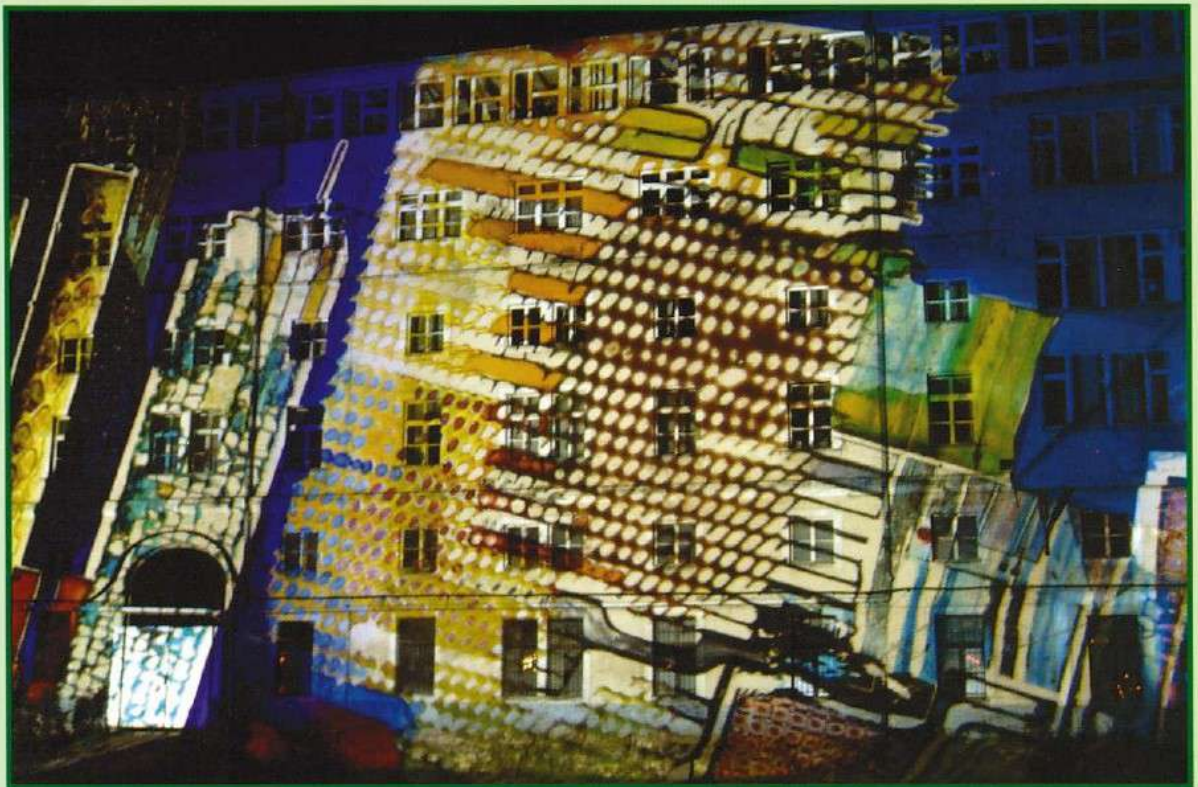


híradástechnika

1945 VOLUME LX. 2005

hírközlés ■ informatika



Multimédia szolgáltatások vezeték nélküli hálózatokon

Mobil szolgáltatások – a negyedik G felé

MPEG multimédia-folyamok továbbítása

Személyi hálózatok

2005/11

**A Hírközlési és Informatikai Tudományos Egyesület folyóirata
a Nemzeti Hírközlési és Informatikai Tanács együttműködésével**

Tartalom

<i>ELŐRETEKINTÉS</i>	1
Gál Zoltán, Karsai Andrea, Orosz Péter IPv6 kapcsolatok elemzése mobil WiFi környezetben	2
Péterfalvi Gábor, Pozsár Balázs, Simon Vilmos, Huszák Árpád, Imre Sándor A következő generációs mobil hálózatok fejlődési tendenciái	10
Huszák Árpád, Imre Sándor Videó folyam szelektív újraküldése IP hálózatokban	17
Erdősi Ferenc A mobiltelefon elterjedésének regionális jellemzői Európában	25
Rónai Miklós Aurél, Fodor Kristóf, Biczók Gergely, Turányi Zoltán, Valkó András MAIPAN – middleware személyi hálózatokban működő alkalmazások összekapcsolására	33
Éltető Tamás, Vaderna Péter Időben változó sorbanállási rendszer vizsgálata mátrix-geometrikus eljárások segítségével	38
Benkő Borbála Katalin, Pándi Zsolt Dokumentumkategorizálás távközlési alkalmazásban	45
<i>Felhívás cikkek írására a Híradástechnika cílszámaiba</i>	50

Védnökök

SALLAI GYULA a HTE elnöke és DETREKŐI ÁKOS az NHIT elnöke

Főszerkesztő

SZABÓ CSABA ATTILA

Szerkesztőbizottság

Elnök: ZOMBORY LÁSZLÓ

BARTOLITS ISTVÁN
BÁRSONY ISTVÁN
BUTTYÁN LEVENTE
GYŐRI ERZSÉBET

IMRE SÁNDOR
KÁNTOR CSABA
LOIS LÁSZLÓ
NÉMETH GÉZA
PAKSY GÉZA

PRAZSÁK GERGŐ
TÉTÉNYI ISTVÁN
VESZELY GYULA
VONDERVISZT LAJOS

Előrettekintés

szabo@hit.bme.hu

Bár novemberi számunkra továbbra is a sokszínűség jellemző és a cikkek több részterületet képviselnek, már most szeretném jelezni, hogy jövő év januártól kezdve rendszeresen tematikus lapszámokkal fogunk jelentkezni. Ezek mindegyikében egy-egy téma területet mutatunk majd be a felkért szerzők áttekinthető jellegű, tehát szándékunk szerint T. Olvasóink széles köre számára érthető és érdekes cikkekkkel. A szám gerincét képező tutoriális rész mellett helyet adunk a témakörhöz kapcsolódó, hozzánk beküldött kutatási jellegű cikkeknek is. Ezt a lehetőséget ajánljuk potenciális szerzőink figyelmébe, és ehhez nyújtunk segítséget felhívásainkkal, melyek közül elsőként – egyszerre két-tővel is – e számunkban jelentkezőnk, a későbbiekben pedig rendszeresen, szinte minden hónapban találkozhatnak.

2006 első felében a következő tematikus számokat tervezzük, a szerkesztőbizottság egy-egy, az adott téma területet „gondozó” tagjának, mint vendégszerkesztőnek a közreműködésével:

Január: Web-technológiák (Vonderviszt Lajos)
Február: Optikai kommunikáció (Paksy Géza)
Március: Beszédfeldolgozás (Németh Géza)
Április: Úrkutatás, úrtávközlés (Kántor Csaba)
Május: Informatikai biztonság (Buttyán Levente)
Június: Válogatás a „Networkshop” legjobb előadásaiból (Tétényi István)

Várjuk az érdekes eredményeket tartalmazó cikkeket és számítunk olvasóink érdeklődésére!

Mostani számunkban egy cikksokorral találkozunk az Olvasó, amelynek közös nevezője: a vezeték nélküli hálózatokon nyújtandó multimédia szolgáltatások. Az első cikk motivációja az új generációs Internet protokollja, az IPv6 feletti mobil szolgáltatások megjelenése és elterjedése. Izgalmas felhasználói és szakmai kérdéskörnek fogalmazódik meg a mobilitás hatásának mértéke a TCPv6, UDPv6 protokollokra épülő szolgáltatások viszonylatában. A szerzők valós, vezeték nélküli hálózati környezetben végzett összehasonlító mérések eredményeivel adnak segítséget a szolgáltatás-minőség biztosításához.

A jövőben egyre inkább a multimédia továbbítása fogja adni a hálózati terhelés nagy részét. Sokak szerint a 3G hálózatok nem lesznek alkalmasak igazán szélessávú multimédia szolgáltatások nyújtására, ezért a kutatás és fejlesztés fókuszában a 4G-nek nevezett

megoldási irányok állnak. Ezekről a fejlődési tendenciákról ad következő cikkünk áttekintést, a 3G hálózatok lehetőségeinek határáról, az új igényekről, melyek szükségessé teszik a jövőben a 4G hálózatok kialakítását, és hogy jelenleg milyen technológiai lehetőségek állnak rendelkezésünkre az ezen elvárásokat kielégítő 4G hálózatok megvalósításához.

Az Interneten elérhető multimédiás alkalmazások egyre népszerűbbé válnak a felhasználók körében. Mobil környezetben történő robbanásszerű elterjedésüket azonban a korlátozott sávszélesség és a vezeték nélküli csatorna jellegzetesen magas hibavalószínűsége korlátozza, hiszen a gyakori hibák jelentősen rontják a multimédiás szolgáltatás minőségét. Cikkünk bemutat egy új algoritmust, amely a DCCP transzport protokoll alkalmazásával jelentősen javítja az audió-videó folyam minőségét. A javasolt szelektív újraküldéses algoritmus kiemelten kezeli az MPEG típusú multimédiás folyamat fontosabb kereteit, így csökkentve a bithibák továbbterjedését. A szolgáltatás minősége így nagymértékben javítható.

A vezeték nélküli terminálok, mint például okostelefonok, digitális személyi asszisztensek (PDA-k) és laptopok növekvő száma miatt egyre inkább szükség van arra, hogy a személyi hálózatokat egyszerű módon lehessen felállítani, valamint be- és újrakonfigurálni. Cikkünk szerzői arra adnak választ, hogy hogyan lehet ilyen hálózatokban egy dinamikus változtatható, de egységes számítástechnikai környezetet létrehozni, mely kiterjed az egymással összekötött vezeték és vezeték nélküli, valamint fix és mobil eszközökre.

A témakört kiegészíti egy kommunikáció-geográfiai cikk, amely bemutatja, hogyan jutottunk el a harmadik generációs rendszerekig. A cikk a mobiltelefonia európai elterjedésének időbeli folyamatával és térbeli mintáival foglalkozik a mindenkor technológiák (analog, digitális) időszakaira vonatkozóan. A jellegzetességeket komplex szemléletmóddal, az egyes hatótényezők (a GDP, a természetföldrajzi adottságok, a népsűrűség, a lakosság kulturális színvonala és a távközlési politika) együttes vizsgálatával mutatja be.

E számunk még további két érdekes cikket tartalmaz: az egyik egy távközlési feladatok vizsgálatára is alkalmas új sorbanállási modellel, a másik a távközlési rendszerek üzemeltetéséhez elengedhetlen hibajegy-rendszerekkel foglalkozik.

Szabó Csaba Attila,
főszerkesztő

IPv6 kapcsolatok elemzése mobil WiFi környezetben

GÁL ZOLTÁN, KARSAI ANDREA, OROSZ PÉTER

Debreceni Egyetem Informatikai Szolgáltató Központ (DISZK)
zgal@cis.unideb.hu, kandrea@fox.unideb.hu, orosp@delphin.unideb.hu

Kulcsszavak: Internet2, WiFi, L2 és L3 roaming, TCP Slow Start algoritmus, TCP Windowing algoritmus

A cikkben egy valós kültéri mobil WiFi rendszer roaming fázisa alatt bekövetkező folyamatok IPv4 és IPv6 kapcsolatokra kifejtett hatását vizsgáljuk meg. A mobil kliens bázisállomásokhoz viszonyított relatív sebessége és a roaming végrehajtásának kölcsönhatása jelentősen befolyásolja a TCP kapcsolatokat, miközben kevésbé hat az UDP átvitelre. Az összehasonlító mérésekből statisztikai módszerekkel nyert eredmények lehetővé teszik, hogy valós képet nyerjünk az IPv4 és az IPv6 mobil átvitel esetén tanúsított viselkedésére vonatkozóan, valamint választ kaphatunk arra a kérdésre, hogy valóban magasabb minőségű mobil adatátvitelt eredményez-e az IPv6 protokoll vezetékek nélküli adatkapcsolati réteg fölött, elődjéhez, az IPv4-hez képest.

1. Bevezetés

Az Internet2 legjelentősebb előnyeként említhető az IPv6 feletti szofisztikált mobil szolgáltatások megjelenése és elterjedése. Izgalmas felhasználói és szakmai kérdéskörnek fogalmazódik meg a mobilitás hatásának mértéke a TCPv6, UDPv6 protollokka épülő szolgáltatások viszonylatában. Ahhoz, hogy ebben a témában a minőségi válasz mennyiségi jellemzőkkel is érzékelhető legyen, szükséges olyan összehasonlító mérések elvégzése, amelyek előtérbe helyezik az IPv4 és IPv6 technológiák közötti különbségeket mobil környezetben.

A szolgáltatások működőképességének biztosítása a csomópont fizikai mozgása közben olyan igény, mely a mai, korszerű hálózatok igény-palettáján joggal jelenik meg. A vezetékek nélküli IP telefonía, a vezetékek nélküli laptop és a PDA számítógépek fejlődése erőteljesen ebbe az irányba mutat. A vezetékek nélküli LAN-ok mobilitás tulajdonsága újabb hasznos eredményekhez vezet:

- *innovatív alkalmazás fejlesztés:* vészjelzések, üzenetek küldése, folyamatos hálózati kapcsolatban álló munkafolyam rendszerek megjelenése;
- *hatékonyság és termelékenység növekedés:* a folyamatos hálózati kapcsolódás lehetővé teszi a munka bárholnani elvégzését időkiesés nélkül;

– *adatok megnövelt hitelessége:* az adatok bármikor és bárholn elérhetők;

– *rendelkezésre állás:* a felhasználó virtuálisan online kapcsolatban maradhat az otthonában, az utcán és a munkahelyén is.

Jelen anyagban a mobil WiFi hálózaton kommunikáló IP terminálok különböző alkalmazásainak mennyiségi összehasonlítását mutatjuk be, magyarázatot adunk a tapasztalt jelenségekre és következtetéseket vonunk le a várható fejlesztési igényekre vonatkozóan.

2. Mobil adatátvitel

Mint ismeretes, az IP protokoll v4 és v6 verziója a hagyományos rögzített, huzalos hálózati kommunikáción túlmenően képes ellátni mobil funkciókat is. A vezetékek nélküli adatátviteli kapcsolatok a hálózati réteg számára ugyancsak képesek biztosítani a keretek továbbítását. Ezáltal az IP protokoll verziójától függően a szállítási rétegben működő szolgálatok részére kisebb vagy nagyobb mértékben történik az alsóbb rétegek viselkedésének figyelembe vétele.

Az IP protokoll mobil funkciója arra vonatkozik, hogy a terminál kommunikáció közben elmozdul fizikai helyéről, minek következtében megváltozik a hármas rétegbeli hálózati környezete. Ehhez a meglátogatott új hely-

1. táblázat Az IPv4 és IPv6 összehasonlítása

Megkülönböztető tulajdonságok	Mobil IPv4	Mobil IPv6
Speciális útválasztó funkció (foreign agent)	Igen	Nem
Útvonal optimalizálás készség	Protokoll része	Kiterjesztés
Szimmetrikus elérhetőség az MT és útválasztója között az aktuális helyen	Nem	Igen
Útválasztási forgalom (overhead) sávszélesség igénye	Sok	Kevés
Lekapcsolódás az L2 rétegről	Nem	Igen
“Tunnel soft” állapot kezelés igénye	Igen	Nem
“Dynamic home agent” cím felfedezés	Nem	Igen

		Protokoll	
		IPv4 / IPv6	Mobil IPv4 / Mobil IPv6
Hálózat	Vezetékes	√	√
	Vezeték nélküli	√	√

2. táblázat Cikkünk témája

színen egy „foreign agent” funkció betöltését biztosító útválasztó egy IP feletti IP alagút segítségével továbbra is kapcsolatban áll a „home agent”, eredeti útválasztóval [1]. Ezáltal a mozgó terminál az új helyszínen továbbra is képes kommunikálni. Fontos kérdés, hogy az „agent” processzek közötti interakció mennyire gyors, illetve milyen további terhelést okoz a hálózaton. Az IP protokoll v4 és v6 verziója ebből az aspektusból is lényegesen eltérő viselkedést mutat. Ezen tulajdonságokat az 1. táblázatban soroljuk fel.

A vezeték nélküli adatkapcsolatok esetén lehetőség van arra is, hogy a mobil terminál ugyanabban az üzenetszórás tartományban maradjon, vagyis a terminál IP címének változása nélkül módosuljon a forgalmazott keretek útvonala. Ez tipikusan az L2 roaming esete, amikor a mobil terminál úgy vált bázisállomást, hogy csak az adatkapcsolati eszközök CAM táblájának tartalma módosul. A cikk valós kültéri mobil WiFi rendszer környezetben egy mozgó járműben elhelyezett terminál által generált rádiós cellaváltás alatt bekövetkező roaming folyamat IPv4, illetve IPv6 kapcsolatokra kifejtett hatását vizsgálja meg (2. táblázat).

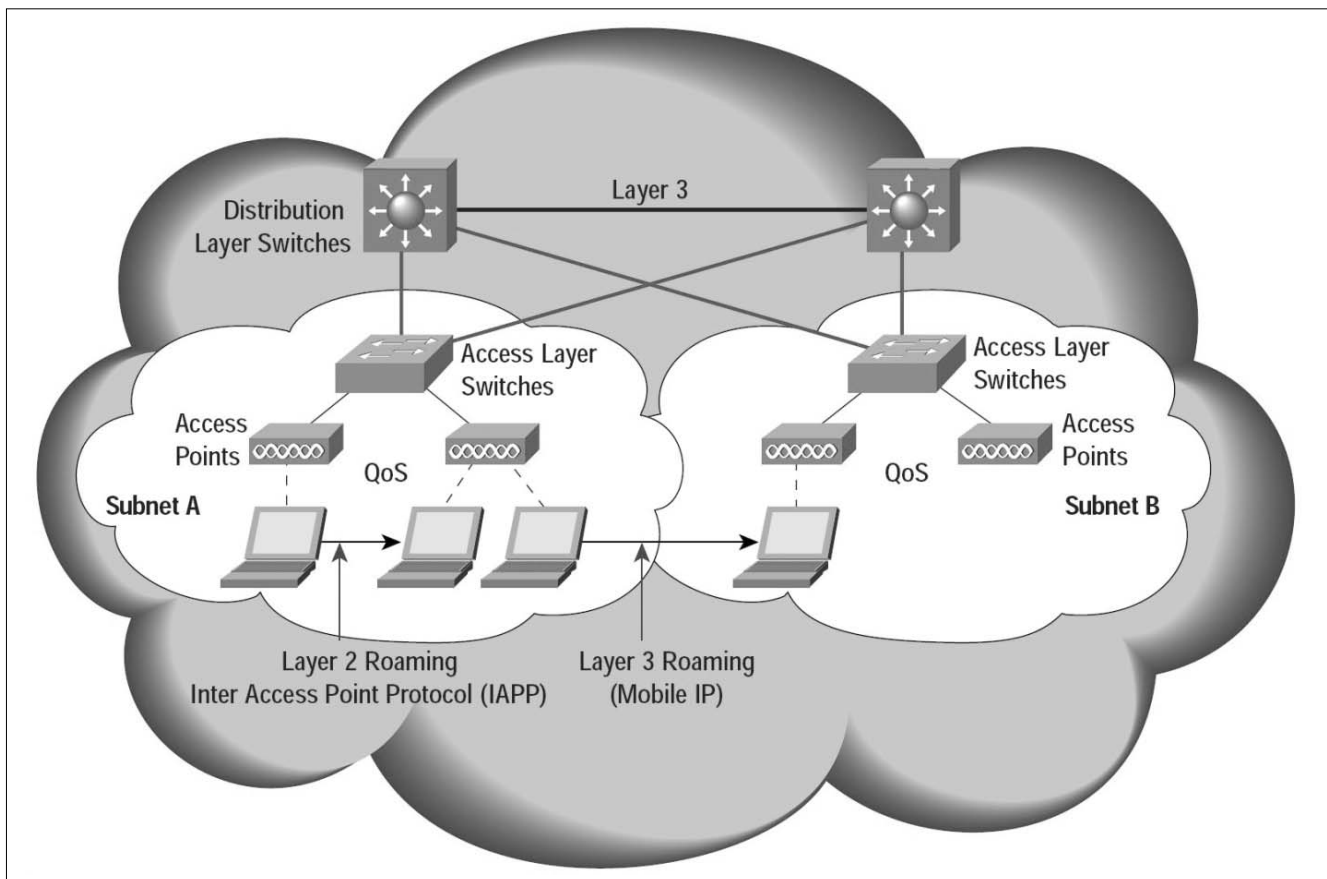
3. Roaming mechanizmusok

A vezeték nélküli LAN-ok lehetővé teszik, hogy a csomópontok a vállalati hálózathoz virtuálisan kapcsolódjanak. A cellaváltás (roaming) olyan időben lejátszódó folyamat, amely során a mobil terminál egyik kiszolgáló AP-bázisállomástól egy másik AP-ra csatlakozik rá. Adatkapcsolati (L2) roamingról beszélünk, ha a folyamat azonos alhálózatba tartozó AP-k között történik (1. ábra).

