

Kapacitásmodellezés és az IT-architektúra-tervezés kihívásainak vizsgálata statisztikai és prediktív analitikai eszközökkel

Fehér Péter

PhD, a Budapesti Corvinus Egyetem egyetemi docense
E-mail: peter.fehér@uni-corvinus.hu

Kő Andrea

PhD, a Budapesti Corvinus Egyetem egyetemi tanára
E-mail: andrea.ko@uni-corvinus.hu

Szabó Zoltán

PhD, a Budapesti Corvinus Egyetem egyetemi docense
E-mail: zoltan.szabo@uni-corvinus.hu

A minőségi internetbankos szolgáltatások biztosításáért felelős informatikai infrastruktúra kapacitásának tervezése összetett feladat. Az egyik alapvető kérdés ezen a területen az, hogy milyen tényezők befolyásolják a szolgáltatások minőségét, milyen kapcsolat van az informatikai infrastruktúra és az infrastruktúrát kiaknázó elektronikus banki szolgáltatások között. A cikkben bemutatott kutatás az elektronikus pénzügyi szolgáltatásokhoz kapcsolódó tényleges kapacitásterhelések modellezésére, előrejelzésére vállalkozott egy hazai bank szolgáltatásainak kontextusában. A modellezés egyik sajátosságát az adta, hogy túl az üzleti elemzésekben megszokott statisztikai modelleken, a prediktív analitika területéhez tartozó neurális hálózatok is felhasználásra kerültek. A vizsgálat további jellemzője a modellezéshez kapcsolódó szakterület egyedisége. A szakirodalmi háttérrel áttekintve kevés olyan modell található, amely komplex informatikai infrastruktúrák kapacitásmodellezését vizsgálja.

TÁRGYSZÓ:
IT-architektúra.
Neurális hálózat.
Kapacitástervezés modellezése.

DOI: 10.20311/stat2016.11-12.hu1149

A hazai internetbanki szolgáltatások egyre népszerűbbek, az ügyfelek számának jelentős növekedése tapasztalható az utóbbi időszakban is. Már 2009 végén a lakossági internetbanki ügyfelek aránya átlagosan 35 százalék volt az internetes szolgáltatást nyújtó bankok összes lakossági ügyfeléhez viszonyítva, a vállalati ügyfelek esetén ez az arány 45 százalékot ért el (*GKIeNET-T-Home-T-Mobil* [2009]). A legnagyobb hazai bank napi internetbanki belépéseinek a száma 2012-ben elérte a 121 ezret, a felhasználók száma a 2 milliót, míg a tranzakciók száma megközelítette a 11 milliót (*OTP* [2014]). Az elektronikus csatornák igénybevételét a szolgáltatók is ösztönzik, sok esetben kedvezményeket adnak azoknak a felhasználóknak, akik a papíralapú adatfeldolgozást elektronikusra cserélik. Az elektronikus szolgáltatásoknak kétségtelenül vannak előnyei, többek között a papíralapú adatfeldolgozás költségeinek elkerülése, a szolgáltatások eredményesebb nyomon követhetősége, az audit követelményeinek való megfelelés támogatása. Az internetbanki szolgáltatások megfelelő szintű biztosításában az informatikai infrastruktúrának meghatározó szerepe van. A háttér infrastruktúra komplex, nehéz behatárolni az egyes elektronikus szolgáltatások erőforrás-felhasználását, holott ez szükséges lenne a ráfordítások tervezéséhez. A technológia komplexitásának növekedése nehezíti a kapcsolódó infrastruktúra komponenseinek teljesítmény- és kapacitásmonitorozását (*Metzler* [2003]). A minőségi szolgáltatások biztosításáért felelős informatikai infrastruktúra kapacitásának tervezése kihívásokkal terhelt, fontos feladat. Az egyik alapvető kérdés ezen a területen az, hogy milyen tényezők befolyásolják a szolgáltatások minőségét. A kapacitásmenedzsment számos tevékenységet lefed, felelős azért, hogy az IT- (informatiótechnológiai) szolgáltatások és az IT-infrastruktúra kapacitása képes legyen a megállapodás szerinti szolgáltatási célkitűzéseket költséghatékonyan és időben nyújtani. Kiterjed minden olyan erőforrásra, ami az IT-szolgáltatás nyújtásához szükséges, valamint a rövid, közép- és hosszú távú üzleti követelményeket megvalósító tervekre is (*itSMF Hungary* [2009]). A kapcsolódó informatikai beruházások (többek között idetartozik az IT-infrastruktúra bővítése is) hasznait nehéz megfogható módon kimutatni, bár ez egyre inkább elvárásként jelentkezik.

A cikk az informatikai infrastruktúra és az infrastruktúrát kiaknázó elektronikus banki szolgáltatások kapcsolatának elemzési tapasztalatait foglalja össze. A vizsgálat két kulcskérdése:

- Az infrastruktúra-terhelések változása hogyan hat az elektronikus banki szolgáltatások viselkedésére?
- Van-e a kimutatható kapcsolat az erőforrás-felhasználás és a szolgáltatási szintek között?

A kutatás az elektronikus pénzügyi szolgáltatásokhoz kapcsolódó tényleges kapacitásterhelések modellezésére, előrejelzésére vállalkozott egy hazai bank szolgáltatásainak kontextusában. A modellezés egyik sajátosságát az adta, hogy túl az üzleti elemzésekben megszokott statisztikai modelleken, a prediktív analitika területéhez tartozó neurális hálózatokat is felhasználtuk. A kutatás további jellemzője a modellezéshez kapcsolódó szakterület egyedisége. A szakirodalmi háttérrel áttekintve kevés olyan modell található, amely komplex informatikai infrastruktúrák kapacitásmodellezését vizsgálja.

