



M Ű E G Y E T E M 1 7 8 2

Budapesti Műszaki és Gazdaságtudományi Egyetem
Villamosmérnöki és Informatikai Kar
Irányítástechnika és Informatika Tanszék

SZAKÉRTŐI- ÉS ADATTUDÁS MODELLEZÉSE BAYES HÁLÓKKAL ÜGYFÉLNYILVÁNTARTÓ RENDSZEREKBE

Tézisfüzet

Vámos Gábor

Témavezető: Dr. Nagy Ákos egyetemi docens
BME Irányítástechnika és Informatika Tanszék

Budapest, 2006

Bevezetés

A közműszolgáltatók, a pénzügyi szolgáltatók, a telekommunikációs társaságok (továbbiakban szolgáltatók) *ügyfélnyilvántartó rendszereiket* elsősorban periodikus számlázási műveletek és egyéb belső-elszámolási feladatok kiszolgálására tartják fenn. Szempontunkból egy általánosan vett ügyfélnyilvántartó rendszer két fő komponensét a szolgáltató által tárolt *adatok*, valamint az ügyfélnyilvántartást vezető *alkalmazottak* képezik. Az adathalmaz egy olyan objektív ügyfélleírásnak tekinthető, amelynek összefüggéseiben kódolódik az *adattudás*. Az alkalmazottak az üzleti folyamatokról humán absztrakciójukkal alkotnak szubjektív képet. Ezt a leírást *szakértői tudásnak* nevezzük. E két tudástípus feltárásával, majd egyesített ábrázolásával képezhető az ügyfélnyilvántartó rendszer *tudásalapú modellje*, amely hatékonyan támogathatja a szolgáltatót munkafolyamataiban, összetett üzleti döntéseinek meghozatalában.

Egy tudásalapú modellel szemben – csakúgy, mint bármely matematikai modellel szemben – alapvető követelmény, hogy legyen a lehető legteljesebb, ugyanakkor ne legyen redundáns. Az ügyfélnyilvántartó rendszerek esetében olyan különmemű tudásforrások egységes kezelését kell megoldani, amelyek esetenként bizonytalanul, egymásnak ellentmondóan írják le a valóságot. Ilyenkor a determinisztikus modellreprezentációs eszközök kevésbé alkalmazhatók, helyettük a valószínűségi megközelítés képezhet releváns alternatívát. Az értekezés a valószínűségi tudásábrázolásra alkalmas univerzális approximátorok közül a *Bayes hálókra* szorítkozik.

A Bayes hálókkal kapcsolatos kutatások az 1980-as évektől indultak, szélesebb körű gyakorlati felhasználásuk az 1990-es évek második felében kezdődött el. Hatékony alkalmazhatóságuk elsőként olyan orvosi és gazdasági tématerületeken mutatkozott meg, amelyek megfigyelhetősége – esetenként a nem teljesen kidolgozott teoretikus háttér miatt - erősen behatárolt volt. Később fokozatosan teret hódítottak olyan adathalmazok kezelésében is, amelyek akár méretük, akár redundáns vagy hiányos jellegük miatt klasszikus statisztikai módszerekkel már nehezen voltak elemezhetőek [Russel03].

Egy döntéstámogató rendszerben a tudásbázis értéke messze a többi szoftverkomponens értéke fölött áll. A Bayes hálók igazi ereje éppen abban rejlik, hogy maga a tudásbázis natív tulajdonsága az *alkalmazás-függetlenség* és az *újrafelhasználhatóság*. Ez elsősorban abból következik, hogy egy Bayes hálóval implementált tudásbázisban nincs rögzített input-output mező. Feladatorientáltságot csak a tudásmodellen futatott

valószínűségi következtetések formájában implementált tudásbázis lekérdezések és e lekérdezések eredményét szemléltető felhasználói felületek mutatnak. A tudásábrázolást a Bayes hálók ráadásul úgy képesek megvalósítani, hogy a kialakított tudásmodell a későbbiekben is szerkesztető marad, hiszen az implementált tudásbázis az emberi gondolkodásmódhoz jól illeszkedik.