Ha a terminál másik alhálózatba tartozó új AP-hoz csatlakozik, akkor hálózati (L3) roamingról beszélünk. Hálózati cellaváltás az adatkapcsolati roaming sikeres lezajlása után következhet be [2]. A cellaváltás mindig a terminál döntésén alapul, amelynek feladata a lehetséges bázisállomások felderítése, az ezekhez tartozó paraméterek értékelése, majd a szóbaeső cellák közül az új kiválasztásának eldöntése. Az adatkapcsolati cellaváltás az alábbi fázisokat foglalja magába:

1) A terminál az „A” cellából elmozdul a „B” cellába. A bázisállomások ugyanabban az alhálózatban vannak, így L2 roamingról beszélünk. Ahogy a terminál ki-

1. ábra Az L2 és az L3 roaming



lép az „A” cellából, az APA bázisállomással fennálló kapcsolat paramétereit közül valamelyik átlépi a megadott küszöbértéket, s ez kiváltja a roaming folyamat indítását.

2) A kliens végigselektálja az összes IEEE 802.11-es csatornát, lehetséges bázisállomást keresve. Megtalálja az AP_B-t, lezajlik a fizikai rádiós csatornán a hitelesítés és az asszociáció folyamata.

3) Az AP_B a kliens alhálózatába egy nulla tartalmú multicast üzenetet küld, amelynek forrás fizikai címe éppen a mobil terminál címével egyezik meg. Ez alapján a huzalos LAN hálózatban található switch-ek frissítik kapcsolási táblájukat. Így a terminálnak címzett Ethernet keretek ezután nem az APA, hanem az AP_B bázisállomáshoz kerülnek.

4) Az AP_B a saját forrás MAC címével küld egy multicast üzenetet, amelyben értesíti az alhálózat összes bázisállomását arról, hogy az adott MAC című terminál hozzá asszociált. Ahogy az APA ezt megkapja, törli a mobil terminál MAC címét az asszociációs táblájából.

A roaming folyamatot mindig a kliens kezdeményezi, de a folyamatra vonatkozóan még nem létezik IEEE szabvány [7]. A Cisco gyártmányú terminálok esetében az alábbi események váltják ki a roaming folyamat indítását:

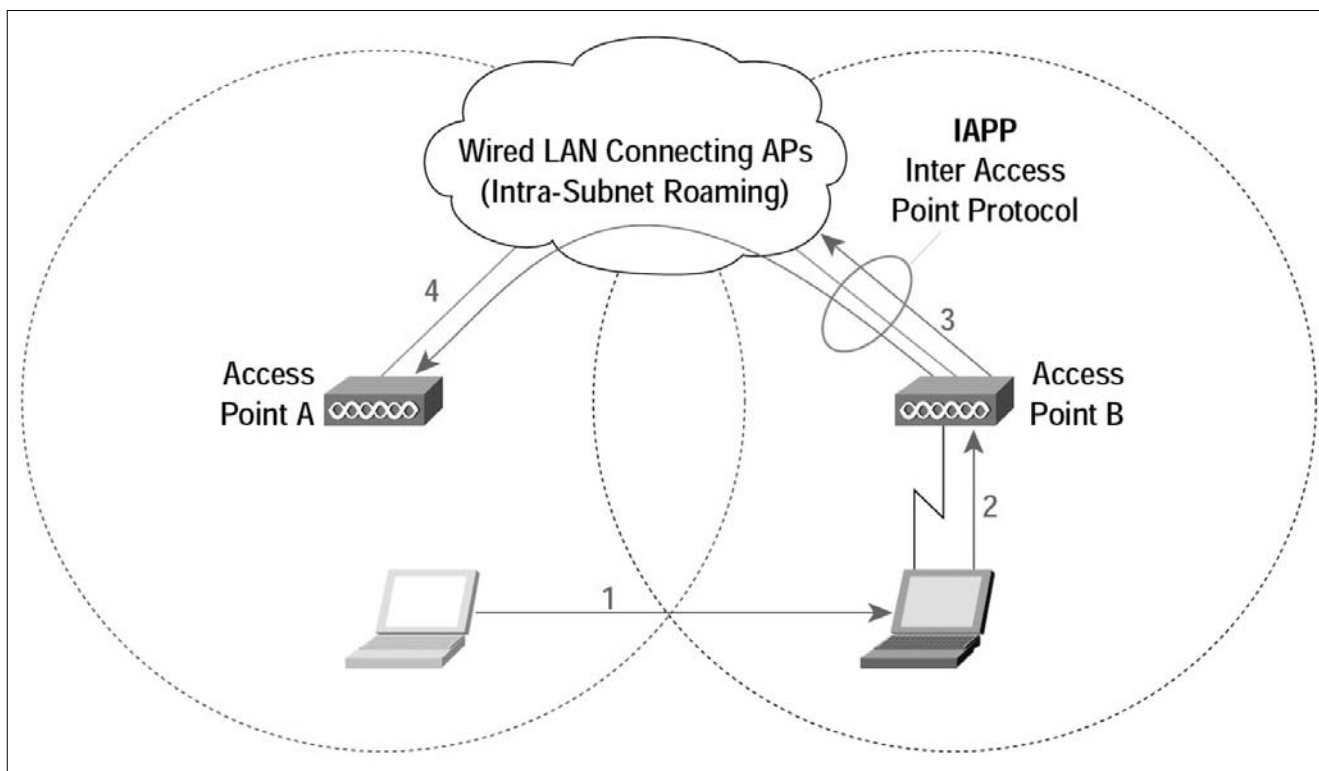
a) *Maximális csomagküldés próbálkozási szám átlépése.* Ha a kliens a maximum data retry-ként megadott számú próbálkozás után sem tudja a csomagot elküldeni, elindítja a roaming folyamatot. A Cisco Aironet kliensben ez az érték alapértelmezés szerint 16, és az Aironet Client Utilityben állítható.

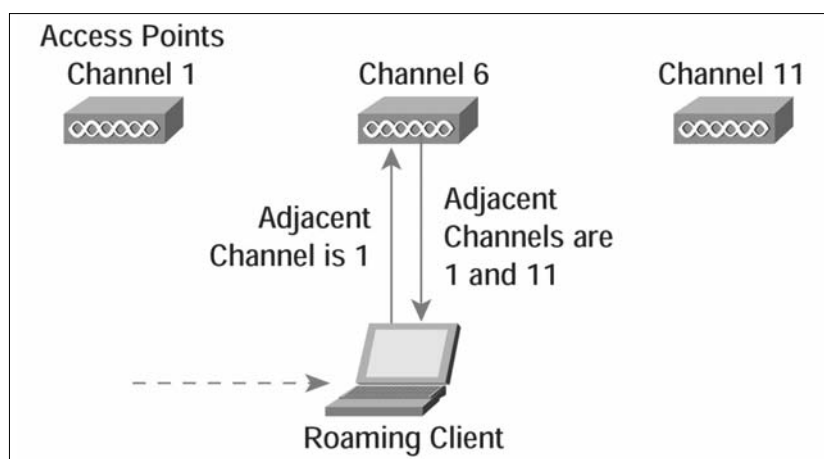
b) *Túl sok „bacon” kihagyása.* Minden, bázisállomáshoz társított kliensgép periodikusan kap „bacon” keretet. Alapértelmezésben 100 milliszekundumonként küld „bacon”-t a bázisállomás. Ez a periódus egyben konfigurációs paraméter is. A terminál a „bacon”-ben található érték alapján megtanulja annak periódusát. Amennyiben a terminál nyolc periódus ideig nem kap „bacon”-t, kezdetét veszi a roaming folyamat. A beérkező „bacon”-ök folyamatos figyelésével – még egy „idle” állapotban levő kliens is – képes érzékelni a vezeték nélküli kapcsolat minőségének romlását, majd pedig roaming-ot kezdeményezni.

c) *Átviteli ráta váltása.* Normál esetben a rádiós keretek átvitele a bázisállomás alapértelmezett adatátviteli sebességével történik. Ez a ráta a legmagasabb átviteli sebesség, amelyet „required” vagy „enable” paraméterként lehet az AP-n beállítani. Minden olyan alkalommal, amikor egy csomagot alacsonyabb sebességgel kell újraküldeni, a „retransmit” számláló hárommal növekszik. Minden olyan csomag esetében, amikor az alapértelmezett átviteli sebességgel sikerült a továbbítás, ez a számláló eggyel csökken egészen addig, míg a nulla értéket el nem éri. Amennyiben a számláló eléri a 12-es felső határt, az alábbi események valamelyike következik be:

- Ha a kliens nem hajtott végre cellaváltást az elmúlt 30 másodpercben, akkor bekövetkezik a gyors cellaváltás (fast roaming).
- Ha az említett időn belül roaming-ot hajtott végre, akkor eggyel alacsonyabb fokozatra csökkenti az átviteli rátát. Az alapértelmezett átvitelnél alacsonyabb rátájú sikeres átvitel esetén, egy rövid idő elteltével ismét visszaugrik az eggyel magasabb sebességű üzemmódba.

2. ábra Az L2 roaming lépései





3. ábra A fast roaming csatorna keresése

d) *Periódikus kliens intervallum (opcionális)*. A Cisco Aironet v6.1-től kezdve konfigurálni lehet, hogy a mobil terminál milyen gyakorisággal, illetve milyen jelerősség mellett keressen jobb vételi minőségű bázisállomást. Ezekkel a beállításokkal a terminál egy jobb térerejű bázisállomást fog keresni feltéve, hogy az alábbi feltételek mindegyike teljesül:

- A terminál már legalább 20 másodperce asszociált az aktuális AP-hoz. Ez a feltétel megakadályozza, hogy a kliens túl gyorsan kapcsoljon a bázisállomások között. Érvényes értékek 5-255 másodperc.
- A térerősség 50%-nál gyengébb. Érvényes intervallum: 0-75%-ig.

e) *Kliens inicializáció*. A terminál bekapcsolásakor és újraindításakor lezajló folyamat.

A roaming folyamathoz új bázisállomás keresése szükséges [3]. Ennek érdekében a terminál a rádiós csatornákon scan technika segítségével meghatározza az elérhető bázisállomások listáját, amelyből a legjobbat választja ki. A scan technika csatornánként egy „probe” teszt üzenet küldését jelenti, amire „probe” válasz vagy „bacon” érkezik a csatornán üzemelő bázisállomástól. Az AP-től érkező „bacon”-öket csak akkor veszi figyelembe a kliens, ha az SSID és a titkosítási beállítások megegyeznek.

A keresés befejezése után a listából kiválaszt egy bázisállomást, hogy az elérési paramétereit összehasonlítsa a lista többi tagjával. Ha a terminál kezdeti „start-up” fázisban van, akkor az új AP a listában elsőként szereplő tag lesz; ha a terminál roaming fázisban van, akkor az új AP a korábbi marad amennyiben választ a teszt „probe” keretekre. Válasz hiánya esetén a lista első tagja lesz az új AP. Az aktuális AP a lista többi elemével összehasonlításra kerül. Ahhoz, hogy egy tag új AP lehessen, minden listabeli AP-nak az alábbi szempontokat kell teljesítenie:

- A potenciális cél AP jelerőssége legalább 20%. Ha a térerő több mint 20%-kal gyengébb, mint az aktuális AP térereje, akkor legalább 50% jelerősséggel kell rendelkezzen.

– Ha a potenciális cél AP repeater módban van, és több rádió hop-ra van a gerinchálózattól, mint az aktuális AP, akkor 20%-kal nagyobb jelerőssége kell, hogy legyen, mint a jelenlegi AP-nak.

– A potenciális cél AP-nál a küldő egység terheltsége maximum 10%-kal lehet nagyobb, mint a jelenlegi AP esetén. A terminál a felsorolt alapkritériumoknak megfelelő bázisállomásokat összehasonlítja a jelenlegi bázisállomással. Ha egy elfogadott AP teljesít egyet az alábbi feltételek közül, akkor azt a terminál új, aktuális AP-nak választja, majd a lista többi AP-ját már ehhez az újonnan

választott AP-hoz hasonlítja a továbbiakban: a jelerősség 20%-kal nagyobb, mint az aktuális bázisállomásé; kevesebb hop távolság a gerinchez; legalább négyel kevesebb a kapcsolódott kliensek száma, mint a jelenlegi AP esetén; legalább 20%-kal kisebb a küldő egység terheltsége.

A 12.2.(11)JA IOS verziótól kezdődően a Cisco „fast secure roaming” implementáció két újabb lehetőséggel bővült: egyrészt növelt hatékonyságú a 802.11-es csatornakeresés a fizikai roaming alatt, másrészt hatékonyabb újra hitelesítési mechanizmus jelenik meg, amely fejlett titkosító kulcs menedzsmentet alkalmaz. Függetlenül az alkalmazott biztonsági módszertől, a hatékonyabb csatornakeresés gyorsabb L2 roaming-ot tesz lehetővé.

Az újrakapcsolás hatékonyságát növelő kulcs menedzsment felgyorsítja a Cisco LEAP hitelesítési folyamatot, így a roaming rövid idő alatt és biztonságosan zajlik le. A Cisco terminálokon és bázisállomásokon az IEEE 802.11 csatornakeresés alapértelmezés szerint egyaránt engedélyezett.

A „fast secure roaming”-ot egy csatornakeresés előzi meg. A 12.2(11)JA előtti IOS verziók esetén a kliensnek 37 ms-ot vett igénybe egy rádiócsatorna ellenőrzése, ami a magyar szabványok szerinti 13 csatorna esetén összességében 481 ms-ot jelent. A kliens minden egyes csatorna esetén az alábbi lépéseket hajtja végre: miután a terminál rádiós hardvere ráhangolódik az adott WLAN csatornára, figyel hogy elkerülje az ütközést, majd „probe” keretet küld és várja a „probe response” vagy a „bacon” jelzést. A fast secure roaming esetén hatékonyabb a csatornakeresés: az újrakapcsolás informálja az új AP-t a korábbi AP-val való kapcsolat elvesztése óta eltelt időről, a csatornaszámról, és az SSID-ről.

Ezeket az információkat felhasználva, az új AP felépít egy listát a szomszédos bázisállomásokról, és az általuk használt rádiócsatornákról. Ha a szomszédos AP-kről információt szolgáltató mobil terminál több, mint 10 másodperce kapcsolódott le az előző AP-ról, akkor az általa küldött információkat nem veszi figyelembe az új AP. A bázisállomások maximum 30 szomszédos AP-ról tárolnak információt. Ez a lista egy egynapos periódus alatt elévül. Amikor a terminál asszociál egy AP-hoz, az

új bázisállomás unicast csomagban visszaküldi számára a szomszédos AP-k listáját.

Ha a kliensnek roaming-ot kell végrehajtania, megvizsgálja az aktuális AP-tól kapott listát, és csak azokat a rádiócsatornákat ellenőrzi, melyeket a szomszédos bázisállomások valamelyike használ. A kliensállomás az elfoglaltságától függően az alábbi három roaming típus egyikét alkalmazza:

- *Normal roam*: a kliens nem kapott és nem küldött unicast csomagot az elmúlt 500 ms-ban. Nem használja az AP-tól kapott listát, ellenőrzi az adott térségben érvényes összes 802.11-es csatornát.
- *Fast roam*: a kliens kapott vagy küldött unicast csomagot az elmúlt 500 ms-ban. A szomszédos AP-k által használt csatornákat ellenőrzi. Ha nem talál új AP-t a lista alapján, átvizsgálja az összes csatornát. A kliens 75 ms-ra korlátozza a keresési idejét, ha legalább egy jobb AP-t tudott találni.
- *Very fast roam*: a kliens kapott vagy küldött unicast csomagot az elmúlt 500 ms-ban, és nullánál nagyobb százalékkal növeli az adott cella terheltségét. A többi esemény a „fast roaming”-gal megegyező kivéve, hogy jobb bázisállomás találata esetén a keresés azonnal befejeződik.

A méréseinkhez olyan eszközparkot használtunk, amely a fenti három roaming típus bármelyikét végre tudja hajtani.