A tanulmány további fejezeteiben áttekintjük a kutatáshoz kapcsolódó szakirodalmat, bemutatjuk a vizsgált szolgáltatásokat és modellezési környezetüket. Ezt követi az elemzések és eredmények ismertetése. Végül összegezzük a tapasztalatokat, és kitérünk a lehetséges további fejlesztési irányokra is.

1. A kutatáshoz kapcsolódó irodalom áttekintése

Ebben a fejezetben bemutatjuk a kutatáshoz kapcsolódó fontosabb szakirodalmi területeket. Ismertetjük az IT-szolgáltatásmenedzsment vonatkozó részterületeit, a banki környezetekben jelenlevő komplex informatikai infrastruktúra jellemzőit. Ismertetjük a prediktív modellezésben alkalmazott neurális hálózatok sajátosságait és kitérünk a lehetséges felhasználási módok korlátaira is.

1.1. IT-szolgáltatásmenedzsment

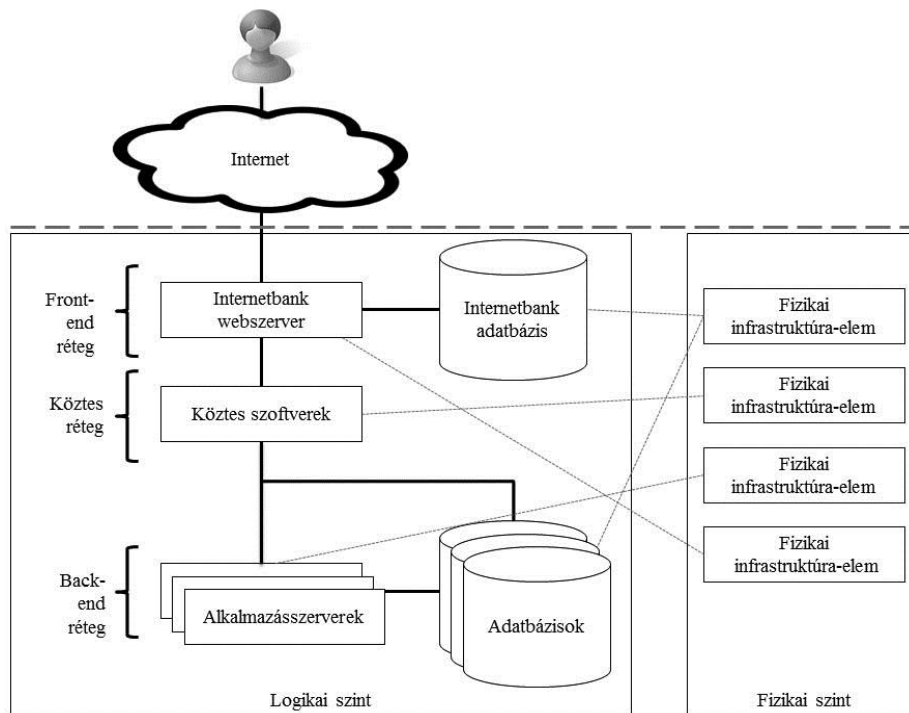
Az informatikai szolgáltatások rendelkezésre állása, illetve a felhasználók által érzékelt minősége a támogató infrastruktúra megbízható működésére, ezen belül egyre inkább az informatikai kapacitások megfelelő kihasználására épül. Emiatt az egyes informatikai szolgáltatások és az infrastruktúra-elemek monitorozása egyre fontosabbá válik az üzletileg kritikus alkalmazások esetében. A banki környezetekben az informatikai háttér elengedhetetlen feltétele a szolgáltatások nyújtásának. A technológiai megoldások komplexitásának növekedésével az egyes teljesítmények monitorozása mind nagyobb kihívást jelent a szervezetek számára (*Metzler [2003]*), alapvetően a következő szakaszban bemutatott okok miatt.

1.1.1. Megosztott környezet

A nagyvállalati környezetekre jellemző az infrastrukturális erőforrások közös, megosztott használata, azaz egyetlen fizikai erőforrást több virtuális szerver, ezen

keresztül szolgáltatás használ fel. Kutatásunk során egy banki környezetben vizsgáltuk a fizikai infrastruktúra, a logikai infrastruktúra és a szolgáltatások kapcsolatát. Jellemzően egy banki szolgáltatást az ügyfél a front-end rendszeren keresztül (kliens szoftver, böngésző) ér el, majd az adatokat, kéréseket egy köztes (middleware) réteg juttatja el a feldolgozó alkalmazásokhoz (back-end réteg). Az alkalmazások a feldolgozás során maguk is meghívhatják más alkalmazások funkcióit.

1. ábra. A nagyvállalati architektúra



Forrás: Saját szerkesztés.

1.1.2. Változó infrastruktúra

A komplex infrastruktúra ugyanakkor jellemzően nem tekinthető állandónak, hiszen annak egyes elemei (például a CPU-k (central processing unit – központi feldolgozóegység) száma) folyamatosan változik. Annak érdekében, hogy az egyes mérési eredmények összehasonlíthatók legyenek, szükséges, hogy a mért eredményeket egységes infrastruktúra-környezetre transzformáljuk (például egy CPU-ra jutó kapacitáskihasználtság). Ennek előfeltétele, hogy a mérési időszakra vonatkozólag teljes mértékben rendelkezésre álljanak az infrastruktúra változásai.

1.1.3. Mérési területek

A vállalatok által használt mérések jellemzően az alapinfrastruktúrára vonatkoznak, ugyanakkor keveset foglalkoznak a végfelhasználók által érzékelt teljesítmény-problémákkal. Az infrastruktúrakapacitás-leterheltség, illetve a szolgáltatások észlelt minősége között ugyanakkor az összefüggés nem egyértelmű, ezért csupán az egyik oldal vizsgálatával nem kapunk teljes képet (*HP* [2008]).