A Bayes hálók bizonyított tulajdonsága, hogy mind *a háló optimális generálása* [Chickering03], mind *a háló optimális kiértékelése* [Cooper90][Dagum93] megkötések nélkül *NP-nehéz* problémákra vezetnek. A széleskörű alkalmazhatóságukat ez a körülmény hosszú ideig gátolta, hiszen csak az ezredvégre fejlődött a számítástechnikai teljesítmény arra a szintre, amellyel a Bayes hálókkal kapcsolatos komplex számítási feladatok - kisebb problémaméretek esetén akár „brute force” megközelítéssel is - kezelhetővé váltak. Nagyobb problémateretek modellezésekor a számítási kapacitásigény heurisztikus hasznossági függvények alkalmazásával tartható kézben [Bouckaert95][Kjaerulff90]. Ilyenkor a heurisztika teremt egyensúlyt a modellezési részfeladat eredményének pontossága és a számításának költsége (idő, kapacitás) között.

Kitűzött kutatási feladatok

Az ügyfélnyilvántartó rendszereket olyan problématereteknek tekintjük, amelyekben a tárgyter minden fogalmára igaz, hogy adattudáson is és szakértői tudáson keresztül is megfigyelhető. A szakértői tudás jellemzően több szakértő ismeretanyagából áll össze, tehát valamely valóságfogalom több szakértő tudásán keresztül más-más aspektusból figyelhető meg.

A disszertáció az ügyfélnyilvántartó rendszerek Bayes hálóval történő modellezésének témakörében az alábbi feladatok megoldását célozza meg:

- több szakértő tudásának integrálása validált módon, az egyes szakértők kompetenciaszintjeinek figyelembevételével;
- a tudásmérnök támogatása - a számítási kapacitásigényt döntően befolyásoló - strukturális modellkomponens (továbbiakban DAG - irányított körmentes gráf) fejlesztésében;
- a tudásmodell létrehozásával és kiértékelésével kapcsolatos folyamatok számításigényének optimalizációja a két folyamat közötti kapcsolat megerősítésével.

Új tudományos eredmények

1. Tézis: Új modellezési eljárást dolgoztam ki az adattudás és a szakértői tudás együttes kezelésén alapuló valószínűségi tudásmodellezés folyamatára [2][12].

A tudásmodellt a létrehozástól az alkalmazásig a tudásmérnöki tevékenység (továbbiakban KEP - Knowledge Engineering Process) kíséri. A tudásmérnöki tevékenység leírására számos KEP módszertant dolgoztak ki [Taboada03]. Ezek többsége a szakértői tudás redundanciamentes, determinisztikus leírására kínálnak jól kiforrott, formalizált eszközkészletet. Az ügyfényilvántartó rendszerek leírásakor azonban bizonytalansággal terhelt tudásforrások kezelésére kell felkészülni, ezért célszerű valószínűségi modellreprezentációs eszközt alkalmazni. Az elmúlt negyed évszázadban a valószínűségi tudásmodellezés területén megjelent szakmai cikkek és dolgozatok nagyjából egy-egy algoritmus-specifikus problémakörre fókuszálnak, átfogó módszertani kérdésekkel csak elvétve foglalkoznak.

Az első tézis - a témakörben hiánypótló leírásként - egy átfogó KEP módszertant ad meg Bayes hálókkal implementált valószínűségi tudásbázisok építésére. Az így adódó modellezési eljárás alapján megfogalmazható algoritmusok számításigénye jelentősen csökkenthető.

A modellezési eljárás leírása korszerű UML eszközökkel történik, főbb elemei:

- tárgyterek osztályozása a fogalmak megfigyelhetőségi szintjei szerint;
- a modellezési folyamat humán szereplőinek és technikai entitásainak osztályozása;
- modellezés részfolyamatainak funkcionális definíciója;
- a részfolyamatok hierarchikus leírása a kapcsolódási pontok megadásával;
- a modellfejlesztés állapotainak és állapotátmeneteinek definíciója és az esetleges visszalépési lehetőségek rögzítése;
- a grafikus modellkomponens elemi modellinkrementumainak leírása.

Az első tézis modellezési eljárásrendszere a további tudományos eredmények tárgyalását alapozza meg.

2. Tézis: Új, több tudásforrás konzisztens kezelésére képes inkrementális algoritmust alkottam a Bayes hálóval reprezentált információban gazdag problématerék strukturális modellkomponensének fejlesztésére [5].

Több tényező is indokolja az adattudás és a szakértői tudás bevonását és együttes kezelését valamely ügyfélnyilvántartó rendszer strukturális modellkomponensének fejlesztésébe. Ezek közül a legfontosabb, hogy kizárólag az adattudásra épített struktúra-meghatározás esetén a keresési tér mérete szuper-exponenciális jellegű lenne a modellváltozók számát tekintve [Robinson76].