4. Mérési környezet

Az IPv4 és IPv6 protokollok viselkedését mobil környezetben úgy vizsgáltuk, hogy egy kültéri teszt WiFi rendszert állítottunk össze. Ez IEEE 802.11b szabvány szerint működő két darab egymástól 100 méteres távolsá-

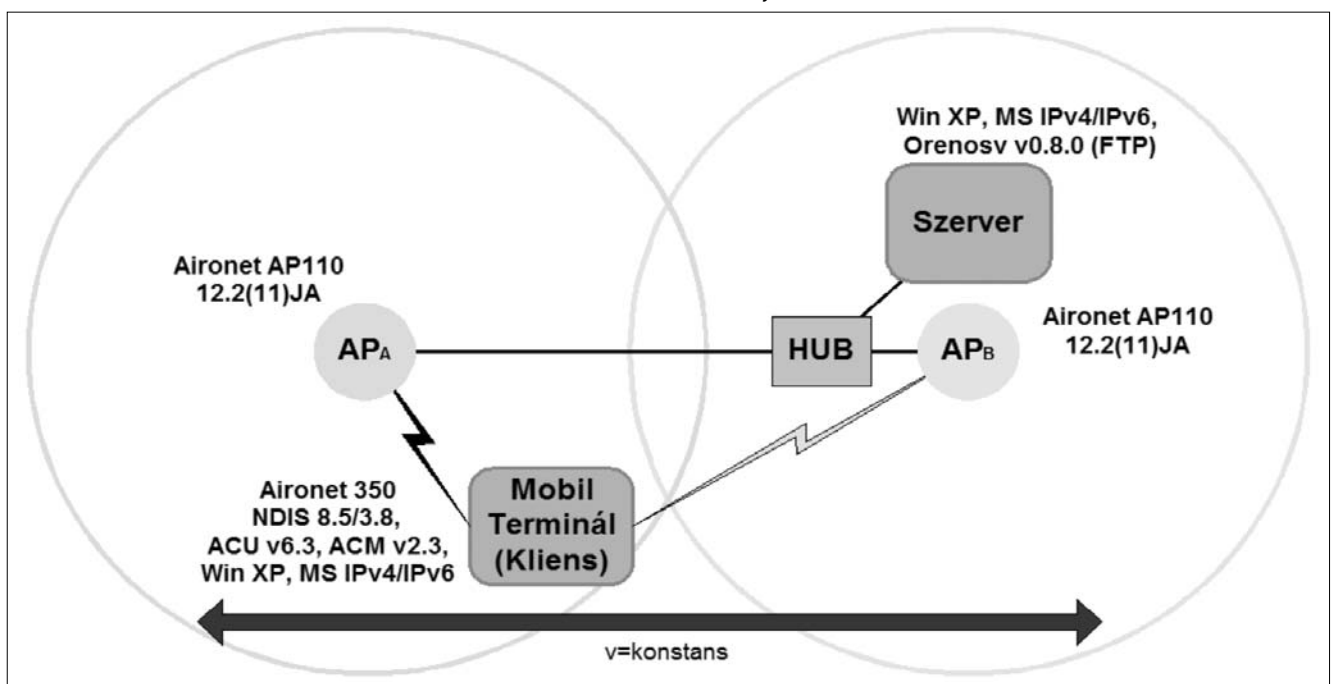
gon huzalos Ethernet kapcsolattal összekötött bázisállomásokból és egy mobil terminálból (kliensből) állt. Mint ismeretes a 11 Mbit/s-os WiFi szabvány is támogatja a roaming funkciót. Ugyanakkor a vezeték nélküli adatátvitel sebessége erőteljesen függ a bázisállomás és a kliens közötti távolságtól. A mobil WiFi terminál mozgás közben közeledik, majd távolodik a bázisállomástól.

Ez az adatkapcsolati rétegben az átviteli sebesség automatikus váltását okozza a 0:1:2:5,5:11 Mbit/s-os értékek között. Alapértelmezés szerint roaming esetén 11:5,5:2:1:0:1:2:5,5:11 Mbit/s-os sebességértékek mellett történik az átvitel. A mi esetünkben az átviteli sebességet 11 Mbit/s-ra rögzítettük, ezáltal a térerő változása kényszerítette a terminált a roaming kezdeményezésére.

A szállítási réteg különböző protokolljainak viselkedését figyeltük, miközben a mobil terminál egy autóban – a roaming idején konstans sebességgel – mozgott a bázisállomások közötti iránnyal párhuzamos útvonalon. A szerver oldalon az Ethereal snoop programot futtatuk, amely az adatkapcsolati réteg minden egyes kérését időbélyeggel letárolta és további analízisra adott lehetőséget. Roaming idején az adaforgalom iránya fontos, hiszen a huzalos (down) vagy a vezeték nélküli (up) oldalon a roaming miatt puffertelést végzése szükséges, ami szignifikánsan befolyásolja a TCP kapcsolatok működését.

Az ICMP üzenetek méretét úgy választottuk meg, hogy a spay ping esetén 64 bájt, 1500 bájt, illetve 32 kilobájt méretű legyen az adatkapcsolati keret mérete. Ennek jelentősége a minimális, maximális keretméret (MTU), illetve az IP csomag szegmentálásánál van. Az autó sebessége a két bázisállomás közötti szakaszon konstans volt és a lakott területen szokásos értékek szerint közlekedtünk (10 km/h, 30 km/h, 50 km/h).

4. ábra A mérési környezet



Paraméterek		Független mérések	
Bázisállomás	Cisco Aironet AP1120	L4 protokollok	TCP (FTP), UDP (Spray)
Mobil Terminál	Cisco Aironet 350 Series	L3 protokollok	IPv4, IPv6
AP IOS (1 mW)	12.2(11)JA	TCP forgalom	MT ⇒ Szerver (Up) Szerver ⇒ MT (Down)
MT és szerver OS	Windows XP	UDP üzenet [B]	18, 1472, 31970
MT Radio Firmware (1 mW)	Win/NDIS Driver 8.5/3.8, ACU v6.3, ACM v2.3	Sebesség [Km/h]	10, 30, 50
FTP szerver (IPv4/IPv6)	Orenosv v0.8.0	D(APA,APB)	100 m

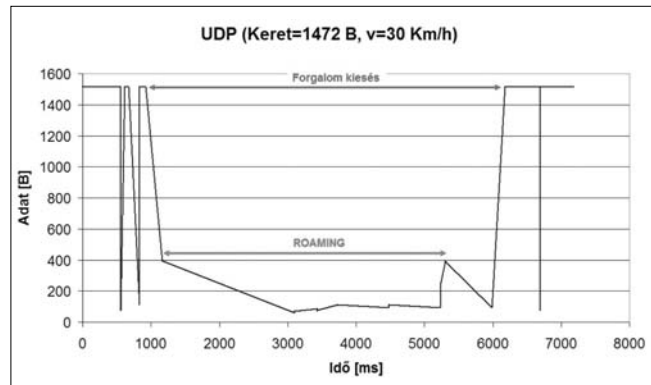
3. és 4. táblázat

5. Mért jellemzők, értelmezés

A capture programmal vételezett keretsorozatból értelmezni lehet a roaming folyamat jelzését, valamint a szállítási réteg forgalmát. Ezáltal mérhetővé vált a roaming $R[ms]$, illetve az ebből származó forgalom kiesés $T[ms]$ időtartama.

A TCP kapcsolatok „Slow Start” és „Windowing” algoritmusai alapján valósul meg a nagyméretű fájlok továbbítása FTP-vel. Az adatkapcsolati réteg átviteli sebességének változása a window méretének szabályozását teszi szükségessé. A WiFi átviteltechnika roaming fázisának időtartama erőteljesen befolyásolja a TCP hatásfokát. Az UDP átvitel a jellegéből adódóan sokkal alkalmazkodóbb természetű. Másodpercenként 100 csomagot küldtünk a spray ping segítségével, amely a csomagmérettől (64 bájt, 1500 bájt, 32 kbájt) függően a rádiós csatornát 0,93%, 21,82%, illetve 100%-ig terhelte. A mért paraméterek alapján a következő megállapításokat tehetjük:

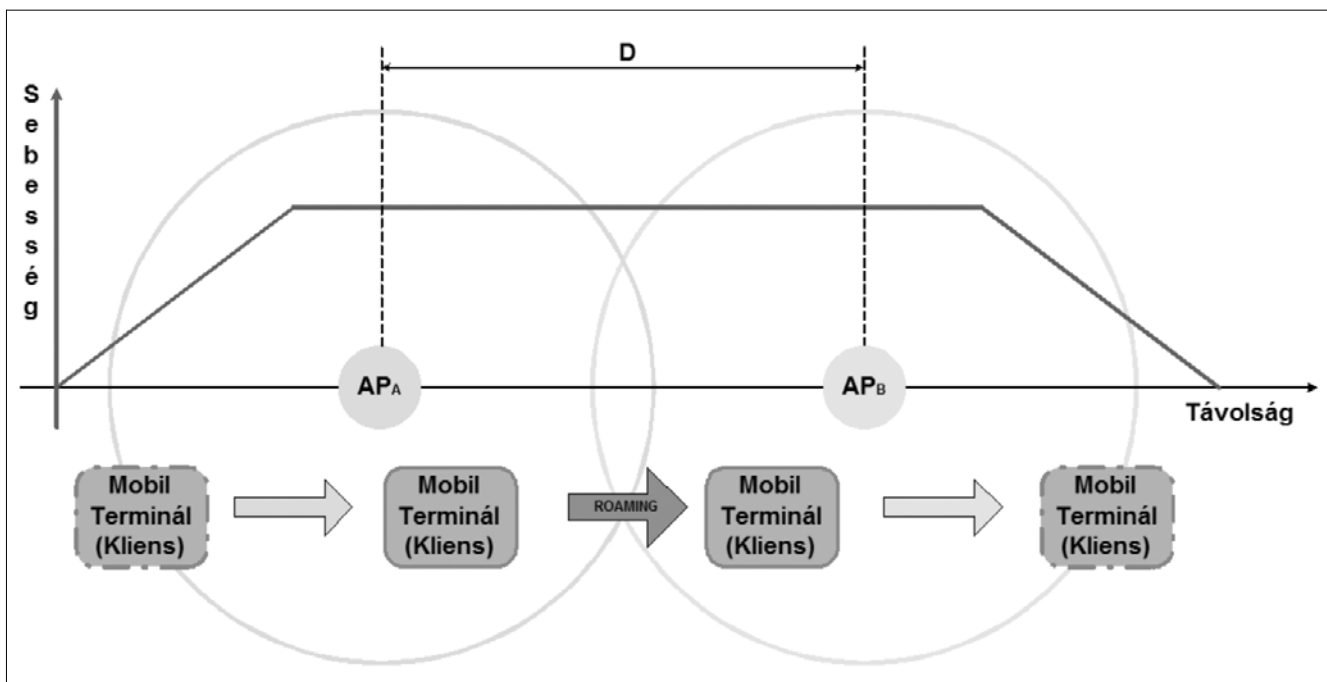
❶ Ethernet alapértelmezett MTU alatti keretméret (1500 bájt) esetén az ICMPv4 roaming ideje csökken a

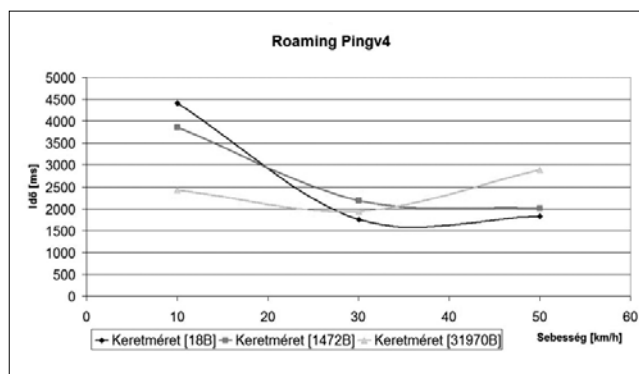
6. ábra Mért paraméterek ($R[ms]$, $T[ms]$)

sebességgel, míg MTU feletti keretméret esetén ugyanez növekvő tendenciát mutat. Szegmentációnál több időbe telik a csomagok sorrendjének visszaállítása. Az ICMPv6 roaming ideje minden keretméret esetén gyakorlatilag csökken a sebességgel. Ez az IPv6 közegérzékelő tulajdonságának tulajdonítható.

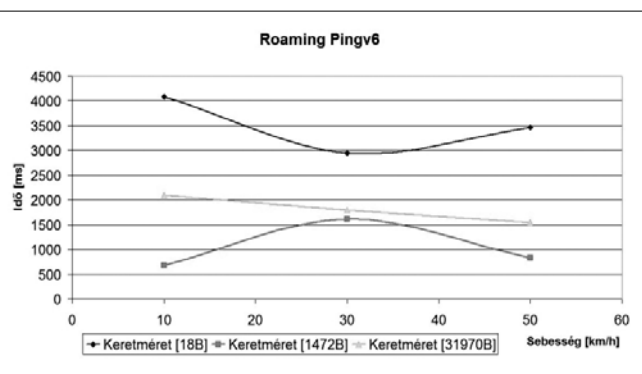
❷ Ethernet alapértelmezett MTU keretméret alatt az ICMPv4 forgalom kimaradása a sebességtől független, de a szegmentáció az időkéssleltetést csökkenti. A lát-

5. ábra A roaming folyamata





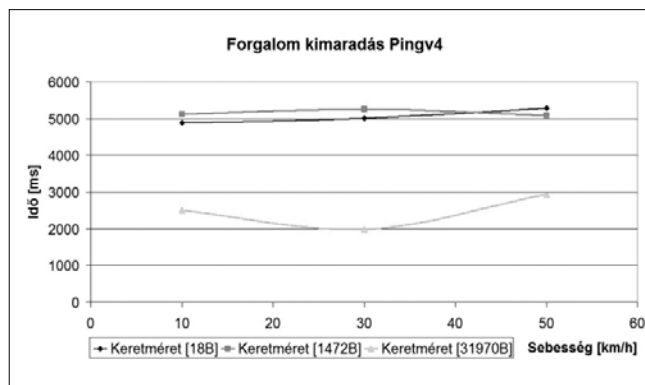
7. ábra Roaming keretméret függése ICMPv4 esetén



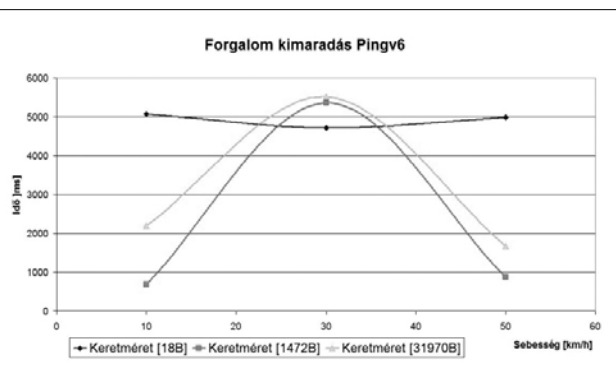
8. ábra Roaming keretméret függése ICMPv6 esetén

szólag ellentmondásos hatás a csatorna folyamatos terhelésével magyarázható. Az ICMPv6 esetén erőteljes ingadozás tapasztalható a sebesség függvényében. Ennek oka az IPv6-nak az L2 rétegtől való lekapcsolódása miatt adódik. Ezt a lekapcsolódást az IPv4 nem teszi meg, így kevésbé érzékeny az ICMPv4.

szignifikánsan nagyobb adatkiesést szenved, mint a feltöltés esetén. Ez abból adódik, hogy a TCPv4 sokkal gyorsabban változtatja a window méretét, ebből következően letöltéskor sok adat elvesz a régi bázisállomás irányába küldött jelentős mennyiségű keretszám miatt.



9. ábra Roaming hatása az UDPv4-re



10. ábra Roaming hatása az UDPv6-ra

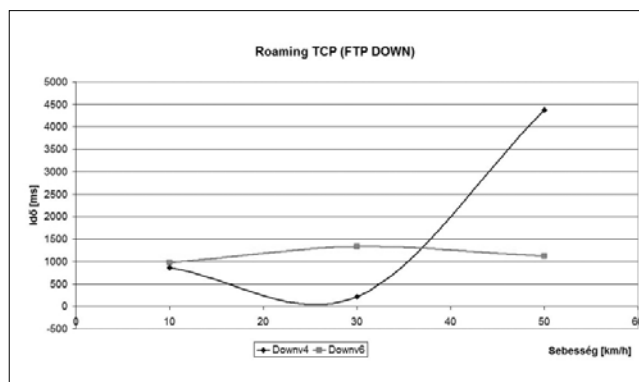
③ A forgalom kimaradás ICMPv4 esetén másfél másodperccel, ICMPv6 esetén pedig csak egy másodperccel hosszabb, mint a roaming időtartama. Kis keretméret esetén kevésbé függ a sebességtől az ICMP forgalom kimaradása, míg nagyobb keretméretek esetén csak az ICMPv6 érzékeny a sebességre.

⑤ A TCPv4 nagyobb window mérettel és lassúbb dinamikával dolgozik, míg a TCPv6 kisebb ablakméreteket alkalmaz és gyorsan szabályozza. Emiatt az TCPv6 jobban viseli a WiFi környezet roaming eseményeit, csökkentve ezáltal a forgalom kieséseket.

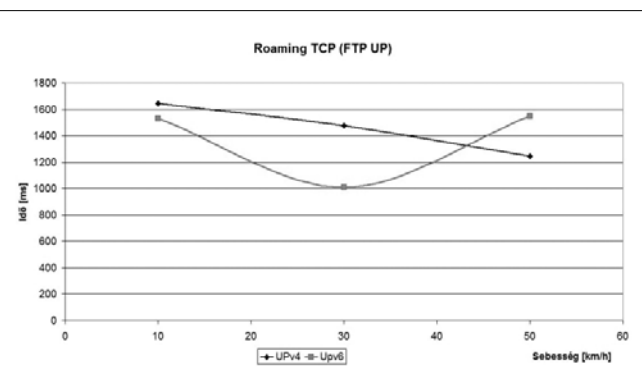
④ A forgalom kimaradás TCPv6 esetén független az adatfolyam irányától. TCPv4 esetén azonban a letöltés

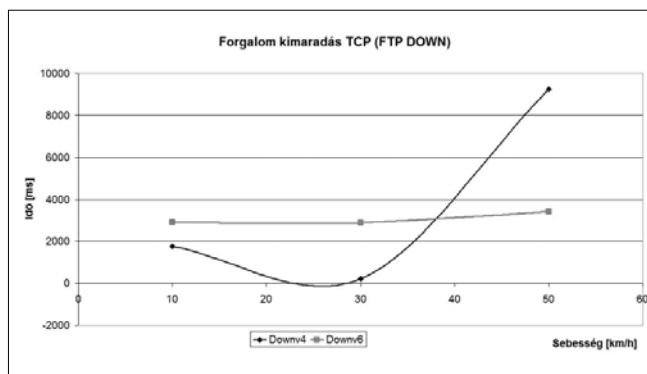
⑥ A TCPv4 forgalom kimaradása lefelé irányú adatforgalom esetén függ a mobil terminál mozgási sebességétől. 50 kilométeres óránkénti sebességnél akár 9,2

11. ábra Roaming TCP letöltésnél



12. ábra Roaming TCP feltöltésnél





13. ábra Roaming hatása a TCP-re letöltésnél

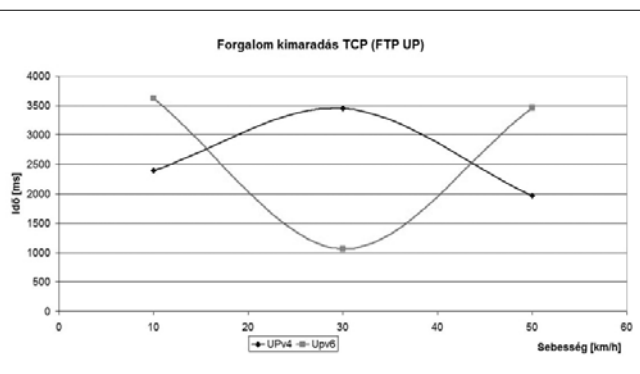
sec kiesést is képes produkálni. Ez lehetetlenné teszi a gyors járművekből történő folyamatos kommunikációt. TCPv6 lefelé forgalomnál ez az érték gyakorlatilag független a mozgási sebességtől és 3,8 sec alatt marad. A TCPv4 forgalom kimaradása felfelé irányú adatforgalomnál kis mértékben módosul a terminál sebességével, míg ugyanez TCPv6 esetén jelentősen változik.