1.1.4. Informatikai szolgáltatások értelmezése

Annak érdekében, hogy a mérések értelmezhetőek legyen, szükséges annak tisztázása, hogy az egyes informatikai szolgáltatásokba milyen műveletek, feldolgozások tartoznak, illetve mely erőforrásokat érintik (*HP* [2012]).

A hálózati szolgáltatások esetében kérdés, hogy az infrastruktúra mely szeleteire van az informatikának közvetlen befolyása (így képes beavatkozni incidensek esetén), és melyek függetlenek az informatikai szolgáltató szervezettől.

A mérések nem tartalmazhatnak az informatikai szolgáltató szervezettől független elemeket, mivel ezekkel kapcsolatban sem megfelelő mérés, sem beavatkozás nem lehetséges (például interneten keresztül nyújtott szolgáltatás esetében a válaszidőt torzíthatja az internetszolgáltatás minősége, a felhasználó saját hálózati környezete vagy például a felhasználó által igénybe vett számítógép).

Ennek megfelelően a szolgáltatási mérések felosztása szükséges (*HP* [2012]). A mérések kiterjesztése lehet annak vizsgálata, hogy a független elemek milyen mértékben torzítják egy adott szolgáltatás minőségét, és melyik terület jelenti a szűk keresztmetszetet.

1.2. A neurális hálózatok és modellezési sajátosságaik áttekintése

Az ANN (artificial neural network – mesterséges neurális hálózatok), az emberi idegsejtek működését modellezve, egymással kommunikáló egyszerű processzorokból (mint mesterséges neuronokból) felépülő párhuzamos működésű architektúrák (*Anthony–Bartlett* [2009]). A neurális hálózatok irodalma gazdag, topológiájukat, tanítási algoritmusait és felhasználási lehetőségeiket is sokan elemzik (*Aggarwal* [2015], *Altrichter et al.* [2006], *Borgulya* [1998], *Zeng–Wang* [2010], *Karayiannis–Venetsanopoulo* [1992], *Sharda–Delen–Turban* [2015]). Az első mesterséges neurális hálózatot bemutató tanulmány *McCulloch* és *Pitts* [1943] nevéhez fűződik. Ez a hálózat még számos korlátozó tényezővel rendelkezett, nem tudták jól hasznosítani a gyakorlatban sem, de inspirálta a további kutatásokat. *Werbos* [1974] fejlesztette ki a back-propagation algoritmust, amely az egyik legnépszerűbb tanulási algoritmus lett,

és jelentősen hozzájárult a mesterséges neurális hálózatok népszerűsítéséhez. Hopfield [1982] megoldást adott az input és output adatok közötti nemlineáris kapcsolatok kezelésére, míg Zupan és Gasteiger [1991] munkáiban megjelent a többrétegű neurális hálózat koncepciója.

A neurális hálózatok mind topológiájukban, mind tanítási paramétereikben testre szabhatók. Osztályozhatók a kezelt probléma típusa, a topológia (rétegek és a rétegekben tárolt neuronok száma), az őket alkotó neuronok kapcsolódási jellemzői, a tanítási algoritmusuk és a tanítás módja szerint (Kő [1992]). Egyik legnépszerűbb sajátosságuk, hogy tanuló rendszerek lévén, a változó struktúrájú kapcsolatrendszerket természetes és hatékony módon kezelhetjük velük. Nemlineáris jellegükből fakadóan gyakorlatilag tetszőleges bemenet-kimenet leképezés leírására alkalmazhatók. A neurális hálózatokban az információk a neuronok hálózatában elosztva, a kapcsolódó súlyok segítségével dinamikusan őrződnek. Az új információk ezeknek a súlyoknak a megváltozásával – tanulással – hozzáadódnak a régiekhez (Rojas [1996]).

Egy neurális hálózat bemenő vektorát $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$ módon jelölve, az \mathbf{x}_i az i -edik bemenő érték ún. aktivációs szintjét jelenti. Minden összekötött neuronpárhoz tartozik egy változtatható (vagy változó) érték, amelyet súlynak neveznek. Szokás ezt a kapcsolat erősségének vagy hosszú távú memóriának is nevezni. A j -edik neuron kimenetének kialakításában részt vevő súlyok egy $\mathbf{w}_j = (\mathbf{w}_{1j}, \mathbf{w}_{2j}, \dots, \mathbf{w}_{nj})$ vektort alkotnak, amelyben a \mathbf{w}_{ij} értékek az i -edik (külső bemenetből vagy neurontól származó) \mathbf{x}_i értéknek a j -edik neuron y_j kimenetére gyakorolt hatását fejezik ki. Gyakran használatos még a q_j küszöbérték, amely a hozzá tartozó \mathbf{w}_{0j} súllyal módosítja a j -edik neuron kimenetét. Ez az érték gyakran tekinthető olyan aktivációs küszöbnek, amelyet a beérkező jelek együttes szintjének meg kell haladnia ahhoz, hogy kimenő jel keletkezzék. A j -edik neuron y_j kimenő értékét tehát az \mathbf{x} bemenő vektor, a kapcsolódó \mathbf{w}_j súlyvektor (ha van, a q_j küszöbérték a hozzá tartozó \mathbf{w}_{0j} súllyal) és egy függvény (leképezés) határozzák meg. A számítás menete legtöbbször a következő: az \mathbf{x} és \mathbf{w}_j vektorok skaláris szorzatának eredményéből levonjuk a súlyozott küszöbértéket ($\mathbf{w}_{0j} q_j$), majd az eredményt egy nemlineáris f függvénytranszformációnak vetjük alá. Az f ún. aktivációs függvény alkalmazásával a kimenet (y_j) így a következő lesz:

$$y_j = f\left(\sum_{i=1, n} \mathbf{x}_i \mathbf{w}_{ij} - \mathbf{w}_{0j} q_j\right) \text{ vagy}$$

$$y_j = f\left(\mathbf{x} \mathbf{w}_j - \mathbf{w}_{0j} q_j\right).$$

A jelölést egyszerűsíti, ha -1 -nek választva a q_j -t beolvasztjuk az \mathbf{x} bemeneti vektorba, a \mathbf{w}_{0j} -t pedig a \mathbf{w}_j súlyvektorba. Ekkor az \mathbf{x} és \mathbf{w}_j vektorok skaláris szorzata az aktivációs függvény argumentuma:

$$y_j = f\left(\sum_{i=1, n} \mathbf{x}_i \mathbf{w}_{ij}\right) \text{ vagy}$$

$$y_j = f(\mathbf{x} \mathbf{w}_j).$$

Természetes általánosítás lenne, ha a bemenetek lineáris kombinációja helyett másfajta aggregálást választanánk – például valamilyen más átlagból kiindulva, mint a számtani átlag – viszont akkor az egyszerű szerkezetből származó előnyöket elveszítenénk. A tipikus átviteli, más néven aktivációs függvények lineárisak, lépcsősek vagy szigmoid jellegű függvények (Sántáné-Tóth *et al.* [2007], Kő [1992]).

A neurális hálózatok felépítéséből következik, hogy jó hibatűrő képességgel rendelkeznek. Mivel az információk a neuronok között elosztva tárolódnak, viszonylag sok elem kiesése sem veszélyezteti a rendszer működését. A neurális hálózatoknak nincs irányító központjuk, minden művelet szigorúan lokális, csak a közvetlen kapcsolatban levő elemektől függ (Sántáné-Tóth *et al.* [2007]). A különböző paradigmákkal az osztályozási – felügyelet nélküli és a felügyelt modellekkel –, a függvényillesztési feladatokat hatékonyan lehet elvégezni (Cheng–Titterington [1994]). A felügyelt tanulás (supervised learning), amelyet prediktív vagy következtető elemzésnek is neveznek, egy „tanító” által irányított tanulási folyamat, vagyis a tanulási folyamatot leíró algoritmus paramétereit a „tanító” adja meg. A tanulás tanító példák halmaza (training set) alapján történik, ahol a példák attribútumai közül az egyik a többi attribútum értékétől függő ún. célattribútum. Formálisan a felügyelt tanulás algoritmusai egy függvényt (mint hipotézist) határoznak meg (Rojas [1996]).

– Legyen adott m számú tanító példa (x_i, y_i) változópárok formájában $i = 1, \dots, m$.

– A tanulás során előállítandó az az $y = f(x)$ függvény, amelyre minden (x_i, y_i) változópárra $y_i = f(x_i)$ $i = 1, \dots, m$.

– A kapott f hipotézistől elvárjuk azt, hogy az adott szituációt jellemző további változópárokra is jó becslést adjon. Amennyiben több kielégítő hipotézist kapunk, dönthetünk a legegyszerűbb függvény választásáról.

Az x_i értéke helyett gyakran a bemenő adatok kifejezést használják, míg az y_i értékét kimenő adatnak, célattribútumnak vagy hipotézisnek nevezik. Amennyiben a

függvény nem folytonos, és az y_i értékek száma kevés, szokás ezeket az értékeket osztályoknak nevezni; ekkor azt mondjuk, hogy az f függvény minden x_i értékhez hozzárendel egy osztályt – vagyis osztályoz (*Sántáné-Tóth et al.* [2007]).

A mesterséges neurális hálózatok ismertebb felhasználási lehetőségei a beszéd- és alakfelismerés, robottechnika, kódolás, kódok megfejtése, gépi fordítás, osztályozási feladatok megoldása, optimalizálások, térbeli interpoláció, különböző gazdasági problémák megoldása, tőzsdei és piaci előrejelzések (*Tsai-Wu* [2008], *Fajszí-Cser-Fehér* [2010], *Sántáné-Tóth et al.* [2007]).

A neurális hálózatok közgazdasági, modellezési alkalmazásának vannak korlátai és hátrányai is, például az, hogy a megtalált strukturális kapcsolatokat neuronokra szétosztottan tárolja, ezeknek nincs könnyen értelmezhető funkcionális formája. A statisztikai alapokra épülő ökonometriai modellekkel ellentétben nem kapunk automatikusan statisztikákat a modell paramétereire és változóinak előre jelzett értékeire. Éppen ezért az üzleti életben kevésbé elterjedt modellezési megközelítésnek tekinthetők, összehasonlítva a statisztikai modellek jelentős részével (*Cheng-Titterington* [1994]).

2. A vizsgált szolgáltatások és modellezési környezetük

A kutatás során 13 internetbankos szolgáltatást vizsgáltunk az egyes szolgáltatáscsoportokhoz tartozó kapacitást leíró adatok és a szolgáltatási szintek (válaszidők) közötti összefüggések elemzése céljából. A legnépszerűbb szolgáltatások kiemelt figyelmet kaptak, ezek a következők voltak:

– *Bankszámla-áttekintő*: az ügyfél itt tekintheti meg az internetbank szolgáltatásban szereplő számláinak adatait, a kiválasztott számlájának forgalmát, kivonatait és az utalások részleteit. Ezen az oldalon láthatja a banknál vezetett folyószámláit, illetve amennyiben rendelkezik, a betétlekötéseit, hitelkártyáját, illetve jelzáloghitelének állapotát.

– *Számlatörténet-lekérdezés*: a kiválasztott számla forgalmának megtekintésére ad lehetőséget.

