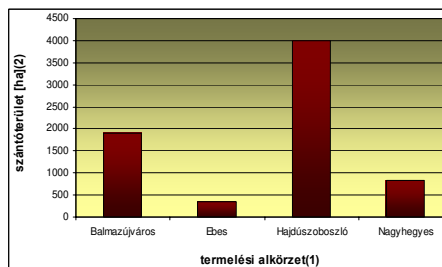


Figure 3 The size of arable land possessed by the firms producing for Eastern Sugar Ltd.

Figure 4 Distribution of the size of arable lands in the production subdivisions

Name of the firm(1), Size of arable land [ha](2)

Production subdivision(1), Size of arable land(2)

A jelenleg birtokunkban lévő adatok feldolgozása után elvégeztük az egyes termelési alkörzetek termésátlagainak kiértékelését és összehasonlítottuk a 2004-es évi országos termésátlagokkal (5. ábra). Az egyes alkörzetekre a 8 fő termesztett növény (csemege kukorica, takarmány kukorica, kalászosok, cukorrépa, burgonya, napraforgó, repce, pillangós takarmány) termésátlagait dolgoztuk fel. Az ebben a körzetben termelők eredményei a cukorrépa és a burgonya esetében jelentősen, a napraforgó, a kalászosok és a pillangósok esetében kis mértékben meghaladja az országos termésátlagokat, viszont a csemege kukorica, a takarmánykukorica és a repce alatta maradtak annak.

A cukorgyár által végzett kérdőíves felmérés kiterjedt a gépellátottság vizsgálatára is. Az adatok alapján kiszámoltuk a gépek átlag életkorát az egyes termelési alkörzetekre vonatkozóan. A teljes műszaki ellátottság bemutatásától eltekintően, ezért csak az általunk 2 legfontosabbnak vélt gépcsoport, az erőgépek és a betakarítógépek adatainak kiértékelésével nyert grafikont mutatjuk be (6., 7. ábra). Az erőgépek átlag életkorából kitűnik, hogy az ebesi alkörzetben tevékenykedő cégek erőgépei a legidősebbek, ám a betakarítógépek vizsgálata azt mutatja, hogy ebben a körzetben a legfiatalabbak a gépek. Hasonló, de fordított összefüggés látható a nagyhegyesi körzet grafikonján. Itt a legfiatalabbak az erőgépek, viszont a betakarítógépek a legidősebbek. Az adatokból jól látható, hogy a cégek többsége nem képes egyszerre minden géptípust fejleszteni, így a gépbeszerzések során általában mindössze csak 1 géptípus fejlesztésére, új gép vásárlására van anyagi kapacitásuk. Az ebesi és a nagyhegyesi termelési alkörzetre a kevés számú, de nagy összefüggő területekkel rendelkező termelők a jellemzőek. A többszempontú optimalizálás során várhatóan inkább a kisebb termelők közül fognak kikerülni azok, akik cukorrépa termesztésük feladására kényszerülnek, hiszen a kis gazdálkodások számára egy új gép beszerzése, különösen ha speciális cukorrépa betakarítógépről van szó, anyagiilag rendkívül megterhelő.

Az adatbázis kialakítását követően rögzítjük térképen az adatokat. Logikai földrajzi modellt alkotunk, majd elemzést hajtunk végre. A kitűzött cél érdekében szükséges adatforrások meghatározását követően azok térképi rétegekre való bontását végezzük el (8. ábra) és annak meghatározását, hogy az egyes rétegekben milyen objektum típusokat

fogunk elhelyezni. Elvégezzük a szükséges adatintegrációt az egységes vektoros vagy raszteres környezet kialakításához. A hagyományos tudományokban az egyes komplex jelenségeket egymástól elkülönítve kellett vizsgálni, míg a térinformatika ezek komplex egymásra hatásában tudja elemezni az egyes jelenségeket. A térinformatikai modellek rendkívül jól tudják támogatni a különböző döntéshozatali folyamatok megértését és szimulálását. Egy egyszerű adatbázis lekérdezés eredményeként gyakran a döntéshozó nem megfelelő információval rendelkezik a döntés előkészítéséhez (Tamás, 2001). Egy komplex döntéstámogatási rendszer viszont, ahol a döntéshozatali szempontoknak megfelelő többszörös feltételrendszert és elfogadható kockázati szinteket tudunk beépíteni a környezeti modellbe, megalapozottabb, és jóval alacsonyabb döntési kockázattal járó döntési alternatívát hozhatunk, mint egy egyszerű lekérdezés révén.

5. ábra

A termelési alkörzetek termésátlagainak összehasonlítása az országos termésátlagokkal

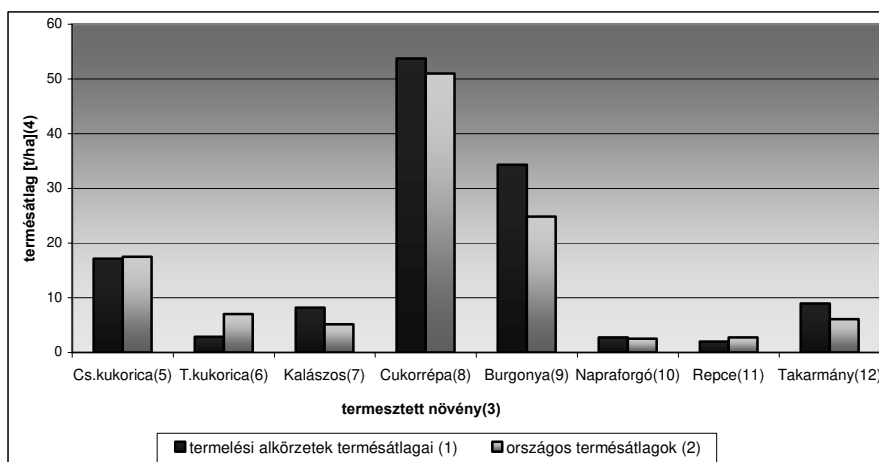


Figure 5 The average yields of the production subdivisions, compared by country yields

Yield of the production subdivisions(1), Country yields(2), Grew plant(3), Yield [t/ha](4), Sweet corn(5), Fodder corn(6), Corn in the ear(7), Sugar beet(8), Potato(9), Sunflower(10), Rape(11), Fodder(12),

A GIS adatbázis nagy előnye az, hogy képes számos más formátumú adat fogadására, megjelenítésére, feldolgozására, interpretálására. A jövőben a folyamatosan beérkező termelési adatokon kívül a következő adatokkal fogjuk kiegészíteni az termeltetési adatbázist: talaj-térképek (AgroTopo, üzemi genetikus és földértékelési térképek), légifotók (hagyományos, valamint a tanszékünkön található multispektrális Tetracam felvételek), műholdas felvételek (MODIS adatsorok, heti vagy havi összesítésben, NDVI index; igény és anyagi lehetőségek szerint Spot vagy Landsat képek). Rendelkezésünkre áll a Corine földhasználati kategóriák térképe. Ezen elemzést fogunk végrehajtani arra vonatkozólag, hogy mennyire változott meg a terület hasznosítása, főleg a szántóföldi területek vonatkozásában, a cukorgyári körzetek területén. Ezen felül az adatbázis bővíthető GPS-es terepi mérések (talajminták, földterületek pontosítása, meteorológiai mérőállomások adataival).

6. ábra

Erőgépek átlag életkor szerinti megoszlása termelési alkörzetenként

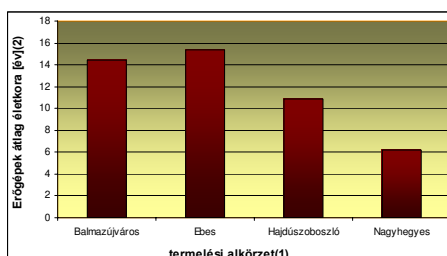


Figure 6 Average year of farm engines by production subdivisions

Production subdivision(1), Average age of farm engines [year](2)

7. ábra

Betakarítógépek átlag életkor szerinti megoszlása termelési alkörzetenként

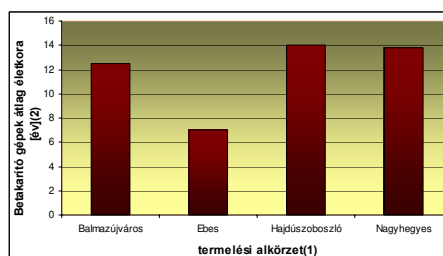


Figure 7 Average year of harvesting machines by production subdivisions

Production subdivision(1), Average age of harvesting machines(2)

8. ábra

A termelési alkörzetek és szántóterületeik

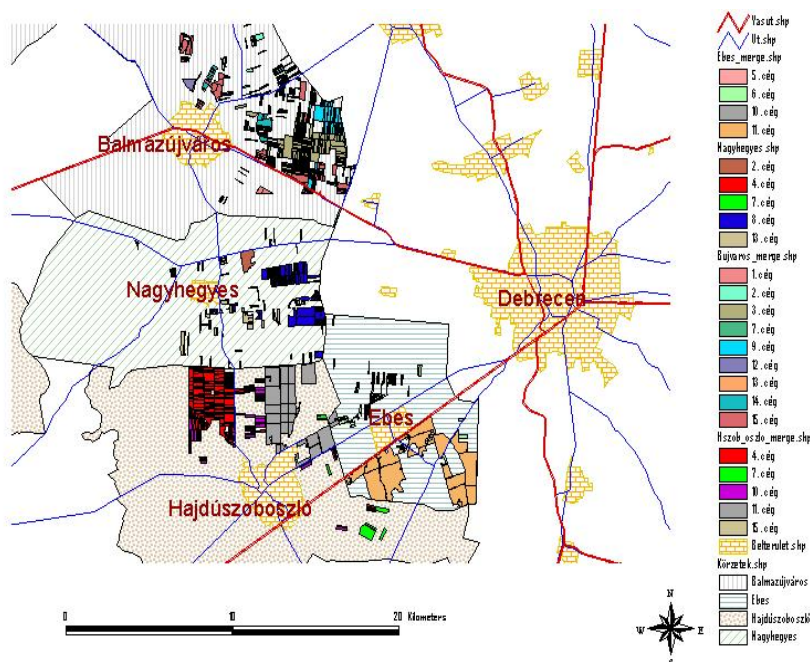


Figure 8 The location of production subdivisions, indicating the arable lands of the agricultural firms

KÖVETKEZTETÉSEK

A termőhely igen intenzív folyamatos elemezhetősége minimalizálja a termesztési kockázatot. Jelen munkánkban a cukorgyári termelési rendszer kezdeti lépéseiről számoltunk be. A minél szélesebb körű termesztéstechnológiai és egyéb, az egyes termelési alkörzetekre vonatkozó adatokból kialakított, térinformatikai környezetbe integrált adatbázis alkalmas olyan területi és termelési optimalizálásra, melynek hatására leszűrhetőek egy termeltetési rendszer gyenge pontjai, és ezáltal nemcsak a termeltetés hatékonysága növelhető meg, de jelentős költségcsökkenés is realizálható.

Az Eastern Sugar Rt. termesztéstechnológiai adatbázisa folyamatosan bővül. A közeljövőben kerül sor a tápanyagvisszapótlási, vetési, gyomirtási adatok feldolgozására, melynek révén a cukorgyár olyan adatbázis birtokába jut, amelynek segítségével termeltetési rendszere egyszerűbbé, kezelhetőbbé, és legfőképp a piacon versenyképesebbé válhat. A kibővített adatbázist ArcMap környezetbe, korszerű geoadatbázis formátumúvá is átalakítjuk. Az adatok folyamatos frissítésével a speciális szempontoknak megfelelő, a piaci igényekhez alkalmazkodó optimalizálás könnyen és egyszerűen végrehajthatóvá válik, ezáltal biztosítható a cukorgyár hatékony működése.

IRODALOM

- Buttenfield, B.P., Mackaness, W.A. (1991). Visualization. 9-20. Maguire, D.J. et al. (ed.): Geographical Information System: principles and application, Longman, London, 1. 428-443.
- Detrekői, Á., Szabó, Gy. (1995). Bevezetés a térinformatikába. Nemzeti Tankönyvkiadó : Budapest. 15-17.
- Kesselyák, P. (2001). Új minőségi elvárások az információs társadalomban. Magyar Minőség, 10. 2. Budapest, 8-11.
- Kettinger, Gy., Ördög, V. (2004). Fejezetek a magyar vetőmagszakma magyaróvári történeteiből (1867-1945). In: Mag Kutatás, Fejlesztés és Környezet. 2004. 12. 2-11.
- Németh, T., Jolánkai, M. (2002). A precíziós növénytermesztés elemei. In: Nagy, J. (szerk.) Mezőgazdaságtudomány, EU konform mezőgazdaság és élelmiszerbiztonság. DE ATC, Debrecen. 12-21.
- Tamás, J. (1999): Precíziós mezőgazdaság kiépítésének feltételrendszere. Tiszántúli Mezőgazdasági Tudományos napok '99. DATE, 1999. október 28-29.
- Tamás, J. (2001): Precíziós mezőgazdaság. Mezőgazdasági Szaktudás Kiadó, Budapest. 50-65.

Levelezési cím (*Corresponding author*):

Nagy Ildikó

Debreceni Egyetem, Agrártudományi Centrum, Mezőgazdaságtudományi Kar
Víz- és Környezetgazdálkodási Tanszék
4032 Debrecen, Böszörményi út 138.

*University of Debrecen, Centre of Agricultural Sciences, Faculty of Agriculture,
Department of Water- and Environmental Management
H-4032 Debrecen, Böszörményi street 138.*

Tel.: 36-52-508-444/88236

e-mail: inagy@gissserver1.date.hu



A Mobile 3D technológia lehetőségeinek kihasználása a térbeli tájékozódásban

Tukora B.

Pécsi Tudományegyetem, Pollack Mihály Műszaki Kar, Műszaki Informatika Tanszék, 7624 Pécs, Rókus utca 2.

ÖSSZEFOGLALÁS

A legújabb technológiai fejlesztések lehetővé teszik, hogy a globális helymeghatározó rendszerrel (GPS) nem rendelkező mobil eszközök (mobiltelefonok, PDA-k) a beépített kamerájuk által szolgáltatott vizuális információk alapján meghatározzák saját helyzetüket a térben, környezetüket virtuális objektumokkal egészítsék ki, és az ily módon létrehozott világot megjelenítsék a képernyőjükön. E koncepció egyik lehetséges felhasználási területe a régészeti ásatások, alaprajz szintjén megmaradt utca- vagy városrészletek eredeti állapotának bemutatása úgy, hogy a szemlélő mobil eszközének kijelzőjén mindig az aktuális helyzetének és orientációjának megfelelően rajzolódik ki a régvolt környezet. Előnye a már meglévő, hasonló célú rendszerekkel szemben az olcsósága, rugalmassága és széles körű használhatósága. A mobil eszköznek mindeközben igen összetett feladatot kell elvégeznie. A dolgozat a rendszer fejlesztése során felmerülő problémákat tárgyalja és ad rájuk megoldási javaslatokat.

(Kulcsszavak: mobil technológia, kamera kalibráció, 3D rekonstrukció, modellezés egyetlen kép alapján)

ABSTRACT

Utilizing the Mobile 3D technology's potentials in the spatial orientation

B. Tukora

University of Pécs, Pollack Mihály Faculty of Engineering, Department of Information Technics, H-7624 Pécs, Rókus u. 2.

Recent developments make it possible for the mobile devices without GPS system (such as mobile phones and PDA-s) to determine their spatial situation on the visual informations given by the embedded camera, complete the environment with virtual objects, and display the created world. This conception can be used at the visualisation of archeological fossils and the original condition of ruined sections of streets or towns, where the bygone environment is shown according to the viewer's position and orientation. Comparing with the similar existing systems the advantages are cheapness, flexibility and the possible wide range of usage. The mobile device used for this purpose must accomplish a very complex task. This paper discusses the arising problems during the development and gives proposals for solving them.

(Keywords: mobile technology, camera calibration, 3D reconstruction, single view modelling)

BEVEZETÉS

A korszerű mobiltelefonok a Mobile 3D technológia segítségével képesek az OpenGL szabvány által definiált 3 dimenziós objektumok, világok megjelenítésére Java platformon. A szabványosított fejlesztői környezetet a JSR-184 Java specifikáció rögzíti. Szintén

rögzített az eszközökhöz tartozó digitális fényképezőgépek kezelésének (JSR-135) illetve a kliens-szerver kommunikáció (JSR-120) programozásának módja is. A dolgozat egy olyan, mobil eszközökön futó alkalmazás fontosabb funkcióit tárgyalja, amely

- a környezetében jelen lévő referenciapontokat kellő biztonsággal azonosítja,
- a referenciapontok alapján meghatározza saját pozícióját és orientációját a térben,
- ennek megfelelően a memóriájában tárolt 3 dimenziós objektumokat beilleszti a környezetről készült képbe,
- végül megjeleníti virtuális elemekkel kiegészített helyszínt a kijelzőjén.

ANYAG ÉS MÓDSZER

A referenciapontok azonosítása

A mobil eszközön futó alkalmazás első feladata az, hogy beépített fényképezőgépével képet készítsen a helyszínről, felismerje rajta a környezet térbeli helyzetéről áruklódó elemeket, és ezek pozícióját képernyő-koordináta-rendszerben (2D) tárolja. Egy olyan algoritmus lenne az ideális, amellyel egyértelműen azonosítani tudnánk a képen megjelenő objektumokat: házakat, köztéri tárgyakat. Sajnos ilyen – általánosan használható, egyszerű betanítási funkcióval bíró – módszer nem létezik. A megoldást külön e célra készített referenciátárgyak, -pontok, vagy -ábrák környezetbe helyezése jelenti. Az ezekkel szemben támasztott követelmények: a könnyű felismerhetőség, a képen szereplő tárgyakról való biztos megkülönböztethetőség. Ezek kielégítéséhez kell megtalálnunk a megfelelő szint (színeket) és formát.

A beépített fényképezőgéppel készült képeket a Java-specifikáció szerint raszteres formában, RGB színrendszerben kódolva kapjuk meg. Ez azt jelenti, hogy minden egyes képpontot három szám jellemez: a pont színének vörös (Red), zöld (Green) és kék (Blue) komponensei. Amennyiben az összetevők mennyiségét a 0..1 skálán adjuk meg, egy tiszta vörös képpont az 1;0;0 értékkel bír. A színelméleti alapoknak megfelelően a 0;0;0 a fekete, az 1;1;1 a fehér színre jellemző érték. Amennyiben a vörös színt világosítani szeretnénk, a vörös mellett a zöld és kék összetevők arányát is növelnünk kell (azonos mértékben): pl. 1;0,2;0,2. Egy tárgyat a természetben mindig más és más színűnek látunk: a környezeti fényeknek megfelelően sötétebbnek vagy világosabbnak, eltérő árnyalatúnak. A mobil eszközök egyszerű automata kameráinál fokozottan érvényesül ez a jelenség: egy vörös színű tárgy fehér háttér előtt szinte feketének látszik, mivel az automatika az egész kép átlaga alapján számítja ki a középtónust, így a fehér „elszürkül”, a sötétebb tárgyak feketék lesznek. Egy mobiltelefonnal tehát nem lehet „egy bizonyos” szint megkeresni, csak „hozzá hasonlót”.

Az RGB rendszerben a környezettől leginkább megkülönböztethető színek azok, amelyekben valamelyik komponens jóval kisebb vagy nagyobb arányban szerepel, mint a másik kettő. Ilyenek az alapszínek – a vörös (1;0;0), a zöld (0;1;0) és a kék (0;0;1) – mellett az ezek additív keveréséből származók: a sárga (1;1;0), a lila (1;0;1) és a zöldeskék (0;1;1). Ha a referenciapontokat középtónusú szürke alapon az említett színek valamelyikéből állítjuk elő, nagy esélyünk van arra, hogy különböző fényviszonyok mellett is felismerhetők maradnak. A meglévő bizonytalanságok miatt az eljárás sikere jelentősen múlik azon, hogy a lehető legkevesebb színt használjuk fel az azonosításhoz. Természetesen más tárgyaknak (ruháknak, hirdetéseknek) is lehet a referenciapontokhoz hasonló színük. Az egyértelmű azonosítás végett a pontokat mintázatba, összetett ábrába (tárgyra) kell helyezniük. A mintázatoknak könnyen dekódolhatóknak, felismerhetőknak kell lenniük. Egy szürke alapon vörös négyzet például egyszerre négy pontot definiál, egyszerűen kiszámolhatók a sarokpontok koordinátái, alakja bizonyos

határok között könnyen ellenőrizhető, a szürke és vörös közötti csekély tónuskülönbség egyensúlyt ad a fényviszonyokban –, viszont a képen más vörös téglalapok is előfordulhatnak, és négy egyenrangú pont nem ad információt a négyzet orientációjáról. Ahhoz, hogy konkrétan milyen referenciatárgy vagy -ábra kerüljön felhasználásra, tudnunk kell, hogy a pozíció és orientáció meghatározását végző algoritmusnak hány, és milyen elrendezésű referenciapontra van szüksége –, így erre a kérdésre választ csak a következő fejezet után adhatunk.

A kamera pozíciójának és orientációjának meghatározása

A feladatunk az, hogy a referenciapontok alapján meghatározzuk a mobil eszközünk helyzetét és irányultságát a környezet tárgyaihoz viszonyítva. Ehhez definiálnunk kell egy világ-koordinátarendszert a helyszín egy tetszőleges pontján, tetszőleges (de logikus) irányultsággal. A referenciapontok egyértelműen utalnak e koordináta-rendszer origójának helyzetére és egységvektorainak irányára, nagyságára. Mindeközben az eszköz memóriájában egy „virtuális” koordináta-rendszerben definiált objektumok várakoznak. Amennyiben meg tudjuk határozni a világ-rendszerben a kamera helyzetét, és ugyanezen paraméterekkel állítjuk be a kamera virtuális koordináta-rendszerhez viszonyított helyzetét, a két koordináta-rendszer fedésbe kerül egymással: a „külső” és „belső” világ összeolvasható.

Legegyszerűbb esetben a referenciapontok világ és virtuális koordináta-rendszerbeli helyzete megegyezik, így a világ-koordinátarendszerhez viszonyított helyzet meghatározásával automatikusan megkapjuk a virtuális-rendszerbeli helyzetünket is.

A 3 dimenziós pontok a középpontos (perspektívus) projekció szabályai szerint vetülnek a képernyő síkjára (1. ábra). Középpontos vetítés során minden egyes térpont, illetve a vetítési síkon előállt képe a kamera gyűjtőpontjából kiinduló egyazon egyenesre esik. Mivel a vetítéskor a mélységre vonatkozó adat elveszik, a perspektívus projekció elvileg nem invertálható folyamat: ez adja a számítás nehézségét.

1. ábra

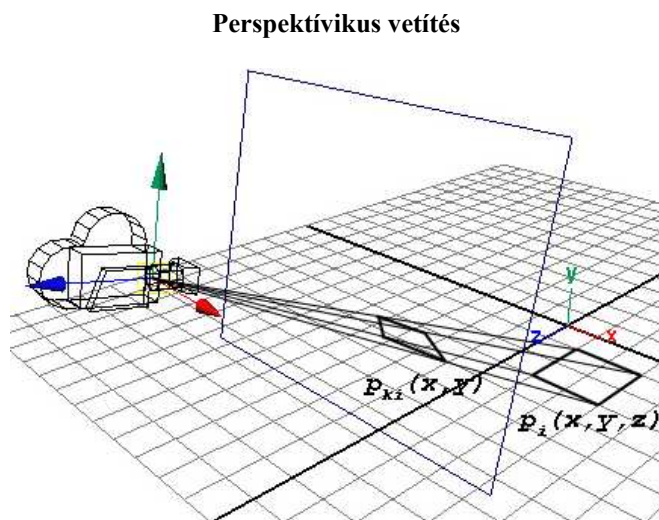


Figure 1: Perspective projection

Számításunk kiinduló értékei az 1. ábra szerint:

$$\underline{\underline{p}} = [\underline{p}_1, \underline{p}_2, \dots, \underline{p}_n] = \begin{bmatrix} p_{1x} & p_{2x} & \dots & p_{nx} \\ p_{1y} & p_{2y} & \dots & p_{ny} \\ p_{1z} & p_{2z} & \dots & p_{nz} \end{bmatrix}, \text{ vagyis a referenciapontok 3 dimenziós}$$

koordinátái a világ-koordináta-rendszerben, illetve

$$\underline{\underline{p}}_k = [\underline{p}_{k1}, \underline{p}_{k2}, \dots, \underline{p}_{kn}] = \begin{bmatrix} p_{k1x} & p_{k2x} & \dots & p_{knx} \\ p_{k1y} & p_{k2y} & \dots & p_{kny} \end{bmatrix}, \text{ a referenciapontok 2 dimenziós}$$

koordinátái a képernyő-koordináta-rendszerben.

A kamera helyzetét leíró paramétereket homogén transzformációs mátrix formájában keressük:

$$\underline{\underline{H}} = \left[\begin{array}{ccc|c} \underline{\underline{R}} & \underline{\underline{k}} \\ \hline 0 & 0 & 0 & 1 \end{array} \right] = \left[\begin{array}{ccc|c} \underline{e}_1 & \underline{e}_2 & \underline{e}_3 & \underline{k} \\ \hline 0 & 0 & 0 & 1 \end{array} \right] = \left[\begin{array}{ccc|c} e_{1x} & e_{2x} & e_{3x} & k_x \\ e_{1y} & e_{2y} & e_{3y} & k_y \\ e_{1z} & e_{2z} & e_{3z} & k_z \\ \hline 0 & 0 & 0 & 1 \end{array} \right], \text{ ahol}$$

$\underline{\underline{R}}$ a kamerához rendelt koordináta-rendszer forgatómátrixa az alap-koordináta-rendszerben, $\underline{e}_1$, $\underline{e}_2$, $\underline{e}_3$ a kamera-koordináta-rendszer egységvektorai az alap-koordináta-rendszerben, $\underline{k}$ pedig a kamera-koordináta-rendszer origójának pozíciója az alap-koordináta-rendszerben (2. ábra).

2. ábra

Kamera koordináta-rendszer

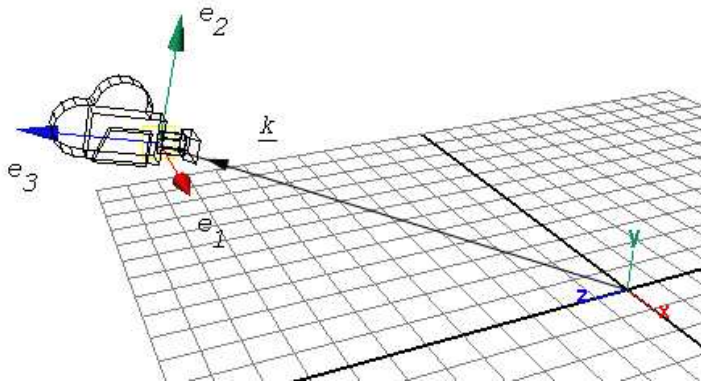


Figure 2: Camera coordinate-system



HEC-RAS alapú geoadatbázis vizsgálata az EU Vízügyi Keretirányelv előírásai alapján I.

Pregun Cs., Tamás J., Takács P., Bíró T.

DE ATC MTK Víz- és Környezetgazdálkodási Tanszék, 4032 Debrecen, Böszörményi út 138.

ÖSSZEFOGLALÁS

A HEC-RAS hidrológiai program alkalmas a felszíni természetes és mesterséges vízfolyások és vízfolyásrendszerek modellezésére, beleértve a különböző műtárgyak, szükségeltározók hatásait, árhullámok levonulását, árvízi elöntéseket, töltésszakadásokat stb. A program megbízhatóan számolja a mederre jellemző aktuális hidraulikai adatokat, valamint ezeket helyzetazonosítottan, ArcView GIS környezetben tudja kezelni és menteni. A Tanszéken rendelkezésünkre álló térinformatikai adatbázisban a Bihari-sík földtani, talajtani, domborzati, földhasználati, lefolyási stb. sajátosságairól találhatóak részletes információk. A hidrológiai és a térinformatikai modellek közötti adatcsere révén becsléseket tehetünk a környezet- és tájvédelmi célú, szükséges és elégséges vízrendezési beavatkozásokról
(Kulcsszavak: hidrológiai és térinformatikai modell, környezet- és tájvédelem, vízrendezés)

ABSTRACT

Analyzing of HEC-RAS based geodata basis from the aspects of the EU Water Framework Directive

Cs. Pregun, J. Tamás, P. Takács, T. Bíró

University of Debrecen, Center for Agriculture, Department of Water- and Environmental management, H-4032 Debrecen,
Böszörményi út 138.

The HEC-RAS hydrological software is suitable for modeling natural and artificial water bodies and river networks, including hydraulic structures. The program able to simulate the effects of emergency water reservoirs, dam breaks; visualize, track and calculate floods and possible flooded areas at given water level. The program calculates several actual hydraulic data of the riverbed, and handles the both inputs and outputs in georeferenced ArcView GIS environment. The complex GIS database in our Department contains detailed information on the geographic, geologic, land use, slope, runoff, soil, etc. features of the Bihar-Plain. The data exchange between the GIS and the hydrologic database offers the opportunity to calculate predictions for the suitable and necessary surface watershed control for environmental and landscape protection purposes.

(Keywords: hydrological and GIS model, environmental and landscape protection, watershed management)

BEVEZETÉS

Az EU Vízügyi Keretirányelvvel kapcsolatos főbb hazai feladatok

- Az EU Vízügyi Keretirányelve (2000/60/EK) célul tűzte ki, hogy a tagországok víztereiben el kell érni a jó ökológiai állapotot az érvénybelépéstől (2000. október

- 23.) számított 15 éven belül. Ennek elérése érdekében először minden tagországban fel kell mérni a felszíni vizek és a vízgyűjtő területek ökológiai állapotát.
- A jó ökológiai állapotot megfelelő hidrológiai jellemzőkkel kell alátámasztani. A DE ATC MTK Víz- és Környezetgazdálkodási Tanszékén e kapcsolatrendszer feltárása és koncepcionális modellezése céljából folynak kutatások (*Tamás és Bíró, 2005*).
 - Az EU Vízügyi Keretirányelv alapján az osztott vízgyűjtőn való fenntartható vízgazdálkodásnak ki kell elégítenie a természet- és környezetvédelmi, a vízgazdálkodási, az árvíz- és belvízvédelmi, a mezőgazdasági, az ipari, a lakossági stb. igényeket egyaránt.
 - Az Európai Nitrát Irányelv (91/676/EGK) többek között javasolja bizonyos mezőgazdasági területeknek a vizes élőhelyek rehabilitációjával párhuzamosan történő művelés alóli kivonását a felszíni és felszín alatti vizek érdekében.
 - A Kárpát-medence vízkészlet-gazdálkodására a jellegzetes földrajzi viszonyok alapvető hatást gyakorolnak. A felszíni vizek túlnyomó része a medence peremén, Magyarország határain kívül ered, ezért a medencébe érkező felszíni vízkészletek minőségét és mennyiségét csak a határokon belül áll módunkban ellenőrizni és szabályozni (*Somlyódy, 2000*).
 - A Berettyó folyó és vízgyűjtője a folyamatos vízrendezések során mára szinte teljesen mesterséges mederben és vízszabályozási körülmények között található. A folyó természetes hidrogeológiai viszonyaira alapozva saját önszabályozó mechanizmusát igyekszik követni, amely az antropogén környezetben számos vízmennyiségi és minőségi problémát vet fel.
 - Az általunk végzett hidroökológiai vizsgálatok és a folyamatok modellezése egy természetközeli, önfenntartó, és környezete számára kisebb kockázatot jelentő folyókörnyezet kialakítását célozza meg. Ennek keretében elvégeztük a folyó hidrológiai felmérését, és ökológiai mintatereteket válogattunk le. Az EU VKI javaslatait is figyelembe véve kialakítottuk a folyó és környezetének virtuális modelljét a HEC-RAS program segítségével, és megalapoztuk egy dinamikus térbeli döntéstámogatási rendszert kialakítását. Felhasználásra került a VKI CIRCA a Víz Keretirányelv Közös Végrehajtási Stratégiájával kapcsolatos információcsere eszköze. Ezt a szerepét többek között azzal tölti be, hogy tartalmazza mindazokat a dokumentumokat, amelyeket a Víz Keretirányelv Közös Végrehajtási Stratégiája alapján az EU szinten létrehozott különböző szervezeti egységek – munkacsoportok, fórumok, stb. – kidolgoznak. Ez lehetővé teszi a fenntartható földhasználat és vízkészlet-gazdálkodás ok-okozati rendszerének hatékony megértését és fejlesztését.
 - A hazai és a nemzetközi elvárások kielégítésének meghatározó eszköze a komplex hidrológiai-hidrodinamikai modellek alkalmazása (*Tamás, 2004*).

A hidrológiai és térinformatikai modellezés szerepe a környezetvédelemben

A különböző környezeti, hidrológiai és hidroökológiai stb. modellek és módszerek szerepe egyre inkább növekszik mind a vízkészlet-gazdálkodásban, mind a környezetvédelemben. Ennek okai a következők:

- Nő az igény a hidrológiai folyamatok minél pontosabb, különböző léptékű térbeli és időbeli leírására
- Fokozódik a szükséglet az emberi tevékenység vízminőségre és a hidrológiai ciklusra gyakorolt hatásának részletes elemzésére
- Az informatika használata egyre szélesebb körben terjed el a társadalom minden működési területén és szerveződési szintjén
- A távérzékelési adatok felhasználása egyre nagyobb hangsúlyt kap a környezeti modellezésben. Ezek nagymértékben hozzájárulnak a hidrológiai paraméterek pontosabb leírásához (*Tamás, 2004*).

- A mederben és a vízgyűjtőn lejátszódó folyamatok (a hidrológiai folyamatok sztochasztikus jellege, a vízgazdálkodás és a földhasználat változásai, az erdőgazdálkodás során végbemenő intercepciós tározódási-kapacitás bővülése ill. csökkenése stb.) nagyban megnöveli az elemzések és előrejelzések térbeli és időbeli bizonytalanságát.
- Az „ismeretlen” bizonytalanság és a meglepetés a hagyományos mérnöki gyakorlattól idegen kezelési elveket követel: kulcsszerepet kap a megelőzés és az ökológiai rendszer visszacsatolásának beágyazása a tervezésbe és a működtetésbe (Istvánovics és Somlyódi, 2000).
- A vízgyűjtők integrált, fenntartható fejlesztésének kulcsa az anyagforgalom minél zártabbá tétele, az irreverzibilis veszteségek minimalizálása. Ahhoz, hogy ez megvalósítható legyen, olyan mérőszámokra, fenntarthatósági mutatókra van szükség, amelyekkel nyomon lehet követni a nagy területre kiterjedő és különböző ágazatok hatáskörébe tartozó beavatkozások hatásait (Bíró és Tamás, 2002).
- A környezeti problémák mellett a társadalmi-politikai-gazdasági kihívások is új igényekkel lépnek fel a környezeti modellekkel szemben. A véges vízkészletekkel való gazdálkodás fenntarthatósága, a vizes és vízi élőhelyek, a biodiverzitás védelme, az EU Víz-Keretirányelv szigorú direktívái felértékeltek a hidrológiai modellek jelentőségét is.
- A modell integrációk különböző fokai (adatcsere, közös interface, natív módon integrált rendszerek) képessé teszik a felhasználót arra, hogy a szükségleteknek megfelelő léptékben hozhasson létre valós idejű előrejelzési és ellenőrzési, illetve előrejelzési és tervezési modelleket (1. ábra).

A HEC-RAS program rövid áttekintése

A HEC-RAS (Hydrologic Engineering Center – River Analysis System) egy interaktív használatra kifejlesztett integrált szoftver-rendszer, amely egy több modulú felhasználó környezetben működik. A program kifejlesztője az US Army Corps of Engineers (USACE), Hydrologic Engineering Center (HEC) (Davis, California). Egydimenziós modell, amely a keresztiszelvények geokoordináta-helyes sorba rendezésével képes pseudo-3-D-s kép előállítására.

A HEC-RAS modell input alapadat-szükséglete:

- A meder geometriai adatai (keresztiszelvények, X-Y-Z EOY koordináták)
- A mederérdességi tényezők (Manning – féle n-koefficiensek)
- A kontrakciós és expanziós koefficiens megadása
- Határfeltételek megadása (permanens állapotra vízállás, vízhozam, vízhozam-görbe, vízfelszín-esés adatok, nem permanens állapotra észlelt és számított vízállás és vízhozam idősorok, vízhozam-görbe adatok)

A modell a határfeltételek megadása után a különböző vízhozamoknak és vízállásoknak megfelelően a következő adatokat szolgáltatja minden keresztiszelvényről :

- Az energiavonal magassága
- Sebességmagasság
- A vízfelszín magassága
- Kritikus magasság
- Az energiavonal lejtése
- Vízhozam
- Legnagyobb víztükörszélesség
- Átlagos vízsebesség
- A maximális vízhozam a keresztiszelvényben
- A meder max. mélysége

- A szakasz hossza (a soron következő keresztmetszelynyig)
- Minimális medermagasság
- Sűrűlódási energia-veszteség
- Kontrakciós és expanziós energia-veszteség
- Nyírófeszültség
- Nedvesített keresztmetszelynyi felület
- Nedvesített keresztmetszelynyi terület
- Hidraulikai mélység stb. (~250 adat) (Warner et al., 2002)

1. ábra:

Egy komplex környezeti modell kialakításának vázlata

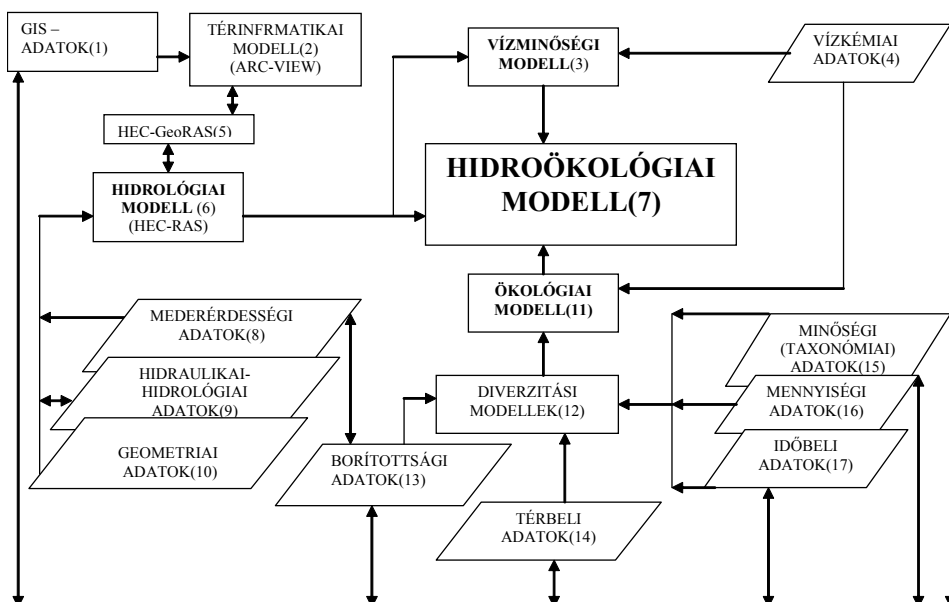


Figure 1: A possible environmental model construction (scheme)

GIS-data(1), GIS-model(2), Water-quality model(3), Water-chemical model(4), HEC-GeoRAS extension(5), Hydrological model(6), Hydroecological model(7), Manning-n data(8), Hydraulic-hydrological data(9), Geometric data(10), Ecological model(11), Diversity models(12), Covering data(13), Spatial data(14), Taxonomic data(15), Quantitative data(16), Temporal data(17)

Célkitűzések

1. A Bihari sík vízfolyásainak rendezése, és az azt kísérő vizes élőhelyek rehabilitálása sürgős feladat (2. ábra). A tervezés során alapvető fontosságú azoknak a medergeometriai és hidrológiai – hidraulikai jellemzőknek az ismerete, amelyek hatással lehetnek a vízi életközösségekre, és amelyek megváltoztatásával a legkedvezőbb eredményt érhetjük el (Spencer et al., 1998)

2. A környezeti állapot felméréséhez szükséges élettelen környezeti tényezők elemzésében nagy szerepet játszhatnak a számítógépes modellek, mivel sok időigényes és nehézkes terepi mérést válthatnak ki.

Jelen tanulmánynak az a célja, hogy felbecsülje ezeket a lehetőségeket.

2. ábra

A Bihari sík főbb vízfolyásai

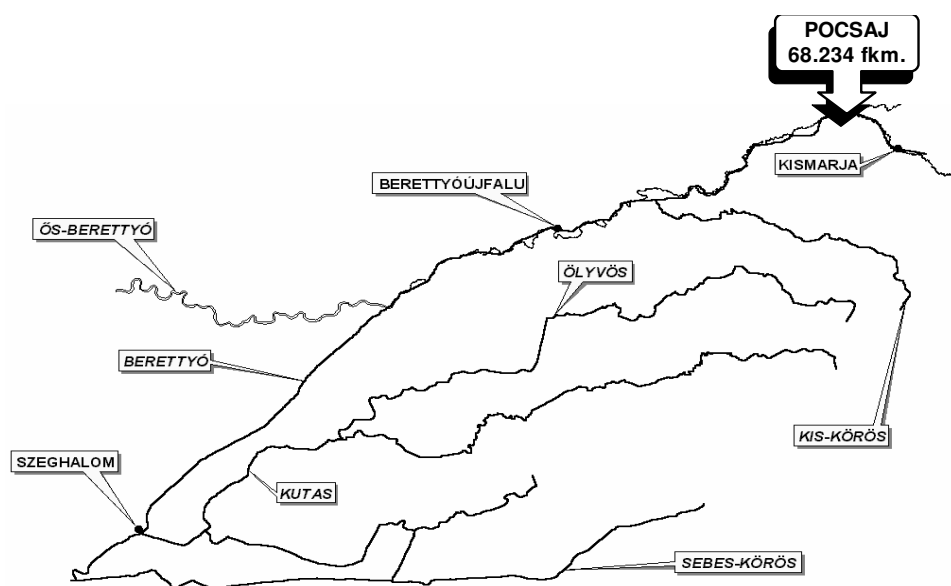


Figure 2: The major streams feeding the Bihari Plain

ANYAG ÉS MÓDSZER

A keresztshelvény hidraulikai jellemzőinek összefüggései

A munka során a 2005. február 12. és április 12. közötti vízhozamok modellezését végeztük el. A vizsgálati helyek közül a Pocsaji, 136. keresztshelvényre vonatkozó eredményeket ismertetjük (3., 4., 5. ábra).

Elsősorban azoknak a hidrológiai paramétereknek az összefüggéseit vizsgáltuk, amelyeket terepi körülmények között is megbízhatóan és pontosan lehet mérni, és amelyek kiindulásul szolgálnak a hidrológiai számításokhoz szükséges alapvető lezármaztatott jellemzők számára.

Mérhető alapadatok:

- Medergeometria
- A folyószakasz hossza (F_m)
- Vízükörszélesség (D)
- A meder legmélyebb pontja (mB_f)

- Vízsebesség (v)
- Vízükör magassága (B) (mBf)
- Mederérdesség (n)
- A vízszint (~energiavonal) esése (I)

3. ábra

**Vízhozamok a Berettyón (m^3/s) – Pocsaj,
2005. február 12. – április 12. (68.234 fkm, 136. VO).**

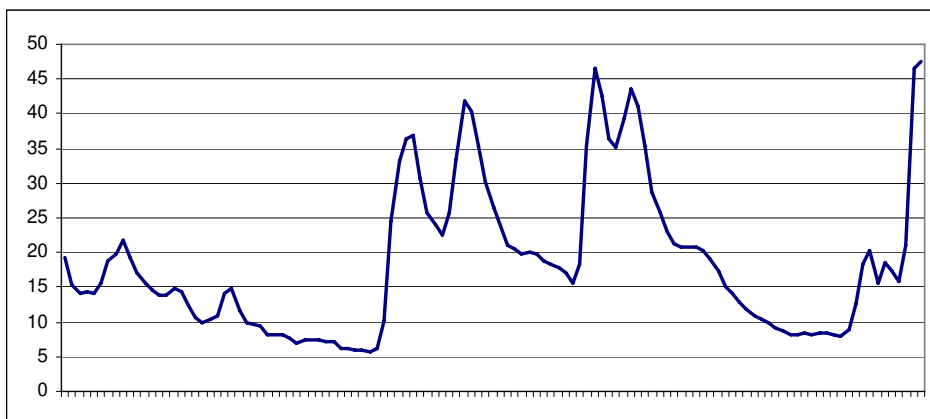


Figure 3: Average Run-off on the Berettyó, Pocsaj 12. February 2005. – 12. April 2005. 68.234 Running kilometer, 136. cross-section

4. ábra

**A legalacsonyabb és legmagasabb vízállások a Berettyón (m)
Pocsaj, 2005. február 12 – április 12, (mBf; 68.234 fkm, 136. VO).**

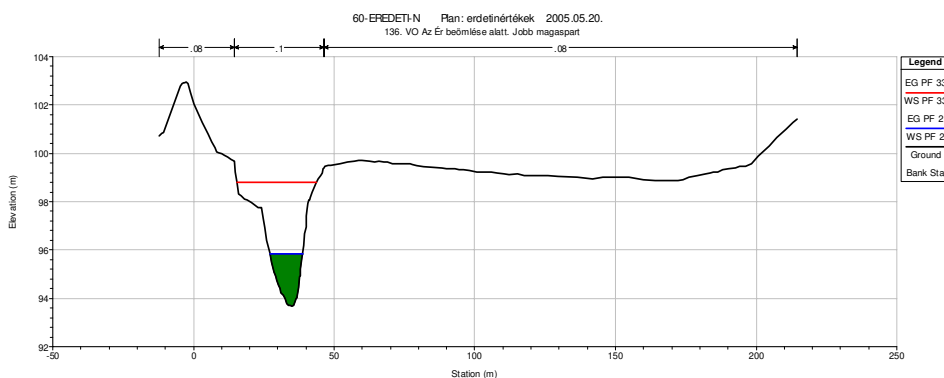


Figure 4: The 136. cross-section (max. and min. stage) 12. February, 2005. – 12. April, 2005. 68.234 Running kilometer, Pocsaj

5. ábra

A Berettyó magyarországi szakaszának hosszszelvénye (y= mBf, x=fm)

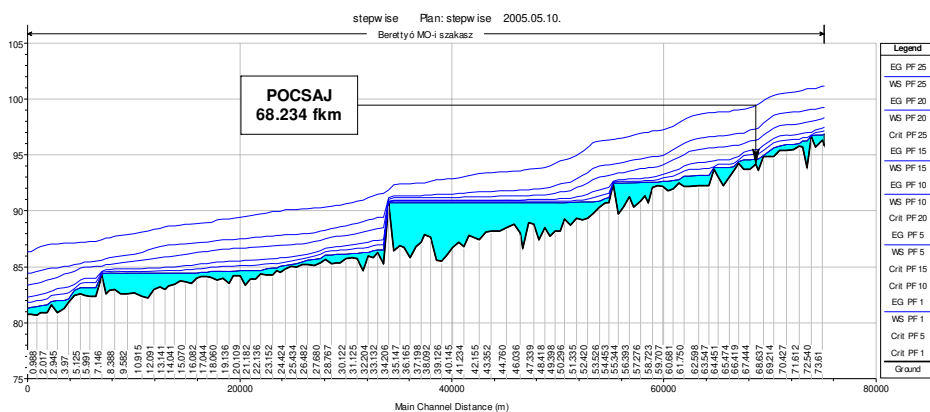


Figure 5: Longitudinal section of the hungarian stage of the Berettyó River

Számítható alapadatok

- Nedvesített keresztzeti terület (A)
- Nedvesített keresztzeti terület = periméter (P)

Leszármaztatott mennyiségek

- Vízhozam (Q) $Q = v \times A$
- Hidraulikai sugár (R) $R = \frac{A}{P}$
- Hidraulikai mélység (L) $L = \frac{A}{D}$
- A vízszint (~energiavonal) esése (I) $I = \frac{\Delta B}{\Delta Fm}$
- Froude-szám (F) $F = \frac{v^2}{2g}$
- Reynolds-féle szám (Re) $Re = \frac{vl}{v(nü)}$
- Átlagos sebesség $v = C\sqrt{R \times I}$ illetve $v = \frac{1}{n} \times \sqrt[3]{R^2 \times \sqrt{I}}$
- Átlagos sebesség $n = \frac{1}{v} \times \sqrt[3]{R^2 \times \sqrt{I}}$
- Manning-féle mederérdesség (n) $C = \frac{1}{n} \times \sqrt[6]{R}$
- Chézy-féle sebességállandó (C)

A mederérdességek megállapítása

A modellezés legtöbb problémával járó input adata a mederérdesség. Vagy indirekt módon határozhatjuk meg, vagy tapasztalat alapján, táblázat segítségével becsülhetjük

meg (1. táblázat). A mederérdességek értékei rendkívül változatosak lehetnek, és nagymértékben függenek a legkülönbözőbb környezeti tényezőktől, pl. a mederfelszín érdességétől, a meder benőttségétől, a meder vonalvezetésétől, a mederben elhelyezkedő természetes ill. mesterséges akadályoktól és műtárgyaktól, a leülepedett és a lebegtetett hordalék mennyiségi és minőségi jellemzőitől, a hőmérséklettől stb.

1. táblázat

A Manning-féle mederérdességi tényezők értékei néhány különböző medertípus esetében

A meder típusa és jellemzése(1)		Manning-féle mederérdességi értékek (2)		
		Min.	Norm.	Max.
A. Természetes vízfolyások(3)				
1. Főmeder(4)				
a.	Tiszta, egyenes, telt, nincsenek árkok és mély gödrök(5)	0.025	0.030	0.033
b.	Ua. mint fent, kövekkel és növényzettel(6)	0.030	0.035	0.040
c.	Tiszta, kanyargós, gödrökkel és zátonyokkal(7)	0.033	0.040	0.045
d.	Ua. mint fent, kövekkel és növényzettel(8)	0.035	0.045	0.050
e.	Ua. mint fent, alacsonyabb lépcsőkkel, több inefektív /vízhozam szempontjából közömbös/ lejtővel és szelvénnel(9)	0.040	0.048	0.055
f.	Ua. mint "d", kövekkel(10)	0.045	0.050	0.060
g.	Lassú folyású, benőtt ágak, mély gödrökkel(11)	0.050	0.070	0.080
h.	Erősen benőtt ágak, mély gödrök vagy elágazások, sűrűn álló fatörzsekkel és bozóttal(12)	0.070	0.100	0.150
2. Árterek(13)				
a.	Nyílt síkság (14)			
	1. Alacsony gyepek (15)	0.025	0.030	0.035
	2. Magas gyepek (16)	0.030	0.035	0.050
b.	Művelt terület (17)			
	1. Növények nélkül (18)	0.020	0.030	0.040
	2. Kifejlett kapásnövények, sorokban (19)	0.025	0.035	0.045
	3. Kifejlett szántóföldi növények (20)	0.030	0.040	0.050
c.	Bozót (21)			
	1. Szórványos bozót, sűrű aljnövényzet (22)	0.035	0.050	0.070
	2. Gyér bozót és facsoport, télen (23)	0.035	0.050	0.060
	3. Gyér bozót és facsoport, nyáron (24)	0.040	0.060	0.080
	4. Közepesen sűrű bozót, télen (25)	0.045	0.070	0.110
	5. Közepesen sűrű bozót, nyáron (26)	0.070	0.100	0.160
d.	Fák (27)			
	1. Kitisztított talaj törzsekkel, sarjadékok nélkül (28)	0.030	0.040	0.050
	2. Ua. mint fent, sűrű sarjadékokkal (29)	0.050	0.060	0.080
	3. Sűrűn álló fatörzsek, kidőlt fák, az áramlás a koronaszint alatt (30)	0.080	0.100	0.120
	4. Ua. mint fent, az áramlás a koronaszintet eléri (31)	0.100	0.120	0.160
	5. Sűrű füzes, nyáron, egyenes vonalban (32)	0.110	0.150	0.200

Forrás (Source): Open-Channel Hydraulics, Chow, 1959

Table 1: Manning „n” values of some types of channels

Type of Channel and Description(1), Manning „n” values(2), Natural Streams(3), Main Channels(4), Clean, straight, full, no rifts or deep pools(5), Same as above, but more stones and weeds(6), Clean, winding, some pools and shoals(7), (Same as above, but some weeds and stones(8), (Same as above, lower stages, more ineffective slopes and sections(9), (Same as "d" but more stones(10), Sluggish reaches, weedy. deep pools(11), Very weedy reaches,

deep pools, or floodways with heavy stands of timber and brush(12), Flood Plains(13), Pasture no brush(14), Short grass(15), High grass(16), Cultivated areas(17), No crop(18), Mature row crops(19), Mature field crops(20), Brush(21), Scattered brush, heavy weeds(22), Light brush and trees, in winter(23), Light brush and trees, in summer(24), Medium to dense brush, in winter(25), Medium to dense brush, in summer(26), Trees(27), Cleared land with tree stumps, no sprouts(28), Same as above, but heavy sprouts(29), Heavy stand of timber, few down trees, little undergrowth, flow below branches(30), Same as above, but with flow into branches(31), Dense willows, summer, straight(32)

A becslések pontosítása érdekében elengedhetetlen a rendszeres terepi bejárás, a műholdas felvételek, és egyéb távérzékelési adatok elemzése.

A hidrológiai adatok összefüggéseit kétmintás-t próbával, valamint többváltozós lineáris regressziós analízis (stepwise analysis), nem lineáris összefüggések esetében a regressziós görbék becslésének segítségével elemeztük. A számításokhoz az SPSS 12.0 statisztikai szoftvert alkalmaztuk. A terepi mérések szempontjaihoz alkalmazkodva független változókként elsősorban a mérhető alapadatokat, míg függő változókként a lezármaztatott mennyiségeket kezeltük.

Gyakorlati szempontból a vízhozam független változóként ill. alapadatként is szerepelhet, mivel a vízügyi igazgatóságok az országos vízügyi adatbázisban a mért vízállás idősorok mellett a számított vízhozam idősorokat is közzéteszik.

A nyírófeszültség és a vízáramlás jellegének megállapítása

A vízrendezési és természetvédelmi célú beavatkozások szempontjából alapvető fontosságú a nyírófeszültség és a vízáramlás jellege, amelyek a modell adataiból és a terepi mérésekből egyaránt számíthatóak a következő egyenletek szerint:

$$\text{Nyírófeszültség: } \tau = \eta \frac{dv}{dy}$$

ahol

τ = nyírófeszültség (Nm^{-2})

η = dinamikai viszkozitási tényező (Nm^{-2}s)

v = sebesség

y = a sebesség irányára merőleges hossz

Turbulens ill. lamináris áramlás meghatározása a Reynolds-szám (Re) alapján: $Re = \frac{vR}{\nu}$

turbulens – $Re > 2320$

lamináris – $Re < 2320$

ahol

ν = sebesség

ν = kinematikai viszkozitás (20°C -on $10^{-3} \text{ m}^2/\text{s}$)

R = hidraulikai sugár

Rohanó ill. áramló áramlás meghatározása: $v_{grav} = \sqrt{gh}$

rohanó: $V_{v\acute{z}} > V_{grav}$

áramló: $V_{v\acute{z}} < V_{grav}$

ahol

V_{grav} = a gravitációs hullámok sebessége

g = grav. állandó

h = vízmélység

$V_{v\acute{z}}$ = aktuális vízsebesség

EREDMÉNY ÉS ÉRTÉKELÉS

A keresztszelvényben a mért és a modell által számított vízmagasságok időszora 0.05 szignifikanciaszint mellett azonosnak tekinthető, így a napi vízállás adatokból, és az azok alapján számított vízhozamokból megbízható közelítéseket lehet tenni az egyéb hidraulikai jellemzőkre is (6. ábra). Példánkban a hiteles vízmerce adatai és a modell által szolgáltatott adatok összefüggéseit mutatjuk be (2. táblázat).

6. ábra

Vízállások a Berettyón
(Pocsaj, 2005. február 12 – április 12, mBf m, 68.234 fkm, 136. VO).

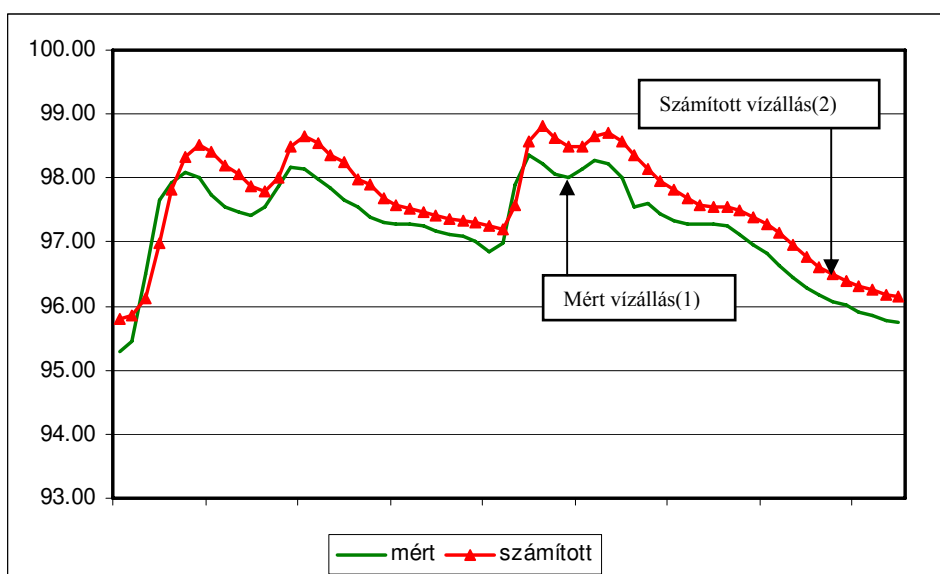


Figure 6: Stages of hydrograph, Berettyó, Pocsaj 12. February 2005. – 12. April 2005. 68.234 Running kilometer, 136. cross-section

Measured stages(1), Calculated stages(2)

A determinációs együtthatók értékéből látszik, hogy a vízállásadatok ismeretében elég nagy pontossággal lehet következtetni bizonyos hidraulikai jellemzőkre. Ez elegendőnek látszik az árvízi becslésekre és a vízrendezési beavatkozások tervezésére. Figyelemre méltó, hogy a vízszintesítés és a vízállás esetében milyen alacsony a determinációs együttható értéke. Ez arra utal, hogy a kis vízszintesítés (a magyarországi szakaszon ~ 0.0002) nem játszik nagy szerepet a vízhozam és vízsebesség alakulásában, lényeges az egyéb hidrológiai tényezők (mederérdesség, nyíróerők stb.) szerepe is.

A hidrobiológiai kutatásokban is jól felhasználható adatok számítása érdekében a modell további fejlesztése szükséges.

2. táblázat

A jellemző hidraulikai paraméterek összefüggései a Berettyó 136. keresztszelvényében

Függő változó (1)	Független változó: mért tengerszint feletti magasság mBf (m) (2)	R² (3)
Számított mBf (m) (4)	$y=1,14+0,99x$	0.91
Vízsebesség (5)	$y=7,4 \times 10^{-10} \times e^{0,21x}$	0.72
Vízhozam (6)	$y=1,5 \times 10^{-27} \times e^{0,67x}$	0.98
Periméter (7)	$y=6,1 \times 10^{-13} \times e^{0,32x}$	0.81
Víztükrörszélesség (8)	$y=5,4 \times 10^{-13} \times e^{0,32x}$	0.77
Vízzintesés (9)	-	0.32
Nedv. ker. terület (10)	$y=2 \times 10^{-18} \times e^{0,46x}$	0.92

Table 2. Relationships of the hydrological characteristics at the 136. cross-section of Berettyó River

Dependent Variable(1), Independent Variable: Water Surface Elevation(2), Regression Square(3), Computed Water Surface Elevation(4) Average Velocity(5), Flow(6), Wetted Perimeter(7), Max. Depth(8), Energy Gradeline Slope(9), Wetted Flow Area(10)

KÖVETKEZTETÉSEK

A hidrológiai modell segítségével meg lehet becsülni az általunk kiválasztott vizsgálati helyek hidraulikai paramétereit, de a helyszínen végzett mintavételeket és méréseket nem pótolja. A digitális és terepi módszerek együttesen viszont eredményesen alkalmazhatóak a környezeti állapot értékelésében.

A modell fejlesztésével és megbízhatóságával kapcsolatos jövőbeli feladatok:

- Bővíteni az eddigi 148 keresztszelvény számát
- Pontosítani a földrajzi és geometriai adatokat
- Fel kell becsülni a műtárgyak hidraulikai hatásait
- A beömlő vízfolyások bevonása a modellbe
- A mederérdességi tényezők pontosabb, a szezonális és vízhozam változásokat is tükröző felmérése
- A távérzékelési módszerek alkalmazása a mederérdesség meghatározásába
- Több, hitelesített vízmérce létesítése a rendszer csomópontjaiban

IRODALOM

- Bíró, T., Tamás, J. (2002). Vízfolyások vízrajzi adatbázisa és hidrodinamikai modellezése. In: Proc. Informatika a Felsőoktatásban Konferencia. Debrecen, 2002. augusztus 28-30. 259. CD.
- Chow, V.T. (1959). Open-Channel Hydraulics. McGraw-Hill, Inc. Singapore. International Edition, 1973. 98-114.
- EU Víz Kereterv (2000/60/EK)
- Istvánovics, V., Somlyódy, L. (2000). Az ökológia, a természetvédelem és a vízgazdálkodás kapcsolata. In: Vízügyi Közlemények, 2000/3-4. Budapest. 525-551.
- Európai Nitrát Irányelv (91/676/EGK)

- Somlyódy, L. (2000). A hazai vízgazdálkodás és stratégiai pillérei. In: Vízügyi közlemények, 2000/3-4. Budapest. 376-417.
- Spencer, C., Robertson, A.I., Curtis, A. (1998). Development and testing of a rapid appraisal wetland condition index in south-eastern Australia. In: Journal of Environmental Management 54. 143-159.
- Tamás, J. (Szerk.) 2004: Vízkészlet-modellezés. Egyetemi tankönyv. Debreceni Egyetem. 4-6.
- Tamás, J., Bíró, T. (2001). Határvizek -különös tekintettel a Berettyó folyó vízgyűjtőjére- vízminőségi célállapota és ökológiai vízigénye meghatározásának tudományos megalapozása. In: KöM tudományos eredmények. 51.
- Warner, J., Brunner, G., Wolfe, B., Piper, S. (2002). User's Manual. HEC-RAS River Analysis System Version 3.1. (2002) US Army Corps of Engineers, Institute for Water Resources Hydrologic Engineering Center

Levelezési cím (*Corresponding author*):

Pregun Csaba

Debreceni Egyetem, Agrártudományi Centrum, Mezőgazdaságtudományi kar
Víz- és Környezetgazdálkodási Tanszék
4032 Debrecen, Böszörményi út 138.

*University of Debrecen, Centre of Agricultural Sciences, Faculty of Agriculture
Department of Water- and Environmental Management
H-4032 Debrecen, Böszörményi street 138.*

Tel.: 36-52-508-444/8250

e-mail: cpregun@gissserver1.date.hu



Az üledék-toxicitás térbeli eloszlásának vizsgálata a Kis-Balaton Vízhatal Rendszer II. ütemének területén

Kováts N.¹, Borbély G.¹, Magyar I.¹, Szép Zs.¹, Paulovits G.²,
Pomogyi P.³

¹Veszprémi Egyetem Környezetmérnöki és Kémiai Technológia Tanszék, 8200 Veszprém Egyetem u. 10

²MTA Balatoni Limnológiai Kutatóintézet, 8237 Tihany Klebersberg Kunó u. 3

³Nyugat-dunántúli Vízügyi Igazgatóság Kis-balatoni Üzemeltetés, 8360 Keszthely Csik Ferenc sétány 1.

ÖSSZEFOGLALÁS

1863 előtt a Kis-Balaton a Balaton előtt természetes szűrőmezőt képezett, amely a Zálával érkező tápanyag jelentős részét visszatartotta. 1863-ban a Sió-zsilip megépítésével megeremlődtek a feltételek a Balaton vízszint-szabályozásához. A vízszint szabályozása és alacsony szinten tartása következtében a Kis-Balaton mocsarai összezsugorodtak, a szűrőmező funkciója csökkent. Mindennek következményeképpen a tó vízminősége az 1960-as évektől kezdődően megnőtt tápanyagterhelés hatására romlásnak indult, az eutrofizáció felgyorsult, 1970 után a tó eutróf-hipertróf állapotba jutott. A Kis-Balaton Vízhatal Rendszer (KBVR) fő célja a Balaton tápanyagterhelésének csökkentése volt, két tározórész segítségével. Bár a rendszer kialakításának elsődlegesen vízminőségvédelmi okai voltak, a terület egyben jelentős természetvédelmi értéket képvisel. 1979-ben felkerült a Ramsar Vizes Élőhelyek listájára. A KBVR hatékonyságát, a bemenő és kimenő vízminőségi mutatókat rendszeres monitoringvizsgálat keretében figyelik. Nincs lehetőség viszont arra, hogy a diffúz szennyezőforrások (pl. mezőgazdasági területek) által okozott terhelést kimutassák. Célunk annak a vizsgálata volt, hogyan alakul az 1992-ben részlegesen elárasztott II. tározó üledékének szennyezettsége a tározót érő külső terhelés következtében, részletesen elemezve a szennyezettség térbeli eloszlását, az eloszlásban mutatkozó trendeket. Az üledékszennyezettség térképezéséhez a Nyugat-dunántúli Vízügyi Igazgatóság Kis-balatoni Üzemeltetésétől származó digitális ortofotót használtuk. (Kulcsszavak: Kis-Balaton Vízhatal Rendszer, üledéktoxicitás, ToxAlert, kockázattérképezés)

ABSTRACT

Analysis of spatial distribution of sediment toxicity in the 2nd reservoir of the Kis Balaton Water Protection System

N. Kováts¹, G. Borbély¹, I. Magyar¹, Zs. Szép¹, G. Paulovits² and P. Pomogyi³

¹University of Veszprém, Dept. of Environmental Engineering and Chemical Technology H-8200 Veszprém Egyetem Str. 10

²Balatoni Limnological Research Institute of the Hungarian Academy of Sciences H-8237 Tihany Klebersberg Kunó Str. 3

³West Transdanubian Water Authority, Dept. Kis-Balaton, H-8360 Keszthely Csik Ferenc Str. 1

Prior to 1863 the Kis-Balaton functioned as a natural filter zone, retenting most of the nutrients carried by River Zala. In 1863 by the construction of the Sió sluice control of the water level of Lake Balaton became feasible. As a consequence, marshlands of Kis-Balaton

began to disappear and the filtering function started to degrade. Due to these facts water quality of the lake turned into eutrophic by the 1960's. Main function of the Kis-Balaton Water Protection System is the protection of the water quality of Lake Balaton by retaining most of the nutrients and suspended solids carried by the River Zala and other small watercourses in two reservoirs. Besides this natural cleaning function, however, Kis-Balaton is also a significant nature conservation area under the Ramsar Convention since 1979. Efficiency of the System is continuously monitored. However, load presented by diffuse pollution sources such as agricultural runoff cannot be determined. Our goal was to reveal how ecological risk in the stage 2 is determined by external sources, analysing spatial distribution and trends in contamination. For mapping contamination digital orthophoto provided by the Western-Transdanubian Water Authority was used.

(Keywords: Kis-Balaton Water Protection System, sediment toxicity, ToxAlert, risk mapping)

BEVEZETÉS

A Kis-Balaton története szorosan összefügg a Balaton történetével, mely világviszonylatban is az egyik legjobban kutatott tavak egyike. Az ősi Balaton vízszintje magasabb volt a mainál. Az előfordult legmagasabb vízszintet 113.5 mAf értékre becsülik (Bendefy, 1968), a jelenlegi mintegy 104,8 m-rel szemben.

A Sió szabályozása 1831-ben kezdődött, ekkor a Balaton vízszintje 1 méterrel csökkent, és a Kis-Balaton partvonala is beljebb húzódott. 1863-ban készült el a Sió-zsilip, mely a Balaton vízszint szabályozásához teremtette meg a feltételeket. A szabályozás nem csak a Balatonban eredményezett még alacsonyabb vízszintet, de a kis-balatoni mocsarak is összezsugorodtak, szinte teljesen kiszáradtak. 1922-ig a Zaláról érkező árvizek miatt még időszakosan nyílt vízfelszínek jelentek meg. Ezt követően a folyómeder kotrási munkálatai mellett árvízvédelmi töltéseket is építettek, a Kis-Balaton sorsát ezzel végleg megpecsételve. Csupán fél négyzetkilométeres szabad víztükör maradt, a többit elborították a vízinövények (Fekete et al., 1991).

1910-ben már Rieger Antal azt javasolta, hogy a volt Kis-Balaton területén található mocsarat árásszák el, ezzel tisztítva a Balatonba jutó vizet. Cholnoky Jenő 1942-ben újra felvetette ezt az ötletet. Azonban még hosszú ideig csak a viták folytak erről a lehetőségről, és eközben a Balaton vízminősége folyamatosan romlott. A 60-as években a tó vizének eutrofizációja következtében szükségessé vált a probléma kezelése (Magyarics et al., 1999).

