



Térbeli és időbeli elemzések multidimenzionális szemléletben

Kottyán László
PhD hallgató

Székesfehérvár, 2007.03.13.



A döntéstámogató rendszerek típusai

Azokat az információs rendszereket, amelyek segítik a döntési tevékenységeket összefoglalóan döntéstámogató rendszereknek nevezzük. A döntéstámogató rendszereket az alábbi általános kategóriákba sorolhatjuk:

- kommunikáció vezérelt: két vagy több felhasználó közötti kommunikációt, információcserét tesz lehetővé. Ide sorolhatók többek között a video- és audiókonferenciát biztosító eszközök, az elektronikus levelezés vagy az állományok, alkalmazások megosztását lehetővé tevő eszközök.
- adatvezérelt: hozzáférést és elemzési lehetőséget nyújt egy szervezet adataihoz. Ilyen eszköz az OLAP vagy az SDSS
- dokumentum vezérelt: a nem struktúrált audió, video és írott dokumentumok kezelésére helyezi a hangsúlyt.
- tudás vagy ismeretvezérelt: elsősorban a vezetőknek szolgáltat információkat heterogén adatforrásokból, speciális feladatok megoldására. Ide tartoznak az adatbányászati eszközök.
- modell-vezérelt: különböző modellek – pénzügyi, statisztikai, szimulációs – kezelésével támogatja a problémamegoldást. Ehhez a típushoz az OLAP tartozik.

(Power, 2002)



A térbeli döntéstámogató rendszerek, interaktív számítógép alapú rendszerek, amelyek félig strukturált térbeli döntési problémák megoldására alkalmasak. A térbeli döntési problémák jellemzői:

- a nagy számú döntési alternatívák,
- a célja vagy következménye a döntési alternatívának valamilyen térbeli változó,
- az alternatívák kiértékelése több kritérium alapján történik,
- a kritériumok mind mennyiségi, mind minőségi természetűek is lehetnek,
- tipikusan több döntéshozó vagy csoport vesz részt a döntési folyamatban,
- a döntéshozók különböző szempontokat helyeznek előtérbe,
- a döntésekre gyakran a bizonytalanság jellemző. (Stefanakis, 2002)
- Az SDSS-re jellemző, hogy félig strukturált térbeli adatokkal is dolgozik, ami azzal jár, hogy ezek az adatok nem feltétlenül szabályosak, esetleg hiányosak vagy nem illeszthetők szükségszerűen egy kötött sémába. (Stefanakis, 2002)



- Cél: a felső vezetés számára a stratégiai döntések meghozatalának, prognózisok felállításának lehetővé tétele, nagy mennyiségű – egy a vállalat napi működésében keletkező - adathalmaz elemzésével.
- OLAP rendszerek az adatok kezeléséhez adattárházakat használnak. Az adattárházak a multidimenzionális adatmodellen alapuló adatbázis-kezelő rendszerek.
- A multidimenzionális adatmodell alapeleme az adatkocka, amely celláiban a tényeket, a vizsgált adatokat tárolja. Az adatkocka jellemzője a dimenziói, amelyek függvényében elemezhetők a tényadatok. A kívánt információ az adatkockán végrehajtott OLAP műveletekkel nyerhető ki.

OLAP: On-line Analytical Processing, On-line elemzés-orientált rendszerek



A multidimenzionális adatmodell

Az adatmodellel kapcsolatos fontosabb fogalmak:

- **Cella** (cell): egy adat tároló egysége
- **Dimenzió** (dimension): Az információ kategóriája, például az idő dimenzió.
- **Tulajdonság** (attribute): A dimenzióon belüli . Például a Hónap az Idő dimenzió tulajdonsága.
- **Hierarchia** (hierarchy): egy dimenzió tulajdonságai között létrehozott kapcsolat, a hierarchia szintek meghatározzák az elemzés granularitását, részletezettségét. Pl.: az Idő dimenzió egy lehetséges hierarchiája Nap→Hónap→Év
- **Tény** (fact, keyfigure): az adatkockában tárolt, elemzett adatok
- **Műveletek:**
 - Roll-up: adatok összegzése vagy dimenzió redukció
 - Drill-down: a nagyobb szintű összesítés részletezése vagy új dimenzió bevezetése
 - Slice, dice: szelekció és projekció
 - Pivot: a megjelenítendő eredmény forgatása
 - Drill-across: egynél több ténytábla használata, kockák közötti elemzés
 - Drill-through: visszaadja az eredmény cella forrásadatait