6. Összefoglalás

A mobil kliens bázisállomásokhoz viszonyított relatív sebessége és a roaming végrehajtásának kölcsönhatása jelentősen befolyásolja a TCP kapcsolatokat, miközben kevésbé hat az UDP átvitelre [5,6]. Az összehasonlító mérésekből statisztikai módszerekkel nyert eredmények lehetővé teszik, hogy valós képet kapjunk az IPv4 és az IPv6 mobil átvitel esetén tanúsított viselkedésre vonatkozóan, valamint választ kaphatunk arra a kérdésre, hogy valóban magasabb minőségű mobil adatátvitelt eredményez-e az IPv6 protokoll vezeték nélküli adatkapcsolati réteg fölött elődjéhez, az IPv4-hez képest.

A hagyományos elektronikus alkalmazások az IPv4 protokoll „best effort” jellege miatt lassúbb átvitelt biztosítanak mobil WiFi környezetben, míg az IPv6 protokoll az alsóbb rétegekhez történő gyors adaptáció miatt hatékony átvitelt képes biztosítani. Az időérzékeny alkalmazások (IP telefon, videokonferencia stb.) az IPv4 protokoll QoS korlátai miatt mobil WiFi környezetben nagy kieséseket szenvednek, így a minőség elfogadhatatlan. Az IPv6 gyors adaptációja miatt a kiesések kisebbek, ezért a jelenlegi mobil WiFi környezetben fast roaming esetén közel elfogadható minőségű infokommunikációs szolgáltatások használhatók [4].

A témával kapcsolatosan további elemzési lehetőség a lakott területen kívüli környezetben, nagyobb mozgási sebességgel haladó mobil terminálok (autópályán, vonaton) adatkommunikációs szolgáltatásainak minőségét befolyásoló tényezők feltárása és értelmezése. Egyértelműen körvonalazódik, hogy a vezeték nélküli kapcsolatok mobil funkcióinak kiaknázásához halaszthatatlan egyrészt a roaming folyamat gyorsítása, másrészt pedig az IPv6 feletti speciális alkalmazások kifejlesztése.



14. ábra Roaming hatása a TCP-re feltöltésnél

Irodalom

- [1] Microsoft TechNet, The Cable Guy – Sept-2004: „Introduction to Mobile IPv6”, <http://www.microsoft.com/technet/community/columns/cableguy/cg0904.msp>
- [2] Charles E. Perkins, Sun Microsystems: „Nomadcity: How mobility will affect the protocol stack”, <http://www.computer.org/internet/v2n1/nomad.htm>
- [3] Microsoft Corporation: „Understanding Mobile IPv6”, <http://www.microsoft.com/downloads/details.aspx?FamilyID=f85dd3f2-802b-4ea3-8148-6cde835c8921&displaylang=en>
- [4] Ye Tian, Kai Xu, Nirwan Ansari: „TCP in Wireless Environments: Problems and Solutions”, IEEE Radio Communications, March 2005.
- [5] Gál Zoltán, Karsai Andrea: „Videokonferencia rendszerek minőségi garancia jellemzőinek elemzése”, NetworkShop'2004 konferenciakiadvány, Széchenyi István Egyetem, Győr, 2004. április 5-7.
- [6] Zoltán Gál, György Terdik: „Multifractal Study of Wireless and Wireline Datanetworks”, 8th International Conference on Advances in Communications and Control, Telecommunications/Signal Processing – Proceedings, Crete, Greece, 25-29 June 2001.
- [7] Cisco Systems, Inc.: „Cisco Fast Secure Roaming” – technical paper.

A következő generációs mobil hálózatok fejlődési tendenciái

PÉTERFALVI GÁBOR, POZSÁR BALÁZS, SIMON VILMOS, HUSZÁK ÁRPÁD, IMRE SÁNDOR

BME Híradástechnikai Tanszék, Mobil Távközlési és Informatikai Laboratórium
{petergab, pozsy, svlmos, huszak, imre}@mcl.hu

Kulcsszavak: következő generációs mobil hálózatok, mobil multimédia, QoS

A mobil távközlési piac nagy mértékű fellendülése túlnyomórészt a mobil beszédszolgáltatás növekvő népszerűségének volt köszönhető az elmúlt évtizedben. Elérve a lehetséges mobil felhasználók számának felső korlátját, már nem várhatjuk a hálózati forgalom további növekedését csupán a beszéd alapú szolgáltatástól, így a jövőben egyre inkább a multimédia továbbítása fogja adni a hálózati terhelés nagy részét. A multimédia ilyen arányú megjelenése a forgalomban, nagy mennyiségű adat továbbítását követeli meg a hálózat peremén, a hozzáférési hálózatban is, ami által egyre nehezebb lesz garantálni a megfelelő szolgáltatás-minőségi paramétereket (QoS). Sok szakvélemény szerint a 3G hálózatok nem lesznek alkalmasak igazán szélessávú multimédia alkalmazások használatára, ezért már sok kutató a 4G irányában fejt ki érdeklődését. Ezekről a tendenciákról adunk most áttekintést; olyan új igényekről, melyek szükségessé teszik a jövőben a 4G hálózatok kialakítását, és hogy jelenleg milyen technológiai lehetőségek állnak rendelkezésünkre ezen hálózatok megvalósításához.

1. Bevezetés

Az előző évtizedben tanúi lehettünk a mobil telekommunikációs hálózatok hihetetlen gyors fejlődésének. A kezdetben analóg rendszereket (1G) hamar felváltották a második generációs (2G) hálózatok, mint például a GSM vagy az IS-95, melyek a jó minőségű hangtovábbítás mellett már kis sebességű adattovábbítást is lehetővé tettek. A 2G technológia, élen az európai GSM megvalósítással, a 90-es években igazi sikertörténet volt, és az új évezred hajnalán élte fénykorát. Napjainkban folyik az áttérés a 2G-ről az Európában UMTS, az USA-ban CDMA-2000, Japánban W-CDMA néven futó harmadik generációs mobil hálózatokra (3G), melyek új kódolási, mobilitási megoldásokkal, nagyobb sebességgel próbálják kielégíteni a 21. századi társadalom megnövekedett igényeit.

A fejlődés nem állhat meg. Annak ellenére, hogy a 3G hálózatokat még csak néhány országban valósították meg, egyre inkább előtérbe kerülnek a következő, negyedik generációs hálózatok (4G) tervezési-megvalósítási kérdései is. Ennek okai a sávszélesség, a mobilitás, a minőség további növelésének alapvető igénye, új ultra-szélessávú szolgáltatások bevezetése, valamint olyan fejlődési trendek, mint például a „mindenütt jelenlét” (ubiquity), vagy a globalitás, melyek a 3G fejlesztésénél még nem voltak tervezési szempontok.

Cikkünkben szó lesz azokról a fejlődési tendenciákról melyek napjainkban és a közeljövőben a mobil hálózatok területén irányadónak mutatkoznak; a 3G hálózatok lehetőségeinek határáról; olyan új igényekről, elvárásokról melyek szükségessé teszik a jövőben a 4G hálózatok kialakítását; végül bővebben kifejtjük, hogy jelenleg milyen technológiai lehetőségek állnak rendelkezésünkre ezen elvárásokat kielégítő 4G hálózatok megvalósításához.

2. A mobil hálózatok fejlődési iránya ma és a közeljövőben

A mobil telekommunikációs piac óriási fejlődése a 90-es években a mobil beszédszolgáltatás futótűszerű terjedésének volt köszönhető világszerte. Egyre többen vásároltak mobil készüléket, és éltek a mobilitás nyújtotta lehetőségekkel. Ennek eredményeképpen például Japánban 2000-re a mobil előfizetések száma meghaladta a vezetékes előfizetéseket. Azonban, ha már egy ország minden lakosának van mobil készüléke, akkor nem várhatjuk az előfizetések, vagy a hálózati forgalom további növekedését csupán a beszéd alapú szolgáltatástól. A mobil hálózatok fejlődésének fenntartásához új irányelvekre van szükség. Ezek a következők:

2.1. Új valósídejű, multimédiás szolgáltatások bevezetése

Ha egy szóval kellene kifejeznünk, hogy manapság mi a mozgatórugója, lendkereke a vezetékes és vezeték-nélküli hálózatok fejlődésének, akkor valószínűleg sokunk azt mondaná: a *multimédia*. Valóban, a hang-, kép- és videófájlok hatalmas mennyiségű forgalma kezd uralkodóvá válni az Interneten, valamint a vezeték-nélküli hálózatokban is kezd megjelenni az alapvető hálózati forgalom mellett a kifelbontású képek és videók továbbítása.

Az erre épülő szolgáltatások – például az MMS – már megjelentek a 2G mobil hálózatokban. A multimédiás szolgáltatások nagyban növelik a hálózati forgalmat, és ez által a szolgáltatók bevételeit is, így finanszírozzák a jövőbeli befektetéseket. A 3G hálózatokban ez a tendencia egyre inkább jellemző lesz, és folytatódni fog a jövőben is. A szabványosított interfészeknek (Parlay) köszönhetően bárki fejleszthet majd olyan alkalmazásokat, melyek a mobil terminálokra futva igénybe veszik a hálózat szolgáltatásait. Egyre újabb, már

valósídejű multimédiás szolgáltatások jelennek meg, melyek azonban már jóval nagyobb követelményeket támasztanak a kiszolgáló hálózattal szemben. A hűtött nagyfelbontású mobil videokonferencia szolgáltatás például 2G rendszerekben semmiképpen, de lehet, hogy még 3G rendszerekben sem lesz megvalósítható.

2.2. Mindenütt jelenlevő mikro-mobil terminálok

Egy másik lehetőség a bevételek növelésére a „mindenütt jelenlevő” mobil terminálok térhódítása lehet. A jövőben már nem csak közvetlenül emberek használhatják majd a hálózatot, hanem minden, ami mozog, például gépek vagy állatok. Az autókba épített fedélzeti számítógép például magától letölti a legújabb programokat az éppen elérhető mobil hálózatról, az intelligens bőröndünk jelzi, hogy éppen melyik repülőtéren várakozik, vagy a kutyánk nyakába akasztott mikro-mobil terminál tájékoztat majd bennünket, hogy a házi kedvenc éppen merre kóborol.

A hálózathoz csatlakozó terminálok száma így nagyságrendileg is nőhet, ami teljesen új kihívásokat jelent a jövőbeli hálózatok tervezésénél [1].

2.3. Globális és lokális hálózati átjárhatóság

Globális hálózati átjárhatóságon itt azt értjük, hogy például az Európában használt mobil készülékek használhatóak legyenek Amerikában, vagy Japánban is. Ha például valaki felszáll Londonban egy repülőgépre és elutazik New York-ba, akkor ugyanúgy használhatja azt az USA-ban is. A globális, kontinenseken átívelő barangolás (roaming) megvalósítására, például UMTS és CDMA-2000 között IP szinten már születnek megoldási javaslatok [2]. Az IP alatti rétegekben, többek között a különböző kódolási, sávkiosztási, menedzselési eljárások eltérése miatt ez a probléma remélhetőleg szoftver-rádió segítségével megoldhatóvá válik. Az átjárhatóság feltétele tehát a jelenleg még kontinensenként különböző hálózati technológiák teljes konvergenciája. Ehhez globális, mindenki által elfogadott hálózati architektúrák és protokollok szükségesek, melynek megvalósítása manapság még futurisztikusnak tűnik.

A lokális hálózati átjárhatóság azt jelenti, hogy egy mobil terminál egyidejűleg hozzáférhessen különböző technológiákat alkalmazó vezeték-nélküli hálózathoz. Ez arra ad lehetőséget, hogy akár a felhasználó, akár a mobil készülék önállóan kiválassza az adott pillanatban igénybe vett szolgáltatáshoz a sávszélességben, költségben leginkább megfelelő hálózatot. Például, ha az előfizető megérkezik a munkahelyére, és ott ingyenes WLAN áll rendelkezésére akkor mobil készüléke váltson át az éppen használt cellás hálózatról az ingyenes, nagyobb bitsebességű WLAN-ra, anélkül hogy az éppen igénybe vett szolgáltatás megszakadna.

A különböző hálózati technikák közötti váltást Vertical Handover-nek (VHO) nevezzük. A VHO lehetővé teszi a mindig a legmegfelelőbb hálózathoz való hozzáférést (Allways Best-Connected Network, ABCN) megjelölését.

Tekintsük most át, hogy a fentiekben bemutatott fejlődési tendenciák milyen általános technológiai elvárások elé állítják a jövő hálózatait.

3. Technológiai elvárások a jövő vezeték-nélküli hálózataival szemben

Mint azt az első pontban láttuk, a jövőben egyre inkább a multimédia továbbítása fogja adni a hálózati terhelés nagy részét. A multimédia ilyen arányú megjelenése a forgalomban, nagy mennyiségű adat továbbítását követeli meg a hálózat peremén, a hozzáférési hálózatban is. Mobil hálózatok esetében ez a rádiós interfész (RI) nagymértékű terhelését vonja maga után, és így az RI sávszélessége szűk keresztmetszete lehet a multimédiás szolgáltatásoknak. Alapvető elvárás tehát a sávszélesség drasztikus növelése.

Egy másik hasonlóan fontos elvárás az újgenerációs hálózatokkal szemben a szolgáltatás minőségének (QoS – Quality of Service) biztosítása. Azonban a hagyományos IP alapú, csomagkapcsolt hálózatok, ahogy az Internet is csak best-effort minőségű szolgáltatásokat biztosítanak. Ez azt jelenti, hogy mindent megtesznek a csomagok gyors továbbításáért, de nem várhatjuk el a hálózattól, hogy megfeleljen olyan konkrét minőségi paramétereknek, mint például a csomagok maximális késleltetése. Azonban főleg a valósídejű szolgáltatások, mint a fent említett videokonferencia is, kifejezetten érzékenyek a sávszélesség ingadozására, vagy a csomagok késleltetésére. Ahhoz tehát, hogy ilyen szolgáltatásokat biztosíthassunk az előfizetőknek, magas szintű QoS-t kell biztosítanunk az adatforgalom számára [3].

Vezeték nélküli hálózatok esetében a QoS biztosítása szorosan összefügg a mobilitás egyszerű és gyors kezelésével mind RI, mind gerinchálózati, IP szinten. Cellás hálózatok esetében például a cellaváltások (handover, HO) kezelése súlyos problémákat vet fel a QoS biztosításában. Cellaváltásnál meg kell szakítani az összeköttetést a régi bázisállomással, és kapcsolatot kell létesíteni az újjal. Ez a hálózati szolgáltatás rövid idejű kiesését is jelentheti. Soft HO esetén is – amikor az új kapcsolat felépüléséig a régi megmarad, tehát elvileg nincs kiesés – a HO-t kezelő jelzésüzenetek az új cím megszerzéséig komoly késleltetést, és bitsebesség-csökkenést okozhatnak a hasznos forgalomban. A valósídejű csomagkapcsolt szolgáltatások bevezetéséhez tehát csökkenteni kell a HO okozta járulékos hálózati terhelést az RI-n.

Ahhoz, hogy a vezetékes hálózatoknál jól bevált IP protokollt mozgó eszközök esetében is használhassuk, IP szinten kell megvalósítani a mobilitást. A már létező megoldások a Mobil IPv4 és Mobil IPv6 protokollok azonban még továbbfejlesztésre szorulnak, például az alhálózat-mobilitás és a személyi mobilitás terén.

A IP szintű mobilitás kezelésének gyorsítására több javaslat is született, például a Hierarchikus Mobil IP, és a Regionális regisztráció [4].

A különböző hálózati architektúrák konvergenciája napjainkban egyre inkább felgyorsul. A konvergencia kulcsa az IP protokoll, mely összekapcsolja a különböző célokra, különböző technológiákkal és protollokkal megvalósított hálózatokat. A jövőben megvalósulnak olyan *All IP hálózatok*, amelyekben mind az adat, mind a beszéd IP csomagokban lesz továbbítva. Ehhez többek között mobil hozzáférés esetén is meg kell valósítani a *VoIP* (Voice over IP)-t. Az All IP hálózatok kialakítása alapvető fontosságú a globális és lokális hálózati átjárhatóság megvalósításához.

Az All IP hálózatok megjelenése azonban még nem elegendő a globális átjárhatóság és a vertical handover megvalósulásához. Ehhez ugyanis az IP alatti rétegekben továbbra is technológiai átváltásra van szükség, hiszen az elérhető rádiós interfészek fizikai és adatkapcsolati paraméterei különbözhetnek. Ezt a problémát szoftver-rádió alkalmazásával lehetne megoldani.

Végül ahhoz, hogy a második pontban bemutatott „mindenütt jelenlevőség” kialakulhasson, amellet, hogy *csökkenten kell az előfizetői díjakat, nagy mennyiségű új hálózati elemet kell telepíteni a nagyszámú terminál kiszolgálásához*. Ehhez nagy kapacitású, olcsó átviteli technikákat kell alkalmazni. Az átlagos átviteli bitköltség csökkentése tehát alapvető feltétele annak, hogy az új szolgáltatások széles körben elterjedhessenek.

A fejlődési irányok és a szükséges technológiai elvárások kapcsolatát az 1. ábra mutatja.

4. Mit nyújt a 3G?

Ebben a fejezetben áttekintjük, hogy a fent felsorolt technológiai elvárásokból – a 3G hálózati architektúrák és szolgáltatások alapján – mi valósult meg idáig.

4.1. 3G hálózati architektúrák, és technikák

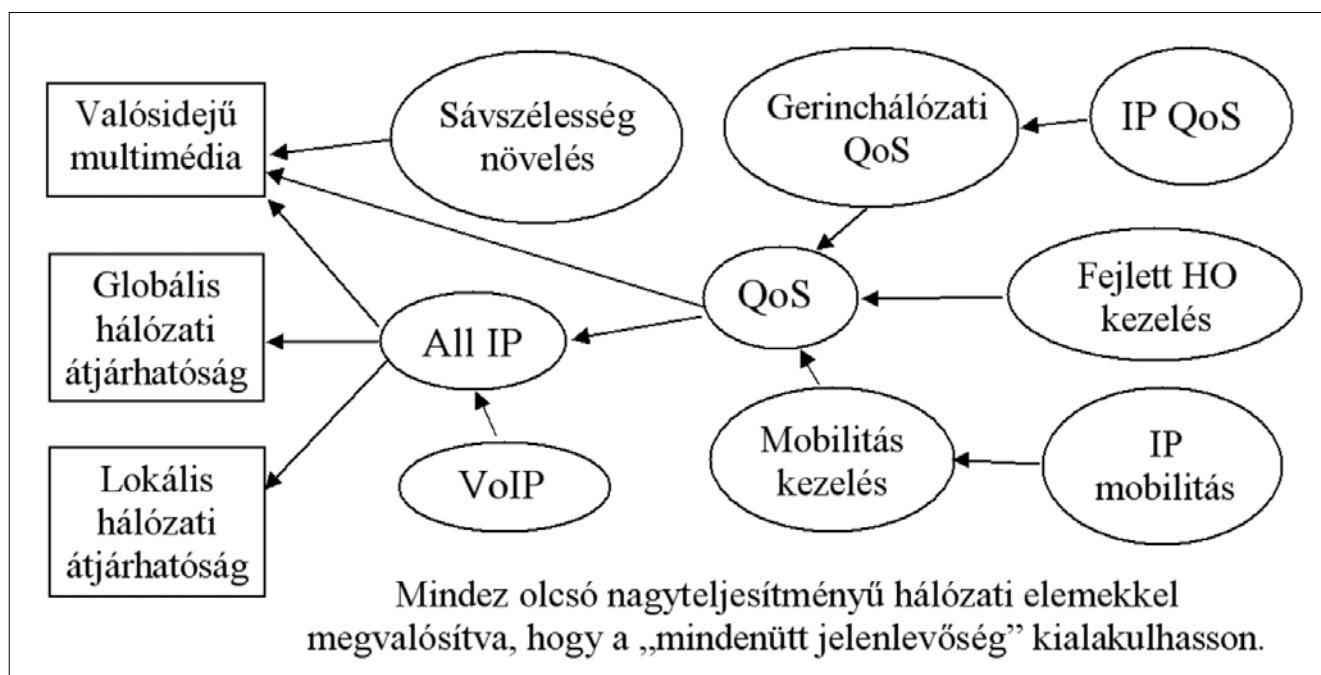
A 90-es évek végén kezdődött a 3G hálózatok szabványosítása, melyre két nemzetközi szervezet alakult: a 3GPP (Third-Generation Partnership Project), és a 3GPP2. A két megvalósítási javaslat több szempontból *eltért, az alkalmazott rádiós hozzáférési technológiában* például a 3GPP a *WCDMA*, míg a 3GPP2 a *CDMA-2000* technológiát javasolta.