A KBVR kiépítésére azért került sor, hogy a Balatonba jutó szennyező anyagokat, foszfor és lebegőanyag terhelést csökkentsék és a tóban az eutrofizálódást megállítsák. Ezt úgy képzelték el, hogy a területen található nádasok és sásosok természetes szűrő rendszert képezve, átlagosan kb. másfél hónap tartózkodási idő alatt (amit a terelőtöltések biztosítanak), felveszik a többlet tápanyagot. Ehhez a területet úgy kellett kialakítani, hogy a víztisztítási folyamatokban szerepet játszó fajok ökológiai igényeit is kielégítsék.

A Kis-Balaton Védőrendszer (KBVR) kialakítását két lépcsőben valósították meg. A Felső (Hídvégi-tó) tározó 1981-85 között került kiépítésre. Itt két különálló teret alakítottak ki, a fő tározót és a Kazettát. Ez utóbbi a havária jellegű szennyezések „csapdázására” 5 millió m³ kapacitással rendelkezik. A Kis-Balaton ma Fenéki-tó néven említett tározója (valójában az eredetileg tervezett II. ütem egy része) 1992 őszén került elárasztásra.

1986 és 1997 között az I. ütem, a Hídvégi-tó mintegy 78000 t lebegőanyagot, 290 t összes foszfort (TP), 250 t ortofoszfát-foszfort (PO₄-P) és 800 t összes nitrogént (TN)

tartott vissza. A II. ütem, a Fenéki-tó az I. ütemről érkező lebegőanyag mintegy 75%-át távolított el (Tátrai *et al.*, 2000).

Elsődleges célunk annak a vizsgálata volt, hogy a visszatartott szennyezés milyen kockázatot jelent a tározótérre, pontosabban a II. ütemre. Ennek elemzése már csak azért is fontos, mivel a Kis-Balaton vízminőségvédelmi Rendszer természetvédelmi szempontból kiemelt terület, elsősorban madárfaunája érdemel figyelmet (a Kis-Balatont 1979-ben Ramsar-területté nyilvánították).

ANYAG ÉS MÓDSZER

Mintavétel

A minták a Kis-Balaton II. ütemének területéről, valamint az előtte és az utána lévő Zala szakaszból származnak. A 2003. októberi mintavétel során Hargrave-mintavevőt, míg 2004. szeptemberében és októberében csőmintavevőt használtunk, az üledék felső 30 cm-es rétegéből vettünk mintát. A mintavételi helyek pontos meghatározása e-Trex Vista GPS készülékkel történt.

Az üledékminták előkészítése a vizsgálatokhoz

A minták előkészítése az MSZ 21470/2-81 (1982) szabvány alapján történt. Az ökotoxikológiai tesztek elvégzéséhez kivonatot készítettünk, kivonószerként 0,1%-os dimetil-szulfoxidot (DMSO) használtunk.

Toxicitásvizsgálat

A *Vibrio fischeri* (régebbi nevén *Photobacterium phosphoreum*) baktérium által kibocsátott fény (biolumineszcencia) gátlása az alapelve annak a tesztrendszernek, amelynek több kereskedelembe hozzáférhető változata létezik (a Merck által gyártott ToxAlert, az Azur által fémjelzett Microtox, ill. a LUMISTox). A mérgező anyag változásokat idéz elő a sejtalkotók állapotában – sejtfal, sejtmembrán, az elektrontranszport-rendszer, enzimek, a citoplazma alkotói – amelyek a biolumineszcencia csökkenésében mutatkoznak meg.

A tesztet bár eredetileg elsősorban szennyvíz minősítésére dolgozták ki, sikeresen alkalmazták szennyezett üledékek toxicitásának értékelésére is (pl. Bennett and Cubbage, 1992; Svenson *et al.*, 1996; Johnson and Long, 1998). Doherty (2001) szerint a teszt során kapott eredmények korrelálnak más, az üledék minősítésére szolgáló ökotoxikológiai tesztek eredményeivel.

A ToxAlert®100 luminométer minden, a toxicitás értékeléséhez szükséges adatot automatikusan kiszámít. Először az f_{kt} korrekciós faktort számolja ki a mért lumineszcencia-intenzitásból ([1] egyenlet).

$$f_{kt} = I_{kt} / I_0 \text{ (t=30 perc)} \quad [1]$$

ahol:

f_{kt} az expozíciós (kontakt) időre megadott korrekciós faktor

I_{kt} lumineszcencia intenzitás a kontrollban, RLU-ban (relative luminescence units) mérve, az expozíciós idő után

I_0 a kontroll tesztszuspenzió lumineszcencia intenzitása közvetlenül a hígító (2%-os NaCl) oldat hozzáadása előtt.

A korrekciós faktor alkalmazásával az egyes tesztminta küvetákra kiszámolja I_0 korrigált értékét ([2] egyenlet).

$$I_{ct} = I_0 \times f_{kt} \quad [2]$$

ahol:

f_{kt} az f_{ikt} értékek átlaga

l_0 a kontroll tesztszuszpenzió lumineszcencia intenzitása közvetlenül a hígító (2%-os NaCl) oldat hozzáadása előtt

l_{ct} az l_0 korrigált értéke a tesztminta küvetákra közvetlenül a tesztminta hozzáadása előtt.

Ezek után a tesztminta H_t inhíbeáló effektusát számítja ki ([3] egyenlet).

$$H_t = [(l_{ct} - l_{Tt} / l_{ct})] \times 100 \quad [3]$$

ahol

H_t a tesztminta inhíbeáló effektusa az expozíciós idő után, %-ban megadva

l_{ct} az l_0 korrigált értéke a tesztminta küvetákra közvetlenül a tesztminta hozzáadása előtt

l_{Tt} a tesztminta lumineszcencia intenzitása az expozíciós idő után.

Toxikusnak a 20% feletti gátlást mutató mintákat tekintjük (Suter, 1996).

Az eredmények térbeli értékelése

Alaptérképként a Nyugat-Dunántúli Vízügyi Igazgatóság által rendelkezésünkre bocsátott digitális ortofotó szolgált. A térinformatikai feldolgozáshoz az ArcView 3.2 verzióját, a szennyezettség térbeli eloszlásának feltüntetésére a Golden Software által kifejlesztett Surfer8 kontúrozó és felületszerkesztő szoftvert használtuk.

EREDMÉNY ÉS ÉRTÉKELÉS

Az 1. táblázat tartalmazza a 2003-ban, a Kis-Balaton II. fokozat elárasztott területén (Ingói-berek) vett minták koordinátáit, a 15 és 30 perc expozíciós idővel mért valamint az átlagolt toxicitás értékeket. A 2004. folyamán vett minták értékei a 2. és 3. táblázatokban láthatók.

1. táblázat.

Mért toxicitás értékek a KBVR II. ütemén, Ingói-berek, 2003. X. 16.

<Ht>: a tesztmintában tapasztalt gátlás az expozíciós idő után, %-ban megadva

Jel	Koordináták		15 min	30 min
			<Ht>	<Ht>
	EOV Y	EOV X	%	
1	507690,46	150477,90	14,55	11,70
2	507878,57	150182,47	61,60*	67,50*
3	508056,50	150102,26	25,85	21,15
4	508078,03	149934,96	20,00	13,60
5	508087,13	149565,95	19,80	14,00
6	508093,19	149284,12	14,60	13,85
7	508169,47	149111,82	17,95	14,20
8	508526,73	148953,24	37,15*	34,35*

* 20%-os toxicitást meghaladó minták (toxicity values exceeding 20%)

Table 1: Calculated toxicity values in the II. reservoir of the KBWPS, 16.10.2003.
<Ht>: inhibitory effect in the test sample after the exposure, in %

A KBVR II. ütem területén 2004. őszén két mintavételezés történt. Az első esetben 7 minta került megvételre. A második alkalommal 6 minta begyűjtésére került sor. Mint ahogy a koordinátákból is látható, a mérési pontok az egészhez képest túl kicsi területet fednek le ahhoz, hogy azokból egy eloszlástérképet legyen érdemes szerkeszteni. Ezeknek a mintavételi pontoknak az esetében egy-egy mintavételi terület pontjainak adataiból számított toxicitás-értékeket átlagoltuk, ezeket az 1. ábrán tüntettük fel.

2. táblázat

Mért toxicitás értékek a KBVR II. ütemén, 2004. IX. 20.

<Ht>: a tesztmintában tapasztalt gátlás az expozíciós idő után, %-ban megadva

Jel (1)	Koordináták (2)		15 min	30 min
			<H _t >	<H _t >
	EOV Y	EOV X	%	
1	513173,85	152063,99	63,10	58,85
2	513167,26	152054,88	72,30	65,95
3	513108,65	152019,16	69,35	62,35
4	513098,32	151995,30	73,85	67,95
5	513079,55	151899,37	73,50	66,80
6	513052,19	151762,86	78,55	71,85
7	513127,67	152107,67	71,80	63,85

Table 2: Calculated toxicity values in the II. reservoir of the KBWPS, 20.09.2004. <Ht>: inhibitory effect in the test sample after the exposure, in %

Signal(1), Coordinates(2)

3. táblázat

Mért toxicitás értékek a KBVR II. ütemén, 2004. X. 18.

<Ht>: a tesztmintában tapasztalt gátlás az expozíciós idő után, %-ban megadva

Jel (1)	Koordináták (2)		15 min	30 min
			<H _t >	<H _t >
	EOV Y	EOV X	%	
1	506783,80	145104,87	74,95	72,15
2	506719,18	145124,95	75,15	73,10
3	507813,60	145358,27	53,95	50,25
4	507809,47	145291,66	67,70	63,35
5	507733,32	145310,15	55,50	51,10
6	511479,89	150051,57	75,15	69,70

Calculated toxicity values in the II. reservoir of the KBWPS, 18.10.2004. <Ht>: inhibitory effect in the test sample after the exposure, in %

See Table 2

1. ábra

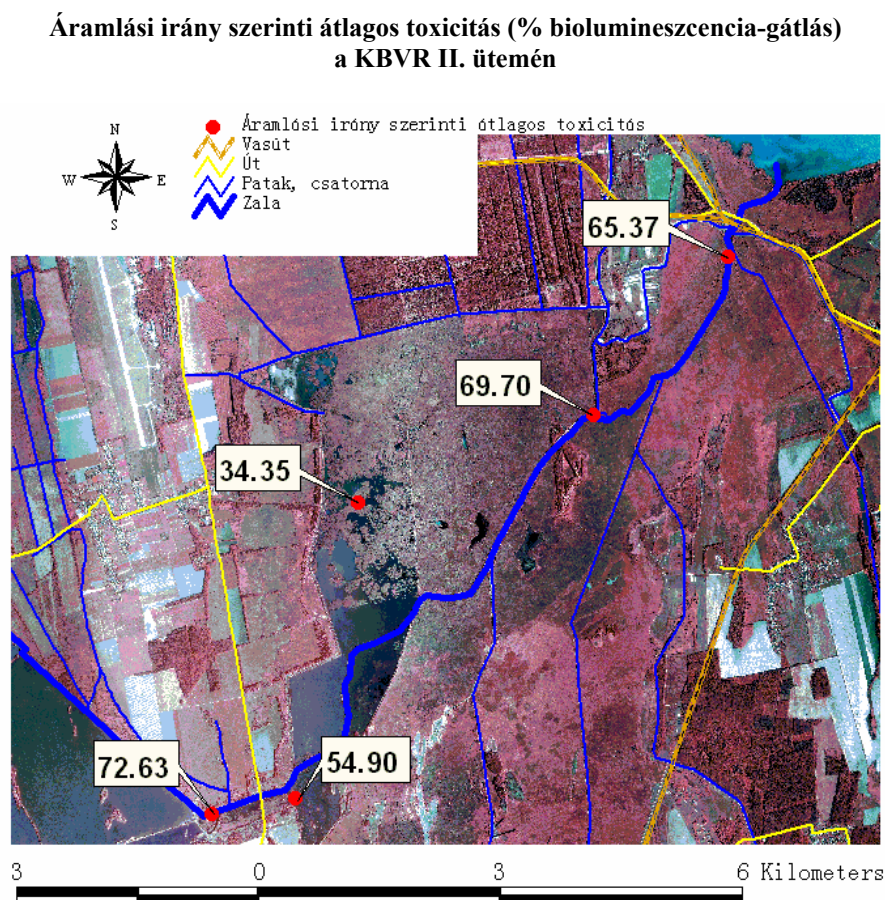


Figure 1: Average toxicity values (%bioluminescence inhibition) along water direction in the 2nd reservoir of the KBWPS

Az Ingói-berek 2003-as vizsgálati eredményeiből viszont létrehozható értékelhető eloszlástérkép (2. ábra). Ez esetben fontos, hogy az ábra értelmezésekor vegyük figyelembe, hogy nem csak az Ingói-berek nyílt vizét ábrázolja, hanem az azt Ny-ról határoló erdősávot is, illetve az attól szintén Ny-ra lévő vizenyős terület egy részét is. Nem tudtuk csak és kizárólag az Ingói-berek nyílt vizes területét ábrázolni, mivel a Surfer a vizsgált terület koordinátaminimumaira és -maximumaira fektetett téglalappal dolgozik.

Ha megnézzük az 1. ábrát, azt látjuk, hogy a terület Hidvégi-tó felőli végétől a kelet felé haladva – tulajdonképpen a folyásirányt követve – a toxicitás csökken. E tendencia két helyütt látszik megszakadni:

- Az Ingói-berekben a sármelléki repülőtér és környékének kisvízeinek betorkollási pontja környékén, és
- A Hévíz-Páhoki csatorna betorkollásánál.

2. ábra

**Toxicitás-eloszlás (% biolumineszcencia-gátlás) az Ingói-berek Ny-i oldalán,
2003. X. 16. a.) t=15 min b.) t=30 min expozíciós idő után**

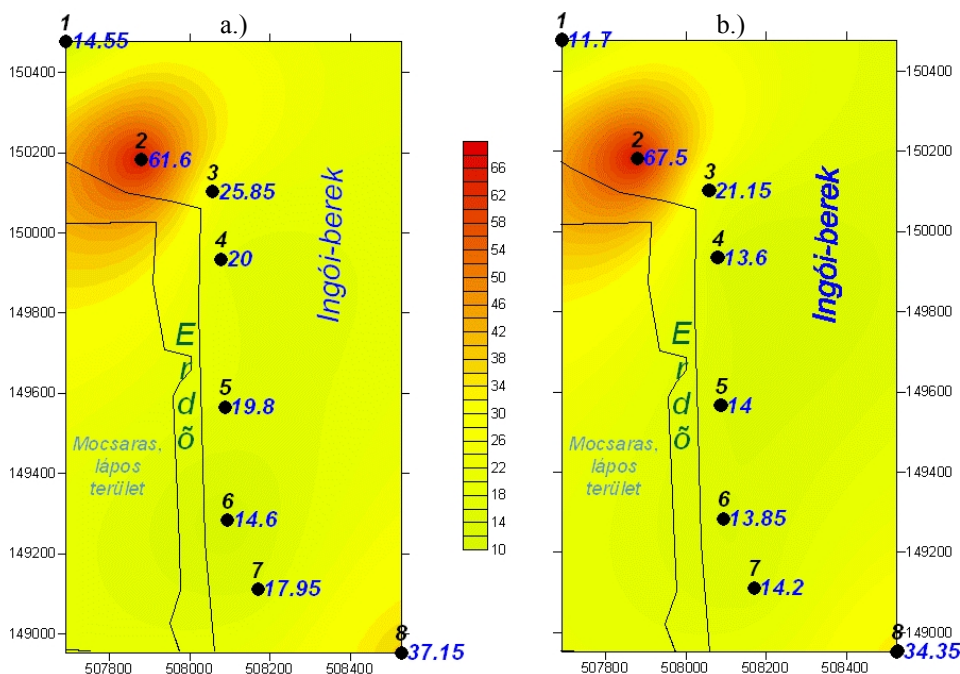


Figure 2: Distribution of toxicity (% bioluminescence inhibition) in the west of Ingói-berek (16.10.2003) a.) t=15 and b.) 30 min exposure

Azt láthatjuk, hogy folyásirány szerint a Hídvégi-tó végénél az iszap meglehetősen toxikus. A Fenéki-tóban aztán erősen csökken az üledék toxicitása, egészen addig, amíg a Hévíz-Páhoki-csatorna a Zala medrébe nem torkollik. Ez a toxicitás jelentős emelkedését okozza; az üledék toxicitása ismét majdnem olyan magas értéket ér el, mint közvetlenül a Hídvégi-tó után. Innen a Balatonig a toxicitás már nem csökken jelentősen.

KÖVETKEZTETÉSEK

Az általunk vizsgált 3 mérési sorozat alapján elmondható, hogy az üledék toxicitása, melyet a biolumineszcencia-gátlás reprezentál, a KBVR I. ütemről (Hídvégi-tó) a II. ütemre (Fenéki-tó) történő átlépéskor magasnak számít. Az eredményekből látszik, hogy a vízáramlás irányában kezdetben a toxicitás csökkenése figyelhető meg. Feltételezhető ugyanakkor, hogy a Kis-Balaton környezetéből beérkező kisvízfolyások (a sármelléki repülőtér irányából befolyó valamint a Hévíz-Páhoki-csatorna) jelentős többletterhelést jelentenek a tározótérre nézve. Mindennek ismeretében a jövőben szükséges lehet a Fenéki-tó területére beérkező vízfolyások vizsgálata, különös figyelmet fordítva az esetlegesen szennyezettebb területekről érkezőkre, tehát a potenciálisan nagyobb kockázatot jelentőkre.

A mért ökototoxicitás értékek önmagukban nem adnak információt a szennyezés mibenlétére. Eleve a mintavételi módszerekből, mintatartósításból, ill. a kivonatok elkészítéséből számos bizonytalansági tényező adódik (Burton, 1991). Így például az üledékminta oxidációja során megváltozhat a fémek felvehetősége. Becker and Ginn (1995) arról számolt be, hogy az üledékminták tárolásának ideje jelentős mértékben befolyásolta a Microtox tesztek eredményét.

A vizsgált terület környékén jelenleg, ill. régebben jellemző tájhasználat ismeretében felvetődik a repülőtérrel származó szénhidrogén-származékok, növényvédőszer használatból visszamaradt vegyületek (PAH, PCB, rézgalic Cu tartalma), ill. esetleg közlekedésből származó Pb előfordulása. Ezek közül a jövőben a repülőtér környezeti kockázatát tervezzük vizsgálni, immár analitikai mérések elvégzésével. Overton et al. (1997) eredményei szerint a teszt megfelelően detektálja a szénhidrogén-szennyezést.

Az antropogén eredetű szennyező komponensek mellett felvetődhet a természetes eredetű toxicitás problémája, az ún. mátrix-hatás is (pl. Jacobs et al., 1992). Mindezek figyelembevételével az ökototoxicitás méréseket előszűrésre, az ún. „forró pontok”, azaz a környezetegészségre magas kockázatot jelentő területek azonosítására lehet ajánlani (pl. Bennett and Cubbage, 1992).

IRODALOM

- Becker, D.S. and Ginn, T.C. (1995). Effects of storage time on toxicity of sediments from Puget Sound, Washington. In: Environmental Toxicology and Chemistry. 14. 5. 829-835.
- Bendefy, L. (1968). A Balaton vízszintjének változásai a neolitikumtól napjainkig. In: Hidrológiai Közöny, 6. 257-263.
- Bennett, J., Cubbage, J. (1992). Review and Evaluation of Microtox® Test for Freshwater Sediments. Environmental Assessment Program Report, 92-e04
- Burton, G.A. Jr. (1991). Assessing freshwater sediment toxicity. In: Environ. Toxicol.Chem. 10. 1585-1627.
- Burton, G.A. Jr., Baudo, R., Beltrami, M. and Rowland, C. (2001). Assessing sediment contamination using six toxicity assays. J. Limnol. 60(2): 263-267
- Doherty, F.G. (2001). A Review Of The Microtox® Toxicity Test System for Assessing the Toxicity of Sediments and Soils. In: Water Quality Research Journal of Canada, 36. 3. 475-518.
- Fekete, E., Szabó, S.A., Tóth, Á. (1991). A vízzennyezés ökológiája, Pro Natura Kiadó : Budapest
- Jacobs, M.W., Delfino, J.J., and Bitton, G. (1992). The Toxicity of Sulphur to Microtox® from Acetonitrile Extracts of Contaminated Sediments. In: Environ. Toxicol. Chem., 11: 1137-1143.
- Johnson, B.T., Long, E.R. (1998). Rapid toxicity assessment of sediments from estuarine ecosystems: A new tandem in vitro testing approach. In: Environ. Toxicol. Chem. 17. 6. 1099-1106.
- Magyarics, A., Pomogyi, P., Pék T. (1999). A Kis-Balaton Vedőrendszer kialakítása, működésének eredményei. Vízügyi Közlemények, LXXXI. Évfolyam, 4. füzet.
- Suter, G.W. (1996). Risk Characterization for Ecological Risk Assessment of Contaminated Sites. Prepared by the Environmental Restoration Risk Assessment Program, Lockheed Martin Energy Systems, Inc., Oak Ridge, Tennessee, ES/ER/TM-200

- Svenson, A., Edsholt, E., Rickling, M., Remberger, M., Rottorp, J. (1996). Sediment contaminants and Microtox toxicity tested in a direct contact exposure test. In: Environ. Toxicol. Water Qual. 11. 4. 293-300.
- Tátrai, I., Mátyás, K., Korponai, J., Paulovits, G., Pomogyi, P. (2000). The role of the Kis-Balaton Water Protection System in the control of water quality of Lake Balaton. In: Ecological Engineering 16. 73-78.
- Overton, E.B., Jr., C.B. Henry and I. Mendelssohn (1997). Application of Microtox assay to establish and evaluate the efficacy of *in situ* burning of oiled marshes. Louisiana Applied Oil Spill Research and Development Program, OSRADP Technical Report series 96-009.

Levelezési cím (*Corresponding author*):

Kováts Nóra

Veszprémi Egyetem, Környezetmérnöki és Kémiai Technológia Tanszék
8200 Veszprém, Egyetem u. 10.

University of Veszprém

Department of Environmental Engineering and Chemical Technology

H-8200 Veszprém, Egyetem Str. 10.

Tel.: 36-88-624-655, Fax: 36-88-624-533

e-mail: kovats@almos.vein.hu



Talajvíz ionkoncentrációjának vizsgálata (Közelítő módszerek összehasonlítása)

Gimesi L.

PTE TTK Informatika és Általános Technika Tanszék, 7624 Pécs, Ifjúság útja 6.

ÖSSZEFOGLALÁS

A pécsi Hőerőmű (PANNONPOWER Rt.) zagytározójának rekultivációja során felmerült az igény a talajvízben oldott ionok koncentrációjának vizsgálatára (monitorozására). Mivel felszín közeli vízről van szó, ezért az oldott anyagok hatással lehetnek a felszíni és felszín közeli élővilágra. A talajvíz kapcsolatban van más vizekkel is (folyók, tavak), ezért szükséges a víz mozgásának vizsgálata, amire az ionkoncentráció változásából következtethetünk. Szemléletesség érdekében a mérési eredményeket térben (grafikusan), háromdimenziós digitális terepmodell segítségével ábrázoltuk, amihez a lehető legtöbb mérési adat kell. Mivel a vízmintákat figyelőkutakból nyerjük, megoldhatatlan feladat, hogy tetszőleges számban és helyről gyűjtsünk információt. Ezért a terepmodell elkészítéséhez szükséges adatokat – az ismert mérések alapján – becsléssel (közelítő eljárással) határoztuk meg. A feladat elvégzéséhez a szakirodalmakban számos közelítő eljárás megtalálható, amelyek közül néhány módszer (krigelés, spline, inverz távolság és neurális hálózat) összehasonlítását is elvégeztük.

(Kulcsszavak: digitális terepmodell, neurális hálózat, krigelés, spline, inverz távolság)

ABSTRACT

Examination of ion concentration of ground water (Comparison of approximation methods)

L. Gimesi

University of Pécs, Faculty of Science, Department of Informatics and Technology, Pécs, H-7624 Ifjúság u. 6.

During the recultivation of the slurry reservoirs of Pannonpower Co. felt the need for examination (monitoring) of the soluted ion concentration in ground water. It had to take account of the fact that the water in question is close to the surface which could have effected to ecosystem on the surface. Ground water is strongly related to other sources of water like rivers, lakes, so the monitoring of the water flow had to be considered. Examination of the change in the ion concentration level of the water directly shows the water currents. The paper will display the data of examination with the help of a three dimensional digital graphical model which was chosen to cover the most of the data. Data collection is not an easy task especially if we aim to get multilevel informations, we were restricted to use mainly monitoring wells. Some of the data are based on estimation (approximation) procedures. To achieve the best possible approximation method we compared the efficiency of several methods, namely kriging, spline, inverse distance weighted, neural network.

(Keywords: digital surface model, neural network, kriging, spline, inverse distance weighting)

BEVEZETÉS

"A Dél-dunántúli régió környezetterhelésének csökkentésére irányuló komplex hulladékkezelési és rekultivációs technológia, valamint monitoring rendszer kifejlesztése és alkalmazása" című pályázat keretében sor került a pécsi Hőerőmű (PANNONPOWER Rt.) zagytározó rekultivációjának előkészítésére. E projecten belül felmerült az igény a talajvízben oldott ionok koncentrációjának vizsgálatára (monitorozására). Mivel felszín közeli vízről van szó, ezért az oldott anyagok hatással lehetnek a felszíni és felszín közeli élővilágra. (NKFP-3/050/2001 pályázat, 2003)

További feladat a talajvíz áramlási irányának meghatározása. Mivel a talajvíz kapcsolatban van más vizekkel (folyók, tavak), a szennyezőanyag eljuthat más területekre is. Erre az ionkoncentráció változásából próbáltunk meg következtetéseket levonni (Gimesi, 2003). (A feladat elvégzéséhez az adatokat a PANNONPOWER Rt. biztosította számunkra.)

Az emberi gondolkodáshoz közelebb áll a képi megjelenítés, mint a nagy, numerikus adatokat tartalmazó táblázatok. Ezek ugyan pontos információt adnak, de nehezen kezelhetők, és az összefüggések bemutatására nem alkalmasak. Az időbeni változások követéséhez, azok szemléltetéséhez a számsorok már teljesen alkalmatlanok, így monitoring rendszerek esetében kizárólag grafikát alkalmaznak.

Az adatok térbeli elhelyezkedéséről a 3 dimenziós digitális terepmodell (felületmodell) biztosítja a legszemléletesebb megjelenítést. Ekkor a mérési eredmények a környezettel együtt jelennek meg, így az összefüggések is jól értelmezhetők. Monitoring esetén a digitális terepmodell ábráit egymásra helyezve, esetleg animáció készítésével szemléletessé tehetjük a változásokat.

3D FELÜLETMODELLEK

A háromdimenziós felületmodell egy közönséges háromdimenziós geometriai tér (skalár- vagy vektormező), amelyről feltesszük, hogy a skalár és vektor helyfüggvények egyértékűek, folytonosak és a helykoordinátáinak ($r=xi+yj+zk$) alkalmasan differenciálható függvényei. (Korn, 1975) Továbbá feltételzzük, hogy a z koordináta csak x és y -től függ: $z=f(x,y)$.

A skalártér megjelenítéséhez (lerajzolásához) egy képzeletbeli hálót fektetünk a felületre, amely celláit homogénnek tekintjük. (A cella helye adja meg a rajz x és y koordinátáit, a cella értéke pedig a z koordinátát.) A rajz elkészítéséhez szükségünk van az összes cellaérték ismeretére.

Feladatunk talajvízben oldott anyagok (ionok) koncentrációjának vizsgálata. A mintákat megfigyelő kutakból nyerjük, amely erősen behatárolja a mintavételi helyek számát. (Tehát nem rendelkezünk tetszőlegesen nagy számú mintavételi hellyel.) Ebből következik, hogy nem áll rendelkezésünkre minden egyes cellához adat.

Az ismeretlen adatok meghatározásához közelítő eljárást (becslést) kell alkalmaznunk. A szakirodalmak (pl. Steiner, 1990; Iványi, 2004) több módszert is ajánlanak: statisztikai módszerek, spline, inverz távolság módszer, 3d evolúciós algoritmusok, neurális hálózatok, Fuzzy algoritmusok, fraktálok, stb.

Egy domborzati terepmodell elkészítésénél, a felmérés során láthatóak a felszín jellemzői (minimum, maximum helyek, a terepre jellemző, szemmel látható változások, stb.). Ezért az adatgyűjtés során ezekről a jól meghatározott helyekről vehetünk mintát. Egy korrekt felmérés alapján a valóságot jól közelítő domborzati modellt állíthatunk elő.

Feladatunk elvégzése során más a helyzet, ugyanis a kutakból vett mintákról nem tudjuk megállapítani, hogy azok mennyire jellemző értékek. Ezért is fontos, hogy a becslési módszert helyesen válasszuk ki.

A grafikus megjelenítéshez az ESRI Inc. által készített ArcMap – ArcView programot használtuk, amely a krigelés, az inverz távolság és az spline algoritmusokat használja. A felsorolt módszerek mellett, a már korábban publikált (Gimesi, 2004b) neurális hálózat segítségével is elkészítettük a felületmodellt.

Krigelés (Kriging)

A módszert Krige professzor dolgozta ki hagyományos statisztika alkalmazásával. Lényege: a keresett értéket az ismert adatok súlyozott átlagából számítjuk úgy, hogy az eredmények szórása minimális legyen (Steiner, 1990).

Az eljárás azt vizsgálja, hogy a térbeli pontok szórása milyen gyorsan változik. Ez a variogram (szemivariogram) nevű függvény segítségével határozható meg, amely a pontok közötti távolság függvényében adja meg az értékkülönbségek négyzetösszegének a felét:

$$\gamma(h) = \frac{1}{2n(h)} \sum_{i=1}^{n(h)} (Z_{P_i} - Z_{P_{i+h}})^2$$

ahol P_i az összes olyan mérési pont, amelytől h távolságra még található ismert ($Z_{P_{i+h}}$) érték.

Gyakorlatban a variogram meghatározásához közelítő modellt használnak. Az ArcView programnál használt modellek:

szférikus modell (*spherical model*):

$$\gamma(h) = C \left[1,5 \frac{h}{H} - 0,5 \left(\frac{h}{H} \right)^3 \right], \quad \text{ha } 0 \leq h \leq H,$$

$$\gamma(h) = C, \quad \text{ha } h > H;$$

kör modell (*circular model*):

$$\gamma(h) = C \frac{2}{\pi} \left[\frac{h}{H} \sqrt{1 - \left(\frac{h}{H} \right)^2} + \arcsin \left(\frac{h}{H} \right) \right], \quad \text{ha } 0 \leq h \leq H,$$

$$\gamma(h) = C, \quad \text{ha } h > H;$$

exponenciális modell (*exponential model*):

$$\gamma(h) = C \left[1 - \exp \left(- \left(\frac{h}{A} \right) \right) \right], \quad A \cong \frac{H}{3};$$

Gauss-modell (*Gaussian model*):

$$\gamma(h) = C \{ 1 - \exp [- (h / A)^2] \}, \quad A \cong \frac{H}{\sqrt{3}}.$$

A fenti modellek grafikus összehasonlítását az 1. ábra mutatja, ahol megfigyelhetők az eltérések a különböző eljárások között. Látható, hogy a modell megválasztása hatással van az eredmények jóségára.

Az ismert adatok hatása a vizsgált pontra a köztük lévő távolsággal csökken, ezért a számolás során a H hatótávolságon túli pontokat már nem vesszük figyelembe. E távolságban lesz a variogram maximális (telítési) értéke C .

1. ábra

Variogram modellek

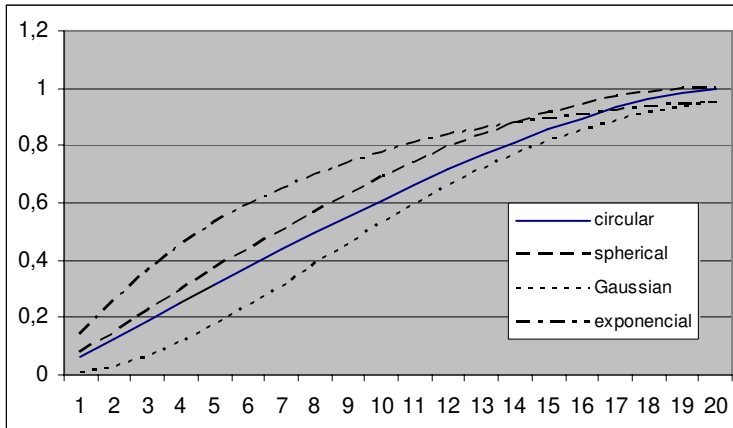


Figure 1: Models of variogram

A számítandó P_0 pontbeli $Z(P_0)$ értéket n darab közeli P_i pont $Z(P_i)$ értékének súlyozott átlagaként becsüljük meg. (Steiner, 1990)

$$Z(P_0) \approx \sum_{i=1}^n s_i Z(P_i),$$

ahol: $\sum_{i=1}^n s_i = 1$, ugyanis ekkor lesz a becslés szórása minimális.

S_i -ket az alábbi mátrixegyenletből határozhatjuk meg:

$$S_0 = K^{-1} C_0$$

ahol: K (Krige-mátrix) és a C_0 elemei a variogramból számíthatók.

A módszerrel készült koncentráció eloszlást az 2. ábra szemlélteti.

Inverz távolság (Inverse Distance Weighting)

A módszer lényege, hogy az ismeretlen adatokat a szomszédos (ismert) adatok alapján úgy határozza meg, hogy minél távolabb van egy pont (P_i), annál kisebb a hatása az eredményre. Ezt egy súlytényezővel vesszük figyelembe:

$$Z_e = \frac{\sum_{i=1}^n \frac{1}{d_i^w} Z_i}{\sum_{i=1}^n \frac{1}{d_i^w}}$$

ahol:

- Z_e a keresett érték,
- d_i távolság az ismert i . ponttól,
- Z_i az ismert i . pont érték,
- w súlytényező.

Kétdimenziós esetben:

$$Y_e = \frac{\sum_{i=1}^n \frac{1}{|X_e - X_i|^w} Y_i}{\sum_{i=1}^n \frac{1}{|X_e - X_i|^w}}$$

A súlyozás következtében a távoli pontok hatása elhanyagolható, ezért a gyakorlatban egy bizonyos távolságon (hatótávolságon) túl már nem vesszük figyelembe az ismert adatok hatását.

E módszerrel kapott 2D-s eredményt a 3. ábra szemléltet, 3D-s alkalmazásra példa a 4. ábra.

2. ábra

Mn-ion koncentráció Krigeléssel



Figure 2: Concentration of manganese with Kriging

3. ábra

Interpolációs módszerek összehasonlítása

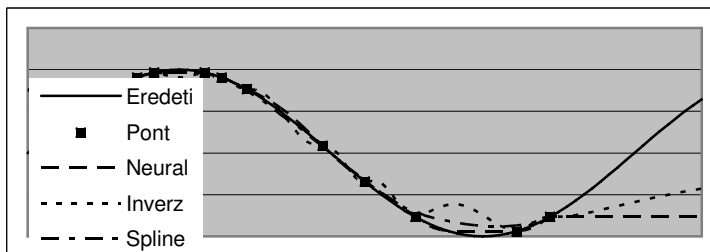


Figure 3: Compare different interpolation methods

4. ábra

Mn-ion koncentráció inverz távolság módszerrel

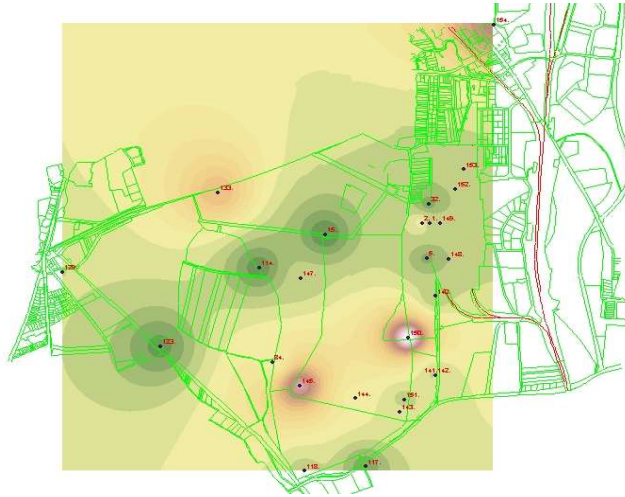


Figure 4: Concentration of manganese with Inverse Distance Weighting

Spline

Amennyiben feltételezzük, hogy a skalármező megfelelően sima, akkor az interpoláláshoz jól használható az úgynevezett „rugalmas vékonylemez” modell, ahol a felület harmadrendű polinomokkal közelíthető (Stern, 1990).

Egyváltozós esetre vizsgálva határozzuk meg n darab ismert pont (P_i) közelítő görbéjét. Legyen P_i harmadfokú polinom az $[X_i, X_{i+1}]$ intervallumon ($i=1 \dots n$).

Biztosítanunk kell a polinomok folytonosságát:

$$P_i(X_{i+1}) = P_{i+1}(X_{i+1}) \quad i=2 \dots (n-1),$$

valamint azt, hogy a polinomok végpontjának érintője megegyezzen a következő polinom kezdőpontjának érintőjével, így biztosítható a sima átmenet:

$$P_i'(X_{i+1}) = P_{i+1}'(X_{i+1}) \text{ és}$$

$$P_i''(X_{i+1}) = P_{i+1}''(X_{i+1}) \quad i=2 \dots (n-1),$$

továbbá:

$$P_1''(X_1) = 0 \text{ és } P_{n-1}''(X_n) = 0$$

(A görbe az X_0 illetve X_n pontban érinti az első és az utolsó két pontot összekötő szakaszt.)

Az ismertetett eljárás számítógépes megoldására a Cox-deBoor algoritmust használják, miszerint a görbét lineáris számítások sorozatával konstruálják meg (Iványi, 2004).

A $P(u)$ görbét az ismert $n+1$ pont paramétereivel (P_i) határozzuk meg (Newman, 1985).

$$P(u) = \sum_{i=0}^n P_i N_{i,k}(u)$$

A k paraméter a foksám, amely megadja a görbe simaságát. Az 2. ábrán lévő görbe $k=3$ értékkel készült.

A súlyfüggvényt rekurzív módon definiáljuk:

$$N_{i,k}(u) = \frac{(u - t_i)N_{i,k-1}(u)}{t_{i+k-1} - t_i} + \frac{(t_{i+k} - u)N_{i+1,k-1}(u)}{t_{i+k} - t_{i+1}},$$

Ha $t_i \leq u < t_{i+1}$, akkor $N_{i,1}(u)=1$,

egyébként $N_{i,1}(u)=0$.

$t_0 \dots t_{i+k}$ csomóértékek, amelyeket a következő szabály szerint számítjuk:

ha $i < k$, akkor $t_i=0$,

ha $k \leq i \leq n$, akkor $t_i=i-k+1$,

ha $i > k$, akkor $t_i=n-k+2$.

Felületmodellezésnél (5. ábra) a felületet két görbe Descartes-szorzataként állítjuk elő:

$$P(u, v) = \sum_{i=0}^n \sum_{j=0}^m P_{i,j} N_{i,k}(u) N_{j,l}(v).$$

5. ábra

Mn-ion koncentráció spline módszerrel

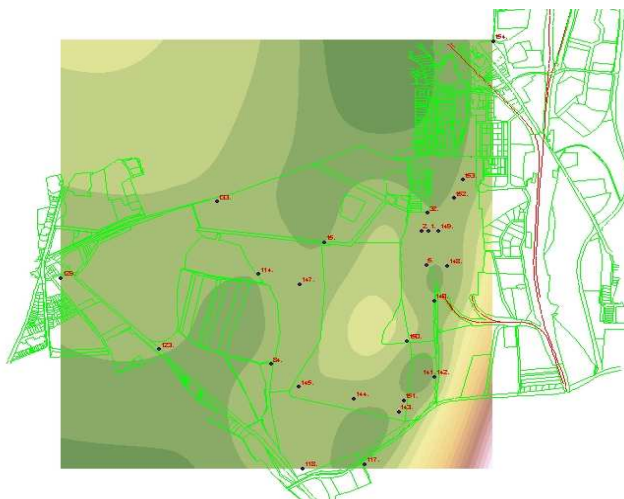


Figure 5: Concentration of manganese with Spline

Neurális háló (NN)

Az idegsejt modellezésével az irodalom (Gimesi, 2004a) részletesen foglalkozik, itt csak a lényegét emeljük ki.

Az idegsejt modelljét a 6. ábra szemlélteti. Az ingerület (i) szinapszisok segítségével jut el az idegsejtre. A szinapszisokban az ingerület erősödhet (ingerlés) vagy gyengülhet (gátlás), ennek mértékét egy súlyszámmal w -vel jelöljük. Az idegsejtre

– a szinapszison keresztül – érkező ingerület nagyságát az $i \cdot w$ szorzat adja. Az idegsejt felületén a beérkező ingerületek összegződnek, a kialakult eredő:

$$x = \sum_{i=1}^n i_i w_i$$

6. ábra

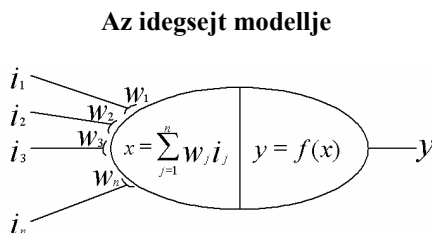


Figure 6: Model of the nerve-cell

Ha az eredő inger (x) eléri a küszöbszintet, akkor kialakul az ingerület (y), amely átadódik a következő idegsejtnek. Az idegsejt „átviteli függvénye” legyen $f(x)$, így a kialakult ingerület:

$$y = f(x)$$

A legegyszerűbb esetben $f(x)$ egy egységugrás:

ha $x \geq \text{küszöb}$, akkor $y = 1$,

ha $x < \text{küszöb}$, akkor $y = 0$.

Egy egyszerű idegrendszermodellt (neurális hálót) a 7. ábra mutat be.

7. ábra

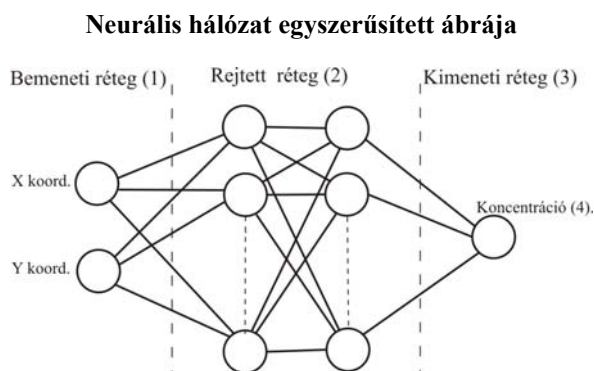


Figure 7: Structure of the simple neural network

Input layer(1), Hidden layer(2), Output layer(3), Concentration(4)

Az ingerfelvevő (érzék-) sejtek alkotják a bemeneti (input) réteget. Itt annyi neuron található, ahány bemenő adat (változó) tartozik egy feladathoz. (Esetünkben ez az x és y (EOV) koordináta pár.)

A következő (rejtett) réteg reprezentálja az idegrendszert, ahol a neuronok a legváltozatosabb módon kapcsolódhatnak össze. E rétegben több alréteg is definiálható. A szükséges neuronok számának meghatározására több elmélet is született. A tapasztalatok azt mutatják, hogy egzakt módon ez nem határozható meg, ezért a korszerű szimulációs (modellező) programok az alrétegek és a neuronok számát is képesek automatikusan változtatni.

Az ingerekre adott válasz – ahogy az idegrendszerben is – a kimeneten jelenik meg. A kimeneti rétegben (output) annyi neuron van, ahány kimeneti (eredmény) érték. (Jelen feladatban ez a talajvízben oldott anyag koncentrációja.)

Az input neuronról érkező jel a következő szint mindegyik neuronjára rákerül w_i -vel (súlyszámmal) való szorzás után. A neuronra érkező jelek összegződnek, majd az átviteli függvénynek megfelelően megjelennek a neuron kimenetén. Innen továbbjutnak a következő réteg (alréteg) neuronjaira megszorozva az összeköttetésre jellemző súlyszámmal. Ez addig folytatódik, amíg a kimeneti réteget el nem érjük. Az output neuronokban csak összegzés történik.

A bonyolultabb modelleknél (ahogy a valóságban is) a kapcsolatok nemcsak a következő réteg neuronjaival alakulhatnak ki, hanem bármelyik rétegben lévővel, sőt visszacsatolás is lehetséges, vagyis a kimeneten megjelenő jel visszajuthat egy előző rétegbe.

A neurális háló használatához meg kell terveznünk a hálózatot. A bemeneti és kimeneti réteg neuronjainak száma a rendelkezésünkre álló és a számítandó adatoktól függ. A rejtett rétegben alkalmazott neuronok mennyisége – különböző elméleti megfontolások alapján – próbálgatással vagy automatikusan határozható meg. Továbbá szükség van a neuronok átviteli függvényének és a súlyszámok megadására is.

Súlyszámokat a szimulációs program határozza meg ismert adatok (mérési eredmények) alapján. Ez a tanulási folyamat (Gimesi, 2004a).

A neurális hálóval készült két- és háromdimenziós eredményeket a 3. ábra, illetve a 8. ábra mutatja be.

8. ábra

Mn-ion koncentráció neurális hálózattal

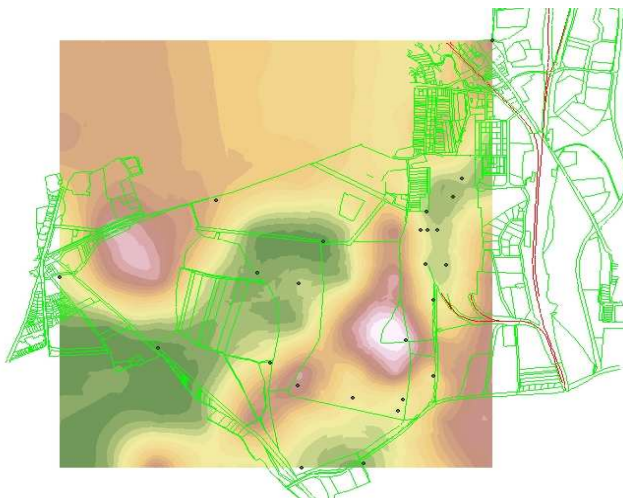


Figure 8: Concentration of manganese with NN

EREDMÉNY ÉS ÉRTÉKELÉS

A 3. ábrán egy kétdimenziós függvény közelítését mutatjuk be az ismertetett eljárások segítségével, néhány megadott pont alapján. Az ábrából látható, hogy a *spline* használatakor – a súlyszámtól függően – az eredetihez képest egy simább függvényt kapunk. E megállapításunkat a 5. ábra is alátámasztja.

Az *inverz távolság* modell viszont kiemeli a szélsőértékeket. Esetünkben ez nem szerencsés, mivel nem tudjuk, hogy a mért értékek maximum (minimum) helyei a valóságban is a mérési pontokban találhatók-e. A 3D-s modell a 4. ábrán látható.

Krigelés esetén (2. ábra) kevésbé hangsúlyosak a szélsőértékek, így egy simább felületet kapunk.

A kétdimenziós modellt a *neurális hálózattal* sikerült a legjobban megközelíteni.

A háromdimenziós modellek összehasonlításakor látható, hogy ahol a mintavételezési helyek közel vannak egymáshoz, az ábrák (*spline* kivételével) hasonló jellegű mutatnak.

Érdemes összehasonlítani azokat a helyeket is, ahol a közelben nincs mért adatunk. Az *inverz távolság* modell és a *krigelés* ezeken a helyeken – elsősorban az ábra széleinél – egy átlagos értékhez közelít. A *neurális háló* ezzel szemben – a mért értékek alapján ezekre a pontokra is – az átlagtól eltérő becslést ad.

A *neurális hálózat* modell hátránya, hogy a hálózat kialakítása és a tanulási folyamat hosszú időt, jelentős erőforrást és nagy figyelmet igényel.

A feladat elvégzése során több esetben is tapasztaltuk, hogy *krigeléssel* nem lehetett elvégezni az interpolációt. Erre a problémára utal Steiner is (Steiner, 1990), hogy a *krigelés* akkor használható, ha az ismert adatok nem függetlenek egymástól, valamint a meghatározandó értékek is korrelálnak az ismert értékekkel. (Közelebbiekkal jobban, a távolabbiakkal kevésbé.) Azaz a *krigelés* viszonylag sima felület esetén alkalmazható.

A bevezetőben említett project keretében megkíséreltük megállapítani a talajvíz áramlásának irányát, a koncentráció változásának segítségével. Az ábrák kiértékelése során arra a következtetésre jutottunk, hogy a rendelkezésünkre álló adatok alapján ez egyértelműen nem határozható meg, mivel a mérési eredményeket nagymértékben befolyásolja a csapadék mennyisége is. A talajban jelentős mértékben találhatók olyan sók, amelyeket a csapadék kiold, és ezzel folyamatosan változtatja a talajvíz összetételét.

KÖVETKEZTETÉSEK

A Pannon Hőerőmű Rt. zagytározójának területén korábbi (2000-2002) mérésekből származó talajvíz adatokat dolgoztunk fel az ArcView program illetve *neurális hálózat* segítségével.

Az adatok kiértékeléséhez segítséget nyújtott, hogy az eredmények egy térképre készültek, amely tartalmazza a felszíni azonosítási pontokat. Ezzel lehetőség van például a terepen végzett mérések és a talajvíz-adatok összehasonlítására, azok közötti összefüggések feltárására.

A feladat elvégzése során összehasonlítottunk néhány, a térinformatikában használt interpolációs eljárással, valamint *neurális hálózattal* készített terepmodellt. Az a tény, hogy modellezéshez szükséges adatokat kutakból nyertük, behatárolja a mintavételi pontok helyét és számát. Így nem állt módunkban annak ellenőrzése, hogy a különböző interpolációs eljárások mennyire közelítik meg a valóságot. Ehhez az elemzéshez olyan

adatokra lenne szükség, ahol a mintavételi helyek kötetlenebbül megválaszthatók. Ilyen például a felszíni vizek vagy a talaj összetételének vizsgálata.

További vizsgálatokat kívánunk végezni a talajvíz mozgására, összefüggéseket keresünk a csapadék és a koncentrációváltozás között. Elemzésre vár az extra magas, illetve alacsony koncentrációk eredete, a koncentráció eloszlás, és annak kapcsolata a talajösszetétellel és a felszínnel.

A későbbiekben vizsgálni kívánjuk a különböző interpolációs módszerek stabilitását. Ennek kapcsán arról szeretnénk információt kapni, hogyha kiveszünk egy-egy mérési helyet a rendszerből, akkor milyen mértékben változnak meg az interpolált adatok.

IRODALOM

- Gimesi L. (2003). Az NKFP-3/050/2001 pályázat térinformatikai feldolgozó rendszerének továbbfejlesztése a Pannon Hőerőmű Rt. korábbi abiotikus adatainak a térinformatikai rendszerbe vitele és feldolgozása érdekében. Szakvélemény. Pécsi Tudományegyetem, Pécs.
- Gimesi L., Béres Cs. Z., Bérczi Sz., Hegyi S., Cech V. (2004). Teaching planetary GIS by constructing its model for the test terrain of the Hunveyor and Hussar. 35th Lunar and Planetary Science Conference, Houston. 1140.
- Gimesi L. (2004a). Neurális hálózatok és térinformatikai alkalmazásuk. Térinformatika No 7, 13-15.
- Gimesi L. (2004b). Mesterséges intelligencia alkalmazása a rekultivációban. Acta Agraria Kaposváriensis, Kaposvári Egyetem, Vol 8 No 3, 1-9.
- Iványi A. (2004). Informatikai Algoritmusok I. ELTE Eötvös Kiadó, Budapest.
- Korn, G.A. – Korn T. M. (1975). Matematikai kézikönyv műszakiaknak. Műszaki könyvkiadó, Budapest.
- Newman, W.M. – Sproull, R.F. (1985) Interaktív számítógépes grafika. Műszaki könyvkiadó, Budapest.
- NKFP-3/050/2001 pályázat. (2003). A Dél-dunántúli régió környezetterhelésének csökkentésére irányuló komplex hulladékkezelési és rekultivációs technológia, valamint monitoring rendszer kifejlesztése és alkalmazása.
- Steiner F. (1990). A geostatistika alapjai. Tankönyvkiadó, Budapest.

Levelezési cím (*Corresponding Author*):

Gimesi László

PTE, TTK, Informatika és Általános Technika Tanszék

7624 Pécs, Ifjúság útja 6.

University of Pécs, Faculty of Science

Department of Informatics and Technology

H-7624 Pécs, Ifjúság u. 6.

Tel.: 36-72-503-697, Fax: 36-72-503-697

e-mail: gimesi@ttk.pte.hu



Szakértő szoftverágens a diszlexia lehetőségének megállapítására

Starkné W. Á., Fogarassyné V. Á., Csoma Á.

Pannon Egyetem, Matematikai és Számítástechnikai Tanszék, 8200 Veszprém Egyetem u. 10.

ÖSSZEFOGLALÁS

A történelem kezdetétől fogva elbűvöli az embereket a nem humán eredetű ágensok ötlete. Maga a fogalom, mint android, humanoid, robot stb. áthatják kultúránkat, formálják tudatalattinkat, attól függően, hogy épp melyik szoftverágenst észleljük. Mindennapi értelemben egy szoftverágenstől azt várjuk, hogy valaki helyett végrehajtsa egy bizonyos feladatot, amelyre felhatalmazták. Az ágenstől elvárjuk, hogy képes legyen egy átfogó szaktudás kifejtésére úgy, hogy számításba veszi a felhasználó és a szituáció sajátosságait is. A cikk alapvető célja, hogy bemutassa, hogyan lehet korszerű eszköz segítségével olyan szakértő szoftverágenst építeni, amely egy olyan diszfunkció korai felderítésében segíthet, amely igen sok kisgyermek esetében tapasztalható ma Magyarországon, ez a betegség a diszlexia. A szakértő ágens kialakítása során nagy gondot kellett fordítani arra, hogy a rendszer jövőbeni feltételezett használói nem feltétlenül számítástechnikai szakterületen dolgoznak, hanem a vizsgálat szempontjából fontos korosztállyal, az elsősztályosokkal foglalkoznak. Továbbá kihívást jelentett az is, hogy olyan diszlexia-veszélyeztetettségi területeket találjunk, amelyek vizsgálatokkor nem szükséges a gyermek közvetlen, aktív részvétele, hanem a gyermek mindennapi szokásait figyelve válaszolhatóak meg a kialakított témacsoportokban feltett kérdések. Tizenegy megfigyelési terület került definiálásra, amelyekhez különböző kérdéscsoportokat rendeltünk. A területeket többszörös normalizálással tettünk összehasonlíthatóvá, így az értékelés utáni következtetés konzekvenssé vált. A rendszer meghatározza azon területeket, amelyeknél a vizsgált gyermek készségeinek kialakultsága nem megfelelő. Ezen területi készségek meglétéből vagy nem létéből lehet meghatározni, hogy a gyermek diszlexia-veszélyeztetett lehet-e avagy sem. A teszt az értékelés során felhívhatja a figyelmet a gyermek esetleges téri tájékozódási zavarára, vizuál-motoros-koordinációs zavarára, motoros-koordinációs zavarára, figyelem-, emlékezet zavarára, általános képességeinek elmaradására, beszédfejlődésének megkésésére, rossz olvasási technikájára, illetve megadhatja még azon területeket, ahol a jövőben jobban oda kell figyelni a gyermek készségeinek fejlődése folyamán.

(Kulcsszavak: szoftverágens, mesterséges intelligencia, diszlexia)

ABSTRACT

Expert software agent for the diagnosis of possible dyslexia

Á.W. Starkné, Á.V. Fogarassyné, Á. Csoma

Pannon University, H-8200 Veszprém Egyetem u. 10.

Since the beginning of recorded history, people have been fascinated with the idea of non-human agencies. Popular notions about androids, humanoids, robots and cyborgs creatures permeate our culture, forming the unconscious backdrop against which software agents are

perceived. As in the everyday sense, we expect a software agent to act on behalf of someone to carry out a particular task which has been delegated to it. The best agents would not only need to exercise a particular form of expertise, but also take into account the peculiarities of the user and situation. Our fundamental aim in writing this article is to provide a general and clear definition of dyslexia as a concept and as a group of symptoms. In the article, we also aim at highlighting the major areas the examination of which could serve as a basis of designing an expert software agent that analyses possibility to dyslexia. During the development of such an expert agent, we had to pay special attention to the fact that the probable future users of the system would not necessarily be information technology experts, but instructors and others dealing with the targeted age group, i.e. first grade students in elementary school. Another challenge during system development was to find different areas where susceptibility to dyslexia could be examined, areas for the examination of which the active participation of children was not necessary. The test under different headings can be answered simply by monitoring the everyday behaviour and habits of the child. During expert software agent evaluation, attention might be called to any possible orientation, visual-motor-coordination, motor-coordination, concentration or memory disorders, inadequately developed general abilities, delayed development of speech or a bad reading technique.

(Keywords: software agent, artificial intelligence, dyslexia)

BEVEZETÉS

Napjaink egyik gyakran használt kifejezése: diszlexiás a gyermek. Tapasztalataink alapján, akik ezzel a szókapcsolattal élnek, nem látják át a diszlexia komplexitását. Természetesen ilyenkor nem a szakemberekről beszélünk, hanem olyan szülőről, aki nagy figyelmet fordít a gyermekek fejlődésére. A tanulási zavarokkal, mint például a diszgráfia, diszlexia, diszkalkúlia nagyon sok szakirodalomban foglalkoznak (lásd például Boldureva, 1997; Ligeti, 1967; Lohmann, 1996; Selikowitz, 1997; Torda, 1999 stb.), hiszen nagyon sok gyermeket érintenek ezek a problémák. Sokáig kutattunk olyan alkalmazás után, amely a diszlexia-veszélyeztetettség felmérésére alkalmas. Valószínű a terület komplexitása miatt eddig nem készült ilyen jellegű alkalmazás.

A diszlexia okait kutatva az egyes részterületek (agy dominancia, vizuális észlelés...) fejlődésével ismerkedhetünk meg ezen írás második fejezetében. Mindezek után az olvasással foglalkozunk figyelembe véve a normális, megfelelő olvasást, s amelyhez viszonyítjuk az olvasási nehézséget. Ahhoz, hogy meg lehessen érteni egy-egy diszlexiás gyermek állapotát, aki átlagos intelligenciával rendelkezik, tudnunk kell többek között, hogy milyen módszerrel is tanult meg olvasni, s ez az ő esetében milyen hátrányokkal, előnyökkel járhatott. Úgy éreztük, foglalkoznunk kell a diszlexiás vagy diszlexia-veszélyeztetett gyermek természetével, különös tekintettel a koncentrálásra. A harmadik fejezetben az alkalmazás informatikai hátterének néhány fogalmi ismertetésén keresztül a felhasznált technikát vázoljuk fel. A negyedik fejezetben tizenegy megfigyelési területet definiáltuk, amelyekhez különböző kérdéscsoportokat rendeltük. A területeket többszörös normalizálással tettük összehasonlíthatóvá, így az értékelés utáni következtetés konzekvenssé vált. A kialakítandó alkalmazás olyan fejlesztői környezet megválasztását tette szükségessé, amely integrálja a kliens-szerver programozást, webes alkalmazást és egy honlap mesterséges intelligenciával, tudással való ellátását.

A DISZLEXIA, MINT A TANULÁSI ZAVAROK EGYIKE

Kirk és Bateman (1962) meghatározásában a tanulási zavar „olyan elmaradás, rendellenesség vagy megkésett fejlődés a beszéd, olvasási, írási, számolási folyamatokban

vagy más iskolai tantárgyakban, amelyet lehetséges agyi diszfunkció és/vagy emocionális vagy viselkedési zavar által okozott pszichológiai hátrány eredményez. Nem értelmi fogyatékos, érzékszervi hiányosság vagy kulturális vagy oktatási tényezők okozzák.”

A meghatározás hangsúlyozza a tanulási zavarnak az iskolával kapcsolatos viselkedésben való megjelenését és sérült pszichológiai folyamatokkal való kapcsolatát. A tanulási zavarok kifejezés átfogó meghatározása azon rendellenességek heterogén csoportjának, amelyek az intelligencia szinttől függetlenül jelentős nehézséget okoznak az írás olvasás és/vagy számolás elsajátításában, és a zavarok visszavezethetők a központi idegrendszernek a szokásostól eltérő fejlődésére, működésére. A szindróma kialakulásának hátterében a szenzomotoros képességek, illetve sajátos információfeldolgozásbeli eltérések azonosíthatók. A zavarok megjelenhetnek testmozgászavar, gyenge egyensúlyérzék és gyenge téri orientáció terén, valamint a vizuális, auditív és taktilis észlelésbeli és szekvenciális információfeldolgozásbeli deficitekben.

A legtöbb esetben a tanulási zavarok az információ feldolgozásának egy speciális módjának tekinthetők. Megfelelő ingerkörnyezet és oktatás esetén a zavarok kialakulása megelőzhető/enyhíthető.

Bár a tanulási zavarok megjelenhetnek más hátrányt okozó rendellenességekkel (érezékszervi, mozgásszervi rendellenességek, értelmi fogyatékos, társas-érzelmi zavarok) vagy környezeti hatásokkal (szocio-kulturális hátrány, nem megfelelő oktatás), ezeknek nem egyenes következménye a tanulási zavarok kialakulása (*Gyarmathy, 1996*). A diszlexia különböző időpontokban, rendszerint az iskolában fellépő részleges teljesítményzavar, amely általában a gyermekkel szemben támasztott túlzott követelmények eredményeképpen válik láthatóvá. A teljesítményzavar az agy egy részének késedelmes érése miatt jön létre, amely visszavezethető domináns örökletességre, laterális (oldaliság) problémára, észlelési nehézségekre, anatómiai károsodásra.

Az éres során tulajdonképpen az integráció fejlődéséről beszélünk, hiszen ekkor az észlelési folyamatban résztvevő különböző ingerek összefüggésük szerinti „egésszé” rendeződése zajlik le. Az idegrendszer teljes érettsége csak a 10-12. életévben fejeződik be. Minél egyértelműbben integrálódnak az érzékek az idegrendszerben és minél jobban összevethetők a korábban gyűjtött tapasztalatokkal, annál zavartalanabb a fejlődés, az éres folyamata. Csak az egyes érzékelési modalitások (azonos érzékeléstípusok), vagy a különböző érzékelési modalitások integrációjának belső, vagy külső zavara lassíthatja le, vagy szakíthatja meg az éresi folyamatot. Ez pedig különböző típusú tanulási zavarokhoz vezethet.

Az észlelések az éresi folyamattal párhuzamosan mindig differenciáltabban működnek és egyre komplikáltabb kapcsolatok épülnek fel. Egyre nagyobb az állandó információcsere lehetősége a két agyfélteke között.

A laterális fejlődése az észlelés és a kompenzálás az éres során funkcionális agyspecializálódáshoz vezet, amely az egyes emberekben jellegzetes. Az emberek különbözőségét az agyfunkcióikban lévő változatos eltérések és a kompenzálás különböző kombinációi adják (*Lohmann, 1996*).

A szülők a sajátos biológiai adottságok ismeretének hiányában mindig azon fáradoznak, hogy gyermeküket a normális kép alapján alakítsák, attól félve, hogy gyermeküknek különben nem lesz jövője. Minél biztosabbnak és elismertebbnek érzi magát a gyermek, annál kevésbé mutat regressziót (korábbi gyermeki magatartásformába való visszaesést), annál inkább rendelkezik erővel ahhoz, hogy új utakat keressen. A diszlexiás gyermekek gyakran másságuk miatt érdekesek, kezdeményezők, kreatívak, eredetiek. De az ilyen gyermek a környezet hatására kudarcossá vagy különőssé is válhat. Az iskola és sok szülő számára ezek a gyermekek kényelmetlenek, kifejezetten nehezen kezelhetők, és alig van esélyük az egyoldalú tanulási módszerek mellett.

Számos kutató foglalkozott a gyermekek tanulási nehézségeivel az agy információfeldolgozási munkáját figyelve elsődlegesen, s ők alakítottak ki bizonyos elméleteket az információfeldolgozási képességek hiányosságairól (*Gyarmathy és Smythe, 2000; Kirk és Bateman, 1962; Boldureva és Isekova, 1997; Sarkady és Zsoldos, 1992* stb.).

Ahhoz, hogy a gyermekek 8-10 éves korukban elérjenek egy olvasási szintet, két szakaszon kell túljutniuk, mégpedig a vizuális memória és a fonológiai szakaszon.

Az első részhez nem tartozik hozzá a lexikális rendszer (a lexikon üres). Így a gyermekek a szavakat úgy ismerik fel, mint a tárgyakat, azaz a kinézetüket, rajzukat jegyzi meg. Éppen ezért a „fáj” és a „jaj” szavak könnyen összetéveszthetőek számukra. Ebből kifolyólag ez a módszer olvasásnál hatástalan, hiszen számos olyan szavunk van, amelyek hossza és alakja nagyjából azonos, így a vizuális rendszer gyorsan nem tud különbséget tenni közöttük.

A második szakasz a fonológiai stádium, amelyet normális gyermekek 6-7 éves korukban érnek el. Ekkor kezdődik az alaplexikon feltöltése a későbbi automatikus stádium biztosításához. A használt rendszer egy lehetséges út a lexikális rendszerhez. Ezt fonológiai rendszernek nevezzük, mert a szavakat feldarabolja (szótagolja) az azt összetevő hangokra.

A hang legkisebb egysége a fonéma. Az ábécé kevesebb betűből áll, mint ahány fonéma létezik, ezért sok fonéma jelölésére betűkombinációkat használunk. Grafémának nevezzük minden olyan betűt vagy betűcsoportot, amelynek egy fonéma felel meg.

Ebből adódik, hogy az olvasástanulás során a pontos fonéma ismeretéhez az agynak meg kell tanulnia a nyelv graféma-fonéma kapcsolatát. A szó egységnek tekintendő, és a betűket a fonémák során megtanult szabályok fényében kell értékelnünk. Ahogy a gyermekek egyre jobban elsajátítják a grafémák lefordítását, a lapon a korrekt fonémát látják, amely természetesen függ a szövegkörnyezettől is, és elkezdik az agyban lévő lexikont szavakkal feltölteni. Ha ez megtörtént, akkor kezdik a fonológiai rendszert megkerülve az ismerős szavakat a lexikonból kikeresni. Így a relatíve lassú fonológiai rendszer használata helyett előtérbe kerül az automatikus olvasás, amihez a lexikont használják.

Számos teszt áll rendelkezésre, hogy egy adott gyermek olvasási képességeit felmérjük, és eredménye összehasonlíthatóvá váljon kortársaiéval (ilyen módszereket említ például *Smythe és Everatt, 2000; Gyarmathy és Smythe, 2000*).

Az olvasás során fel kell jegyezni az elkövetett hibák számát az olvasási pontosság megállapítása érdekében, amit szintén az életkornak megfelelő állandó értékekkel hasonlítunk össze. Természetesen ezek még mindig nem elegendő adatok, hiszen meg kell győződnünk megfelelő kérdések segítségével arról, hogy a gyermek valóban értette-e a szöveget. A vizsgálat során a vizsgálatot vezető személynek arra is figyelnie kell, milyen típusú hibák a leggyakoribbak.

Dr. Ligeti Róbert az olvasásban előforduló tévesztések jellegét csoportosította (*Ligeti, 1967*). Az olvasni kezdő és a diszlexiás gyermekek hasonló hibákat követnek el. A gyermekek nagy többsége túljut, a diszlexiások megrekednek a hibás olvasás fokán. Ezen leggyakoribb hibák a következők:

- Szavak, szótagok vagy betűk felcserélése, betűk átrendezése a szavakon belül.
- Szótagok, hangok kihagyása.
- Hozzátoldások.
- Ismétlések.

További specifikus tesztre van szükség ahhoz, hogy a probléma pontos természetét fel lehessen tární. Például meg kell vizsgálni a vizuális percepciót: az agy képes-e megérteni azt, amit a szem lát? Természetesen az értelmes és értelmetlen szavakat össze kell

hasonlítani a fonológiai jártasság eldöntése érdekében. Mivel a nyelv és a tanulás szorosan összefügg egymással, érthető, hogy a nyelvi nehézségek gyakran együtt járnak a tanulási nehézségekkel is.

Gyakran lustasággal vádoljuk a diszlexiás gyermeket, mivel nem mutatnak érdeklődést és dekoncentráltak. Írásuk közel olvashatatlan. A gyermek fél, bizonytalan, s görccsökkel teli. A félelem akadályozza meg az emlékezést és a lényeges-lényegtelen közti különbség felfedését (emelkedett adrenalin elválasztás, amely teljes gátláshoz vezethet).

Az idődrukk is nagy szerepet játszik, hiszen a sorrendiség megállapításához, a differenciáláshoz a diszfunkció zavarban szenvedő gyermeknek nincs megfelelően elég ideje a normális olvasási tempóhoz viszonyítva, így egy normálisan olvasó gyermekhez képest sokkal nagyobb erőfeszítésre van szüksége olvasás közben. A diszlexiás gyermeknek szüksége van arra, hogy bármennyire is kicsi a felolvasni kívánt szó, azt hangosan kimondja és elemezze. Ha siet, akkor fél szavakat, fél mondatokat ír le. A koncentrációerő lassan elfogy, mindinkább elfárad és a munkája kaotikus lesz az elkezdett, de be nem fejezett szavaktól, mondatoktól (kétségbeesési hibák).