Termelési adatok adatkocka

A termény dimenzió

G1
G2

Cellák a tényadatokkal: termelési mennyiségek

10	12	13	0	24
12	0	11	21	0

Az idő dimenzió

2005
2004
2003

A „hely” összetett dimenzió, parcellák és termőföldek

11	12	13	21	22
T1			T2	

- az SDSS-nek:
 - Hol van a...?
 - Mi változott ott...?
 - Melyik az optimális útvonal...?
 - Hol történt ajelenség?
- az OLAP-nak
 - Sikeresek voltak az elvégzett feladatok?
 - Hogyan alakult a termelés az elmúlt 3 hónapban?
 - Milyen kereskedelmi eredmények várhatóak?
 - Milyen volt az elmúlt üzleti év?

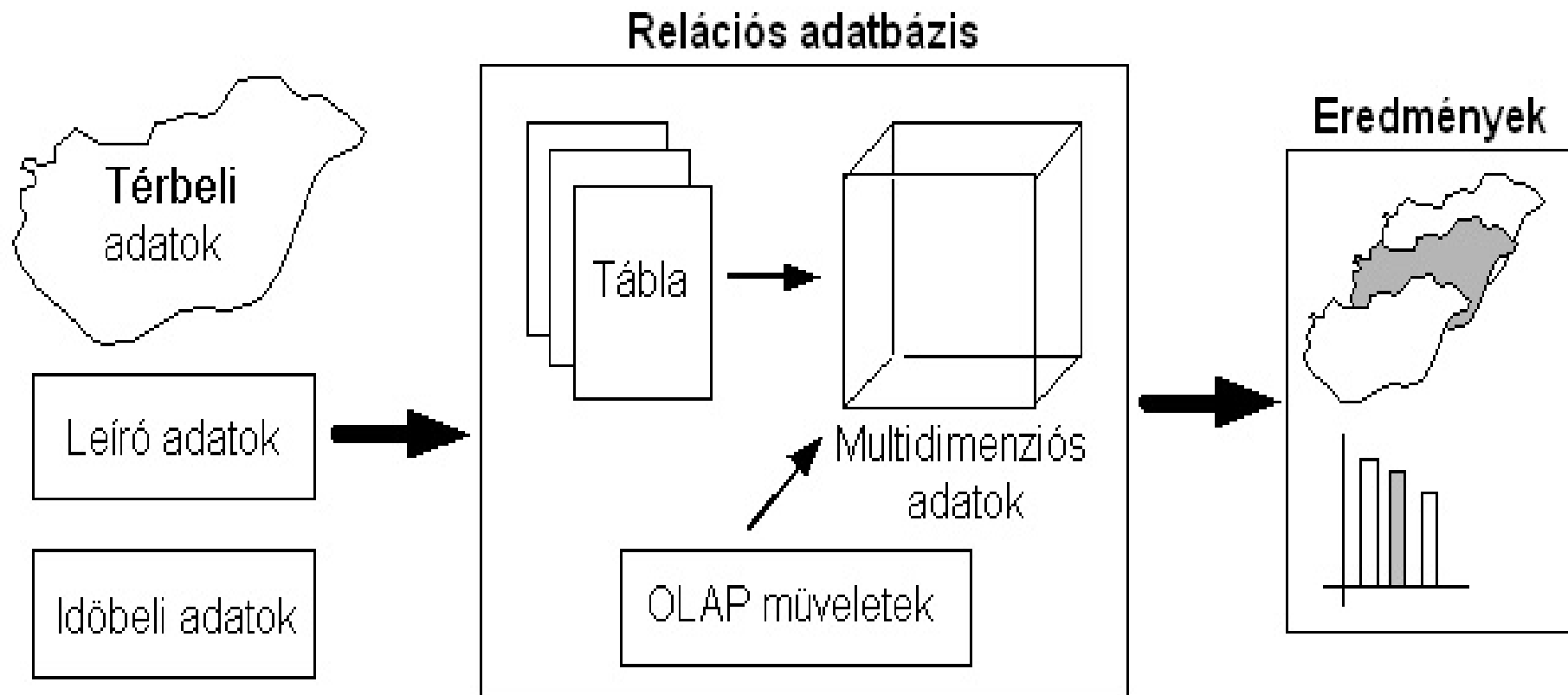


- Az OLAP és a GIS funkcionalitásának integrálására a 90-es években történtek az első kezdeményezések. A kezdeti megoldások azonban kimerültek abban, hogy az OLAP elemzések eredményeit térképen megjelenítsék, a GIS csupán térképnéző funkcióként szolgált. Később korlátozott képességekkel rendelkező GIS és OLAP kombinációk készültek, amelyek vagy az ún. OLAP-centrikus vagy a GIS-centrikus alkalmazások csoportjába tartoztak, attól függően, hogy mely funkciókra helyezték a hangsúlyt az alkalmazásokban (Bédard et al., 2007).
- Dr. Yvan Bédard (Laval Egyetem, Kanada) 1997-ben vezette be a SOLAP, Spatial OLAP kifejezést. A definíció szerint a *SOLAP egy vizuális platform, amely a multidimenzionális szemléletet követve elsősorban a gyors és egyszerű térbeli-időbeli elemzéseket támogatja különböző aggregációs szinteken, térképi, táblázatos valamint diagramos megjelenítéssel.*



- 2005-ben, 8 éves kutatási munka eredményeként egy készült el a kutató csoport saját fejlesztésű kereskedelmi alkalmazása a JMap SOLAP. A JMAP esetén, szemben a korábbi gyakorlattal, az objektumokat, mint példányokat az adatbázisban és a tényeket, mint példányokat a kockában tárolták. (Bédard, levelezés)
- A JMap SOLAP, Java környezetbe integrált alkalmazás jellemzői:
 - multidimenzionális adatstruktúra
 - kartográfiai és nem-kartográfiai megjelenítés (táblázat, diagram)
 - térbeli dimenziók
 - OLAP műveletek