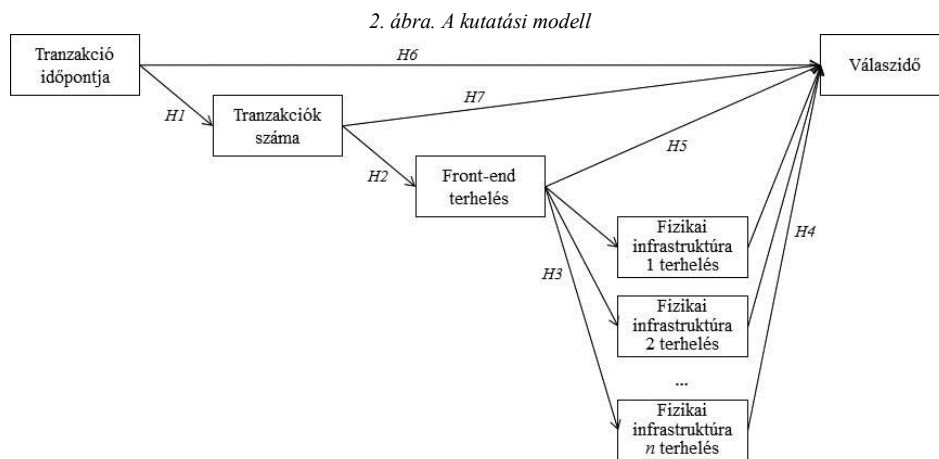
– *Pénzmozgáscsoportok*: az ügyfél láthatja a számláján történő pénzmozgások arányát kategóriák szerint, és a szolgáltatás segítségével elemezheti a számláján/kártyáján történt pénzmozgásokat.

– *Forintátutalás*: az ügyfél átutalási és átvezetési megbízásokat indíthat. Kötegel utalásokat is kezdeményezhet, illetve jóváhagyhatja az engedélyezésre váró tételeket. Lehetősége van megadni rendszeres átutalási megbízást.

– *Mobiltelefonegyenleg-feltöltés*: az ügyfél mobiltelefon-számhoz tartozó egyenleget tölthet fel.

3. Összefüggések statisztikai vizsgálata

A kutatási modell kialakításakor a szolgáltatások ügyfél által értelmezett minőségét a szolgáltatás válaszidejével jellemeztük. A válaszidőt a banki szolgáltatáson belül mértük (kérés beérkezésétől a válasz kiadásáig), azaz figyelmen kívül hagytuk a banki hálózaton kívüli válaszidőt, lévén erre a banknak nincs ráhatása. A szolgáltatás minőségére ható változók közül vizsgáltuk a szolgáltatás igénybevételének idejét, a tranzakciószámot és a különböző rétegek kapacitásfelhasználását.



A kutatás folyamán statisztikai módszerekkel vizsgáltuk az architektúra és a magyarázóváltozók közötti összefüggéseket. Kiemelten egy szolgáltatás esetén a 2. ábrán látható modellt alakítottuk ki, ahol a következő (a kutatás során tesztelt) hipotéziseket fogalmaztuk meg:

- H1: A tranzakciók számát meghatározza a tranzakciók ideje (nap-tári/heti időpontja)
- H2: A front-end réteg terhelését meghatározza a tranzakciók száma
- H3: A fizikai infrastruktúra terhelését meghatározza a front-end réteg terhelése

– H4–H7: A szolgáltatás válaszidejét meghatározzák a modell további tényezői (napszak, tranzakciósám, front-end terhelés, fizikai infrastruktúra-terhelés).

A hipotézisek tesztelése (lásd a táblázatot) alapján megállapítható, hogy azok jól határoznak meg nyilvánvaló összefüggéseket (például a tranzakciók idejének és számának összefüggését), ugyanakkor a válaszidőt tekintve már messze nem egyértelműek a kapcsolatok, és nem építhető megfelelő előrejelző modell.

A hipotézisek tesztelése

Megnevezés	Érték	Megnevezés	Érték
Hipotézis 1		Hipotézis 4	
Munkanap	Hétfő–péntek	Parciális korreláció (1. feldolgozási szint; válaszidő; kontrol: front-end)	–0,074 (0,000)
Szóráshomogenitás-, Levene-próba	72,882 (0,000)	Parciális korreláció (2. feldolgozási szint; válaszidő; kontrol: front-end)	0,063 (0,000)
ANOVA, <i>F</i> -próba	152,080 (0,000)	Parciális korreláció (3. feldolgozási szint; válaszidő; kontrol: front-end)	–0,079 (0,000)
Hipotézis 2		Hipotézis 5	
Korreláció (tranzakciósám; front-end)	0,498 (0,000)	Parciális korreláció (front-end; válaszidő, kontrol: tranzakciósám)	–0,063 (0,000)
Hipotézis 3		Hipotézis 6	
Parciális korreláció (front-end; 1. feldolgozási szint; kontrol: tranzakciósám)	0,240 (0,000)	Munkanap	Hétfő–péntek
Parciális korreláció (front-end; 2. feldolgozási szint; kontrol: tranzakciósám)	0,364 (0,000)	Szóráshomogenitás-, Levene-próba	162,84 (0,000)
Parciális korreláció (front-end; 3. feldolgozási szint; kontrol: tranzakciósám)	0,583 (0,000)	ANOVA-, <i>F</i> -próba	45,176 (0,000)
		Hipotézis 7	
		Korreláció (tranzakciósám; válaszidő)	0,432 (0,000)

Összességében az eredmények azt mutatták, hogy a vizsgált tényezők között csak nagyon gyenge összefüggés van, azaz a szolgáltatások végrehajtására kapott válaszidő tekintetében az alapstatistikai eszközcsoport nem ad megfelelő előrejelzést.