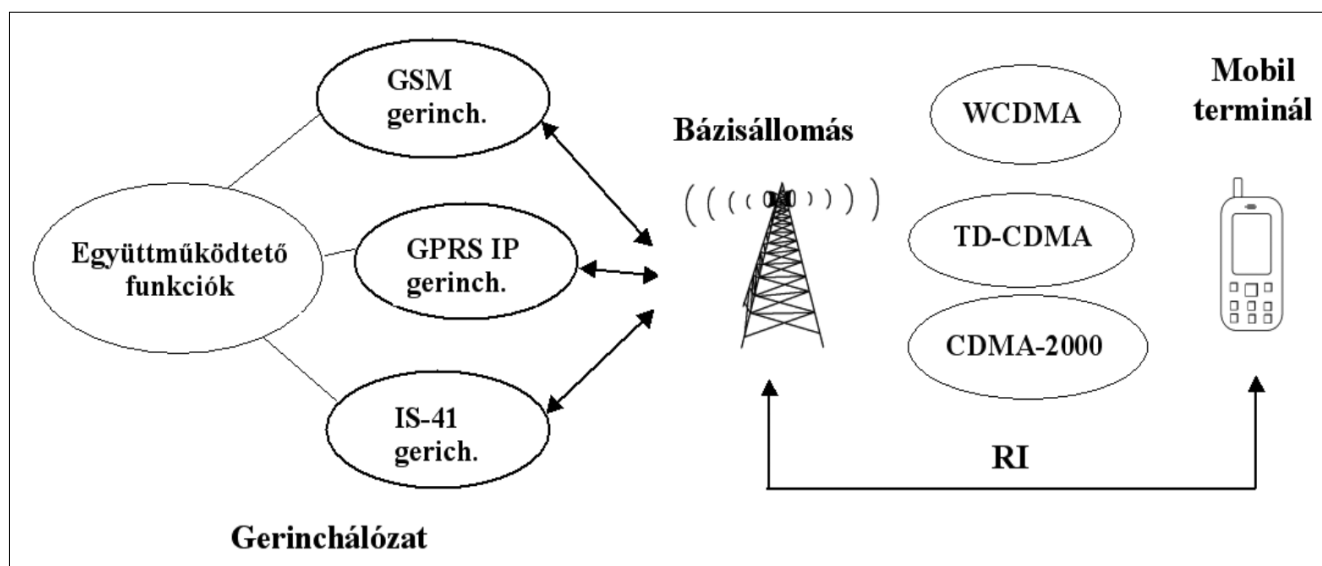
Az *UTRAN (UMTS Terrestrial Radio Network)-ban alkalmazott WCDMA esetében a többszörös hozzáférést az ortogonális Walsh-Hadamard kódok által megvalósított szórás (spreading) biztosítja*. A mobil terminál által adni kívánt bitsorozat egy ortogonális kódsorozattal szorozzuk meg, melynek frekvenciája SF (Spreading Faktor) egészszámu többszöröse az eredeti bitsorozat frekvenciájának. A megszorított kód egy bitje a chip. Adáskor a chip-sebesség állandó 3.84 Mchip/sec, tehát a bitsebességet az SF szám határozza meg. Mennél több terminál akar egyszerre adni a frekvenciasávban, annál nagyobb mértékű (nagyobb SF) szórást kell alkalmazni, hogy a jel-zaj arány ne romoljon, és a vétel még lehetséges legyen, ez pedig a bitsebesség csökkentését jelenti a terminálok számára. Ennek köszönhetően nincs éles felső korlát az egyidejűleg aktív terminálok számára, mivel az csak a még elfogadható jel-zaj aránytól függ. *Az elérhető legnagyobb bitsebességek vonalkapcsolt átvitelnél 384 kbit/sec, míg csomagkapcsolt átvitel esetében 2 Mbit/sec* [5,6].

A CDMA-2000 technológiában az egy vivőre jutó chip-sebesség kisebb: 1.2288 Mchip/sec; itt a bitsebességet több vivőfrekvencia (maximum három) használatával (Multi Carrier üzemmód) lehet növelni.

Mint láttuk a 3GPP és a 3GPP2 által javasolt rádiós technikák – elsősorban üzleti megfontolások miatt – eltérnek. Abban azonban a két szervezet egyetértett, hogy

1. ábra Fejlődési irányok és technológiai elvárások





2. ábra UMTS és CDMA-2000 architektúrák

a 3G hálózatoknak *folyamatos fejlődés útján kell létrejönniük*. Ez azt jelentette, hogy a 3G minél többet használjon fel a már létező 2G hálózati architektúrából (GSM, GPRS, és az amerikai IS-41). Erre elsősorban gerinchálózati szinten volt lehetőség, mivel a bitsebesség növelése érdekében az RI-n mindenképpen változtatni kellett. Ennek értelmében például Európában az UMTS kiépítésénél nem sokat változtattak a vonalkapcsolt GSM, és az IP alapú csomagkapcsolt GPRS gerinchálózati alapokon. Az UMTS és CDMA-2000 hálózati architektúrák váza a 2. ábrán látható.

4.2. 3G szolgáltatások

A 3G szabványosításakor a következő szolgáltatások megvalósítását tűzték ki célul:

- új multimédiás és valósídejű szolgáltatások bevezetése;
- a mobil Internet-hozzáférés lehetőségeinek kiterjesztése (például Internetes vásárlás);
- lokális információt nyújtó, adatszóró ill. streaming szolgáltatások bevezetése;
- a vezetékes telefonszolgáltatással ekvivalens beszédátviteli minőség nyújtása;
- váltás különböző rádiós technológiákat használó hálózatok között (VHO).

Az első 3G szolgáltatás-csomag: a FOMA

Az első kezdetleges 3G szolgáltatásokat 1999. februárjában vezették be Japánban. A FOMA (Freedom

of Mobile Multimedia Access) szolgáltatás-csomag – mely már W-CDMA átviteli technológiát alkalmazott – akkor nagy újdonságnak számított világszerte, mára pedig több mint 36 millió előfizetője van Japánban. Az azóta továbbfejlesztett FOMA nyújtotta szolgáltatásokat az 1. táblázat mutatja. Mint az a táblázatból látható, a FOMA már képes többféle multimédiás és valósídejű szolgáltatás nyújtására, például a kislebontású videotelefon szolgáltatásra. Beszédátviteli minősége megegyezik a vezetékes telefonszolgáltatás minőségével, és nagysebességű adathozzáférést biztosít.

UMTS szolgáltatások

Az UMTS-ben a szolgáltatások *négy forgalmi osztályba sorolhatók*, az alapján, hogy megvalósításukhoz milyen minőségű adattovábbításra van szükség:

1. A *valósídejű* szolgáltatások osztályába tartozik természetesen a hagyományos beszéd szolgáltatás, a videotelefonia, valamint a valósídejű Internetes játékok. Ezek a szolgáltatások különösen érzékenyek a késleltetésre, és a sáv szélesség ingadozásra. Jelenleg ezek a szolgáltatások, – ahogy a FOMA esetében is – még csak viszonylag alacsony bitsebességen vehetők igénybe. Ez a beszéd szolgáltatás esetében nem nagy probléma, mivel az átlagos bitsebesség-igénye nem nagy, ezért az alkalmazott AMR (Adaptive Multi-Rate) hangátviteli technológia el is éri a kitűzött PSTN minőséget. Videó-telefonia esetében azonban a mozgóképátvitel jelenleg csak kis felbontásban oldható meg.

i-mode	Nagyméretű e-mail (max. 5000 japán karakter) továbbítása, képekkel és hanggal, i-apli alkalmazások (pl. JAVA játékok a hálózaton)
Videó-telefon	Max. 64 kbit/s sebességű, kislebontású mozgókép-továbbítással
Nagysebességű adathozzáférés	Downlink: max. 384 kbit/s, Uplink: max. 64 kbit/s
Beszédátviteli minősége	A vezetékes telefonszolgáltatással ekvivalens

1. táblázat
FOMA szolgáltatások

2. *Streaming*-en általában azt értjük, amikor egy multimédia file lejátszását már a letöltés alatt megkezdjük. Ez elsősorban azért kényelmes, mert nem kell megvárni, míg sikerül letölteni az egész fájlt, valamint gyengén valósidejű (több másodperces késleltetés sem zavaró) VoD (Video on Demand) műsorszórásra is alkalmas. Mivel ennél az osztálynál a valósidejűség már nem olyan szigorú elvárás, ezért például buffereléssel kiküszöbölhető a sáv szélesség kismértékű ingadozásának hatása.

3. Az *interaktív* osztályba tartoznak a web-böngészés, a lokális szolgáltatások, a távoli adatbázis hozzáférés, és különféle egyéb kliens-szerver szolgáltatások. Ezek a szolgáltatások általában a kérés-válasz (request-response) kiszolgálási sémára épülnek, a körbefordulási késleltetés tehát kritikus lehet ezeknél a szolgáltatásoknál is, bár valósidejűnek már nem nevezhetők.

4. A *hátér* osztályba olyan szolgáltatások tartoznak, melyeknek adatátviteli igényeivel elég akkor foglalkozni, amikor a magasabb szintű szolgáltatások nem terhelik a hálózatot. Tipikusan ebbe az osztályba tartoznak az e-mail, SMS, és elektronikus képeslap vagy MMS továbbítás. Ezeknél a szolgáltatásoknál a késleltetés lehet több perc vagy óra is.

4.3. További fejlesztési lehetőségek a 3G-ben

A Parlay csoport

A Parlay Csoport 1998-ban több nemzetközi vállalat összefogásával alakult, melynek célja olyan API-k (Application Program Interface) definiálása volt, melyek támogatják a külső alkalmazásokat a hálózatban. A Parlay Csoport egy olyan biztonságos és gazdag API csomagot fejlesztett, mely biztosítja a skálázhatóságot és a bővíthetőséget. A Parlay API lehetővé teszi a hálózati operátorok, szolgáltatók, általános software-fejlesztők számára, hogy telekommunikációs szolgáltatásokat integráljanak bármely IT szoftverjükbe, így biztosítva a titkos és valósidejű kommunikációt. A Parlay Csoport célja a telekommunikációs piac lehetőségeinek kiterjesztése volt, ugyanúgy, ahogy a PC megjelenése teret nyitott a szoftver-fejlesztőknek új kreatív és innovatív programok készítésére. A Parlay API technológia független interfészek egy olyan halmazát definiálja, mely metódusokat, eseményeket, paramétereket, és ezek szemantikáját határozza meg, úgy hogy külső (nem megbízható harmadik fél), és belső (megbízható hálózat operátor) alkalmazásfejlesztők hozzáférhessenek a gerinchálózat erőforrásaihoz és lehetőségeihez. A Parlay API-k tehát lehetővé teszik az új szélessávú multimédiás szolgáltatások gyors megjelenését, azáltal, hogy kisebb, vagy nem mobil-szolgáltatással foglalkozó szoftver-fejlesztő cégek is fejleszthetnek ilyen szolgáltatásokat [7].

Szoftver-rádió

A kommunikációs technológiák gyors fejlődésének köszönhetően a rádiós rendszerek egyre nagyobb részét valósítják meg szoftver alapon. A szoftver-rádió egy olyan rádiós eszköz, melynek csatorna-modulációs hullámformái szoftverben definiáltak. Tehát a hullámfor-

mák szoftveresen generáltak, melyekből egy szélessávú DAC (digitális-analóg átalakító) készít analóg jeleket. A vevő egy szélessávú ADC-n (analóg-digitális átalakító) alapul, és az így kapott digitális jel konverzióit és demodulációját már a processzoron futó software végzi.

A szoftver-rádiókban a különböző mobil hozzáférési szabványoknak az alkalmazott függvények különböző paraméterlistái felelnek meg. Ezáltal lehetővé válik a különböző hozzáférési technológiák közötti gyors váltás, ami nélkülözhetetlen a VHO megvalósításához. Az így kapott software által meghatározott rádió lényegében egy jól programozható hardware, amely a programozhatósága által nagyon rugalmas. Ez a rugalmasság egyszerűsíti az architektúrát, növeli a kompatibilitást, és lehetőséget ad új technológiák könnyebb bevezetésére [8].

4.4. A 3G határa

Láttuk, hogy az elsődleges elvárás, az RI sáv szélességének növelése megvalósult a 3G-ben, köszönhetően a W-CDMA és CDMA-2000 technológiáknak. Ez természetesen minimális követelménye volt a multimédiás szolgáltatások bevezetésének. Azonban annak ellenére, hogy az új kódolási megoldásokkal akár 2 Mbit/sec adatátviteli sebesség is elérhető csomagkapcsolt esetben, *a 3G valósidejű szolgáltatások bitsebessége általában 384 kbit/sec alatt marad*. Ez azért van így, mert a valósidejű szolgáltatásokat a 3G-ben *továbbra is vonalkapcsolt átvittel oldják meg*, a hagyományos 2G vonalkapcsolt gerinchálózati architektúrán. Ennek oka pedig az, hogy 3G-ben még *nem sikerült megfelelően megoldani a valósidejű csomagkapcsolt átvitelhez szükséges QoS biztosítását, és az ezzel szorosan összefüggő mobilitást (HO-k) sem*. Tehát a valósidejű továbbítást igénylő adatokat nem tudjuk a nagyobb teljesítményű, IP alapú, csomagkapcsolt gerinchálózaton továbbítani.

A globális és lokális hálózati átjárhatóság szinte teljes hiánya a 3G hálózatokban ugyanerre a problémára vezethető vissza. Napjainkban kezdenek megjelenni olyan mobil készülékek, melyek már több különböző RI-t támogatnak, ezekkel a VHO elviekben már megvalósítható lenne. Több szolgáltatás azonban továbbra sem vehető igénybe, bármilyen hozzáférési hálózat alkalmazásával. Ha például telefonálunk egy UMTS hálózaton keresztül, és közben beérünk egy ingyenes IP alapú WLAN hálózatba, akkor át kellene állni a vonalkapcsolt hangátvitelről a csomagkapcsolt VoIP-ra, úgy hogy a beszélgető felek azt ne vegyék észre, ezt pedig nagyon nehéz lenne megoldani. A helyzet sokkal egyszerűbb lenne, ha a celluláris hálózatban is csomagkapcsolt módon IP fölött menne a teljes forgalom, beleértve a valósidejű szolgáltatásokat is, vagyis ha megvalósulna az *All IP hálózati struktúra*.

Összefoglalva tehát a fentieket: a 3G sok mindent elért a kitűzött célokból, de *igazi változást nem hozott a 2G-hez képest*, elsősorban azért, mert *szorosan a 2G-re épül*. A 3G elsősorban a hozzáférési hálózaton

változtatott, ezáltal próbál az új elvárásoknak megfelelni több-kevesebb sikerrel. A 4G hálózatok valószínűleg nagyobb lépést jelentenek majd a 3G-hez képest, mert **RI és gerinchálózati szinten is változtatásokra lesz szükség.**

5. Elvárások a 4G hálózatoktól

Sokak szerint a 3G hálózatok valójában nem lesznek alkalmasak igazán szélessávú multimédia alkalmazások kiszolgálására. A 4G hálózatok azonban az olyan jó minőségű videó adások továbbítására is alkalmasak lesznek, mint a HDTV, köszönhetően az akár 100 Mbit/s letöltési sebességnek, mely a sokszorosa a 3G-beli értéknek. Ez már ma is elérhető WLAN környezetben, de a fejlesztés alatt álló hozzáférési technológiák segítségével még jobb eredmények érhetőek el.

Az ilyen sávszélességet nyújtó hozzáférési technológiához nagyon jó alapot nyújt az *OFDM (Orthogonal Frequency Division Multiplexing)*, mely különösen jó interferenciátűrővel rendelkezik. Ennek egyik továbbfejlesztése a *W-OFDM (Wide-band)*, mellyel sikerült már a bázisállomástól több kilométeres távolságban lévő, gyorsan mozgó (100 km/h) terminálra is 32 Mbit/s letöltési sebességet elérni. Az NTT DoCoMo cég a *VFS-OFCDM (Variable Spreading Factor Orthogonal Frequency and Code Division Multiplexing)* és *VFS-CDMA (Variable Spreading Factor Code Division Multiple Access)* technológiákkal kísérletezik, előbbi a downlink irányban nyújt nagyon nagy sávszélességet, míg az utóbbi mindkét irányban gyors.

A harmadik generációs hálózatok szabványait igyekeztek egységesíteni, ám ez mégsem sikerült. A kontinensek közti roaming feltétele, hogy a hozzáférési technológia egységes legyen, vagy legalábbis a készülékek alkalmasak legyenek az elterjedt technológiai megoldások igénybevételére. Úgy tűnik, hogy az egységes szabvány a 4G hálózatokban sem fog megvalósulni, ám ennek elsősorban gazdasági, politikai okai vannak. Azonban a szoftver-rádió technológia segítségével lehetővé válhat az IP alatti rétegek különbségeinek áthidalása is. A 4G rendszerekben tehát alapkövetelménynek válhat a nagy flexibilitású szoftver-rádiós technológiák alkalmazása.

Az eddigi hálózatok mind kültéri mobil szolgáltatásra koncentráltak. A következő lépés a hibrid megvalósítások irányába kell, hogy történjen, melyek magukba foglalják a mobil hálózatokat éppúgy, mint a pillanatnyilag elérhető WLAN hálózatokat. Így nem csak a lefedettség nőne, hanem a kapcsolat sávszélessége és átviteli minősége is mindig a lehető legjobb lenne, azaz ténylegesen megvalósulna az *always best connected* elv. A hálózatok közötti vándorlásoknál nagyon fontos, hogy ezek gyorsan és a felhasználó számára észrevétlenül történjenek, hiszen például egy videó stream nézése közben nagyon zavaró lehet a lassú handover. Ehhez természetesen több rétegbeli átkapcsolásra lehet szükség.

A leendő negyedik generációs hálózatok egyik legfontosabb alapköve az *All IP*, pontosabban az *All IPv6* megvalósulása. Ez azt jelenti, hogy a teljes hálózat a gerinctől a terminálokig, illetve az ezen megvalósított szolgáltatások, beleértve a beszédet (VoIP) és tetszőleges más, akár szigorú QoS-t igénylőket, csomagkapcsoltan, IPv6 alapon működnek. Ennek a letisztult képnek a megvalósulása rengeteg előnnyel jár, mind a gyártók, mind a szolgáltatók, mind a végfelhasználók számára, továbbá elengedhetetlen a globális mobilitás megvalósításához, de természetesen ahhoz további egységesítésekre is szükség van.