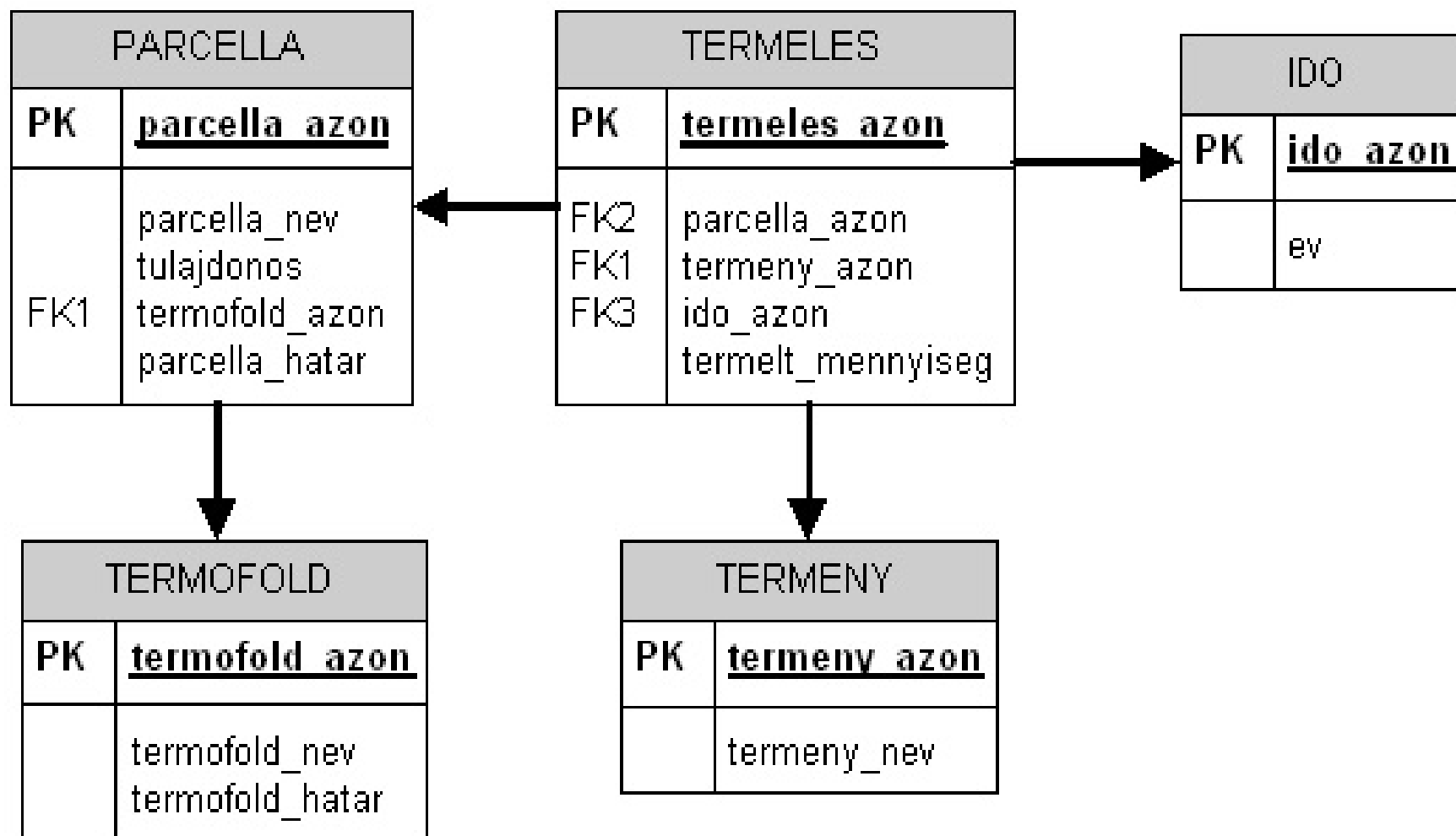
Feltételezésem szerint azoknál a feladatoknál, ahol a térbeli információk mellett fontos az időbeliség - például a hosszabb távon megvalósítható agrárprojektek vagy a környezeti monitoring rendszerek -, ***a multidimenzionális adatmodellre épülő térbeli-időbeli elemzések*** alkalmasak lehetnek arra, hogy a döntéshozók számára minőségi szempontból újabb, értékeesebb információk kinyerésével segítsék a döntéshozatalt.

Ezt szemlélteti a következőkben ismertetésre kerülő egyszerű példa, amelynek műveletei olyan információkat szolgáltatnak, ami ***kizárólag térbeli vagy SQL műveletekkel nem kinyerhető.***



- A termőföldek termelési adatainak elemzése és térképi megjelenítése a multidimenzionális adatmodellre alapozva.
- A termelési adatok elemzésére alkalmas adatmodell központi eleme a ténytábla, amely a **termelési mennyiségeket** tárolja.
- A termelési adatokat **három dimenzió**, a **termény**, az **idő** és a **helyre** vonatkozó, összetett - parcella és termőföld - dimenzió tükrében szemlélhetjük.
- A dimenziók az adatok jellegének megfelelően szöveges, szám és dátum típusú adatokat tartalmaznak, a PARCELLA táblában a **parcella_hatar**, a TERMOFOLD táblában a **termofold_hatar** attribútumok **geometriai adatok**.
- Fontos tulajdonsága a sémának a granularitása, amely a mintapéldában a térbeli dimenzió mentén teszi lehetővé az adatok összesítését és részletezését.