Egyes szolgáltatásokra, így a kiemelten kezelt forintátutalásra, a hagyományosan jól alkalmazható regressziós elemzések sem adtak kielégítő eredményt.

A tapasztalt hiányosságok miatt, a mesterséges intelligencia területéhez tartozó modellezési lehetőségeket megvizsgálva, a neurális hálózatok alkalmazására került sor a modellezés következő fázisában. A választást elsődlegesen a neurális hálózatok előző fejezetben ismertetett előnyös tulajdonságai indokolták, elsődlegesen a következők: képesek nemlineáris input-output kapcsolatok leírására, jó hibatűrő képességgel rendelkeznek, és jól támogatják az előrejelzési feladatok megoldását.

4. A vizsgált szolgáltatások neurális hálózatos modellezése

Neurális hálózatos modellek alkalmazása akkor célszerű, ha ismeretlen a magyarázóváltozók és a magyarázott változó közötti összefüggésrendszer, valamint a magyarázóváltozók között is feltáratlan összefüggések vannak. A kapacitásmodellezés problémaköre az ismertetett jellemzők miatt illeszkedik a neurális hálózatos elemzési területhez. A kutatás feltételezése szerint egyes szolgáltatások esetében, a szolgáltatási szintet jellemző válaszidőt több tényező is befolyásolhatja (*Fehér-Kő-Szabó* [2010], *Kő-Fehér-Szabó* [2014]):

- szezonális hatások figyelembevételével a hét bizonyos napja (hétfő, kedd, ...), mely kihat a tranzakciók számára;
- az adott szolgáltatást jellemző tranzakciószám, mely feltételezhetően hat az érintett rendszerek kapacitására;
- maguk az infrastruktúraterhelési-adatok, melyek befolyásolják a válaszidőt.

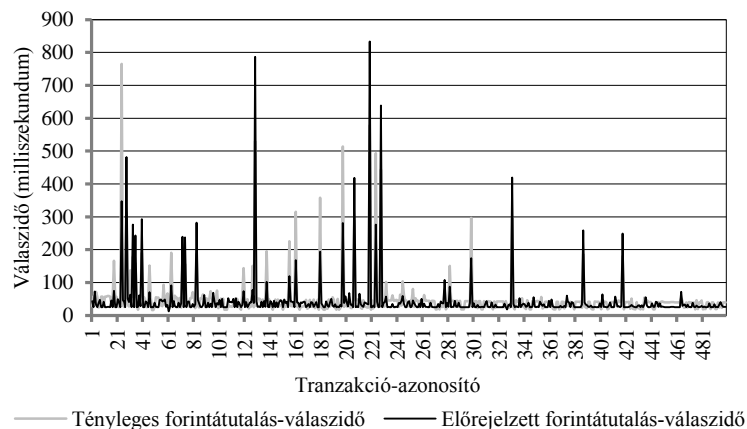
A neurális hálózatos modellek esetében az egyes internetbankos szolgáltatásokra a felsorolt adatokat használtuk fel a modellek kialakításához.

A neurális hálózatos modellek felépítése hasonló volt, ezért a modellek közül a forintátutalási szolgáltatás modelljét mutatjuk be, amely egy bemeneti, egy rejtett és egy kimeneti réteget tartalmazott. A rejtett rétegben a neuronok száma hat volt. A forintátutalási modellben használt bemeneti változók: 1. dátum:óra:perc; 2. dátum-weekday; 3. tranzakciószám; 4–7. a szolgáltatás nyújtásában meghatározó infrastrukturális elemek CPU-terhelése. A forintátutalási modellben használt kimeneti változó: forintátutalási szolgáltatás válaszidő.

A forintátutalási-modell tanítása során tanító és validáló mintaként 11 807 rekordot tartalmazó szolgáltatási adathalmazt használtunk fel, míg a modell tesztelését

13 622 rekordszámú mintán végeztük. A modell értékelésénél a következő hibafüggvényeket használtuk: korreláció, abszolút hiba, RMSE (root mean squared error – átlagos négyzetes hiba). A forintátutalási-szolgáltatás esetében egyes hibafüggvényértékek elmaradtak az elvárásoktól (például a korreláció értéke 0,21 volt), de a problémát a minta egy kis része okozta (a számlaátekintő szolgáltatásnál az adatok 8 százaléka). Amennyiben a tesztmintából kivesszük azokat a kiugró értékeket, amelyek a válaszidőre az általános tendenciától eltérő csúcsokat adják, a maradék 12 485 rekordra a hibafüggvények alapján elfogadható modelleket kaptunk (korreláció értéke 0,89-re módosult). A neurális hálózatos modellezési feladathoz a RapidMiner nyílt forráskódú adatbányászati szoftvert választottuk, elsődlegesen gazdag adatbányászati modellezési lehetőségei, végfelhasználói támogatása (intuitív grafikus felhasználói felülettel rendelkezik) és vizualizációs megoldásai miatt. A RapidMiner szoftvernek több éve vezető helye van az adatbányászati feladatokat támogató szoftverek között (Piatetsky-Shapiro [2016]).