Szükség van az IP protokoll megfelelő adaptációjára is. A Mobil IPv6 az IP protokoll egy elfogadott kiterjesztése, azonban még nem elég ahhoz, hogy világ méretű hálózatok kérdéseit megoldja, ehhez további fejlesztésekre illetve kutatásokra van szükség [9]. Még fontosabbá teszi a mobilitás kérdését az is, hogy a jövő generációs mobil hálózatok cellaméretei csökkennek, így a hálózat még érzékenyebb lesz a handover kezelésre [10].

5.1. Cross Layer Design

További kérdéseket vet fel az IP alkalmazása a mobil készülékekben. A klasszikus Internet egyik alapja a protokoll architektúrája. Ebben az egyes rétegek egy-egy jól meghatározott feladatot oldanak meg, az alattuk lévő réteg segítségével, szolgáltatást nyújtva ezáltal a felette lévő rétegeknek. A 4G hálózatokban előre látható új elvárások azonban megkövetelik, hogy az architektúra QoS-t és mobilitás kezelést is nyújtson. A QoS természetesen az egyre növekvő valósidejű alkalmazásokhoz elengedhetetlen, a mobilitás menedzsment pedig nemcsak az azonos hálózaton belüli vándorláshoz, hanem akár a különböző technológiájú hálózatok közötti vándorláshoz is szükséges. *A szigorú modularitás fenntartása, ahol ez egyes hálózati rétegek csak a szomszédosakkal kommunikálnak, erősen csökkentheti a hatékonyságát a QoS szolgáltatásoknak, illetve egyéb fontos szempontoknak, mint például az energiafelvétel, melyek így összességében negatív hatással lennének a 4G használatra.*

A megoldás az úgynevezett **Cross-Layer Design** [11] használata lehet, *mely a biztonság, a QoS és a mobilitás kérdéseiben egyaránt adaptálja az IP-t a mobil használatra.*

Biztonság

Jelenleg a titkosítás több különböző rétegben is megvalósítható és megvalósított, sőt ezek együttes használata egyre terjed. Ilyen például az SSH, SSL, PGP az alkalmazási rétegben, az IPSec a végpontok közötti titkosításhoz, a WEP (mely tervezési hiba miatt valójában nem alkalmas valódi titkosításra) egy adott hozzáférési hálózati forgalom titkosítására, továbbá a Bluetooth és az UMTS is definiál titkosítási eljárásokat. Nem csak főleges számítás- és energiaigényt jelenthet a többszörös titkosítás, de a késleltetést is számottevően növelheti. A rétegek közötti kommunikáció segítségével

vel megoldható, hogy pontosan egyszeres legyen a titkosítás, azaz egy adott szolgáltatás eléréséhez tartozó forgalom például csak az alkalmazási rétegben legyen titkosítva, mivel ekkor már fölösleges ugyanazt a hálózati rétegben is még egyszer titkosítani. Természetesen figyelembe kell venni az adott körülményeket, és mindig a megfelelő réteget kell választani.

QoS

A 4G hálózatokban az egyik legnagyobb kihívás a kis energiafogyasztás megoldása. Az energiafogyasztás az adási bithiba-arányon keresztül szorosan összefügg a QoS biztosításával. Ha túlságosan kis teljesítménnyel adunk, akkor megnő a hibaarány és a valószínű alkalmasok használhatatlanná válnak. Az energiafogyasztás csökkentésének egyik lehetőség az adóteljesítmény csökkentése, illetve a kivárás. A közegre figyelve, a csomagot akkor kell küldeni, amikor a legkisebb az interferencia, így kisebb energiával is elküldhető, de lehet hogy csak lassabban. A másik lehetőség a bithiba-arány (BER) növekedésének megengedése, így az adóteljesítmény szintén csökkenthető azonos bitsebesség mellett. Természetesen ezeket a technikákat körültekintően kell alkalmazni, és itt lesz fontos megint a rétegek közötti kommunikáció. A jelenleg elterjedt 802.11 alapú kártyák nagy része ma is képes olyan információk közlésére, mint a jelminőség, jel-erősség illetve zajerősség.

Mobilitás

Az IP hálózatok mobil rendszerekre történő kiterjesztése új problémákat vet fel, mivel az Internetet nem a mobilitásra tervezték eredetileg. A legszembetűnőbb példa a TCP esete cellaváltáskor: a régi hozzáférési ponttal kezd megszakadni a kapcsolat, a bithibák nőnek, a TCP algoritmus elkezd eldobni a csomagokat. A kapcsolat egy bizonyos ponton megszakad, létrejön a kapcsolat az új hozzáférési ponttal. Ekkor ismét jó minőségű kapcsolat áll rendelkezésre, azonban a TCP algoritmus ehhez csak lassan alkalmazkodik, így feleslegesen lassan kezdi el kihasználni a rendelkezésre álló sávszélességet. Itt is segíthet a Cross Layer Manager, mely a link layerben történt változás után értesíti a TCP réteget, és az ezzel a plusz információval sokkal jobban tud alkalmazkodni a megváltozott hálózati feltételekhez.

A skálázhatóság a 3G majd a 4G hálózatokban egyre fontosabb kérdéssé válik. A 4G egyik víziója, az *ubiquity*, azaz a „mindenütt jelenlét”, amely nagyszámú növekedést jelent, hiszen az egy felhasználó – egy mobil elv helyett az egy eszköz – egy végpont elven szinte minden elképzelhető szeretnének bevonni a hálózatba. Ez a 3G-ben még szóba sem jöhetett az örökölt 2G átviteli technológiák magas átviteli költségei miatt. A 4G percdíjak körülbelül a 3G percdíjainak tizede lesz, ennek azonban egyik feltétele az új, sokkal olcsóbb, nagy teljesítményű, könnyen felügyelhető hálózati architektúra telepítése.

6. Összefoglalás

A 4G hálózatoknak jelentős különbséget és kézzelfogható előnyöket kell nyújtaniuk a felhasználóknak a 3G hálózatokhoz képest. Nem szabad csak úgy tekinteni a 2G, 2.5G, 3G, 4G kifejezésekre, mint egyre nagyobb kapacitással rendelkező hálózatokra. A felhasználóknak az új, jobb minőségű szolgáltatások és ezek integrációja fontos, mivel ezekre épülhetnek a népszerű alkalmazások. Igyekeztünk megvilágítani, hogy a következő generáció legfontosabb feladata annak elérése, hogy telefonunkkal bárhol a világon kommunikálhassunk, bármikor bármilyen internetes szolgáltatást elérhessünk, – mindezt olcsón, gyorsan és kényelmesen. Röviden, a 4G egy olyan intelligens technológia-együttes kell hogy legyen, mely határok nélkül összeköti az egész világot.

Irodalom

- [1] Keiji Tachikawa: A Perspective on the Evolution of Mobile Communications, NTT DoCoMo, Inc., IEEE Communications Magazine, October 2003.
- [2] Ai-Chun Pang, Jyh-Cheng Chen, Yuan-Kai Chen, Prathima Agrawal: Mobility and Session Management – UMTS vs CDMA2000, IEEE Wireless Communications, August 2004
- [3] Hai Jiang, Weihua Zhuang: Quality-of-Service Provisioning in Future 4G CDMA Cellular Networks, University of Waterloo, IEEE Wireless Communications, April 2004.
- [4] M. Chiussi, D.A. Khotimsky, S. Krishnan: Mobility Management in 3G All-IP Networks, IEEE Communications Magazine, September 2002.
- [5] Harri Holma, Antti Toskala: WCDMA for UMTS – Radio Access for 3G Mobile Communications
- [6] Domenico Porcino, Walter Hirt: Ultra-Wideband Radio Technology – Potential and Challenges Ahead, IEEE Communications Magazine, July 2003.
- [7] Opening Up Networks with JAIN Parlay, Simon Beddus, Steve Davis, IEEE Communications Magazine, April 2000.
- [8] Anne Wiesler, Friedrich K. Jondral: A Software Radio for 2G and 3G Mobile Systems, IEEE Transactions on Vehicular Technology, vol. 51, No.4. July 2002.
- [9] Huszák Á., Kiefer T., Simon V., Tilk G. L., Dr. Imre Sándor, Szabó Sándor: Mobilitás kezelés az IP alapú hálózatokban, Híradástechnika 2003/4.
- [10] Vilmos Simon, Dr. Sándor Imre: A Domain Forming Algorithm for the Next Generation, IP Based Mobile Networks, SoftCom 2004, Croatia-Italy.
- [11] G. Carneiro, J. Ruela, M. Ricardo: Cross-Layer Design in 4G Wireless Terminals, INESC Porto, IEEE Wireless Communications, April 2004.

Videó folyam szelektív újraküldése IP hálózatokban

HUSZÁK ÁRPÁD, IMRE SÁNDOR

BME Híradástechnikai Tanszék, Mobil Távközlési és Informatikai Laboratórium
{huszak, imre}@hit.bme.hu

Kulcsszavak: szelektív újraküldés, DCCP, MPEG folyam

Az Interneten elérhető multimédiás alkalmazások egyre népszerűbbé válnak a felhasználók körében. Mobil környezetben történő robbanásszerű elterjedésüket azonban a korlátozott sávszélesség és a vezeték nélküli csatorna jellegzetesen magas hibavalószínűsége korlátozza, hiszen a gyakori hibák jelentősen rontják a multimédiás szolgáltatás minőségét. Munkánkban bemutatunk egy új algoritmust, amely a DCCP transzport protokoll alkalmazásával jelentősen javítja az audió/videó folyam minőségét. Az általunk kidolgozott szelektív újraküldés algoritmus kiemelten kezeli az MPEG típusú multimédiás folyamatok fontosabb kereteit, így csökkentve a bithibák továbbterjedését. A szolgáltatás minősége így nagymértékben javítható, amit analitikus vizsgálatokkal is alátámasztottunk.

1. Bevezetés

Az egyre inkább tért hódító multimédiás alkalmazások Interneten való elterjedését a rendelkezésre álló sávszélesség korlátozza. Az újonnan megjelenő hozzáférési hálózatok azonban képesek lesznek kielégíteni az új igényeket. Ilyen technológia például a nemrégiben megjelent WiMAX (IEEE 802.16) [1], amely akár ötven kilométer sugarú körben, 70 Mbit/s sebességen kínál vezeték nélküli internetes hozzáférést. Ennek okán mindenképpen nagymértékű növekedésre kell számítani az IP alapú audió-videó alkalmazások terén.

Ugyanez a növekedési tendencia figyelhető meg a mobil eszközök elterjedése terén is. Jelenleg előrehaladott fejlesztések folynak adat-szolgáltatások biztosítására vezeték nélküli hálózatokban, és a végső cél a mobil kommunikáció és az Internet teljes konvergenciája, a mobil Internet kifejlesztése. Az IP alapú hálózatot használva, a vezetékes hálózatokra írt jól bevált alkalmazások működhetnek a vezeték nélküli hálózatokon is.

A vezeték nélküli hálózatok tulajdonságai nagymértékben eltérnek a vezetékes hálózatokétól, hiszen a rádiós csatorna sokkal érzékenyebb a zavarokra, a környezeti hatásokra vagy az időjárásra. Egy mozgó állomás esetén pedig a cellaváltás (handover) miatti késleltetések, adatvesztés is jelentősen ronthatja a minőséget.

A feladat tehát egy változó tulajdonságú és jelentős hibavalószínűséggel rendelkező vezeték nélküli összeköttetésen garantált minőségű adatfolyam-átvitel, vagyis QoS (Quality of Service) biztosítása. Ez az adatfolyam lehet szigorúan érzékeny a hibákra, mint például FTP, WWW alkalmazásokhoz tartozó adatátvitel, de lehet a hibákra érzéketlenebb, mint a hang-, vagy képátvitel. Természetesen az utóbbiaknál is törekedni kell arra, hogy minimális legyen a hibaarány. A kisebb hibákat az emberi érzékszervek nem észlelik, ezért bizonyos mértékig megengedhető a hibázás, annak érdekében,

hogy a vevő oldalon folyamatos legyen feldolgozás, más szóval folyamatos mozgóképet illetve hangot érzékeljen a felhasználó.

A következő fejezetben megismerhetjük a jelentősebb szállítási rétegbeli protokollokat, amelyek alkalmazások lehetnek az audió-videó folyam továbbítására. Ezt követően pedig az MPEG formátum jellemzőit mutatjuk be. A 4. fejezetben az általunk kidolgozott algoritmust mutatunk be, mely egy új transzport protokoll (DCCP) alkalmazásával javítani képes az MPEG típusú multimédiás adatfolyam minőségén a szelektív újraküldési technika segítségével, majd pedig analitikus módszerekkel vizsgáljuk az algoritmus hatékonyságát. A 6. fejezetben levonjuk a következtetéseinket és ismertetjük a továbblépési lehetőségeket.

2. Transzport protokollok

A szállítási réteg két nagy protokollja a TCP (Transmission Control Protocol) [2] és az UDP (User Datagram Protocol) [3]. Ezek a protokollok a mai napig meghatározói a számítógépek valamint számítógép-hálózatok közötti adattovábbításnak, pedig immár 25 éves szabványok. Mindkét szabványt vezetékes hálózatra dolgozták ki, azonban a ma egyre szélesebb körben használt vezeték nélküli hálózatok jellemzői jelentősen különböznek vezetékes hálózatok adatátviteli tulajdonságaitól. Emiatt a régi protokollokat felül kell vizsgálni, és megfelelően módosítani illetve kiterjeszteni.

A mai Internet protokollokat olyan vezetékes összeköttetésekre dolgozták ki, melyeknek a jellemzőik a következők: nagy sávszélesség, kis késleltetés, kis hibavalószínűség.

Az elmúlt időszakban megjelent és egyre népszerűbb vezeték nélküli hálózatok átvitelére azonban pont az ellenkező tulajdonságok jellemzőek: kisebb sávszélesség, nagy késleltetés, nagy hibavalószínűség.

A két protokoll, a TCP és az UDP, valójában nagyon különböző célt szolgálnak. A TCP egy kapcsolatorientált és megbízható adatforgalmat biztosít a felhasználónak, míg az UDP megbízhatatlan. A TCP esetében a megbízhatóság azt jelenti, hogy az elküldött csomagok biztosan megérkeznek, de az esetleges újraküldések miatti késleltetésre nincs garancia, míg UDP esetén a küldő elküldi a csomagot és ezután a hálózaton múlik, hogy megérkezik-e. Újraküldésből adódó késleltetéssel ekkor nem kell számolni.

Ezen jellemzői miatt az UDP-t kizárólag olyan esetekben alkalmazzák, ahol a küldött adatok egy részének elvesztése nem okoz működési problémát, sőt, esetleg kívánatos is (például torlódás esetén), mint a műsorszórás; vagy ahol ezen hibák korrigálásáról egy magasabb szintű protokoll gondoskodik. Az 1. táblázat tartalmazza a különböző alkalmazásokhoz leggyakrabban használt transzport protokollokat.

1. táblázat Alkalmazások és protokolljaik

Alkalmazás	Alkalmazás réteg protokollja	Szállítási réteg
e-mail	SMTP	TCP
távoli hozzáférés	Telnet	TCP
Web	HTTP	TCP
file átvitel	FTP	TCP
távoli file server	NFS	UDP
Multimédia	egyedi	UDP
IP telefonía	egyedi	UDP
hálózat menedzsment	SNMP	UDP
útvonalválasztás - routing	RIP	UDP

A közelmúltban több új transzport protokollt fejlesztettek ki, melyek megpróbálják kiküszöbölni a régebbi protokollok hibáit. Ilyen az UDP módosított változata az UDP Lite (Lightweight User Datagram Protocol) [4] és a DCCP (Datagram Congestion Control Protocol) [5].

2.1. User Datagram Protocol (UDP)

Az IP protokoll csak két gép közötti adattovábbítást biztosítja. Nem teszi lehetővé az alkalmazások vagy a felhasználók azonosítását. Az UDP szállítási protokoll biztosítja, hogy egy gépen egyidejűleg futó több alkalmazói program egymástól függetlenül küldhessen és fogadhasson csomagokat.

Az UDP sokkal gyorsabb protokoll, mint a TCP protokoll, viszont nem megbízható adatátvitel szempontjából. Nem kapcsolat-orientált, nincs hibajavítás, nincs nyugtázás. Tulajdonképpen az IP szint által biztosított szolgáltatásokat nyújtja felfelé. Akkor szokták használni, ha az adatátvitel sebessége a legfontosabb, minden többi feladatot a felette elhelyezkedő réteg lát el. Tipikusan a DNS-ek (Domain Name Server), real-time alkalmazások, játékok szokták használni (egy játékban vagy real-time hang átvitele esetén, ha egy csomag rossz, akkor ott legfeljebb döccen egyet, de ez még mindig kisebb baj, mintha az adott pontnál megállna és onnantól elkezdené újra adni a csomagokat). A szegényesebb szolgáltatásból adódóan sokkal egyszerűbb az UDP felépítés.

Az UDP esetében is felmerül a mobilitásból adódó változó hibaarány, ami időnként a kapcsolat teljes megszakadásához vezet. A nyugtázásra ugyan nem kell

várni, de a változó csatornaminőség és a cellaváltások itt is komoly gondot okoznak.

A hibaarány növekedéséből adódó csomagvesztésre legegyszerűbb módon úgy lehet védekezni, hogy a hatékonyabb FEC (Forward Error Connection) hibavédekező kódolást alkalmazunk. Ezzel azonban növeljük az átvendő adatmennyiséget, nagyobb lesz az overhead, és a hálózati terhelés is növekszik.

2.2. Lightweight User Datagram Protocol (UDPLite)

Az UDP módosításával létrehoztak egy újabb protokollt, az UDP-Lite-t [RFC 3828] amely valójában az új valósídejű multimédiás szolgáltatások kiszolgálására jött létre. Az új protokoll csupán annyiban változott az eredeti UDP-hez képest, hogy egy új részleges ellenőrzőösszeget vezettek be. Amennyiben a csomagnak abban részében keletkezik bithiba, amelyet a részleges ellenőrzőösszeg lefed, akkor a vevő érzékeli a hibát és eldobja a csomagot, míg ha a hiba olyan helyen van, amit a részleges ellenőrzőösszeg nem fed le, akkor nem dobja el. Ebben az esetben az alkalmazásnak kell kezelnie a hibás csomagot. Ezzel az eljárással körülbelül 40%-kal csökken a csomagok eldobásának száma. Abban az esetben, ha az UDP ellenőrzőösszege az egész csomagra kiterjed, az UDP Lite működése megegyezik a hagyományos UDP működésével.

Az egyik legjelentősebb ok, amiért létrehozták az UDP Lite protokollt az, hogy az alkalmazások egy csoportja kezelni tudja a hibás csomagokat is. A felhasználó által megfigyelhető minőség jobb lesz, ha a hibás csomagok nem kerülnek eldobásra, hanem az alkalmazásig eljutnak. Több hang és videó codec is ehhez az alkalmazás csoportba tartozik: ITU-H.263, ITU-H.265, MPEG-4 videó codec (ISO-14496). Ezek a kódolók jobb minőséget nyújtanak hibás csomagok kezelésével, mint ha egyáltalán nincs is csomag.