További fontos érv, hogy a szakértői tudás önmagában való alkalmazása nehézkessé tenné a modellezés folyamatát, hiszen az egyedi szakértők a valóságfogalmakról redundáns leírásokat adhatnak a humán interpretációs képességük különbözősége miatt [Feigenbaum84]. A kifejlesztett algoritmus az adattanításban tapasztalt szuper-exponenciális jelleg és a szakértői tudás-kinyerés során jelentkező bizonytalanság kézbe tartását célozza meg.

Lényeges tulajdonságok:

- az algoritmus a modellképzést iterációs lépésekre bontja: egy-egy iterációban a strukturális modellkomponens fejlesztése történik a szakértői tudás kinyerésével és reprezentációjával;
- minden iteráció végén a generált modellstruktúra objektív értékelése zajlik: ezzel funkcionálisan a kinyert tudás relevancia-vizsgálata valósul meg;
- a struktúra objektív értékelése az AIC [Akaike74] - Akaike Information Criterion - adattanító metrikán alapul: az AIC-ben az illeszkedést mérő komponens és a komplexitást értékelő komponens élesen elkülönül [Friedman91], így arányuk finomhangolható a szakértői egyedek korábban mutatott kompetencia-szintjeinek tekintetében;
- több szakértő, vagy több szakértői csoport tudásának együttes kezelése futási időben történik: esettanulmányokkal bizonyított, hogy az eltérő tudásforrások tervezési idejű összevezetése inkonzisztenciát okoz a függetlenségi viszonyok reprezentációjában [Richardson03][Williamson01].

A kifejlesztett algoritmus további előnye a szakirodalomban közölt eljárásokhoz képest az, hogy a szakértői tudás alapján az aktuális modellállapotot globálisan képes analizálni.

3. Tézis: Új, a releváns modellinkrementumok grafikus megjelenítésén alapuló algoritmust dolgoztam ki a Bayes hálóval reprezentált valószínűségi tudásmodell strukturális komponensének élhozzáadással történő fejlesztésére [1].

A szakértői tudás kinyerése és beépítése a strukturális modellkomponensbe történhet közvetett módon a tudásmérnök segédletével, illetve közvetlenül a szakértők által. Mindkét esetben célszerű egy grafikusan megjelenített nézetten dolgozni, mintsem közvetlenül a matematikai reprezentáción (szomszédsági mátrixon, szomszédsági listákon stb.).

A Bayes háló modellkomponenseinek megjelenítésére és szerkesztésére számos eszköz alkalmazható. Általánosan tapasztalt funkcionális hiányosságuk, hogy nem jelenítenek meg előzetes információt a kezdeményezett élhozzáadás műveletek irányított körmentességet befolyásoló hatásairól. A DAG kritérium ellenőrzése csak a már specifikált új jelölt élnek megfelelő csomópont-párra történik meg a gráf - valamely algoritmus szerinti részleges vagy teljes - bejárásával. Mivel a kezdeményezett szerkesztési műveletekre nem láthatók a lehetséges továbblépési alternatívák, ezért a modellfejlesztési munkafolyamat próbálkozás jelleggel zajlik. Különösen nagy csomópontszámú modellek esetében okoz ez problémát, mert tudásmérnök és a szakértők közötti együttműködés hatékonyságát erőteljesen csökkentheti az élberajzolás kísérletek gyakori kudarca.

A javasolt algoritmus a Bayes háló strukturális modellkomponensének szerkesztését támogatja. Az élberajzolás műveletek közben megjeleníti azon élpárokat, amelyek nem generálnak irányított kört; ezáltal a releváns modellfejlesztési alternatívák egyértelműen áttekinthetővé válnak. A modellinkrementumok megjelenítéséhez definiáltam az elérhetőségi mátrix adatstruktúra fogalmát. Az elérhetőségi mátrix létrehozására algoritmust fejlesztettem, tárgyárányára felsőkorlátot adtam. Futási tesztekkel megmutattam, hogy az algoritmus alapját képező adatstruktúra generálásának számításigénye a modellfejlesztés grafikus interakcióit nem gátolja.

Az így kialakuló eljárás az ismert módszerekhez képest lényegesen kisebb számításigényű, miáltal jelentősen csökkenthető a válaszidő.

4. Tézis: Új algoritmust dolgoztam ki a Bayes hálós tudásmodell strukturális komponensének éldékrementummal történő fejlesztésére [6][11].