3. ábra. Tényleges és előrejelzett válaszidő a forintátutalási szolgáltatásra

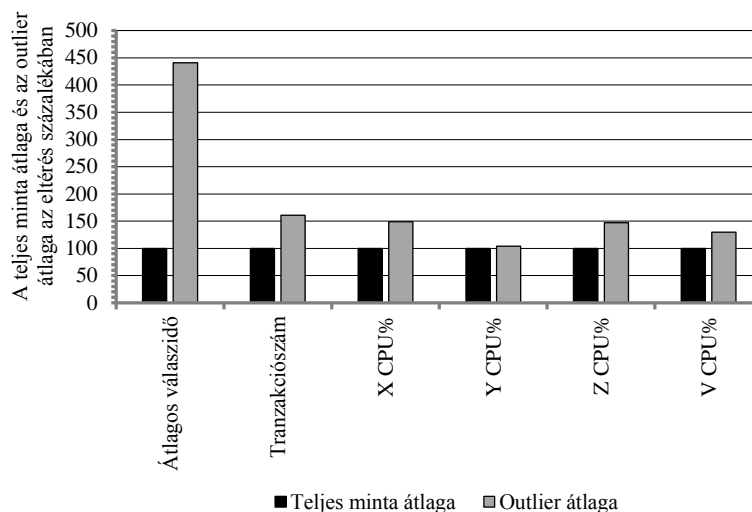


Forrás: Saját szerkesztés.

4.1. A kiugró értékek vizsgálata

A kiugró értékekre a neurális hálózatos modell nem adott megfelelő előrejelzést. Arányuk a teljes mintában változó volt, általában 2 és 4 százalék közé esett; kivéve a forintátutalásnál, ahol ez az érték 8 százalék volt. Egy lehetséges meghatározási módjuk, hogy az 5 százaléknál nagyobb relatív hibával rendelkező rekordokat soroljuk ebbe a csoportba. A forintátutalás területén 1 137 rekordot jelöltünk ki, mint kiugró érték. Az kiugró értékek vizsgálata során a cél az eltérések elemzése és lehetséges magyarázatának feltárása volt (Fehér-Kő-Szabó [2010]).

4. ábra. Forintátutalás-eltérések



A vizsgálat első lépéseken összehasonlítottuk a teljes minta és az kiugróérték-adatok jellemzőit. A vizsgálat megmutatta, hogy az kiugró értékkel kapcsolatos esetekre magasabb terhelési értékek jellemzők, melyek magasabb tranzakciószámmal és magasabb válaszidővel jellemezhetők. (Lásd a 4. ábrát.) A további elemzés arra is rámutatott, hogy a forintátutalás esetében meghatározó szerepet játszó négy infrastruktúrális elem közül három esetében az alacsony és közepes CPU% terhelés a jellemző, addig az ötödik infrastruktúra kiugró érték esetében kimondottan a magas CPU% terhelés (100 százalék közeli) a legjellemzőbb.

A vizsgálat során kiderült, hogy az infrastruktúrális elemeket tekintve egyedül az ötödik infrastruktúra-elemre volt jellemző, hogy szignifikánsan magasabb értékek kerültek előtérbe a kiugró értékekhez köthető eseteknél, mint a teljes mintát véve.

A kiugró értékek vizsgálata során megállapítható, hogy a problémás esetek oka az volt, hogy a meghatározott fizikai erőforrások leterheltsége elérte a maximumát vagy közel a maximumát. A fizikai erőforrások teljes kihasználásakor a feldolgozásra váró tranzakciók között sorban állás alakul ki, mely a válaszidőket kiszámíthatatlanná és előrejelezhetetlenné teszi. Az kiugró értékek esetében a válaszidők jelentős növekedése volt tapasztalható, szinte minden kiszolgáló érintett infrastruktúrális elem terhelésének növekedése mellett. Ugyanakkor a válaszidők ilyen mértékű növekedéséért elsősorban a csúcsterhelések lehetnek felelősek. A csúcsterheléses fizikai erőforrás a vizsgálat alapján beazonosíthatóvá vált. Az elemzés rámutat, hogy mindkét erőforrás esetében a csúcsterhelés gátja lehet a szolgáltatások megfelelő színvonalú biztosításának.

5. Összefoglalás

A kutatás számos értékes javaslattal zárult. A kialakított modellt a kutatást követő féléves időszakban változatlanul, az utána következő félévben pedig aktualizálva alkalmazták. A neurális hálózatokra alapuló modellek a rejtett, alapstatisztikával nem vizsgálható összefüggéseket is megmutatják a szolgáltatási szintek (válaszidők) és a magyarázóváltozók (hét napja, tranzakciószám, infrastruktúra-leterheltségek) között. A tesztelés alapján kijelenthető, hogy a modellek megfelelő becslést nyújtanak.

A kialakított modell több lehetséges felhasználási területét sikerült azonosítani:

- Válaszidők előrejelzése a paraméterek változtatásával (például tranzakciószám), trendelemzés, alkalmazásméretezés vagy egyéb becslés adataira alapozva.

- A modellezési megközelítés alkalmazása az alkalmazásméretezés esetében, mellyel előrejelezhető az éles környezetben megnyilvánuló hatás.

- A megvalósított fejlesztések összevetése az előrejelzésekkel: a modellek futtatása által elvárt válaszidők és a ténylegesen megvalósított fejlesztés által szolgáltatott válaszidők összevetése (az implementálást követő áttekintés részeként).

- Nem engedélyezett infrastruktúra változások felderítése: amennyiben a valós mért értékek és a modell által előrejelzett értékek között jelentős eltérés tapasztalható.

A jövőbeli kutatási terveink között szerepel további prediktív analitikai megoldások kifejlesztése a vizsgált problémakörre. A banki átutalásokra (átutalási időre) vonatkozó új szabályozási elvárások változásai kihívásokat jelentenek a kapacitásmenedzsment területén, további teret nyitva a bemutatott modellezési környezet alkalmazására és továbbfejlesztésére.

Irodalom

- AGGARWAL, C. C. [2015]: *Data Mining: The Textbook*. Springer. Heidelberg. <http://dx.doi.org/10.1007/978-3-319-14142-8>
- ALTRICHTER M. – HORVÁTH G. – PATAKI B. – STRAUZS GY. – TAKÁCS G. – VALYON J. [2006]: *Neurális hálózatok*. Panem Kiadó. Budapest.
- ANTHONY, M. – BARTLETT, P. L. [2009]: *Neural Network Learning: Theoretical Foundations*. Cambridge University Press. Cambridge. <http://dx.doi.org/10.1017/CBO9780511624216>
- BORGULYA I. [1998]: *Neurális hálók és fuzzy-rendszerek*. Dialóg Campus Kiadó. Pécs.