Hangos olvasás során a kudarcból való félelemből adódik a koncentrációhiány. Általában egy diszlexiásnak nehezebb esik a balról jobbra való olvasás hangosan. Ha a gyermek pontosan akar olvasni, akkor ehhez neki dupla vagy esetleg háromszor annyi időre van szüksége, mint olvasni jól tudó társainak.

Ez természetesen a tanítókat és az osztálytársakat türelmetlenné teszi. Ha megpróbál normális tempóban olvasni, úgy a felolvasása pontatlan lesz, és kombinálja a szöveg tartalmát. Mindkét eset elégtelen eredményhez vezet.

A diszlexia a gyermek számára gond, s mindemellett nem tudja, hogy osztálytársai könnyebben és egyértelműbben érzélik a dolgokat és ezáltal könnyebben tanulnak. Így nem tudja megérteni a kudarcot, csalódik önmagában, és a kritikával, a hibáztatással szemben túlérzékenyvé válik. Általában próbálja elkerülni a sikertelenséggel fenyegető írásbeliséget, mert az bosszússá teszi, félelemmel tölti el.

Egyúttal tartós motivációvesztés jön létre. A gyermek magára hagyottnak érzi magát, és mivel a rutinvizsgálatok során a szemnél és fülnél elváltozásokat nem tárnak fel, a jelenséget nem érti a környezet sem.

A vizuális percepció és az emlékezet két szorosan összefüggő terület, hiszen akkor tudjuk felismerni az elénk kerülő formákat, ha azokra emlékezünk, van róla tapasztalatunk. De akár egy új formát is könnyen fel tudunk térképezni agyunkkal, hiszen összetevőire tudjuk bontani, azaz függőleges és vízszintes vonalakra, körívekre és élekre. Három-hat éves korra a szenzomotoros intelligencia időszakát az észlelés szempontjából a látás vezető szerepe váltja fel. Hat éves életkorban alakul ki a vizuális észlelésnek, az alak, forma, nagyság és szín differenciálásának, felismerésének a képessége, amely az olvasás szempontjából nélkülözhetetlen. Diszlexia-veszélyeztetett gyermekeknél gyakran találkozunk e részfunkciók fejlődési késésével, sérülésével.

Már az óvodai foglalkozások során tapasztalhatjuk, hogy néhány gyermek nem ismeri fel a kívánt formákat, vagy csupán akkor nem ismeri fel, ha azok más helyzetben, más környezetben, nagyságban kerülnek a szeméi elé. Nehézséget okozhat az irány szerinti tájékozódás azokon a feladatlapokon, amiken ezen formák, ábrák szerepelnek. Nem biztos, hogy egy diszlexiás gyermek felismeri az egy irányba néző tárgyakat a lapon.

A nem diszlexiás gyermek pontosabban tud elemezni, mint a diszlexiás. Pl. ha egy P betűt kell felismerniük, akkor a nem diszlexiás azonnal realizálja, hogy milyen elemekre lehet bontani a formát a már előzőekben említett szempontok szerint (függőleges és vízszintes vonalak, körívek, élek, a betű szára lefele, vagy felfele mutat). A gyors és pontos vizuális analízis segíti a nem diszlexiás gyermeket a felismerésben.

A diszlexiás gyermekek vizuális percepciója gyenge. Nehezen tudják a lényeges és lényegtelen elemeket megkülönböztetni, a részeknek az egészszel való összefüggését helytelenül ítélik meg, s a részeknek az egymáshoz való viszonyát is rosszul látják. Így a vizuális érzékelés nem igazán fejleszti az emlékezést, nincs kapaszkodójuk vizuális emlékezetük számára, amely a forma megjegyzését, könnyed felidézését segítené, differenciálását biztosítaná a hasonló formáktól. A bizonytalan megjegyzés, felidezés csak bizonytalan emléknymokat hagy maga után, s így pl. a hasonló p, b, d betűk olvasása is könnyen átcúsúzhat egymásba, összekeveredhet.

A motorikus, akusztikus és vizuális észlelés, emlékezés funkcióinak összerendezetten, integráltan kell működniük ahhoz, hogy az iskolai előrehaladás sikeres legyen.

INFORMATIKAI ALAPOK

Szoftverágensek

Az American Heritage Dictionary az ágenseket a következőképpen definiálja: „valami, ami cselekszik, hatalma vagy hatásköre van a cselekvésre... mást képviselve”. A kifejezés a latin „agere” melléknévi igenév formájából származik, melynek jelentése: vezetni, irányítani, cselekedni, tenni. Egy, sok ágenskutató számára elfogadható, specifikusabb definíciója az ágensnek: szoftver entitás, amely folyamatosan és függetlenül működik egy bizonyos környezetben, gyakran tartalmaz más ágenseket és folyamatokat (Shoham, 1997). A folyamatosság és függetlenség követelménye abból származik, hogy egy ágens képes legyen a környezeti változásokra válaszoló tevékenységek végrehajtására rugalmasan és intelligens módon, emberi irányítás vagy beavatkozás nélkül. Ideális esetben egy ágens, mely hosszú időn keresztül működik egy környezetben, képes tanulni a tapasztalataiból. Ezen kívül egy olyan ágenstől, amely más ágensekkel és folyamatokkal egy környezetben működik, elvárhatjuk, hogy képes legyen kommunikálni és együttműködni a többi ágenssel (Bradshaw, 1997).

Szakértő szoftverágensek

A szakértő szoftverágens egy olyan programrendszer, amely az emberi problémamegoldást próbálja meg modellezni a tárolt ismeretanyag alapján. Ez az ismeretanyag egy szűk területre, jelen esetben a diszlexia lehetőségének megállapítására vonatkozik. Az emberi problémamegoldáshoz hasonlóan a programrendszer végeredményként valamilyen szakvéleményt, döntést vagy elemzést ad.

Minden szakértő ágens három fő komponensből áll, úgymint *tudásbázis, következtető gép, felhasználói interfész*. A tudásbázis tartalmazza mindazon ismereteket, tényeket, szabályokat, heurisztikákat, amelyek szükségesek ahhoz, hogy helyes következtetéssel éljen a rendszer. A következtető gép határozza meg azon új ismereteket, amelyeket a tudásbázisban tárolt ismeretekből és a felhasználtól kapott információkból nyerhetünk ki. A felhasználói interfész pedig segít a kommunikációban, a magyarázatadásnál és a rendszer használatánál.

Több előny is származik abból, ha egy problémacsoport megoldására szakértő ágenszt alkalmazunk. Például, hogy a rendszer több feladatot képes egymás után megoldani illetve, hogy ugyanazon feladat többszöri megoldása mindig ugyanazt az eredményt szolgáltatja az emberi végkövetkeztetések sokszínűségével ellentétben. Továbbá előnynek tekinthető, hogy a rendszer alkalmazása nem ütközik igazán sem helyi, sem időbeni korlátokba.

A szakértő ágensekkel szemben támasztott követelmények első helyén az kell álljon, hogy a rendszernek ugyanazt a tudást kell reprezentálnia, mint az adott szakterület

képviselőinek. A második elvárás, hogy a rendszer hibás vagy hiányos adatok esetén is megfelelő reakcióval él magas szintű képességek birtokában.

Széles körben alkalmazhatóak a szakértő szoftverágensek, hiszen következtetéseket, megállapításokat az ellenőrzés, tervezés, diagnosztizálás, oktatás, értelmezés, magyarázat, monitorozás, előrejelzés szelekció során képesek meghatározni (*Russel és Norvig*, 2000). Jelen cikk éppen a diagnosztizálásra koncentrált, hiszen az egyes, a későbbiekben kifejtett megfigyelési területek válaszaiból képes előre jelezni a diszlexia meglétének lehetőségét.

A diszlexia lehetőségének vizsgálatát ellátó szakértő szoftverágens létrehozásánál a fő szempont az volt, hogy az elsőosztályos gyermekekkel foglalkozó tanítók nem feltétlenül képesek kiszűrni az olyan tanulókat, akiknél az olvasás a későbbiekben gondot okozhat, így tehát egy olyan ágens létrehozása volt a cél, amely segítőkezetet nyújt a pedagógustársadalom e korosztállyal foglalkozó rétegének, esetleg a szülőknek.

A DISZLEXIA LEHETŐSÉGÉNEK MEGÁLLAPÍTÁSÁRA SZOLGÁLÓ PROTOTÍPUS LÉTREHOZÁSA

Tudásbázis kialakítása a diszlexia felmérésének vizsgálatához

A 2. fejezetben a diszlexia jeleinek előfordulási területeit azonosítottuk be, amelyekből jól meghatározható tizenegy jelcsoport, melyeket egy tanító könnyen megfigyelhet a környezetében élő gyermekeken.

A tizenegy terület az alábbi:

1. *Nagymozgások*: Ez a terület a gyermek jól megfigyelhető "nagy" mozgásaira vonatkozik. Figyelembe kell venni a gyermek koordinációs képességeit, mennyire ügyes a labdajátékok során. Ha a gyermek bizonytalan már ezeknél a nagy ívű mozdulatoknál is, akkor valószínűsíthető, hogy a kisebb, úgynevezett finommotoros mozgásai is koordinálatlanok lesznek.
2. *Oldaliság (Laterális)*: Az oldaliság kérdése szintén fontos, hiszen ha az oldaliság nem alakul ki megfelelően, - azaz a gyermek például bal kézzel ír, azonban jobb kézzel nyúl a tárgyakért - akkor a gyermek a tapasztalatok alapján könnyen diszlexia-veszélyeztetetté válhat. A legsúlyosabb eset a keresztoldaliság, mint például a bal láb és a jobb kéz preferenciája. A veszélyeztetettség kérdésénél nem lehet nagy súlya annak a keresztoldaliságnak, mint például a jobb szem és bal kéz használata.
3. *Térorientáció*: Ezen koordinációs zavarok arra engednek következtetni, hogy a gyermek az olvasás vagy írás tanulása során könnyen felcserélheti a képileg közelálló betűpárokat, úgy mint pl. d-b, p-q, u-n.
4. *Koncentráció, figyelem*: A gyermek olvasás órán való figyelmetlensége is utalhat diszlexiára. A kérdéskört nagyon óvatosan kell kezelni, mivel a figyelem elkalandozása akár a gyermek téma iránti érdektelensége miatt is lehet. Azonban a figyelem elkalandozása jelentheti azt is, hogy a gyermeknek nehezebb esik az egyes betűket felismerni vagy éppen megformálni.
5. *Finommotoros mozgások*: A gyermek finom, apró mozgásaira koncentrált ez a kérdéskör. Például a ceruzafogás vagy akár a reggeli öltözködésnél a ruha begombolása figyelmeztető jellé válhat.
6. *Ritmusrészék*: Talán a legfontosabb terület ez. A legújabb kutatások szerint a diszlexiás gyermekek nagy részénél tapasztalható az ütemérzék hiánya (*Smythe*, 2004).
7. *A beszéd*: A beszéd területe az, ahol a legtermészetesebb közegben érhetjük tetten a diszlexiára utaló jeleket, hiszen a gyermek mindennapi kommunikációja során

megtapasztalhatjuk, hogy az ismert szavakat hogyan is tárolta el emlékezetében, vajon az egyes hangzók tiszták-e. Továbbá megtapasztalhatjuk, hogy a gyermek szókincse a kortársaihoz mérten mennyire megfelelő.

8. *Formaészlelés:* Ha egy gyermek formaészlelése nem megfelelő, könnyen előfordulhat, hogy amikor a betűket tanulja, akkor a betűket alkotó formákat rosszul jegyzi meg. Ennek következménye lehet a helytelen betűfelismerés az olvasás során.
9. *Emlékezet:* Ha a gyermek nem képes a megfelelő felidézésre, akkor nehezebbé eshet megjegyezni az egyes betűk formáját, képét. Továbbá fontos tudni, hogy mennyire könnyen vagy nehezen tanul verseket, vagy vissza tud-e mondani 4-5 számot egymás után.
10. *Olvasás:* Az olvasás során jelentkezhetnek leginkább a diszlexia jelei. Azonban mint az előzőekben oly sok kérdést, ezt is óvatosan kell kezelni, hiszen lehet, hogy kezdetben csupán a gyakorlatlanság okozza, hogy a gyermek a többiekétől eltérő tempóban, több tévesztéssel olvas. Ez önmagában nem diszlexia, további vizsgálatokat igényel a kérdés eldöntése.
11. *Magatartás:* Ez a külső viselkedési tünetegyüttes mutathatja a gyermek olvasáshoz, íráshoz való kötődését. Ha a gyermeket sok kudarc éri olvasás közben (pl. nem ismer fel egyes betűket) próbálja érthető módon kerülni az olvasási feladatokat.
12. Az elkészült alkalmazás jelen állapotában a fenti tizenegy területre adott válaszokból következtet a diszlexia lehetőségére. A tények leírására a kijelentő kalkulus módszerét választottuk, mivel a gyermekkel foglalkozó személy megfigyeléseit kell a következtetésnél alapul venni. Így minden egyes területhez kérdéscsoportokat rendeltünk hozzá úgy, hogy a kérdéscsoporton belül is az egyes kérdések súlyai eltérőek.

A tudásbázist kialakító teszt felépítése, kérdései és azok értékelési szempontjai

A teszt szerkezetét úgy építettük fel, hogy az előzőekben felsorolt területek mindegyikéhez tartozik egy általános magyarázó szöveg, amivel a felhasználó találkozik a teszt kitöltése során. A felhasználónak minden területen kérdésekre kell válaszolnia. A feltett kérdések egyszerűek, a hétköznapi élet során könnyen megfigyelhetők. Területenként eltérő a kérdésszám, súlyozva ezzel is a kérdéscsoportok fontosságát. Az egyes kérdések is különböző súllyal vesznek részt az értékelés, a következtetés folyamatában. Hosszas megfontolás után három fokozatot különböztettünk meg, így az egyes kérdéscsoportokban feltett eldöntendő kérdések fontosságát egymáshoz viszonyítva lehet kifejezni a diszlexia lehetséges meglétének tekintetében. A három fokozat biztosítja azt, hogy a rendszerbe nem kerültek be olyan kérdések, amelyek bagatelizálhatóak a fontosabb információkhoz képest, így minden egyes válasz a következtetés szempontjából mérvadó lehet.

Most pedig következzenek a feltett kérdések területenkénti csoportosításban. A zárójelben látható szám az adott kérdés súlyát fejezi ki oly módon, hogy minél nagyobb a megadott szám, annál nagyobb súllyal szerepel a következtetés során. Ha a válasz egy normál képességű gyermek szintjét mutatja, úgy ezt a választ neutrálisnak tekintjük az értékelés során.

A kérdésekre adható válaszok közül az szerepel eleve megjelölten, ami egy gyermek normális fejlődése során elvárható lenne. Ha a másik lehetőséget választja a felhasználó, akkor az adott válasz az előzetesen meghatározott súllyal fog szerepelni a kiértékelés során. Tehát a teszt feltételezi, hogy egy normál képességű gyermekről szeretnénk a megfigyelések által eldönteni a diszlexia fennállásának lehetőségét. Éppen ezért az igen-nem válaszpárok helyei folyamatosan cserélődnek akár az egyes kérdéscsoportokon belül is.

1. Nagymozgások:
 - a. A gyermek mozgásokat ügyesen utánoz, mozgása jól koordinált, biztos? (2)
 - b. Gyakran elesik, beüti magát, nekimegy asztalnak, széknek? (2)
 - c. Ügyesen rúg, kap el, dob, pattogat labdát? (2)
2. Oldaliság (Lateralitás):
 - a. Az első kérdés: a gyermek jobb vagy bal kezes? (A válasznak megfelelően folytatódik a tesztsor.)
 - i. Tárgyakért bal kézzel nyúl? (3)
 - ii. Tárgyakért jobb kézzel nyúl? (3)
 - b. Írásnál vagy rajzoláskor váltogatja a jobb és bal kezét? (2)
3. Térorientáció:
 - a. Téveszti a bal és a jobb irányokat? (2)
 - b. Nehezen találja meg a feladatot könyvekben, munkafüzetekben? (2)
 - c. Tornaórán, játékok közben nehezen érti meg az utasításokat? (1)
/Pl.: Állj el! Menj a szekrény és a bordásfal közé.../
 - d. Fordítva ír le számokat, betűket? (2)
/Pl.: A kilences szára másik irányba néz.../
4. Koncentráció, figyelem:
 - a. Sok tévesztéssel keres egy bizonyos betűt szövegben, újságban? (3)
/Pl.: k vagy u vagy b betű keresése/
 - b. Gyakran elkalandozik figyelme, mással foglalkozik olvasás órán? (1)
 - c. Írás, olvasás során könnyen, hamar elfárad? (2)
5. Finommotoros mozgások:
 - a. Ceruzafogása megfelelő, vonalvezetése határozott, nyomatéka egyenletes? (1)
 - b. Írásképe rendezett és a betűk sorközből maradnak? (2)
 - c. Ügyesen öltözködik? (1)
/Tud cipőfűzőt kötni, inget gombolni.../
6. Ritmusérzék:
 - a. Nehezen tapsol vissza egyszerű ritmust? (3)
 - b. A szavakat tudja szótagolva helyesen letapsolni? (3)
 - c. Nehezen folytatja a ritmikus mintákat kirakni, rajzolni? (3)
 - d. Hosszú-rövid zörejeket, hangokat, magánhangzókat nehezen különbözteti meg hallás után? (2)
7. A beszéd:
 - a. A gyermek nehezen fejezi ki magát? (3)
/Válaszai rövidek, gyakran keresgéli a szavakat./
 - b. Hangjai nem tiszták, még beszédhibája van? (3)
 - c. Nehezen különböztet meg hasonló hangzású szavakat hallás után? (3)
/domb-gomb; vonal-fonal; toll-tol/
 - d. Nehezen alkot fogalmakat? (3)
/ásó, gereblye → szerszám/
 - e. Nehezen alkot mondatokat képekről, inkább csak felsorolást ad? (3)
 - f. Nehezen választja le, hogy milyen hanggal kezdődik és végződik egy szó? (3)
8. Formaészlelés:
 - a. Keveri a hasonló alakú betűket? (2)
/b-d; u-n; f-t/
 - b. Nehezen másol le egyszerű formákat vagy rosszul rakja ki őket pálcikából? (2)

9. Emlékezet:

- a. Vissza tud mondani 4-5 hangot vagy számot helyes sorrendben? (2)
- b. Nehezen jegyez meg mondókákat, verset, mesét? (2)
- c. Összetett utasításokat részben jegyez csak meg? (2)
- d. Nem veszi észre a változást tárgyak, képek csoportjában (7-8 elemű halmaz)? (3)

10. Olvasás:

- a. A többiekhez mérten lassan olvas? (2)
- b. Gondot jelent a betűk összeolvasása szótagokká, szavakká?(3)
- c. A betűket gyakran téveszti? (2)
- d. Kihagy, betold, ismételi betűket, szótagokat? (2)
- e. Megérti amit olvasott? (3)

11. Magatartás:

- a. Nem szereti az olvasást, tiltakozik az olvasási, írási feladat ellen? (2)
- b. Gyakran panaszkodik hasfájásra vagy fejfájásra? (2)
- c. Dühösen reagál kudarcélmény esetén olvasásnál, írásnál, hamar feladja a próbálkozást? (3)

A teszt értékelési rendszere

Mint azt az előzőekben említettünk, tizenegy fontos területet lehet meghatározni a diszlexia lehetőségének megállapítása szempontjából, azonban ezek a területek nem egyenrangúak a végkövetkeztetést illetően. A legnagyobb gondot az okozta az értékelési rendszer kialakítása során, hogy hogyan lehet ezeket a különböző súllyal szereplő területeket egymással összehasonlíthatóvá tenni annak érdekében, hogy meg lehessen állapítani, hogy az adott válaszok függvényében a kérdéscsoportok válaszaiból kapott értékek mennyire szórnak. Ez azért lényeges szempont, mivel a diszlexia lehetősége nagyon nagy abban az esetben, ha a tesztben szereplő területeken elért eredmények közel állnak egymáshoz, tehát például a vizsgált gyermek koordinációs képessége hasonló mértékben marad el a normálístól, mint az olvasás vagy akár a finommotoros mozgások területén.

Első lépésben meghatároztuk, hogy átlagosan milyen értéket kaphat egy válasz bármely kérdéscsoportban. A következő módon jártunk el: meghatároztuk az egyes területeken szereshető átlagértékeket. (Például a *Nagymozgások* területén három kérdésből 6 pont szereshető, így az átlagérték kettő lett. Ugyanezzel az eljárással az *Emlékezet* területén a négy kérdésre adható maximális pontszám kilenc, így az átlagérték 2,25 lett.) Nevezzük ezt az értéket az egy kérdésre adott átlagértéknek és jelöljük „K”-val.

Második lépésben megadható, hogy egy területen átlagosan mennyi kérdésre válaszolt a felhasználó a kérdéscsoportban szerzett érték alapján. Megadása a következő módon történt: az adott területen szerzett értéket osztottuk el az első lépésben meghatározott egy kérdésre adott átlagértékkel. Ekkor megkaptuk, hogy az adott kérdéscsoportban mennyi kérdésre is válaszolhatott a felhasználó (ez nem feltétlenül egyezik meg a ténylegesen megválaszolt kérdések számával, több is lehet annál). Ha az adott területen szerzett értéket jelöljük „SZ_n”-nel, akkor a „V_n”-nel jelölt átlagos kérdésszámot így adhatjuk meg: $V_n = SZ_n / K$, ahol „n ∈ [1..11]”. Tehát „n” jelöli, hogy éppen melyik területre vonatkoztatjuk az értékeket.

Ahhoz, hogy az összes területet összehasonlíthatóvá tehesük, ezt az adott területen megválaszolt átlagos kérdésszámot kellett tovább normalizálni az egyes területeken lévő összes kérdések számával. Ha ez utóbbi számot „O_n”-nel jelöljük, akkor az adott területre számított normalizált érték („H_n”) az alábbi módon számítható ki:

$$H_n = \frac{SZ_n / K}{O_n}$$

A teszt kiértékelése és a következtetések megadása az összes területre számított normalizált értékek meghatározása után következhet. Ezekből az értékekből származtatunk egy egyszerű átlagot („A”), amelyhez a későbbiek folyamán hasonlítjuk az elért eredményeket. Feltűnhet, hogy ez az átlagérték nem egy olyan szám, amely minden gyermek esetén azonos hasonlítási alapot szolgáltat, hiszen az elért értékekből származtatjuk. Mivel a kérdéses területeket úgy határoztuk meg, hogy azok a gyermek képességeit társaiéhoz viszonyítják, így a közösséghez való viszonyítás már ekkor megtörténik. A diszlexia lehetőségének megadása pedig abban az elemzésben rejlik, hogy ezek a területek hogyan viszonyulnak egymáshoz, így itt nem egy normális képességű gyermek adja a standardot, hanem ez mindig változik a vizsgált személyhez kötve.

Hogy a szórást meg tudjuk adni, meg kell határoznunk az elért normalizált értékek közti átlagos különbséget („D”).

$$D = \frac{\sum_{n=1}^{10} \text{abs}(H_{n+1} - H_n)}{11}$$

A kiértékelés során azok a területek, amelyek az átlagérték felett vannak a szóráson túl, nagyon fontosak a következtetés levonását tekintve. Az alábbi öt terület különösen fontos a következtetés során: *Olasás, Beszéd, Ritmusérzék, Térorientáció, Oldaliság*.

Ha ezek közül már egy-két terület kiemelkedik, úgy már a diszlexia lehetősége nagyon erős. Azonban figyelmesen kell kezelni a kérdést, hiszen ha csupán az *Olasás* és a *Térorientáció* területe kimagasló, akkor a téri tájékozódási zavar lehetősége áll fenn és ez irányban kell a vizsgálatokat tovább folytatni szakember bevonásával. Azonban ha ehhez a két területhez csatlakozik még egy a fentebb felsorolt kiemelt területek közül, akkor már a diszlexia lehetősége áll fenn. Ugyanígy meghatározható a vizual-motoros-koordinációs zavar lehetősége, ha a következő területen elért eredmények ugranak ki: *Figyelem, Nagymozgás, Ritmus, Térorientáció, Formaészlelés*. A figyelem-, emlékeztetési zavar lehetősége is megállapítható akkor, ha az eredmények ezt igazolják a *Figyelem, Ritmus, Formaészlelés, Emlékeztetési* területén. Olyan eset is előfordulhat természetesen, amikor a vizsgált személy csupán a *Beszéd* területén mutat elmaradást. Ez további vizsgálatokat kell, hogy maga után vonjon, azonban nem egyértelműen adható meg a diszlexia lehetősége. Sokan gondolnák úgy, hogyha csak az *Olasás* készsége nem alakult ki eléggé, akkor már diszlexiáról beszélhetünk. Ez nem így van, hiszen ez lehet akár a rossz olvasási technika hibája vagy akár a kevés gyakorlás eredménye. Éppen ezért a teszt végén javaslatokat talál a felhasználó, hogy mely területek kiemelkedőek, melyekkel kell komolyabban foglalkozni, illetve azon területeket is láthatja, amelyekkel valószínűleg a közeljövőben problémája lesz a vizsgált alany. Kiemelkedő területeknek azok számítanak, ahol az elért eredmények magasabbak az átlagérték és az átlagkülönbség összegénél ($H_n > A + D$). A veszélyeztetett területeket azon területek jelentik, amelyeknél elért értékek az átlagérték, illetve az átlagérték és átlagkülönbség összege között szerepelnek ($A < H_n < A + D$). A rendszer figyelmeztetéssel is él abban az esetben, ha csupán egy-egy terület ugrik ki a többitől. A teszt legvégén megjelennek az elért eredmények és az értékelés a következtetéssel együtt. Ezt a végső összegzést akár ki is nyomtathatja a felhasználó, s a további vizsgálatokat elvégző szakembernek adhatja át, aki az eredmények ismeretében már koncentráltabban tudja a vizsgálatokat elkezdni.

Implementálás

A kialakítandó program fejlesztői környezetének megválasztása azért jelentett nehéz feladatot, mivel az újszerű megoldást alkalmazó rendszer integrálja a kliens-szerver programozást, webes alkalmazást és egy honlap mesterséges intelligenciával, tudással való ellátását. Erre a feladatra a *DecisionScript* programcsomag bizonyult a legalkalmasabbnak (*Vanguard Soft Corp* (2004)).

A *DecisionScript* új irányvonalat képvisel abban a tekintetben, hogy a fejlesztőnek a program írásával egyidőben lehetősége van kialakítani egy egységes felhasználói interfészt. Ugyanakkor az egyes programrészek fejlesztése mellett átláthatja az eddig elvégzett programozási feladatok szerteágazó egészét. (1. ábra)

1. ábra

A DecisionScript-környezet integrálja az egyes fejlesztői felületeket

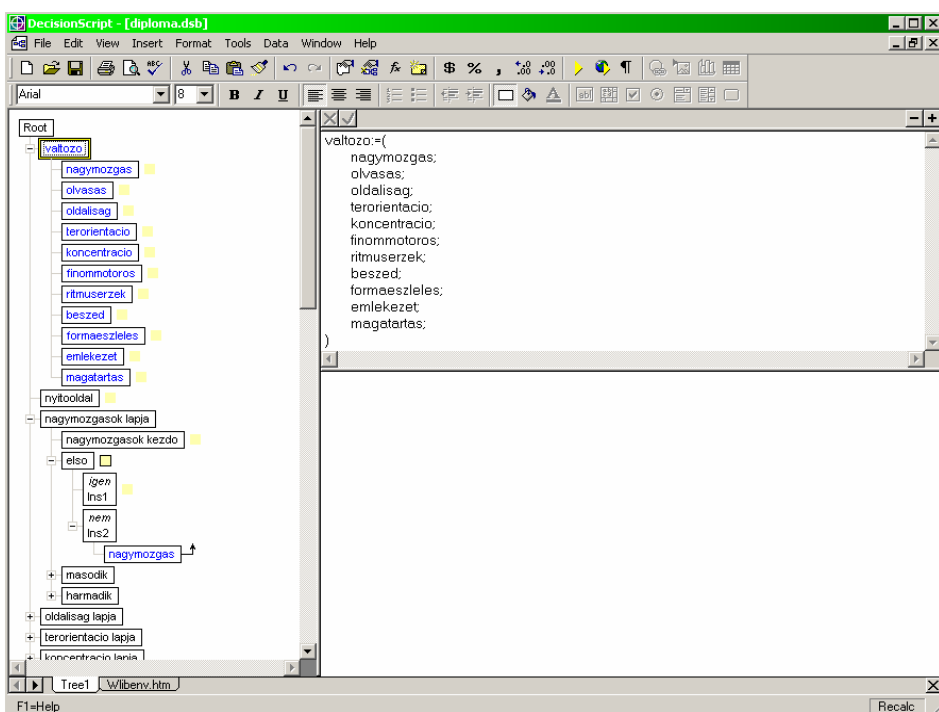


Figure 1: The DecisionScript context integrates individual developer interfaces

A 1. ábra bal oldalán látható függőleges ablakban a programozó áttekintheti eddigi munkájának szerkezetét. A Rejtés/Kibontás opció lehetőséget biztosít arra, hogy az éppen fejlesztési szempontból feleslegesnek tartott részek ne zavarják a munkát. Programozás szempontjából fontos kiemelni, hogy ebben a fejlesztői környezetben objektumokat hozunk létre és programozzuk őket. Egy-egy objektum megfeleltethető egy-egy függvénynek is, amelyek paraméterei újabb objektumokat, függvényeket eredményeznek. Minden egyes ilyen objektumhoz tartozik egy felhasználói felület is, amelyet vagy megjelenítünk az adott helyen, vagy nem.

A szerkezeti ablakban jól látható, hogy mely objektumok mely más objektumokkal állnak kapcsolatban. Egy, a dobozból kiinduló folytonos vonal szimbolizálja, hogy ahhoz az aktív objektumhoz mely objektumok tartoznak szorosan. Ha egy objektum programozásakor hivatkozunk egy már meglévő objektumra annak azonosítójával, úgy azt a rendszer egy, a dobozból jobbra kiinduló, felfele mutató nyilacskával jelzi.

Ennek a hivatkozás-jelző rendszernek köszönhetően maga a szerkezeti ábra, struktúra nem válik túlzottan bonyolulttá, hiszen ha egy már megalkotott szerkezeti egységre (döntési fa részre) szeretnénk hivatkozni, úgy azt nem kell erre a helyre beszúrni. A rendszer így a hivatkozott helyen folytatja a működését abban az esetben, ha az objektum nem csupán értéket hordoz. Egy objektum tárolhat csak értéket is, ekkor a program működésének alkalmával nem vehető észre semmi a felhasználó számára az objektum lefutása során, csupán a munkamemóriában változhatnak meg az adatok.

A WYSIWYG (What You See Is What You Get) editor (2. ábra) ablakban adhatjuk meg az adott objektumhoz rendelni kívánt felhasználói felületet, ami a program futása során megjelenik a képernyőn.

2 ábra

A DecisionScript program WYSIWYG editor ablaka



Figure 2: The WYSIWYG editor window in DecisionScript

A programhoz külön környezeti paraméterek állíthatók be (3. ábra).

Az első opció (*HTML envelope*) tulajdonképpen egy template kiválasztását teszi lehetővé, amely egységes keretrendszerbe foglalja a felhasználói felületet. Egy-egy ilyen template egy html forráskód, gerinc, amelyet a DecisionScript tölt fel tartalommal futás

közben. A második opció (*Include library files*) különálló, már előre megírt script-állományok importálását teszi lehetővé, amelyeket aztán, mint beépített függvényeket lehet használni. Ez lehetőséget nyújt a moduláris fejleszthetőségre. A harmadik opció esetében (*Title header and footer*) meghatározható a program futás közben megjelenő neve, illetve minden egyes oldalon megjelenő élőfej és élőláb. Ez is az egységes felhasználói interfész kialakítására szolgál.

3. ábra

Módosítani kívánt környezeti paraméterek kiválasztása

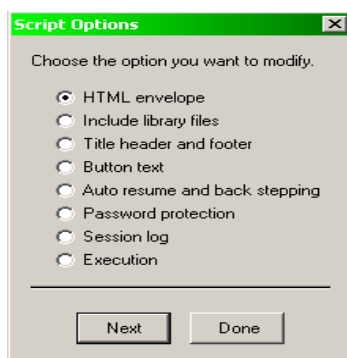


Figure 3: Selection of environmental parameters to be modified

A negyedik lehetőségnél (*Button text*) módunkba áll változtatni az előre definiált gombok feliratozását (például magyaríthatunk). A következő választásnál (*Auto resume and beack stepping*) beállítható, hogy egy megszakadt kapcsolat esetén a felhasználó a programot folytathassa a szakadás előtti állapottól, avagy sem.

Megadhatunk jelszavas védelmet is (*Password protection*), hogy az esetlegesen fejlesztés alatt álló programot csak bizonyos felhasználók tekinthessék meg. Módunkban áll *naplófájl* létrehozása is, amely tárolja a kapcsolat alatt megtekintett kérdéseket és azokra adott felhasználói válaszokat is. Ezzel lehetővé téve az egyes esetek nyomkövetését. A utolsó opció (*Execution*) szolgál arra, hogy megadjuk azt az időt, ami után a program megszakítja a kapcsolatot a klienssel, ha nincs adatforgalom a két gép között.

A fejlesztési lehetőségek legnagyobb részét az ODBC-n keresztüli kapcsolatteremtés jelenti más adatbázisokkal, ahonnan SQL nyelv segítségével nyerhetünk ki információkat táblázatokból, adatállományokból (például Access, Excel, dBase, FoxPro, Oracle).

A program futása során minden egyes olyan objektum le fog futni, amely közvetlenül csatlakozik a „gyökérhez” (4. ábra).

Ezek az objektumok lehetnek szülők (amit egy + jel jelez az azonosítójuk előtt), vagy gyermekobjektumok, amelyek természetesen csak a megfelelő esetben futhatnak le.

A fejlesztői környezet lehetőséget nyújt arra, hogy különálló szerver-program segítségével nélkül lehessen tesztelni az építés alatt álló rendszert. Azonban alapvetően a nyelvet kliensszerver alapú működésre fejlesztették. Éppen ez adja a *DecisionScript* egyik legnagyobb erősségét, hogy az általa kifejlesztett rendszer semmilyen különleges környezetet nem igényel a futtatáshoz a felhasználó részéről. A felhasználó egy böngészőn keresztül csatlakozhat a szerverhez -, ami egy webes kiszolgáló - és használhatja a rendszert.

4. ábra

A „gyökérhez” kapcsolódó objektumok a szerkezeti ablakban

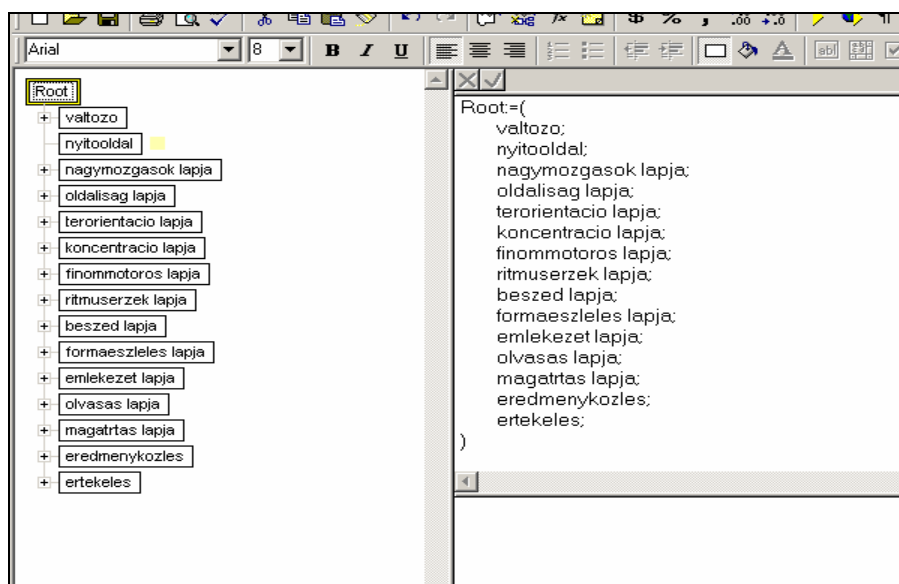


Figure 4: Objects related to the root in the structural window

Amikor egy felhasználó csatlakozik az adott oldalra, ahol elhelyeztük előzőleg a *DecisionScript*-tel megírt rendszerünket, a szerver feldolgozza a beérkezett kérést és egy megfelelő html kódot generál, amelyet bármely böngésző meg tud jeleníteni. Ez a kód tartalmazhatja a html szabvány összes lehetséges elemét.

Amikor fejlesztőkörnyezetben teszteljük a készülő programunkat webes megjelenítéssel, úgy a saját, kliens gépen alapértelmezettként beállított böngészőben fog futni a rendszer. Azonban lehetőségünk van arra is, hogy kikerülve a böngészős tesztelést, egy egyszerű, szokványos ablakban futtassuk az alkalmazást.

KONKLÚZIÓ

Azt várnánk, hogy a rendszer a diszlexia lehetőségének létét vagy nem létét adja meg végeredményképpen. Ez az elvárás egyértelmű, azonban a vizsgált terület nem engedi meg az ilyen egyértelmű kijelentéseket. A diszlexia-veszélyeztettség kialakulása igencsak összetett folyamat, amelynek egy része a fentebb felsorolt tizenegy terület. De ne feledjük, hogy ezen területek jól megfigyelhetőek a gyermek aktív közreműködése nélkül is! Tehát ezen tizenegy kérdéscsoport vizsgálatának eredménye csupán az lehet, hogy a gyermek mennyire lehet veszélyeztetett és nem az, hogy veszélyeztetett-e.

A kiértékelés során a rendszert felkészítettük arra, hogy meghatározza azon területeket, amelyeknél a vizsgált gyermek készségeinek kialakultsága nem megfelelő (5. ábra). Ezen területek meglétéből vagy nem létéből lehet meghatározni, hogy a gyermek diszlexiás lehet-e avagy sem.

5. ábra

Kiértékelés és végső következtetés oldala



Figure 5: The page of evaluation and final conclusion

A teszt az értékelés során felhívhatja a figyelmet a gyermek esetleges téri tájékozódási zavarára, vizual-motoros-koordinációs zavarára, motoros-koordinációs zavarára, figyelem, illetve emlékezet zavarára, általános képességeinek elmaradására, beszédfejlődésének megkésésére, rossz olvasási technikájára. Továbbá a rendszer megadhatja azon területeket is, amelyekre a jövőben fokozottan oda kell figyelni a gyermek készségeinek fejlődése folyamán.

A kialakított rendszer előnye, hogy a diszlexia lehetőségének vizsgálata során felfedezhető egyéb, más tanulási nehézségekkel kapcsolatos jeleket is felismeri, és erre figyelmezteti is a felhasználót (6. ábra). Így adódhat olyan eset is, amikor a vizsgálat diszlexia lehetőségének megállapítása miatt kezdődik el, azonban eredményeképpen nem diszlexiára utaló következtetések születnek.

6. ábra:

Figyelmeztetés valamely terület elmaradottságára vagy tünet együttesek egyidejű létre

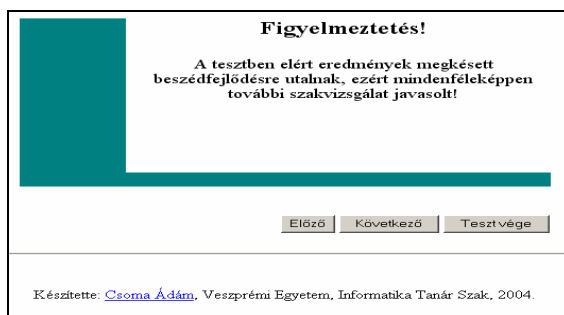


Figure 6: Warning about underdevelopment in an area or existence of joint symptoms

A rendszer továbbfejlesztése során lehet összetettebb vizsgálati struktúrát is összeállítani, például az anamnézis felvételével, annak kiértékelésével. Az anamnézis megléte esetén, illetve az intelligencia felmérésével már megvalósíthatónak tűnik egy komplex döntési fa struktúrájának a kialakítása is, hiszen ezen területekre vonatkozó adatokból már levonhatóak azok a következtetések, hogy a vizsgált gyermeknek mely tanulási területeken lehetnek a jövőben nehézségei. Ebből adódóan egy olyan irányvonal kialakítása is elképzelhető, hogy a gyermek komplex tanulási nehézségekre vonatkozó vizsgálatát egy ilyen szakértő ágenssel lehessen elvégezni, amelynek egyik eleme lehetne a most elkészült modul is.

Az alkalmazás validálása több iskolában megtörtént. Jelenleg a megfigyelt gyermekeket hagyományos diszlexia-vizsgálatnak is alávetik, a tesztben elért eredmények hitelesítése végett.

IRODALOM:

- Bradshaw, J.M. (1997). Software Agents, Menlo Park, Calif.: AAAI Press, 3-46.
- Boldureva, T., Isekova, O.R. In. (Eds.) Salter, R. and Smythe, I. (1997). The International Book of Dyslexia. London, WDNF and EDA, 148-152.
- Gyarmathy É. (1996). Tanulási zavarokkal küzdő tehetséges gyerekek azonosítása, Ph.D. thesis, Kossuth Lajos Egyetem, Debrecen.
- Gyarmathy É., Smythe, I. (2000) Többsz nyelvűség és az olvasási zavarok. In: Erdélyi Pszichológiai Szemle, December, 63-76.
- Kirk, S. A., Bateman, B. (1962). Diagnosis and remediation of learning disabilities. In: Exceptional Children, 29. 73-78.
- Ligeti R. (1967). Gyermekek olvasászavarai, Akadémiai Kiadó : Budapest, 111-123.
- Lohmann, B. (1996). Diszlexiások az iskolában, Akkor Kiadó : Budapest, 15-26.
- Russel, S. J., Norvig, P. (2000). Mesterséges intelligencia modern megközelítésben (Artificial Intelligence A Modern Approach) Panem-Prentice Hall, 69-82.
- Sarkady, K., Zsoldos, M. (1992). Konceptuális kérdések a tanulási zavar körül. In: Magyar Pszichológiai Szemle, 3-4. 259-270.
- Selikowitz, M. (1997). Diszlexia, Medicina Könyvkiadó Rt. : Budapest, 49.
- Smythe, I., Everatt, J. (2000). Dyslexia Diagnosis in Different Languages. In: Peer L and Reid G. (Eds), Multilingualism, Literacy and Dyslexia (2000), David Fulton Publishers : London.
- Shoham, Y. (1997). An Overview of Agent-oriented Programming. In: Software Agents, (Ed.: J. M: Bradshaw) Menlo Park, California : AAAI Press, 271-290.
- Torda, Á. (1999). Korai kutatások az olvasás- és írászavar tárgykörében, Fejlesztő pedagógia, Különszám, 20-22.
- Vanguard Software Corporation, 21. April, 2004, <http://www.vanguardsw.com>.
- Smythe, I., 21. April, 2004, <http://web.ukonline.co.uk/wdnf>.

Levelezési cím (*Corresponding author*):

Starkné Werner Ágnes

Pannon Egyetem, Matematikai és Számítástechnikai Tanszék,
8200 Veszprém Egyetem u. 10.

*Pannon University, Department of Mathematics and Computing
H-8200 Veszprém, Egyetem u. 10.*

Tel.: 36-88-422-022/4787 Fax: 36-88-624-521

e-mail: wernera@almos.vein.hu



Karbantartási költségkeret-felosztás számítógépes támogatása

Heckenast¹ T., Csonka¹ B.Gy., Nagy² V.

¹Széchenyi István Egyetem, Informatika Tanszék, 9026 Győr, Egyetem tér 1.

²Széchenyi István Egyetem, Közúti és Vasúti Járművek Tanszék, 9026 Győr, Egyetem tér 1.

ÖSSZEFOGLALÁS

A cikkben egy nagy közlekedési vállalat járműpark karbantartási költségeinek elemzését és a rendelkezésre álló költségkeret optimális felosztását támogató szoftvermegoldás kerül ismertetésre. A megoldás egy úgynevezett könyvelési modell alapú költségbecslési, elemzési döntéstámogató rendszer, a megrendelő igényeihez igazodó domén-modellre építve. A cikk ismerteti az általunk kidolgozott speciális hierarchikus break-down analízis technikát, melyet költségfa modellnek neveztünk. Bemutatjuk, hogyan alkalmazható ez a modell az adott alkalmazási környezetben - kapcsolódva a specifikus normatív domén-modellhez. A modell operacionalizálásához kifejlesztettünk egy döntéstámogató szoftvert. Bemutatjuk, hogyan képezhető le a költségfa modell egy relációs adatbáziskezelő platformon megvalósított, több-dimenziós adatkockára épített lekérdező rendszerre. A döntéstámogató rendszer bizonyos adatai a megrendelő vállalatirányítási integrált rendszeréből származnak, és egy speciális objektum adatbázist alkotnak. Így az általunk kifejlesztett rendszer az adatalapú döntéstámogató (adatbányászati) eszközökhöz hasonlóan rendelkezik adattisztító és előfeldolgozó megoldásokkal. Ezek a megoldások fontos szerepet játszanak a domén-modell paramétereinek kialakításában. Részletesen ismertetjük a rendszer felhasználói felületét, melyen keresztül felépíthető a költségfa és elvégezhető a normatív modell paramétereinek exploratív érzékenységvizsgálata, valamint az objektum adatbázis elemzése. A felhasználói felület részét képezi többfajta grafikus megjelenítési technika is.

(Kulcsszavak: költségbecslés, költségfa, döntéstámogatás, lekérdezés)

ABSTRACT

Computer supported maintenance cost frame division

T. Heckenast¹, B.Gy. Csonka¹, V. Nagy²

¹Széchenyi István University, Department of Informatics, H-9026, Győr, Egyetem square 1..

²Széchenyi István University, Department of Automotive and Railway Engineering, H-9026 Győr, Egyetem square 1.

In this paper a software solution supporting the analysis of maintenance costs and the optimal division of the available cost frame for a large transport company is presented. The solution is a bookkeeping cost estimation and analysis decision support system custom built for the domain of the customer. The paper describes the special hierarchical break-down analysis technique, which was called cost tree model. It will be demonstrated how the model can be used in the given application context including the nominative domain model. The model is operationalized in a decision support system software. We will present how to map the cost tree onto a multi-dimensional data-cube implemented in a relational database management system. Most of the data come from

the ERP system of the customer and is stored in a domain specific database. So our system can be categorized as a data-oriented decision support system and has data transformation and data cleaning tools as well. These tools play special role in the formation of the parameters of the domain model. We will describe the user interfaces of the system, which enable the construction of the cost tree and the explorative sensitivity analysis for the parameters also can be carried out. The user interface incorporates graphical display techniques.

(Keywords: cost estimation, cost tree, decision support, query)

BEVEZETÉS

A következőkben egy költségelemzési és költség-keret felosztási módszer kerül ismertetésre. A kidolgozott módszer célja, hogy egy rendszer működési és fenntartási költségeinek meghatározása, becslése alapján lehetővé tegye az elkövetkező pénzügyi időszakra rendelkezésre álló költségkeret minél ésszerűbb felosztását. A rendelkezésre álló költség-keret felosztását az általános gyakorlatnak megfelelően a tényleges költségigények arányában célszerű elvégezni.

Kíváncsi, hogy a költségek becslésére, megadására alkalmazott módszer feleljen meg a következő általános szempontoknak: tartalmazza mind a direkt és mind az indirekt költségeket, a költségeket a szervezeti struktúrához igazítva vegye figyelembe, rugalmasan kezelje az érintett időszakra várható költségviselkedést, tegye lehetővé a tényadatokból kirajzolódó ráfordítások valamint szakértői domén modellekkel meghatározott normatív költség igények összehasonlítását.

A költségigények meghatározásának vagy becslésének egy lehetséges módja a hierarchikus költségfa használata. A módszer lehetővé teszi a költségek aggregálását egy meghatározott költségstruktúra mellett. A költségfa (költség hierarchia) a költségelemek egy olyan átfogó reprezentációja, mely a fő költségkategóriákból kiindulva lebontja a költségeket a legszűkebb kategóriák szintjére. Tartalmazza mind a közvetlen, mind a közvetett költségeket. Gyakorlatilag a költségek egy olyan strukturált váza, mely a költségeket addig dekomponálja, míg azok megfelelő pontossággal meg nem adhatók. Ez a megközelítés lehetővé teszi, hogy a költségeket az elemi, jól meghatározott és ténylegesen ismert költségekből aggregáljuk a magasabb szintekre. Az elemzéshez használt költségfa meghatározása és rögzítése a felsőpénzügyi vezetés dolga, de az egyes alsóbb ill. elemi szinten a költségek megadását (meghatározását, becslését) a megfelelő információkkal és tudással rendelkező szervezeti egységek végzik.

A cikkben egy olyan szoftver megvalósítást mutatunk be, mely a megrendelő specifikus doménjéhez igazított költségfa alapú elemzést és költségkeret felosztást tesz lehetővé.

KÖLTSÉGELEMZÉS, BECSLÉS

Működési és fenntartási költségek elemzése

A működési és fenntartási költségek elemzésének célja egy rendszer működtetése során felmerülő költségek racionális felhasználásának meghatározása. A megfelelő megközelítés az, hogy a rendelkezésre álló korábbi működési adatok felhasználásával kialakítsunk egy olyan keretet, mely kielégítően leírja az adott rendszer várható működési feltételeinek megfelelő költséghatásokat. Egy olyan analitikus megközelítést

és reprezentációt kell választani, amely alkalmazkodik az adott rendszer tulajdonságaihoz és képes figyelembe venni az összes releváns költségtényezőt.

A referencia rendszer egy olyan hasonló vagy azonos rendszer melynek működésével kapcsolatban már rendelkezésre állnak adatok. Ez legtöbbször megegyezik a vizsgálat tárgyát képező rendszerrel.

Az elemzés időbeli hatásköre általában a pénzügyi tervezés és elszámolás időintervallumaihoz igazodik. Ritkábban fordul elő, de magában foglalhatja egy rendszer teljes életciklusát is. Az elemzés tartományában tehát a rendszer vagy annak bizonyos részrendszerei az életciklusuk tetszőleges szakaszában lehetnek.

A költségbecslés kritériumai

Mivel a költségmeghatározási helyzetek nagyon különbözőek lehetnek, tekintsünk néhány alapvető szempontot: teljesség, észszerűség, konzisztencia, dokumentáltság. Ezek a kritériumok a költségbecslési modellel szemben is követelményeket támasztanak: konzisztens költségelem struktúra, konzisztens adatok felhasználása, flexibilis becslési technika, felhasználhatóság a tervezési folyamatban.

Költségbecslési technikák

A legtöbb működési és fenntartási költségelemzés három becslési technika kombinációját használja: parametrikus-, analóg- és mérnöki becslés.

A *parametrikus becslés* lényege, hogy az egész rendszert, megvalósítást, működtetést jellemző költségeket egyenletekkel, képletekkel határozzuk meg. Ez a technika általában egy rendszer beindításakor vagy akkor használatos, ha csak korlátozottan állnak rendelkezésre technikai és költség adatok. A felhasznált relációknak (rendszerjellemzők és költségvonzataik között) alkalmazhatónak kell lenniük az adatokra. Ahogy a rendelkezésre álló adatok változnak, újra kell értékelni a relációk érvényességét. A nem megfelelően alkalmazott költségbecslési relációk komoly becslési hibára vezethetnek.

Az *analóg becslés* esetén egy már üzemben lévő, megvalósításban és működésben hasonló rendszer költségeinek igazításával becsüljük a költségeket. Sokszor előfordul, hogy a tervezett rendszernek csak bizonyos részrendszereihez találunk referencia (rész)rendszereket. Ennek a módszernek az alkalmazásával elkerülhetjük a parametrikus becslés negatív aspektusait, de alkalmazása erősen függ az analóg rendszereket ismerő szakértők véleményétől.

A *mérnöki megközelítés* egy "bottom-up" jellegű becslést eredményez. A három technika közül ez a leginkább időigényes. A becslés lényege, hogy a teljes rendszer költségeit kisebb működési egységek költségeiből határozzuk meg. A cél az, hogy a valódi működés minél több részletét minél pontosabban vegyük figyelembe. A rendszert általában alsóbb szintű komponensekre kell bontani, és ezek költségeit külön-külön meghatározni. Ez a módszer nagy munkaigényű, valamint sok és részletes működési adatra van szükség. A módszer használhatóságát az érintett rendszer életciklusbeli helye, és az adatok megbízhatósága befolyásolja.

A gyakorlatban az úgynevezett *könyvelési modellt* használják költségbecslésre. Ez a modell az előző három technika kombinációjából áll. Különféle képletekkel meghatározzák a rendszer elemeinek költségeit (egyszerűrelációkkal illetve közvetlen költség adatok bevitelével), majd aggregálják őket. Bizonyos elemeknél például egységárat, máshol különféle modelleket használnak.

Adatforrások, paraméterek

A költségbecslésekben háromféle adatot lehet felhasználni: működési/műveleti adatok (paraméterek), technikai adatok (attribútum adatok) és költség adatok. Ezek az adatok

rendelkezésre állhatnak nyers vagy feldolgozott formában. A nyers adatok használhatók működő rendszerek historikus költségelemzésére. A nyers adatok általában az egység vagy műhely szintű karbantartási napló adatbázisokban állnak rendelkezésre. A feldolgozott adatok általában már költség jellegű, aggregált adatokat jelentenek. Ilyen adatokat a működtetéssel kapcsolatos tranzakció feldolgozó rendszerek adatbázisai felett kialakított sok szempontú elemző vagy adatbányászati alkalmazásokból nyerhetünk.

A KÖLTSÉGFA MODELL

Költség meghatározás költségfával

A mérnöki elemzési/becslési megközelítés esetén szóba jöhet, jól használható költségbecslési modell (költségelem struktúra) a költségfa. Ez a költségbecslési módszer tulajdonképpen a könyvelési modellt valósítja meg.

A költségfa a költségelemzés alapját képező objektumok, tevékenységek (költségelemek) attribútumai alapján teszi lehetővé a költségek összesítését. A költségstruktúra hierarchikus lebontása után, az elemi költségtényezők ismert vagy megfelelően becsült adatai alapján a hierarchia magasabb szintjeire, költség kategóriákra meghatározhatók az aggregált költségek. A módszer kellően rugalmas többféle becslési módszer egyidejű alkalmazására, így alkalmas különböző életciklus-fázisban lévő részrendszerek együttes figyelembe vételére. A rendszer költségeinek aggregálása kétféle csomóponton keresztül valósul meg: egyrészt a költségelemek attribútumaiból valamilyen formula és paraméterek segítségével (formula-aggregáció) meghatározásra kerül az adott elemi költség, majd ezen költségekből összesíthetők a magasabb szintek költségei (tétel-aggregáció). A különböző típusú elemi költségek kezelésére létrehozhatunk az adott költség kiszámítására használható paraméterezhető költség modulokat. Az ilyen módon kidolgozott modulok (sztereotípiák) újrafelhasználhatók az azonos típusú, de a hierarchia különböző pontjain megjelenő és esetleg eltérő paraméterekkel rendelkező költségek figyelembevételénél. A költségek meghatározásánál figyelembe veendő objektumok tetszőleges költség meghatározási alapot jelentő egységek lehetnek (pl. járművek, raktárak, műhelyek). Ezen objektumok költség szempontból meghatározó attribútumainak tárolására egy úgynevezett objektum-adatbázis szolgál. Az adatbázis tartalmazza az objektumok azonosításához és a költségek becsléséhez szükséges adatokat.

A költség modulokban felhasznált, működési adatokat jelentő paramétereket szintén el lehet helyezni egy paramétertárba, mely biztosítaná ezen adatok egyszeres tárolását, egységes, centralizált és kényelmes menedzselését.

Paraméterek meghatározása

Ha az üzemeltetés és fenntartás során keletkező működési adatok lekérdezhetők, aggregálhatók a költség meghatározás szempontjából fontos attribútumok szerint, akkor a paraméterek meghatározásának legjobb módszere az adatok feldolgozása egy tudás kinyerő vagy adatbányászati alkalmazással. Ezek az eszközök képesek létrehozni a több szempontú elemzés megkívánta adatkockákat, és képesek hatékonyan kezelni őket. Komplex költségbecslési feladatoknál ez különösen fontos, ugyanis általában nagy adathalmazokból kell bonyolult aggregálással paramétereket meghatározni, vagy nagyszámú elemzést kell elvégezni sok paraméter meghatározására, változó feltételek mellett.

Költségkeret felosztás

Ha a költségfa módszert használjuk a tényleges költségek meghatározására, akkor ez a rendszer kielégítő modellként szolgálhat egy meghatározott költségkeret felosztására is.

Ebben az esetben a felosztásban érdekelt egységeket szerepeltetni kell az aggregáló hierarchiában, így a költségkeretet a tényleges igények arányában lehet felosztani. Ilyenkor természetes követelmény, hogy a hierarchián belül konzisztens módon határozzuk meg a költségigényeket a részhierarchiák számára. A felosztáskor gyakran előfordul, hogy figyelembe kell venni egyéb korrekciós tényezőket (például költségigény prioritások, biztonsági tartalékok stb.) is.

A DÖNTÉSTÁMOGATÓ SZOFTVER

Speciális követelmények

Az előzőekben ismertetett költségbecslési módszer alkalmazása valódi környezetben csak annak szoftveres implementációja esetén képzelhető el komplex alkalmazási környezet esetén. Jelen cikkben ismertetett implementáció nem egy általános célú megvalósítása a módszernek, hanem egy speciális területen való alkalmazása. Az implementáció továbbra is nyílt és rugalmas, csak a teljesítmény optimalizálás szempontjából és a megvalósítás egyszerűsítése miatt tartalmaz speciálisan kötött elemeket. Az általunk megvalósított alkalmazás felhasználója egy nagy közlekedési vállalat, amelynek járműpark karbantartási, valamint felújítási költségkeretének optimális felosztását kell támogatni. A három költségbecslési eljárás ötvözetét kellett implementálni. Ahogy a későbbi leírásból látható, ezek a követelmények leginkább az objektum-adatbázis szerkezetére, valamint a parametrikus becslések esetén kialakítható formula-aggregációs megoldásokra voltak hatással.

Az objektum-adatbázis

A konkrét feladat jellegéből fakadóan az objektum adatbázis elemei a közlekedési vállalat járműveivel kapcsolatos adatokat tartalmazzák. Ezek az adatok az előzőpénzügyi időszakra vonatkozó, a járművek karbantartási ráfordításait valamint a becslési formulákban szerepet játszó műszaki paramétereket tartalmazzák, járművenként egy bejegyzést.

Az aggregált adatok előállítását a cég SAP rendszerében történik, és off-line módon importálhatók a döntéstámogató rendszer saját adatbázisába. A bejegyzések száma 10000-es nagyságrendű. Az adatok importálása után a szoftver lehetővé teszi a klasszikus adatbányászati adatkarbantartási műveletek elvégzését, adattisztítást: adatok törlése, javítása nagy tömegben.

Az objektum-adatbázis tulajdonképpen egy adatkockát képez a következődimenziókkal: járműkategória, beszerzési év, km futás, telephely, felújítás éve, költségadatok. Ezen dimenziók mellett hat szabadon definiálható numerikus érték rendelhető az egyes bejegyzésekhez. Ezek az értékek a formula-aggregációkban használhatók (főleg paraméteres becslés) és az adatkarbantartás során tölthetők fel értékekkel. Ezek az adatok megjelenhetnek az adatkocka dimenzióiként is. Értékük lehet megadott konstans, vagy tetszőleges nem string dimenzióra végrehajtott leképezés. A leképezések a dimenziók értékészletét képezik le szakaszonként konstans vagy lineáris, a felhasználó által definiálható függvények segítségével. Ez természetesen korlátozást jelent, de a mérnöki gyakorlathoz elegendő (domén modelljeink csak ilyen jellegű függvényeket tartalmaztak) és effektíven megvalósíthatók. A költségfa és csomópontjai A költségfa kétféle csomópontból áll: elemi költségek meghatározására szolgáló csomópontból és a költségek összesítésére szolgáló aggregáló csomópontból (*1. ábra*).

Az elemi költségszámok a jelen megvalósításban tetszőleges számú tétel összeszorozását tudják elvégezni. A szorzat elemei ad hoc konstansok, névvel ellátott és a

paramétertárban tárolt konstansok vagy az objektum-adatbázison végrehajtott, numerikus eredményre vezető lekérdezések lehetnek (2. ábra).

1. ábra

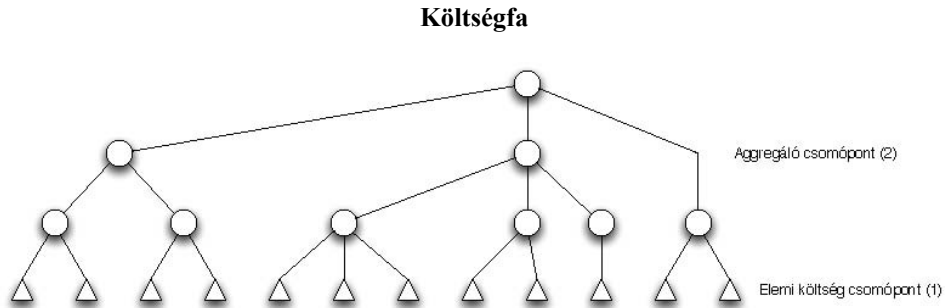


Figure 1: Cost tree

Elementary cost node(1), Aggregation node(2)

2. ábra

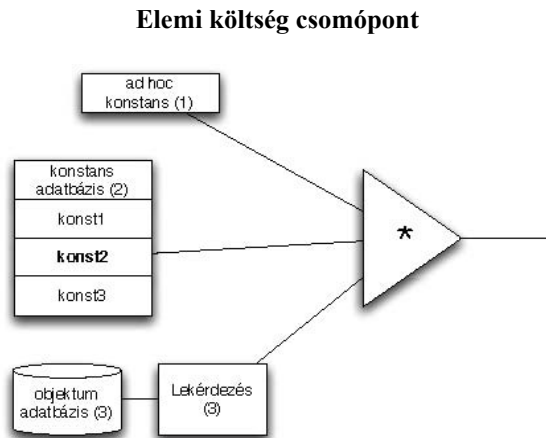


Figure 2: Elementary cost node

Ad hoc constant(1), Constant database(2), Object database(3), Query(3)

Az aggregáló csomópontok kettős szerepet töltenek be: egyrészt összesítik a hozzájuk tartozó részfák költségeit, valamint a csatlakozó elemi költségcsomópontok lekérdezéseinek feltételeit egészítik ki. Ez azért kényelmes, mert maguk a csomópontok adott dimenzió menti hierarchikus felbontást jelentenek általában. Így az egyes költségcsomópontok változatlan formában felhasználhatók a hierarchia különbözőpontjain (3. ábra).

3. ábra

Aggregáló csomópontok szűrési feltétellel

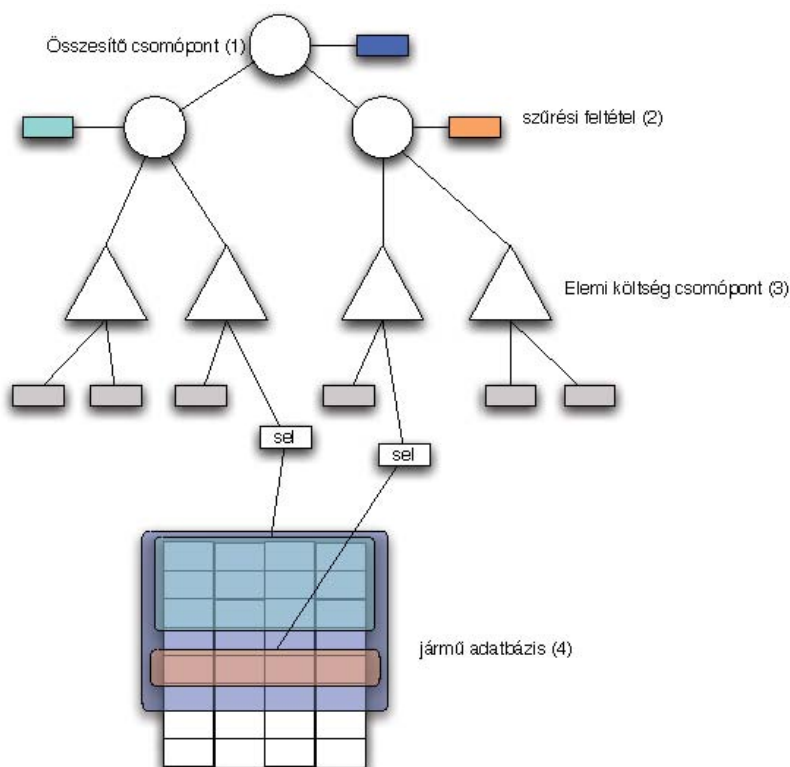


Figure 3: Aggregation node with filtering condition

Aggregation node(1), Filtering condition(2), Elementary cost node (3), Vehicle database(4)

Elemzési lehetőségek

Mivel a szoftverrel szemben elvárás volt, hogy a problématulajdonos a szakértők által javasolt becslési összefüggések mellett az adatokat elemezni is tudja, a költség csomópontokban használható lekérdezés-definiáló felületről tetszőleges, az adatkockára vonatkozó lekérdezést lehet megadni és kiértékelni, valamint vizuálisan egy paraméter (dimenzió) szerinti exploratív érzékenységvizsgálatot is el lehet végezni. Normatív költségbecslési modell Az objektum-adatbázis szerkezet, valamint a szoftverben megvalósított funkcionalitás kialakítása a domén szakértők által létrehozott költségbecslési modellre épült, de a rugalmas használat miatt bizonyos kereteken belül a lehetőségek szerint független, abban az értelemben, hogy a költségfa struktúrája szabadon kialakítható, újabb költségtényezők és paramétereik beépíthetők. A legszorosabb megkötések az objektum-adatbázisra érvényesek, mely ebben az esetben egy jármű adatbázis.

A szakértők által megadott modellek normatív jellegűek, a nemzetközi tapasztalat alapján szükségesnek ítélt működtetési költségekből indulnak ki. A költségek becslésére felhasznált jármű-és egyéb műszaki paraméterek a következők: járműtípusa, életkora, km-futása, üzemanyag felhasználása, váratlan meghibásodások száma, a telephelyre jellemző járműkiadási ráta és túlfuttatási arány. Ezek az adatok jelennek meg természetesen a jármű adatbázisban is.

A becslési egyenletek a különböző tényezők által implikált költségeket adják meg általában a beszerzési érték arányában. Két tényező, a járművek életkora, valamint a km-futás játsza a legfontosabb szerepet. A járműéletkora továbbá jelentősen meghatározza a többi tényező által generált költségeknek a súlyát. Mindezen költségeket a modellek különböző járműcsoportokra vonatkoztatva adják meg, szakaszonként lineáris függvények formájában.

A modell leképezése a szoftverben

A modell operacionalizálása a szoftverben a szoftverarchitektúra különböző komponenseiben van elosztva. A modell függvényei által, a járművek műszaki paramétereiből származó értékek kiszámítása az adatbázis feltöltésekor végezhető. Külön tevékenység a járművek típusok szerinti csoportokba sorolása.

A járműadatbázisból származó aggregált értékek (min, max, átlag, darabszám) lekérdezésekkel állíthatók elő. Ezek az értékek közvetlenül felhasználhatók az elemi költségcsomópontok számításaiban (szorzások). A költségek aggregálását járműcsoportokra, telephelyekre, üzemigazgatóságokra a költségfa összesítő csomópontjai végzik el. A szoftver képes több elemzési adatbázist kezelni, így különbözőbecslési modellek eredménye összehasonlítható.

A felhasználói felület

A szoftver felhasználói felületének kialakításakor az volt a cél, hogy a megrendelő nem informatikus, műszaki háttérrel rendelkező, a költségelemzési, felosztási eljárásban részvevő munkatársai könnyen, intuitíven tudják használni. Ezt a célt egy viszonylag egyszerű, grafikus Windows MDI felülettel sikerült elérni.

A szoftver menürendszere keskeny és sekély, a műveletek elvégzése inkább a kapcsolódó ablak nyomógombjaival lehetséges, ez közelebb áll a direkt manipulációs interfészek szelleméhez. Az objektum adatbázis megjelenítése egy a táblázatkezelőkéhez hasonló táblázatban történik. Erről a felületről érhetők el az adattisztítási és leképezési funkciók is. A leképezési függvények megadása a töréspontok koordinátáinak megadásával lehetséges. Az ablak képes a megadott függvény grafikus megjelenítésére a koordináták ellenőrzéséhez.

Az adatok elemzésére, illetve a költségcsomópontokban felhasználható lekérdezések kialakítására külön ablak szolgál (4. ábra). A lekérdezések feltétel megadása egy táblázattal történik, megfelelő támogatást nyújtó dialógus ablakok segítségével. Erről a felületről végezhető el a lekérdezés egy feltételének automatikus változtatásával a lekérdezés vizuális paraméterérzékenység vizsgálata is.