A hagyományos megközelítés szerint a modellképzés eredménye a következtetési művelettől független Bayes háló. Így olyan modell alkotható, amely az applikációk közötti hordozhatóságot tökéletesen támogatja. Gyakran azonban a modellképzés pillanatában ismert és kötött a következtető algoritmus típusa. Ilyenkor nem követelmény az általános használhatóság, helyette hasznosabb egy olyan tudásmodell megalkotása, amely a rögzített következtetési algoritmusra optimalizált.

Az éltörléssel végrehajtott modellfejlesztési művelet során kifejezetten hasznos a következtetési algoritmus rögzítése. Az éltörlés jellemzően az inkrementális modellfejlesztés utolsó fázisaiban indokolt, amikor a valóság fogalmaira jellemző összefüggések feltárása teljessé válik. Ekkorra ugyanis néhány, már a kezdeti struktúrában feltűnő élről kiderülhet, hogy az újabb élekhez képest csak gyenge valószínűségi kapcsolatot reprezentál. Ha ezzel párhuzamosan a háló komplexitása az előre meghatározott maximumához közelít, akkor az újabb, erősebb élek berajzolásához a gyengének ítélt élek törlése szükséges. Ilyenkor a tudásmérnök feladata annak eldöntése, hogy mely élek távolítandók el. A közel azonos fontosságúnak ítélt gyenge élek közül történő választást támogathatja egy olyan mérőszám, amely élenként méri a következtetési műveletben okozott komplexitás-növekedést. A tudásmérnök így az éltörléssel párhuzamosan a következtetési művelet számításigényét is képes minimalizálni.

A tézis a Bayes hálóval reprezentált tudásmodell fejlesztési és felhasználási folyamataiban vizsgált optimalitási kritériumok között teremt kapcsolatot, feltételezve a szakirodalomban széleskörűen alkalmazott AIC minősítő függvény [Friedman91] és a PPTC következtető algoritmus [Gaag97] felhasználását.

Az új modellfejlesztő algoritmusban minősítő függvényként a triangularizációs heurisztikák által generált klikkméretet használtam, amely az aktuális modellinkrementum hatására bekövetkező modellkomplexitás-változást a következtetés várható számításigény-változásával arányos mérőszámmal méri. Ezáltal már a modellfejlesztés fázisában jelentősen csökkenthető a modell későbbi kiértékelésének számításigénye.

Az eredmények alkalmazása

Az első és a második tézis a BME Irányítástechnika és Informatika Tanszék és a DSS Consulting Kft. együttműködésével sikeresen lezárt „IKTA4-042 Nagy adatbázisok valószínűségi struktúrájának feltárása és humán interfész kiépítése” projekt keretében került kidolgozásra és felhasználásra [11][12][13][14][15]. Az általunk javasolt megközelítés szerint a Bayes hálós modell építése és a modell felhasználása (vagyis a következtetés) nem vizsgálható és végezhető önmagában. Ehelyett egy iteratív modellépítési folyamatot követ, amelynek bizonyos fázisaiban a majdani következtetési folyamat számításigénye előrevetítésre kerül; így lehetőség nyílik arra, hogy a modell bonyolultsága és a modell gyakorlati használhatósága között optimális egyensúly jöjjön létre.

A harmadik és negyedik tézis algoritmusai Matlab R14 környezetben implementálásra kerültek [6]. Az elkészült grafikus felület alkalmas a szomszédsági mátrixal reprezentált kvalitatív modellkomponens és a feltételes valószínűségeloszlás táblázatokkal megadott kvantitatív modellkomponens együttes kezelésére. A vizualizáció a széles körben alkalmazott document/view architektúrának megfelelően történik, így a valószínűségi modellezés számítási funkcióit szolgáló adatok (csomópontok adatai, struktúra adatok) és a megjelenítést támogató adatstruktúrák (elnevezések, színjelölések, pixelszintű információk, lokalitás) élesen elkülönülnek. A forráskód kommentekkel és függvénykönyvárral közelítőleg 4000 sorból áll.

A dolgozat módszertani megközelítésében több szolgáltató ügyfélnyilvántartó rendszere is sikeresen modellezésre került a 2002-2007-es időszakban [2][5][8][17][16][18]. Az alkalmazások elsődleges célja a fogyasztói visszaélések detektálásának támogatása. Ezúton is szeretnék köszönetet mondani partnereinknek bizalmukért és támogatásukért: ÉMÁSZ Nyrt., ELMŰ Nyrt., E.ON Zrt.