- CHENG, B. – TITTERINGTON, D. [1994]: Neural networks: A review from statistical perspective. *Statistical Science*. Vol. 9. No. 1. pp. 2–30. <http://dx.doi.org/10.1214/ss/1177010638>
- FAJSZI B. – CSER L. – FEHÉR T. [2010]: *Üzleti haszon az adatok az adatok mélyén – Az adatbányászat mindennapjai*. Aliena Kiadó. Budapest.
- FEHÉR P. – KŐ A. – SZABÓ Z. [2010]: *Szolgáltatási szint menedzsment és kapacitásmenedzsment összefüggéseinek kutatás-fejlesztési feladatai, kutatási jelentés*. Corvinno Kft. Budapest.
- GKIENET–T-HOME–T-MOBIL [2009]: *Jelentés az internetgazdaságról*. Budapest.
- HOPFIELD, J. J. [1982]: Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences*. Vol. 79. Issue 8. pp. 2554–2558. <http://dx.doi.org/10.1073/pnas.79.8.2554>
- HP (HEWLETT PACKARD) [2008]: *The Top 10 Misconceptions about Performance and Availability Monitoring*. http://www.ipsus-it.ie/images/pdfs/top%2010_white%20paper.pdf
- HP [2012]: *When Application Performance is Better, Business Works Better*. http://www.hp.com/hpinfo/newsroom/press_kits/2012/HPDiscover2012/HP_APM_9.2_Six_Steps_White_Paper.pdf
- ITSMF HUNGARY (IT SERVICE MANAGEMENT FORUM HUNGARY) [2009]: *Glossary of ITIL V3*. Budapest.
- KARAYIANNIS, N. – VENETSANOPOULO, A. N. [1993]: *Artificial Neural Networks: Learning Algorithms, Performance, and Applications*. Springer Science+Business Media. New York. <http://dx.doi.org/10.1007/978-1-4757-4547-4>
- KŐ A. [1992]: *Intelligens neurális hálózatok és alkalmazási lehetőségeik a közgazdaságban és a gazdálkodásban*. PhD-értekezés. Budapesti Közgazdaságtudományi Egyetem. Budapest.
- KŐ, A. – FEHÉR, P. – SZABÓ, Z. [2014]: *Challenges of Capacity Modelling in Complex IT Architectures*. ACM, SCITEPRESS Digital Library. Barcelona. <http://dx.doi.org/10.5220/0004851505430550>
- MCCULLOH, W. S. – PITTS, W. [1943]: A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*. Vol. 5. pp. 115–133.
- METZLER, A. [2003]: *The Mandate to Implement Unified Performance Management*. http://www.ashtonmetzler.com/Performance_Mngt.pdf
- OTP (ORSZÁGOS TAKARÉKPÉNZTÁR ÉS KERESKEDELMI BANK NYRT.) [2014]: *2 millió OTPdirekt ügyfél*. <https://www.otpbank.hu/static/portal/sw/flash/2milliodirekt/index.html>
- PIATETSKY-SHAPIO, G. [2016]: *KDnuggets 2016 Software Poll Results*. <http://www.kdnuggets.com/2016/06/r-python-top-analytics-data-mining-data-science-software.html>
- ROJAS, R. [1996]: *Neural Networks*. Springer-Verlag. Berlin. <http://dx.doi.org/10.1007/978-3-642-61068-4>
- SÁNTÁNÉ-TÓTH E. – KŐ A. – LOVRICS L. – BÍRÓ M. [2007]: *Döntéstámogató rendszerek*. Panem Kiadó. Budapest.
- SHARDA, R. – DELEN, D. – TURBAN, E. [2015]: *Business Intelligence and Analytics: Systems for Decision Support*. 10th Edition. Pearson Education Limited. London.
- TSAI, C.-F. – WU, J.-W. [2008]. Using neural network ensembles for bankruptcy prediction and credit scoring. *Expert Systems with Applications*. Vol. 34. Issue 4. pp. 2639–2649. <http://dx.doi.org/10.1016/j.eswa.2007.05.019>

- WERBOS, P. [1974]: *Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences*. PhD thesis. Harvard University. Cambridge.
- ZENG, Z. –WANG, J. [2010]: *Advances in Neural Network Research and Applications*. Springer. Berlin. <http://dx.doi.org/10.1007/978-3-642-12990-2>
- ZUPAN, J. – GASTEIGER, J. [1991]: Neural networks: A new method for solving chemical problems or just a passing phase? *Analytica Chimica Acta*. Vol. 248. Issue 1. pp. 1–30. [http://dx.doi.org/10.1016/S0003-2670\(00\)80865-X](http://dx.doi.org/10.1016/S0003-2670(00)80865-X)

Summary

Planning a complex IT infrastructure responsible for reliable internet banking services is a complex task. With increasing technological complexity, it is more and more difficult to monitor and relate component performances and capacities. The questions are: What is required to provide quality services? Which factors have impact on service quality? While some IT environments, like internal clouds and related technologies create flexibility for customers, the limited IT resources arise the problems of providing capacities that have impact on IT service quality. The paper presents these problems. The authors performed statistical analysis and predictive modelling to describe the relationship between IT service quality and background infrastructure capacity usage. Their models were tested in a pilot area of a widely used electronic banking service. As multivariate statistical modelling and hypothesis testing had limited results in phase 1, in phase 2 further predictive modelling opportunities were explored, resulting in the development and use of a neural-network-based model.