A költségfa szerkezetének megadására és a fa kiértékelésére külön dokumentumablak szolgál (5. ábra). A fa csomópontjait egy TreeView megjelenítő segítségével lehet kialakítani. Az egyes csomópontok szerkesztése dialógus ablakokban történik. A fa kiértékelésekor egy táblázatot kapunk, mely a költség értékeket, az összköltséghez viszonyított abszolút, valamint szintenkénti relatív súlyokat tartalmazza. Ezek az eredmények megjeleníthetők egy TreeMapként is, ahol a megjelenített szintek száma szabályozható.

4. ábra

Lekérdezések megadása

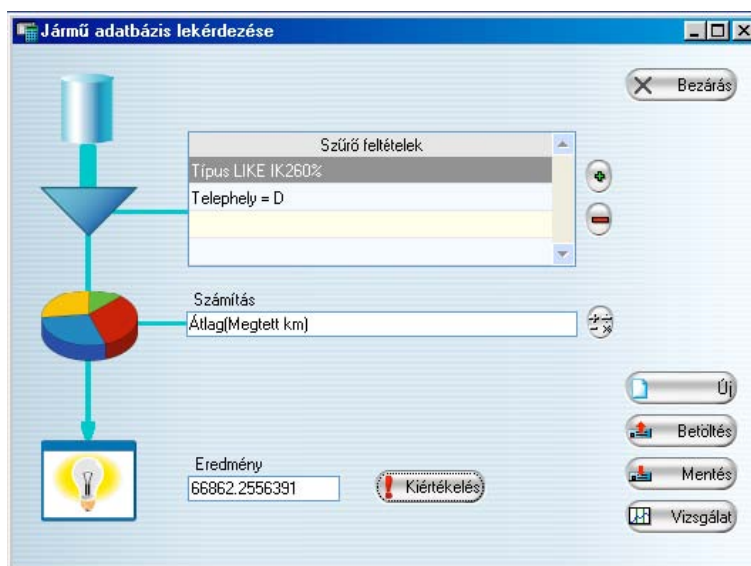


Figure 4: Query specification

5. ábra

Költségfa ablak

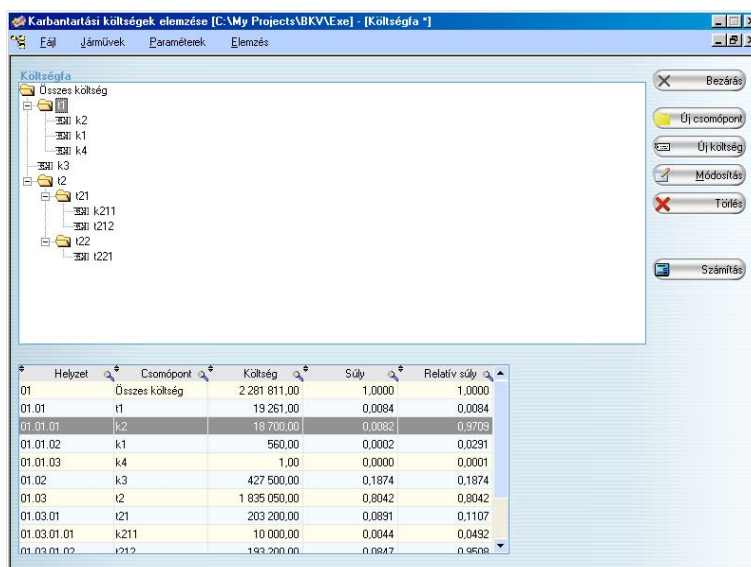


Figure 5: Cost tree window

IRODALOM

- Su S.Y.W., J. Dujmovic, D.S. Batory, S.B. Navathe, R. Elnicki (1987). A Cost-Benefit Decision Model: Analysis, Comparison, and Selection of Data Management Systems, *ACM Transactions on Database Systems*, 12. 3. Sept. 1987, 472-520.
- Office of the Secretary of Defense Cost Analysis Improvement Group (1992). *Operating and Support Cost-estimating Guide*,
- Asahi, T., Turo D., Shneiderman B. (1995). Visual decision-making: Using treemaps for the Analytic Hierarchy Process, In: *Proceedings of CHI'95: Conference on Human Factors in Computing Systems: Mosaic of Creativity*, 1995, Denver CO.
- Bozóky, L., Nagy V. (2003). Karbantartási és felújítási költségkeretek elosztási módszereinek meghatározására javaslatok kidolgozása, *Kutatási jelentés*, 2003, Győr
- Nagy, V., Bozóky, L. (2001). TENDER, "Komplex rendszerek összemérési vizsgálata" probléma megoldására szoftverrendszer fejlesztése, *Kutatási jelentés*, 2001, Győr

Levelezési cím (*Corresponding author*):

Heckenast Tamás

Széchenyi István Egyetem, Informatika Tanszék

9026 Győr, Egyetem tér 1.

Széchenyi István University, Department of Informatics

H-9026, Győr, Egyetem square 1.

Tel.: 36-96-613-617

e-mail: heckenas@sze.hu



Vállalkozók adózási stratégiáinak szimulációja generikus kétrétegű háló modellel

Varga M., Bitáné Bíró B., Bánkuti Gy., Csukás B.,

Kaposvári Egyetem, Gazdaságtudományi Kar, 7400 Kaposvár, Guba S. u. 40.

ÖSSZEFOGLALÁS

Az egyéni vállalkozói réteg egyre fontosabb szerepet tölt be Magyarország gazdasági életében, már csak számát tekintve is. Fontos odafigyelni rájuk, elemezni, tervezni adózásukat. Célunk éppen ezért az egyéni vállalkozók jövedelmének optimalizálása, a választható adónemek közti választás megkönnyítése. A cél elérése érdekében egy új módszer, a kétrétegű háló modell alkalmazását választottuk. Elsőként az ide vonatkozó, a törvényben meghatározott adók és járulékok kétrétegű háló modellbe történő adaptálását végeztük el, az aktív és passzív elemeket határoztuk meg. Az így felépített kétrétegű háló egyszerűen leképezhető dinamikus adatbázis elemeire. Ezzel az adatbázissal kommunikáló keretprogram segítségével dinamikus szimulátor készíthető, mely havi időlépésenként számol. Néhány input (bevétel, költség, kivét) megadása után gyorsan és hatékonyan számolja a vállalkozót érintő outputokat (fizetendők, nettó eredmény). A vállalkozói adónemek (VJA, átalányadó, EVA) közti választás megkönnyítésére egy lépésben szimuláljuk a vállalkozó adóévét, és az így kapott együttes adó- és járulékkösszegeket hasonlítjuk össze. Legnagyobb eredménynek az elkészített generikus kétrétegű háló modell alapú dinamikus szimulátor tekinthető, melybe az aktuális törvényi szabályozókat beépítettük. Törvényváltozás esetén változtatni csak az Excel adatbázisban szükséges. Következtetésképpen a szimuláció bármikor egyszerűen elvégezhető a vállalkozó adottságaihoz igazítottan, egyéni esetekre vonatkozóan is. (Kulcsszavak: jövedelem optimalizálás, generikus kétrétegű háló modell alapú dinamikus szimulátor)

ABSTRACT

Tax simulation of private entrepreneurs based on generic bi-layered net model

M. Varga, B. Bitáné Bíró, Gy. Bánkuti, B. Csukás

University of Kaposvár, Faculty of Economic Science, Kaposvár, H-7400 Guba Sándor u. 40.

In the past 10-15 years in Hungary, the sphere of the private entrepreneurs is more and more important in our economic conditions due their growing number. Therefore it is a substantial question to analyse and plan their taxation. Our goal is to optimize the entrepreneur's receipts and to give assistance by choosing the right tax category. To achieve the purpose of our study we apply a new method, simulation based on generic bi-layered net model. First of all we adapted the concerning orders of the benefits and taxes in our net model and determined the passive and active elements. The built model is communicating with a framesoftware and calculates in monthly timesteps. The simulator calculates very effectively the needed outputs (taxes, benefits and net income) from a few inputs (returns, costs, income) and helps by choosing the right tax category (

Entrepreneur income tax, Flat-rate tax or Simplified entrepreneur income tax) with the simulation of the full taxyear in one timestep. The result of the study is the generic tax-calculator based on the generic bi-layered net model, which contains the actual legal orders in a simple Excel worksheet, therefore it is easy to swap the tax- and benefit rates – by the user - in case of changed rules. Based on the results we came to the conclusion that the generic tax-calculator can be applied for every individual case by entrepreneurs, bookkeepers or financial advisors.

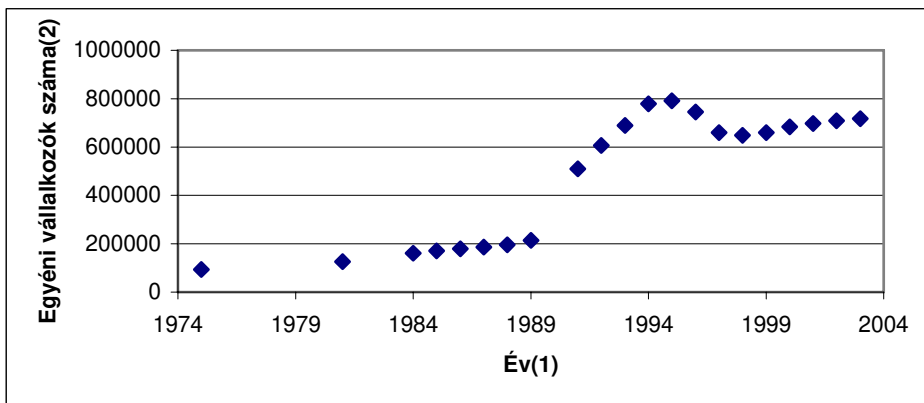
(Keywords: optimize receipts, generic tax-calculator based on the generic bi-layered net model)

BEVEZETÉS

Az adózás egy olyan fontos tényezője a mai gazdasági életnek, mely nem csak az adó megfizetéséből, hanem számos olyan aspektusból tevődik össze, amelyeket nem szabad, illetve nem lehet figyelmen kívül hagyni. Különös tekintettel kell lennünk az egyéni vállalkozókra, mely réteg az utóbbi egy-másfél évtizedben növekvő szereppel rendelkezik a gazdasági életben. Ezt igazolják az 1. ábrán látható statisztikai adatok.

1.ábra

Regisztrált egyéni vállalkozók száma



Forrás (Source): Statisztikai zsebkönyvek (1975-2003), *Paperback of Statistics (1975-2003)*

Figure1: Number of registered private entrepreneurs

Year(1), Number of private entrepreneurs(2)

Az egyéni vállalkozások jövedelmezőségi és adózási helyzetét bonyolítja, a rendszerváltás óta eltelt időben hozott számos törvényváltozás, illetve újabb adónemek bevezetése. Jelenleg egy vállalkozó háromféle adónem közül választhat:

- Vállalkozói jövedelemadó – továbbiakban VJA
- Átalányadó
- Egyszerűsített vállalkozói adó – továbbiakban EVA

Elsősorban az adott adónem törvényi feltételeinek kell megfelelni ahhoz, hogy annak alanyává válhasson, ezen kívül viszont egy sürgető gazdasági kérdés a vállalkozó szempontjából, hogy a saját esetében milyen mértékű fizetési kötelezettségek terhelik.

A háromféle adónemre vonatkozó konkrét törvényeken kívül a táblázatos összesítéshez következő törvények ismerete is szükséges: 1990. évi V. tv, 1992. évi LXXIV. tv, 1998. évi LXVI. tv, 2003. évi XCII. Tv.

A VJA szerint adózó, fizetendő terheit az 1. táblázat tartalmazza az 1995. évi személyi jövedelemadóról szóló CXVII. törvény alapján.

Az átalányadó szerint adózó vállalkozó kötelezettségeit a 2. táblázat tartalmazza az 1995. évi személyi jövedelemadóról szóló CXVII. törvény alapján.

EVA választása esetén a 3. táblázatban látható mértékűek a kötelezettségek, ezen adónem a 2002. évi XLIII. törvény hatálybalépésétől alkalmazható.

Vizsgálatainkban célul tűztük ki, hogy – egy új módszer alkalmazásával – kísérletet tegyünk az egyéni vállalkozók jövedelmének optimalizálására és a választható adónemek közti választás megkönnyítésére.

ANYAG ÉS MÓDSZER

A jövedelemoptimalizálásra és az adónemek közti választás megkönnyítésére szolgáló kalkulációkat generikus kétrétegű háló modell alapú dinamikus és generikus szimulációval végeztük el. A kétrétegű háló egy olyan új módszer amely bonyolult folyamatok számítógépes leképezését teszi lehetővé. A törvényi előírásoknak megfelelő adó- és járulékszabályokat egy Excel táblázat munkalapjain passzív és aktív elemek formájában tettük összekapcsolhatóvá egy Visual Prolog nyelven írt keretprogrammal.

A generikus kétrétegű háló modell kétféle gráfélből épül fel. A gráfpontok a tulajdonképpeni passzív és aktív elemek, míg a gráfélek a leolvasási és módosítási csatornák. (Csukás és Bánkúti, 2003)

1. táblázat

Vállalkozói jövedelemadó szerinti közterhek

Adók és járulékok (1)	Alap (2)	%
VJA	Bruttó bevétel- összes költség	16%
Szja	Vállalkozói kivét, általában az éves minimálbér 2004-ben 2005-ben	18, 26 és 38% 18 és 38%
Osztalékadó	Eredményből a vállalkozói kivét 30%-ának megfelelő összeg 2004-ben 2005-ben	20% 25%
	Vállalkozói osztalékalap fennmaradó része	35%
Társadalombiztosítási nyugdíjbiztosítási járulék	Vállalkozói kivét v. minimálbér	18%
Társadalombiztosítási egészségbiztosítási járulék	Vállalkozói kivét v. minimálbér	11%
Nyugdíjjárulék	Vállalkozói kivét v. minimálbér	8,50%
Egészségbiztosítási járulék	Vállalkozói kivét v. minimálbér	4%
tételeLes egészségügyi hozzájárulás	Fix összeg: 41.400 Ft/év	3450 Ft/hó
Vállalkozói járulék	vállalkozói kivét v. minimálbér 2005-től	4%

Table 1: Rates and taxes in case of entrepreneur income tax

Taxes and benefits(1), Taxable income(2)

2. táblázat

Átalányadó választása esetén fizetendő adók és járulékok

Adók és járulékok (1)	Alap (2)	%
Átalányadó	Az egyéni vállalkozó 40 százalék költséghányad levonásával állapítja meg.	2004-ben 18, 26 és 38%
	Az adóév egészében kizárólag az alább* -törvény szerinti 53. & (3) és (4) bekezdésben- felsorolt tevékenységet folytató egyéni vállalkozó 80 százalék költséghányad levonásával állapítja meg.	
	Az adóév egészében kizárólag az üzletek működéséről és a belkereskedelmi tevékenység folytatásának feltételeiről szóló kormányrendelet alapján kiskereskedelmi tevékenységet folytató egyéni vállalkozó 87 százalék költséghányad levonásával állapítja meg.	
	A c) pontban említett egyéni vállalkozó – feltéve, hogy a tevékenysége külön-külön vagy együttesen az adóév egészében kizárólag a (4) bekezdésben felsorolt üzlet(ek) működtetésére terjed ki – 93 százalék költséghányad levonásával állapítja meg.	2005-ben 18 és 38% sávosan
	A mezőgazdasági kistermelő 85 százalék, illetőleg a bevételének azon részéből, amely állattenyésztésből vagy állati termék előállításából származik, 94 százalék költséghányad levonásával állapítja meg.	
Társadalombiztosítási járulék	Átalányadó alap illetve minimálbér	18%+11%=29%
Nyugdíjjárulék		8,5%
Egészségbiztosítási járulék		4%
Tételes eü. hozzájárulás	Fix összeg: 41400 Ft/ év	3450 Ft/ hó
Vállalkozói járulék 2005-től	Átalányadó alap illetve minimálbér	4%

Table 2: Rates and taxes in case of flat-rate tax

Taxes and benefits(1), Taxable income(2), Percentage(3)

3. táblázat

EVA hatálya alá tartozók fizetendő adója, járulékai

Adók és járulékok (1)	Alap (2)	%
Adó mértéke	Pozitív adóalap	15%
Társadalombiztosítási járulék	Minimálbér vagy választás szerinti minimálbérnél nagyobb összeg	29%
Nyugdíjjárulék	Minimálbér vagy választás szerinti minimálbérnél nagyobb összeg	8,50%
Egészségbiztosítási járulék	Minimálbér vagy választás szerinti minimálbérnél nagyobb összeg	4%
Tételes egészségügyi hozzájárulás	Havonta fizetendő, mindig azonos összegben	3 450 Ft
Vállalkozói járulék 2005-től	Minimálbér vagy választás szerinti minimálbérnél nagyobb összeg	4%

Table 3: Rates and taxes in case of simplified entrepreneur income tax

Taxes and benefits(1), Taxable income(2)

A passzív elemek mértékek, fizikai mennyiségek vagy jelek, míg az aktívak mérték változások vagy szabályok, melyek lényegében a „két réteget” alkotják. Köztük a kommunikációs csatornák teremtenek kapcsolatot. A passzívtól az aktív elem felé tartó a

leolvasási, az aktív - passzív irányú pedig a módosítási csatorna. A vállalkozói adónemekre adaptált struktúrában a passzív elemek az adók, járulékok, adóalapok mértékei, az aktív elemek pedig az aktuális törvény szerint felépített szabályok, melyek az adót számolják.

A kétrétegű háló modellel leírt folyamat szimulációja lényegében úgy játszódik le, hogy az aktív elem leolvassa a szükséges passzív elemek értékeit. Minden aktív elem mögött egy kis „program” működik, mely kiszámítja a megváltozást, mely a csatornákon visszajut a passzív elemekig szabályozott időlépésenként. Ennek bemutatására választottuk a 2. ábrán szereplő nettó bevételt és az ÁFA-t számító hálórészletet.

2. ábra

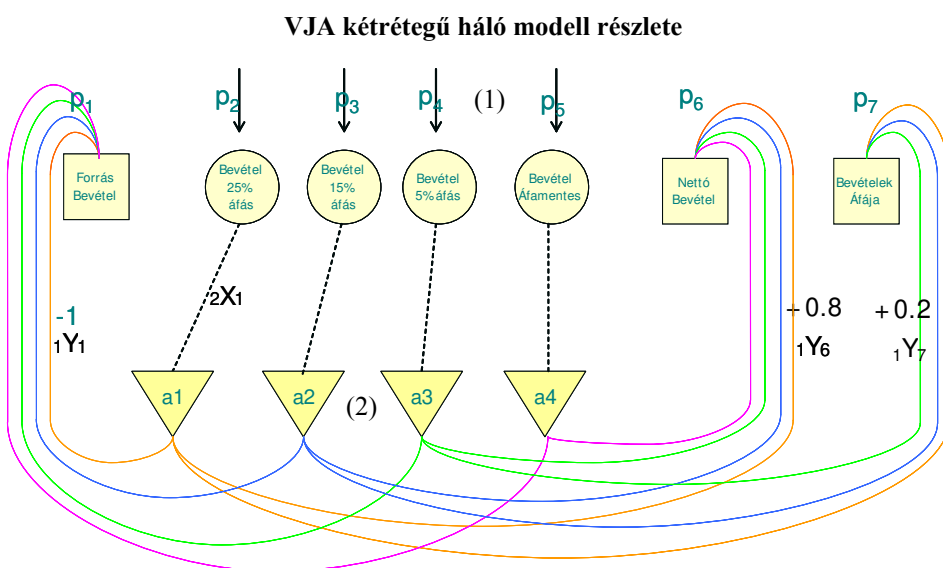


Figure 2: Fragment of generic bi-layered net model for entrepreneur income tax

$p_1, p_2, p_3, p_4, p_5, p_6, p_7$ – passive elements(1), a_1, a_2, a_3, a_4 – active elements(2)

A passzív-aktív elemek funkcióit, a köztük lévő kapcsolatrendszer matematikailag legáltalánosabban leképezéssel írhatjuk le, melynek leolvasásai a belépő kommunikációs csatornán érkező adatok, kép elemei pedig a kimenő kommunikációs csatornák tartalmát meghatározó jellemzők.

A 2. ábra a_1 -es aktív elemének a leképezése a következő módon néz ki:

$$a_1 \rightarrow \text{Bevetel25} = \begin{bmatrix} \text{Be25} = 2X_1 \\ 1Y_1 = -\text{Be25}, 1Y_6 = 0,8\text{Be25}, 1Y_7 = 0,2\text{Be25} \end{bmatrix}$$

Röviden összefoglalva, az a_1 -es elem működése így írható le:

Az ' a_1 ' aktív elem feladata, hogy a 25% Áfát tartalmazó bruttó bevételt leolvassa annak tartalmát a Forrás_Bevétel nevezetű p_1 -es passzív elemből elvegye, majd az adott összeg 80%-át a p_6 -os passzív elemhez (Nettó_bevétel), 20%-át pedig a p_7 -es „Bevételek_Áfája” nevezetű passzív elemhez adja. Ennek általános jelentése, hogy az áfás bevételt két részre bontja, az áfára és az áfát már nem tartalmazó nettó bevételre.

A háló összes aktív és passzív elemére hasonló séma szerint készült leképezés.

A modell nagy előnye abban mutatkozik, hogy az alapvető problémát, vagy folyamatot, kis részekre bontja, a közöttük fennálló kapcsolatokat külön jelöli, ezért bármikor, bármelyik eleme könnyen megváltoztatható, kicserélhető (Bánkuti, 2004; Csukás, 2000).

EREDMÉNYEK

Munkánk eredményeként kétféle szimuláció készült el: egy dinamikus szimulátor VJA adójának és lehetséges jövedelmezőségének havi időlépésben történő számítására, illetve egy generikus szimulátor, mely egy adóévre vonatkozóan, egy lépésben az adónemek összehasonlíthatóságát teszi lehetővé.

A dinamikus szimulátor 2004-es évre vonatkozó eredményeit a 4. táblázatban alacsony (4 millió Ft) és közepes (10 millió Ft), az 5. táblázatban pedig magasabb (15 millió Ft) árbevétel kategória esetére (40%-os költséghányad és 25%-os áfa mellett) foglaltuk össze. A feltüntetett „Fizetendő” sor magában foglalja az év folyamán felmerülő összes adót és járulékot.

Az általános tendencia felvázolása érdekében 4-féle kivét esetén végeztettük el a számolást a szimulátorral a különböző árbevétel kategóriákra. Az egyes árbevétel kategóriák esetén a vizsgált kivétek közül a legjobb a minimálbérrel megegyező összegű.

A táblázatokból az is megállapítható, hogy a bevétel kategóriák növelésével a fizetendő kötelezettségek százalékos arányai közti eltérések nem nőnek tovább, hanem inkább csökkennek.

A számítások ugyanígy elvégezhetők tetszőleges árbevétel, költséghányad és vállalkozói kivét esetén. Az elkészített adó- és járulékszámító programmal a vállalkozó egész éve reprezentálható, tervezhető, mivel az Excel táblában a vállalkozó által megadható paraméterek beírása után a program az általunk előre meghatározott módon számolja a kért outputokat.

A folyamatos kimutatást a havi időlépés biztosítja (3. ábra).

4. táblázat

A dinamikus szimulátor eredményei (I.)

Terhek (3)	4 millió Ft br. bevétel (2)				10 millió Ft br. bevétel			
	Kivét éves összege, e Ft-ban (1)				Kivét éves összege, e Ft-ban			
	0	0	636	1500	2000	636	1000	1200
ÁFA	480	1200	1200	1200	1200	480	480	480
SzJA	0	0	114.48	326	516	114.48	196	248
TB (29%)	184.44	184.44	184.44	435	580	184.44	290	348
Eü és Ny (12,5%)	79.5	79.5	79.5	187.5	250	79.5	125	150
Teho	41.4	41.4	41.4	41.4	41.4	41.4	41.4	41.4
VJA alap	1694.16	4574.16	3938.16	2823.6	2178.6	1058.16	588.6	330.6
VJA	271.0656	731.8656	630.1056	451.776	348.576	169.3056	94.176	52.896
Osztalékba	1423.094	3842.294	3308.054	2371.824	1830.024	888.8544	494.424	277.704
20% oszt. alap	0	0	190.8	450	600	190.8	300	277.704
20% oszt. adó	0	0	38.16	90	120	38.16	60	55.5408
35% oszt. alap	1423.094	3842.294	3117.254	1921.824	1230.024	698.0544	194.424	0
35% oszt. adó	498.083	1344.803	1091.039	672.6384	430.5084	244.319	68.0484	0
Fizetendő adók és járulékok összege	1074.489	2382.009	2179.125	2204.314	2286.4844	871.6046	874.6244	895.8368

Table 4: Results of the dinamic simulator (I.)

Amount of income in thousand Ft(1), Gross income(2), Rates and taxes(3)

5. táblázat

A dinamikus szimulátor eredményei (II.)

Terhek (3)	15 millió Ft br. bevétel (2)			
	Kivét éves összege, e Ft-ban (1)			
	0	636	1500	3000
ÁFA	1800	1800	1800	1800
SzJA	0	114.48	326	896
TB (29%)	184.44	184.44	435	870
Eü és Ny (12,5%)	79.5	79.5	187.5	375
Teho	41.4	41.4	41.4	41.4
VJA alap	6974.16	6338.16	5223.6	3288.6
VJA	1115.866	1014.106	835.776	526.176
Osztalékba	5858.294	5324.054	4387.824	2762.424
20% oszt. alap	0	190.8	450	900
20% oszt. adó	0	38.16	90	180
35% oszt. alap	5858.294	5133.254	3937.824	1862.424
35% oszt. adó	2050.403	1796.639	1378.238	651.8484
Fizetendő adók és járulékok összege	3471.609	3268.725	3293.914	3540.4244

Table 5: Results of the dinamic simulator (II.)

See Table 4

3. ábra

Havi cash-flow alakulása 6 mFt bruttó árbevétel esetén

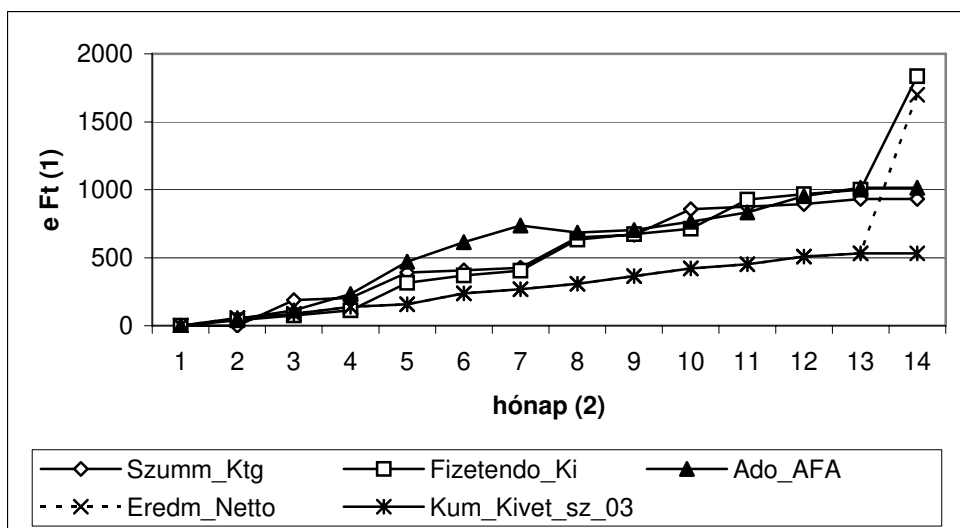


Figure 3: Monthly cash-flow

Thousand Ft(1), Month(2)

Az 3. ábra esetén a 13. hónap az elszámolási időszakot jelenti. Az időlépésenkénti számolás révén jól látható például, hogy a fizetendők negyedévenként jelentősen emelkednek, hiszen a negyedéves VJA előlegfizetés ekkor esedékes.

A nettó eredmény év végi hirtelen növekedését pedig az elszámolási időszak végén az osztalékból származó jövedelem adja. A mozgások dinamikája azonban minden esetben a bevételek, költségek és a kivét mértékétől függ.

A dinamikus szimulátorral lehetővé vált a VJA szerint adózó főfoglalkozású egyéni vállalkozó adóévének modellezése havi időlépésben, továbbá adott szituációban a lehető legjobb összegű vállalkozói kivét meghatározása.

A generikus adókalkulátor a választható háromféle adónemet modellezi egy adóévre vonatkozóan, egy időlépésben. A VJA, az átalányadó és az EVA szabályait egyszerre modellezve, ugyanazokat az inputokat felhasználva, a szoftver, outputjában azonnal összehasonlítható eredményeket produkál.

A generikus szimulátor eredményeinek alakulását a 4. ábra mutatja, a 2005. évi adószabályoknak megfelelően, 4, 15 és 22 millió forintos bruttó árbevétel esetén.

4. ábra

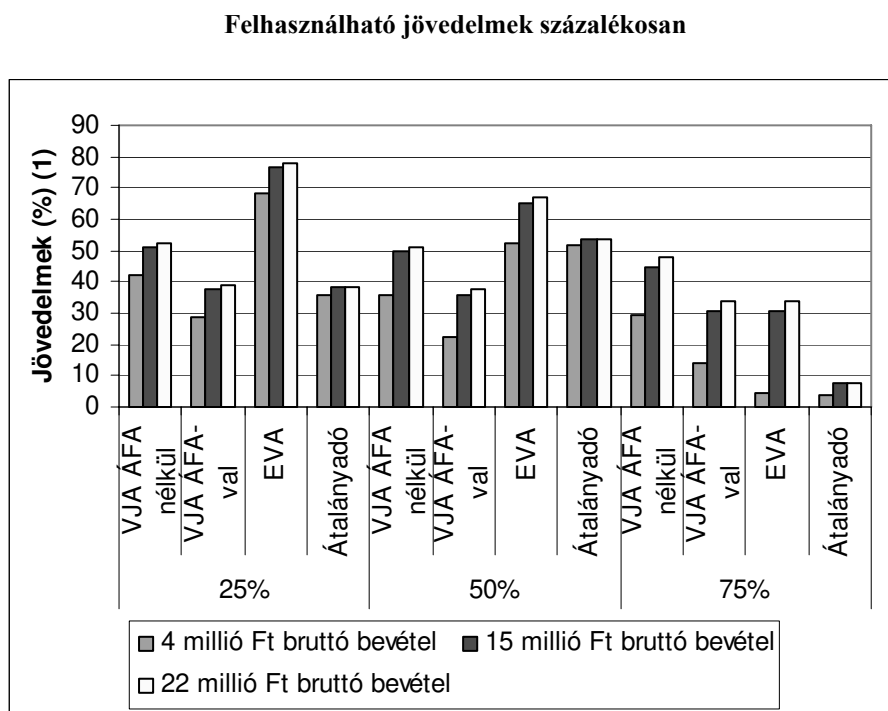


Figure 5: Useable income percentage

Income (%) (1), Costrate - Tax categories (2)

Az adókalkulátor eredményeiből, konkrét következtetést tudunk levonni adónem választási nehézségek esetén. Természetesen a munkánk során alkalmazott bevételi- és

költségtényezők csak útmutatást adnak, viszont a program használatával konkrét, a vállalkozó által megadott összegeknél is meg lehet mondani, hogy az adott estre mely adónemet célszerű választani.

A felhasznált adatok és elemzések tükrében elmondható, hogy az EVA választása 70% körüli költséghányadig 15 millió forint fölötti bevétel mellett még mindig a legjobb eredményt adja, a többi adónemhez viszonyítva.

Az olyan esetekben, ahol a tevékenység ÁFA-körös és a fizetendő ÁFA rendszeresen meghaladja a levonható áfa összegét, egyértelműen az EVA a legjobb választás, különösen ha kevés az elszámolható költség. Ide tartoznak a tipikusan alacsony ráfordítással operáló és az ÁFA-törvény hatálya alá tartozó tevékenységek, például: ügyvéd, üzletkötő, tervező, pénzügyi tanácsadó, tehát olyan szellemi tevékenységek, ahol nem keletkezik számottevő számlával igazolható költség.

A VJA-nál az alanyi/tárgyi adómentes tevékenységek előnye, a jövedelemarányban, az ÁFA-s vállalkozással szemben egyértelműen 75% költséghányad fölött mutatkozik, de alacsonyabb költség esetén is pár százalékkal jobb eredményt ad.

Az átalányadó választása akkor helyes döntés, ha a törvény által meghatározott lehető legmagasabb költséghányadot tudjuk érvényesíteni a bevételekkel szemben. A 40%-ot érvényesíteni tudó vállalkozó akkor jár jól, ha a tényleges költségeit 25% körül tudja tartani.

A generikus adókalkulátorral sikerült megvalósítani a három adónem összehasonlíthatóságát, amely alapján a vállalkozók egyéni eseteire vonatkozóan lehet javaslatot tenni.

KÖVETKEZTETÉSEK

Az eredményeink alapján azok a következtetések vonhatók le, hogy az elkészült dinamikus szimulátor és generikus adókalkulátor rendkívül hatékony segítséget nyújthat a potenciális felhasználók számára (adótanácsadók, könyvelőirodák, vállalkozók). Az elkészült szoftverek gyors eredményt szolgáltatnak és előrejelzésre alkalmasak, mert az adószabályok változása esetén könnyen kicserélhetők a módosult adó- és járulékkulcsok, végül de nem utolsósorban időt takarít meg, ugyanis alkalmazása esetén nem lesz szükség hosszadalmas kézi számításokra.

IRODALOM

1990. évi V. törvény - Az egyéni vállalkozásról
1992. évi LXXIV. törvény - Az általános forgalmi adóról
1995. évi CXVII. törvény - A személyi jövedelemadóról
1997. évi LXXX. törvény - A társadalombiztosítás ellátásaira és a magánnyugdíjra jogosultakról, valamint e szolgáltatások fedezetéről, egységes szerkezetben a végrehajtásáról szóló 195/1997.(XI. 5) kormányrendelettel
1998. évi LXVI. törvény - Az egészségügyi hozzájárulásról (módosítás: 2002/XLII tv. 315.&.)
2002. évi XLIII. törvény - Az egyszerűsített vállalkozói adóról
2003. évi XCII. törvény - Az adózás rendjéről
Bánkuti Gy. (2004). Kétrétegű háló modell alapú szimulációs módszer metabolikus folyamatok modellezésére. Alkalmazott Informatika Konferencia, Kaposvár, 2004 május 20.

- Csukás B., Bánkúti Gy. (2003). Generic Bi-layered Net Model Based Direct Computer Mapping of Process Models. Seminar. Carnegie Mellon University, Pittsburgh, PA, January 17.
- Csukás B. (2000). Megmaradás, információ, evolúció – a folyamatmérnöki tudomány alapjai. In: Környezettudomány. Szerk.: Somogyi Ferenc. Veszprémi Egyetemi Kiadó : Veszprém, 55-86.
- Magyar Statisztikai Zsebkönyv (1973-2003). Statisztikai Kiadó Vállalat : Budapest.

Levelezési cím (*Corresponding author*):

Varga Mónika

Kaposvári Egyetem, Gazdaságtudományi Kar

4700 Kaposvár, Guba S. u. 40.

University of Kaposvár, Faculty of Economic Science

H-7400 Kaposvár, Guba S. u. 40.

Tel.: 36-70-581-5673

e-mail: varga@matinf.gtk.u-kaposvar.hu



Új termék elfogadásának determinisztikus és sztochasztikus modelljei

Orova¹ L.-né, Komáromi² A.

¹ zent István Egyetem GK Informatika Tanszék, 2100 Gödöllő, Páter K. u 1.

²Budapesti Corvinus Egyetem, 1093 Budapest IX., Fővám tér 8.

ÖSSZEFOGLALÁS

A piac szereplői újabb és újabb termékek kifejlesztésével tudják versenyképességüket megőrizni. A gyártó cégeknek termelési stratégiájuk, a forgalmazóknak pedig választékpolitikájuk kialakításában nagy segítséget nyújt, ha megalapozott módon becsülhetik meg, mekkora keresletre számíthatnak, és hogy ez időben hogyan alakul. A viszonylag hosszú távú előrejelzés determinisztikus matematikai modell segítségével kísérelhető meg, mely elsősorban a gyártásnak nyújthat támpontot, forgalmazás szempontjából azonban előtérbe kerülhet a váratlan, véletlenszerű események hatása is, melyekre a rövidtávú előrejelzést lehetővé tevő sztochasztikus matematikai modell alkalmazható. Dolgozatunkban a modellek rövid összefoglalása után egy élelmiszeripari példát elemezzünk. Havi vállalati adatokat tartalmazó idősorokat dolgoztunk fel, melynek során új ízesítésű alkoholos ital értékesítésének időbeni alakulásával kapcsolatos vizsgálatokat folytattunk. A vizsgálatok során felhasználtuk az új termékek bevezetésére széles körben alkalmazott determinisztikus, diffúziós Bass modellt, valamint a (főleg árfolyamok, tőzsdei papírok erősen ingadozó idősorainak rövidtávú modellezésében alkalmazott) sztochasztikus modellt, mellyel az új termék bevezetését AutóRegreszív és MozgóÁtlagolású (ARMA) folyamatként kezeltük.

(Kulcsszavak: Új termék elterjedése, Bass modell, ARMA)

ABSTARCT

Deterministic and Stochastic Models of New Product Adoption

L.-né Orova¹, A. Komáromi²

¹Szent István University FME Department of Informatics, H-2100 Gödöllő, Páter K. u 1.

²Corvinus University of Budapest, H-1093 Budapest IX., Fővám tér 8.

Participants of the market retain their good position on the market by developing new and modern products. Valid estimation of the demand can help manufacturers to develop the output strategy and dealers to find the proper product selection policy. Long-term forecasts can support production by deterministic mathematical models above all but unexpected and random effect come to the front in the case of bringing in the good. Stochastic mathematical models are available for short-term forecast in the latter case. In this paper, we analyse an example from food industry after a short summary of different models. The time series of monthly data of a company has been processed and a study was made about the sale rate of a kind of new spiced alcoholic drink in time. We applied the well known deterministic new product diffusion model of Bass and the stochastic models of autoregressive and moving average processes which are widely applied for modelling the strongly fluctuating time series of rates and stock market in our research.

(Keywords: diffusion of new product, Bass model, ARMA)

BEVEZETÉS

Az új termék, technológia elterjedésének ismerete nélkülözhetetlen a termelő cégek számára, ezért külföldi és hazai kutatók már évtizedek óta tanulmányozzák az innováció diffúzióját. Az innováció fogalmkörébe az új terméken, technológián kívül az új piac, beszerzési forrás, valamint új szervezet létrehozását vagy felfedezését már *Schumpeter* (1930) is beleértette.

A technológiai újítások terjedésének vizsgálata a magyar kutatókat is régóta foglalkoztatja, példa erre Eötvös Lőránd báró és Bucsy Béla munkássága, akik e témakörben Fejlődéelmélet címmel közös egyetemi jegyzetet írtak 1919-ben.

Az innováció diffúziójaként *Rogers* (1983) egy olyan folyamatot definiál, melynek során az innováció a társadalom tagjai között bizonyos kommunikációs csatornákon keresztül idővel ismertté válik. A diffúziós folyamat négy alapvető összetevőjének az innovációt, az időt, a kommunikációs csatornát, valamint az adott társadalmi rendszert tekinti. (*Rogers* az innovációt szélesebb körben értelmezi, mint *Schumpeter*).

Egy innováció elfogadásának különböző szakaszai figyelhetők meg: megismerés, érdeklődés, értékelés, kipróbálás, elfogadás. *Rogers* különböző empirikus kutatások alapján az egyéneket csoportokba osztotta

- I. Innovátorok,
- II. Korai elfogadók,
- III. Korai többség,
- IV. Késői többség,
- V. Lemaradók,

annak függvényében, hogy milyen ütemben adoptálják az újítást (1. ábra).

1. ábra

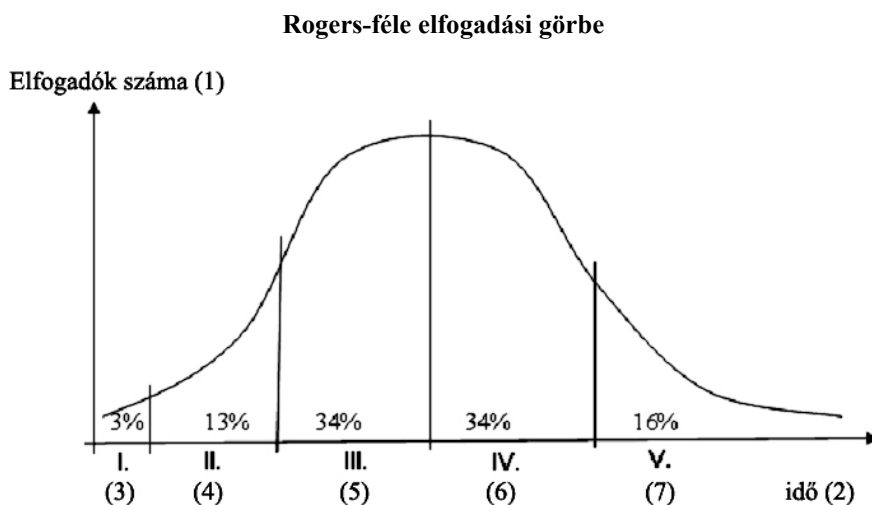


Figure 1: Innovation adoption curve (Rogers, 1962)

Number of adopters(1), Time(2), I. Innovators(3), II. Early Adopters(4), III. Early Majority(5), IV. Late Majority(6), V. Laggards(7),

Diffúzió terjedésének modellezése

A diffúziós modellek célja, hogy egyszerű matematikai összefüggéssel adja meg az innováció bevezetésétől eltelt idő függvényében az innováció terjedésének mértékét a várható „elfogadók” adott körében.

Az elfogadó az, aki megvásárolja és használja az új terméket. Tartós fogyasztási cikkeknel az első vásárlás veendő számításba, más cikkeknel természetesen a vásárlás meg is ismétlődhet. Az elfogadó lehet egy személy, család, vagy egy csoport is, az innovációtól függően. Egy általános diffúziós modell tartós fogyasztási cikk egyszeri megvásárlásával foglalkozik. A teljes piac, $TM(t)$, három részre, szegmensre tagolódik:

1. Jelenlegi piac, $N(t)$, azok az elfogadók, akik az adott időpontig már vásároltak az új termékből.
2. Potenciális piac, $P(t)$, tud a termékről s megvásárolja azt az adott időpontban. Valójában azon halmaz része, akik tudnak a termékről, érdeklődnek iránta, meg is tudják venni a terméket: $TN(t)$, s az adott időpontban meg is vásárolják azt. $P(t) = TN(t) - N(t)$
3. Az érintetlen piac, $TM(t) - TN(t)$, mely vagy nem tud az innováció létezéséről, vagy valamilyen oknál fogva nem számítható az innováció lehetséges fogyasztójának az adott időpontban.

A diffúziós folyamat során az érintetlen piac tagjaiból válnak ki a potenciális piac tagjai, akikből aztán a vásárlás során a jelenlegi piacot bővítő egyének kerülnek ki. Az idők folyamán minden szegmens nagysága változik.

Különböző elképzelések vannak a felől, hogy mi a szegmensek változásának mozgatórugója különös tekintettel arra a mozgásra, mely az egyik szegmensből a másikba irányul. Tapasztalatok alapján az innováció-elfogadás mértékének időbeli alakulása S alakú görbét eredményez, azaz kezdetben kevés új terméket vásárolnak az innovátorok (a Rogers-féle terminológia alapján), majd a korai elfogadók és a korai többség vásárlásai miatt erősen megemelkedik a görbe, a késői többség, de különösen a „lemaradók” vásárlása következtében a görbe ellaposodik.

Modellezésre már a francia *Trade* 1903-ban is ezt az S alakú diffúziós görbét ajánlotta Láng (2003). Véleménye szerint a vezető, vélemény-irányító egyéneket utánozzák a többiek, (opinion leader), de később az S görbén alapuló diffúziós modellek több csoportja is kifejlődött: strukturált, valamint kapcsolati diffúziós hálózatok, melyek az egyén társadalmi kapcsolatain nyugszanak és küszöb- és kritikus tömeg modellek, melyek az elfogadók ill. elfogadások mértékétől teszi függővé az újabb elfogadásokat. Valente (1995)

A strukturált diffúziós hálózatok feltevése, hogy az egyes társadalmi csoportok közötti innováció átvételére a kiterjedt gyenge kapcsolatokkal rendelkező személyek alkalmasak. A kapcsolati diffúziós modell szerint az innováció terjedési sebességét a társadalom tagjainak személyes kapcsolatai határozzák meg. Ezen belül vannak olyan feltevések, hogy az információk először egy meghatározott réteghez jutnak el, csak utána a többiekhez, vagy a csoporton belül az információ állandó sebességgel halad, illetve az ego hálók mérete határozza meg az innováció terjedését, valamint az úgynevezett személyes érintettségi modell, melynek feltételezése, hogy az egyén elfogadását az határozza meg, hogy a személyes környezetében vannak-e jó tapasztalatok. Új termék bevezetésének ilyenfajta vizsgálatához számos szociológiai tanulmány kapcsolódik.

A küszöb- és kritikus tömeg modellek vizsgálatához globális adatok szükségesek, melyek a gyakorlatban jobban rendelkezésre állnak, a klasszikus modellek és e modellek kiterjesztései ilyen típusúak, alapjukat a piaci szegmensek méretének időbeni változása adja Mahajan és Müller (1979).

Bass modell

Bass modell (1969) feltételezi, hogy egy vásárló csak egy egységet vásárol és a teljes, valamint a potenciális piac mérete állandó. ($TM(t)=\text{const.}$, $TN(t)=m$). A tömegkommunikáció hatására vásárolnak kezdetben az innovátorok, és személyes beszélgetés hatására az imitátorok. (Sikeres erőfeszítések történtek a Bass kiterjesztésére a megszorítások enyhítése érdekében. Heeler (1980)

Az alapfeltételezés szerint annak a valószínűsége, hogy új vásárló vásárol egy adott időpontban, az addigi vásárlások lineáris függvénye.

$$P(t) = \frac{f(t)}{1 - F(t)} = p + qF(t),$$

ahol:

$f(t)$ a t időpontban történő vásárlás feltétel nélküli valószínűsége,
 $F(t)$ a t ideig történő összes vásárlás valószínűsége.
 p az innovációra jellemző paraméter (az első vásárlás valószínűsége $t=0$ -ban)
 q az imitációra jellemző paraméter és

$$F(0)=0, \text{ valamint } F(T) = \int_0^T f(t)dt.$$

Adott időpontban a vásárlás valószínűsége:

$$f(t) = (1 - F(t)) * (p + qF(t)) = p + (p - q)F(t) - qF(t)^2,$$

ahol:

adott időpontban a vásárlások: $Y(t) = mf(t)$ és
 az adott időpontig az összes vásárlás: $N(t) = mF(t)$, ahol m potenciális piac vásárlása a termék teljes élettartama alatt.

Az adott időpontban a vásárlások a fenti egyenletek alapján:

$$Y(t) = mp + (q - p)N(t) - \frac{q}{m}N(t)^2$$

A termék sikeres, $q > p$ esetén.

A legnagyobb elfogadás mértékét, $Y(t_{\max})$, és időpontját, $t_{\max}$, a kiindulási feltételezésen alapuló összefüggésekből levezetve alábbiak szerint adódik:

$$Y(t_{\max}) = \frac{m(p + q)^2}{4q} \quad t_{\max} = \frac{1}{p + q} \ln\left(\frac{q}{p}\right),$$

Több adatsor vizsgálata alátámasztotta a gyakorlatban a fenti összefüggések helyességét.

Az idősorok vizsgálatakor a folytonos modell helyett célszerű diszkrét időintervallumokat vizsgálni, így a T-edik időintervallumban az új elfogadás valószínűsége:

$$\frac{Y(T)}{m - N(T - 1)} = p + q \frac{N(T - 1)}{m}$$

ahol:

$Y(T)$ az új elfogadások száma a t-edik időintervallumban.

$N(T-1)$ a $t < T-1$ időtartamban a cumulatív elfogadások száma.

A fenti képletből az alábbi, gyakorlatban is jól használható összefüggés adódik:

$$Y(T) = mp + (q - p)N(T - 1) - \frac{q}{m}N(T - 1)^2$$

A diffúzióra jellemző paraméterek. p , q , és m meghatározhatók a valós adatsorokból nyert $Y(T)$ és $N(T-1)$ értékpárookra a legkisebb négyzetek módszerével illesztett másodfokú polinommal, hiszen a fenti egyenlet felírható az alábbi formában:

$$Y(T)=a+bN(T-1)+c[N(T-1)]^2,$$

ahonnan

$$q=-mc, \quad p=a/m \quad \text{és} \quad m = \frac{-b \pm \sqrt{b^2 - 4ac}}{2c},$$

$m>0$, $b \geq 0$, $c>0$ feltétel teljesülése mellett.

Valós adatsor alapján meghatározott paraméterek lehetővé teszik a termék jövőbeni elfogadásának előrejelzését.

Az innováció diffúziójának Bass modelljét eredményesen alkalmazták többféle termékre, különböző országokban, a modell elterjedtségének egyik mutatója, hogy a különböző országokban, többféle termékre meghatározott p és q értékek alapján tanulmányozzák a diffúzió sebességének változását országok és termékek között Blute (2002).

A Bass modellt alkalmazva egy termékre, a megfigyelés időtartamától függően eltérő eredmények adódnak a főbb paraméterekre. A jelenség magyarázható azzal, hogy a körülmények változnak, ezért elméleti kutatások folynak a Bass modell sztochasztikus megfogalmazására Niu (2002).

Sztochasztikus idősortmodellek

Gyakorlati alkalmazásokban műszaki területeken évtizedek óta használják a sztochasztikus idősortmodelleket, melynek lényege, hogy egy független változót viszonyítanak saját múltbeli értékeihez és véletlen hibáihoz. Az utóbbi évtizedekben más területeken, pl. az ökonometriában is sikeresen alkalmazzák rövidtávú előrejelzésekre, például árfolyamok, tőzsdei papírok erősen ingadozó idősorainak vizsgálatában. Felmerül a kérdés, lehetséges-e az innováció diffúzióját ilyen módon eredményesen modellezni.

Az idősortmodelleknél az idősort trend (T_t), szezonális (S_t) és sztochasztikus (u_t) komponens alkotja Ramanathan (2002):

$$Y_t = T_t + S_t + u_t$$

Az effajta modellezés célja megállapítani, hogy mi határozza meg a folyamatot:

- a vizsgált jelenség függ-e korábbi értékétől, megfigyelhető-e öröklődés, amit időben független véletlen változó zavar, (ami maga is beépül a folyamatba, autoregresszív modell: AR)
- valamilyen véletlen külső lökések határozzák meg a folyamatot, melyek hol feltűnnek, hol elhalnak és a folyamat egyáltalán nem emlékszik ezekre a külső hatásokra (mozgóátlag modell: MA)

$$Y_t = \nu_t - \beta_1 \nu_{t-1} - \beta_2 \nu_{t-2} - \dots - \beta_q \nu_{t-q}$$

- ezen hatások kombinációja figyelhető meg (autoregresszív és mozgóátlagolású modell: ARMA), esetleg szezonális hatásokkal.

Célunk a determinisztikus Bass modell és a sztochasztikus ARMA modell alkalmazhatóságának vizsgálata előrejelzésre új élelmiszeripari termék esetén.

ANYAG ÉS MÓDSZER

Az innováció diffúzióját egy élelmiszeripari termék, új ízesítésű alkohol kapcsán vizsgáltuk. Rendelkezésünkre a gyártó havi kiszállítási álltak 1995. októberétől 2003. októberéig, mely jelen esetben majdnem egy teljes termék-életgörbét jelent.

Az adatsorra adaptáltuk a Bass modellt, 4 év adatait vettük figyelembe, s a meghatározott paraméterek alapján 4 évre előrejelzést végeztünk.

Sztochasztikus modell felírásához a Box-Jenkins módszert, a számolási munkák egyszerűsítése érdekében az Eviews ökonometriai programcsomagot alkalmaztuk. Az ex post előrejelzés mintája a 7 éves adatsor havi adatai, az előrejelzési periódus 1 év. A stacionaritást az Eviews program beépített lehetőségével teszteltük. A folyamat valószínű jellegét korrelogram alapján határoztuk meg, majd különböző trendet és folyamatot (ARMA), feltételezve megvizsgáltuk az illeszkedés és a modell „jószágát” különböző kritériumok alapján.

A sztochasztikus előrejelzést összehasonlítottuk a valós adatokkal, valamint a Bass modell alapján kapott előrejelzéssel.

EREDMÉNY ÉS ÉRTÉKELÉS

A determinisztikus Bass modell paramétereire a 4 éves havi adatokat alapul véve (2. ábra): $m=1\,425\,144$, azaz ennyi a potenciális piac összes felvevőképessége, az innovációs tényező, $p=0,0042$ és az imitációs tényező $q=0,043$.

Átfogó felméréseket végeztek különböző termékek elterjedésének vizsgálatára, melynek során meghatározták különböző termékek Bass-féle paramétereit *Lilien* (1999). Körülbelül 10 évre vonatkozó adatokat dolgoztak fel s az innovációs tényezőre átlagosan 0,040, az imitációs tényezőre pedig 0,398-at kaptak. Ezek az értékek egy nagyságrenddel nagyobbak, mint a jelen modellezés eredménye, de a tényezők érzékenyek a mintavételi gyakoriságra, amiben 12 szerez különbség van, mivel mi havi adatokkal dolgoztunk a minél pontosabb előrejelzés érdekében, s nem utolsó sorban azért, hogy a havi adatokat alapul vevő sztochasztikus előrejelzéssel össze lehessen hasonlítani a determinisztikus modell eredményét. (Éves adatokon $p=0,048$ és $q=385$, s ez megfelel *Lilien* felmérésének is.)

2. ábra

Négy éven alapuló Bass-féle előrejelzés

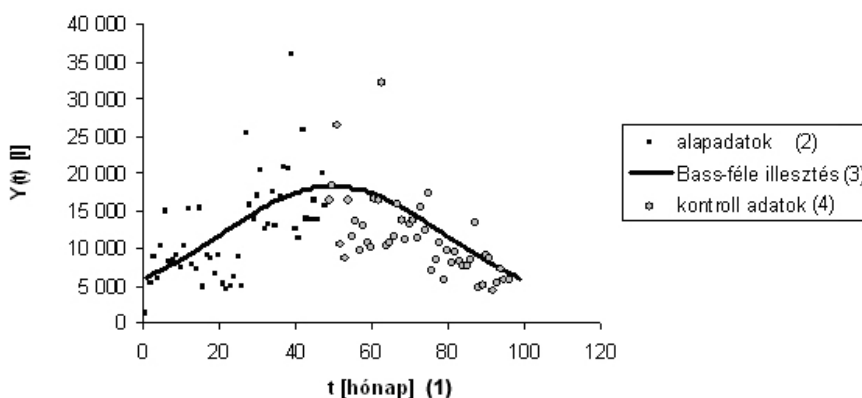


Figure 2: Bass model forecast based on four years data

Time [month](1), Base data(2), Fitted Bass function(3), Control data(4),

A legjobbnak talált sztochasztikus modell a másodrendű autóregrszív és másodrendű mozgóátlag modell: SARIMA(0,12,0)(2,0,2), melynél 12 hónapos szezonális hatás kiszűrésével sikerült a modellezés érdekében stacionárissá tenni az adatsort. A folyamat Megfigyelhető továbbá, hogy 1997. 12. hó és 2001. 12. hó között magasabb volt a kibocsátások átlaga, mint az idősor elején és végén, ezt egy „dummy” változóval kezeltük, s így a trendfüggvény egy három szakaszból álló konstansfüggvény lett, amire jó illeszkedést kaptunk. A eredeti adatsor és az előrejelzés a 3. ábrán látható.

3. ábra

ARMA előrejelzés

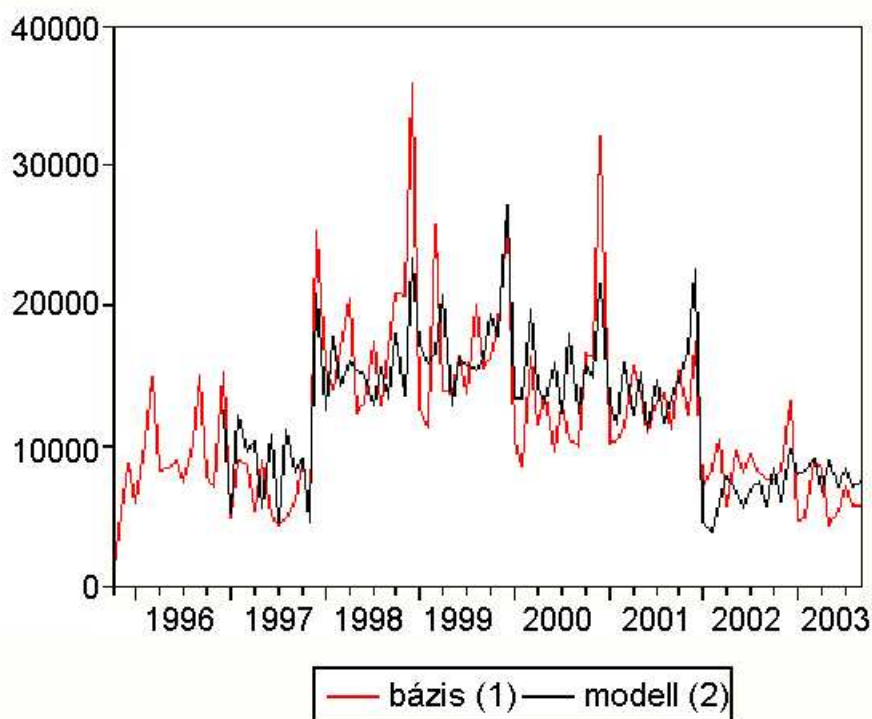


Figure 3: ARMA trend prediction

Base(1), Model(2)

A kétféle előrejelzés eredményének szemléletes összehasonlítása: 4. ábrán látható. Megfigyelhető, hogy az előrejelzett értékek nagyobbak mindkét módszernél, mint a ténylegesek. A sztochasztikus előrejelzés az első négy hónapra jobb közelítést eredményez, mint a Bass modell, tekintetbe véve azonban azt a tényt, hogy a Bass modellnek ez a negyedik előrejelzett éve, megállapítható, hogy hosszútávú előrejelzésre ezen élelmiszeripari termék esetén sikeresen alkalmazható a tartós fogyasztási cikkek vizsgálatára kifejlesztett Bass modell.

4. ábra

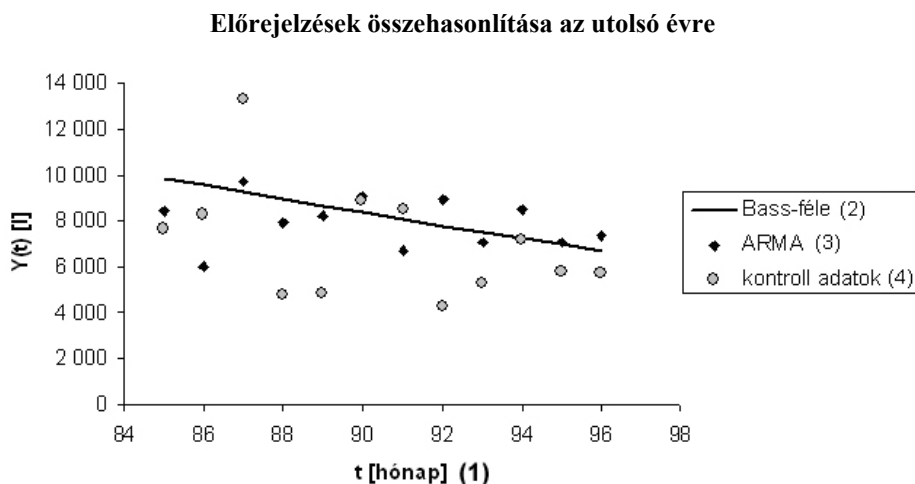


Figure 4: Comparison of last year forecasts

Time [month](1), Bass(2), ARMA(3), Control data(4)

KÖVETKEZTETÉSEK

A viszonylag hosszú távú előrejelzés a determinisztikus Bass modell segítségével eredményesnek bizonyult, ami elsősorban az új termék gyártóinak nyújthat támpontot, forgalmazás szempontjából azonban a sztochasztikus ARMA modell nyújt megbízható előrejelzést, a havi ingadozásokra érzékenyen.

IRODALOM

- Bass F. (1969). A New Product Growth Model for Consumer Durables Management Science 15 (January), 215-227.
- Heeler R.M., Hustad T.P. (1980). Problems in Predicting New Product Growth for Consumer Durables, Management Science, 26(10)
- Láng S., Letenyi L., Siklós V.(2003). Információs technológia és helyi társadalom. Budapest, BKÁE, II. kötet 5-28. www.socialnetwork.hu
- Lilien G.L., Blute M. (1999). Diffusion Models: Managerial Application and Software, ISBM Report 7-1999, Institute for the Study of Business Markets, 7.
- Mahajan V., Müller E. (1979). Innovation Diffusion and New Product Growth Models In: Marketing, Journal of Marketing, 233-247.
- Ramanathan R. (2002). Introductory Econometrics with Application, Panem Könyvkiadó : Budapest, 531-541.
- Rogers E. (1995). Diffusion of Innovations Fourth Edition, The Free Press, New York, 263-268.
- Schumpeter J. (1987). Theorie der wirtschaftlichen Entwicklung, Berlin, Drucker & Humblot.

- Shun-Chien Niu (2002). A stochastic Formulation of the Bass Model of New-Product Diffusion, In: Mathematical Problems in Engineering, 2002, 8. 3. 249-263.
- Valente T. (1995). Network Models of the Diffusion of Innovations. New Jersey : Hampton Press

Levelezési cím (*Corresponding author*):

Orova Lászlóné

Szent István Egyetem, GK, Informatika Tanszék

2100 Gödöllő, Páter K. u. 1

Szent István University, FME Department of Informatics

H-2100 Gödöllő, Páter K. u. 1.

Tel.: 36-28-522-051

e-mail: orova.laszlone@gek.szie.hu

Kocsor A., Bánhalmi A., Paczolay D.

MTA-SZTE Mesterséges Intelligencia Tanszéki Kutatócsoport, 6720 Szeged, Aradi vértanúk tere 1.

ÖSSZEFOGLALÁS

*Az automatikus beszédfelismerési technológiák jelent
sztohasztikus szegmentális megközelítés. Mindkét modell felépítéséhez egy nagym*

ABSTRACT

Technological background of a speech recognition system for the dictation of thyroid gland medical reports

A. Kocsor, A. Bánhalmi, D. Paczolay

Research Group on Artificial Intelligence of the address: Hungarian Academy of Sciences and University of Szeged,
H-6720 Szeged, Aradi vértanúk tere 1.

*With the considerable development of speech recognition technologies in several
administration-requiring professions the demand for the so called speech-based documentation
has grown. The use of this procedure is of great importance in the case of medical reports with
special linguistic features few systems for dictating medical reports have been developed so far*

which fact can be attributed to linguistic specialties and high development expenses. In Szeged we developed a core module capable of automatic recognition of the Hungarian language on which several domain oriented systems can be built. The core module contains the so called acoustic model, which is suitable for the recognition of the Hungarian phoneme set and representative modeling. For the building of the model we used two significantly different approaches. One is the Hidden Markov Model, well known in speech recognition, the other is the novel stochastic segmental approach developed in Szeged. For the development of both models we used a large speech corpus with 500 speakers, and then the performance of the modules was tested on test databases. To accompany the core module we built a language module (for Windows environment) suitable for the dictation of thyroid gland medical reports in order to justify the applicability of the developed methods. The module was built on 9231 written thyroid medical reports and over 2500 word forms. We present the structure of the developed system for dictating thyroid gland medical reports, the technology of the built language and acoustic models, the test results describing the efficiency of the models, furthermore we mention the different aspects of the application and technology of the software.

(Keywords: continous speech recognizer, ASR, automatic speech recognition, dictation system, HMM, Hidden Markov Model, MSD, Morphosyntactical descriptor, grammar, accoustic model, N-gram)

BEVEZETÉS

Az informatika gyors fejlődése az utóbbi évtizedekben egyre hatékonyabb beszédfelismerő rendszerek létrehozását tette és teszi lehetővé. A cikkünkben betekintést adunk egy magyar nyelvű diktáló rendszer felépítésének elméleti háttérébe és technikai részleteibe. Leírást adunk az egymástól eltérő akusztikai felismerő modellekről és összehasonlítjuk ezeket különböző tesztadatbázisokon. Bemutatjuk a folyamatos beszéd felismeréséhez kötődő problémákat és azokat a megoldási lehetőségeket, amelyek segítségével egy hatékony diktáló rendszer építhető fel. Egy folyamatos beszédet felismerő rendszer alapvető fontosságú része a megfelelő nyelvi modell. Kutatócsoportunk az elmúlt években kutatási célokra kifejlesztett egy általános célú beszédfelismerőt, amelynek most elkészült egy scintigráfias leletek diktálására alkalmas változata. Az új rendszer segítségével teszteljük le a különböző nyelvi modellek szerepét.

ISMERTEBB DIKTÁLÓ RENDSZEREK

A beszédfelismerő rendszerek nemzetközi piacán mára szinte uralkodóvá vált a Scansoft, amely több felismerő rendszert vásárolt fel és kínál különféle néven. Ilyen termék angol nyelvre a Dragon Naturally Speaking, vagy a IBM ViaVoice (Smith, IBM, 2002). Ezek a rendszerek több éves múltra tekintenek vissza, tapasztalataink szerint igen megbízhatóak a felismerési pontosságot tekintve. A Dragon Naturally Speaking-nek létezik egy speciális, orvosi szakszöveg diktálására alkalmas változata is. A rendszer hét nyelvre vásárolható meg.

A Scansoft piacvezető cég még nem adta ki a magyar változatot, ami azért sem meglepő, mert az agglutináló nyelvekre lényegesen nehezebb felismerőt készíteni. A magyar nyelvben a problémát az okozza, hogy nagyon magas a szóalakok száma a ragozás miatt.

Magyar nyelvre az első „ipari” beszédfelismerőt a Philips készítette el 2004. júliusában, éppen orvosi diktálásra. A Philips közel 10 éve foglalkozik ilyen rendszerekkel, a magyar változat kiadása előtt már 21 nyelvre készített beszédfelismerőt (Nyers, 2004). A szoftvert – igen magas ára miatt – még nem állt módunkban tesztelni, a

leírások alapján hosszabb adaptáció után a szavakra számított felismerési pontosság akár 95%-os eredményt is elérhet.

Kutatócsoportunk közel 10 éve foglalkozik beszédfelismerő rendszerek kifejlesztésével, valamint ehhez kötődő alapkutatásokkal. Mivel kiindulási alapok nem álltak rendelkezésünkre, szükség volt nagy méretű beszédatbázisok létrehozására, azok felszegmentálására, valamint különböző felismerő és nyelvtani modulok elkészítésére egyrészt a szakirodalom, másrészt a saját kutatásaink alapján.

KÉT ELTÉRŐ MEGKÖZELÍTÉS A BESZÉDFELISMERÉSBEN

Jelenleg a beszédfelismerésnek – mint a Mesterséges Intelligencia kutatás egyik leginkább kutatott ágának – egyik legfontosabb elérendő célja egy általános diktáló rendszer létrehozása, amely a mikrofonba bemondott mondatokat helyesen alakítja át írott szöveggé. Egy ilyen diktáló rendszer létrehozása nagyon sokféle matematikai (*Vapnik*, 1998), informatikai valamint nyelvészeti problémát vet fel. Figyelembe kell venni az adott nyelv fonetikai-akusztikai sokszínűségét, nyelvtani és szórendi sajátosságait. Emellett a másik nagyon fontos problémakör az, hogy a rendszernek fel kell ismernie a bemondott szavakat, illetve azokat a kisebb akusztikai egységeket, amikből a szavak felépülnek.

A közép- és nagyszótáras felismerők mindegyikének gyakorlati megvalósításakor a legkisebb egység, amelyet a rendszernek fel kell ismernie, az a fonéma. A szavak felismerése ezen alkotóelemek felismerésén keresztül valósul meg. A téma kutatása során több különböző gépi tanuló-osztályozó algoritmus (*Duda et al.*, 2001) létrehozására volt szükség a nyelvi alapegységek és az összetettebb struktúrák (szavak, mondatok) felismerésének érdekében (*Huang et al.*, 2001). Az ilyen algoritmusoknak két legfontosabb ága a HMM (Rejtett Markov Modell) alapú (*Becchetti and Ricotti*, 2000) illetve a szegmens alapú megközelítés (*Kocsor et al.* 1999).

A két irányvonalban közös, hogy a mikrofonból érkező digitalizált beszédjelből kis időközönként megfelelő méretű mintát veszünk, és minden ilyen kis jeldarabból bizonyos számú jellemzőt vonunk ki, amelyekkel az adott jeldarabot jól jellemezni tudjuk (*1. ábra*). A beszédfelismerésben használt jellemzőkinyerő algoritmusok száma igen nagy, amelyek közül az összetettebbek a hallás és a központi idegrendszer jelfeldolgozásának vizsgálatából származó tudományos eredményeket is figyelembe veszik (*Rabiner and Schafer*, 1978; *Moore*, 1997).