Hasznosulás az oktatásban

A disszertáció módszertani eredményei a BME Villamosmérnöki és Informatika Kar alábbi tárgyaiban kerül hasznosításra:

Valósídejű rendszerek és hálózatok (VIFO4364)

Informatika Szak, 8. Szemeszter Főszakirány, 5 Kredit, heti 4 előadás, a témakör a tárgy 50%-át teszi ki.

URL: <https://www.vik.bme.hu/kepzes/targyak/VIFO4364/>

Integrált energetikai rendszerek tervezése és irányítása (VIIM409)

Informatika Szak, 8. Szemeszter Főszakirány, 5 Kredit, heti 4 előadás, a témakör a tárgy 50%-át teszi ki.

URL: <https://www.vik.bme.hu/kepzes/targyak/VIIM409/>

Vállalat és termelésirányítási laboratórium (VIFO5305)

Informatika Szak, 9. Szemeszter Főszakirány, 3 Kredit, heti 2 LABOR, a témakör a labortárgy 33%-át teszi ki.

URL: <https://www.vik.bme.hu/kepzes/targyak/VIFO5305/>

Hivatkozások

- [Russel03] S. J. Russel, P. Norvig: *Artificial Intelligence: A Modern Approach*, Prentice Hall Inc., 2nd edition, 2003.
- [Chickering03] D. M. Chickering, C. Meek and D. Heckerman: Large-Sample Learning of Bayesian Networks is NP-Hard, *In Proceedings of Nineteenth Conference on Uncertainty in Artificial Intelligence*, pp 124-133, Morgan Kaufmann, Acapulco, Mexico, 2003.
- [Cooper90] G. F. Cooper: The Computational Complexity of Probabilistic Inference Using Bayesian Belief Networks, *Artificial Intelligence*, 42 (2-3), pp. 393--405, 1990.
- [Dagum93] P. Dagum, M. Luby: Approximate Probabilistic Inference in Bayesian Networks is NP Hard, *Artificial Intelligence*, 60, pp. 141-153, 1993.
- [Bouckaert95] R. R. Bouckaert: Bayesian Belief Networks: From Construction to Inference, *PhD Thesis, University to Utrecht*, 1995.
- [Kjaerulff90] U. Kjaerulff: Triangulation of Graphs - Algorithms Giving Small Total State Space, *Technical Report TR R 90-09, Department of Mathematics and Computer Science, Strandvejen, Aalborg, Denmark*, 1990.
- [Taboada03] M. Taboada, M. Argüello, J. Des, J. Mira: Building Knowledge-Based Intelligent Systems by Reusing, *In Innovations in KE, International Series on Advanced Intelligence*, Vol. 82, pp. 1-30, 2003.
- [Robinson76] R. D. Robinson: Counting Unlabeled Acyclic Digraphs, *In Proceedings Australian Conference on Combinatorial Mathematics*, Vol. 5, pp. 28-43, 1976.
- [Feigenbaum84] E. A. Feigenbaum: Knowledge Engineering: The Applied Side of Artificial Intelligence, *Annals of the New York Academy of Sciences*, 426, pp. 91-107, 1984.
- [Akaike74] H. Akaike: A new Look at the Statistical Model Identification, *IEEE Transactions on Automatic Control* 19, pp. 716-722, 1974.

- [Friedman91] J. Friedman: Multivariate adaptive regression splines (with discussion), *Annals of Statistics* 19, pp. 1-67, 1991.
- [Richardson03] M. Richardson, P. Domingos: Learning with Knowledge from Multiple Experts, *In Proceedings of the Twentieth International Conference on Machine Learning (ICML-2003)*, pp. 624-631, Washington DC, 2003.
- [Williamson01] J. Williamson: Foundations for Bayesian networks, *In eds. David Corfield & Jon Williamson Foundations of Bayesianism*, pp. 75-115, *Kluwer Applied Logic Series*, 2001.
- [Gaag97] L. C. van der Gaag, H. L. Bodlaender: Comparing Loop Cutsets and Clique Trees in Probabilistic Inference, *Technical Report UU-CS-1997-42, Utrecht University: Information and Computing Sciences*, 1997.