A csomagoknak nem minden bájta egyformán fontos, ezért a fontosabb bájtokat szükséges jobban védeni. Az Unequal Forward Error Correction is ezt a célt szolgálja. Egy rádiós csatornán, ahol a hibavalószínűség már jelentős, olyan szállítási rétegbeli protokollra van szükség, amely alkalmazkodik a csatorna tulajdonságaihoz éppúgy, mint az alkalmazásokhoz.

A szállítási réteg hibaérzékelő mechanizmusának, tehát védenie kell az alapvető információkat, mint a fejléc, de szükség szerint az adatok egy részét vagy egészét is. Annak eldöntése, hogy mely adatok fontosak, és melyek nem, azt az alkalmazásoknak kell eldönteniük. Annak meghatározása, hogy mely adatokat kell lefedni ellenőrző összeggel, a küldő oldali alkalmazás feladata.

2.3. Datagram Congestion Control Protocol (DCCP)

A DCCP egy megbízhatatlan transzport protokoll, amely torlódásssabályozási algoritmus használatára, valamint sorrendhelyes csomagtovábbításra is alkalmas a TCP-hez hasonlóan. UDP esetén a torlódás elhárítása védekezést az alkalmazásoknak kellett megoldaniuk, míg DCCP esetén ez már a protokoll szerves része.

A DCCP-t úgy próbálták alakítani, hogy a TCP és az UDP előnyeit egy protokollként valósítsák meg. A DCCP fejlécében így ráismerhetünk az előbbi protollokból ismert fejlécmezőkre (1. ábra).

A fejléc hossza minimálisan 12 bájt, maximálisan pedig a 1024 bájtot is elérheti, ha az opcionális mezőket, és az egyes csomagtípusok esetén használt pótlólagos mezőket is használjuk. Általános esetben az ellenőrzőösszeg (Checksum), az összes adatot lefedi, de az UDP Lite-hoz hasonlóan, a DCCP is lehetővé teszi az adatok részleges lefedését ellenőrzőösszeggel. Az ábrán egy olyan esetet mutattunk be, ahol nem a teljes csomag volt lefedve az ellenőrzőösszeg által (a lefedett terület szürke). Természetesen az is lehetséges, hogy az adatok egyáltalán ne legyenek lefedve, csupán a fejléc. Ez lehetőséget ad arra, hogy azok az alkalmazások, amelyek képesek kezelni a sérült adatokat, hatékonyan működjenek. Arra azonban mindig ügyelni kell, hogy csupán a fejléct követő adatok fedhetők le.

Source port				Dest port	
Data offset		CCVal	CsCov	Checksum	
Res	Type	X	Reserved	Sequence Number (high bits)	
Sequence Number (low bits)					
Payload					

1. ábra A DCCP fejléce

A DCCP működését tekintve összetettebb, mint az UDP, de valójában egy nagyon hatékony szállítási rétegi protokoll. A különböző működési módoknak köszönhetően a felhasználói alkalmazásokhoz kiválóan alkalmazkodó protokoll. Igen fontos tulajdonsága a protokollnak, hogy a vevő oldal nyugtát küld az adó oldalra, így a csomagvesztésre fény derül, hiszen e nélkül a torlódásszabályozó algoritmus sem működne.

A kapcsolatorientált DCCP a kapcsolat felépítése során megbízható protokollként működik. A torlódásszabályozással kapcsolatos üzenetek szintén megbízható adatfolyamként kerül továbbításra. Jelenleg két torlódásszabályozó algoritmust specifikáltak: TCP-like Congestion Control [CCID 2] és TFRC (TCP-Friendly Rate Control) Congestion Control [CCID 3]. A torlódásszabályozási algoritmust a kapcsolat felépítés során kerül meghatározásra. Az adatfolyam során azonban mind a vevő, mind az adó oldal kezdeményezheti a torlódásszabályozási algoritmus megváltoztatását. Jelenleg ugyan csak két ilyen algoritmus van specifikálva, de a DCCP protokollt felkészítették esetleges újabb torlódáskezelési módszerek bevezetésére is. Az adatfolyam ugyan megbízhatatlan maradt, de az adó oldal értesül a vevő által fogadott csomagok helyes megérkezéséről. A vevő nyugtát küld az érkezett csomagokról. A torlódáskezelő algoritmus egyben azt is meghatározza, hogy milyen gyakran érkeznek nyugták. A TCP-like (CCID 2) esetben körülbelül két elküldött csomag után érkezik nyugta, míg a TCP-Friendly Rate Control (CCID 3) esetén körbefordulási időnként (Round Trip Time) küld egy nyugtát a vevő. A DCCP képes annak meghatározására is, hogy milyen okból történt csomagvesztés. Ez az opció fontos lehet a torlódásszabályozó algoritmus számára, hiszen abban az esetben, ha például bithiba keletkezik, vagy a vevőoldali buffer túlszordulása miatt kerül sor csomageldobásra, nincs szükség a torlódásszabályozó algoritmus beavatkozására.

A DCCP-t olyan alkalmazások számára fejlesztették ki, mint például a streaming médiaalkalmazások, amelyek ki tudják használni a DCCP beépített szabályozási módszereit. Annak érdekében, hogy a DCCP hatékonyan vegye fel a versenyt a gyors UDP-vel, a DCCP csomagok fejlécét próbálták a lehető legkisebbre méretezni. A protokoll bizonyos feladatok estén még így is túlságosan bonyolult, ezért kifejlesztettek egy egyszerűsített DCCP protokollt is, melynek neve DCCP Lite [6].

3. MPEG formátum

Az MPEG (Moving Picture Experts Group) egy széles körben alkalmazott szabvány a videó (kép és hang) digitális adatfolyamának tömörítésére, az ehhez kapcsolódó specifikációkat tartalmazza. A különböző MPEG formátumok kidolgozása és fejlesztése folyamatos, az MPEG formátumokról kiforrott, de nem lezárt szabványok rendelkeznek.

Az MPEG tömörítés a videó/audió adatfolyam mértékét csökkenti, rendkívül számításgényes, veszteséges tömörítő eljárással (meghatározott keretek között skálázhatóan), kezelhető mértékűre. Az MPEG videó képértékelési frekvenciája és az MPEG audió mintavételi frekvenciája szigorúan meghatározott és a különböző formátumok a különböző felhasználási területekhez igazodó adatátvitelt és minőséget célozzák meg:

- MPEG-1: CD-ROM alkalmazások, videó-CD, általában 1,5 Mbit/sec-ig;
- MPEG-2: DVD, broadcast videó, távközlés, többnyire 4-9 Mbit/sec;
- MPEG-3: HDTV alkalmazások 1920x1080 felbontásig, 20-40 Mbit/sec;
- MPEG-4: alacsony adatátviteli igényű alkalmazások 176x144 felbontásig, 4,8-64 kbit/s.

Az MPEG videó képcsoportokból (GoP – Group Of Pictures) épül fel (2. ábra), amelyekben három különböző képtípus előre meghatározott sorrendben követi egymást:

2. ábra GoP struktúra



I-képek: Intra frame coded – csak képkockán belül kódolt. Csak olyan információt használ fel, amely a képkockán belül megtalálható. Az egyes képkockákban belül JPEG tömörítést alkalmaz és tömörítési arányuk viszonylag csekély.

P-képek: a P (predicted) képkocka az őt megelőző „I” vagy „P” képen alapul, azokat használja referenciaként. Ezt nevezik „forward prediction”-nek. A P-keret a megelőző I- vagy P-keret képrészleteinek elmozdulását, illetve a képtartalmak közti különbséget rögzíti. A videó egymást követő képkockáiban az objektumok alakja általában nem változik. Az MPEG a P (és B) képekben az objektumok elmozdulását mozgásvektorokkal írja le, a mozgáskompensációs eljárás segítségével. Az eltérő alak- és színinformációk meghatározása a megelőző (I, P) kép megfelelő információira épül, a különbséget, változást kódolja. A P-képek tömörítési aránya nagyobb, mint az I-kereté. Mivel a P-képek P-képből is származtathatók, előfordulhatnak közöttük hibásak is, és további hibaforrásul szolgálhatnak a hibás keretet referenciaként használó képkockák kódolásakor.

B-képek: a B (bidirectional) kép a megelőző és rákövetkező „I” vagy „P” képkockákat is felhasználja referenciaként. A kódolás a mozgáskompensációs technika felhasználásával, a P-keretekhez hasonló módon, de két irányból történik. Ezt az eljárást „bidirectional prediction”, két irányból történő jóslásnak nevezik. A B-képeknek a tömörítési aránya a legnagyobb és nem szolgál referenciaként más képek kódolásakor, ezért nem terjeszt hibákat sem.

A GoP-n belül az I és P képtípusok számát (melyekből következik a B-k száma is) a kódoló határozza meg. Az I, P vagy B képek különböző aránya nagyban befolyásolja az elérhető képminőséget az adott adatátviteli sebesség mellett. A képcsoportban (GoP) az I képek frekvenciáját, másképpen fogalmazva a köztük lévő intervallumot jelölik N -nel. P illetve a megfelelő B képkockák frekvenciájának jele: M , a különféle tömörítők ezt nevezhetik *SubGoP*-nak is. Ha $N=1$, ekkor minden képkocka I-frame. Néhány példa a 2. táblázatban látható.

GoP	N	M
IBBPBPBPBI...	10	2
IPPPPPPPPPPPPI...	15	1
IBBPBPBPBPBPBI...	15	3
IPPPPI...	5	1
IBPBI...	4	2
IBBI...	3	3
IPPI...	3	1

2. táblázat Keretek gyakorisága

A következő fejezetekben a képcsoportok méretével fogunk dolgozni. A következő összefüggéssel (1) számolható a GoP átlagos mérete, az N és M valamint az I-, P-, B-keretek átlagos méreteinek felhasználásával:

$$size(GoP) = size(I) + \left(\frac{N}{M} - 1\right) \cdot size(P) + \left(N - \frac{N}{M}\right) \cdot size(B)$$

4. Szelektív újraküldési algoritmus

Napjaink mobil eszközeinek teljesítménye már lehetővé teszi nagyobb erőforrás-igényű multimédiás alkalmazások használatát. Azonban az általuk használt rádiós csatorna rendkívül változékony, nagy hibavalószínűséggel rendelkezik, amely komoly gondok elé állítja a tervezőket.

Az IP alapú csomagkapcsolt hálózatokban történő multimédiás adatátvitel esetén nem mindegyik szállítási rétegbeli protokoll felel meg. Rendkívül fontos ugyanis, hogy hiba esetén az audio-videó folyam ne álljon le, várva a hibás csomagok újraküldésére. A nagy hibaránnal rendelkező rádiós csatornán ugyanis ez élvezhetetlenné tenné a szolgáltatást.

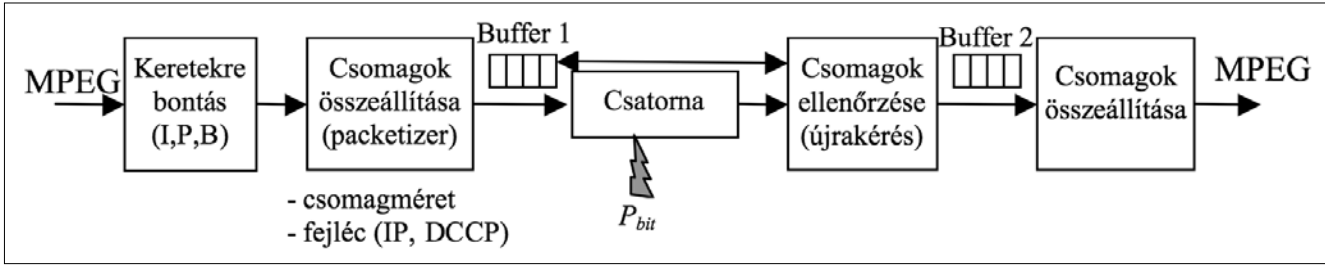
Ez az oka annak, hogy a TCP nem alkalmas erre a feladatra, hiszen ennek a protokollnak az alapkövetelménye, hogy minden egyes csomagot hibamentesen eljuttasson a vevőhöz, nem törődve azzal, hogy az mennyi ideig fog tartani. A TCP sorszámokkal látja el a csomagokat, így téve lehetővé az újraküldést. Az UDP esetében már jobb a helyzet, hiszen itt pont az ellenkezője történik. A protokoll nem törődik azzal, hogy a csomag helyesen érkezik-e meg a vevőhöz, sokkal inkább az a fontos, hogy folyamatos legyen az adatátvitel. Az UDPLite annyiban javít a helyzeten, hogy a sérült csomagok nem kerülnek eldobásra, hacsak nem a fejlécben történt a sérülés. Ezt a részleges ellenőrzőösszeg (partial checksum) használata teszi lehetővé. Az újraküldés sem az UDP, sem az UDPLite esetén nem lehetséges, hiszen a csomagokra nem lehet hivatkozni, ugyanis nincsenek ellátva a csomag azonosítási sorszámmal.

A DCCP fejlécben megtalálható a csomag sorszám mező, így a csomagokra való hivatkozás is lehetséges. A sorszámozás a vevő oldalon történő sorbarendezéshez nélkülözhetetlen, hiszen nem garantált, hogy a csomagok sorrendben érkeznek meg. Bár a DCCP nem írja elő a csomagok újraküldését, ugyanakkor értesül a csomagvesztésekről, a vevő által küldött nyugtákból. A szelektív újraküldési technika ezeket az előnyöket használja ki MPEG típusú multimédiás adatátvitel során.

Az alapötlet az volt, hogy az MPEG formátum I-kereteit kiemelten kezeljük, hiszen az ezt követő keretek ezekre épülnek egészen a következő I-keretig. Amennyiben ezeket a keretek hibátlanul továbbítanánk a vevő oldalra, jelentősen javíthatnánk az audio-videó minőséget. Algoritmusunk ezt oldja meg, vagyis amennyiben egy I-kép meghibásodik, újraküldjük azt.

Ezt a problémát leghatékonyabban a DCCP protokoll segítségével tudjuk megoldani, hiszen sorszámozott csomagokkal rendelkezik, és nem alapfeladata a hibás csomagok újraküldése, valamint alkalmas a hibás csomagok továbbítására is, az UDPLite-hoz hasonlóan. Ebből a szempontból a DCCP összesíti a többi transzport protokoll előnyét, ezért választottuk ezt a protokollt a probléma megoldásához.

Az általunk kidolgozott modell a 3. ábrán látható.



3. ábra A szelektív újraküldés modellje

A legelső modul az MPEG formátumú forrás adatait I-, P-, B-keretekre bontja, amelyeket a következő modul megadott méretű csomagokká állít össze. Itt történik a csomagok ellátása a fejlécekkel is (DCCP, IP), majd azok a csomagok, amelyek tartalmaznak I-kereteket, a *Buffer1* tárolóba kerülnek, hogy az esetleges újraküldés esetén innen lehessen kiolvasni azokat. A csomagokat a P_{bit} hibavalószínűségű csatornán továbbítjuk. A következő modul feladata a továbbított csomagok ellenőrzőösszegének vizsgálata és amennyiben hibát érzékel az I-keretet tartalmazó csomagban, újrakéri azt. A meghibásodás nélkül megérkezett csomagok a *Buffer2* tárolóba kerülnek. Az újraküldött csomagokat is ebbe a tárolóba illesztjük be, kihasználva a DCCP csomagok sorszámát. Mivel a sorszám egyértelműen azonosítja a csomagot, könnyedén tehetjük ezt meg. A következő modulban pedig összeállítjuk az MPEG folyamat a *Buffer2*-ből kiolvasott csomagok alapján. Ezáltal egy jobb minőségű audio-videó folyamatot kapunk, de cserébe körbefordulásnyi idővel (RTT – Round Trip Time) késleltetjük a multimédiás folyamatot. Természetesen semmilyen garancia nincs arra, hogy az újraküldött csomag hibátlanul érkezik meg, tehát e csomagok egy része ismét hibás lesz. A helyesen megérkezettek azonban javítják a médiafolyam minőségét.

Általános esetben egy csomag meghibásodásának valószínűségét a következő egyenlet adja meg, ahol P_{bit} a bithiba valószínűsége, amelyet a rádiós csatorna okoz, és az alsóbb rétegbeli protokollok nem tudnak javítani, CsM pedig a csomag mérete bitekben kifejezve:

$$P_{csomaghiba} = 1 - (1 - P_{bit})^{CsM} \quad (2)$$

A $P_{csomaghiba}$ UDP esetében a csomagvesztés, míg TCP esetében a csomag újraküldés valószínűségét jelenti.

Az algoritmusunk használata esetén, DCCP protokoll alkalmazásával és annak részleges ellenőrzőösszeg (partial checksum) lehetőségét kihasználva, szintén ez lesz a csomag meghibásodásának valószínűsége, azonban a csomag újraküldésének valószínűségét a következő képlet adja meg, ahol a *Fejléc* a csomag fejlécének méretét jelenti (DCCP adatátvitel során 12 vagy 16 bájt), míg az *Iframe* az I-kerethez tartozó információmennyiséget jelöli egy csomagon belül:

$$P_{újraküldés} = 1 - (1 - P_{bit})^{Fejléc + Iframe} \quad (3)$$

A képlet csak azokra a csomagokra vonatkozik, melyben van I-kerethez tartozó adat. Természetesen a

csomagon belül más kerethez tartozó adatok is lehetnek, azonban csak az I-kép adatait követően. A módszerünk alkalmazásához elengedhetetlen, hogy minden I-keret új csomagban kezdődjön, hiszen a DCCP protokoll részleges ellenőrzőösszege csak ebben az esetben használható.

Annak a feladatnak az ellátását, hogy a fejléct követően I-kép adata következzen, a 3. ábrán ismertetett csomag-összeállító modul a végzi. A részleges ellenőrzőösszeg (partial checksum) alkalmazása csak ebben az esetben teszi lehetővé, hogy a fejléct és az azt követő I-keretekhez tartozó információkat védjük, a többi adattal pedig ne foglalkozzunk. Ekkor az I-képet tartalmazó csomag újraküldési valószínűsége alulról és felülről is korlátozott:

$$1 - (1 - P_{bit})^{CsM} \geq P_{újraküldés} > 1 - (1 - P_{bit})^{Fejléc} \quad (4)$$

Ahol a CsM a csomagméretet jelöli, amelynek értéke a torlódáskezelő algoritmustól függ. A vizsgálataink során ennek értékét a számítások egyszerűsítése miatt állandónak tekintjük.

Az előbb leírtakból következik, hogy minél ritkábban vannak I-keretek, vagyis N (I-képek gyakorisága) minél nagyobb, annál kisebb valószínűséggel lesz szükség a csomag újraküldésére. Az N paraméter növelésével azonban a hang és kép minősége romlik, cserébe viszont nagyobb tömörítést érhetünk el.