További közös jellemzője a beszédfelismerésben használatos tanuló algoritmusoknak az, hogy a gépi tanulásnak az ún. felügyelt tanulás ágába tartoznak. Ez azt jelenti, hogy például az "a" fonéma tanulásakor megfelelő mennyiségű "a" fonéma példánynak kell előre rendelkezésre állnia, maga a tanuló algoritmus ezeket a példákat "tanulja meg". Ezért a tanító adatbázisokat úgy kell kialakítani, hogy a rögzített beszéden kívül tartalmaznia kell az egyes fonémák időbeli helyét a jelen belül. A fonetikai adatbázis ilyen módon történő előállítását szegmentálásnak nevezzük (*2. ábra*). Egy megfelelő minőségű kisszótáras beszédfelismerő szoftver létrehozásához a tapasztalatok szerint több órányi beszédet tartalmazó felszegmentált adatbázisra van szükség.

HMM alapú megközelítés

A HMM (Hidden Markov Model) egy statisztikai alapú tanuló algoritmus, amelyet beszédfelismerésre az 1970-es évek közepétől alkalmaznak, és manapság a módosított változataival együtt a leginkább elterjedt megoldásnak számít (*Becchetti and Ricotti*, 2000). Lényegében arról van szó, hogy minden egyes fonémához létrehozunk egy – általában – három állapotú, balról jobbra elrendezésű rejtett markov modellt (*3. ábra*).

1. ábra

Jellemzőkből álló vektor előállítása a szignálból.

3. ábra

Három állapotú HMM

szegmenshatárok időbeli pozíciói kvantáltak, elég például milliszekundum pontossággal meghatározni. Ennek az egyik oka az, hogy nem lehetséges a fonémák pontos elkülönítése, mert a folyamatos beszéd során a fonémák folytonosan kapcsolódnak egymáshoz, nincsen közöttük éles határ. A határok bizonytalanságából eredő hibát a sok tanítópélda szerencsére képes ellensúlyozni. A másik oka annak, hogy a határokat csak pár milliszekundumos pontossággal keressük, hogy ez nagymértékben gyorsítja a felismerőt, anélkül, hogy hatékonysága észrevehetően csökkenne. További gyorsítási lehetőség minimális illetve maximális hossz megadása a beszédhangokra (Tóth et al., 2000).

További jellemzője a szegmens alapú megközelítésnek, hogy a HMM-nél általánosabb modellnek tekinthető, mert a minimum, maximum és a lépésköz paramétereknek bizonyos beállítása mellett a HMM-nek egy jó közelítését kapjuk meg speciális esetként.

4. ábra

Egy lehetséges HMM modell hierarchia.

A kétféle irányvonal kísérleti összehasonlítása

A következő fejezetben összehasonlítjuk a HMM és a szegmens alapú akusztikus modelleket. Először ismertetjük az összehasonlításunk módszertani hátterét, majd kiértékeljük a kapott eredményeket.

Az összehasonlítás metodikája, felhasznált beszédatadatok

A különböző beszédfelismerő algoritmusok akusztikus modellező képességének összehasonlításakor a következőket szokás elvégezni:

- A fonéma felismerés pontosságának mérése.
- Izolált szó felismerés pontosságának mérése.

Magasabb szintű felismerési hatékonyság mérése (mondatra illetve szövegre) ebben a kontextusban nem szükséges, mivel ez az összehasonlítás inkább az összetett felismerő rendszer, a nyelvtan, a vágás, stb. jószágát méri, és kevésbé a felismerő mag szerepét.

A kiértékelés során fontos, hogy az összehasonlítandó felismerő modulok ugyanazokon az adatbázisokon legyenek tanítva, illetve ugyanazokon legyenek tesztelve is.

A tanításhoz és teszteléshez a következő adatbázisokat használtuk fel (ezek jellemző statisztikai adatait is feltüntetjük az 1. táblázatban):

1. táblázat

A tanításhoz és a teszteléshez felhasznált adatbázisok

Tanító adatbázis (1)	Fájlok száma (2)	Fonéma szám (3)	Teszt adatbázis (4)	Fájlok száma	Fonéma szám
Szám (5)	2728	20236	Szám	1224	9075
Mtba (6)	3199	154014	Mtba-városok (8)	431	nem szegmentált (9)
Beme-laptop (7)	3315	23762	Beme-laptop	1110	7981
Mtba	3199	154014	Mtba-teszt (10)	687	33159
Timit	3696	142910	Timit	1344	51681

Table 1: Database used for training and testing

Database for training(1), The total number of wave files(2), The total number of phonemes(3) Database for testing(4) Numbers(5), MTBA: Hungarian Telephone Speech Database(6) Our special speech database (7) MTBA cities(8), Not segmented(9), MTBA test(8),

A szám adatbázisok 1 és 1000000 közötti számok bemondását tartalmazzák. Az Mtba adatbázis különböző telefonon keresztül bemondott szövegek állományait tartalmazza. Ezen belül az Mtba-városok városnevek bemondásaiból áll. A Beme-laptop adatbázis a Beszédmester saját fejlesztésű beszédterápiás szoftverhez készített, férfiak, nők és gyermekek bemondásait tartalmazó adatbázis. A Timit adatbázison kívül az összes adatbázis magyar nyelvű bemondásokat tartalmaz, és teljesen, vagy részben Szegeden készült.

Fonéma felismerés hatékonysága

A fonéma felismerés hatékonyságának összehasonlításakor a felismerő modulokkal szavakat ismertetünk fel, majd összehasonlítjuk a felismert fonémasorozatot a referencia fonémasorozattal. Az általánosan elfogadott összehasonlító eljárás az un. szerkesztési

távolság (Edit Distance). Az Edit Distance (Crochemore and Ryller, 1994) eljárás azt méri, hogy mennyi módosítással lehet megkapni egyik sorozatból a másikat úgy, hogy csak a törés, a beszúrás és az átírás műveleteket használhatjuk. Legyen például a bemondott szó a "scintigram", ekkor a referencia fonéma sorozat (a szabványos SAMPA kódrendszer segítségével megadva): "s - 'ts i n - 't i + 'g r O m". A felismerő által jelzett sorozat lehet például "s - 'ts s i E m - 't i E i + 'g O m". Jelen esetben 4 törlés, 1 átírás és 1 beszúrás művelettel kaphatjuk meg az eredeti sorozatból a felismert sorozatot, így a kettő távolsága 6. A teljes tesztadatbázisra vonatkozó hibát két mennyiséggel szokás jellemezni: $correction=(N-d-r)/N$, illetve $accuracy=(N-i-d-r)/N$, ahol N a referencia sorozatok teljes hossza, i a beszúrás (insertion), d a törlés (deletion), r az átírás (replace) műveletek száma. A 2. táblázatban foglaljuk össze a kétféle irányvonallal kapcsolatos eredményeket:

2. táblázat

A HMM és szegmens alapú fonéma felismerés hatékonysága

Adatbázis (1)	correction	accuracy	insertion	deletion	replace	N
Mtba-teszt (2)	53.67%/55.99%	43.46%/46.93%	3266/3129	3524/3536	12734/11659	34532
Timit	62.38%/63.09%	53.08%/54.20%	4681/4596	4809/4858	14759/14215	51681

Table 2: The efficiency of the segment and the HMM based phoneme recognition

Databases(1), Hungarian Telephone Speech Database Databases used for testing(2)

A táblázatból az olvasható le, hogy a szegmens alapú felismerő valamennyivel jobb felismerési eredményeket ér el fonéma szinten, mint a HMM. A fonémák illetve fonémacsoportok tanulására a szegmens alapú felismerő ANN (Artificial Neural Network, Bishop, 1995) osztályozó algoritmust használt, míg a HMM az alapértelmezett GMM-et (Becchetti and Ricotti, 2000).

Izolált szó felismerés hatékonysága

Izolált szavas teszteléskor a felismerőnek nem fonéma sorozatot kell megadnia, hanem el kell döntenie, hogy melyik szó lett bemondva. A lehetséges szavakat egy szótár tárolja. A teszt eredménye a felismerési pontosság, amelyet a 3 táblázatban foglalunk össze.

3. táblázat

A HMM és a szegmens alapú szó szintű felismerés hatékonysága

Adatbázis (1)	HMM	Szegmens alapú felismerő (2)
Szám (3)	99.50%	98.56%
Mtba-városok (4)	93.35%	89.96%
Beme-laptop (5)	98.47%	93.83%

Table 3: The efficiency of the HMM and the segment-based word level recognition

Databases(1), Segment based recognition(2) Numbers(3), MTBA cities(4), Our special speech database(5)

Izolált szó felismerésénél határozottan jobb a HMM alapú felismerő a szegmens alapúnál, csak 1 helyen értek el közel azonos felismerési pontosságot. Ennek az lehet az oka, hogy a szegmens alapú modellek nem mindig találják meg a szegmens határokat, illetve a HMM egy szegmenst több minta alapján értékel ki, így stabilabb működésű. A fonémák illetve fonémacsoportok tanulása az előző bekezdésben alkalmazott tanítással egyezik meg.

A PAJZSMIRIGY LELETEK DIKTÁLÁSÁRA ALKALMAS RENDSZER FELÉPÍTÉSE

A két magmodul felismerési képessége nem tér el lényegesen, de a HMM alapú rendszer folyamatos felismerésre alkalmazható változata könnyebben implementálható. A további vizsgálatainkat ezért csak a HMM alapú akusztikus magmodul használatával végeztük el. Ebben a fejezetben leírjuk a folyamatos felismerés technikai hátterét és a nyelvi modell felépítésének módját.

A nyelvi modell tanításához használt adatbázisok

Az orvosi beszédfelismerő nyelvtani modelljének létrehozásához egy pajzsmirigy szcintigráfias leletekből álló szövegkorpuszt használtunk. Az írásos vizsgálati anyagokat 1998 és 2004 között rögzítettük. A közel 9000 leletet a különböző formátumokból egy közös szöveges formátumra kellett konvertálnunk, majd több lépéses javítási folyamat következett. A vizsgálatokról készült minden egyes lelet a következő részeket tartalmazza:

- a) fejléc
- b) klinikai adatok
- c) kérdés
- d) előző vizsgálata
- e) jelen vizsgálata
- f) összefoglaló vélemény
- g) aláírás

A szövegkorpuszból töröltük a hiányos leleteket az átvizsgálása során. A szöveges adatbázis létrehozásakor az a) és g) részek nem lettek felhasználva. A végleges adatbázis 8546 szövegből áll. A szövegben 2500 szóalak fordul elő (számok és dátumok nélkül). Átlagosan 11 mondatot, és mondatonként 6 szót tartalmaz egy-egy lelet. A mondatonkénti szavak számának eloszlása nem normális eloszlást mutat, ami annak tudható be, hogy a közel 95000 mondat közül mindösszesen 12500 különböző mondatot tartalmazott az adatbázis.

Folyamatos beszédfelismerés gyakorlati megoldása

Egy diktáló rendszertől egyrészt azt várjuk el, hogy a felismert szöveg a bemondottal megegyezzen, másrészt azt, hogy az írott szöveggé alakítást a bemondással szinkronban végezze. Ha a rendszer úgy működne, hogy minden szó után minden szó következhet, és minden lehetséges szósortozatot a felismerés végéig megtartunk, mint hipotézist, akkor a hipotézisek száma rövid idő alatt kezelhetetlen méreteket öltene, a felismerés sebessége pedig messze elmaradna a bemondás ütemétől. Ennek a problémának a gyakorlati megoldásaként az ún. Multi Stack Decoding-ot használhatjuk (*Rabiner and Juang, 1993*). Ennek egy lényeges eleme az, hogy minden időpontban valószínűségük alapján rangsoroljuk az aktuális hipotéziseinket (amelyek tulajdonképpen fonémasorozatok). A rangsorolás valamilyen érték, általában a valószínűség alapján történik. Ezután csak azokat a hipotéziseket tartjuk meg, illetve folytatjuk, amelyek a sorban elől állnak. A

legelterjedtebb két ilyen vágási technika az un. N-Best (az első N legjobb megtartása) illetve a Viterbi (a maximális értéktől való eltérés maximálása). Ahhoz, hogy a folyamatos beszédfelismerő megfelelő hatékonysággal működjön, elegendő, de nem túl sok hipotézist kell időről időre megtartani és tovább vinni, el kell kerülni az azonos hipotézisek ismétlődését, és meg kell találni a megfelelő rangsoroló függvényt. A felsorolt feltételeket a nyelvi modell segítségével teljesíthetjük.

Nyelvi modell fontossága

Mint korábban láttuk, alapvetően a felismerők a legvalószínűbb fonémasorozatokat adják vissza. A nyelvtan szerepe e sorozatok szűrése menet közben, aminek célja a hipotézisek számának minél hatékonyabb korlátozása. Erre van lehetőség, hiszen feltételezhető, hogy csak értelmes szavakat kell felismernie a rendszernek. A szűrés azt jelenti, hogy nem következhet minden szó után minden szó, illetve minden fonéma után minden fonéma egy adott időpontban, a nyelvtan leszűkíti a lehetséges következő fonémák és egyben a következő szavak halmazát, és lehetőség szerint megadja azok valószínűségét.

A szó N-gramm

A legegyszerűbb nyelvtan nem más, mint egy szótár, amelyben a lehetséges szavak és azok kiejtései szerepelnek. Az egyszerű szótár alapú rendszerek hátránya, hogy nem tárolnak semmilyen hosszabbtávú információt a felismert előzményről. A szó N-gramm egy olyan módszer, ami ezt a problémát oldja meg részben. Ez egy statisztika, amely megadja, hogy milyen valószínű egy adott szó az N-1 darab előtte álló szó ismeretében. A gyakorlatban csak kis N-re van esély akkora adatbázis összegyűjtésére, hogy jól közelítsük a beszélt nyelvet. Mivel az N-gramm egy véges korpuszból leszámolt statisztika, így az ezen az alapon működő nyelvtan túl szigorú lehet. Ha egy szó-N-es nem szerepelt az adatbázisban, akkor a nyelvtan azt adja, hogy az ilyen bemondás helytelen. Különböző módszerekkel lehet változtatni az N-gramm szigorúságán, például, ha egy kis pozitív ϵ konstanszt rendelünk a hiányzó szó-N-esek valószínűségéhez (Huang et al., 2001). Az ϵ érték megválasztása nem triviális, mert, ha túl nagy, akkor rossz sorozatok valószínűbbek lehetnek, mint a helyesek. Egy jobban parameterezhető rendszert kapunk, ha ϵ értékét nem konstansnak választjuk. Az ϵ értékét más-más szó-N-esekre az ismert szó-gramm valószínűségi értékek segítségével tudjuk megbecsülni. Tapasztalat szerint a kizárólag N-grammokra épülő nyelvtanok kötött szórend esetén jobban működnek, mint szabad szórend esetén. Kötetlen szórendű nyelv használatakor sajnos további probléma az, hogy kis N-re akár olyan köröket is megenged az N-gramm, ami nagyobb N használatakor nem fordulhatna elő. Azonban az N növelésével exponenciálisan kellene növelni a tanító adatbázis méretét, és ezzel párhuzamosan nőne a tárigény is. Sajnos a magyar nyelv szabad szórendűsége miatt mégis modelleznünk kell hosszú távú kapcsolatokat is, ezért más modellekre is szükség van.

MSD kód alapú szabályok

Az MSD kódolás (morfoszintaktikai kódok) minden szóhoz hozzárendel egy vagy több, 8 hosszúságú karaktersorozatot, amelyek a szavak lehetséges mondattani szerepeit írják le. Az MSD kódok alapján létrehozott szabályokat a következőképpen kapjuk meg: a szövegkorpusz mondataiban lévő szavakat lecseréljük azok mondattani szerepét leíró MSD kód első karakterére. A cserével párhuzamosan előállítjuk a szócsoportokat is, amelybe tarozó szavak ugyanabban a nyelvtani szerepben állhatnak a mondatban (pl. jelző, határozó). Mivel egy-egy szónak más-más mondattani szerepe lehet a különböző mondatokban, így a csoportok nem feltétlenül diszjunktak. A folyamat végén minden

mondathoz egy kódsorozatot rendelünk, ezek összessége adja meg az MSD kódos nyelvtani szabályokat (az azonosak összevonása után). A megvalósítás szempontjából két lehetőség nyílik az elkészült nyelvtan felépítésére. Az egyik lehetőség, hogy a csoportokból keletkezett szótárakat egymás után összefűzzük a szabályok alapján, a másik lehetőség, hogy a szótárakat nem helyettesítjük be, csak a kódját tároljuk el. Az utóbbit beágyazott megoldásnak nevezzük.

Az MSD kódos leírásnál a szavak jelentése eltűnik, csak a szavak mondattani szerepe marad meg. Egy szemléletes példa erre, az hogy, ha tanítás során a „piros alma”, „sárga banán” is előfordul, akkor a nyelvtan előállítja a „piros banán”-t is. Az ismertetett N-gramm használatával viszont elvárható, hogy az ilyen jellegű problémák csökkennek, mert megmondja, milyen valószínűséggel követik egymást a szavak. A nyelvtan általánosító ereje az ϵ paraméterrel szabályozható.

A hasonulás

Folyamatos beszédben igen gyakori a szavak közötti hasonulás, ami akusztikailag megváltoztathatja a szavak végét vagy elejét. A hasonulás kezelését bonyolítja az is, hogy nem tudjuk, hogy a beszélő tart-e szünetet a szavak között vagy sem, ezért ezeknek a hasonulást leíró szabályoknak, mint alternatíváknak kell megjeleníteniük a nyelvtanban. A gyakorlati megvalósítás szempontjából az a kevésbé bonyolultabb, ha előre meghatározzuk a hasonulással keletkező fonéma sorozatokat. Ebben az esetben nem alkalmazható a beágyazott nyelvtani megvalósítás. A hasonulást leíró szabályok összetettsége és nagy száma miatt a hipotézisek bővítése közben (azaz nem előre eldöntve, hanem aktuálisan meghatározva) a hasonulás kezelése nem oldható meg megfelelő sebességgel.

Id

arányosan nő a memóriaigény. Általános esetben ($N > 2$) az N -gramm memóriaigénye exponenciálisan nő az előzmény vektor hosszával ($N-1$). A gyakorlatban a 3-gramm vagy ennél összetettebb nyelvtan esetén már nem lehet előre eltárolni a táblázatot minden egyes fonéma folytatáshoz. Sajnos valós időben sem lehet megkeresni a maximális értékeket a hipotézis generálás közben. Az utolsó fejezetben egy áthidaló megoldást adunk meg erre a problémára.

Teszteredmények a pajzsmirigy adatbázison

Az első tesztben azt vizsgáljuk, hogyan változik a felismerés pontossága MSD kódos nyelvtan használatával. A teszthez véletlenszerűen kiválasztottunk 100 leletet a pajzsmirigy szövegadatbázisból (részletes leírást a *A nyelvi modell tanításához használt adatbázisok* fejezetben találunk). Minden leletet három részletben olvastak be a beszélők. Egy rész több mondatból is állhat, ezeket 5 beszélőtől rögzítettük, minden beszélő 20-20 leletet olvasott be. Az első részbe a „Klinikai adatok”, a „Kérdés” és az „Előző vizsgálat” szakasz került. A második részben a „Jelen vizsgálata”, a harmadik részben pedig az „Összefoglaló vélemény” bemondások találhatók. A szótárakat és a nyelvtanokat mindhárom részre külön-külön is elkészítettük. A teszteket az „Összefoglaló vélemény” részekhez tartozó szótárral, nyelvtannal, hanganyaggal végeztük el (4. táblázat).

4. táblázat

A HMM alapú folyamatos felismerő hatékonysága

Nyelvi modell (1)	correction	accuracy	insertion	deletion	replace
A, szó követ szót (2)	64,18%	52,03%	108	45	273
B, MSD kódos váz (3)	65,20%	63,17%	18	55	254

Table 4: The efficiency of the HMM-based continous speech recognizer on sentences

Language model(1), A: Word by word(2) B: Class N-grams based on MSD codes(3)

A táblázatból látszik, hogy ha kis mértékben is, de az MSD kódos szabályokat tartalmazó nyelvtan jobb eredményt ér el. Részletesebb vizsgálat után arra a megállapításra jutottunk, hogy habár a nyelvtani szabályok segítenek a nem helyes hipotézisek eldobásában, mégis ha rossz szót szúr be a felismerő, mert aktuálisan az tűnik valószínűnek, akkor nincsen esély a mondat helyes befejezésére.

A következő teszt sorozatban azt vizsgáljuk meg, hogy hogyan javít az N -gramm az előző két nyelvtanon. Az $\varepsilon=0$ beállítás egy szigorú vágást ad, mert ha egy szónál a nyelvtan „eltéved”, akkor a hiba kijavítását maga az N -gramm nehezíti meg. Az $\varepsilon=10^{-5}$ beállítás esetén könnyebben helyreállhat a felismerés szótévesztés után (5. táblázat).

A táblázatból kiolvasható, hogy mindkét nyelvtan által elért eredményeken javít a szó- N -gramm. A két nyelvtan eredménye szó-3-gramm és $\varepsilon=10^{-5}$ beállítás mellett közelíti meg egymást a legjobban.

Az alábbi tesztekben azt vizsgáljuk, hogyan módosul a felismerés pontossága, ha a hasonulási szabályokat is felhasználjuk a nyelvtan készítésekor (6. táblázat).

A hasonulási szabályok alkalmazásával az eredmények csak kis mértékben javulnak, mert a hipotézisek számát jelentősen megnövelheti, miközben a vágás mértéke nem változott.

5. táblázat

**A HMM alapú folyamatos felismerő hatékonysága mondatokon
szó N-gram használatakor**

Nyelvi modell (1)	correction	accuracy	insertion	deletion	replace
A + 2 gramm $\varepsilon=0$	79,95%	79,84%	1	151	27
A + 2 gramm $\varepsilon=10^{-5}$	80,51%	75,45%	45	23	150
A + 3 gramm $\varepsilon=0$	80,63%	75,56%	45	27	145
A + 3 gramm $\varepsilon=10^{-5}$	81,64%	76,80%	43	19	144
B + 2 gramm $\varepsilon=0$	80,96%	80,85%	1	63	106
B + 2 gramm $\varepsilon=10^{-5}$	82,54%	82,20%	3	58	97
B + 3 gramm $\varepsilon=0$	86,93%	86,71%	2	30	86
B + 3 gramm $\varepsilon=10^{-5}$	92,11%	92,00%	1	5	65

Table 5: The efficiency of the HMM-based continous speech recognizer on sentences using word N-grams

Language model(1) A and B signs the language model from Table 4.

6. táblázat

**A HMM alapú folyamatos felismerő hatékonysága mondatokon szó N-gram és
hasonulás használatakor**

Nyelvi modell	correction	accuracy	insertion	deletion	replace
A + h	69,03%	65,31%	33	68	207
A + h + 2 gramm $\varepsilon=0$	85,81%	82,32%	31	24	102
A + h + 2 gramm $\varepsilon=10^{-5}$	86,71%	77,81%	79	17	101
A + h + 3 gramm $\varepsilon=0$	87,83%	85,36%	22	24	84
A + h + 3 gramm $\varepsilon=10^{-5}$	88,06%	80,63%	66	17	89
B + h	68,24%	67,00%	11	104	178
B + h + 2 gramm $\varepsilon=0$	91,21%	91,10%	1	58	20
B + h + 2 gramm $\varepsilon=10^{-5}$	93,35%	93,02%	3	41	18
B + h + 3 gramm $\varepsilon=0$	94,14%	94,03%	1	48	4
B + h + 3 gramm $\varepsilon=10^{-5}$	95,15%	95,04%	1	40	3

Table 6: The efficiency of the HMM-based continous speech recognizer on sentences using word N-grams and assimilation rules

Language model (1) A and B denotes the language model from Table 4, here h means that assimilation rules were used.

A 7. táblázat azt szemlélteti, miként változik az egyes nyelvtanokkal a felismerés pontossága, amikor azok tízszer nagyobb leletadatbázist írnak le. A szóanyag és a nyelvtan a fent ismertetett módon készült, de most 1000 leletet választottunk véletlenszerűen úgy, hogy ezek tartalmazták az előbbi 100 leletet is.

7. táblázat

A HMM alapú felismerő hatékonysága mondatokon különböző nyelvi modellek használata esetén

Nyelvi modell (1)	correction	accuracy	insertion	deletion	replace
C (2)	55,96%	48,54%	66	68	323
C + 3 gramm (3) $\varepsilon=0$	71,05%	65,32%	51	128	129
C + h (4) gramm $\varepsilon=10^{-5}$	72,63%	67,34%	47	119	124
C + h + 3 gramm $\varepsilon=10^{-5}$	72,07%	70,83%	11	134	114
D (5)	57,32%	55,97%	12	112	267
D + 3 gramm $\varepsilon=0$	73,76%	68,47%	47	109	124
D + 3 gramm $\varepsilon=10^{-5}$	74,55%	73,65%	8	82	144
D + h + 3 gramm $\varepsilon=10^{-5}$	74,55%	74,10%	4	107	119

Table 7: The efficiency of the HMM-based continous speech recognizer on sentences using different language models

Language model(1) C: denotes the word by word language model(2) D: denotes the model using class N-grams based on MSD codes(3) h means that assimilation rules were used (4) "3 gramm" means that word 3-gram model was used

A várakozásainknak megfelelően a szavak számának jelentős növelése minden esetben rontott a felismerési pontosságon.

A teszteredményeket úgy foglalhatjuk össze, hogy a szó-N-gramm javít leginkább a rendszeren, míg a hasonulás és az MSD kódos szabályok alkalmazása kevésbé, így ezek még további vizsgálatra szorulnak.

GYAKORLATI ALKALMAZHATÓSÁG

A folyamatos beszédfelismerő demonstrálására egy általános, könnyen használható diktáló rendszert fejlesztettünk ki. Az akusztikai és nyelvi modulok egymástól függetlenül módosíthatók, így könnyen újabb modellekre cserélhetők. A rendszer alapjaiban támogatja a különböző beszédstílusú felhasználókat, hasonló hangú beszélők meghatározott csoportján tanított akusztikai modellekből kiindulva a felhasználók profájlokat hozhatnak létre. Ilyen kiinduló csoportok lehetnek például: felnőtt nő, férfi vagy gyerek. Későbbiekben a felhasználóhoz tartozó akusztikus modell adaptáció segítségével még pontosabbá tehető, személyre szabható. A különféle szakszöveg diktálására alkalmas nyelvtanok külön is telepíthetők a rendszerhez. A profájl a felhasználóhoz rendelt, telepített nyelvi modellek személyre szabását, illetve adaptációját is tartalmazza. A tervezett nyelvi adaptáció az N-gramm súlyok módosítását jelenti, illetve az egyszerűbb nyelvtanok esetén új szót is lehet majd felvenni a szótárba.

A diktáló rendszer egy toolbart jelenít meg a képernyő felső részén. Itt könnyen választhatunk az elmentett profájlok között, illetve kiválaszthatjuk azt, hogy hol jelenjen meg a bediktált szöveg írott változata: a diktáló rendszer tartalmaz egy egyszerű szövegszerkesztőt, de lehetőség van a Microsoft Word szövegszerkesztő programjába is diktálni. Ha úgy kívánjuk, akkor tetszőleges aktív alkalmazásban is képes megjeleníteni a szöveget a diktáló rendszer. A feldolgozás tetszőleges időben felfüggeszthető, illetve lehetőség van korábban rögzített hanganyag szöveggé konvertálására is.

EREDMÉNYEK ÖSSZEFOGLALÁSA, TOVÁBBI KUTATÁSI ÉS FEJLESZTÉSI TERVEK

A demonstrációs célokra kifejlesztett diktáló rendszerünk jelenleg kis, illetve közepes méretű szótárakkal hatékonyan képes valós időben működni. Célunk egy nagyszótáros piacképes rendszer előállítása. A felismerő magmodul hatékonyságának növelése érdekében az adaptáció megvalósítása elkerülhetetlen. Emellett, a felismerési pontosság növelése érdekében, tervezzük a különböző gépi tanuló algoritmusok kombinációjával (Felföldi *et al.*, 2002) történő felismerésnek a megvalósítását, illetve a HMM és szegmens alapú módszerek ötvöztetését.

Az MSD kódokon alapuló nyelvtanok a jelenlegi formájukban a tanításhoz használt adatbázisok méretének növelésével olyan memóriaigényűvé válnak, hogy egy átlagos teljesítményű PC-n már nem alkalmazhatók. Az sem elég hatékony módszer, amikor a szabályok szerepét 4, illetve 5 szó-N-grammokkal váltanánk fel. Olyan más – nagyobb nyelvtani szerkezeteket leíró – megoldásokon dolgozunk jelenleg, amelyek kevésbé memóriaigényesek. A szó-N-grammok valószínűségének eltárolásához szükséges memória nagyságrendekkel csökkenthető, ha a szavakat osztályokba soroljuk, és osztály-N-grammot használunk. Ebben az esetben több hipotézist használunk a szükségesnél. Egy lehetséges megoldás az, hogy az osztályokra készítünk egy nagyobb (4-, vagy 5-gramm) és szavakra egy kisebb (2-, vagy 3-gramm) szótár alapú nyelvtant. Így egyben azt is megoldjuk, hogy előrehozottan kiértékelhessük a nyelvtant (lásd 0). Amennyiben garantálható, hogy a csoportok száma kicsi – pl. 16, ami MSD kódok első karaktere esetében igaz –, akkor lehetőség nyílik minden – a nyelvtan által adott – lehetséges fonémához előre megadni a megfelelő valószínűséget. Ezek után a nyelvtannak az osztály-N-gramm része szintaktikai szabályokat, míg a szó-N-gramm része inkább szemantikai szabályokat szolgáltatna. E probléma megoldása több más probléma kivizsgálását is magával hozza, így ez egy későbbi cikkben kerülhet publikálásra.

IRODALOM

- Becchetti, C., Ricotti, L.P. (2000). *Speech Recognition*, John Wiley & Sons LTD, Chichester, England
- C.M. Bishop, (1995). *Neural Networks for Pattern Recognition*, Oxford University Press
- Duda, R.O., Hart, P.E., Stork, D.G. (2001). *Pattern Classification*, Wiley
- Felföldi, L., Kocsor, A., Tóth, L. (2002). Classifier Combination in Speech Recognition, Conference of PhD students on Computer Sciences, Volume of Extended Abstracts, Szeged, Hungary, 30-31.
- Huang, X., Acero, A., Hon, H. (2001). *Spoken Language Processing*, Prentice Hall, New Jersey
- Kocsor, A., Tóth, L., Kuba Jr., A., Kovács, K., Jelasity, M., Gyimóthy, T., Csirik, J. (2000a). A Comparative Study of Several Feature Space Transformation and Learning Methods for Phoneme Classification, *International Journal of Speech Technology*, 3. 3/4. 263-276
- Kocsor, A., Kuba, A., Tóth, L. (2000b). Phoneme Classification Using Kernel Principal Component Analysis, *Periodica Polytechnica*, 44. 1. 77-90.
- Kocsor, A., Kuba, A., Tóth, L., Jelasity, M., Felföldi, L., Gyimóthy, T., Csirik, J. (1999). A Segment-Based Statistical Speech Recognition System for Isolated/Continuous Number Recognition, *Proceedings of the FUSST'99*, Aug. 19-21, Sagadi, Estonia, 201-211.

- Crochemore, M., Ryller, W. (1994). Text Algorithms, Oxford University Press, Oxford
- Nyers, Á., (2004). Beszédfelismerés az orvosi dokumentáció korszerűsítésére, IME (Informatika és Menedzsment az Egészségügyben) 3. 5. 39-43.
- Moore, B.C.J. (1997). An Introduction to the Psychology of Hearing, Academic Press
- Rabiner, L.R., Juang, B.H. (1993). Fundamentals of Speech Recognition, Prentice-Hall, Englewood
- Rabiner, L.R., Schafer, R.W. (1978). Digital Processing of Speech Signals, Prentice-Hall, Englewood
- Smith, J., IBM, (2002). ViaVoice and Dragon Naturally Speaking XP, ANWALT S.32
- Tóth, L., Kocsor, A., Kovács, K. (2000). A Discriminative Segmental Speech Model and its Application to Hungarian Number Recognition, Springer Verlag, TSD'2000, 307-313.
- Vapnik, V.N. (1998). Statistical Learning Theory, Wiley

Levelezési cím (*Corresponding author*):

Kocsor András

MTA-SZTE Mesterséges Intelligencia Tanszéki Kutatócsoport

6720 Szeged, Aradi vértanúk tere 1.

Research Group on Artificial Intelligence of the address: Hungarian Academy of Sciences and University of Szeged

H-6720 Szeged, Aradi vértanúk tere 1.

Tel.: 36-62-544-126, Fax: 36-62-425-508

e-mail: kocsor@inf.u-szeged.hu



A Pécsi Tudományegyetem Botanikus Kertjének térinformatikai rendszere

Bornemisza I.

Pécsi Tudományegyetem, Informatika és Általános Technika Tanszék, 7624 Pécs, Ifjúság u. 6.

ÖSSZEFOGLALÁS

A Pécsi Tudományegyetem Informatika és Általános Technika Tanszékén, a számítástechnika-technika szakos tanárképzés keretén belül évek óta folyik a térinformatika tantárgy oktatása. A gyakorlati képzésre kiválasztott feladat célja kettős: a hallgatók életközeli feladattal szereznek tapasztalatot és továbbadható tudást, a leírt projekt végeredményeként kialakított térinformatikai rendszer pedig a Botanikus Kert munkáját segíti majd.
(Kulcsszavak: térinformatika oktatása, botanikus kert)

ABSTRACT

The geoinformational system of the Botanical Garden at the University of Pécs

I. Bornemisza

University of Pécs, Department of Computer Science and General Technics, H-7624 Pécs, Ifjúság u. 6.

At the Department of Computer Science and General Technics of the University of Pécs geoinformatics has been taught for years to students of computer science and technics. The task chosen for practical training has two goals: on the one hand, the students experience lifelike situations and gain knowledge to pass on, further more, the result of the accomplished project, i.e. the geoinformatics system will help the work of the botanical garden.

(Keywords: education of geoinformatics, botanical garden)

BEVEZETÉS

A PTE TTK Informatika és Általános Technika Tanszékén, a számítástechnika-technika szakos tanár hallgatók részére több év óta folyik a térinformatika tantárgy oktatása. A térinformatikát az érdeklődő hallgatók választható, kiegészítő tantárgyként vehetik fel. Tekintettel arra, hogy a számítástechnika-technika szakos tanár hallgatónak sok témához kell értenie, nem várható el, hogy minden területeken a legképzettebb szakember tudásszintjét birtokolja – általános ismereteire elsősorban azért van szüksége, hogy leendő diákjainak továbbadhassa. A gyakorlati képzés keretében olyan témát kell választani, amely életszerű, de belefér a megadott óraszámba. Nem kizárt, hogy általános iskolában jobb képességű gyerekekkel térinformatikai feladatok megoldhatók legyenek, de reálisan kell az oktatott elméleti ismeretek mennyiségét és a gyakorlati feladatok nehézségi fokát megválasztani (Kapunári, 1999).

Az elmúlt években egyetemünkön több tanszék is részt vett abban a kutatásban, amely a Dél-Dunántúli régió környezetvizsgálatát célozta (Béres et al., 2002). A projekt célja egy gazdaságos rekultivációs technológia kifejlesztése, annak demonstrációs alkalmazása, valamint komplex monitoring rendszerének kidolgozása volt. A kutatás

során a társtanszékek részéről felmerült adatmegjelenítési igényeket, a mérési adatok ábrázolását ArcView térinformatikai rendszerrel elégitettük ki. A munka során sok olyan részfeladat merült fel, amelynek megoldási módját célszerűnek látszott a hallgatókkal is megismertetni. Ebben a kutatásban az érintett cégek sok esetben nem publikus adataival dolgoztunk, emiatt oktatási célra egy másik területet kellett választani. A kiválasztott téma szakmailag nem túl bonyolult, az alapok oktatása céljára azonban megfelelő, tekinthetjük iskolapéldának.

ANYAG ÉS MÓDSZER

Mivel tanár szakos hallgatóink a későbbi tanítványaik kedvéért tanulnak, célszerű a témát úgy bevezetni, hogy azt majd egy tizenéves is megértse, sőt, az érdeklődését is felkeltse. Érdemes tehát a tematikát a megjelenítésre, a látványra „kihegyezni”. Megjegyzendő, hogy szerencsés, ha a tanulás keretében elkészített rendszernek gyakorlati haszna is van, de nem kell feltétlenül a valós használhatóságra törekedni, hiszen az elsődleges cél az oktatás. Szemléletes példaként megemlíthető, hogy tanszékünkön egy kutatócsoport kísérleti gyakorló űrszonda építésével, fejlesztésével foglalkozik, szintén pedagógiai céllal, s természetesen nem tervezik az űrszonda kilövését.

A 2004/2005-ös tanév őszi félévében megkezdett, várhatóan 2-3 éves kifutású hallgatói projekt kitűzött célja: az egyetem területén működő Botanikus Kert digitális térképének előállítás, majd a növények nyilvántartási adatainak adatbázisba konvertálása, végül ezen adatbázis összekapcsolása a térképpel.

A szemeszter első felében az elméleti alapokkal ismerkednek meg a hallgatók. Az oktatás tematikájának összeállításakor Dr. Katona Endre interneten elérhető jegyzetét (Katona, 2003) vettem alapul:

- Térinformatikai alapfogalmak
- Térinformatikai szoftverek. Raszteres és vektoros adatábrázolás
- Speciális hardver és szoftver eszközök
- Térképészeti alapok. Térképtípusok, vetületi rendszerek, felmérés, térképkészítés
- Műholdas helymeghatározás (GPS)

A gyakorlati feladat során a műholdas helymeghatározás, valamint az adatbázis és az ArcView együttes kezelésének megismerése volt a kitűzött cél. A megoldás első szakaszának lépései tehát:

- A kiválasztott terület jellemző pontjai koordinátáinak meghatározása (GPS)
- Mérési adatok bevitele, kezelése adatbázisban (MS-Access)
- A felmért terület pontjainak ábrázolása adatbázisból (ArcView)

A körülbelül 300×400 m-es terület felmérése GPS (Garmin GPS 72) műszerrel történt, amely önálló eszközként használva alig éri el az 10-15 m-es pontosságot. Az adatokat a könnyebb szerkeszthetőség és az egyes adatfeldolgozó eszközök együttműködésének gyakorlása érdekében MS-Excel táblázatba vittük, majd ezt importáltuk MS-Access-be (1. ábra).

Mivel egy működő rendszernél ennél nagyobb pontosságra van szükség, a GPS-t csak a mérés technikájának oktatására használjuk, s a térképi ábrázolásra a Földhivatal által készített térképet alkalmazzuk, így a gyakorlati feladat második szakaszában a térkép szkennelése, illesztése, digitalizálása is előkerül. A digitalizáláshoz OCAD szoftvert használtunk (2. ábra). A kész térképet DXF formátumban mentettük, ez megfelelő csereformátumnak bizonyult az OCAD és ArcView között

1. ábra

A műholdas helymeghatározás eszköze és eredménye

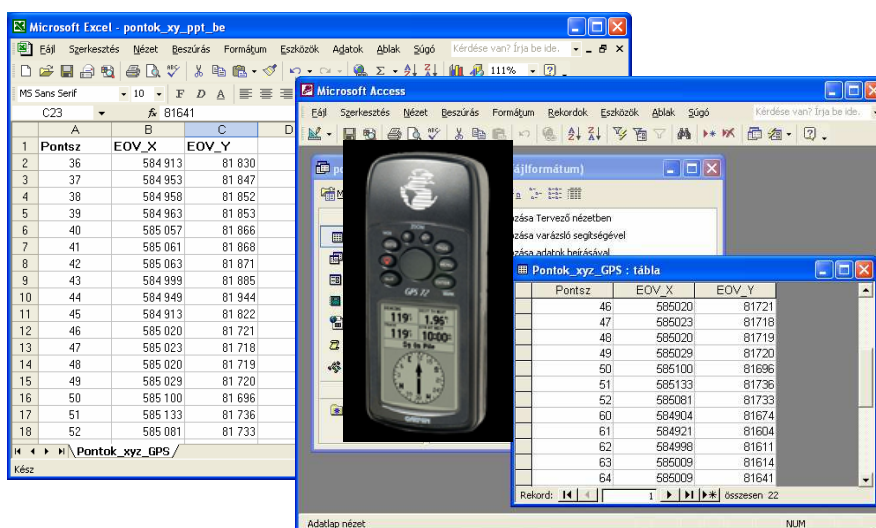


Figure 1: The tool and result of the global positioning

2. ábra

Digitalizálás OCAD térképrajzoló szoftverrel

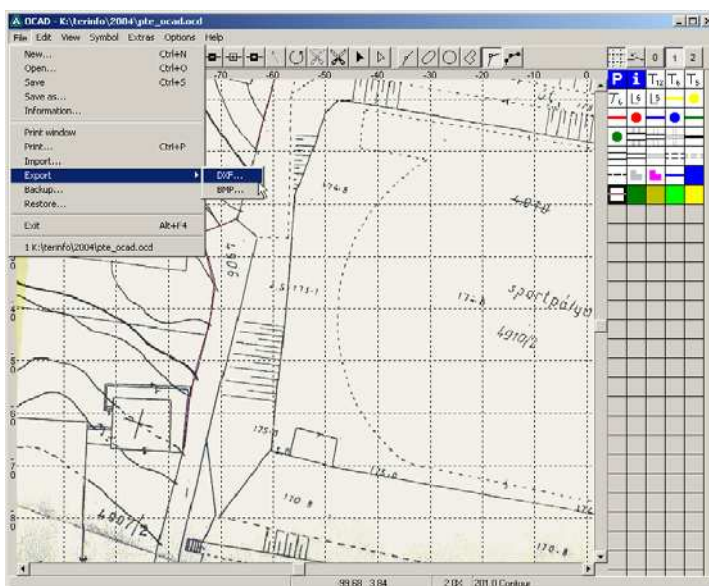


Figure 2. Digitizing with OCAD drawing program for cartography

A rendszer elkészítésének végső szakaszában az egyes növények koordinátáit kell majd meghatározni, s a többi, jelenleg papíron elérhető adattal együtt (azonosító szám, magyar és latin megnevezés, méretek) adatbázisba bevinni.

EREDMÉNY ÉS ÉRTÉKELÉS

Első lépésben eredménynek könyvelhető el a hallgatói érdeklődés, mely abban is megnyilvánult, hogy több hallgató szakdolgozati témának választotta a térinformatikát. Feltételezhető, hogy ők tanítványaiknak is szívesen oktatják majd informatika óra vagy szakkör keretén belül a megismert módszereket.

Szerencsés együttműködésre van kilátás a Növényteni Tanszékkel. A tanszék egyik hallgatója TDK-munka keretében készített egy internetes adatbázist a botanikus kertek maglistáinak (*Index Seminum*) nyilvántartásához, a magcsere támogatása céljára (Darabos, 2005). Fogalmazhatnánk divatos kifejezéssel úgy is, hogy a két projekt összefogásával folyamatban van egy *eBotanikusKert* kialakítása.

KÖVETKEZTETÉSEK

A leírt projekt mintegy húsz másod- és harmadéves hallgató egy szemeszteres tevékenységét mutatja. Nem tekinthető magas hatásfokúnak a tevékenység még akkor sem, ha tudjuk, hogy két csoportban, párhuzamosan végezték ugyanazt a munkát. Egy néhány fős, szakértő csapat feltehetőleg néhány hét alatt megoldaná a teljes feladatot, de itt az elmélet és az eszközkezelés megismerése volt a valódi feladat. A további félévekben más hallgatók veszik fel ugyanezt a tantárgyat, így a tevékenység nagy részét újra kell tanulni, s csak a szemeszter második szakaszában használhatjuk fel a korábban elért eredményeket. Reményeink szerint ennek ellenére 2-3 éven belül működőképes rendszert adhatunk át a Botanikus Kertnek, s az sem elhanyagolható siker, hogy közben több tucat hallgató szerzett némi tapasztalatot a térinformatika területén.

IRODALOM

- Béres Cs.Z., Bornemisza I., Csapó L., Gimesi L., Markó T. (2002). Összefoglaló az NKFP-3-050/2001 számú, „A dél-dunántúli régió környezetterhelésének csökkentésére irányuló komplex hulladékkezelési és rekultivációs technológia, valamint monitoring rendszer kifejlesztése és alkalmazása” című projekt 2.8 alprojekt informatikai rendszeréről. Pécsi Tudományegyetem, Pécs
- Darabos P. (2005). BotLine - Internetes adatbázis botanikus kerteknek, TDK dolgozat, PTE Növényteni Tanszék, Pécs
- Kapuvári B. (1999). ATOM-Program: Alapfokú Térinformatikai Oktatási Módszertan. Térinformatika oktatása az általános iskolában. InCo: első magyar internetes folyóirat az információs korról 99/2. <http://www.inco.hu/inco2/infoert/cikk5.htm>
- Katona E. (2003). Térinformatika. Előadási jegyzet, Szegedi Tudományegyetem, Alkalmazott Informatikai Tanszék, Szeged

Levelezési cím (*Corresponding author*):

Bornemisza Imre

PTE, TTK, Informatika és Általános Technika Tanszék

7624 Pécs, Ifjúság útja 6.

University of Pécs, Faculty of Natural Sciences

Department of Computer Science and General Technics

H-7401 Pécs, Ifjúság u. 6.

Tel.: 36-72-503-697, Fax: 36-72-503-697

e-mail: born@gamma.ttk.pte.hu



Milyen kompetenciákkal rendelkezzen egy „térinformatikus”?

Honfi V., Micsinai R. P., Barabás T.

Kaposvári Egyetem, Gazdaságtudományi Kar, Informatika Tanszék, Kaposvár, 7400 Guba Sándor út 40.

ÖSSZEFOGLALÁS

A Kaposvári Egyetemen 2001. szeptembere óta folyik térinformatika oktatás, elsősorban gazdasági mérnök szakos hallgatóknak. Cél, hogy a végzős hallgatókat olyan használható ismeretekkel ruházzuk fel, melyek segítségével könnyen megállják helyüket a gyakorlati életben, ezáltal biztosítva intézményünk jó hírnevét. Ennek az eléréséhez elengedhetetlen, hogy folyamatosan ellenőrizzük oktatásunk minőségét, korszerűségét és alkalmazkodjunk a piaci igényekhez. Ezért végeztünk kutatást, igényfelmérést, hogy hallgatóink lehetséges későbbi munkaadói milyen ismereteket, kompetenciákat várnak el a hozzájuk felvételiző munkavállalótól. A kérdőívek összesítése után arra jutottunk, hogy szorosabb együttműködés kialakítása szükséges a térinformatikával foglalkozó vállalkozásokkal, az oktatás gyakorlatiasabbá tétele érdekében.

(Kulcsszavak: térinformatika, készségek, képességek, elvárások)

ABSTRACT

What Skills does a GIS Analyst require?

V. Honfi, R. P. Micsinai, T. Barabás

University of Kaposvár, Faculty of Economic Science, Department of Information Technologie, Kaposvár, H-7400 Guba S. u. 40.

In the University of Kaposvár GIS educating is on since September 2001, primarily for students specialized as economic engineers. The object is to give the graduating students such usable knowledge, that can help them to stand their ground easily in practical life, hereby guaranteeing our institutions good reputation. To reach this it is indispensable to check the quality, modernity of our education permanently and to adapt to the demands of the market. Therefore we have made research, survey of demands with our students possible employers, to find out what knowledge, skills they expect from the employees applying to them. After totalizing the results we got clear image on which points of the education is the change necessary.

(Keywords: GIS, skills, ability, expectations)

BEVEZETÉS

A technikai, technológiai fejlődés megköveteli az egyetemi oktatóktól is a folyamatos tájékozottságot, naprakészséget, melynek természetes következménye a tananyagfejlesztés. Eközben azonban nem szabad elfeledkezni arról, hogy a felsőoktatási intézmények is a piacról, az oktatási piacról élnek, így szükségszerű, hogy a végzős hallgatók olyan használható ismereteket kapjanak, melyek segítségével könnyen megállják helyüket a gyakorlati életben.

Ennek a célnak az eléréséhez elengedhetetlen az oktatás minőségének folyamatos ellenőrzése, hogy biztosítható legyen a korszerűség és alkalmazkodjon a piaci igényekhez. Ezért végeztünk egy kutatást, igényfelmérést, hogy hallgatóink lehetséges későbbi munkaadói milyen ismereteket, kompetenciákat várnak el a hozzájuk jelentkező munkavállalótól. Az eredmények összesítése után világos képet kaphatunk arról, hogy az oktatás mely pontjain szükségesek változtatások.

ANYAG ÉS MÓDSZER

A vizsgálat elvégzésekor elsősorban arra a kérdésre kerestük a választ, hogy a jelenleg a Kaposvári Egyetemen folyó oktatás mennyire felel meg a munkáltatói igényeknek. Abból a feltevésből indultunk ki, hogy nem elég csak a szakmai ismeretekre koncentrálni, mivel a térinformatikus több szakma tevékenységét is integrálja a munkájában, ezért az emberi tényezők is nagy súllyal esnek latba (Latham, 2001). A munkáltatói igények megismeréséhez legalkalmasabbnak a zárt kérdőívek alkalmazását találtuk, mivel, mindenképpen primer adatgyűjtésre volt szükség. Az egyéb megkérdezési módszerek nem, vagy csak nehezebben biztosították volna a kitűzött cél teljesülését.

A kérdőíveken megfogalmazott kérdések négy csoportba sorolhatóak. A kérdések egyik csoportja arra kereste a választ, hogy a szakmai ismeretek közül melyik részterületet, mennyire tartják fontosnak az egyes munkáltatók. Ezekről vártuk a térinformatika tantárgy oktatásához kidolgozott tematika megerősítését, illetve azt, hogy a hiányosságokra rámutassanak.

A második csoportba soroljuk azokat a kérdéseket, melyek arra ösztönözték a válaszadókat, hogy súlyozzák, illetve jelezzék azokat a képességeket, melyekkel szerintük rendelkeznie kell egy térinformatikával foglalkozó munkavállalónak.

A kérdések harmadik csoportja azokat az emberi tulajdonságokat igyekezett összegyűjteni, melyek fontosak lehetnek egy munkáltató szempontjából a munkaerő kiválasztás során. A vizsgált tulajdonságok munkaerő közvetítő és fejvadász cégek álláshirdetési alapján kerültek kiválasztásra. A tulajdonságok kijelölése előtt 200 az informatika legkülönbözőbb területeire vonatkozó álláshirdetést vizsgáltunk meg, és választottuk ki az öt leggyakrabban előforduló elvárt tulajdonságot. Sem a cégek, sem pedig az álláshirdetések kiválasztása nem tekinthető reprezentatív mintavételnek, de ez a vizsgálat célja szempontjából irreleváns.

A negyedik kérdéscsoport kiegészítő információkat igyekezett összegyűjteni egyrészt a válaszadókról, másrészt arról, hogy milyen változtatásokkal, finomításokkal tudná az egyetem hatékonyabban segíteni az itt végzett hallgatókat a munkaerőpiacon. A válaszadókra vonatkozó információk későbbi, nagyobb mintán elvégzett vizsgálatoknál lehetővé teszik majd számunkra, hogy a működési terület illetve a vállalkozás nagysága szerint mélyebb összefüggéseket keressünk, differenciáljuk, illetve súlyozzuk az elvárt ismereteket.

Az ismeretekre, készségekre és képességekre vonatkozóan feltett 7 alkérdés közül egy kivételével, ötös skálán kellett jelölniük a válaszadóknak, hogy az adott ismeretet mennyire tartják fontosnak, illetve a vizsgált kijelentéssel mennyire értenek egyet. A hetedik kérdés, mint egyetlen nyílt kérdés biztosított lehetőséget arra, hogy a válaszadó kiegészítse a kérdéseket azokkal a tudományterületekkel, melyeknek az ismeretét még fontosnak tartaná, de valamely okból nem tértünk ki azok vizsgálatára.

A kiértékelés során, minden esetben kiszámításra került az abszolút pontszám, amelyet a válaszokhoz rendelt pont értékek és a válaszadók számának szorzata határoz meg. Az abszolút pontszámok alapján kiszámoltuk az egyes kérdéscsoportokon belül az átlagos pontszámokat, és ennek alapján került meghatározásra a relatív fontosság, ahol 100%-nak a

kérdéscsoporton belüli átlagos pontszámot vettük. Erre azért került sor, mert az abszolút pontszámok helyett a relatív fontosság, jelen vizsgálatban jobban hasznosítható, értelmezhető információval szolgál. Ennek alapján került kialakításra a sorrend, melyből kiderül, hogy melyik elemet, elemeket tartották a válaszadók a legfontosabbnak egy-egy kérdéscsoporton belül. Minden egyes tulajdonságra meghatározásra került az is, hogy a válaszadók száma hogyan oszlik meg a fontosság megítélésében.

Az első adatgyűjtés lezárásáig, 2005. április 30-ig érkezett válaszokat dolgoztuk fel. Ez a minta semmiképpen nem tekinthető reprezentatívnak, de arra elegendő, hogy első körben az elképzelések helyességét igazolja, vagy azok átgondolására ösztönözzön. Az adatgyűjtés és a kérdőívek kitöltése elsősorban Somogy megyére illetve a Dél-dunántúli régióra koncentrált, de a későbbiekben folytatni kívánjuk országos szinten is.

EREDMÉNY ÉS ÉRTÉKELÉS

Az első kérdéscsoportban arra kerestük a választ, hogy az oktatáshoz előzőleg összeállított tematika szerint, az egyes részterületeket mennyire tartják fontosnak a válaszadók. Ez alapján 7 szakmai ismeret csoportot kellett véleményezniük, úgymint: adatbázis tervezési ismeretek, információ technológiai jártasság, térképészeti ismeretek, relációs adatbázisok működési elvei, matematikai és statisztikai ismeretek, legalább 2 szoftver ismerete, programozási ismeretek. Az 1. táblázat tartalmazza a válaszadók megoszlását.

1. táblázat

Szakmai képességek fontosságának megítélése

	Különösen fontos (1)	Nagyon fontos (2)	Fontos (3)	Nem fontos (4)	Egyáltalán nem fontos (5)
Adatbázis tervezés (6)	23%	62%	15%	0%	0%
IT jártasság (7)	31%	38%	15%	15%	0%
Térképészeti ismeretek (8)	15%	46%	38%	0%	0%
Relációs adatbázisok (9)	23%	46%	15%	15%	0%
Matematika és statisztika (10)	8%	54%	23%	15%	0%
2 szoftver ismerete (11)	8%	46%	31%	15%	0%
Programozási ismeretek (12)	8%	31%	23%	23%	15%

Table 1: Importance of professional abilities

Especialy important(1), Very important(2), Important(3), Not important(4), Not in the least(5), Database designing(6), IT experting(7), Tophological knowledge(8), Relational databases(9), Mathematics and statistics(10), Cognition of 2 softwares(11), Programming(12)

A szakmai ismeretek közül egyértelműen fontosnak ítélték a munkaadók az adatbázis tervezési, térképészeti ismereteket és fontosnak tartják az információ-technológiai jártasságot, bár ennél már jobban megoszlak a válaszadók véleménye. Viszonylag kicsi különbség mutatkozik a relatív pontszámokban, de a programozási ismeretek tekintetében az a tény, hogy a válaszadók 38%-a nem tartja fontosnak, már figyelembe veendő. Ennek valószínűleg az az oka, hogy a válaszadók legtöbbször nem saját fejlesztésű, hanem készen vásárolt szoftver üzemel, így az elemzések elvégzéséhez is inkább felhasználóra, mint fejlesztőre van szükségük. Az 1. ábra mutatja, hogy az egyes ismereteket a munkaadók mennyire ítélik fontosnak. Az ábrán 100% az összes válaszadó alapján számított átlagos pontszám.

1. ábra

Szakmai képességek fontossága

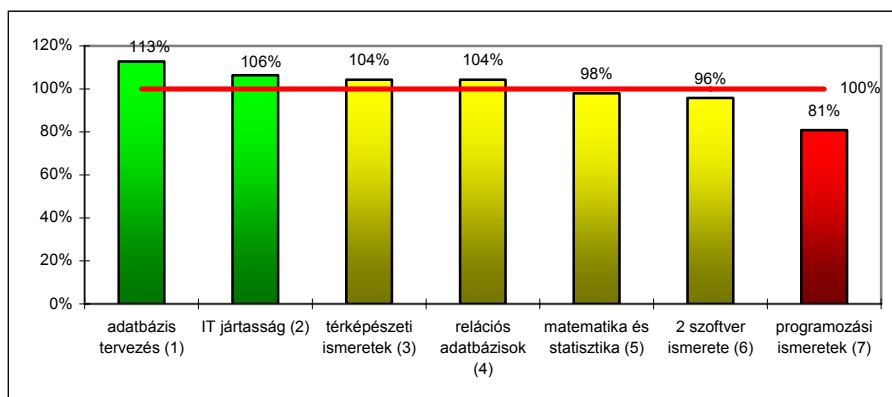


Figure 1. Importance of professional abilities

Database designing(1), IT experting(2), topographical knowledge(3), relational databases(4), mathematics and statistics(5), cognition of 2 softwares(6), programming(7)

A munkavállalói képességek tekintetében arra kértük a foglalkoztatókat, jelöljék meg, hogy a rendszerben gondolkodás képességét, a kitűnő írásbeli és szóbeli képességeket, a jó elemző és problémamegoldó, valamint a jó kapcsolatteremtő és a produktív munkavégzésre való képességet mennyire tartják fontosnak egy leendő munkatárs, beosztott esetén. A számított relatív fontosságok nagyon közel esnek egymáshoz, így nem jelenthetjük ki egyértelműen, hogy egyik vagy másik képesség fontosabb, vagy kevésbé fontos, mint a többi. Ha megnézzük a válaszok megoszlását, akkor viszont már felállítható egyfajta sorrend az egyes képességek fontosságának megítélése alapján. (2. táblázat.)

Ez alapján kijelenthető, hogy a széles látókörűséget, a rendszerben gondolkodást és a jó elemző és problémamegoldó képességet a válaszadók túlnyomó többsége különösen fontosnak tartja (2. ábra). Meglepő lehet, hogy a második helyre sorolt rendszerben gondolkodás esetén a válaszadók 15%-a ezt a képességet nem tartja fontosnak. Ha jobban megvizsgáljuk ezt a kérdést, rögtön kiderül, hogy ők jellemzően csupán egy konkrét feladathoz, mások által elkészített elemzések megjelenítéséhez használják a térinformatikát, és azt szolgáltatásként veszik igénybe.

A vizsgált képességek közül a kitűnő írásbeli és szóbeli képességek megléte, illetve a jó kapcsolatteremtő képesség elsősorban a szocializáció során alakul ki, melynek megléte az egyetemi tanulmányok során is előnyös, de külön fejlesztése nem feladata az oktatásunknak. A széles látókörűség, rendszerben gondolkodás, a jó elemző és problémamegoldó képesség és a produktív munkavégzésre való alkalmasság jól átgondolt oktatás-szervezéssel, megfelelő feladatok kiválasztásával és szemléletalkotással fejleszthető.

Kíváncsiak voltunk arra is, hogy a foglalkoztatók mennyire tartanak fontosnak egyes emberi tulajdonságokat. A vizsgált öt tulajdonság a kreativitás, a pozitív kisugárzás, magabiztosság, új dolgokra való fogékonyság, valamint a vállalkozó típus (konkrét tartalommal nehezen felruházható) tulajdonságok voltak. A felmérés eredményét a 3. táblázat foglalja össze.

2. táblázat

Egyéb képességek fontossága

	Különösen fontos (1)	Nagyon fontos (2)	Fontos (3)	Nem fontos (4)	Egyáltalán nem fontos (5)	Relatív fontosság (6)
Széles látókör (7)	54%	38%	8%	0%	0%	104%
Rendszerben gondolkodik (8)	62%	23%	0%	15%	0%	101%
Jó elemző és problémamegoldó (9)	62%	8%	31%	0%	0%	101%
Termékeny munkavégzés (10)	38%	54%	8%	0%	0%	101%
Jó kapcsolatteremtő (11)	31%	62%	8%	0%	0%	99%
Kitűnő írásbeli és szóbeli képesség (12)	31%	54%	8%	8%	0%	95%

Table 2: Importance of talents

Especialmente importante(1), Muy importante(2), Importante(3), No importante(4), No en lo menos(5), Relativa importancia(6), Amplio horizonte intelectual(7), Pensamiento en sistema(8), buen analista y solucionador de problemas(9), Productivo en el trabajo(10), Buen creador de relaciones(11), Excelente habilidad verbal y de escritura(12)

2. ábra

A válaszadók megoszlása a képességek fontossága alapján

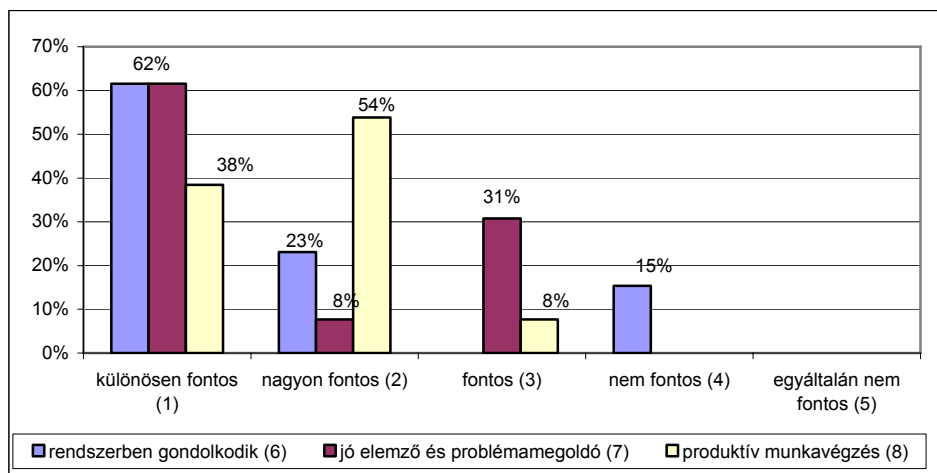


Figure 2. Importance of abilities at respondents discretion

Especialmente importante(1), Muy importante(2), Importante(3), No importante(4), No en lo menos(5), Pensamiento en sistema(6), Buen analista y solucionador de problemas(7), Productivo en el trabajo(8)

3. táblázat

Emberi tulajdonságok fontossága

	Különösen fontos (1)	Nagyon fontos (2)	Fontos (3)	Nem fontos (4)	Egyáltalán nem fontos (5)	Relatív fontosság (6)
Kreatív (7)	38%	46%	15%	0%	0%	108%
Vállalkozó típus (8)	31%	62%	8%	0%	0%	108%
Új dolgokra fogékony (9)	23%	54%	23%	0%	0%	102%
Pozitív kisugárzás (10)	23%	38%	31%	8%	0%	96%
Magabiztos (11)	8%	31%	46%	15%	0%	85%

Table 3: Importance of human properties

Especialy important(1), Very important(2), Important(3), Not important(4), Not in the least(5), Relative importance(6), Creativity(7), Adventurous(8), Feeling for new objects(9), Positive attitude(10), Confidence(11)

Érdekes és elgondolkodtató eredményként könyvelhetjük el, hogy bár a munkaadók nagy hányada (93%-a) nagyon, vagy különösen fontosnak tartja, hogy a leendő munkatárs vállalkozó típus legyen, addig a magabiztosságot mindössze 39%-uk sorolja ugyanebbe a kategóriába (3. ábra).

3. ábra

Emberi tulajdonságok fontossága

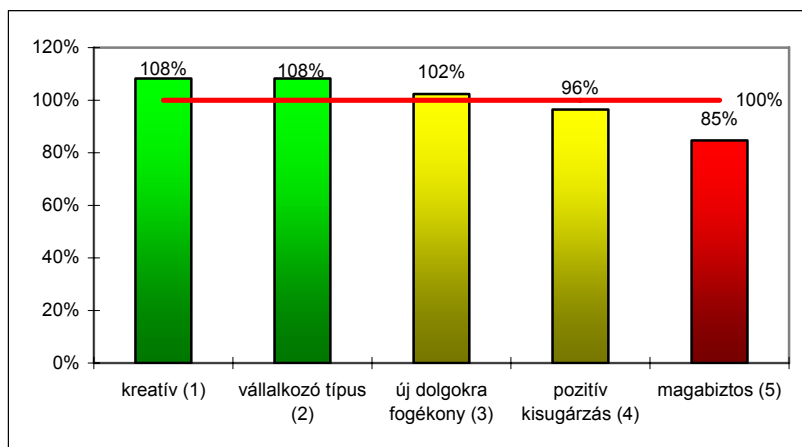


Figure 3. Importance of human properties

Creativity(7), Adventurous(8), Feeling for new objects(9), Positive attitude(10), Confidence(11)

A fenti tulajdonságok közül a kreativitás bármelyik tantárgy keretén belül fejleszthető, az új dolgokra való fogékonyságra a hallgatók motiválhatóak. Az, hogy valaki vállalkozó típus részben alkat kérdése, de az biztos, hogy a hallgató tanulmányi munkája

során elért kisebb-nagyobb sikerek elősegíthetik, hogy a hallgató magabiztosan, pozitív kisugárzással lépjen a munkaerőpiacra.

Érdekes eredményt mutattak a negyedik kérdéscsoportra adott válaszok, ahol arra voltunk kíváncsiak, milyen elvárásoknak feleljen meg, milyen egyéb képességekkel, tulajdonságokkal, dokumentumokkal rendelkezzen egy, az egyetemről kikerülő, frissdiplomás munkavállaló.

A legtöbb válaszadó szerint az ideális jelölt esetében jó, ha már kipróbált egy-két dolgot, de nem kell, hogy mindenhez értsen. Ezek a jellemzőket az átlagosnál fontosabbnak tartják a munkáltatók. Az átlaggal esik egybe két elvárás, a legalább egy idegen nyelv tárgyalási szintű ismerete, valamint a hivatalosan is dokumentált térinformatikai végzettség. Ezeknél valamivel kevésbé tartják fontosnak a munkáltatók az igazolt szakmai gyakorlatot, valamint az igazolt részvételt korábbi projekteken.

KÖVETKEZTETÉSEK

A válaszok alapján kijelenthető, hogy helyesnek bizonyult az elképzelés, miszerint az elméleti képzésben hangsúlyosabban kell oktatni a térképészeti ismereteket. Az adatbázis kezelés külön tantárgyként is szerepel az érintett hallgatók tantervében, amely ismereteket az elméleti és gyakorlati órákon konkrét térinformatikai problémák megoldása során elmélyítünk. A gyakorlatokon igyekszünk bemutatni legalább 2 térinformatikai szoftver használatát, alapvető szolgáltatásait. Ez helyes elképzelésnek bizonyult, bár az időkeret szűk a teljes és rendszerszerű áttekintéshez.

Szintén külön tantárgy keretében szerzik meg információtechnológiai és matematikai, statisztikai ismereteiket a hallgatók, melyek elmélyítésére a jelenlegi időkeret nem biztosít lehetőséget.

A jelenlegi rendszerben a térinformatika tantárgy oktatása különböző szakokon, különböző évfolyamokon történik. Így a fejleszthető képességek tekintetében újra átgondolandó és megvizsgálandó egyes tantárgyak egymásra épülésének logikája, hogy kellőképpen biztosítják-e a rendszerszemlélet, széles látókörűség kialakulásának lehetőségét.

Tapasztalataink alapján a hallgatók kreativitása jól fejleszthető, ha gyakorlaton olyan feladatokat kapnak, melyek önmagukban is érdekesek, és gyakorlati hasznosulásuk is legalább részben biztosított. A feladatok megoldásához a technikai környezetet és az oktató tutori szerepkörben való jelenlétét célszerű biztosítani, a felmerülő problémákat konzultációs jelleggel megbeszélni. Mindenképpen szükségesnek látszik, hogy a jelenleginél jóval több gyakorlati hellyel célszerű felvenni a kapcsolatot annak érdekében, hogy a hallgatók őket érdeklő, valós problémák megoldása során érjenek el sikereket.

Annak érdekében pedig, hogy hallgatóink jobban megfeleljenek a munkáltatói elvárásoknak, a közeljövőben meg kell keresni a módját, hogyan lehet a hallgatók térinformatikai ismereteit valamilyen dokumentummal is igazolhatóvá tenni.

Jó megoldást jelentene, ha az egyetem tudásbázisát felhasználva, piaci körülmények között működő vállalkozásokkal szorosabban együttműködve, valós problémák megoldása során alkalmazhatnák és megszerezhetnék a hallgatók a megszerzett elméleti ismereteket.

IRODALOM

G. Letham (2001). What Skills does a GIS Analyst require? <http://careers.geocomm.com>

Levelezési cím (*Corresponding author*):

Honfi Vid

Kaposvári Egyetem, Gazdaságtudományi Kar

Informatika Tanszék

7401 Kaposvár, Pf. 16.

University of Kaposvár, Faculty of Economic Science

Department of Information Technologie

H-7401 Kaposvár, P.O.Box 16.

Tel.: 36-82-314-155

e-mail: honfi@mail.atk.u-kaposvar.hu



Hiszterézises hődiffúziós probléma megoldása multigrid módszerrel

Jancskár I., Iványi A.

University of Pécs, Department of Information Technology, H-7624 Pécs, Rókus u. 2.