Geometriai adat: termofold_hatar,
parcella_hatar



Termelési adatok adatkocka

A termény dimenzió

G1
G2

Cellák a tényadatokkal: termelési mennyiségek

10	12	13	0	24
12	0	11	21	0

Az idő dimenzió

2005
2004
2003

A „hely” összetett dimenzió, parcellák és termőföldek

11	12	13	21	22
T1			T2	

Kiinduló adatok

T1

11	A
12	B
13	C

T2

21	22
A	C

Parcella	Termény
11	G1, G2
12	G1
13	G1, G2
21	G2
22	G1

Év
2003
2004
2005

	2003		2004		2005	
	G1	G2	G1	G2	G1	G2
11	10	12	9	11	13	13
12	12	-	11	-	14	-
13	13	11	12	10	13	12
21	-	21	-	20		23
22	24	-	24	-	25	-



- ***Mennyi volt az egyes termények hozama a T1 termőföldön összességében?***
- A térbeli drill-down művelet kiinduló állapotát a következő ábra mutatja. A térképen a termofold_hatar réteg jelenik meg, amely tartalmazza a termények összesített adatait az egyes területekre vonatkoztatva.



Kiinduló állapot

T1

G1=107

G2=69

T2

G1=73

G1=64

A művelet
végrehajtása a **T1**
termőföld
kiválasztásával
hajtható végre.





A drill-down eredménye térképen

Térbeli művelet eredményként a térképen megjelenik a parcella_hatar réteg a mennyiségi adatokkal valamint elkészül egy táblázat is .

T1

11	G1=32 G2=36
12	G1=37 G2=0
13	G1=38 G2=33



A drill-down eredménye táblázatban

Drill-down

	G1	G2
T1	107	69
T2	73	64



	G1	G2
11	32	36
12	37	0
13	38	33
21	0	64
22	73	0



Összetett lekérdezés 1.

- Az „A” tulajdonoshoz tartozó parcellák termelése évenkénti és terményenkénti bontásban

	2003		2004		2005	
	G1	G2	G1	G2	G1	G2
11	10	12	9	11	13	13
21	0	21	0	20	0	23

Térképi megjelenítés: 11-es parcella adatai (következő dia)



Összetett lekérdezés 2.

2003

11

G1=10 G2= 12

2004

11

G1=9 G2= 11

2005

11

G1=13 G2= 13



- A fenti példát megvalósító alkalmazás kialakítása még nem teljes, egyes részfeladatok azonban megoldottak.
- A kutatáshoz felhasznált eszközök:
 - PostgreSQL + PostGIS,
 - Feature Manipulation Engine,
 - Különböző formátumú térképi állományok,
 - ArcGIS.
- A megoldott feladatok:
 - térbeli adatok tárolása relációs adatbázis-kezelő rendszerben,
 - csillag és hópehely sémák alkalmazása, térbeli dimenziókkal,
 - egyszerűbb SQL lekérdezések és az eredmények térképi megjelenítése
- További kutatást igénylő kérdések:
 - OLAP műveletek, funkciók implementálása,
 - térbeli adatok (mértékek) szervezése, kezelése,
 - elemzési eredmények megjelenítése.





Felhasznált irodalom

1. Power, D. J., 2002, Decision Support Systems: Concepts and Resources for Managers, Westport, CT: Greenwood/Quorum Books, 2002.
2. Stefanakis, E., 2002, Semi-structured Data and XML in Geographic Data Modeling and Handling, Tutorial, Join International Symposium on Geospatial Theory, Processing and Applications, Ottawa, Canada, July 8, 2002.
3. Pastor, J., 2004. Conception d'une légende interactive et forable pour le SOLAP. Unpublished M.Sc. Thesis, Geomatics Sciences Department, Laval University., <http://www.theses.ulaval.ca/>, Accessed: May 2006.
4. Rivest, S., Y. Bédard, M.-J. Proulx, M. Nadeau, F. Hubert & J. Pastor, 2005, SOLAP: Merging Business Intelligence with Geospatial Technology for Interactive Spatio-Temporal Exploration and Analysis of Data, Journal of International Society for Photogrammetry and Remote Sensing (ISPRS) "Advances in spatio-temporal analysis and representation, Vol. 60, No. 1, pp. 17-33.
5. Bédard, Y., S. Rivest, & M.-J. Proulx, 2007, Spatial On-Line Analytical Processing (SOLAP): Concepts, Architectures and Solutions from a Geomatics Engineering Perspective, Dans: Robert Wrembel & Christian Koncilia (ed(s)), Data Warehouses and OLAP : Concepts, Architectures and Solutions, Chap. 13, IRM Press (Idea Group), London, UK, pp. 298-319.