Publikációk

Idegen nyelvű folyóiratcikk

- [1] Vámos: A Matlab Graphical Tool to Support Knowledge Engineering, *Production Systems and Information Engineering, Vol. 4*, pp. 81-93, Miskolc University Press, 2006, ISSN 1785-1270

Magyar nyelvű folyóiratcikk

- [2] Vámos-Beck: Fogyasztók viselkedésének tudásalapú modellezése, *Elektrotechnika* 99. évf. 2006/3, pp. 17-18, 2006, HUISSN 0367-0708

Nemzetközi konferencia-kiadványban megjelent idegen nyelvű cikk

- [3] Kovács-Kiss-Nagy-Vámos: Early Detection System for Vegetation Fire in the Aggtelek National Park, *Proceedings of IEEE TEHOSS'2005*, pp. 493-498, Gdańsk, 2005.09.28-30. ISBN 83-917681-9-8.
- [4] Kovács-Kiss-Nagy-Vámos: Autonomous Fire Detection and Warning System for Early Recognition of Vegetation Fires, *Proceedings of MicroCad 2005 International Scientific Conference*, pp. 241-246, Miskolc, 2005.03.10-11. ISBN 963-661-646-9.
- [5] Arató-Kiss-Vajta-Vámos: Fraudulent Consumer Behaviour Analysis and Detection for Utility Companies, *Proceedings of 2nd Conference on Information Technology*, pp. 37-43, Gdańsk, 2004.05.16-18. ISBN 83-917681-5-5.
- [6] Vámos-Nagy-Kiss: GUI Tool for Manipulate and Refine Bayesian Networks, *Proceedings of Second Hungarian Conference on Computer Graphics and Geometrics*, pp. 65-70, Budapest, 2003.06.30.-2003.07.01. ISBN 963-420-766-9.
- [7] Vámos-Nagy-Kiss: Bayesian Network Based Modelling for Flaw Detection in Metallic Fusion Welds USING X-Ray Images, *Proceedings of 4th Workshop on European Scientific and Industrial Collaboration Promoting Advanced Technologies in Manufacturing*, pp. 181-186, Wesic, Miskolc, 2003.05.28-30. ISBN 963-961-570-5.

- [8] Vámos-Nagy-Kiss: Creating Bayesian network based probabilistic models for practical applications, *Proceedings of MicroCad 2003 International Scientific Conference*, pp. 163-168, Miskolc, 2003.03.6-7. ISBN 963-661-547-0.
- [9] Vámos-Loványi-Nagy-Kiss: Flaw Detection in Metallic Fusion Welds on X-ray Images Using Bayesian Networks, *Proceedings of First Hungarian Conference on Computer Graphics and Geometrics*, pp. 118-123, Budapest, 2002.05.28-29. ISBN 963-420-718-9.
- [10] Vámos-Nagy: Cardholder Identification Systems based on Behaviour Analysis and Biometric Methods, *Proceedings of MicroCad 2001 International Scientific Conference*, pp. 177-182, Miskolc, 2001.03.01-02. ISBN 963-661-457-1.

Technical Report

- [11] Vámos-Kiss: Nagy adatbázisok valószínűségi struktúrájának feltárása és humán interfész kifejlesztése – Részletes Kutatási Jelentés
IKTA4-042, 2002.12.15
- [12] Vámos-Kiss-Aszalós: Nagy adatbázisok valószínűségi struktúrájának feltárása és humán interfész kifejlesztése - Módszertani Tanulmány
IKTA4-042, 2002.07.1.

Csak szóban elhangzott előadás

- [13] Vámos-Nagy: Nagy adatbázisok valószínűségi struktúrájának feltárása és humán interfész kifejlesztése – Projektzáró prezentáció, IKTA4 FÓRUM, Budapest, 2004.04.15
- [14] Kiss-Nagy-Egri-Vámos: Nagy adatbázisok valószínűségi struktúrájának feltárása és humán interfész kifejlesztése – Projektbemutató prezentáció, IKTA4 FÓRUM, Budapest, 2002.11.20.
- [15] Aszalós-Vámos: Nagy adatbázisok valószínűségi struktúrájának feltárása és humán interfész kifejlesztése – Projektbemutató prezentáció, IKTA4 FÓRUM, Budapest, 2002.02.07.

Könyvtárakban el nem helyezett kutatási jelentés

- [16] Vámos-Kiss-Vajta: A lakossági ügyfélkör ellenőrzési címlistáinak generálása, Projektzáró dokumentum, 2004.11.25.
- [17] Vámos-Kiss-Vajta: Az áramszámlázási adatbázis hihetőség-vizsgálata, Projektzáró dokumentum, 2003.09.25.

Csak kivonatban megjelent konferencia-előadás

- [18] Vámos-Beck: Szabálytalan vételezés felderítésének támogatása az ügyfelek viselkedésének valószínűségi modellezésével, *Magyar Elektrotechnikai Egyesület 52. Jubileumi Vándorgyűlés Konferencia kiadványa*, 2005