ÖSSZEFOGLALÁS

Tranziens hővezetési problémák numerikus megoldása véges differenciák módszerével, hőmérséklettől függő diffuzivitást feltételezve nagyobb mértékű felbontás esetében igen jelentős memória és CPU idő felhasználással jár. Implicit diszkretizálási sémát alkalmazva a diffuzivitás hőmérséklet függése miatt a hőmérsékletmező meghatározása minden egyes időlépésben iterációt igényel. Kihasználva a multigrid technika adta lehetőségeket, olyan hierarchikus algoritmust mutatunk be, mellyel a futási idő jelentősen lerövidíthető. Összehasonlító vizsgálatokkal elemezzük a javasolt módszer alkalmazhatósági feltételeit. Bemutatunk néhány példát a nemlineáris diffúzióra különböző kezdeti és peremfeltételek esetében.

(Kulcsszavak: nemlineáris hődiffúzió, multigrid technika, hiszterézis)

ABSTRACT

Multigrid hysteretic heat diffusion

I. Jancskár, A. Iványi

University of Pécs, Department of Information Technology, H-7624 Pécs, Rókus u. 2.

In this paper a transient non-linear heat diffusion problem is studied. Considering hysteretic type temperature-dependence of thermal diffusion, the numerical solution is very resource- and time-consuming. Utilizing the advantages of the multigrid technique a coarse level iterating algorithm is introduced. Applying a hierarchical algorithm the solution time could be reduced. The proposed method has been verified by comparing to other numerical schemes. The thermal heating-cooling asymmetry resulted by the hysteresis and the memory character of diffusivity are proved by numerical simulations.

(Keywords: non-linear heat diffusion, multigrid technique, hysteresis)

BEVEZETÉS

Az utóbbi időkben egyre több tudományos munka foglalkozik változó termikus tulajdonságokkal rendelkező kompozit anyagokban lejátszódó termikus folyamatok analizálásával. Számos kísérleti mérés igazolja, hogy bizonyos összetett anyagok hővezetési tényezője illetve diffuzivitása nagyobb hőmérsékleti skálát tekintve nem tekinthető konstans értékűnek, sőt egyes esetekben hiszterézises jelleget mutat. Például Santos at al. (2000) különböző SiC-tartalmú kerámiák tanulmányozása során a hőmérséklettől függő, hiszterézis jelleget mutató diffuzivitás változást regisztrált. A hiszterézist mikrostrukturális változások idézik elő, például mikropedések miatti különböző expanziós jelenségek, expanziós anizotrópiák, és/vagy kristályszerkezet

módosulások. Calcogenid félvezetők vizsgálata során *Wawryk at al.* (2000) hiszterézises hőmérsékletfüggő hővezetéstől számolt be. Hasonló jelenséget tapasztalt kompozit anyagokon végzett mérések során mágneses térben *Kashiwada at al.* (2003).

A hőmérséklettől függő hővezetési tulajdonságokkal rendelkező anyagokban lejátszó termikus folyamatok modellezése során rendszerint polinomiális vagy exponenciális közelítésekkel találkozhatunk a szakirodalomban (*Dai at al.*, 1998; *González-Fernández at al.*, 1998). Jelen munkánkban olyan modellt mutatunk be, amely figyelembe veszi a diffuzivitás hőmérséklettől függő hiszterézis jellegét. A numerikus megoldás a nemlineáris jelleg miatt igen erőforrás és időigényes művelet, ezért olyan algoritmust javasunk, amely kihasználva a multigrid technika adta lehetőségeket, rövidebb CPU idő elérésére törekszik. A módszert a durvább felbontáson iteráló multigrid módszernek neveztük el, segítségével a futási idő a finom felbontáson végzett iterációhoz képest lényegesen lecsökkenthető. Teszteredmények bemutatásával elemezzük a módszer pontosságát, a mért normák alapján eldönthető, hogy egy adott probléma estén elfogadható-e a javasolt technika. A munka befejező részében a hiszterézises és a konstans hődiffúzió közötti eltéréseket mutatjuk be hevítési/hűtési folyamatok modell-eredményeinek összevetésével.

MATEMATIKAI MODELL

A tranziens hődiffúzió egyenlete hőmérsékletfüggő diffuzivitást feltételezve kétdimenziós esetben az alábbi parciális differenciálegyenlettel írható le:

$$\frac{\partial \Theta}{\partial t} = \frac{\partial}{\partial x} \left[a(\Theta) \frac{\partial \Theta}{\partial x} \right] + \frac{\partial}{\partial y} \left[a(\Theta) \frac{\partial \Theta}{\partial y} \right] + q^*(\Theta, t) \quad (1)$$

ahol $\Theta(t)$ a dimenziómentes hőmérséklet, $\Theta = (T - T_0) / (T_{\max} - T_0)$, $a(\Theta)$ a hődiffuzivitás, $q^*(\Theta, t)$ a forrástag. A vizsgálatokhoz $x, y \in \Omega$, L oldalhosszúságú $\Omega = [0, L] \times [0, L]$ négyzet alakú tartományt tételezünk fel. Az (1) parabolikus differenciálegyenlet megoldása megkívánja a kezdeti $\Theta(x, y, 0) = \Theta_{\text{init}}$, $x, y \in \Omega$, és a peremfeltételek megadását. Elsőrendű peremfeltételek estében $\Theta(x, y, t) = \Theta_w$, $x, y \in \Gamma(\Omega)$ ismert, $\Gamma(\Omega)$ jelenti a tartomány peremét. A továbbiakban azt az esetet tanulmányozzuk, amikor a diffuzivitás hőmérsékletfüggése hiszterézises jelleget mutat.

NUMERIKUS MEGOLDÁS

A nemlineáris modell numerikus megoldásához diszkretizáljuk az (1) differenciálegyenletet térben a centrális differencia sémát és időben a súlyozott differencia sémát alkalmazva. Az Ω tartományt felosztó háló pontjainak távolsága mindkét irányban legyen $h = L/n$, a pontok száma $(n+1) \times (n+1)$, a peremek legyenek hálópontokon. Az 1. ábra felső részén a tartomány finomabb felbontását, alatta ugyanazon tartomány egy durvább, $H = 2h$ felosztását láthatjuk. Az egymásnak megfelelő pontokat bekereteztük.

A súlyozott differenciasémához a hőmérsékletfüggő diffuzivitásokat mindkét idősíkon azonos, átlagolt értékkel közelítjük. Bevezetve egy középső időszintet, ahol a hőmérséklet a k és $k+1$ időlépések átlaga, meghatározzuk a diffuzivitás értékét az $a(\Theta)$

hiszterézis modell felhasználásával: $\bar{a}_{ij}^{k+1/2,l} = a \left(\frac{1}{2} (\Theta_{ij}^{k+1,l-1} + \Theta_{ij}^k) \right)$. A diffuzivitás

referencia pontjait a 2. ábra szemlélteti. Láthatjuk, hogy a hálópontok közé átlagolt diffuzivitással számolunk a k és a $k+1$ időpillanatban is.

1. ábra

Diszkrétizálás térben: finomabb és durvább felbontás

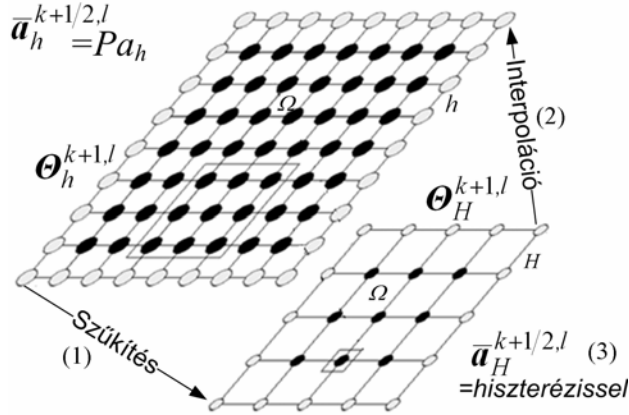


Figure 1: Discretising in space: finer and coarser resolutions

Interpolation(1), Restriction(2), Diffusivity is determined on the coarser grid with hysteresis function(3)

2. ábra

A diffuzivitás vonatkoztatási pontjai

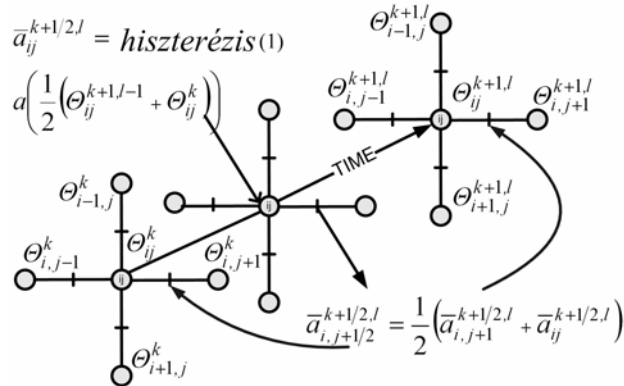


Figure 2: The reference points of the diffusivity in a time-centered method

Hysteresis(1)

A térbeli diszkretizálást két lépésben végezzük el. Az elsőrendű deriváltak jelentik a hőfluxus különbségeket. A hőfluxust a hálópontok közötti képzelte felületre vonatkoztatjuk:

$$\frac{\theta_{ij}^{k+1,l} - \theta_{ij}^k}{\Delta t} = \frac{1}{h} \left(\sigma (\Delta J_{i\pm 1/2,j}^{k+1,l} + \Delta J_{i,j\pm 1/2}^{k+1,l}) + (1 - \sigma) (\Delta J_{i\pm 1/2,j}^{k,l} + \Delta J_{i,j\pm 1/2}^{k,l}) \right) + q^{*k+1/2} \quad (2)$$

A diffuzivitást két szomszédos pont közötti átlagolással kapjuk meg, pl. az (i,j) és $(i,j+1)$ pontok között (2. ábra):

$$\bar{a}_{i,j+1/2}^{k+1/2,l} = \frac{1}{2} (\bar{a}_{i,j+1}^{k+1/2,l} + \bar{a}_{ij}^{k+1/2,l}) \quad (3)$$

A hőfluxus a $k+1$ -ik időpontban, a nemlinearitás miatti iteráció l -ik lépésében, az $(i+1/2,j)$ határfelületen

$$J_{i+1/2,j}^{k+1,l} = \bar{a}_{i+1/2,j}^{k+1/2,l} (\theta_{i+1,j}^{k+1,l} - \theta_{ij}^{k+1,l}) / h \quad (4)$$

Ugyanezen a felületen a k -ik időlépésben: $J_{i+1/2,j}^{k,l} = \bar{a}_{i+1/2,j}^{k+1/2,l} (\theta_{i+1,j}^k - \theta_{ij}^k) / h$. A többi helyen a fluxust hasonló módon számíthatjuk. Behelyettesítve a (4) hőfluxusokat a (3) egyenletbe és átrendezve, az alábbi differenciasémát kapjuk:

$$\begin{aligned} (1+p)\theta_{ij}^{k+1,l} - \frac{\Delta t}{2h^2} (\bar{a}_{i\pm 1/2,j}^{k+1/2,l} \theta_{i\pm 1,j}^{k+1,l} + \bar{a}_{i,j\pm 1/2}^{k+1/2,l} \theta_{i,j\pm 1}^{k+1,l}) = \\ (1-p)\theta_{ij}^k + \frac{\Delta t}{2h^2} (\bar{a}_{i\pm 1/2,j}^{k+1/2,l} \theta_{i\pm 1,j}^k + \bar{a}_{i,j\pm 1/2}^{k+1/2,l} \theta_{i,j\pm 1}^k) + q_{ij}^{k+1/2} \Delta t \end{aligned} \quad (5)$$

ahol

$$\bar{a}_{i\pm 1/2,j}^{k+1/2,l} \theta_{i\pm 1,j}^{k+1,l} = \bar{a}_{i+1/2,j}^{k+1/2,l} \theta_{i+1,j}^{k+1,l} + \bar{a}_{i-1/2,j}^{k+1/2,l} \theta_{i-1,j}^{k+1,l}, \quad (6)$$

stb., és

$$p = \frac{\Delta t}{2h^2} (\bar{a}_{i-1/2,j}^{k+1/2,l} + \bar{a}_{i+1/2,j}^{k+1/2,l} + \bar{a}_{i,j-1/2}^{k+1/2,l} + \bar{a}_{i,j+1/2}^{k+1/2,l}). \quad (7)$$

A bemutatott módszer tulajdonképpen a DuFort–Frankel séma kétdimenziós kiterjesztésének tekinthető, így csak $\Delta t \bar{a}_{\max} / h^2 \ll 1$, $\bar{a}_{\max} = \max(\bar{a}_{ij})$, $i,j=1,\dots,n-1$ korlátozással konzisztens az eredeti differenciálegyenlettel (Stojan, 2005).

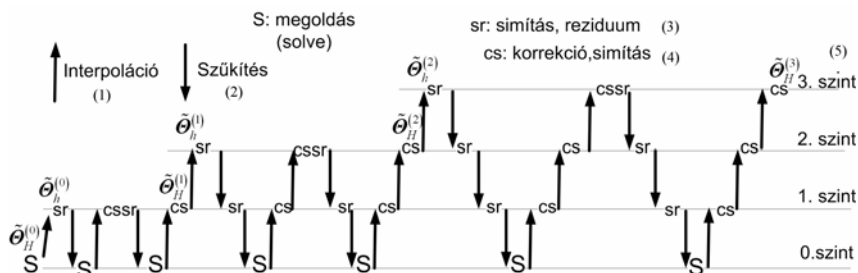
A belső hálópontokra az (5) egyenlet az alábbi mátrix alakban írható fel:

$$A_h^{k+1,l} \boldsymbol{\theta}_h^{k+1,l} = B_h^{k+1,l} \boldsymbol{\theta}_h^k + Q_h^{k+1/2} \quad (8)$$

ahol $A_h^{k+1,l}$ együtthatómátrix, $\boldsymbol{\theta}_h^{k+1,l}$ az ismeretlen hőmérsékletvektor, $B_h^{k+1,l}$ a $\boldsymbol{\theta}_h^k$ ismert hőmérsékletvektor együtthatómátrixa, $Q_h^{k+1/2}$ a forrástag Δt -szerese. Az alsó indexben található 'h' a térbeli felbontás szintjére emlékeztet. Az iteráció kezdetén $\boldsymbol{\theta}_h^{k+1,0} = \boldsymbol{\theta}_h^k$, $\bar{a}_{ij}^{k+1,0} = \bar{a}_{ij}^k$ és az iteráció során $\bar{a}_{ij}^{k+1/2,l} = a \left(\frac{1}{2} (\theta_{ij}^{k+1,l-1} + \theta_{ij}^k) \right)$, $i,j=1,\dots,n-1, l=1,2,\dots$

A (8)-hoz hasonló nagyméretű rendszerek megoldásának korszerű eszköze a multigrid technika, többek között jó konvergencia tulajdonságai miatt (Stojan, 2005). A multigrid módszerek lényegét ismertnek tételezzük fel (lásd pl. Hackbusch, 1985). A javasolt hierarchikus algoritmusba az ún. Full Multigrid (FMG) és a sima multigrid (MG) technikát is beépítettük, a FMG W -ciklusokat szemlélteti a 3. ábra.

3. ábra

Teljes multigrid módszer (FMG) W -ciklus négy felbontási szintenFigure 3: Full Multigrid W -cycles on four grid levels

Interpolation(1), Restriction(2), Smoothing and residuum(3), Correction and smoothing(4), Grid levels(5)

A diffuzivitás hőmérséklet függése miatt a (8) implicit séma iterációt igényel. Az iteráció nagyméretű rendszerek esetében igen erőforrás és időigényes művelet. A hiszterézis diffuzivitás is jelentősen növeli a megoldási időt, mert minden egyes hálópont saját lokális hiszterézis memóriával rendelkezik. Az általunk javasolt módszerrel, kihasználva a multigrid technika kínálta lehetőségeket, a megoldási idő jelentősen csökkenthető. Az iteráció előtt a (8) problémát redukáljuk egy durvább, $H=2h$ felbontásra (1. ábra):

$$A_H^{k+1,l} \theta_H^{k+1,l} = B_H^{k+1,l} \theta_H^k + Q_H^{k+1/2}, \quad (9)$$

ahol az iteráció $\theta_H^{k+1,0}$ $k+1$ -ik időpontbeli kezdeti hőmérsékletét az ún. szűkítéssel (restriction: R_h^H) határozzuk meg a k -ik időpont hőmérsékletéből: $\theta_H^{k+1,0} = \theta_H^k = R_h^H \theta_h^k$.

A diffuzivitás kezdeti értéke a hiszterézis modellből: $\bar{a}_{ij_H}^0 = a(\theta_{ij_H}^k)$. A $Q_H^{k+1/2}$ forrástagot is leszűkítjük, de $Q_H^{k+1/2}$ a durvább felbontásra közvetlenül is kiszámítható. A diffuzivitás hőmérséklet-függése miatt iterációt a (9) egyenlettel ezen a durvább felbontáson végezzük el, $\bar{a}_{ij_H}^{k+1/2,l} = a\left(\frac{1}{2}(\theta_{ij_H}^{k+1,l-1} + \theta_{ij_H}^k)\right)$. Az iteráció véget ér, ha két iteráció közötti különbség

egy előre definiált ε érték alá csökken: $|\theta_H^{k+1,l} - \theta_H^{k+1,l-1}| < \varepsilon$. Az iteráció eredményeként kapott diffuzivitás- és hőmérsékletértékeket interpoláljuk a finomabb felbontásra. Jelölje P_H^h az ún. prolongációs operátort, I_H^h a bilineáris interpolációs operátort, ekkor $\theta_h^{k+1,0} = I_H^h \theta_H^{k+1,l}$ és $\bar{a}_{ij_h}^{k+1/2} = S_{GS} P_H^h \bar{a}_{ij_H}^{k+1/2,l}$, ahol S_{GS} jelenti a Gauss-Seidel simító eljárást. Az interpolált hőmérséklet jó kezdeti hőmérsékletnek tekinthető a finomabb (eredeti h) felbontáson. A finomabb felbontás megoldásaként keresett hőmérsékletmezőt további multigrid iterációval határozzuk meg a (8) egyenlet együtthatói és forrástagja szerint, de a diffuzivitás értékét már nem változtatjuk. A 4. ábrán láthatjuk a javasolt durvább szinten iteráló multigrid algoritmus vázlatát. A nemlineáris iterációban FMG

algoritmust, a finomabb szinten MG algoritmust futtatunk. Az FMG módszer előnye, hogy a legdurvább szinten pontos megoldással kezdi az iterációt, így gyorsabban, egyetlen ún. W-ciklus sorozattal jut el a megfelelően pontos megoldáshoz. A finomabb felbontáson erre nincs szükség, hiszen az alacsonyabb szintről már egy jó kezdeti becsléssel lépünk fel.

4. ábra

A javasolt durvább szinten iteráló multigrid algoritmus vázlata

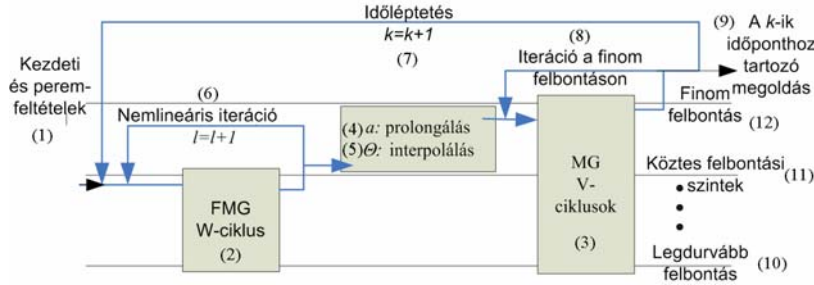


Figure 4: The coarse level iterating multigrid algorithm

Initial conditions(1), Full multigrid W-cycles on the coarser grid(2), Multigrid V-cycles on the finest grid(3), Prolongation(4), Interpolation(5), Iteration of nonlinearity(6), Time stepping(7), Iteration on finest grid level(8), Solution at the time index k (9), Coarsest grid(10), Medial grid levels(11), Finest grid(12)

Az $A_H^{k+1,l}$ és $B_H^{k+1,l}$ együtthatómátrixok konvolúciós magjait, a θ_H^k hőmérsékletet, a $Q_H^{k+1/2}$ forrástágot minden egyes felbontási szintre meg kell határoznunk. Az együtthatómátrixok magjai pontról pontra változnak a diffuzivitás függvényében, pl.:

$$A_H^{k+1,l} = \begin{bmatrix} 0 & -c\bar{\kappa}_{i,j-1/2_H}^{k+1/2,l} & 0 \\ -c\bar{\kappa}_{i-1/2,j_H}^{k+1/2,l} & (1+p)_H & -c\bar{\kappa}_{i+1/2,j_H}^{k+1/2,l} \\ 0 & -c\bar{\kappa}_{i,j+1/2_H}^{k+1/2,l} & 0 \end{bmatrix} \quad (10)$$

ahol $c=\Delta t/2H^2$. A javasolt algoritmus tehát minden szinten tárolja az előző időpont hőmérsékleteit, a forrástágot, a diffuzivitást, a konvolúciós magokat pedig minden hálópontra és minden felbontási szinten újból generálja. Mivel az FMG minden felbontási szinten előállítja a megoldást, a kerekítési hiba becslhető és felhasználható a konvergencia meghatározására. A MG algoritmusban V-ciklusokat iterálunk, itt csak a legfinomabb felbontáson kapjuk meg a hőmérsékleteket. Az iteráció stop feltétele ha a reziduum normája a kiindulási norma ε -od részére csökken:

$$\|r_h^{k+1}\| < \varepsilon \|A_h^{k+1,l} I_H^h \theta_H^{k+1,l} - B_h^{k+1,l} \theta_h^k - Q_h^{k+1/2}\|, \quad \varepsilon < 1.$$

A hiszterézis numerikus közelítése

A diffuzivitás hőmérsékletfüggését leíró numerikus modellhez a mágneses hiszterézis leírására elterjedt diszkrét Preisach modellt alkalmaztuk. A diszkrét Preisach modell alapja az ún. Preisach háromszög (Iványi, 1997), amely véges számú hiszterézis relét tartalmaz

(5. ábra). A modell bemeneti változója a hőmérséklet, kimenete a diffuzivitás. Az eredeti modellt úgy módosítottuk, és a $\mu(\Theta_\alpha, \Theta_\beta)$ eloszlásfüggvényt úgy választottuk meg, hogy a hiszterézis alakja feleljen meg a Santos *et al.*, (2000) munkájában bemutatott, mérésrel felvett hiszterézis görbének. A diffuzivitás hőmérsékletfüggését leíró rekurzív formula:

$$a(\Theta^{k+1}) = a(\Theta^k) \pm \left(b^v \frac{T_p}{m} \sum_j \sum_i \mu(\Theta_{\alpha_i}, \Theta_{\beta_j}) \right), i, j = m, i, j \in P \quad (11)$$

Ahol T_p a P tartomány területe, amely felett az utolsó lépésben megváltozott a hiszterézis relék állapota, m a P tartományban található relék száma (5. ábra). A második rész előjele a hőmérsékletváltozás irányától függ. Emelkedő hőmérséklet esetén pozitív, csökkenő hőmérséklet esetében negatív az előjel. v a fűtési-hűtési ciklusok száma, b^v egy megfelelő szorzótényező, mindkettő értéke függhet a ciklusszámtól. A diffuzivitás alapértéke: $a(\Theta^0) = a^v$ szintén a ciklusszám függvénye. A 6. ábrán láthatunk néhány, a fenti modell segítségével előállított hiszterézis görbét.

Hevítés során a hőmérséklet emelkedésével a mikrostrukturális változások miatt a diffuzivitás hirtelen emelkedést mutat. További hőmérsékletemelkedésre a diffuzivitás alig változik, majd a magasabb hőmérséklet tartományban újabb meredek emelkedés tapasztalható. A hőmérséklet csökkentésével a diffuzivitás is csökken, de magasabb értéken marad, mint hevítés során. Egymást követő termikus ciklusok hiszterézis görbéje nem feltétlenül esik egybe, ezt fejezi ki a modell v paramétere.

A javasolt módszer jellemző tulajdonságai

A javasolt módszer jogosultságát, alkalmazhatóságát és jellemző paramétereit az alábbi tesztprobléma segítségével vizsgáltuk: a kezdeti hőmérséklet-eloszlás legyen: $\Theta = (x, y, 0) = 1$, $t = 0$, a peremfeltétel: $\Theta = (x, y, t) = 0$, $x, y \in \Gamma\Omega$, $t > 0$. Az időlépés: $\Delta t = 500$ s, $L = 0.5$ m, $a_0 = 2.5 \cdot 10^{-7}$ m²/s, a felbontás 33×33 pont. A multigrid paraméterek: az elő- és utósimítások száma 3, a legdurvább felbontás 5×5 pont, a felbontási szintek közötti léptetés bilineáris interpolációval illetve teljes szűkítéssel (full weighting restriction) történt. A legdurvább szinten a fokozatos túlrelaxálás algoritmusát alkalmaztuk megoldóként.

5. ábra

A Preisach háromszög (a) és eloszlásfüggvény 3D-ben (b)

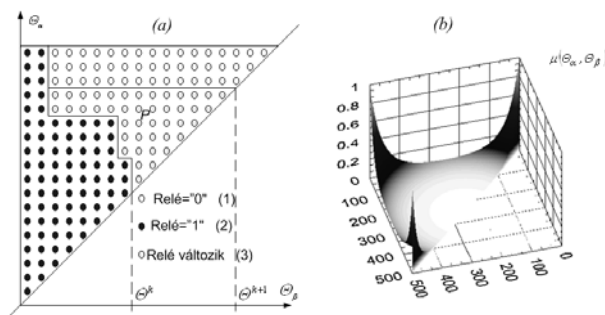


Figure 5: The Preisach triangle (a) and the distribution function (3D) (b)

Relay is off (1), Relay is on(2), Relay is changing to on(3)

6. ábra

A diffuzivitás hőmérsékletfüggése a diszkrét modell szerint

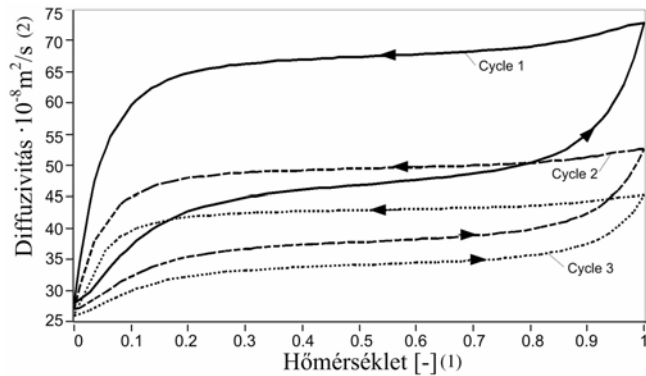


Figure 6: The temperature dependent hysteresis of the thermal diffusivity and effect of the thermal cycling.

Temperature(1), Thermal diffusivity(2)

A javasolt algoritmusban a durvább felbontás iterációja után a diffuzivitás értéke változatlan marad. Kérdés, vajon milyen módon vetítsük fel a finomabb felbontásra? Kétdimenziós esetben a bilineáris interpoláció a legelterjedtebb technika, alkalmazható még a sima prolongáció is. (Mivel célunk a minél rövidebb CPU idő, magasabb rendű interpolációs módszerek nem jöhetnek szóba.) A 7. ábrán látható a hőmérsékleti görbék a vizsgált tartomány középvonalában, három különböző időpontban, három különböző numerikus módszerrel készültek. A 7. ábra a görbéje a finom felbontáson végzett iteráció eredménye. A b görbe mutatja a prolongálással vetített diffuzivitással, a c görbe a bilineáris interpolációval kapott megoldást. Látható, hogy mindhárom időpontban a prolongálással kapott diffuzivitással végzett számítások közelítik meg jobban a finom felbontáson végzett iteráció eredményét.

A teljes hőmérsékleti mező közti különbséget normaképzéssel határozhatjuk meg. A 8. ábra a görbéje a bilineáris interpolációval vetített diffuzivitás és a finom felbontáson végzett iteráció, mint referencia hőmérséklet-eloszlás közötti különbség normáját mutatja az időben. A b görbe a prolongáció és a referencia hőmérsékletmező közti különbség normája. Látható, hogy az b görbe mindvégig az a görbe alatt marad egy nagyságrenddel. Közelítve az állandósult állapot felé, a normák csökkennek. A módszer alkalmazható, ha az eltérés nagyságrendje tolerálható.

A diffuzivitás időbeli átlagolásának jogosságát különböző időlépésekkel kapott eredmények összehasonlításával teszteltük, azonos térbeli felbontás mellett. A referencia időlépés az eredeti időlépés tizede. A 8. ábra c görbéje mutatja a két időlépés közötti különbség normáját. Az első időpillanattól eltekintve gyakorlatilag nincs különbség a két számítás között, vagyis az időbeli átlagképzés megfelelő, amíg a regularitást és konzisztenciát nem sérti Δt növelése.

A durva felbontáson végzett iterációs módszer futási idejét egységnyiinek véve, a finom felbontáson végzett iteráció CPU ideje négy. Tárigény tekintetében a hiszterézis számításának durvább felbontási szintre történő helyezése jelenti a legnagyobb

megtakarítást. Minden hálóponthoz tartozik egy lokális memória, amelyben az adott ponthoz tartozó $\Theta_{ij}(k\Delta t)$ sarokpontjait tároljuk. Feltételezve, hogy minimálisan 30 értéket tárolunk és dolgozunk fel pontonként és ciklusonként, az iterációt áttéve alacsonyabb felbontásra, kb. $23n^2$ adatnak megfelelő memóriát takaríthatunk meg. Nagyméretű feladatok esetében ez jelentősnek tekinthető és kihatással lehet a CPU időkre is.

7. ábra

Hőmérsékletek a tartomány középvonalán, 3 különböző időpontban, a diffuzivitás iterálása a finom felbontáson (a), iteráció a durva felbontáson, vetítés prolongálással (b), vetítés bilineáris interpolációval (c)

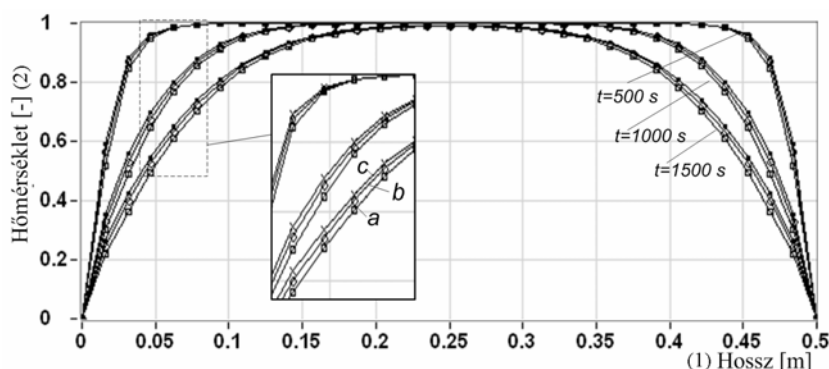


Figure 7: Temperature profiles along the axis of the domain at three time steps: fine grid iteration (a), coarse level iteration with prolongation (b), with bilinear interpolation (c)

Length(1), Temperature(2)

8. ábra

A számítási algoritmusok különbségéből származtatott normák

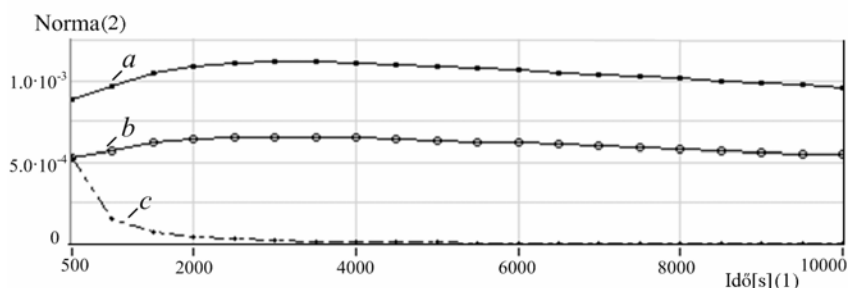


Figure 8: Norm of differences of the calculation schemes: differences between the fine grid iteration and coarse grid iterations with bilinear interpolation (a), with prolongation (b), differences between two distinct time steps (c),

Time(1), Norm(2)

SZIMULÁCIÓS EREDMÉNYEK

A szimulációs algoritmus segítségével a nemlineáris diffúzió fenti modelljének néhány érdekes tulajdonságát szeretnénk bemutatni, összevetve a konstans diffúziót feltételező modellel. Példáink olyan eseteket elemeznek, ahol a diffuzivitás a fő hevítési és hűlési görbe mentén változhat, például a peremek monoton hűtésével vagy fűtésével. Az első példában $\Theta=(x,y,0)=1$, $t=0$, a peremfeltétel: $\Theta=(x,y,t)=1$, $x,y \in \Gamma\Omega$, $t>0$. A második esetben: $\Theta=(x,y,0)=1$, $t=0$, a peremfeltétel: $\Theta=(x,y,t)=0$, $x,y \in \Gamma\Omega$, $t>0$. A többi paraméter megegyezik az előző fejezet adataival. A 9. ábrán látható, hogy a vizsgált tartomány közepén hogyan alakul a hőmérséklet az első ötven időlépésben, a) és c) konstans b) és d) hiszterézises diffuzivitással.

9. ábra

Hevítési (a, b) és hűlési (c, d) hőmérsékleti profilok a tartomány középvonalán, konstans (a, c) és hiszterézises (b, d) diffuzivitást feltételezve

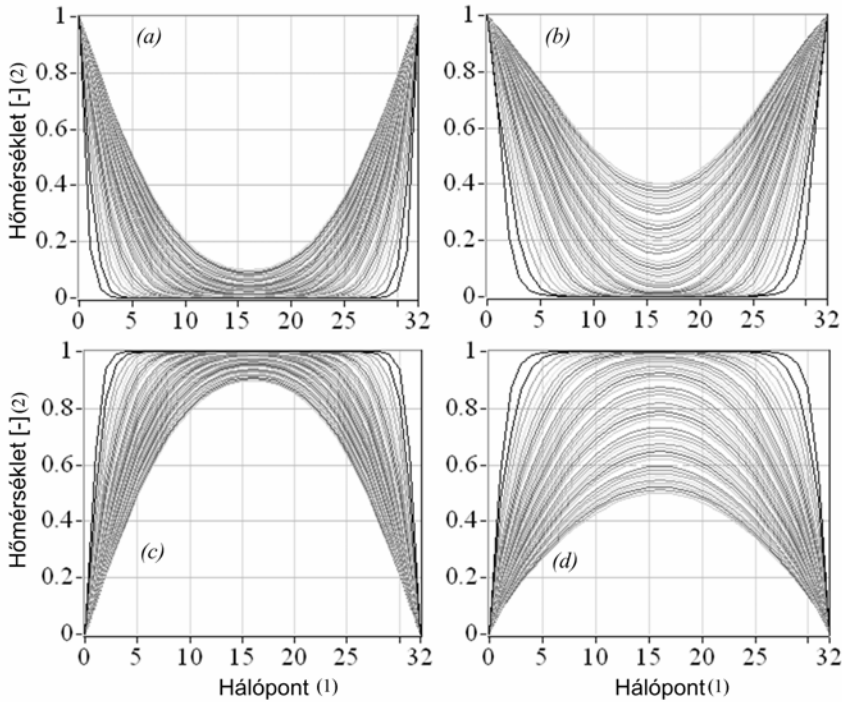


Figure 9: Heating-cooling profiles of the middle axes of the domain: with constant (a) and (c), with hysteretic diffusivity (b) and (d),

Grid point(1), Temperature(2)

A 10. ábrán a tartomány középpontjának hőmérsékleti tranzienseit láthatjuk a négy vizsgált esetre. Konstans diffuzivitással a hevítés-hűtés tranziensek szimmetrikusak. A hiszterézises diffuzivitás miatt a hevítési-hűtési tranziensek eltérőek.

10. ábra

Hőmérséklet-tranziensek a tartomány közepén, konstans és hiszterézises diffuzivitással

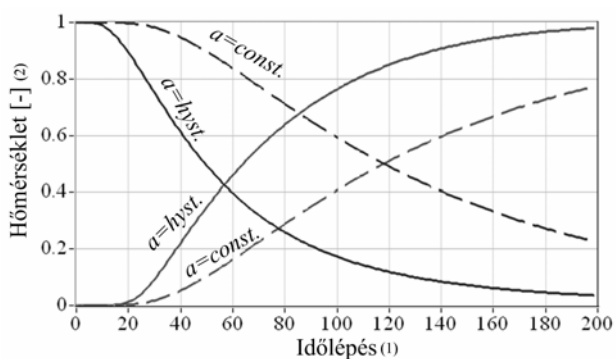


Figure 10: Temperature transients of the middle of the domain with hysteretic and constant diffusivity

Time index(1), Temperature(2)

11. ábra

A hűtési folyamat hőmérséklet-eloszlása a 20. időpillanatban és a hozzá tartozó diffuzivitás

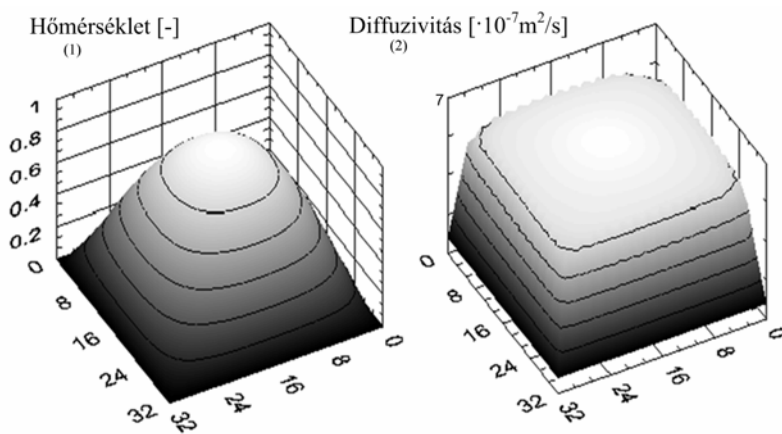


Figure 11: Cooling process: temperature field (a) and the corresponding thermal diffusivity at time-index $k=20$ (b),

Temperature(1), Thermal diffusivity(2)

A termikus folyamat felgyorsult, a hűlési folyamat magasabb diffuzivitással indul, alacsonyabb hőmérsékleteken is magasabb értéken marad, mint hevítés során. A 11.

ábrán látható a huszadik időpillanathoz tartozó hőmérsékleti mező és a hozzá tartozó diffuzivitás. A diffuzivitás alakja még emlékeztet a kiindulási hőmérséklet-eloszlásra.

A teljes felületen a két fajta diffúziót az eltérések normájával hasonlíthatjuk össze. A 12. ábra *a* görbéje a hevítési folyamatra, *b* görbéje a hűtési folyamatra vonatkozóan mutatja a konstans és a histerézises diffúzió közötti eltérés normájának alakulását. A norma az idő előrehaladásával nő, elér egy maximum értéket, majd lassan csökken, ahogy közelít az állandósult állapothoz.

12. ábra

Az eltérések normája a hevítési (*a*) és a hűtési (*b*) folyamat során

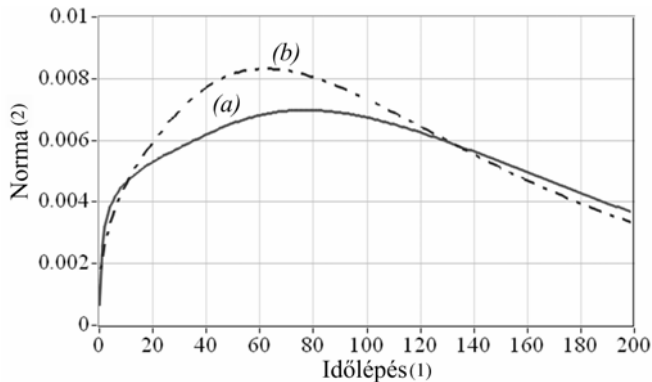


Figure 12: Norm of differences between hysteretic and normal heat diffusion: in the heating(*a*) and cooling processes(*b*),

Time index(1), Norm(2)

A hűtési folyamat a magasabb diffuzivitások miatt kezdetben jobban eltér a konstans diffuzivitású folyamattól, és mivel a folyamat így felgyorsult, az állandósult állapotot is gyorsabban közelíti meg. Az itt bemutatott nemlinearitás azon kompozit anyagok jellemzője, amelyeknek diffuzivitása az ismertett histerézis görbékhez hasonló tendenciát mutat. A diffúziós modell megfelelő paraméterezésével a javasolt algoritmus más jellegű, és más irányú folyamatok szimulálására is alkalmas. A modell lehetőséget nyújt az ún. histerézises minor hurkokon változó diffuzivitás hatásának tanulmányozására is.

KÖVETKEZTETÉSEK

A hőmérséklettől függő diffuzivitású anyagokban lejátszódó termikus folyamatok tranziensei jelentősen eltérhetnek a konstansnak tekinthető vezetési tulajdonságú anyagokétól. Munkánkban ezen folyamatok modellezésére alkalmas olyan hierarchikus algoritmust mutattunk be, amely a multigrid technika adta lehetőségekre építve, rövidebb futási idő elérésére törekszik. Az algoritmus igen jó konvergencia tulajdonságú, elegendő minden időlépésben néhány iterációs ciklus a DuFort-Frankel sémához hasonló implicit sémával. Az időlépés nagysága az implicit séma ellenére nem választható meg

tetszőlegesen, a séma csak korlátozásokkal konzisztens az eredeti differenciálegyenlettel, annak ellenére, hogy a megoldás konvergens, túl nagy időlépéssel nem reguláris megoldást kaphatunk. A hiszterézises diffúzió hűtés/fűtés tranziense aszimmetrikus jellegű. A bemutatott példák egy bizonyos típusú hiszterézises diffúzióra jellemzőek. A hiszterézis modell paramétereinek, eloszlásfüggvényének módosításával az algoritmus más jellegű, más irányban változó hiszterézises diffúziós modellek megoldására is alkalmas.

KÖSZÖNETNYILVÁNÍTÁS

Az alkalmazott kutatás a Nemzeti Fejlesztési Terv Gazdasági Versenyképesség Operatív Program „Infokommunikációs technológia kidolgozása és hasznosítása az energia-elosztás területén ” című pályázata (GVOP-3.1.1.-2004-05-0125/3.0.) keretében valósul meg.

IRODALOM

- Dai, W., Woodward, P.R., (1998). Numerical Simulation for Nonlinear Heat Transfer in a System of Multimaterials. In: J. Comp. Phys. 139. 58-78.
- González-Fernández, C.F., Alhama, F., Alarcón, M., López-Sánchez, J.F. (1998). Digital simulation of transient heat conduction with polynomial variable thermal conductivity and specific heat. In: Comp. Phys. Comm. 111. 53-58.
- Hackbusch, W. (1985). Multigrid Methods and Applications, Springer, Berlin.
- Iványi, A. (1997). Hysteresis models in electromagnetic computation. Budapest : Akadémiai Kiadó
- Kashiwada, Y., Fujishiro, H., Ibeke, M. (2003). Thermal conductivity of $\text{Pr}_{0.65}(\text{Ca}_{1-z}\text{Sr}_z)_{0.35}\text{MnO}_3$ under applied field. Physica B. 329-333, 924-925.
- Santos, W.N., Baldo, J.B., Taylor, R. (2000). Effect of SiC on the thermal diffusivity of silica-based materials. In: Material Research Bulletin. 35. 2091-2100.
- Stojan, G., Takó, G. (1995). “Numerical methods”, (in Hungarian) ELTE. Budapest.
- Wawryk, R., Marucha, Cz., Balcerek, K., Terzijska, B.M., Ivanova, Z., (2000). Thermal Conductivity of polycrystalline and amorphous Se-Te-Cu system. In: Cryogenics. 40. 749-752.

Levelezési cím (*Corresponding author*):

Jancskárné Anweiler Ildikó

Pécsi Tudományegyetem, Műszaki Kar, Műszaki Informatika Tanszék
7624 Pécs, Rókus u. 2.

*University of Pécs, Department of Information Technology
H-7624 Pécs, Rókus u. 2.*

Tel.: 36-72-501-534 Fax: 36-72-501-534

e-mail: jai@morpheus.pte.hu



Városi vízgőzhálózat modellezése és számítógépes felügyeleti rendszerének kidolgozása

Szakonyi L.

Pécsi Tudományegyetem, Pollack Mihály Műszaki Kar, Műszaki Informatika Tanszék, 7624 Pécs, Rókus u. 2.

ÖSSZEFOGLALÁS

Az előadás új információs és kommunikációs technológia regionális hasznosítását jelentő projektet mutat be, mely városi vízgőzhálózatot számítógépes felügyeleti, szakértői és döntéstámogató rendszerének kidolgozását célozza az energetikai veszteségek csökkentése érdekében. Az alkalmazott kutatási tevékenység a kiépítendő infokommunikációs technológia felhasználásával működő energiaáram-hálózat kísérleti identifikációjának elvégzésére, az energiaáram-hálózat számítógépes szimulációjának kiértékelésével az energiatakarékos üzemvitelt biztosító irányítási stratégia meghatározására irányul.

(Kulcsszavak: kétfázisú áramlás, nem-lineáris hálózatok, identifikáció, folyamatirányítás)

ABSTRACT

Identification of the municipal water-steam network, elaboration of its computer controlled system

L. Szakonyi

University of Pécs, Department of Information Technology, H-7624 Pécs, Rókus u. 2.

The presentation introduces a project utilizing the new information and communication technology in a regional perspective and its objective is to elaborate the control-, expert- and decisionmaking support-system in the city's water-steam network informatics network in the interest of reduction of power loss. The applied research activity is aimed at the completion of the experimental identification of the energy-current network utilizing the infocommunication technology to be developed and the definition of the control strategy ensuring energy-saving operation based on the assessment of computer simulation.

(Keywords: Two-phase flow, film condensation, persistence equations, non-linear networks, identification, process control)

ELŐZMÉNYEK

Műszaki technológiai folyamatok identifikálásánál az elsődleges feladat a folyamat absztrahált leírása matematikai modell formájában (Szakonyi, 2002a). A vizsgált anyag – és energiaáram hálózat (városi vízgőzhálózat) üzemviteli paramétereinek helyes beállítása, az energia fogyasztói igényekhez igazodó, folyamatosan kontrollálható és módosítható elosztása igényli a matematikai modellen végzett vizsgálatot, a számítógépes szimulációt. Az elmúlt időszakban az alkalmazott informatika területén megkezdett projektünk (GVOP-3.1.1-2004-05-0125/3.0) bázisrendszerül a Pécs városi vízgőzhálózat szolgál, mely éves szinten ~130 ezer tonna vízgőz elosztásával, ~550 mFt

energiaköltség kiszámlázásával a város jelentős részére terjed ki, s a jelenleginél jóval nagyobb vízgőzigény ellátására alkalmas. Az energiaveszteségek csökkentésének igénye indokolja a veszteségek felmérését, helyi és időbeli pontosítását, továbbá a javaslatot az energiatakarékosabb üzemvitelre. Az előbbi tevékenységek az energiaelosztásban érdekelt valamennyi partner, az energiát előállító, a szolgáltató és a felhasználók érdekeinek és lehetőségeinek figyelembevételével valósulhatnak meg.

E megállapodások sikerét

- a működő rendszeren elvégzett megbízható mérések, a kísérleti identifikáció;
- a fogyasztói igényekhez igazodó, változtatható üzemállapotok és beavatkozások lehetősége;
- az előbbi üzemállapot-alternatívákat megvalósító, energiaáram elosztás modellezésére szolgáló, nemlineáris hálózati matematikai modellek kidolgozása;
- az optimális üzemvitel és számítógépes irányítási stratégia felállítása;
- az energetikai veszteségek csökkentése mellett az energiaellátás biztonságát is szolgáló szakértői és döntéstámogató rendszer kimunkálása

segítheti elő.

Tanszékünk regionális célok megvalósítása érdekében a termelői-szolgáltatói szférával (a PÉTÁV Távfűtő Kft.) együttműködve, a műszaki informatika szak gesztor tanszékétől elvárható kutatási-fejlesztési irányoknak (nagy rendszerek komplexitás-kezelése, mesterséges intelligencia alkalmazások, infokommunikációs technológiák kifejlesztése, hasznosítása) is megfelelően törekszik a projekt megvalósítására egy regionális energiaellátó hálózat energiaveszteségeinek csökkentése, a gazdaságosabb üzemvitel biztosítása érdekében (Szakonyi, 2002b).

AZ ALKALMAZOTT KUTATÁS CÉLJA, FŐBB TEVÉKENYSÉGEI

A projekt fő célja a konzorciumi partner által működtetett energiaáram hálózatra telepítendő számítógépes monitoring és irányítási rendszer létrehozásának elősegítése, mely megteremti a lehetőségét:

- az energiaelosztás folyamatos ellenőrzése és szabályozása révén az időszakosan jelentkező „többletenergia” célirányos hasznosításának, a gerinchálózaton és a fogyasztónál jelentkező energiaveszteségek csökkentésének;
- az elszámolás alapját jelentő ellenőrzött és dokumentált energiaáram mérések, ezek igények szerinti feldolgozása és kiértékelése által az energiát előállító, szolgáltató és felhasználó közötti szerződésteljesítések ellenőrzésének;
- az előbbi szerződő felek közötti megállapodásoknál az optimális energiaelosztást és felhasználást jelentő üzemeltetési stratégia kidolgozásának.

A projekt betervezett, elvégzendő tevékenységei

- a kísérleti identifikáció elvégzését, az energiaelosztás ellenőrzését biztosító monitoring rendszer;
- az anyag- és energiaáram hálózat modelljeként, változó „forrásokkal”, ellenállásokkal, fogyasztóhelyekkel működtethető szimulációs rendszer;
- a változó topológiák, üzemállapotok követésére alkalmas üzemviteli program-csomag;
- az előbbi matematikai modellekhez illesztett irányítási stratégia, szakértői és döntéstámogató rendszer

létrehozását célozzák meg. Az így megvalósuló teljes infokommunikációs technológia (a kialakított szoftvercsomag) alkalmas lehet egy régió meghatározó energia-elosztó rendszerei (városi távfűtő, gázszolgáltató, vízmű hálózat stb.) informatikai infrastruktúrájának kialakítására, továbbfejlesztésére a komplexitás-kezelés új, mesterséges intelligencia alapú módszereit követve, tekintettel a fogyasztói (lakossági) kívánalmakra is.

Az alkalmazott kutatásunk bázisrendszerül szolgáló anyag- és energiaáram hálózatra telepítendő számítógépes monitoring és irányítási, valamint a szakértői és döntéstámogató rendszer kiépítése előfeltételezi az alábbi tevékenységeket, melyeket a projekt előrehaladása során az alábbi ütemben kívánunk megvalósítani :

- a főbb energia-felhasználók (mintegy 20 cég) fogyasztói igényeinek felmérése;
- az energiatermelőnél és a fogyasztóknál kialakított helyi mérőhelyekre támaszkodva javaslattétel a konzorciumi partner által üzemeltetendő monitoring rendszer kiépítésére;
- a meglévő, illetve ideiglenesen telepített távadókkal, mérési-adatgyűjtő rendszerrel az energiaveszteségek feltárására irányuló identifikációs mérések megtervezése, elvégzése, kiértékelése;
- az energiaáram-hálózat matematikai modellezése, számítógépes szimulációja változó energiaeosztás és energiafelhasználás esetén, illetve különböző üzemállapot-alternatíváknál;
- az objektumalapú energiaáram hálózati szoftvercsomag elkészítése a változó topológiákra és üzemállapot-alternatívákra felállított algoritmusok programozásával;
- a fogyasztói igényekhez igazodó, energiatakarékos üzemvitelt biztosító irányítási stratégia kidolgozása az ellenőrző- és beavatkozási helyek, az irányítóberendezések (irányítási algoritmusok) megválasztásával;
- az energetikai veszteségek csökkentését, az energiaellátás biztonságát szolgáló szakértői és döntéstámogató rendszer kifejlesztése (*Szakonyi, 2002c*).

A PROJEKT VÁRHATÓ REGIONÁLIS HATÁSAI

Az projekt alábbiakban megnevezett eredmény-mutatói:

- a vízgőzfogyasztások megbízható mérése, a fogyasztóhelyek folyamatos és egyidejű megfigyelhetősége (a megvalósult monitoring rendszer),
- a vízgőzhálózat energiatermelőhöz és fogyasztókhoz igazodó működtetése változó energiaeosztással, beavatkozó szervek alkalmazásával (a kialakítandó irányítási rendszer),
- a biztonságos üzemvitelt, gazdasági és egyéb szempontokat is támogató szakértői rendszer kialakítása

elsősorban a konzorciumi partnert (PÉTÁV Távfűtő Kft.), de az energiaeosztásban érintett, jelenlegi nagyfogyasztókat (Zsolnay Porcelángyár Rt., Pécsi BŐRISZ Kft., BAT Pécsi Dohánygyár, Honvéd Kórház stb.) is befolyásolják. Ugyanakkor a betervezett ütemezéssel a teljes infokommunikációs rendszer kifejlesztése (ennek mindhárom eleme) egyben hatásmutató szerepet is betölt. Ugyanis a teljes infokommunikációs rendszer létrehozásával, s ennek - az energiaeosztásban érdekelt valamennyi partner (termelő, szolgáltató, felhasználók) számára biztosított - elérhetőségével vélhetően további egyeztetések indulnak meg az egyes partnerek között a tényleges energiafelhasználások, illetve az arányosan leosztott veszteségek kiszámlázása érdekében a termelő és a fogyasztók igényeinek és lehetőségeinek összehangolása

céljából. Az egyes partnereknél a kialakított és elérhető infokommunikációs technológia, az üzemvitelről (a rendelkezésre álló hőenergiáról, annak állapotáról, kalorikus jellemzőiről), az üzemviteli változásokról (a beavatkozások, az energiaelosztás módosításának lehetősége) mindenkor rendelkezésre álló információ új fejlesztéseket, a fajlagos energiafelhasználás, illetve az energiaveszteségek csökkentésére irányuló innovatív tevékenységeket indít el valamennyi partnernél.

A gőzhálózat - fogyasztói igényekhez igazodó, az energiaelosztás módosítását biztosító - üzemvitele regionális szinten a gerinchálózaton jelentkező energiaveszteségek csökkentésével költségmegtakarítást eredményezhet. A vízgőzhálózat jelenlegi kiépítettségét és kapacitását alapul véve a számítógépes felügyeleti rendszer kialakítása és üzembe helyezése növeli esélyét új nagyfogyasztók bevonásának és kiszolgálásának is, ami ugyancsak új beruházásokhoz, a kapacitások jobb kihasználásához vezethet.

A projekt-célok és mérőföldkövek rögzítését szakunk lehetőségeinek és kihívásainak figyelembevétele mellett alapvetően befolyásolta előnyös helyzetűnek nem tekinthető régióink infrastruktúrája, a végbement szerkezeti átalakulás, a korábbi meghatározó iparágak megszűnése, a megmaradt cégek és az új vállalkozások nyereségérdekeltsége. Rövidtávon az energiaelosztásban résztvevő partnerek egyéni érdekei, illetve lehetőségei nem szorgalmazzák az energiaáram-hálózat nagy költségekkel járó korszerűsítését. Hosszútávon, illetve regionális szinten azonban mindenképp eldöntendő, hogy az energiaellátás korszerűsítése, az energiaveszteségek csökkentése milyen alternatívákból választva valósítható meg. Ezeket az alternatívákat szeretné feltárni a projekt innovatív informatikai alkalmazásokkal, infokommunikációs technológiák alkalmazásával.

A PROJEKT KÖRNYEZETVÉDELMI SZEMPONTJAI, KOCKÁZATI TÉNYEZŐI, TOVÁBBVITELE

A regionális vízgőzhálózat jelenlegi működtetése során a hálózat gerincvezetékén elhelyezett kondenzleválasztóknál jelentős kondenzáció, a kondenzvíz (vígőz) automatikus kifúvatása, a környezetbe (léggörbe) való vezetése tapasztalható. A meglévő energiahálózat jelenlegi üzemvitele – különösen az egyes fogyasztók csökkentett energiafelhasználásainak egybeesésekor – ugyanis a szállított vízgőz részbeni kondenzálódását idézi elő. Tervezett energiaveszteség feltáró méréseink részben e kondenzveszteségek meghatározására, számszerűsítésére vonatkoznak. A környezetbe kilépő és hasznosítás nélkül a talajba bekerülő forróvíz, a léggörbe kerülő vízgőz jelentős energiaveszteséget jelenthet. Felderítendő, hogy éves szinten a mintegy 75 ezer tonna vízgőzt, ~220 ezer GJ hőenergiát jelentő különbözetből (mérések alapján a kiadott és a felhasznált energia közötti eltérés) a veszteségek milyen mértékben bonthatók elemeikre. A vízgőzhálózat gerincvezetékén telepített kondenzleválasztókon, illetve az egyes energia-felhasználó cégek telephelyén javarészt hasznosítatlanul a környezetbe (léggörbe) távozó hőenergia közvetett mérési eredmények szerint ~8,5 t/h tömegáramot, ~25 GJ/h energiaáramot feltételez, s ez megfelel a négy legnagyobb vízgőzfelhasználó együttes átlagos vízgőzfogyasztásának. E jelentős, mérésekkel alátámasztandó, jelenleg veszteségként elkönyvelt energiaáram részbeni hasznosítási lehetőségein túlmenően környezetre gyakorolt hatása is elemzést igényel. Kísérleti vizsgálataink tervezésekor, a változó topológiákhoz és üzemállapotokhoz illeszkedő irányítási stratégiák kidolgozásakor a fenntartható fejlődés szempontjaira is tekintettel kell lennünk.

Az anyag-energiaáram hálózat jelenlegi üzemvitelét jellemző energiaáramok és kalorikus jellemzők dokumentálása hitelesített mérőeszközök használatát, a mérési

módszerek, a nyomás- és hőmérséklet korrekciók korrekt alkalmazását feltételezi. Az előbbiek részleges hiánya is kétségesse teheti a megbízható energiavesztés feltáró vizsgálatokat. E kockázat kezelését szolgálja a meglévő mérőhelyek, eszközök és módszerek felülvizsgálata. Az említett vízgőzáram- és hőenergia- különbözet (energiavesztés) indokolja, hogy az energiaellátásban érdekelt igényeit és lehetőségeit elemezzük. Kockázati tényező lehet a nagyszámú felhasználó egyedi igényeinek összehangolása, a fogyasztók- termelő, szolgáltató közötti vélhető érdekellentétek kompromisszumokkal járó feloldása (az energia-előállítónak, szolgáltatónak érdeke a hőenergia kiadás növelése, a fogyasztók vélhető érdeke energiaköltségeik szinten tartása). A kockázatkezelési stratégia kezdeti lépéseit a négy nagy – a jelenleg hasznosított vízgőzfelhasználás ~80%-át fogyasztó – cég igényeinek felmérése jelentené. Ezt követi a hőenergiát előállító üzemviteli stratégiájának megismerése.

Kisebb kockázatot jelenthetnek a konzorciumi partnerek közötti adatkapcsolat kiépítéskor az eltérő alkalmazói rendszerekből adódó konvertálási, adatátviteli problémák, melynek megoldása a hibaellenőrzés és a hibaelhárítás mielőbbi elhárítása. A partnerek közötti, a monitoring rendszer közös használatával kapcsolatos Megállapodás a biztosíték, hogy az identifikációs mérések során a mérés-előkészítés, az irányított beavatkozások, az adatátvitel, a monitoring rendszer üzemeltetése tervezett módon, ütemterv alapján, az energiafogyasztók tudtával és beleegyezésével valósuljanak meg. A megállapodás hiánya kétségesse tenné a kísérleti identifikáció megvalósulását. A hálózatüzemeltetési stratégia megválasztása a konzorciumi partnerek és az energia-előállító közötti konzultációt, egyeztetést igényel, melynek eredménye meghatározza a projekt további folytatását. A hálózatüzemeltetési stratégia eldöntése kockázati tényezőt jelent a projekt hosszú távú fenntarthatóságát tekintve. A kockázatkezelést az alternatívák meggyőző, számadatokkal alátámasztott bemutatása jelentené.

Hosszútávon, regionális szinten mindenképpen szükségesnek látszik a regionális energiaellátó rendszer korszerűsítése, az energiavesztések, a környezetet távlatokban károsító felesleges hőkibocsátás csökkentése. Az energiaáram-hálózat megvalósítandó infokommunikációs rendszere, s ennek kiterjesztése az energiaelosztásban a jelenlegi partnerekre és a későbbiekben jelentkező új fogyasztókra új munkahelyek, munkakörök létrejöttét eredményezi, melyek magasan kvalifikált munkaerőt, informatikai felkészültséggel rendelkező szakembereket igényelnek.

Az innovatív informatikai alkalmazást jelentő feladat megoldása és regionális hasznosítása olyan konzorcium létrehozását igényelte, melyben az ipari partner biztosítja az alkalmazott kutatáshoz szükséges háttérrel, hogy a kutatás során megszerzett tudásanyag a már meglévő eljárásokban, szolgáltatásokban jelentős javulást eredményezzen. Tanszékünk rendelkezik ugyan professzionális informatikai ismertekkel bíró szakemberekkel, de nem tartozik tulajdonába olyan összetett műszaki objektum (technológia rendszer, hálózat), mely komplexitásából adódóan igényelné az informatika eszközeit és módszereit. A műszaki informatikai ismeretek gyakorlatban való hasznosítása, a legújabb kutatási irányok (nagy rendszerek komplexitás-kezelése, mesterséges intelligencia-alkalmazások stb.) követése és gyakorlati alkalmazása nagymértékben indokolja, hogy a partner biztosítsa a technológiai rendszer elérhetőségét, az alkalmazott kutatás-fejlesztés művelésének, az infokommunikációs rendszer lehetőségének kiépítését a non-profit szervezet számára. Ugyanakkor az ipari partner számára is meghatározó igény energiavesztéseinek csökkentése, mely remélhetően a számítógépes monitoring (a későbbiekben irányítási, szakértői és döntéstámogató) rendszer kiépítésével válhat valóra.

KÖSZÖNETNYILVÁNÍTÁS

Az alkalmazott kutatás a Nemzeti Fejlesztési Terv Gazdasági Versenyképesség Operatív Program „Infokommunikációs technológia kidolgozása és hasznosítása az energiaelosztás területén ” című pályázata (GVOP-3.1.1.-2004-05-0125/3.0.) keretében valósul meg.

IRODALOM

- Szakonyi, L. (2002). Jelek és rendszerek I.-II. jegyzet. Pécs : PTE PMMK. 253.
Szakonyi, L. (2002). A műszaki informatika képzés paradoxonai. International Symposium Anniversary of Pollack Mihály College of Engineering. Pécs, 2002. máj. 31-jún. 1. 852-897.
Szakonyi, L. (2002). Műszaki rendszerek és hálózatok jegyzet. Pécs : PTE PMMK. 95.

Levelezési cím (*Corresponding author*):

Szakonyi Lajos

Pécsi Tudományegyetem Pollack Mihály Műszaki Kar

Műszaki Informatika Tanszék

7624 Pécs, Rókus u. 2.

University of Pécs, Department of Information Technology

H-7624 Pécs, Rókus u. 2.

Tel.: 36-72-503-650/3741, Fax: 36-72-501-534

e-mail: szakonyi@morpheus.pte.hu



DSP-k gyakorlati alkalmazása a folyamatirányításban

¹Gyeviki J., ²Csiszár A.

¹SZTE SZÉF Műszaki és Informatika Tanszék Szeged Mars tér 20.

²SZTE TTK Elméleti Fizika Tanszék Szeged Tisza Lajos krt. 84-86.

ÖSSZEFOGLALÁS

Munkánkban egy Texas Instruments TMS320LF2407 típusú DSP-t használtunk egy csúszó-mód szabályozású pneumatikus pozicionálás megvalósítására. A pneumatikus munkahengereket, mint fontos munkavégző elemeket széles körben alkalmazzák az ipari automatizálás területén. Ez a munkahengerek számos előnyös tulajdonságának köszönhető. Nevezetesen egyszerűek, tiszták, olcsók, nagy sebességre képesek, nagy a teljesítmény-tömeg viszonyuk, könnyű a karbantartásuk és eredendően rugalmasak. A levegő összenyomhatóságának a munkahengerben fellépő sűrűlődnak és a levegő szervoszelepen történő nemlineáris átalakulásának köszönhetően a pneumatikus szervorendszer erősen nemlineáris, variáns rendszer. A pneumatikus rendszer nemlinearitása miatt robusztus szabályozást kell alkalmazni. A csúszó-mód szabályozást a hetvenes évek végén kezdték alkalmazni robot manipulátorok irányítására.

(Kulcsszavak: DSP, csúszó-mód szabályozás, pneumatikus rendszer, pozicionálás)

ABSTRACT

Practical applications of DSPs in Process Control

¹J. Gyeviki, ²A. Csiszár

¹University of Szeged, College Faculty of Food Engineering Technical and Informatics Department, Szeged Mars tér 20.

²University of Szeged, Faculty of Natural Science Department of Theoretical Physics, Szeged Tisza Lajos krt. 84-86

In this paper, a sliding mode control is implemented on the Texas Instruments TMS320LF2407 digital signal processor (DSP) for control a pneumatic positioning system. As an important driving element, the pneumatic cylinder is widely used in industrial applications for many automation purposes thanks to their variety of advantages, such as: simple, clean, low cost, high speed, high power to weight ratio, easy maintenance and inherent compliance. The pneumatic servo-system is a very nonlinear time-variant control system because of the compressibility of air, the friction force between the piston and the cylinder, air mass flow rate through the servo-valve, etc. Because on control difficulties, caused by the high nonlinearity of pneumatic systems, a robust control method must be applied. Sliding mode control was introduced in the late 1970's as a control design approach for the control of robotic manipulators.

(Keywords: DSP, Sliding mode control, Pneumatic system, Positioning.)

BEVEZETÉS

A munkánk célja az volt, hogy megismerjük a digitális jelfeldolgozó processzorok (DSP – Digital Signal Processor) működését, felépítését és programozását, mivel a digitális jelfeldolgozás az egyik leggyorsabban fejlődő ága a modern elektronikának. Célunk volt tovább-

bá egy szabályozástechnikai feladat megoldása. Két módszert, a PID (Proportional Integral, Differential - arányos-integráló-differenciáló) szabályozást és az SMC (Sliding Mode Control – csúszó-mód szabályozást) alkalmaztunk a pneumatikus pozicionálás területén.

Az elmúlt század második felében számos szabályozási struktúra született, hogy túllépjének a PID szabályozó korlátain. Ennek ellenére az ipari szabályozások 90 százalékában még napjainkban is, az egyszerű felépítésük és könnyű használatuk miatt, PID szabályozót használnak (Gao, 2001). Természetesen napjaink PID szabályozójának felépítése már jelentősen eltér a kezdeti idők analóg PID szabályozójától. Ma a PID szabályozók digitális elven működnek, és számos olyan algoritmust használnak, ami növeli a szabályozó teljesítményét és hatékonyságát (antiwind-up, auto-tuning, adaptive, fuzzy fine-tuning). A tisztán matematikai algoritmusok mellett megjelentek a mesterséges intelligencia módszerek (Mester, 1995).

A robusztus szabályozás elmélete a bizonytalanságokat kezelni képes rendszerek analizését és tervezését jelenti, különösen olyan körülmények között, ahol egzakt matematikai modellezés nem lehetséges. A múltja viszonylag rövid és még nem tekinthető lezárt elméletnek. A változó struktúrájú rendszerek és ehhez kapcsolódva a csúszó-mód szabályozás elméletét a Szovjetunióban dolgozták ki évtizedekkel ezelőtt. Az elmélet elsősorban V.I. Utkin és D.K. Young (1977) nevéhez fűződik.

A pneumatikus munkahengereket, mint fontos munkavégző elemeket széles körben alkalmazzák az ipari automatizálás területén. Ez a munkahengerek számos előnyös tulajdonságának köszönhető. Nevezetesen egyszerűek, tiszták, olcsók, nagy sebességre képesek, nagy a teljesítmény-tömeg viszonyuk, könnyű a karbantartásuk és eredendően rugalmasak. Egy pneumatikus munkahenger dugattyúját hagyományos alkalmazásoknál csak a két véghelyzetben állítjuk meg. Az elmúlt húsz évben a pneumatikus rendszerek nagy fejlődésen mentek át. Ez a fejlődés a szervopneumatikus rendszerek modellezésében végzett intenzív kutatómunkának köszönhető. Ezzel párhuzamosan alkalmazásra kerültek a szabályozásmélet fejlődésének legújabb vívmányai is. Ennek eredményeként javultak a pozicionálási és pályakövetési tulajdonságok és a pneumatikus aktuátorok robottechnikai alkalmazásokra is alkalmassá váltak.

A kezdeti munkák a pneumatikus rendszerek modellezésében és szabályozásában az 50-es évekre vezethetők vissza Shearer (1956), Blackburn *et al.* (1960). A szervopneumatikus pozicionálás szabályozásának fejlődésén keresztül a modern szabályozástechnika fejlődését is nyomon követhetjük. A korai munkákban Shearer (1956), Burrows és Web (1966), Vaughan (1965) lineáris PID szabályozót használtak. Számos megoldás született az automatikus hangolász PID szabályozással. Fok és Ong, (1999) által elért pozicionálási pontosság ± 0.3 mm volt. Jeon *et al.* (1998) genetikai algoritmust használtak a pozíció-, sebesség- és gyorsulás visszacsatolású szabályozó optimális paramétereinek meghatározására. Dugattyúrúd nélküli munkahengerrel ± 0.1 mm pontosságot értek el. Wang *et al.* (1999) tanuló algoritmussal kiegészített PID szabályozót alkalmazott, az elért pontosság ± 1 mm. Wikander (1988) a nemlinearitás kompenzálásával 0.01 mm pozicionálási pontosságot is elért.

Fontos kutatási irány a sliding mode control (csúszó-mód szabályozás) alkalmazásának vizsgálata. Számos munka született a témában, például Song és Ishida (1997) továbbá Drakunov *et al.* (1997) foglalkoztak a csúszó-mód szabályozással és megmutatták, hogy a sliding mode control sikeresen használható a súrlódás kompenzálására. Az elért pontosság ± 0.2 mm.

A legtöbb javasolt megoldásnál az állandósult hiba nagyobb mint 0.1 mm és ez sem biztosított a dugattyú minden pozíciójában, illetve különböző nyomásértékeknél.

A KÍSÉRLETTERVEZÉS ELMÉLETI HÁTTERE

A kísérleti berendezés elkészítése előtt meg kell ismerkedünk a felhasznált eszközökkel, a pneumatikával és az alkalmazott szabályozások elméletével.

A pneumatika elméleti alapjai

A munkahenger (1. ábra) modelljének elkészítéséhez meg kell vizsgálni a levegő, fojtáson történő áramlását, és el kell készíteni egy kamra töltésének modelljét (Elek, 1975).

Az áramlás leírásához négy jellemzőt kell az áramvonal mentén meghatározni:

- sebesség (*velocity*) (w);
- nyomás (*pressure*) (p);
- hőmérséklet (*temperature*) (T);
- sűrűség (*compactness*) (ρ).

A négy jellemző meghatározásához négy egyenletet kell felírnunk:

$$\text{állapotegyenlet} \quad p \cdot v = R \cdot T, \quad (1)$$

$$\text{energiaegyenlet} \quad c_p \cdot T + \frac{w^2}{2} = \text{állandó}, \quad (2)$$

$$\text{Euler-egyenlet} \quad w \cdot dw + \frac{dp}{\rho} = 0, \quad (3)$$

$$\text{folytonossági egyenlet} \quad A \cdot w \cdot \rho = \text{állandó}. \quad (4)$$

Az alapegyenletek figyelembevételével levezethető az összenyomható ideális sűrűdésmentes gázokra érvényes Bernoulli egyenlet, illetve egy tartályból csővezetéken kiáramló levegőre vonatkozó egyenlet.

Egy kamrában uralkodó nyomás meghatározásánál figyelembe kell venni:

- a térbe beáramló és az onnan kiáramló sűrített levegő mennyiségét ($\dot{m}_{be}$, $\dot{m}_{ki}$);
- a bevezetett illetve elvezetett hő mennyiségét (dQ_a , dQ_b);
- a kamra térfogatváltozásának hatását.

1. ábra

A munkahenger vázlata

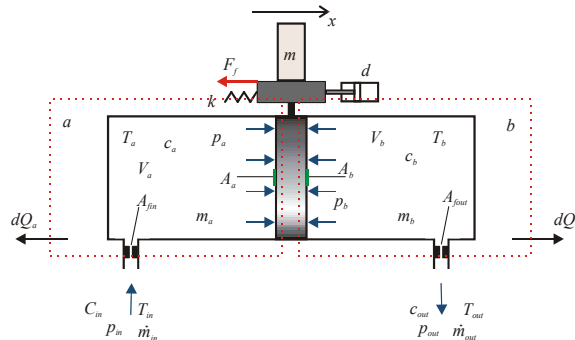


Figure 1: Structure of the pneumatic cylinder

Az energiaegyenlet kimondja, hogy egy rendszerben az energia nem veszt el, s nem is keletkezhet, csak átalakulhat. Azaz a tér egy vizsgált részében a be- és az elvezetett energia különbsége tárolódik. Az $U=c_v \cdot m \cdot T$, $h=c_p \cdot T$ és $dU=c_v \cdot (dm \cdot T + dT \cdot m)$ $\kappa = \frac{c_p}{c_v}$

összefüggések felhasználásával írjuk fel az energiaegyenletet a munkahenger a terére:

$$dU_a = dm_{be} \cdot \left(h_{be} + \frac{w_{be}^2}{2} \right) - dQ_a + W_a \quad (5)$$

ahol

$$dm_{be} \cdot \left(h_{be} + \frac{w_{be}^2}{2} \right) \quad \text{a beáramló levegő által bevitt energia (} h \text{ az entalpia),}$$

$$W_a = -p_a \cdot A_a \cdot x \quad \text{a levegő által végzett mechanikai munka,}$$

$$dQ_a \quad \text{a rendszer határán átlépő összes hőenergia.}$$

Az általunk végzett kísérletekben a levegő áramlási sebessége olyan mértékű, hogy a $h_{be}=c_p \cdot T \gg \frac{w^2}{2}$ feltételezéssel élhetünk. Ha a továbbiakban a kinetikus energiát elhanyagoljuk az elkövetett hiba kisebb mint 2%. Ha a vizsgált kamra teljesen hőszigetelt, akkor $dQ=0$ és adiabatikus állapotváltozásról beszélünk.

$$\frac{c_v(p_a \dot{V}_a + \dot{p}_a V_a)}{R} = c_p T_{in} \dot{m}_{in} - p_a A_a \dot{x} \quad (6)$$

Ebből a nyomásváltozás kifejezhető:

$$\dot{p}_a = RT_{in} \frac{c_p \dot{m}_{in}}{c_v V_a} - p_a A_a \frac{c_v}{c_v V_a} \dot{x} \quad (7)$$

Az energiaegyenletet a munkahenger b terére is felírva, a megfelelő behelyettesítéseket és egyszerűsítéseket elvégezve a hengertér nyomásfüggvénye meghatározható.

$$\dot{p}_b = p_a A_a \frac{c_p}{c_v V_a} \dot{x} - RT_b \frac{c_p \dot{m}_{out}}{c_v V_b} \quad (8)$$

Ahol

$$V_a = V_{a0} + A_a \cdot x$$

$$V_b = V_{b0} + A_b \cdot x$$

Mivel a hengerterekben uralkodó nyomás változási sebessége a be- illetve a kilépő levegő tömegáramától függ, meg kell határoznunk $\dot{m}_{be}$ illetve $\dot{m}_{ki}$ értékét.

$$\dot{m}_{be} = \mu_{fa} \cdot p_{be} \cdot A_{fa} \cdot \sqrt{\frac{2}{R \cdot T_{be}}} \cdot \Psi \quad (9)$$

Ahol μ_{fa} az átfolyási tényező, mely minden olyan hatást figyelembe vesz, mely az ideális viszonytól való eltérésből adódik.

$$\text{Ha } p_a/p_{be} \leq 0.528 \quad \Psi = \left(\frac{2}{\chi + 1} \right)^{\frac{1}{\chi-1}} \cdot \sqrt{\frac{\chi}{\chi + 1}} = 0.484 \quad (10)$$

$$\text{Ha } p_a/p_{be} > 0.528 \quad \Psi = \sqrt{\frac{\chi}{\chi-1} \cdot \left[\left(\frac{p_a}{p_{be}} \right)^{\frac{2}{\chi}} - \left(\frac{p_a}{p_{be}} \right)^{\frac{1+\chi}{\chi}} \right]} \quad (11)$$

Atmoszférába való közvetlen kiáramlásra vonatkozóan a következő egyenleteket írhatjuk fel:

$$\dot{m}_{ki} = \mu_{fb} \cdot p_{ki} \cdot A_{fb} \cdot \sqrt{\frac{\chi}{R \cdot T_{ki}}} \Psi' \quad (12)$$

$$\text{Ha } p_b/p_{ki} \geq 1.885 \quad \Psi' = \left(\frac{2}{\chi + 1} \right)^{\frac{\chi+1}{2(\chi-1)}} \cdot \frac{p_b}{p_{ki}} = 0.578 \cdot \frac{p_b}{p_{ki}} \quad (13)$$

$$\text{Ha } p_b/p_{ki} < 1.885 \quad \text{akkor } \Psi' = \sqrt{\frac{2}{\chi-1} \cdot \left(\frac{p_b}{p_{ki}} \right)^{\frac{\chi-1}{\chi}} \cdot \left[\left(\frac{p_b}{p_{ki}} \right)^{\frac{\chi-1}{\chi}} - 1 \right]} \quad (14)$$

Ezek alapján elkészíthetjük a feladat SIMULINK modelljét.

PID szabályozás

A jelentős átalakulás ellenére a **PID** szabályozási törvény változatlan maradt. A három hangolási paraméter (K_p, K_i, K_d) fizikai összefüggésben áll a hibajellel (e), és nincs szükség a szabályozott szakasz modelljére. Az **I** integráló hatás a múlt ($\int edt$), a **P** arányos hatás a jelen (e) és a **D** differenciáló hatás a jövő hibájával ($\dot{e}$) arányos beavatkozó jelet szolgáltat.

$$u = K_p \cdot e + K_i \int edt + K_d \cdot \dot{e} \quad (15)$$

Az utóbbi időben jelentősen megnőtt az érdeklődés a nemlineáris szabályozástechnika iránt, ugyanis egyre olcsóbbak és egyre hatékonyabbak a szabályozóként használt digitális eszközök, így korábban az iparban szinte kizárólagosan használt PID (arányos integráló differenciáló) szabályozókat egyre bonyolultabb szabályozási algoritmusok váltják fel. Megfigyelhető az a tendencia is, hogy a drága precíziós mechanikát egyre olcsóbb, egyszerűbb és robusztusabb mechanikai szerkezetekkel helyettesítik, és az ebből adódó pontatlanságokat egyre bonyolultabb (nemlineáris, adaptív, robusztus) szabályozási algoritmussal küszöbölik ki. A népszerű PID szabályozók nem alkalmazhatóak a változó paraméterű és változó terhelésű nemlineáris rendszerekre. A terhelés változása, a munkahenger orientációja, lökete, a henger mérete és a henger és dugattyú közötti súrlódás csak néhány tulajdonság, amit negatívumként említhetünk a szervopneumatika kapcsán. A pneumatikus rendszer nemlinearitása miatt robusztus szabályozást kell alkalmazni.

SMC szabályozás

A szabályozás nagy előnye, hogy nem szükséges hozzá a rendszer pontos modelljének ismerete. Elegendő, ha egy egyszerűsített modell alapján mind a zavarok, mind a rendszer-paraméterek változásának korlátait ismerjük. A zavarérzéketlenség ára a végtelen nagy kapcsolási frekvencia, vagyis az ismételt gyors beavatkozás. Ebből következik, hogy a valóságban ideális csúszó-mód szabályozás nem létezik, de az a gyakorlati igénynek megfelelő módon megközelíthető.

Az erősítés a visszacsatoló ágban, két érték között változik, attól függően, hogy az adott pillanatban egy kijelölt állapotváltozó milyen értéket vesz fel. A kapcsolgató szabályozási függvény célja, hogy az állapottérben a nemlineáris szabályozott szakasz állapot trajektóriáját egy – a tervező által előre meghatározott – felületen végigvezesse és a nemlineáris szabályozott szakasz állapotváltozóit minden pillanatban a felületen tartsa. Ezt a felületet kapcsoló felületnek (switching surface) nevezzük. A visszacsatoló ág

erősítése más akkor, amikor a szabályozott szakasz állapot trajektóriája a kapcsolási felület „felett” van és más, amikor az állapot trajektória a kapcsolási felület „alatt” van. Ideális esetben, ha a szabályozott szakasz állapot trajektóriája eléri és átmetszi ezt a felületet, akkor a szabályozás az ezt követő időszakban a trajektóriát a csúszófelületen tartja és vezeti végig.

A csúszó-mód szabályozó megtervezése három fő lépésből áll.

Az első lépés a csúszófelület tervezése, a második lépés egy olyan szabályozási törvény kiválasztása, amely az állapotváltozók trajektóriáját a csúszófelületre kényszeríti, majd azon tartja, a harmadik a legfontosabb lépés a csattogásmentes (chattering free) megvalósítás. Az első fázis kritikus a VSC tervezésében. Fontos, hogy megfelelően definiáljuk a kapcsoló felületet, mivel a szabályozott szakasz ezen a felületen halad végig. A második fázisban azt a kapcsoló szabályozást tervezzük meg, amely a szabályozott szakaszt a csúszófelületre vezeti, majd ott tartja ha a szabályozott szakasz trajektóriája a felületet átmetszette. Ehhez Ljapunov második, vagy direkt módszerét használjuk.

A legegyszerűbb szabályozó elem, amely a csúszó-módot megvalósíthatja, a relé.

$$u_i = k_i \cdot \text{sign}(s_i) \quad (16)$$

A csúszó-mód szabályozó robusztusságát a nagy (ideális esetben végtelen) körerősítésének köszönheti. Erről a relé gondoskodik. Fontos megjegyezni, hogy a valóságban csak véges kapcsolási frekvenciát tudunk megvalósítani. Ez a csattogás jelenségét (chattering) hozza magával. Ez azt jelenti, hogy a szabályozott szakasz állapot trajektóriája a csúszófelület egy kicsiny környezetében perturbál (2. ábra).

2. ábra

Csúszó mód az állapot térben

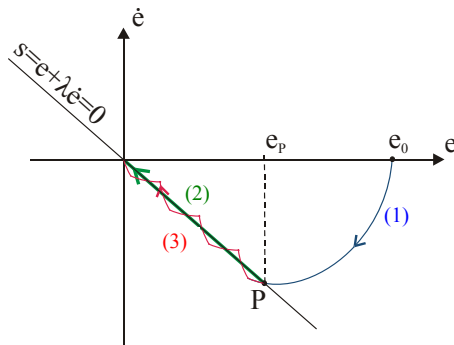


Figure 2: Sliding mode in the state space

Denotes the approaching phase(1), Denotes the sliding phase(2), Denotes the chattering(3)

Ahol (1) a megközelítés fázisa, (2) az ideális csúszó-mód és (3) a csattogás jelensége. A szabályozó realizálásánál figyelembe kell venni azt, hogy ez a perturbáció esetleg nem kívánt jelenségeket okoz a szabályozott szakaszban.

A megközelítési fázisban a rendszer érzékeny a paraméter változásokra, illetve a külső zavarásokra. Amikor a rendszer csúszó-módba kerül, dinamikáját a csúszófelület

határozza meg, és ekkor érzéketlenné válik a paraméter változásokkal, illetve a külső zavarásokkal szemben. Tervezéskor két ellentétes szempontot kell szem előtt tartani: a rendszer trajektóriája minél rövidebb idő alatt érje el a csúszóegyenest, és a csattogást minél inkább elimináljuk. A kutatók nagy energiát szenteltek a kérdés megoldására. *Young et al.* (1987) a megközelítés idejét a visszacsatoló ág erősítésének növelésével csökkentették, de ez fokozta a csattogást. A csattogás csökkentésére *Slotin* (1984) valamint *Yeung és Chen* (1988) határrejteget képzett a csúszóegyenest mentén. *Hwang és Lin* (1992), *Lin és Kung* (1992), valamint *Lin és Chen* (1994) fuzzy logika alkalmazásával csökkentették a csattogást.

A csúszó-módról jó bevezetés találunk *Korondi et al.* (1999) és *Jeon et al.* (1998) munkáiban. Tegyük fel, hogy rendszerünket egy másodrendű differenciálegyenlet írja le:

$$\ddot{x} = f(x, \dot{x}, u) \quad (17)$$

Ahol $x(t)$ a rendszer kimenete, a beavatkozó jelet jelölje $u(t)$, mely jel az $x(t)$ kimenőjelet arra kényszeríti, hogy kövesse az $x_d(t)$ referencia jelet. A követési hibát jelöljük $e(t)$ -vel. A rendszert akkor gyorsítjuk, ha állapota lemaradt a tervezett trajektóriától és akkor lassítjuk, ha megelőzi azt. A következő kérdés az, hogy miként tervezzük meg a hibajel trajektóriáját, és hogyan határozzuk meg a rendszer tényleges állapotának távolságát a tervezett trajektóriától, vagyis azt, hogy mikor kell gyorsítani, illetve lassítani a rendszert. A klasszikus csúszó-mód szabályozásnál (*Yin et al.*, 1998) a hibajel (e) és deriváltjának ($\dot{e}$) lineáris kombinációjával definiálunk egy olyan skalár változót, amely úgy értelmezhető, mint a rendszer állapotának előjeles távolsága egy hiperfelülettől ($n=2$ esetben egyenestől).

$$s = e + \lambda \cdot \dot{e} \quad (18)$$

A (18) egyenlet a fázissíkon egy $-1/\lambda_i$ meredekségű egyenest határoz meg. Ezért szokás csúszóegyenestről beszélni. A 2. ábrán egy képzeletbeli trajektóriát is berajzoltunk. A $t=0$ időpontban a kezdeti értékek: $e=e_0$ és $\dot{e}=0$.

Ahol λ egy időállandó típusú konstans paraméter. A megközelítési fázisban az (19) feltétel

$$s(\dot{e}, e) \cdot \dot{s}(\dot{e}, e) < 0 \quad (19)$$

kielégítésével tudjuk a hiba-trajektóriát a csúszó felület felé vezetni. A csúszási fázisban létezik egy folytonos jel, az ún. u_{eq} egyenértékű (ekvivalens) szabályozójel, amely a csúszófelületen tudja tartani a rendszert, (de nem feltétlenül gondoskodik arról, hogy az állapottrajektória csúszófelülethez konvergáljon, ha az, attól távol van). A két szabályozási törvényt külön-külön kell megtervezni.

Csúszó-módról akkor beszélünk, ha a rendszer állapota a csúszó felület mentén változik:

$$s(\dot{e}, e) = 0 \quad (20)$$

A csúszó-módban történő szabályozáskor az a feladat, hogy a beavatkozójelet mindig úgy kapcsoljuk, hogy a (19) feltétel minden időpillanatban teljesüljön. A szabályozott szakasról és a zavarásról csak annyi információval kell rendelkezni, hogy adott beavatkozójelet mellett a (19) feltétel meglepte egyértelműen eldönthető legyen.

Általában a negatív visszacsatolás miatt a hibajel fázisterében az origó egy meghatározott környezetén belül az u_b beavatkozójelet és az $\dot{s}$ előjele ellentétes. Így gyakran elegendő egy olyan relés szabályozót alkalmazni, amelyik az s skalár változó előjelétől függően kapcsolgatja az u_b beavatkozójelet:

$$u_b = \delta \cdot \text{sign}(s), \quad (21)$$

ahol δ a szabályozó körerősítése által meghatározott pozitív konstans.

A csúszó egyenes meredekségének megválasztásakor két egymás ellen ható követelmény között kell az optimumot megtalálni. Minél kisebb az egyenes meredeksége, annál hamarabb éri el a trajektória a csúszó egyenest, de annál lassabb az egyenes mentén a beállítás. Jeon *et al.* (1998) a gyorsaság és a robusztusság követelményének együttes kielégítése érdekében a csúszó egyenes meredekségének adaptív változtatását javasolják.

A KÍSÉRLETI BEREDEZÉS ISMERTETÉSE

Ahhoz hogy az elméleti megfontolásokat kísérletekkel is alátámasszuk, szükség volt egy olyan pneumatikus pozicionáló szerkezet elkészítésére, mely alkalmas kutatási munkánkhoz szükséges mérések és kísérletek elvégzésére. A berendezéssel szemben támasztott követelmények (Gyeviki, 2004):

- a berendezés kereskedelmi forgalomban beszerezhető alkatrészekből épüljön fel;
- a berendezés tegyen eleget a kutatási program során felmerült követelményeknek;
- a berendezés legyen univerzális és könnyen kezelhető;
- a pneumatikus pozicionálás vizsgálata mellett a berendezés legyen alkalmas hallgatói mérési gyakorlatok végzésére is.

A 3. ábrán látható kísérleti berendezés fő részei: MECMAN 170 típusú 32 mm átmérőjű, 500 mm löketű siklóhenger, melyet egy FESTO MPYE-5-1/8 HF-010B típusú arányos szeleppel vezérelünk. A dugattyú pozícióját egy LINIMIK MSA 320 típusú inkrementális útdóval mérjük. A sebességet és a gyorsulást számítjuk. A kamrák nyomásának mérése nyomás-távadókkal történik.

3. ábra

Kísérleti berendezés

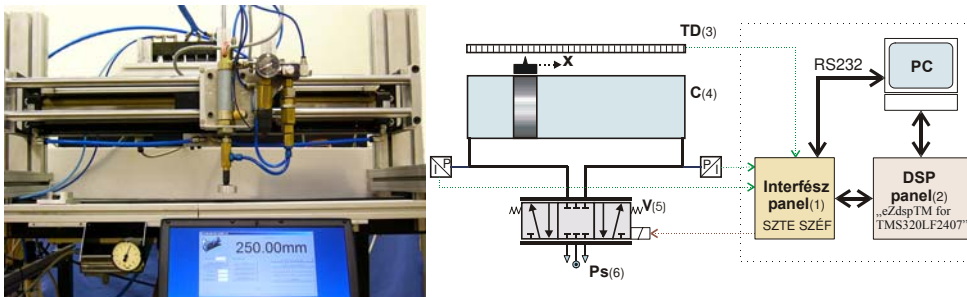


Figure 3: The experimental positioning system

Interface board(1), DSP board(2), Llinear transducer(3), Rodless cylinder(4), Servo-valve(5), Air supply(6)

A berendezés alkalmas függőleges orientációjú munkahengerrel (4. ábra), tömeg- és rugó terheléssel végzett kísérletekre is (Gyeviki, 2004).

A pneumatikus rendszerek nemlinearitása által okozott szabályozási nehézségek miatt egy DSP bázisú csúszó-mód szabályozót terveztünk. Kísérleteinkhez a Spectrum Digital „eZdspTM for TMS320LF2407” DSP kártyáját használtuk. A DSP kezdő-készlet (DSK) a PC paralel portjához kapcsolható, melyen keresztül letölthető a program a DSP-

be, illetve ezen keresztül lépésenként ellenőrizhető a letöltött program. A szabályozó programot „C” nyelven készítettük.

Az eZdspTM LF2407 Starter Kit meghatározó része egy TMS320LF2407-es DSP melyet a Texas Instruments forgalmaz. A modul kiváló lehetőséget biztosít, hogy megismerjük a DSP-t és kisebb alkalmazásokat fejlesszünk.

4. ábra

Függőleges pozicionálás

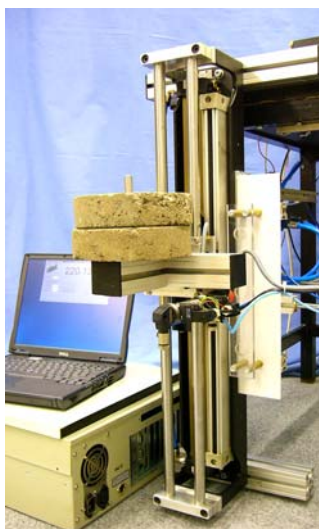


Figure 4: Vertical positioning

Ahhoz hogy a DSP-t szabályozástechnikai feladatok ellátására használjuk, a kimenetek és bemenetek illesztésére van szükség úgy hogy, azok kompatibilisek legyenek a külső eszközök jelszintjével (Csiszár *et al.*, 2004). Az általunk kiválasztott szabályozástechnikai feladat egy pneumatikus munkahenger pozíció szabályozása. Ennek megfelelően terveztük meg az illesztő-panelt (5. ábra).

A DSP analóg bemeneti és digitális ki- és bementi interfésszel rendelkezik. A kísérletünk során szükség volt még analóg kimeneti jelekre is, ezeket D/A konverterrel állítottuk elő. A digitális kimeneti jelszint illesztését relével, a bemenet optikai csatolóval oldottuk meg. A kiegészítő áramkör elkészítésénél szükség volt még a számítógéppel és különböző ipari eszközökkel való kapcsolattartásra.

Kísérleteink során a pozíció érzékelésre egy 0.01 mm-es pontosságú (LINIMIK MSA 320 típusú) inkrementális útdát, a pontosabb 0.001 mm-es mérésekhez pedig Balluff BTL5-S101 abszolút útdát használtunk.

Az általunk választott Balluff távadóval RS485 vagy RS422 soros vonalon kommunikálhatunk. A DSP digitális kimenete, bemenete és az RS485 soros vonal közötti illesztést egy MAX488 típusú IC-vel oldottuk meg.

A számítógéppel való kommunikáció kialakítására leggyakrabban a RS232 kommunikációt alkalmazunk. A DSP és a számítógép közötti RS232-es kommunikációt egy ICL3232IBN integrált áramkörrel valósítottuk meg.

5. ábra

A DSP panel és a illesztő-panel kapcsolata

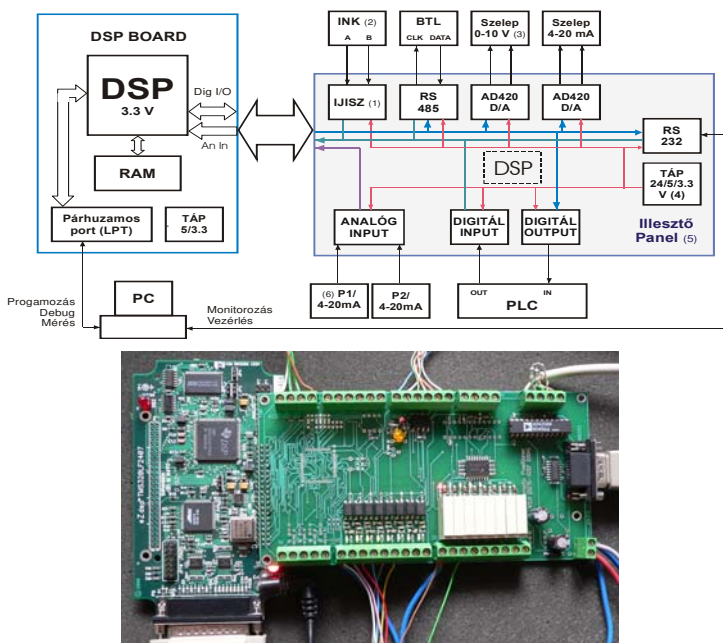


Figure 5: Connection between the DSP Board and interface circuit

A pozíció szabályozásánál nagy hangsúlyt kap a pozíció pontos mérése és gyors beolvasása. A feladat megoldása során mi a DSP saját fejlesztő rendszerét a Code Composer-t alkamtunk. A program a DSP egy kimenetét és egy bemenetét használja a szinkron soros kommunikáció kialakítására. A kimenet a BTL útdónak a szinkron jelet adja, a bemenet az útdótól jövő adatot fogadja.

EREDMÉNY ÉS ÉRTÉKELÉS

A pneumatikus pozicionálási kísérleteink során két szabályozás viselkedését vizsgáltuk. Ezért a PID és az SMC szabályozás mérési eredményeit vetjük össze.

A kísérletek során a nyomás értéke 6 bar, a mintavételezési idő pedig 2 ms volt. A kísérleti eredmények valós idejű gyűjtésére egy adatgyűjtő programot készítettünk, mely egy kommunikációs interfészen keresztül a DSP-ből a PC-be juttatja a mért értékeket. A szabályozó program a DSP programmemóriájában helyezkedik el, így a szabályozás a PC-től függetlenül működik. Mivel a DSP nagy működési sebességgel rendelkezik a mintavételezési frekvencia és a pontosság növelhető. A DSP alkalmas az inkrementális jeladó jelének, illetve analóg és digitális jelek fogadására. A mérés során egy jól behangolt PID szabályozóval és egy csúszó-mód szabályozóval végeztünk összehasonlító kísérleteket.

A pozicionálási időfüggvények a 6. és 8. ábrákon láthatóak. PID szabályozó esetén jelentős 16.3 mm-es túllendülést mértünk. A jól ismert csúszás-megakadás jelensége miatt az állandósult állapotbeli hiba is jelentős: 3.8 mm. A DSP bázisú csúszó-mód szabályozás pozicionálási hibája kisebb, mint 0.01 mm. Ez kisebb mint 1%-a a jól

behangolt PID szabályozó pozicionálási hibájának. A csúszó egyenes meredekségét jelentő időállandót 0.08 sec-ra választottuk. A fázisrajekatóriákat a 7. és a 9. ábrákon láthatjuk. Az 0.01 mm felbontású inkrementális jeladóval csak azt tudtuk megállapítani, hogy a pozicionálási hiba 0.01 mm alatt van. Ahhoz hogy a tényleges pozicionálási pontosságot meg tudjuk határozni, egy BLT5-S101 típusú, 0.001 mm felbontású útdót (Balluff) építettük be a kísérleti berendezésünkbe. A szabályozás finomításával elértük a 10. ábrán látható a pozicionálási pontosságot. A 11. ábrán kinagyítva láthatjuk a túllendülés mértékét, illetve az állandósult hibát. A túllendülés mértéke 0.3 mm alatt, a pozicionálás hibája pedig 1 μ m alatt van.

6. ábra

Pozicionálás PID szabályozóval

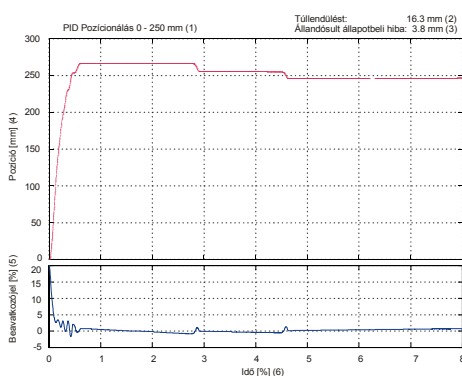


Figure 6: Positioning with PID

7. ábra

PID szabályozás fázis-trajektóriája

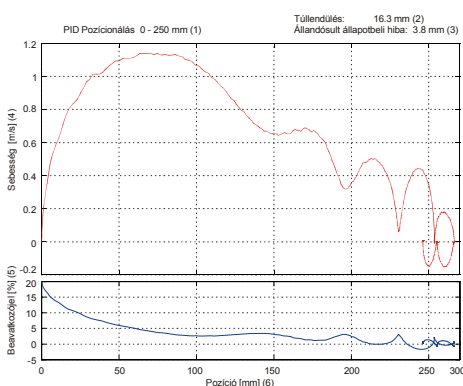


Figure 7: Phase-plot trajectory of PID

8. ábra

Pozicionálás SM szabályozóval

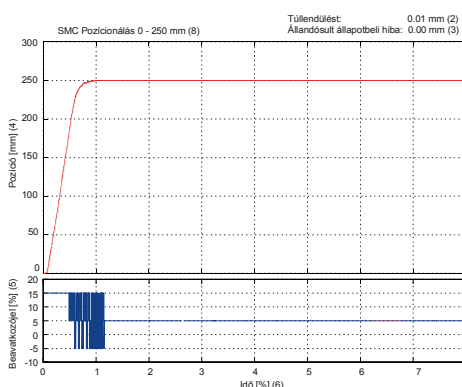


Figure 8: Positioning with SMC

9. ábra

SM szabályozás fázis-trajektóriája

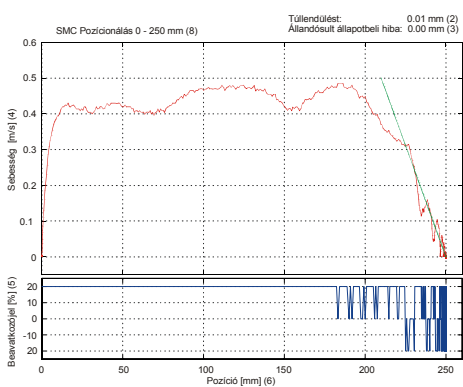


Figure 9: Phase-plot trajectory of SMC

PID positioning(1), Overshoot(2), Steady state error(3), Position(4), Control signal(5), Time(6), Velocity(7), SMC positioning(8)

10. ábra

Az 1 μ m-es pozicionálás pozíció-idő diagramja

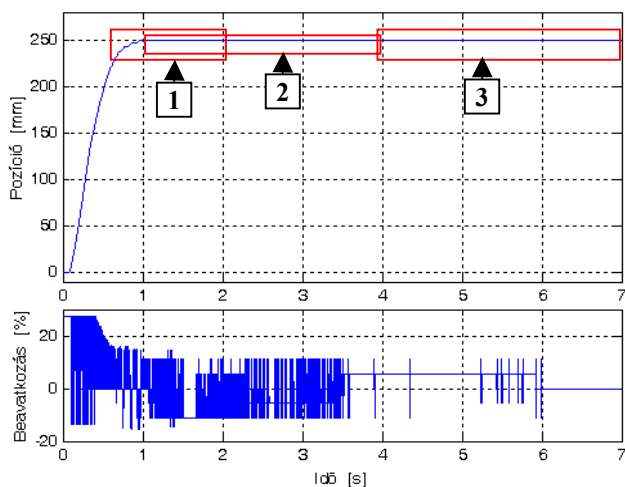


Figure 10: Diagram of the 1 μ m positioning

11. ábra

1 μ m-es pozicionálás pozíció-idő diagramjának részletei

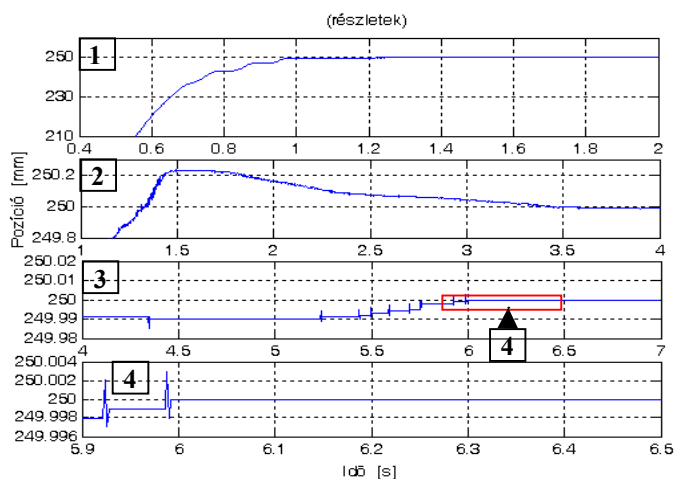


Figure 11: Details of position-time diagram of the 1 μ m positioning

ÖSSZEFOGLALÁS

A laboratóriumi kísérletek alapján megállapíthatjuk, hogy a DSP alapú csúszó-mód szabályozás alkalmas a pneumatikus munkahengerek szabályozására, ugyanis kellően

gyors, robusztus a külső zavarásokkal és belső paraméter-változásokkal szemben. Fontos annak a hangsúlyozása is, hogy a szabályozáshoz nem szükséges a rendszer pontos modelljének ismerete.

IRODALOM

- Blackburn, J.F., Rethof, G., Shearer, J.L. (1960). Fluid Power Control. MIT Press : Cambridge, MA.
- Csiszár, A., Varga, A., Kovács, E. (2004). Servo-pneumatic positioning with sliding mode control (SMC). International Student Experimental Hands-on Project Competition via Internet on Intelligent Mechatronics and Automation, Taiwan
- Drakunov, S., Hanchin, G.D., Su, W.C., Özgüner, Ü. (1997). Nonlinear control of rodless pneumatic servoactuator, or sliding mode versus coulomb friction. In: Automatica, 33. 7. 1401-1408.
- Elek, I., Hudáky J. (1975). Az ipari pneumatika alapjai. Interpress Kiadó : Budapest
- Fok, S.C., Ong, E.K. (1999). Position control and repeatability of a pneumatic rodless cylinder system for continuous positioning. In: Robotics and Computer Integrated Manufacturing, 15. 365-371.
- Gao, Z., Yi, H., Han, J. (2001). An alternative paradigm for control system design. In: Proc. of IEEE Conference on Decision and Control
- Gyeviki, J., Tóth, I.T., Rózsahégyi K. (2004) Sliding mode control and its Application on a Servopneumatic Positioning System. In: Transactions on AUTOMATIC CONTROL and COMPUTER SCIENCE 49. 63. 2004. 99-103.
- Gyeviki, J. (2004) Improving Positioning Accuracy of Pneumatic Systems. In: Gép/A Gépípari Tudományos Egyesület Műszaki Folyóirata, 55. 9. 7-9.
- Hwang, G.C. and Lin, S.C. (1992). A stability approach to fuzzy control design for nonlinear systems. In: Fuzzy Sets and Systems, 48. 279-287.
- Jeon, Y., Lee, C. and Hong, Y. (1998). Optimization of the control parameters of a pneumatic servo cylinder drive using genetic algorithms. In: Control Engineering Practice, 6. 847-853.
- Korondi, P., Hashimoto, H., Utkin V., (1998). Direct Torsion Control of Flexible Shaft based on an Observer Based Discrete-time Sliding Mode. In: IEEE Trans. on Industrial Electronics IE-2. 291-296.
- Korondi, P., Hashimoto, H., (2000). Sliding Mode Design for Motion Control. In: Studies in Applied Electromagnetics and Mechanics 16. 1-12.
- Korondi, P., Hashimoto, H. (1999) Park Vector Based Sliding Mode Control. In: Variable Structure System, Robust and Nonlinear Control. K.D.Young, Ü. Özgüner (eds.), ISBN: 1-85233-197-6, Springer-Verlag.
- Lin, S.C., Kung, C.C. (1992). The fuzzy-slidin mode controller, In: Proceedings of the 15th National Symposium on Automatic Control R.O.C., 359-366.
- Lin, S.C., Chen, Y.Y. (1994). Design of adaptive fuzzy sliding mode for nonlinea system control. In: Proceedings of Third IEEE International Conference on Fuzzy Systems USA, 35-39.
- Mester, Gy. (1995). Neuro-Fuzzy-Genetic Controller Design for Robot Manipulators. In: Proc. IECON'95, IEEE, Orlando, Florida, USA, 1. 87-92.
- Mester, Gy. (1995). Neuro-Fuzzy-Genetic TrackingControl of Flexible Joint robots. In: Proc. I.Intern.Conf. on Adv. Robotics & Intelligent Aut. Athens, Greece 93-98.
- Shearer, J.L. (1956). Study of pneumatic process on the continuous control of motion with compressed air. In: Transactions of ASME, 78. 233-249.

- Slotine, J.J., (1984). Sliding controller design for non-linear systems. In: Int. J. Control, 40. 421-434.
- Song, J. and Ishida, Y., (1997) Robust sliding mode control for pneumatic servo systems. In: International Journal of Engineering Science, 35. 8. 711-723.
- Utkin, V., (1977). Variable Structure Systems with Sliding Mode. In: IEEE Trans. AC-22. 2. 212-222.
- Wang, J., Pu, J. and Moore, P. (1999). Accurate position control of servo pneumatic actuator systems: an application to food packaging. In: Control Engineering Practice, 7. 699-706.
- Wikander, J., (1988). Adaptive Control of Pneumatic Cylinders. Doctoral thesis, Royal Institute of Technology, Stockholm, 1988, ISSN 0282-0048, TRITAMAE-1988-7
- Yeung, K.S. and Chen, Y.P., (1988) A new controller design for manipulators using the theory of variable structure systems. In: IEEE Trans. on Automatic Control, AC-33. 200-206.
- Young, K.D., (1987). Controller Design for Manipulator using Theory of Variable Structure Systems. In: IEEE Trans. On System, Man and Cybernetics, SMC-8. 101-256.
- Yin, Y. and Araki, K. (1998). Modelling and analysis of an asymmetric valvecontrolled single-acting cylinder of a pneumatic force control system. In: Proceedings of the SICE Annual Conference, 109.

Levelezési cím (*Corresponding author*):

Gyeviki János

Szegedi Tudományegyetem Élelmiszeripari Főiskolai Kar
H-6724 Szeged, Mars tér 7.

University of Szeged, Technical and Informatic Department
H-6724 Szeged, Mars tér 7.

Tel.: 36-62-426-549, Fax: 36-62-420-991

e-mail: gyeviki@szef.u-szeged.hu



Desztillációs kolonna R-gráf alapú szuperstruktúrája és MINLP modellje

Czuczai B., Farkas T., Rév E., Fonyó Zs., Lelkes Z.

Budapesti Műszaki és Gazdaságtudományi Egyetem, Vegyipari Műveletek Tanszék, 1521 Budapest, Műegyetem rkp. 3.

ÖSSZEFOGLALÁS

Új modellt fejlesztettünk desztilláló oszlopok szintézisére és optimalizálására. A fejlesztés első lépéseként GDP-modellt alkottunk, amit olyan MINLP-modellé alakítottunk, amely már csak a megengedett struktúrákat reprezentálja, és minimális számú bináris változót használ a struktúrák megkülönböztetésére. A kapott MINLP-modellt összehasonlítottuk a talált irodalmi modellekkel (Viswanathan és Grossmann, 1993; Yeomans és Grossmann, 2000). Az új modell kevesebb bináris változót használ és szignifikánsabban gyorsabb.

(Kulcsszavak: MINLP, desztilláció, szuperstruktúra, folyamatszintézis)

ABSTRACT

Distillation column superstructure and MINLP-model based on R-graph representation

B. Czuczai, T. Farkas, E. Rév, Zs. Fonyó, Z. Lelkes

Budapest University of Technology and Economics, Department of Chemical Engineering

A new model has been developed for distillation column synthesis and optimization based on R-graph represented superstructure. A GDP model is generated first; this is transformed into an MINLP model which represents only the considered structures, and uses minimum number of binary variables to make distinctions between the different structures. The MINLP model is compared to literature models (Viswanathan and Grossmann, 1993; and Yeomans and Grossmann, 2000). The new model uses less number of binary variables, and is significantly faster.

(Keywords: MINLP, distillation, superstructure, process synthesis)

BEVEZETÉS

A vegyipari folyamatok tervezésekor általában a minimális költségű folyamat megtalálása a cél. Komoly ipari jelentőséggel bíró probléma desztillálókolonnnak tartalmazó folyamatok optimalizálása.

A desztillálókolonnnak olyan oszlopok, melyeket többkomponensű folyadékelegyek elválasztására alkalmaznak. Az elválasztás minőségét, de a folyamat költségét is befolyásolják az üzemeltetési paraméterek. Konkrétan az oszlop tetejéről távozó pára lekondenzáltatása után kapott folyadéknak az oszlop tetejébe visszavezetett aránya - az úgynevezett refluxarány -, és az, hogy az oszlop milyen számú ún. egyensúlyi tányért tartalmaz. Ugyanolyan tisztasági követelmények mellett kisebb tányérszámhoz nagyobb refluxarány szükséges, és fordítva. A tányérszámtól függ a fix (beruházási) költség, a

refluxaránytól a változó (üzemeltetési) költség. A kolonna modellezésekor az egyensúlyi tényérok létét ill. nemlétét logikai változókkal írjuk le, ezért a tervezés olyan szintézisfeladat, melyet a folytonos változók mellett egészértékű változók is jellemeznek.

A struktúraorientált folyamatszintézis lényege, hogy egyszerre optimalizáljuk a tervezési és üzemeltetési paramétereket. Ennek során először létrehozunk egy szuperstruktúrát, amely az összes, mérnöki szempontból megengedett struktúrát tartalmazza. Ezután előállítjuk a szuperstruktúra matematikai modelljét, ami általában egy vegyes egészértékű nemlineáris programozási (Mixed Integer Nonlinear Programming, MINLP) feladat. Végül ezt a matematikai modellt optimalizáljuk.

Egy MINLP-feladat általános alakja a következő:

$$\begin{aligned} \min & \left[c^T \cdot \bar{y} + f(\bar{x}) \right] \\ & h(\bar{x}) = 0 \\ & g(\bar{x}) \leq 0 \\ & A \cdot \bar{x} = a \\ & B \cdot \bar{y} + C \cdot \bar{x} \leq d \end{aligned} ,$$

ahol $\bar{x}$ a folytonos változók, $\bar{y}$ a bináris változók vektora, $c^T \cdot \bar{y} + f(\bar{x})$ pedig az optimalizálandó célfüggvény. Az egyenlőség típusú korlátok általában mérlegegyenleteket reprezentálnak, az egyenlőtlenség típusúak különböző, a folyamatra vonatkozó előírásokat. Az ilyen feladatok megoldásának nehézségei általában meredeken növekednek a bináris változók számával. Ezért a bináris változók számának csökkentése kulcsszerepet játszik az ilyen típusú problémák megoldási módszereinek kidolgozásakor.

Az irodalomban két MINLP modell található desztillációs kolonnák szintézisére: *Viswanathan és Grossmann* (1993), valamint *Yeomans és Grossmann* (2000) modellje. Az utóbbi közleményben egy általánosított diszjunktív programozási (Generalized Disjunctive Programming, GDP) modellt publikáltak, de GDP szolver hiányában ezt MINLP modellé alakítottuk, és ezt optimalizáltuk.

Ezek a modellek annak érdekében, hogy minden struktúra reprezentálható legyen, minden tényérhez tartozik egy bináris változó.

Farkas és mtsai. (2004) definiálták a binárisan minimális MINLP reprezentációt, amely minimális számú bináris változót alkalmaz a struktúrák reprezentálására. Ha k darab struktúra létezik, akkor azt a legkisebb n számú bináris változót kell használni, amelyre igaz, hogy $n \geq \log_2 k$.

Munkánk célja az volt, hogy az irodalomban leírt modelleknél jobb numerikus tulajdonságokkal rendelkező modellt alkossunk, mely alkalmas rövid számítási időn belül optimális megoldást nyújtani, és ezzel részmodellként használható legyen desztilláló kolonnákat tartalmazó folyamatok szintézisének és optimalizálásakor.

ANYAG ÉS MÓDSZER

Szuperstruktúra és R-gráf reprezentáció

Egy maximum 15 egyensúlyi tényért tartalmazó hagyományos desztillációs oszlop szuperstruktúrája látható az *1. ábrán*. Az egyensúlyi tényérokat sötétített oválisok, az egységeket téglalapok jelölik. A kis körök az input és output portok, melyeket az áramok kötnek össze. A betáplálás (Feed), a termékek (Dist és Bot), a kondenzátor (Cond), a

visszaforraló (Reb) és a betáp egység állandó egységek, a többi egység feltételes. Alulról számolva egy szekció k -dik egységében lévő egyensúlyi tényérok száma 2^{k-1} . Ezt az elrendezést a binárisan minimális ideális MINLP reprezentáció létrehozásának érdekében alkalmaztuk. Minden egyensúlyi tényérokot tartalmazó egység mellett van egy gőz- és folyadéktovábbító egység. Ha a gőz illetve folyadékáram az egyensúlyi egységek helyett az elkerülő egységeken halad keresztül, a megfelelő egyensúlyi egység nem létezik. Ezáltal a szuperstruktúra bármely egyensúlyi tényérokot tartalmazó egysége megkerülhető.

1. ábra

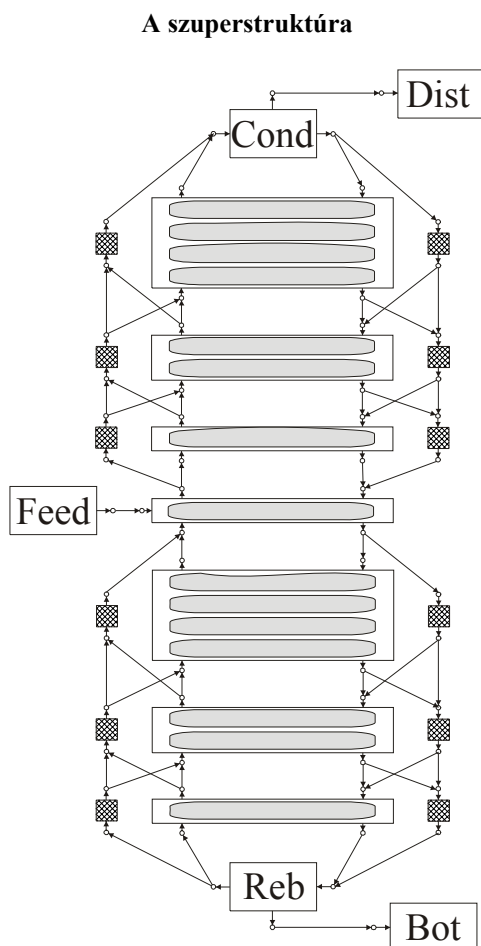


Figure 1: Superstructure

Alap GDP reprezentáció

Az alap GDP reprezentációt (Basic GDP Representation, BGR; lásd *Farkas és mtsai.*, 2004) az R-szupergráf (a szuperstruktúra R-gráfja) alapján állítjuk elő. Ez tartalmazza az egységek egyenleteit, a portok mérlegegyenleteit és a célfüggvényt. Ezekon kívül nem

tartalmaz további összefüggéseket, ezért az összes megvalósítható struktúrát reprezentálja, és külön logikai változót használ minden feltételes egység jelölésére. Az egyensúlyi tényezőket tartalmazó egységek összefüggései tartalmazzák az egységhez tartozó, a fizikai-kémiai egyensúlyt leíró egyenleteket, a tömeg- és hőmérlegeket, továbbá az áramoknak és moláris entalpiáknak a koncentrációból történő számítását. A tényérok fix költségét összevonva egy külön egyenletben számítjuk ki.

$$\begin{array}{l}
 \left[\begin{array}{l}
 Z_{s,3}^U \\
 \wedge \\
 V_i^{in,s,3} + L_{i,j+1}^{s,3} = V_{i,j}^{s,3} + L_{i,j}^{s,3} \quad j \in J4^{first} \\
 V_{i,j-1}^{s,3} + L_{i,j+1}^{s,3} = V_{i,j}^{s,3} + L_{i,j}^{s,3} \quad j \in J4^{inner} \\
 V_{i,j-1}^{s,3} + L_i^{in,s,3} = V_{i,j}^{s,3} + L_{i,j}^{s,3} \quad j \in J4^{last} \\
 L_i^{out,s,3} = L_{i,j}^{s,3} \quad j \in J4^{first} \\
 V_i^{out,s,3} = V_{i,j}^{s,3} \quad j \in J4^{last} \\
 L_{i,j}^{s,3} = LIQ_j^{s,3} \cdot x_{i,j}^{s,3} \quad j \in J4 \\
 V_{i,j}^{s,3} = VAP_j^{s,3} \cdot y_{i,j}^{s,3} \quad j \in J4 \\
 V_i^{in,s,3} \cdot hV_i^{in,s,3} + L_{i,j+1}^{s,3} \cdot hL_{i,j+1}^{s,3} = V_{i,j}^{s,3} \cdot hV_{i,j}^{s,3} + L_{i,j}^{s,3} \cdot hL_{i,j}^{s,3} \quad j \in J4^{first} \\
 V_{i,j-1}^{s,3} \cdot hV_{i,j-1}^{s,3} + L_{i,j+1}^{s,3} \cdot hL_{i,j+1}^{s,3} = V_{i,j}^{s,3} \cdot hV_{i,j}^{s,3} + L_{i,j}^{s,3} \cdot hL_{i,j}^{s,3} \quad j \in J4^{inner} \\
 V_{i,j-1}^{s,3} \cdot hV_{i,j-1}^{s,3} + L_i^{in,s,3} \cdot hL_i^{in,s,3} = V_{i,j}^{s,3} \cdot hV_{i,j}^{s,3} + L_{i,j}^{s,3} \cdot hL_{i,j}^{s,3} \quad j \in J4^{last} \\
 hL_{i,j}^{s,3} = f(T_j^{s,3}) \quad j \in J4 \\
 hV_{i,j}^{s,3} = f(T_j^{s,3}) \quad j \in J4 \\
 hL_i^{out,s,3} = hL_{i,j}^{s,3} \quad j \in J4^{first} \\
 hV_i^{out,s,3} = hV_{i,j}^{s,3} \quad j \in J4^{last} \\
 \sum_{i \in I} x_{i,j}^{s,3} = 1 \quad j \in J4 \\
 \sum_{i \in I} y_{i,j}^{s,3} = 1 \quad j \in J4 \\
 f_{i,j}^{L,s,3} = f_{i,j}^{V,s,3} \quad j \in J4 \\
 f_{i,j}^{L,s,3} = f(x_{i,j}^{s,3}, T_j^{s,3}, p_j^{s,3}) \quad j \in J4 \\
 f_{i,j}^{V,s,3} = f(x_{i,j}^{s,3}, T_j^{s,3}, p_j^{s,3}) \quad j \in J4 \\
 c^{s,3} = f(DC)
 \end{array} \right] \vee \left[\begin{array}{l}
 -Z_{s,3}^U \\
 \wedge \\
 V_i^{in,s,3} = 0 \\
 L_i^{in,s,3} = 0 \\
 L_i^{out,s,3} = 0 \\
 V_i^{out,s,3} = 0 \\
 L_{i,j}^{s,3} = 0 \quad j \in J4 \\
 V_{i,j}^{s,3} = 0 \quad j \in J4 \\
 LIQ_j^{s,3} = 0 \quad j \in J4 \\
 VAP_j^{s,3} = 0 \quad j \in J4 \\
 hL_{i,j}^{s,3} = 0 \quad j \in J4 \\
 hV_{i,j}^{s,3} = 0 \quad j \in J4 \\
 hL_i^{in,s,3} = 0 \\
 hV_i^{in,s,3} = 0 \\
 hL_i^{out,s,3} = 0 \\
 hV_i^{out,s,3} = 0 \\
 c^{s,3} = 0
 \end{array} \right] \quad i \in I, s \in S
 \end{array}$$

Csak a feltételes egységek összefüggéseit részletezzük. A következő halmazokat definiáltuk: S a kolonna szekcióinak halmaza (alsó, felső), K a feltételes egységek halmaza, I a komponensek halmaza, Jn az n tényért tartalmazó egység egyensúlyi

tányérainak halmaza. Jn^{first} , Jn^{last} és Jn^{inner} a Jn halmaznak az első, utolsó illetve közbelső tányérokat tartalmazó alhalmazai.

A feltételes egységek, valamint a gőz- és folyadéktovábbító egységek létezésének leírására a $z_{s,k}^U$, $z_{s,k}^V$ és $z_{s,k}^L$ logikai változókat használjuk. ($s \in S$, $k \in K$).

Az 1. egyenlet a 4 egyensúlyi tényért tartalmazó egység összefüggéseit tartalmazza. Ez az egység a harmadik a szekcióban, $k=3$ ($k \in K$). Ebben az egyenletben V a moláris gőzáram; L a moláris folyadékáram; VAP a teljes gőzáram; LIQ a teljes folyadékáram; hV a moláris gőzentálpia; hL a moláris folyadékentalpia; x a folyadék moltört; y a gőz moltört; T a hőmérséklet; f a fugacitás; P a nyomás; c az egység költsége; DC az oszlop átmérője. Az in és out alsó indexek az egységek be- és kimenő változóit jelölik. A V és L felső indexek a gőz- és folyadékáramokra vonatkoznak. A többi, egyensúlyi tényért tartalmazó egység összefüggései hasonlóan írhatók fel.

A 2. egyenlet a gőztovábbító egységre vonatkozik, ahol tV a moláris gőzáram a gőztovábbító egységben.

$$\left[\begin{array}{l} z_{s,k}^V \\ \wedge \\ tV_{i,s,k}^{in} = tV_{i,s,k}^{out} \end{array} \quad s \in S, k \in K \quad i \in I, s \in S, k \in K \right] \vee \left[\begin{array}{l} \neg z_{s,k}^V \\ \wedge \\ tV_{i,s,k}^{in} = 0 \\ tV_{i,s,k}^{out} = 0 \end{array} \quad s \in S, k \in K \quad i \in I, s \in S, k \in K \quad i \in I, s \in S, k \in K \right]$$

MINLP-reprezentáció

Az alap MINLP reprezentáció (Basic MINLP Representation, BMR) automatikusan előállítható a BGR-ből a logikai változókat binárisakkal helyettesítve (Farkas és mtsai., 2004). A BMR külön bináris változót használ minden feltételes egységre, továbbá minden megvalósítható struktúrát reprezentál.

Mérnöki szempontból azokat a struktúrákat tekintjük megengedettnek, melyek vagy az egyensúlyi egységet tartalmazzák, vagy a hozzá tartozó megkerülő egységeket, de egyszerre kettőt nem. Ezt a következő logikai korlát beszűrésével érhetjük el.

$$z_{s,k}^U \oplus (z_{s,k}^V \wedge z_{s,k}^L) \quad s \in S, k \in K \quad (3)$$

Ez a logikai korlát bináris változókkal is felírható:

$$z_{s,k}^U + z_{s,k}^V = 1 \quad s \in S, k \in K \quad (4)$$

$$z_{s,k}^U + z_{s,k}^L = 1 \quad s \in S, k \in K \quad (5)$$

Ezen korlátokat a BMR-hez adva az már csak a megengedett gráfokat reprezentálja, azaz Ideális MINLP reprezentációvá (Ideal MINLP Representation, IMR) válik.

Egy MINLP-modell megoldhatósága függ az egyenletek alakjától (linearitás, konvexitás, relaxáció, scaling). és a modell bináris változóinak számától. Ezért a bináris változók számának csökkentése csökkenti a MINLP feladat megoldásához szükséges időt.

A 3. egyenlet szerint a $z_{s,k}^V$ és $z_{s,k}^L$ bináris változók redundánsak, ha a $z_{s,k}^U$ változó már létezik, $(1 - z_{s,k}^U)$ kifejezéssel helyettesíthetők. Ezzel a helyettesítéssel a modell bináris változóinak száma a harmadára csökkenthető.

Binárisan minimális MINLP-reprezentáció

Farkas és mtsai. (2004) definíciója szerint a binárisan minimális MINLP reprezentáció (Binarily Minimal MINLP Representation, BMMR) minimális számú bináris változót használ a struktúrák megkülönböztetésére. Könnyen belátható, hogy n db bináris változó 2^n db különböző értéket vehet fel, azaz 2^n struktúra reprezentálható használatukkal. Így ha k számú különböző megkülönböztetendő struktúrát kell kezelnünk, és a bináris változók száma n , akkor a $2^n \geq k$ egyenlőtlenségnek teljesülnie kell, azaz $n \geq \log_2 k$.

Alkalmazzuk ezt a megfontolást egy maximum 31 tényért tartalmazó desztillálóoszlopra, amelyben a felső és alsó oszloprészben külön-külön a tényérok maximális száma 15-15, tehát a megkülönböztetendő struktúrák száma $16 \times 16 = 256$, vagyis a bináris változók minimális száma 8 ($\log_2 256 = 8$).

A k -adik feltételes egység 2^{k-1} egyensúlyi tényért tartalmaz a szuperstruktúrában (1. ábra). Minden ilyen egységhez egy bináris változó tartozik, és csak ezeket a bináris változókat használjuk a modellben. Így a kidolgozott reprezentáció ideális és binárisan minimális egyszerre (Binarily Minimal and Ideal MINLP Representation, BMIMR).

EREDMÉNY ÉS ÉRTÉKELES

A kidolgozott binárisan minimális ideális MINLP modellt (BMIMR) összehasonlítottuk két másik, az irodalomból vett modellel (Viswanathan és Grossmann, 1993; Yeomans és Grossmann, 2000) egy kétkomponensű elválasztási példában (1. táblázat).

Az MINLP reprezentációt Sun Sparc munkaállomáson GAMS (Brooke et al., 1992) egyenletorientált optimalizáló programmal oldottuk meg. Két különböző, MINLP-feladatok megoldására alkalmas szolver használatát teszteltük, az Outer Approximation (OA) módszert alkalmazó DICOPT (2. táblázat), illetve a Branch-and-Bound (B&B) algoritmus elvén alapuló SBB-szolvert (3. táblázat).

Az OA algoritmust alkalmazó DICOPT++ szolverrel kapott eredményeket mutatja a 2. táblázat. Viswanathan modelljének megoldása a 14. iteráció után szolverhiba miatt leállt. Mint azt az 1. táblázat mutatja, az új modell több változót tartalmazott, mint a Yeomans által kidolgozott, megoldása mégis szignifikánsan rövidebb időt igényelt.

1. táblázat

Az MINLP modellek jellemzői

Modell (1)	Egyenletek száma (2)	Nemlineáris egyenletek száma (3)	Változók száma (4)	Bináris változók száma (5)
Viswanathan (6)	1138	388	1105	60
Yeomans (7)	2051	514	1272	60
Új (8)	1592	519	1449	10

Table 1: Characteristics of the MINLP representations

Model(1), Number of equations(2), Number of nonlinear equations(3), Number of variables(4), Number of binary variables(5), Viswanathan model(6), Yeomans model(7), New model(8)

2. táblázat**Az MINLP modellek megoldása DICOPT++ szolვეrrel**

Modell (1)	Iterációk száma (2)	N	D	R	Célfüggvény (3)	Megoldási idő (CPU sec) (4)
Viswanathan	14	20	1,11	1,46	74,07	482
Yeomans	150	13	1,38	2,83	83,15	79 565
Új	150	16	1,18	1,76	72,12	13 643

*Table 2: Solution of the MINLP representations with DICOPT++ solver**Model(1), Number of iterations(2), Objective function(3), Computation time(4)***3. táblázat****Az MINLP modellek megoldása SBB szolვეrrel**

Modell (1)	N	D	R	Célfüggvény (2)	Megoldási idő (CPU sec) (3)
Viswanathan	18	1.137	1.585	72.77	120.7
Yeomans	17	1.151	1.651	72.08	406.5
Új	27	1.117	1.493	82.22	130.9

*Table 3: Solution of the MINLP representations with SBB solver**Model(1), Objective function(2), Computation time(3)*

A B&B algoritmus elvén alapuló SBB szolვეrrel nem sikerült egyértelmű eredményeket kapnunk. Amellett, hogy az SBB szolვეrrel kapott eredmények túl nagy szóródást mutattak, ezzel a szolვეrrel az új modell kevésbé jó eredményeket szolgáltatott, ami annak köszönhető, hogy az egyes bináris változóknak a célfüggvény értékére gyakorolt hatása eltérő, hiszen különböző tényérszámú konfigurációkat reprezentálnak. A struktúrából adódóan bináris fában hasonló célfüggvényértékek egészen más bináris kombinációkhoz tartoznak.

KÖVETKEZTETÉSEK

A kifejlesztett modell alkalmas desztilláló kolonnák tervezésére, az optimális tényérszám és refluxarány meghatározására. Az MINLP-modell mindkét általunk vizsgált szolვეrrel futtatható és megvalósítható megoldást nyújt. Az egyik szolვეrrel (DICOPT++) gyorsabban talál alacsonyabb optimumot, mint a másik két vizsgált modell. A Branch&Bound algoritmust alkalmazó SBB szolვეrrel nem lehet egyértelmű eredményeket kapni. A modell tulajdonságai tovább javíthatók a további fejlesztés folyamán.

IRODALOM

- Brooke, A., Kendrick D., Meeraus, A. (1992). GAMS User's Guide. Scientific Press, USA.
- Farkas, T., Rev, E., Lelkes, Z. (2005). Process flowsheet superstructures: Structural multiplicity and redundancy Part I: Basic GDP and MINLP representations. *Comp. Chem. Eng.*, 29, 2180–2197.
- Friedler, F., Tarjan, K., Huang Y.W., Fan, L.T. (1992). Graph-theoretic approach to process synthesis: axioms and theorems *Chem. Eng. Sci.*, 47, 1973-1988.
- Grossmann, I.E.(1996). Mixed-Integer Optimization Techniques for Algorithmic Process Synthesis. *Advances in Chemical Engineering*, 23, Process Synthesis, 171-246.
- Luyben, M.I., Floudas, C.A.(1994). Analyzing the interaction of design and control—1. A multiobjective framework and application to binary distillation synthesis. *Comp. Eng. Chem.*, 18, 933.
- Farkas, T., Rev, E., Lelkes, Z. (2005). Process flowsheet superstructures: Structural multiplicity and redundancy Part II: Ideal and binarily minimal MINLP representations. *Comp. Chem. Eng.*, 29, 2198–2214.
- Viswanathan, J., Grossmann, I.E.: An alternate MINLP model for finding the number of trays required for a specified separation objective. *Comp. Chem. Eng.*, 17, 949-955.
- Yeomans, H., Grossmann, I.E. (2000). Disjunctive Programming Models for the Optimal Design of Distillation Columns and Separation Sequences. *Ind. Eng. Chem. Res.*, 39, 1637-1648.

Levelezési cím (*Corresponding author*):

Czuczai Barbara

Budapesti Műszaki és Gazdaságtudományi Egyetem

Vegyipari Műveletek Tanszék

1111 Budapest, Műegyetem rkp. 3.

Budapest University of Technology and Economics

Department of Chemical Engineering

H-1111 Budapest, Műegyetem rkp. 3.

Tel.: 36-1-463-34-01

e-mail.: bczuczai@mail.bme.hu



Adaptibility and limitations of the enumerating method

Á. Buza

University of Miskolc, Department of Information Engineering, H-3515 Miskolc, Egyetemváros

ABSTRACT

Nowadays there are a lot of scheduling algorithms which are able to solve different scheduling problems. But the efficiency of these methods is doubtful, because the methods are unable to answer the question, how far is the solution from the optimal solution in the case of Makespan. Only the enumeration can answer the above mentioned question. Enumeration is an NP-hard task, but we can improve its efficiency through parallel processing.

(Keywords: scheduling, heuristics, enumeration)

ÖSSZEFOGLALÁS

A leszámoló algoritmus alkalmazhatósága és korlátai

Buza Á.

Miskolci Egyetem, Alkalmazott Informatikai Tanszék, 3515 Miskolc, Egyetemváros

Jelenleg már számos ütemező algoritmus áll rendelkezésünkre ahhoz, hogy a termelésütemezési feladatokat megoldjuk. Arra a kérdésre azonban, hogy az egyes algoritmusokkal kapott eredmények mennyire közelítik meg az optimális ütemterv jósági indexének értékét az átfutási idő vonatkozásában, egyik sem tud érdemi választ adni – néhány kivételen egyszerű esettől eltekintve. Kizárólag a leszámolással nyílik mód arra, hogy a fenti kérdésre érdemi választ adjunk. Az univerzális leszámoló algoritmus a gyakorlatban – annak hosszú futási ideje miatt – kevésbé elterjedt. Az említett futási idő redukálására számos lehetőség kínálkozik.

(Kulcsszavak: ütemezés, heurisztikák, leszámolás)

INTRODUCTION

Typical scheduling tasks

The scheduling as manufacturing task, arises in every manufacturing plant, where different jobs appear and use the same resources. A job means the completing of finite operations, which appear one after another. The manufacturing control has to allocate the machines to different jobs and calculate the starting time of the different operations on different jobs. There could exist complicated constraints, which determine the non-allowed orders. The mutual contact of the tasks and resources is described by intricate relations and these are stored in the production plans. A scheduling task could be formulated as a discreet optimum problem and in the most cases exact solution does not exist.

A scheduling task could be given generally in the following way: There are n pieces of job, which must be completed in a manufacturing plant and there are m pieces of machine, which are available to finish the jobs. An executive order must be

determined, which satisfies the conditions and has an optimum result from the previously given technological and economical aspect.

If the marks are the following:

- **J** the jobs, which have to be completed ($J=\{J_1,\dots,J_j,\dots,J_n\}$, $n\in\mathbb{N}$),
- **M** the available resources ($M=\{M_1,\dots,M_j,\dots,M_m\}$, $m\in\mathbb{N}$),

the $O_{i,k}(j)$ means k^{th} operation of the i^{th} job, which has to be completed on the j^{th} resource. In the most scheduling tasks i, j and k have an overriding importance, so it is a key-factor to understand their meaning.

There are different task-types, which differ from each other in the number of possible solutions, the length of the solution time and the formulation of the solution method. The following classes can be distinguished by the formulation of the tasks: (Erdélyi 1999)

- Job-shop task:
In the job-shop task n jobs have to be completed on m machines. Every job has to be completed on every machine. The length of the $O_{i,k}(j)$ operation equals to p_{ji} (can be 0). The sequences of the jobs on the different machines are given and can be different at different jobs.
- Open-shop task:
In the open-shop task n jobs have to be completed on the machines. The operations of the different jobs have to be done on different machines and every job must be completed on every machine one time (more than once is prohibited). The sequence of the operations belonging to one job is optional, but different operations of one job can not be processed simultaneously.
- Flow-shop task:
In this task n jobs have to be worked on m machines, but the sequences of the different jobs are identical. This sequence has to be determined at the beginning. In the flow-shop task two different possibilities can be distinguished:
 - the sequence of the jobs at the different machines has to correspond at each machines;
 - or the operations of the different jobs can overtake each other in each machine-buffer.

NUMBER OF THE POSSIBLE SOLUTIONS

The scheduling tasks have finite but numerous solutions in case of finite n and m . The mathematical models which are able to describe the complete task belonging to the circle of combinatorial optimization and they are the typical NP hard problems.

The number of the possible solutions depends on the given mathematical models and the type of the scheduling task. The solution space of the previously mentioned models is huge. Among them the flow-shop task includes the fewest possible solutions. In that case the number of the different possible schedules is $n!$. The solution space of the other models is much higher and the following refer only to the flow-shop tasks (non-overtaking case).

METHODS FOR SOLUTION

Lot of methods are able to solve scheduling problems and the number of them grows continuously. The researchers – mathematicians and engineers – guess more and more methods and some of these could be useful in practice, too.

The methods could be classified into four groups: (Tóth, 1998)

- Exact methods:
They are able to give optimal results, but the adaptabilities of them are crucial. The circumstances determine their usage, and these possible circumstances are very lean. One of them is the enumerating method.
- Rule-based methods:
In the case of these methods different rules are used in the scheduling situations. They are very simple but the efficiency of these rules is doubtful.
- Heuristical methods:
The heuristical methods keep the signature of their developers. Their performance is much better than the rule-based methods, but the efficiency can not be measured and therefore they can not be estimated impartially.
- Searching algorithms:
The searching algorithms follow different strategies. Some of them follow different processes of the nature. Others come from the biology. They work well, but this efficiency can not be measured and for that reason they can not be estimated impartially.

PROBLEMS AND THEIR SOLUTION METHOD

There are a lot of grouping of scheduling algorithms, and they can be classified according to solution methods, usages or their models. One classification can go along the scheduling problems. Then the main classes are: (Erdélyi, 1999)

- Problems, where the optimal solution can be found in polynomial time. In this case the scheduling algorithm can follow the well-known attributes of the problem and it is able to find the optimal solution in polynomial time.
- Small problems, where the optimal solution can be found by enumeration within a reasonable time. This method guarantees the optimal solution and this is the main strength of the enumeration.
- NP-hard problems, where the enumeration can not be used in real time and there are not polynomial solutions which guarantee the optimal solution. In these cases the possibilities are:
 - Heuristics: The solution is fast but the result is just quasi-optimal, so the efficiency of these methods is doubtful.
 - Searching algorithms: These methods follow different strategies and try to find the best solution. The main types are the following:
 - Global strategy: The searching method examines the full solution space but it does not look for the global optimal solution.
 - Local strategy: The local searching methods search for local optimal solutions and they hope that the optimal solution is close to the found solution. The main advantage of these methods is the high velocity.

From the aspect of efficiency the enumeration must be underlined, because this method guarantees the optimal solution. Its price is the relative high running time. At a small scheduling problem (number of jobs = 14) the running time is about 2 days on an average computer.

Counterpoints are the local searching methods and heuristics which are very fast but their results are quasi-optima.

One of the exact methods: the enumerating method

A common property of the exact methods is that their schedules are optimal from the aspect of the examined objective function. Using this common character we are able to create a scheduling model which provides the optimal schedule.

The amount of works, when the exact methods work efficiently, is very small, but contrary to this the importance of these methods is very high from – at least – two aspects:

- The exact methods can be built into several heuristical methods,
- The analysis of the exact solutions can explore the deciding reserves and the structure of the possible but non-optimum solutions.

The number of these methods is slight. There are known only a few – in practice useable – methods because of the complexity of the scheduling task.

At present useful exact algorithms:

- SPT method, for 1 machine,
- Johnson algorithm, for 2 machines (in case of Makespan manager index),
- Extended Johnson algorithm, for 3 machines (in case of Makespan manager index and other criteria!),
- *enumeration*.

From the listed methods the simplest is the enumeration, which makes the other algorithms unnecessary among suitable circumstances and if it is adaptable. The main idea is that all possibilities have to be tried in a model which represents the real manufacturing environment, and after having the necessary data, the most suitable schedule is eligible easily.

The algorithm is very simple but its exclusive usage is excluded by the specialty of the scheduling task. The number of possible events grows factorial according to the number of the jobs. That is why there is no way to examine all possibilities neither at relative little job number, nor at few operations belonging to different jobs in case of job-shop. The opportunity of real time decision supporting is marginal.

The examination run-time can not be reduced to an acceptable measurement even by using the modern high capacity computers.

Theoretically all possibilities must be examined, but fortunately in practice the events which are not allowed because of different technological rules are left out of consideration and its magnitude is invaluable. Hereby the number of allowed schedules could be much fewer.

On the one hand this conception makes the solution easier because of the reduced run-time. On the other hand it excludes the general solution of the original task, because of the prescriptions, which are heuristical and can be different from factory to factory and can change from task to task. These simplifications themselves can be used only in particular industrial cases and they are so task specific that the results coming from the solution will not be useable at another task – with other prescriptions of course.

Tests must examine that the solution space after the simplifications would be so limited that the task could be solved in real time – with the enumerating method, of course.

Remark

This study is the second part of a complex analysis and it depends on the result of the first study. (*Buza, 2004*) Summarizing the first part, it describes the limits on principle and their exemption of parallel processing. Since then a complex system had been

developed on which the written advantages and disadvantages can be experienced. It can provide measuring results which can prove the reason for existence of the method in scheduling.

REQUIREMENTS

The main goal of the research was to develop a new client-server system which can solve scheduling problems parallel. The model was the – well known – flow-shop model and the objective function was Makespan.

The following requirements had had to meet:

- Task-decomposition: the problem has to be broken down to smaller segments which can be divided among the clients.
 - This requirement was kept with limitations, which means that the size of the segments is determined by the user. The part-tasks* must be smaller than this size.
- Size of the part-tasks: the above mentioned task-decomposition has to be done properly, the part-tasks can be neither too small nor too big.
 - The size of the part-tasks depends on the clients' hardware environment. But it can be estimated satisfactorily.
- Communication: the communication structure can be either a strong or a weak point. It must be planned carefully because the order of the communication depends on the structure.
 - The communication is a key-factor, during parallel processing the server-client communication could be so high that it reduces the efficiency of utilization of the resources.
- Security: the security means not the vulnerability of the computers but the errors which have to be treated.
 - The treatment of the mentioned errors can be solved, so the calculated results are in safety in case of an error. Furthermore, "dead" clients do not block the partitioned part-task, it will be divided again among the "living" clients.
- Transparency: the client process must run in background, the user's input must have done in real-time.
 - This is a hard requirement, it can be kept by the size of the part-tasks. But it is hard to predict the type of the computer on which the part-task will be done. That is the reason why it can be a good solution if an average computer is predicted and calibrated the size of the part-task on it.

TECHNICAL DETAILS AND RESULTS

Realisation of distribution

As above mentioned, the problem will be divided into smaller part-tasks. The theoretical description was the theme of the previous study. The practical side of it is following now.

* Part-task is a tightened problem. The user has to determine the size of these part-tasks (p), and the clients check only the solution space of this small task instead of the complete problem (n). The running time is much lower, but the number of these part-tasks is relative high ($\frac{n!}{p!}$ at the flow-shop model).

Let the number of jobs be 5 ($n=5$), and let the size of part-tasks be 3 ($p=3$). The complete solution space consists of 120 different schedules – in the flow-shop model. Using the number of part-tasks describing formula ($\frac{n!}{p!}$), number of part-tasks is 20 in this specific case. These part-tasks must be examined. Schedules of complete problem can be seen on *Table 1*.

Table 1

Schedules of the complete problem

Sequential number of schedules (1)	Schedules (2)
1	1, 2, 3, 4, 5
2	1, 2, 3, 5, 4
3	1, 2, 4, 3, 5
4	1, 2, 4, 5, 3
...	...
119	5, 4, 3, 1, 2
120	5, 4, 3, 2, 1

1. táblázat: A teljes probléma lehetséges ütemtervei

Az ütemtervek sorszáma (1), Ütemtervek (2)

The part-task distribution can be followed on *Table 2*.

Table 2

Part-tasks of the complete solution space

Sequential number of part-task (1)	First schedule in the part-task (2)	Last schedule in the part-task (3)	Number of schedules in the part-task (4)
1	1:1,2,3,4,5	6:1,2,5,4,3	3!=6
2	7:1,3,2,4,5	12:1,3,5,4,2	3!=6
3	13:1,4,2,3,5	18:1,4,5,3,2	3!=6
4	19:1,5,2,3,4	24:1,5,4,3,2	3!=6
5	25:2,1,3,4,5	30:2,1,5,4,3	3!=6
...	...	...	...
18	103:5,2,1,3,4	108:5,2,4,3,1	3!=6
19	109:5,3,1,2,4	114:5,3,4,2,1	3!=6
20	115:5,4,1,2,3	120:5,4,3,2,1	3!=6
Sum:	-	-	20·3!=120

2. táblázat: A teljes megoldási tér részfeladatai

A részfeladat sorszáma (1), A részfeladat első ütemterve (2), A részfeladat utolsó ütemterve (3), Az ütemtervek száma a részfeladatban (4)

After the part-task distribution only the number of the given part-task must have been transferred to the client. It will be enough to identify the specific part-task. Then the client can check the required part-task, calculate the performance index of it and send the best schedule and its performance index (Makespan) to the server application.

With help of the above mentioned method any scheduling problem can be distributed for parallel processing.

The main parameters of the examined tasks can be seen on *Table 3*.

Table 3

The main factors of the problem

Number of jobs (n) (1)	14
Number of machines [†] (m) (2)	10
Size of part-tasks (p) (3)	10
Solution space of the complete problem ($n!$) (4)	87 178 291 200
Solution space of a part-task ($p!$) (5)	3 628 800
Number of part-tasks $\left(\frac{n!}{p!}\right)$ (6)	24 024

3. táblázat: A feladat jellemzői

A munkák száma (1), A gépek száma (2), A részfeladatok mérete (3), A teljes feladat megoldási terének számossága (4), A részfeladatok számossága (5), A részfeladatok darabszáma (6)

Communication conception

The communication conception is one of the most critical elements of the parallel processing. *Figure 1* shows our communication conception with one client. The process of communication is the same by multiple clients, too.

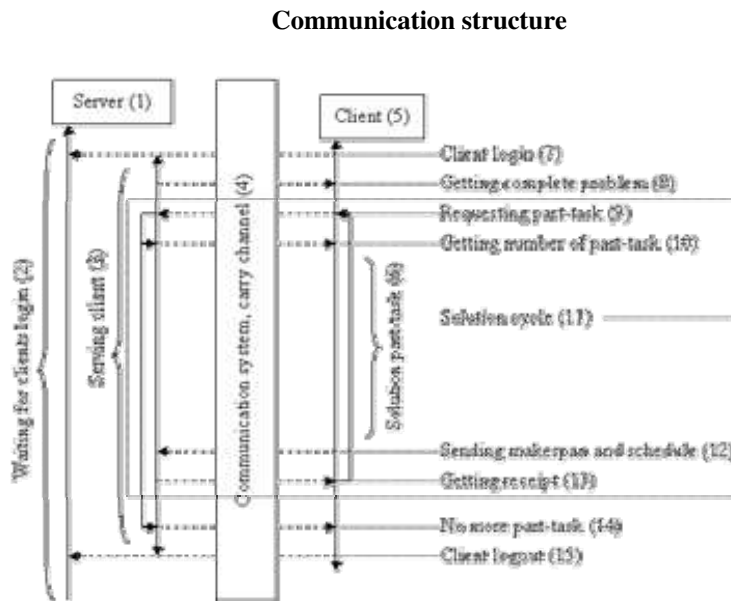
The following steps can be distinguished during the perfect communication:

1. The server process is running permanently and waiting for login of the different clients. If the connection is successful, a new thread is created, which communicates with this client alone. Of course, the server is able to receive the new clients.
2. The client-thread – at the server side – sends the complete problem (including the number of machines (m) and jobs (n), the matrix of machining times and the size of part-tasks (p), which is determined previously) to the client. The client would be able to solve the complete problem but we have an other purpose.
3. The client requests the number of part-task (l), which has to be solved ($1 \leq l \leq \frac{n!}{p!}$,
 l – number of part-task, n – number of jobs, p – size of part-task).
4. The client thread on the server side
5. sends the number of the next part-task waiting for solution (l). (→5.)
6. disconnects if the full problem is solved. (→8.)
7. The client gets the number of the following part-task and checks it.

[†] It must be mentioned, that number of schedules is independent of number of machines.

8. When the client finishes the analysis, it sends the results (makespan and schedule) of the schedule possessing the best performance index to the server.
9. The server gets the results and sends a receipt to the client about the successful communication. (→3.)
10. The communication can be disconnected, because the solution space of the complete problem is examined. The server sends a disconnection message to the client.
11. The client gets the message and destroys himself.
12. The client thread detects this and the connection is over.
13. The server keeps on running.

Figure 1



1. ábra: A kommunikációs struktúra felépítése

Szerver alkalmazás (1), Várakozás a kliens bejelentkezésére (2), A kliens kiszolgálása (3), Kommunikációs közeg, csatorna (4), Kliens alkalmazás (5), A részfeladat megoldása (6), A kliens bejelentkezése (7), A teljes feladat paramétereinek fogadása (8), Részfeladat igénylése (9), A részfeladat sorszámának fogadása (10), Feladat megoldási ciklus (11), Az legrövidebb átfutási idő és a hozzá tartozó ütemterv átküldése (12), Nyugta fogadása (13), A részfeladatok mindegyike megoldásra került (14), A kliens kijelentkezése (15)

The communication is lean, but it is enough to the solution and the safe functioning. There are not any redundant data, which could cause bottle-neck. The complete problem is transmitted so the further communication is very simple.

It means that the clients must have a specific knowledge about the problem and performance indices but this is essential for the efficient communication. If something changes the problem radically, it is enough to change the clients on the server side

because the instances on the client side are downloaded – and refreshed – automatically. So there is no need for any service activities, everything is automatic.

The developer environment: JAVA

In the beginning we planned a system, which does not require any specific knowledge and installation from the users. That is why we decided beside Java. The user does not have to install a complex system, just start his internet browser and – with the help of applet programming – write the given URL into his browser. Our application starts automatically and the analysis is beginning immediately.

There are some problems if the user has an own firewall but this problem is attached to every programming language, not only Java.

So the server and client application have been written in Java and they are really easy to use, according to one of our objectives.

Aspects of the usage

It was mentioned above, that load of the processor can be critical. That is why two possibilities must be distinguished:

- The user is sitting at his computer and working on it. The client is running in the background and does not block the work of the user. In this case the user has to decide how he uses the unnecessary processor time (idle time). The user has to set the priority of the client application.
- There is a specific application for this case, and it can be downloaded and started from the Internet. It connects to the server and begins to solve some part-tasks.
- If the computer is running and nobody uses it, then it can be loaded much better than in the previous. The answer-times could be much larger but it is unimportant. In this case there is an independent client process which detects the idle-time and runs automatically. This application works as a screen-saver, so its realization does not demand extra processor capacity.

CONCLUSIONS AND FURTHER DEVELOPMENTS

Conclusions

Parallel processing is an outstanding tool for improving enumeration efficiency and for reducing running time. There are different limitations but these can be treated by smart distribution. The written applications are able to process parallel and the running time can be reduced significantly. They collect experimental results, which are essential to understand the inner system of the problem.

The communication structure works perfectly and it filters the redundant data-movements. It makes the communication safe and efficient.

Further developments

The main directions of the developments are the systematic use of the applications and the collection of experimental results. The main goal is to develop a brand new scheduling method, which keeps the main advantages of the applications – the shorter running time and the ability to find the optimal solution in the complete solution space in the case of Makespan objective function.

ACKNOWLEDGEMENTS

The researches, presented in this article, were realised in a specific project, which members are academic and industrial partners (Project leader: Dr. Monostori László, MTA SZTAKI). This project is sponsored by the Hungarian Government, too. The author would like to render thanks to the financial and scientific support of the project-members.

REFERENCES

- Buza, Á. (2004). A párhuzamos processzálas alkalmazhatóságának vizsgálata a leszámlálás ütemező algoritmus esetén. Doktoranduszok Fóruma, 2004. 11. 08. Miskolc
- Tóth, T. (1998). Tervezési elvek, modellek és módszerek a számítógéppel integrált gyártásban. Miskolci Egyetemi Kiadó : Miskolc, 150-187.
- Erdélyi, F. (1999). Számítógépes gyártásirányítás. Miskolci Egyetemi Kiadó : Miskolc, 87-123.

Corresponding author (*Levelezési cím*):

Buza Ákos

University of Miskolc, Faculty of Mechanical Engineering
Department of Information Engineering
H-3515 Miskolc, Egyetemváros
*Miskolci Egyetem, Gépészmérnöki Kar,
Alkalmazott Informatikai Tanszék,
3515 Miskolc, Egyetemváros*
Tel.: 36-46-565-111/19-52
e-mail: iitbuza1@uni-miskolc.hu



Intelligent dynamic load balancer for JBoss application server

P. Mileff, K. Nehéz

University of Miskolc, Production Information Engineering Research Team (PIERT) of the Hungarian Academy of Sciences,
Department of Information Engineering, H-3515 Miskolc Egyetemváros

ABSTRACT

The growth of Internet services during the past few years has increased the demand for scalable distributed computing systems. E-commerce systems concurrently serve many clients that transmit a large number of requests. An increasingly popular and cost effective technique to improve server performance is load balancing, where hardware and/or software mechanisms decide which server will execute each client request. Load balancing mechanisms distribute client workload equally among server nodes to improve overall system responsiveness. Load balancers have emerged as a powerful new technology to solve this. This paper focuses on a new generation of adaptive/intelligent dynamic load balancing technique, which based on the J2EE technology and can be practical in J2EE application servers. The paper discusses in detail both the theoretical model of the load balancing and its practical realization. The effectiveness of the new balancing method will be demonstrated through exact measurement results compared with former traditional non-adaptive methods. (Keywords: Distributed systems, Adaptive Load Balancing, J2EE Application server, JBoss)

ÖSSZEFOGLALÁS

Intelligens dinamikus terhelés elosztó JBoss alkalmazás szerverek számára

Mileff P., Nehéz K

Miskolci Egyetem, Magyar Tudományos Akadémia Termelés Informatikai Kutatóhelye (PIERT),
Alkalmazott Informatikai Tanszék, H-3515 Miskolc Egyetemváros

Az Internet nyújtotta lehetőségek nagy arányú növekedésével az elmúlt években rohamosan megnőtt az igény a széles körben skálázható, elosztott számítási rendszerek iránt. A mai E-commerce rendszerek jellemző tulajdonsága, hogy párhuzamosan egyszerre több kliens nagyszámú kérését szolgálják ki. A szerver teljesítmények maximalására napjainkban egyre növekvő népszerűségnek örvend az a költségkímélő, úgynevezett load balancing (terhelés elosztó) technika, amely során különböző hardver és (vagy) szoftver módszerek döntenek arról, hogy melyik szerver szolgálja ki éppen az adott kliens kérését. A technika egyenletesen elosztja a kliensek kéréseit az egyes szerver node-ok között növelve a teljes rendszer hatékonyságát. A terhelés elosztók, mint egy erőteljes, új módszertan jelennek meg ezen problémák megoldására. A cikk egy új generációs adaptív/intelligens, dinamikus terhelés elosztó technikát mutat be, amely J2EE technológiai alapokra épül, és alkalmazása rendkívüli előnyöket jelent J2EE alkalmazásszerverek esetén. Részletesen bemutatásra kerül a terhelés elosztás technikájának mind az elméleti modellje, mind pedig a gyakorlati megvalósíthatóság lehetőségei. Az új technika hatékonyságát egzakt mérési eredményekkel bizonyítjuk, összehasonlítva korábbi hagyományos nem-adaptív módszerek eredményeivel.

(Kulcsszavak: Elosztott rendszerek, Adaptív terhelés elosztás, J2EE alkalmazás szerver, JBoss)

INTRODUCTION

As the number of concurrent requests received by a standalone server increases, the application exceeds the estimated respond time when the work load is too much on a server machine. At this time, there are two options to solve this problem: using faster machines or using multiple machines. The first solution is expensive and limited by the speed of a standalone machine. Second choice is more straightforward: deploy the same application on several machines and redirect client requests to those machines. The system is transparent from outside, which means, client applications perceive a standalone very-fast server with one accessible IP address (*Figure 1*). To achieve the performance and transparency, load balancing algorithms must be utilized.

Load balancing can improve the system performance by providing better utilization of all resources in the whole system consisting of computers connected by local area networks. The objective of load balancing is to reduce the mean response time of requests by distributing the workload.

Theoretical possibilities of realizing load balancing on OSI Layers

The OSI model was developed as a framework for developing protocols and applications that could interact seamlessly. The OSI model consists of seven layers and is referred to as the 7-Layer Networking Model (*Basney and Livny, 1999*). Each layer represents a separate abstraction layer and interacts only with its adjoining layers. Load balancing mechanism can be realized on the Layer 3-7. OSI levels 3 and 4 can be supported balancing mechanisms via network router devices. On layers 5 and 7, 'URL Load Balancing' can be achieved. A lively example of 'URL Load Balancing' can be the following: the URL may be static (such as <http://www.xxx.net/home>) or may be a cookie embedded into a user session. An example of URL load balancing is directing traffic to <http://www.xxx.net/documents> through one group of servers, while sending <http://www.xxx.net/images> to another group. URL load balancing can also set persistence based on the "cookie" negotiated between the client and the server.

Network-based load balancing

This type of load balancing is provided by network router devices and domain name servers (DNS) that service a cluster of host machines. For example, when a client resolves a hostname, the DNS can assign a different IP address to each request dynamically based on current load conditions. The client then contacts the designated server. Next time a different server could be selected for its next DNS resolution. Routers can also be used to bind a TCP flow to any back-end server based on the current load conditions and then use that binding for the duration of the flow. High volume Web sites often use network-based load balancing at the *network* layer (layer 3) and *transport* layer (layer 4). Layer 3 and 4 load balancing (referred to as "switching" *Lindfors et al., 2002*), use the IP address/hostname and port, respectively, to determine where to forward packets. Load balancing at these layers is limited, however, by the fact that they do not take into account the content of client requests. Higher-layer mechanisms – such as the so-called layer 5 switching described above – perform load balancing in accordance with the content of requests, such as pathname information within a URL.

Operating System - based load balancing

This type of load balancing is provided by distributed operating systems via *clustering*, *load sharing*, or *process migration* mechanisms. For instance Microsoft provides a new

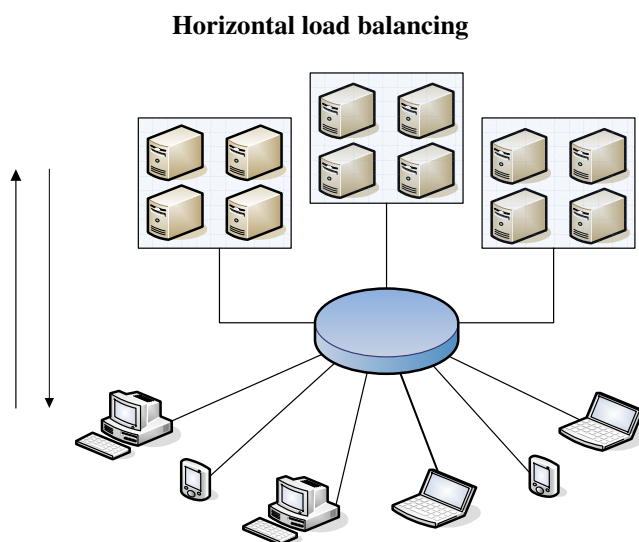
clustering possibility: Microsoft Cluster Server (MSCS). This special Microsoft software provides services such as failure detection, recovery, and the ability to manage the servers as a single system. Clustering is a cost effective way to achieve high-availability and high-performance by combining many commodity computers to improve overall system processing power. Processes can then be distributed transparently among computers in the cluster. Clusters generally employ load sharing and process migration. Balancing load across processors – or more generally across network nodes – can be achieved via *process migration* mechanisms, where the state of a process is transferred between nodes. Transferring process state requires significant platform infrastructure support to handle platform differences between nodes. It may also limit applicability to programming languages based on virtual machines, such as Java.

Middleware-based load balancing

This type of load balancing is performed in middleware products, often on a per-session or per-request basis. For example, layer 5 switching has become a popular technique to determine which Web server should receive a client request for a particular URL. This strategy also allows the detection of “hot spots,” *i.e.*, frequently accessed URLs, so that additional resources can be allocated to handle the large number of requests for such URLs.

Middleware-based load balancing can be used in conjunction with the specialized network-based and OS-based load balancing mechanisms outlined above (*Figure 1*). It can also be applied on top of consumer level (COTS) networks and operating systems, which helps reduce cost. In addition, middleware-based load balancing can provide semantically rich customization possibilities to perform load balancing based on a wide range of application-specific load balancing conditions, such as run-time I/O vs. CPU overhead conditions.

Figure 1



1. ábra: Horizontális terheléelosztás

THE PRACTICAL APPROACH OF BALANCING PROBLEMS

After we have surveyed the theoretical bases of the balancing in a few words we direct our attention to more practical scope of the problem.

A dynamic load balancing can be either *preemptive* or *non-preemptive*. A non-preemptive mechanism transfers only jobs that have just arrived, while a preemptive mechanism transfers jobs at any time, even when the jobs are in execution. Because preemptive mechanism are more costly than non-preemptive one and most of the benefit that can potentially be achieved through dynamic load balancing can be achieved using non-preemptive transfer only, non-preemptive transfers are usually used. Various proposed dynamic balancing methods are based on several policies. Three important ones among them are the transfer policy, the location policy and the selection policy, which decide when, where and what jobs should be transferred respectively. Much work (Basney and Livny, 1999; Schmidt *et al.* 2000) has been published on the design of transfer and location policy but very few on the selection policy.

Balancing policy: When designing a load balancing service it is important to select an appropriate algorithm that decides which server node will process each incoming request. For example, applications where all requests generate nearly identical amounts of load can use a simple Round-Robin algorithm, while applications where load generated by each request cannot be predicted in advance may require more advanced algorithms. In general, load balancing policies can be classified into the following categories:

- *Non-adaptive* – A load balancer can use non-adaptive policies, such as a simple Round-Robin algorithm or a randomized algorithm, to select which node will handle a particular request.
- *Adaptive* – A load balancer can use adaptive policies that utilize run-time information, such as CPU and disk I/O utilization, network loading.

Problem of real-time load balancing

The requests over the network arriving from clients and start a process in memory. Each process runs separated from one another and rivals in gaining resources. The objective of the balancers is to distribute these processes among the individual servers, that response time of processes will be minimal. Because the characteristic of the running tasks can be very various, so it is essential to use an adaptive load balancing algorithm, which try to distribute the tasks in an intelligent way using as it is called *load information*. This is a very difficult objective, because the balancer must accommodate the given job. When we could know in advance what type of task will be arrive, the scheduling algorithm could easily choose the most suitable server for the task, but the type of the tasks knows in general only the client. So the traditional algorithms like Round-Robin or Random access can be usable only with a certain type of tasks.

Leland and Ott (1986) analysed 9.5 million UNIX processes and found that there are three type of processes: CPU intensive processes use great amount of CPU cycles but do a little I/O operations; I/O intensive processes do a great deal of I/O but use a little CPU cycles; canonical processes do a little I/O and use a little CPU cycles. The amount of processes using great amount of CPU cycles and doing a great deal of I/O is extremely small.

Cabrera (1986) analysed 122 thousand processes running on VAX11/785 and found that mean lifetime of processes is 400 ms, the lifetime of 78% of processes is

shorter than one second, 97% of processes terminate within 8 seconds. The author concluded that only long live jobs should be candidates for load balancing due to the overhead costs involved. In a loosely coupled distributed system based on network message passing, a job running longer is often more suitable to transfer than a shorter job since the overhead of transferring a short job may override the benefit.

But not all long running jobs are suitable to transfer. Interactive jobs which constantly need I/O through keyboards and screens, and I/O intensive jobs which heavily access the local file system, will run better locally even though the local CPU load is very heavy (O’Ryan *et al.*, 2000). In a world, only long life and CPU intensive jobs are worthwhile to transfer for remote execution. The question is: how does a scheduler know whether a job is long running and CPU intensive before executing the job? So, a selection policy based on predicting behaviour of a job including it’s lifetime and type is needed to choose which job is suitable for transfer. Little work has been published on this area. The difficult task is to decide the suitability of transferring a job non-preemptively, i.e., predicating the execution time and resource requirements of the job before it is actually executed.

Leland and Ott (1986) found that the residual CPU time needed by a process is linearly related to the amount of CPU time already received by the process (age). The authors developed an assignment scheme which distributes processes based on their age. But an estimate of CPU requirement is unavailable prior to a process’s execution.

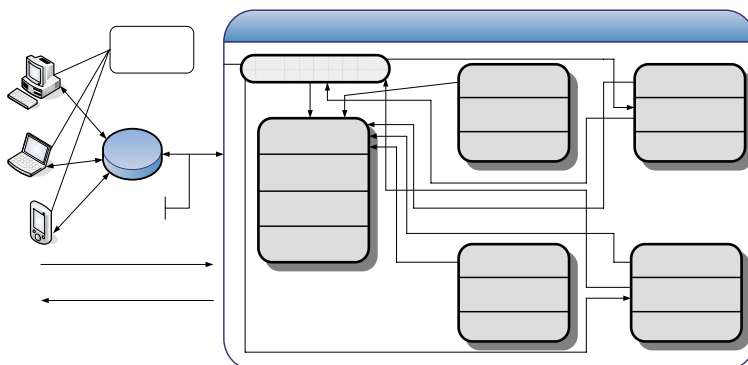
CONCEPT OF AN INTELLIGENT LOAD BALANCER

To create an efficient Load Balancer is a very difficult objective. There are of course many theoretical load balancing solution methods, but many times the practical modell doesn’t make these implementation and efficiency possible. To find the suitable and optimal method for balancing, it is essential to have the most deep knowledge level of the specific system.

Before we go into the details of the Load Balancer, let us examine first the theoretical modell, which is show in *Figure 2*:

Figure 2

JBoss Load Balancer Architecture



2. ábra: JBoss alapú terhelésselosztó felépítése

The theoretical functionality of the balancer is the following: Standalone clients initiate requests over the network through HTTP protocol or RMI to the JBoss cluster. The JBoss cluster can be a complex of homogeneous or inhomogeneous computer on which the JBoss application server runs in cluster mode. Of course, more clients can initiate a request at the same time to the cluster, so the cluster must fulfil more than one request parallel. The incoming requests are received and directed to the compliant node of the cluster by the intelligent load balancer. So it's objective is to choose the most ideal node in term of execution based on the collected *load information* by the Dispatcher. To elect the ideal node is not an easy matter. The main objective of the balancer is to realize a more effective task-division, which response time can be more better than the former algorithm. In additional we concentrate the detailed elaboration of the practical realization.

Components of the Load Balancer

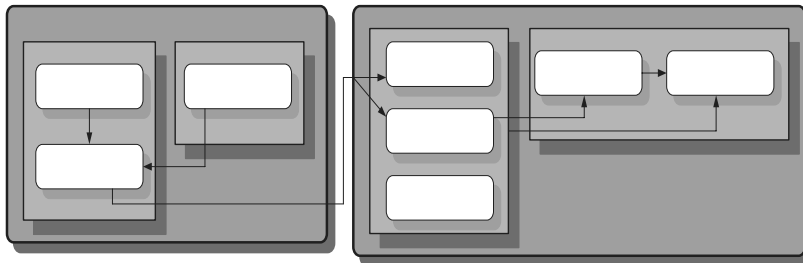
The architecture of our Balancer essentially can be divided into three individual components: the *Statistics Service*, the *Dispatcher*, and the *Scheduler* as well. The individual units are in close communication with one another, none of them can operate without the others. At present the connection of the units works on the concept of the *Remote Method Invocation (RMI)*, but the following objective is to change the entire communication or part of that to the new TreeCache method of JBoss. Utilizing TreeCache, response time may be shorter.

Statistics Service

We can consider from the description above, that the Statistics Service is responsible for the load information. Naturally this unit must run on each node. When a new node come into the cluster, then the Statistics Service start immediately on it. It attempts to find the Dispatcher and provide data to it. *Figure 3* shows the architecture of the Statistics Service and the Dispatcher:

Figure 3

Elements of Statistics Service



3. ábra: A statisztika szolgáltató MBean elemei

Figure 3 shows that Statistics Service is consisted of three parts: *CPU* -, *I/O Statistics* and *Fuzzy Engine*. The functionality arise from those name: CPU Statistics services the CPU usage and I/O Statistics the I/O usage of the specific node. The CPU Statistics and the Fuzzy Logic represent collectively an MBean (*Managed Bean*) unit, however the I/O

Statistics is an another separate MBean unit. In the JBoss system each MBean indicate services. The sufficient node-information are essential to the compliant operating of the balancer. In fact, Java classes are running in a virtual machine on each host, therefore it does not make it possible to query the load information directly from the operating system. For this reason we had to evolve individual methods and had to utilize operating system specific resources. Nevertheless these resources are operating system dependent.

The current version of the balancer works on MS Windows Systems, but further objective is to create Linux/Unix version too. Since the Java 1.5 appeared , it become possible to measure the CPU average usage with the Java Management Extension technology, using the built in `OperatingSystemMXBean` class. It has a function named `getProcessCpuTime()`, which can query the CPU time of the specific JVM (*Java Virtual Machine*) in nanosecond, from which the average CPU usage can be computed. The CPU usage can be query direct from the operating system, but in this case the efficiency of the balancer can degrade to a great extent. The reason for this is that, the MS Windows operating system updates the data of the *Performance Monitor* every 1000 millisecond, on account of which the schedule of the short task becomes impossible. The JBoss system can work with 50 ms sample time, but in this instance the data acquisition is fulfilled in every 100 ms.

The acquiring the I/O information is already much harder task by far. Now Java helps us neither so much as was in case of CPU usage. To get the required infomation we need operating system level methods, to which the C/C++ programming language ensures the suitable environment. The solution was realized by the technology as called JNI (*Java Native Interfaces*), which makes merging the C/C++ and the Java programming language possible. So the survey data of the I/O are realized by native invocation. However the operating system is again a limiting factor, because the data are only updated in every 1000 millisecond. If the client requests are not so frequent, this limit is enough in practice.

Before we change to the consideration of the Fuzzy Engine, it is necessary to make a mention of a relevant feature of the statistics collector MBeans. All the nodes send information to the Fuzzy Engine, when the avarage usage of these are smaller than 100%. This is the most essential condition of the operating of the balancer, what we will detail in the discussion of the Dispatcher.

The Fuzzy Engine is responsible for the part of the adaptivity of the balancer. It gathers the information sented by I/O and CPU services and deducts a *fuzzy* value between 0 and 1 supported by a preset *Fuzzy Engine*. This fuzzy value will be sent to the Dispatcher, that stores it in a hashtable. Current version of Balancer use three fuzzy linguistic variables: one for I/O and an other for CPU utilization and the third one indicates the service capability of a server node. First two variables are considered input variables and third one as output variable. Both input variables are divided into three membership functions, therefore output server capability must be divided into six membership functions. Further aim is to fine the shape of membership functions using a fuzzy-neuro engine. In *Figure 4*, all membership functions of fuzzy variables can be seen.

The dispatcher

The Dispatcher is the second most important part of the Load Balancer. It is also realized by MBean. It's objective is to store the status information sented by each node in a hashtable structure. *Figure 3* shows the architecture of the Dispatcher.

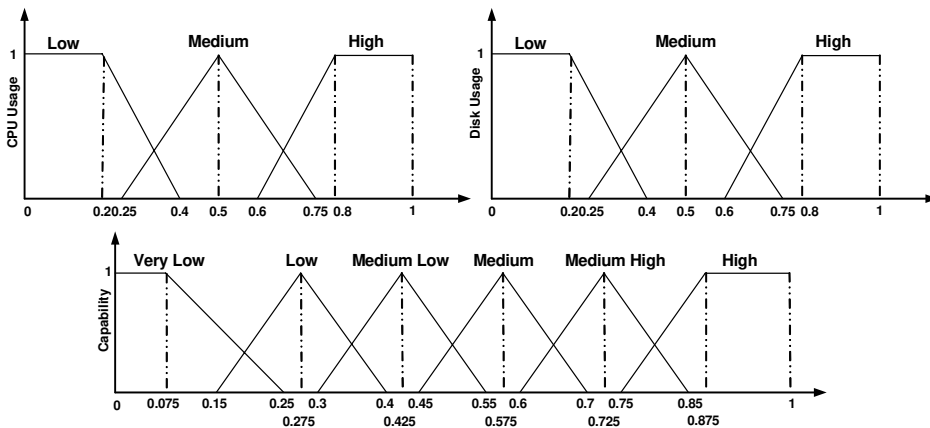
The sent forwarded information consist of two parts: a fuzzy engine value and the IP address of the specific node. The IP address is essential to identify the nodes. The infomormation gets to a hashtable bucket in a vector with the time of arrival together. As *Figure 3* shows, the key of the hashtable is the IP address, because it is individual.

By the discussion of the Statistics Service we have mentioned, that there is a condition, whereas a node only send the information to the Dispatcher, when its load is fewer than 100%. In Dispatcher this effects, that the belonging stored information of the hashtable bucket will not be updated. The *time stamp* of the data is therefore essential, because relying upon these findings will the balancer make a decision to wich information are timely, and which not.

The Dispatcher can be find on only one node in the cluster. It makes no difference on which, but the best thing to do is that, it is started on the fastest node. The connection between the Statistics Services and the Dispatcher is dynamic, that is each node in startup finds and stores the address of the node on which the Dispatcher runs.

Figure 4

The Fuzzy Engine Linguistic Variables



4. ábra: A Fuzzy motorban értelmezett változók

The Balancer

After preparation of the data the work of the balancer is no more so difficult. However we have to pay attention at the optimal implementation, because the least mistake can also cause big response time decrease. The balancer is a java class implemented a *CustomLoadBalancePolicy* interface, which is functionally part of the JBoss base interfaces.

Its theoretical operating is the following: The balancer make a decision on the bases of the status information collected from server nodes. It considers those information valid, which arrived within 150 ms. Those nodes, which are highly loaded, they don't send any information to the Dispatcher, so naturally the balancer doesn't give them a new task. The balancer will choose the node with the best fuzzy engine value. However in case of a big loaded cluster can often occure so, that all of the nodes are loaded fully and none of them makes a sign. Nevertheless in this case the balancer have to choose one of them, but the question is which one.

Many solution methods have sprung up, however by reason of the tests it appeared, that such method needed, which can efficiently distribute the works in case of big loaded

nodes. The first solution is the random distribution. It can be good, or can be very bad because of the random distribution. For instance if the random balancer gives the work to such node, which is slower than the others, and of course also loaded on 100%. It proved a little better that method which gives the work to that node, which average non-response time is the least, if every node are out of time constraint.

A very important element of the balancer is the following: in the current version of the balancer a node can get a work twice one after the other only, if its CPU usage doesn't correspond to the stored value at the giving out of the previous work and also this value is so more little, than the value of all the nodes. This condition came into the balancer therefore, because when almost more clients all at ones give their requests parallel, then without this condition the same node receive the request of more clients, because the requests are so close to one another, that the data of the balancer couldn't update so quickly.

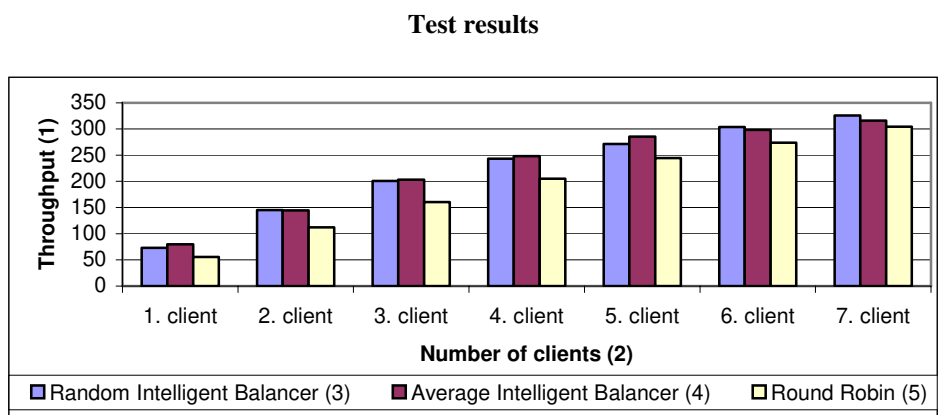
Test and results

The testing process has been carried out on a JBoss cluster consisting seven homogeneous PC-s. Each machine had Pentium III 733 MHz CPU with 256 MByte RAM. Machines were connected via 100Mbps Ethernet network. Utilized operation system was Windows 2000 SP5. Application server version was JBoss 3.2.5 'WonderLand'.

Simulating client requests was carried out with a generic professional simulation environment: *Apache JMeter* (2005). During testing process server machines were slowed-down randomly with a special Loader-MBean emulating I/O or CPU load. Loader-MBean is used for emulating other clients requests and other applications that are parallel launching on the server nodes.

We have started the simulations with a client, then we increased the number of clients to seven. In the course of all simulation we have tested all algorithms three times, then we represented these average results on the *Figure 5*. The diagram shows properly, that in every case the results of the Round-Robin fell short of the results of the Intelligent Balancer.

Figure 5



5. ábra: Szimulációs eredmények

Teljesítmény(1), Kliensek száma(2), Véletlen elosztású adaptív ütemező(3), Átlagos elosztású adaptív ütemező(4), Round-Robin(5).

If we examine the results we can see that, the value of the Throughput increases with the increasing number of the clients, although it is not in direct ratio. The more clients initiate request to the cluster, the more clients share the CPU. Exactly that is why it does no good to more increase the number of the client – like number of the nodes – in the course of the measurement, because at such times the scheduling lose its importance.

Of course, it depends on the type of the task scheduling requested task, that they in what extent require the resources. In the course of seven homogeneous nodes optimal distribution is, if all of them get one. Certainly, we assume that the request of the clients arrive in near time. Because the artificial loads run in random time on the nodes, therefore certain corresponding with the number of nodes or more the nodes become full.

At that time every node are maximum load. Whereas at such time the scheduling is impossible, therefore the best solution is that, if we distribute the tasks optimal among the nodes till then, while the scheduling will be become possible. The Balancer does it in two ways: with random node-choosing and with average response time. One node could not get two tasks one after another.

The *Figure 5* shows both results of the algorithm and with increasing of the number of clients – which means that the more task get into the system – better and better approach the theoretical maximum of the response time of the Round-Robin and the Intelligent Balancer.

In the event of inhomogeneous nodes certainly we can reach much better response time, but of course it depends on the inhomogeneity of the nodes.

The following table summarizes, how much speed increase can be achieved utilizing new Balancer compared with applying Round-Robin algorithm. Results highly depend on the type of tasks: a task to what extent claims the capacity of a node. In our test environment, execution time of a task is 500 ms on a non-loaded server node. Client requests follow each other within 500ms time interval and plus minus 200ms uniform random time. Aim of random interval is to simulate realistic non-predicted client requests. Based on the test results, it is clear that our intelligent balancer algorithm has better performance than Round-Robin algorithm.

During tests intelligent load balancer and only Round-Robin algorithm was compared because Round-Robin algorithm is definitely better than other classic methods like: First Available and Random balancer algorithms. Thus our aim was to outstrip this traditional non-adaptive method.

Table 1.

Balancing Algorithms comparison

Balancer Type (1)	Speed Improvement/client (2)						
	1. client	2. client	3. client	4. client	5. client	6. client	7. client
Random Intelligent Balancer (3)	25%	23%	20%	16%	10%	10%	7%
Average Intelligent Balancer (4)	30%	23%	21%	18%	14%	9%	4%

1. táblázat: Terhelélosztó algoritmusok összehasonlítása

Ütemező típusok(1), Sebesség növekmény/kliens(2), Véletlen elosztású adaptív ütemező(3), Átlagos elosztású adaptív ütemező(4).

CONCLUSION

An intelligent fuzzy-based Load Balancer Application and its test results have been presented in this paper. Continuing work will focus on further developing and implementing more flexible XML based configuration possibilities and redesign communication between server nodes and the dispatched session bean utilizing the new JBoss TreeCache introduced by the latest JBoss version 4.0.

ACKNOWLEDGEMENTS

The research and development summarized in this paper has been carried out by the Production Information Engineering and Research Team (PIERT) established at the Department of Information Engineering and supported by the Hungarian Academy of Sciences. The financial support of the research by the aforementioned source is gratefully acknowledged.

REFERENCES

- J. Lindfors, M. Fleury, The JBoss Group (2002). JMX: Managing J2EE with Java Management Extensions. SAMS Publishing Inc., 56-123.
- J. Basney, M. Livny (1999). Deploying a High Throughput Computing Cluster. In: High Performance Cluster Computing, 1. May 1999.
- C. O’Ryan, F. Kuhns, D.C. Schmidt, O. Othman, J. Parsons (2000). The Design and Performance of a Pluggable Protocols Framework for Real-time Distributed Object Computing Middleware. In: Proceedings of the Middleware 2000 Conference, ACM/IFIP, Apr. 2000. 78-90.
- D. Schmidt, M. Stal, H. Rohnert, F. Buschmann (2000). Pattern-Oriented Software Architecture: Patterns for Concurrent and Networked Objects. Wiley
- L.M. Cabrera (1986). The influence of workload on load balancing strategies, In: Proc. Summer USENIX Conf., June 1986. 446-458.
- W. Leland, T. Ott (1986). Load balancing heuristics and process behavior. In: Proc. ACM SIGMETRICS Conf. Measurement and Modeling of Computer Syst., May 1986.
- J. Shirazi (2003). Java Performance Tuning, Second Edition, O’Reilly, 2003. 34-142.
- JMeter (2005). Generic Simulation Environment <http://jakarta.apache.org/jmeter>, (Apache Jakarta JMeter)
- JBoss (2005). Leading J2EE Open Source Application Server, www.jboss.org
- Fuzzy Logic Systems (2005). <http://www.seattlerobotics.org/encoder/mar98/fuz/flindex.html>

Corresponding author (*Levelezési cím*):

Péter Mileff

University of Miskolc

Department of Information Engineering

Production Information Engineering Research Team (PIERT) of the Hungarian
Academy of Sciences

H-3515 Miskolc Egyetemváros

Miskolci Egyetem, Alkalmazott Informatikai Tanszék

Magyar Tudományos Akadémia Termelés Informatikai Kutatóhelye (PIERT)

Alkalmazott Informatikai Tanszék

3515 Miskolc Egyetemváros

Tel.: 36-46-565-111/19-52, Fax: 36-46-563-405

e-mail: mileff@ait.iit.uni-miskolc.hu



Szoftverfejlesztés a személyi számítógépek ipari körülmények közötti alkalmazhatóságának vizsgálatára

Molnár¹ S., Lágymányosi¹ A., Tímár¹ T., Dezső¹ O.
Szöllősi² Zs., Tokai² Z.

¹Szent István Egyetem, Informatikai és Matematikai Intézet, Informatika Tanszék, 2103, Gödöllő, Páter K. u. 1

²KUKA Robotics Hungária Ipari Kft. 2103, Gödöllő, Páter K. u. 1

ÖSSZEFOGLALÁS

A mesterséges intelligenciára alapozott irányítások egyik lényeges alapeleme az alkalmazott számítógép. A személyi számítógépek rohamos fejlődésének következtében mára a PC-k mind tároló- mind feldolgozó kapacitásukban alkalmassá váltak arra, hogy meghatározott ipari folyamatok irányítását ellássák. Több ipari szereplő ennek jegyében kimondottan ipari körülmények között is alkalmas PC-ket illetve PC alkatrészeket gyárt. Korlátozott klimatikus viszonyok betartása mellett azonban a kereskedelmi kategóriába tartozó PC-k is alkalmassá tehetők ipari körülmények közötti alkalmazásra. Lényeges ismétel, hogy a kereskedelemben kapható PC-k és azok alkatrészai úgy megbízhatóságban, mint klimatikus érzékenységükben heterogén képet mutatnak. Tehát az alkalmazni kívánt számítógépet illetve részeket kellő körültekintéssel kell kiválasztani. A kiválasztást követően pedig, fontos szempont az eszközök klimatikus tesztelésének elvégzése, és ezáltal az alkalmazatlan egységek kiszűrése. Ennek érdekében kidolgozás alatt áll egy számítógépes vizsgálórendszer, melynek feladata az egyes számítógépegységek klimatikus szempontból történő ellenőrzése. A rendszer kidolgozásánál cél az alkalmazási terület klimatikus körülményeinek figyelembevétele, a rendszer hardverfelépítésének meghatározása valamint a működtető szoftverek kifejlesztése mind szerver, mind kliens oldalról. Ezen cikk, a jelenleg is folyó munka egy részét képező és elkészült rendszertervet valamint a szerveroldali szoftverfejlesztést hivatott bemutatni.

(Kulcsszavak: irányítás, teszt, szoftverfejlesztés, adatbázis)

ABSTRACT

Software development for examination of potential industrial application of commercial PC architecture

S. Molnár¹, A. Lágymányosi¹, T. Tímár¹, O. Dezső¹, Zs. Szöllősi², Z. Tokai²

¹ Szent István University, Department of Informatics, H-2103, Gödöllő, Páter K. u. 1.

² KUKA Robotics Hungary Industrial Ltd., H-2103, Gödöllő, Páter K. u. 1.

Applied computer is an essential element of control, based on artificial intelligence. Due to the rapid development of personal computers, nowadays PC-s considering their storage capacity and processing speed are suitable to control particular industrial processes. Several computer manufacturers construct PC-s and hardware components for uses under specially industrial circumstances. On limited climatic conditions

commercial PC-s can be made suitable for application in industrial circumstances. It is worth mentioning that commercial PC-s and their components very different in reliability and climatic sensibility. So selecting parts of applied computers is important. On the other hand making tests under climatic conditions is also essential in filtering unusable units. For solving this problem a computer system is under development. Its basic function is to test hardware components from climatic point of view. The goal in this project is to design the hardware structure of the system, and develop a client-server application for controlling the whole process. This article is a brief introduction into this work containing the rough scheme of the system and making the server side software.

(Keywords: control, test, software development, database)

BEVEZETÉS

Annak érdekében, hogy el lehessen dönteni, miszerint meghatározott ipari alkalmazások elvégzésére alkalmasak-e az egyszerű PC-k először meg kell vizsgálni a PC-k „sebességét”. Az általános PC alkalmazások esetében lényegi kérdésnek tűnik megvizsgálni a központi feldolgozó egység sebességét. Az ipari „gépipari” folyamatok jelentős részénél a rendszerek, pl. a mechanikai méretei miatt korlátozott reakcióidővel és nagy holtidővel rendelkeznek. Ez a gyakorlatban az irányítási feladatok többségénél ms-os esetleg μ s-os, reakció időt jelent, sőt nem ritka a több perces reagálással rendelkező rendszer sem. Az előző szempontból gyorsnak számító berendezések közé sorolandó egy gépjármű ABS rendszere, ahol $n \times 100$ Km/h -s sebesség mellett is legalább $n \times 100 \mu$ s feldolgozási idő áll az irányító rendszer rendelkezésére. Ez azt jelenti, hogy még egy ilyen rendszer esetében is egy i486-os magú processzor is elegendően gyors feldolgozási sebességgel rendelkezik.

Tehát a sebesség kérdését az ipari alkalmazások területén nem a mai processzorok körül kell keresni.

Természetesen egy alap PC –önmagában nem alkalmas arra, hogy egy rendszert irányítson, mivel a mérés és beavatkozás eszközeivel nincs felszerelve. A feladatok megvalósításához tehát kiegészítő eszközökre berendezésekre van szükség. A kiegészítéseket hozzá kell kapcsolnunk a PC-hez. Ehhez valamilyen szabványos illesztő felületet kell használni mely célszerűen a széles körben elterjedt és egységesített szabványokat tekintve vagy ISA vagy PCI rendszert jelent, illetve a „külső csatlakozást” nézve a soros (RS232), párhuzamos (IEEE 1284) valamint az USB vagy az IEEE 1394 interfész.

A PC-k moduláris felépítését követve célszerű egy fejlesztésnél végiggondolni, hogy a belső vagy külső illesztő egységeket célszerűbb-e alkalmazni. A külsőnek előnye a mobilitás és univerzalitás. A belsőnek a kompakt kivitel és a belső tápellátás.

A vizsgálat tárgyát képező rendszer kiválasztásánál a kompakt saját fejlesztésű rendszer került előtérbe.

A fentiek figyelembevételével vizsgáljuk meg, hogy az egyes belső buszok alkalmasak-e a megfelelő mennyiségű adat adott időegység alatt történő átvitelére.

ISA busz

16 adatvezetéke, 24 címvezetéke és 8.33 MHz-es órajele van. Maximum 5 Mbyte adatot lehet átvinni másodpercenként, de ezt az értéket a gyakorlatban csak ritkán lehet elérni.

PCI busz

A PCI sín maximális (elméleti) átviteli sebessége 32 biten 132 Mbyte/s, 64 biten 262 Mbyte/s 33 MHz-es órajel mellett.

Megállapítható, hogy akár még az ISA busz sebessége is kellő egy hétköznapi alkalmazáshoz, de a fejlesztéseknél fontos szempont, hogy manapság megjelenő alaplapok már csak elvélve tartalmazznak ISA csatlakozót ezért a nagyobb sebesség és az elérhetőség is a PCI busz alkalmazása mellett szól.

A sebességek vizsgálata során megállapítható, hogy a feladathoz alkalmasan választott PC –k sebessége nem korlátozza a rendszer működését.

Egy másik lényeges szempont az adott rendszer klimatikus körülményeinek vizsgálata.

Egy gépjárműben alkalmazott berendezésnek a -20 ...-30°C-os hidegben és akár az 50 ... 60 C-os melegben is üzemelnie kell nem is beszélve a hajókon uralkodó körülményekről. Ezt a klimatikus követelményt külön klimatizált dobozolás nélkül egyetlen kereskedelmi PC sem tudja teljesíteni.

Az ipari csarnokokban azonban nem ennyire mostohák a körülmények. Egyetlen üzemben működő berendezés sem alkalmas arra, hogy azonos pontossággal működjön szélsőséges hőmérsékleti viszonyok között. A megfelelő hőmérsékleti környezet biztosítása az üzemben dolgozó emberek munkakörülményeire vonatkozó előírások miatt is szükséges feladat. Kiindulva egy belső üzemi alkalmazásból megállapítható, hogy üzemi esetben a környezeti hőmérséklet szélsőséges esetben sem eshet kívül az +5 ... +55 C-os hőmérsékleti határokon. Bár ez nem tűnik nagyon szélsőséges határnak, de arra kell gondolnunk, különösen a felső 50...55 C-os határ elérése esetén, hogy a PC-k működése során keletkezett hőt is el kell vezetni.

A PC KIVÁLASZTÁSA

A kereskedelembe kapható PC-k és azok alkatelemei úgy megbízhatóságban, mint klimatikus érzékenységükben heterogén képet mutatnak. Tehát az alkalmazni kívánt számítógépet ill. részegységeket kellő körültekintéssel kell kiválasztani. A kiválasztást nem lehet egyszerűen a technikai paraméterek összehasonlításával elvégezni, mivel ebben az esetben a hétköznapi esetekhez képest szélsőségesek a körülmények. A kiválasztott PC-t tartós klímatesztnek kell alávetni. Ha egy gép megfelelt a teszteken és alkalmassága bebizonyosodott, akkor is minden egyes gépet külön le kell tesztelni, mivel a technológia szórása a szerelések, valamint a szállítások is mind olyan meghibásodásokat okozhatnak, amik szobahőmérsékleten esetleg nem kerülnek felszínre.

A fenti vizsgálatok minden lépésén végig kellett menni a KUKA Robotics-nál is, amikor a PC központú robotvezérlés mellett döntöttek (1. ábra).

Az évek során a vezérlő számítógépek folyamatosan jól teljesítettek és beváltották a reményeket. A „PC gyártás” folyamán természetesen a komplett PC mindig teljes klíma és működési teszten megy keresztül annak érdekében, hogy minimálisra csökkenjen a hibás működésű, kiszállított PC. Erre a célra egy egyedi tesztrendszer került kidolgozásra és alkalmazásra.

AZ EGYEDI ALKATRÉSZEK TESZTELÉSÉRE ALKALMAS RENDSZER

Mivel az egyes helyeken alkalmazott PC-k feladatai is megváltozhatnak és természetesen az alkatrészek meghibásodása sem zárható ki, időnként a PC-k egyes részségeit ki kell cserélni.

Az évek során egyre több PC részegységet kellett a már működő rendszerekbe kiszállítani, és ezért felmerült a részségek külön-külön történő klímatesztjének szükségessége (2. ábra).

1. ábra

A KUKA Robotics-nál alkalmazott vezérlő hagyományos PC alkatrészekből

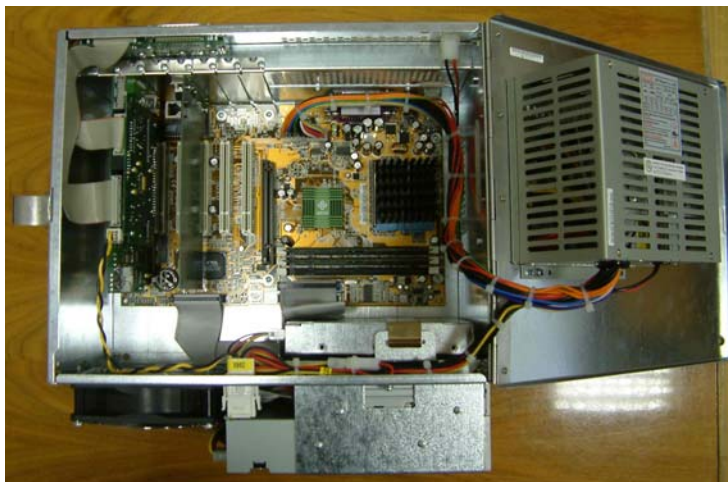


Figure 1: Applied controller PC with commercial elements

2. ábra

A részegységek tesztrendszerének elvi felépítése HDD-k vizsgálatára

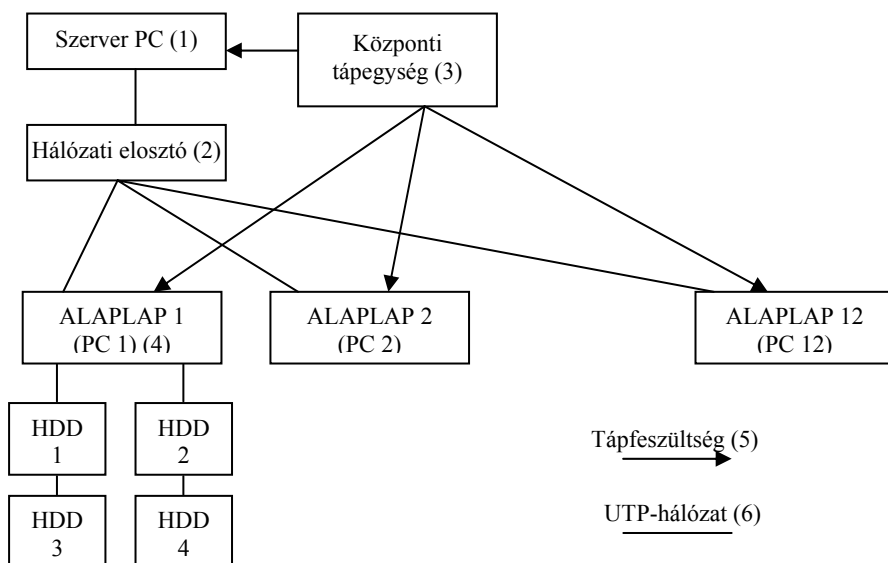


Figure 2: The logical scheme of HDD test

Server(1), Switch(2), Central PowerSupply(3), Motherboard(4), PowerSupply(5), UTP Network(6)

A kidolgozott tesztrendszer alapját egy egyedi linux képzi, a fejlesztés pillanatában létező legfrissebb kernelt (2.6.10) alapul véve. A szerver debian sarge alapokon nyugszik, amelyen a grafikus felhasználói felület fut, ennek alapja a sun 1.5-ös verziójú java futtatókörnyezete. A szerveren található egy DHCP és egy TFTP szerver. Ezen két szoftver felelős a kliensek bootolásáért. A kliensek operációs rendszere egy teljesen egyedileg felépített linux, amely a szerverről töltődik be. A kliens rendszer összesen 4 fájlból áll, egy kernelből, egy az inicializálási folyamatért felelős image-ből, a fő rendszer tömörített image fájljából és a tesztrendszer mini image-éből.

A kliensek indulása, a BOOT folyamat felépítése a következő:

1. a kliens gépek hálózati BOOT-olással indulnak, a szerveren a DHCP szoftvertől kapnak IP címet,
2. a kapott IP címmel a TFTP szervertől megkapják a kernelt, és az inicializáló image-et, majd automatikusan megkezdődik a BOOT folyamat,
3. az init szkript létrehoz egy ramdrive-ot, majd az így létrejött tároló területre http kapcsolaton keresztül letölti a szerverről a rendszer fő állományát, és a tesztrendszer mini image-ét, és felfűzolja ezeket a fájlrendszerbe, majd átadja a vezérlést a rendszerindító folyamatnak,
4. betöltődik a rendszer, elindulnak az alapvető működéshez szükséges folyamatok, és legvégül a rendszer meghívja a teszt szoftver indító állományát,
5. az indító állomány eldönti, hogy milyen tesztípus kell, hogy induljon, és automatán megteszi a szükséges lépéseket.

A tesztrendszer adatai a szerveren található SQL adatbázisban kerülnek eltárolásra, tesztől függően. HDD teszt esetén, a kliensen futó tesztprogram bizonyos (a grafikus felületen meghatározható) időközönként lefuttatja a tesztet, és beírja annak eredményét az adatbázisba, más esetben pl. alaplap teszt esetén pedig a szerveren futó felület bizonyos időközönként lekérdezi a klienst, és a kapott eredményeket ő maga tárolja el.

A tesztek során, egy grafikus felületen keresztül tartható a kapcsolat a szerver számítógéppel. Az aktuális tesztek futtatásán túlmenően a felhasználó a régebbi tesztek eredményeit is értékelheti vagy mentheti. Az aktuális programverzió frissítése is a szerver számítógépen futó program menüjeként érhető el.

A SZERVER OLDALI KEZELŐFELÜLET HDD TESZT ESETÉRE

A szerver program képernyőelemei

Panel

A képernyő egy jól elkülöníthető része. Névvel van ellátva (3. ábra).

Táblázat

Rácshálós ablak sorokkal és oszlopokkal. Jól megkülönböztethető címsora van, melyre való kattintással az adott oszlop szerinti – csökkenő vagy növekvő – rendezettség valósítható meg (4. ábra).

HDD-k panel

Itt kell felvenni a tesztelni kívánt HDD-ket vonalkódjaik megadásával (5. ábra). A teszt során figyelemmel kísérhető az egyes HDD-k összes hibáinak száma.

HDD Teszt panel

Egy kliens teszt adatainak megjelenítése (6. ábra).

S.M.A.R.T. panel

A HDD ún. belső SMART teszt értékeinek megjelenítése (7. ábra).

3. ábra

Tesztek panel

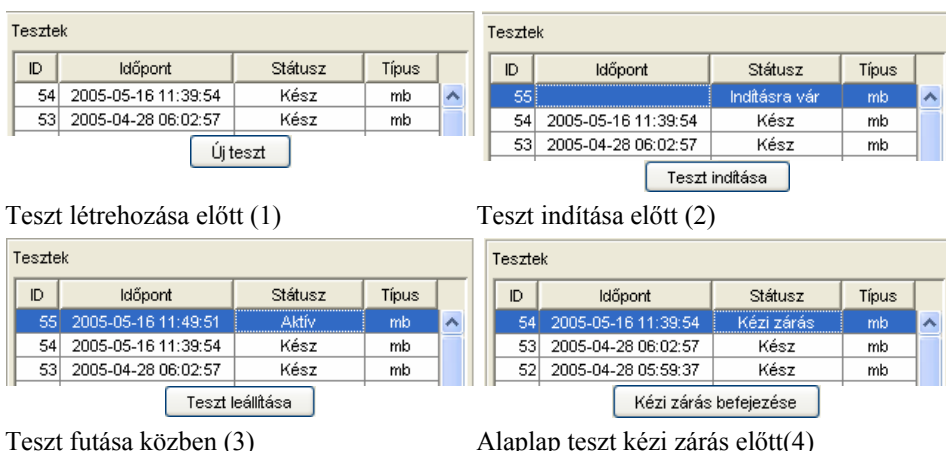


Figure 3: Test panel

Before creation of a new test(1), Before start(2), Under running(3), MB test Manual closing(4)

4. ábra

Táblázat

Státusz
Kész
Kész
Indításra váró

Figure 4: Table

A TESZT LÉPÉSEI

A merevlemez teszt során először mindig kiírásra kerül a merevlemez meghajtó alaplap csatlakozásának sorszáma (0–3), valamint a tesztel HDD modell pontos neve, sorozatszáma, a címezhető LBA szektorok száma és a tárolási kapacitás byte-okban.

Drive_Number: 1
 General_Information
 Model_Name: FUJITSU MHT2030AR
 Serial_Number: NN6QT4813YN1
 Addressable_LBA_Sectors: 58605120
 Storage_Capacity: 30005821440

A teszt következő lépésében a program egy egyszerű írás–olvasási tesztet hajt végre, ami közvetlenül szektor szinten címezi meg, a merevlemez meghajtót, és a kiírt és beolvasott adatok összehasonlítása alapján számlálja a hibásan visszaolvasott szektorok

számát. Az ellenőrzések mellett ez az írás-olvasási folyamat arra is szolgál, hogy a merevlemez meghajtó terhelés alatt legyen, az ezután következő SMART teszt számára.

Read/Write Test of drive 1

Write_Sector: 200000

Read_Sector: 200000

Error_Sector: 0

5. ábra

HDD-k panel

ID	Pozíció	Vonalkód	Modell név	Széria szám	LBA szektorok sz.	Kapacitás	Hibák sz.
3	0		✓ WDC WD100EB...	WD-WMA713982509	19541088	10005037056	2
4	0		✓ FUJITSU MHT2...	NN6QT4813YN1	58605120	30005821440	0

Hibajelzés (1) Megfelelőnek tekintett adat (2)

Hozzáadás Törlés Pozíció 3

Figure 5: HDD test results

Error(1), OK(2)

6. ábra

HDD Teszt panel

ID	Start	End	Hibák ...
5	08 07:17:35	06 23:00:52	✓ 0
7	07 17:07:36	06 23:00:52	✗ 1

Figure 6: HDD Test panel

7. ábra

S.M.A.R.T. panel

S.M.A.R.T.						
Attribútum	Státusz	Akt. ért.	Hiba kús...	Legjobb	Nyers ért.	Jelentés
Read Error Rate	b	✓ 200	51	200	0	
Spin Up Time	7	✓ 109	21	107	1900	Spin_Up_Time:1900,_Average_Spin_Up_Tim...
Start/Stop Count	32	✓ 96	40	96	4518	
Reallocated Sector Count	32	✓ 200	112	200	0	
Seek Error Rate	b	✓ 100	51	253	0	
Power-On Hours Count	32	✓ 89	0	89	8718	Power_On_Time:145_minutes,_18seconds

Figure 7: S.M.A.R.T panel

A teszt harmadik fázisa – a S.M.A.R.T teszt – adja a legtöbb értékelhető információt, amely információk viszont teljesen gyártó specifikusak, sőt egyes gyártókon belül még HDD típus specifikusak is. Így van ez a vizsgálat tárgyát képező két Hitachi és Fujitsu típusú merevlemez esetében is hiszen például a hőmérséklet még nincsen meg a Hitachi HDD esetén, amíg ez a Fujitsu HDD-nél már kiolvasható.

Ha az adott attribútum numerikus értéke eléri a threshold értékét, vagy az alá csökken, akkor a merevlemez már meghibásodás közeli állapotban van. Ezeket az értékeket használjuk a hiba jelzésére, hiszen számunkra a kritikusan alacsony illetve magas hőmérsékleti tartományban is lényeges a megbízható működés.

Az alaplaponk tesztelését az alaplaponk hardware szenzorai által szolgáltatott adatok alapján végezzük. A kiugró értékek alapján lehet megállapítani a hibát.

Mivel a tesztrendszerhez használt klímakamra mérete lehetővé teszi, hogy a vizsgálat során egyszerre akár 48 alaplaponk is működhessen, azaz pl. HDD esetén 192 db tesztelésére is lehetőség van. A hálózaton keresztül történő adatírás a kliensek egyszerre történő futása miatt időnként konfliktus helyzethez vezetett, bár a MySQL szervernek az ilyen eseteket le kellene kezelnie. A probléma kiküszöbölésére a szervernek használt számítógép erőforrásait tovább bővítettük, valamint a programba külön időzítést építettünk be. Az említett módosítások után a szerver maradéktalanul ki tudta szolgálni a klienseket.

ÖSSZEFOGLALÁS

Összefoglalóan megállapítható, hogy a megvalósított rendszer alkalmas a konkrét PC részegységek automatikus monitorozásra és vizsgálatára. A klimatikus tesztet a már régebben kidolgozott és megvalósított klímakamra alkalmazásával képes megvalósítani, az előre programozott körülmények között. A hálózati kapcsolat segítségével a feldolgozást végző számítógép folyamatosan hozzájuthat a vizsgált elemek adataihoz. A jelenleg elkészült programok alkalmasak az alaplaponk valamint a HDD-k tesztelésére. Cél úgy továbbfejleszteni az elkészült programokat, hogy további, előre meghatározott részegységek tesztelésére is alkalmas legyen, és így komplett rendszert alkosson. A fejlesztés során az egyes tesztekre vonatkozóan egyebek között el kell dönteni, hogy a tesztadatok írását a kliens vagy a szerver program fogja kezdeményezni, ugyanis ez a szerveroldali program fejlesztési irányát jelentősen befolyásolja.

KÖSZÖNETNYILVÁNÍTÁS

A projekt a KUKA Robotics Hungária Ipari Kft. támogatásával készül.

Levelezési cím (*Corresponding author*):

Lágymányosi Attila

Szent István Egyetem, Informatikai és Matematikai Intézet, Informatika Tanszék

2103 Gödöllő, Páter K. u. 1

Szent István University, Department of Informatics

H-2103 Gödöllő, Páter K. u. 1.

Tel.: 36-28-522-051 Fax: 36-28-410-804

e-mail: lagymanyosi.attila@gek.szie.hu