A multimédiás folyam átvitele során a csomagok újraküldésére akkor kerül sor, ha a csomagban megsérül a fejléc vagy az I-kerethez tartozó adat, feltéve hogy I-keretet tartalmaz a csomag:

$$P_{újraküldés} = P_{hiba} \cdot P_{Iframe}, \quad (5)$$

ahol a P_{hiba} ,

$$P_{hiba} = 1 - (1 - P_{bit})^{\frac{CsM - Fejléc}{size(GoP)} \cdot size(I) + Fejléc}. \quad (6)$$

Itt azzal a feltételezéssel éltünk, hogy az egy csomagra jutó átlagos I-keret információ:

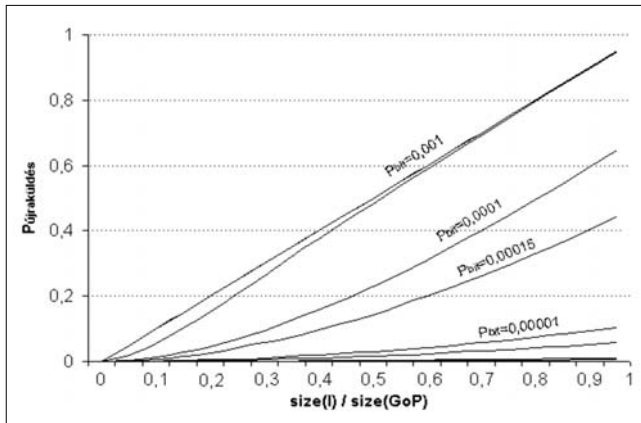
$$(CsM - Fejléc) \cdot \frac{size(I)}{size(GoP)}. \quad (7)$$

Annak valószínűsége,

hogy I-képet tartalmaz a csomag (8):

$$P_{Iframe} = \frac{size(I) + \left[\frac{size(I)}{CsM - Fejléc} \right] \cdot Fejléc}{size(GoP) + \left[\frac{size(GoP)}{CsM - Fejléc} \right] \cdot Fejléc} \approx \frac{size(I)}{size(GoP)}$$

A fejléc és a csomag adatmezőjének aránya kicsi, ezért élhetünk az egyszerűsítéssel.



4. ábra
Újraküldés valószínűsége az I-keret és a GoP méretének függvényében

Összevonva az egyenleteket, a következőt kapjuk (9):

$$P_{\text{újraküldés}} = \frac{\text{size}(I)}{\text{size}(GoP)} \cdot (1 - (1 - P_{\text{bit}})^{\frac{CsM - \text{Fejléc}}{\text{size}(GoP)} \cdot \text{size}(I) + \text{Fejléc}})$$

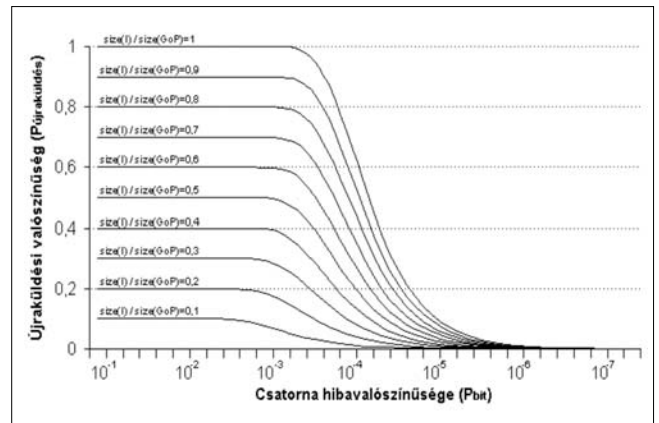
Az utolsó egyenlet felhasználásával kimutatható, hogy hogyan befolyásolja az I-kép és a teljes GoP méretének aránya az újraküldés valószínűségét, az algoritmusunk használata esetén (4. ábra). Az ábra egyértelműen megmutatja, hogy ahogyan növeljük az I-keretek méretét, vagyis a videó folyam minőségét növeljük, úgy növekszik az újraküldés valószínűsége. Szélsőséges esetben, amikor csak I-képből áll a GoP az újraküldés valószínűsége:

$$P_{\text{újraküldés}} = 1 - (1 - P_{\text{bit}})^{CsM} \quad (10)$$

A (9) képlet alapján rá tudunk mutatni a csomagújraküldés valószínűsége és a bithiba valószínűsége közötti összefüggésre, amelyet az 5. ábrán mutatunk be. A bithiba csökkenésével természetesen az újraküldés valószínűsége is csökken, hiszen így kisebb valószínűséggel történik hiba az általunk fontosnak tekintett adatokban, vagyis az I-keretet tartalmazó csomagok, I-képhez tartozó adataiban illetve ezek fejlécében.

Ezen összefüggések felhasználásával akár fix újraküldési valószínűséget biztosíthatunk változó bithiba valószínűség mellett is, oly módon, hogy a fizikai rétegben mérhető P_{bit} paraméter átadásával módosítjuk az MPEG kódoló beállításait (6. ábra).

Ez a kódoló minőségére lesz hatással, hiszen a $\frac{\text{size}(I)}{\text{size}(GoP)}$ arány az N paraméterrel egyenes arányosság-



5. ábra
Csomagújraküldés valószínűsége a bithiba valószínűségének függvényében

ban változik. Az N paraméter az I-keretek gyakoriságát adja meg, ahogyan azt a 3. fejezetben már ismertettük. Természetesen nem tudunk tetszőleges újraküldési valószínűséget biztosítani, mivel egy felső korlátot jelent, amikor csak I-keretekből áll a képcsoport, vagyis

$$\frac{\text{size}(I)}{\text{size}(GoP)} = 1, \quad (11)$$

így tehát minden P_{bit} mellett megadható egy maximális újraküldési valószínűség:

$$P_{\text{újraküldés}} \leq 1 - (1 - P_{\text{bit}})^{CsM}, \quad (12)$$

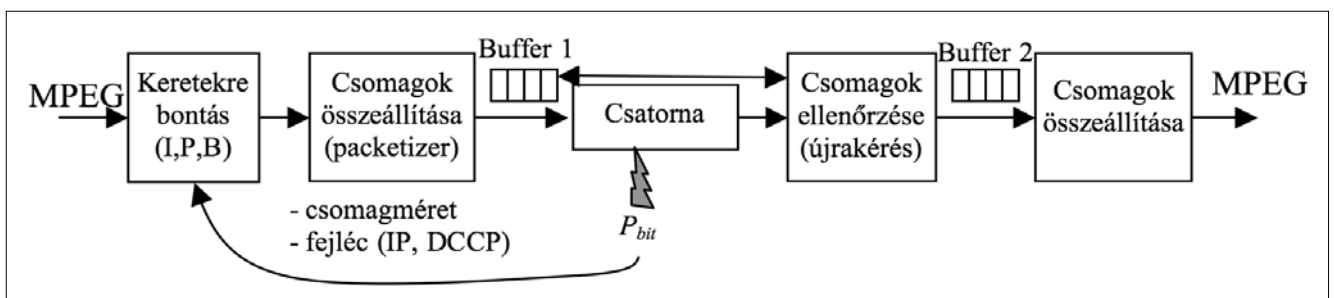
ami a 7. ábrából is jól látszik.

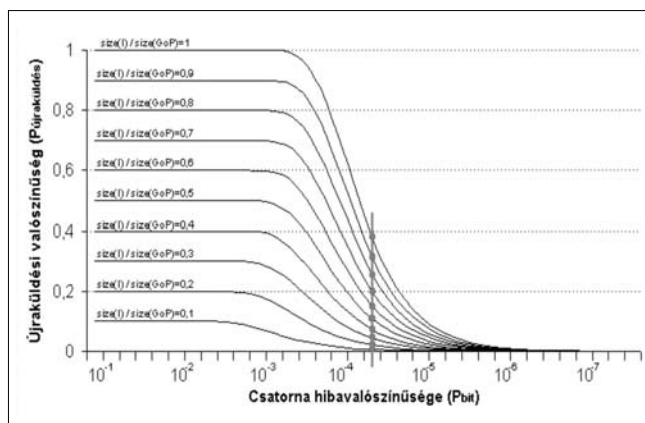
5. Analitikus vizsgálat

Ebben a fejezetben analitikus módszerekkel vizsgáljuk meg a szelektív csomagújraküldés hatékonyságát. Megvizsgáljuk, hogy egy körbefordulási idő (RTT) késleltetés bevetésével – amelyre az egyszeri újraküldés miatt van szükség –, hogyan változik a vevő oldali bithiba valószínűsége. Először a hagyományos UDP által továbbított multimédiás folyamatot vizsgálunk meg, majd pedig az UDP helyett UDP Lite-t vagy DCCP-t alkalmazunk. Legvégül pedig az általunk megtervezett algoritmust hasonlítjuk össze az előzőekkel.

Abban az esetben, ha az UDP protokollt használjuk, az ellenőrzőösszeg a teljes csomagot fedni fogja, emiatt egyetlen hiba a csomagban a teljes csomag eldobá-

6. ábra A P_{bit} átadása az MPEG kódolónak





7. ábra
A maximális újraküldési valószínűség

sát jelenti. Ebből egyértelműen látszik, hogy egy bithiba a csomagban, annyi bit elvesztését, vagyis sérülését okozza, amekkora a csomag mérete volt. Ez azt jelenti, hogy $P_{csomaghiba}$ valószínűséggel veszítünk csomagméretnyi bitet, és $1 - P_{csomaghiba}$ valószínűséggel nem sérül meg egy bit sem. Egy hosszú multimédiás folyam esetén a hibás bitek várható értéke csomagonként (13):

$$E(\text{hibás_bitszám})_{UDP} = P_{csomaghiba} \cdot CsM + (1 - P_{csomaghiba}) \cdot 0 \\ = (1 - (1 - P_{bit})^{CsM}) \cdot CsM$$

Abban az esetben, amikor UDP Lite, illetve DCCP protokollt alkalmazunk, azok részleges ellenőrzőszeg tulajdonságának kihasználásával, természetesen azt várjuk, hogy a hibás bitek számának várható értéke jelentősen kisebb legyen, hiszen a teljes csomag eldobására csak abban az esetben kerül sor, ha a fejléc megsérült. Az adatmezőben keletkező hibák egyszeres hibaként jelentkeznek. Figyelembe kell vennünk, hogy a fejléchihiba esetén már nem számít az adatmezőben keletkező hiba, hiszen ebben az esetben, az adatmezőben történő hiba nem növeli a hibás bitek számát, mert a csomagot teljes egészében hibásnak tekintjük. Így tehát a hibás bitek számának várható értéke a csomagban (14):

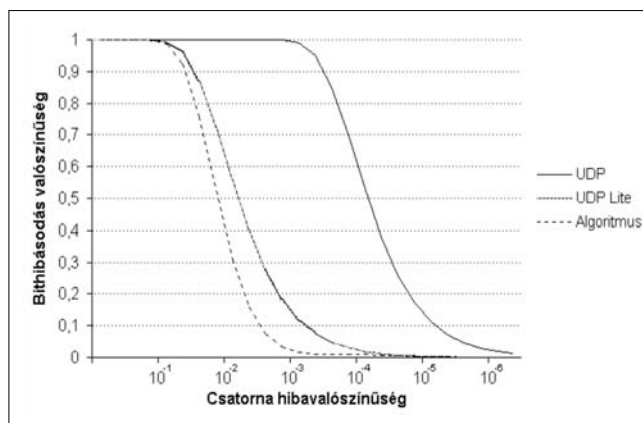
$$E(\text{hibás_bitszám})_{UDPLite} = P_{csomaghiba} \cdot CsM + P_{bit} \cdot (CsM - \text{Fejléc}) - \\ - P_{csomaghiba} \cdot P_{bit} \cdot (CsM - \text{Fejléc}),$$

ahol a $P_{csomaghiba}$
a korábbi egyenletekből már ismert, így

$$E(\text{hibás_bitszám})_{UDPLite} = (1 - (1 - P_{bit})^{Fejléc}) \cdot CsM + P_{bit} \cdot (CsM - \text{Fejléc}) - \\ - (1 - (1 - P_{bit})^{Fejléc}) \cdot P_{bit} \cdot (CsM - \text{Fejléc}) \quad (15)$$

Ezenfelül azt is tudjuk, hogy a csomag mérete (CsM) UDP esetén 8 bájt, míg DCCP esetén minimálisan 12 vagy 16 bájt, attól függően, hogy milyen hosszú csomag-sorszámokat használunk.

Vizsgáljuk meg, mekkora lesz a hibás bitek várható értéke a szelektív újraküldés alkalmazása esetén. Ez annyiban tér el az előző esettől, hogy a kiemelt jelentőségű helyen keletkezett bithiba a csomag újraküldését eredményezi. Természetesen az újraküldött csomag is sérülhet, illetve elveszhet a fejléc sérülése miatt.



8. ábra
Bithiba valószínűség a dekódoló bemenetén

A vizsgált algoritmusban csak egyszeri újraküldéssel számolunk, a többszörös újraküldéssel most nem foglalkozunk. A hibás bitek száma hasonlóan számolható, mint azt UDPLite esetében, vagyis az előző esetben számolt hibás bitek várható értékéhez, hozzá kell adni az újraküldés után továbbra is sérült bitek számát, és ki kell vonni azokon a bitek számát, amelyeket újraküldtünk (16):

$$E(\text{hibás_bitszám})_{Szelektív} = E(\text{hibás_bitszám}) + \\ + P_{újraküldés} \cdot E(\text{hibás_bitszám}) \cdot E(\text{hibás_bitszám}) - P_{újraküldés} \cdot E(\text{hibás_bitszám})$$

Abban az esetben, amikor a nincs I-képhez tartozó adat a csomagban, a képlet alapján kiadódik az UDPLite esetben számolt érték. Megfigyelhető az is, hogy minél nagyobb arányban szerepelnek I-keretek a csomagokban, annál kisebb lesz a hibás bitek várható értéke. Ez azzal magyarázható, hogy ekkor növekszik az újraküldés valószínűsége, és az újraküldött csomagoknak egy része helyes érkezik meg, melyek előzőleg mind hibásak voltak. A 8. ábra a dekódolóba kerülő bitek hiábaarányát mutatja, abban az esetben, amikor a teljes képcsoport méretének felét I-keret teszi ki:

$$\frac{\text{size}(I)}{\text{size}(GoP)} = 0,5 \quad (17)$$

Az ábra jól mutatja, hogy a legtöbb hiba az eredeti UDP esetén keletkezik, ami a teljes csomag eldobása miatt jelentkezik, egyetlen bithiba miatt. Az UDPLite sokkal jobban teljesít, hiszen csak a fejléc hibája okozza a teljes csomag elvesztését. A szelektív újraküldés alkalmazásával tovább csökkenthetjük a hibás bitek számát, kétszeres körbefordulási idő késleltetés árán. Ez a késleltetés azonban elfogadható, a vevő oldali videó minősége azonban jelentősen javul, hiszen nem csupán a hibás bitek számát csökkentettük, hanem az I-keretekben keletkezett hibákat is visszaszorítottuk, melyek a teljes képcsoport (GoP) minőségére hatással vannak. Egyetlen bithiba a kiemelt keretben végiggyűrűzik az összes kereten, egészen a következő I-képig. Ez tehát azt jelenti, hogy egyetlen bit helyreállítása ebben a keretben, N -szeres bithiba csökkenést jelent a képcsoportban, ami már nagyon jelentős minőségi javulást jelent.

Mivel az I-kereten belüli hibajavítás valószínűsége egy csomagra (18):

$$P_{javit} = P_{ujrakuldés} \cdot P_{hibasUDPLite}^2 - P_{ujrakuldés} \cdot P_{hibasUDPLite} = \\ = P_{ujrakuldés} \cdot P_{hibasUDPLite} (P_{hibasUDPLite} - 1),$$

így a teljes képcsoportra vetítve a kijavított bitek várható értéke:

$$E(jav_bit)_{GoP} = \frac{size(I)}{size(GoP)} \cdot P_{javit} \cdot N \quad (19)$$

A javított bitek okozta minőségjavulás számottevő, ami élvezhetőbbé teszi a multimédiás szolgáltatásokat a mobil eszközökben.

6. Összefoglalás

A szelektív újraküldés jelentősen emelheti a multimédiás szolgáltatások minőségét MPEG audió-videó átvitele esetén. Az algoritmus kifejezetten mobil környezetben hatékony, ahol a változó rádiós csatornán való adatátvitel csak magas hibaarány mellett valósítható meg. A MPEG formátum kiemelt I-kereteinek újraküldésével, jelentősen csökkenthető nemcsak e keretek bitjeinek hibaválósínúsége, hanem a teljes képcsoport hibaválósínúsége is, az I-képek hibája továbbgyűrűzik a többi kereten is.

Az analitikus vizsgálatot a szimulációs vizsgálatok fogják követni. Néhány DCCP implementáció már elérhető, ezért a közeljövőben egy valós teszhálózaton kívánjuk analízálni az algoritmust.

Ugyanakkor további kutatási területet jelent annak megvizsgálása, hogy az I-kereteken kívül a P-képeket is újraküldjük, hiszen e keretek hibái a B-képekben is megjelennek. Ezenkívül érdekesnek tűnik egy adaptív MPEG kódoló alkalmazása, annak érdekében, hogy az újraküldési valószínűséget állandó szinten tartsuk, változó rádiós csatornajellemzők mellett. Meg szeretnénk vizsgálni azt is, hogy mekkora késleltetés tudunk elfogadhatónak tekinteni, esetleges többszörös újraküldés esetén.

A szelektív újraküldés alkalmazása jó módszer a mobil hálózaton történő multimédiás forgalom továbbítására, amely a DCCP rendkívüli rugalmasságát használja ki az MPEG folyam minőségének jelentős növelésére.

Irodalom

- [1] <http://www.wimaxforum.org>
- [2] J. Postel: „Transmission Control Protocol”, RFC-793, September 1981.
- [3] J. Postel: „User Datagram Protocol”, RFC-768, August 1980.
- [4] Larzon, Degermark, Pink: „The Lightweight User Datagram Protocol”, RFC-3828, July 2004.
- [5] Kohler, Handley, Floyd: „Datagram Congestion Control Protocol”, draft-ietf-dccp-spec-11.txt, March 2005.
- [6] Phelan: „Datagram Congestion Control Protocol - Lite”, draft-phelan-dccp-lite-00.txt, August 2003.

Hírek

Nemzetközi siker adatbányászati versenyen

Az idén kilencedik alkalommal az ACM (Association of Computing Machinery – www.acm.org) KDD (Knowledge Discovery and Data Mining) csoportja által kiírt „KDD-Kupán” (www.kdnuggets.com/datasets/kddcup.html) a BME Távközlési és Médiainformatikai Tanszék Média Laborjának csapata – Tikk Domonkos, Kardkovács Zsolt Tivadar és Bánsághi Zoltán – két kategóriában (pontosság és kreativitás) is második helyezést ért el.

A verseny az adatbányászattal és gépi tanulással foglalkozó szakemberek legrangosabb megmérettetése, amelyen évről évre egyre több kutatócsoport vesz részt a világ minden tájáról. Az évek során a KDD kupákon számos nagy kihívást jelentő gyakorlati probléma volt a feladat, melyek megoldásai hozzájárultak a tudományterület növekvő sikereihez. Az idei feladat szintén egy nehéz gyakorlati probléma volt: 800.000 internetes kereső-kifejezést kellett automatikusan, tartalmuk szerint, 67 előre megadott kategóriába osztályozni. A nehézségi fokra jellemző, hogy a több, mint 140 regisztrált résztvevő közül csak 32 csapat adott be megoldást. További információk: http://www.tmit.bme.hu/about/awards/acm_kdd_05

Solaris-verseny egyetemek részére

A Sun a felsőoktatással foglalkozó informatikai konferencián, az EDUCAUSE-on bejelentette a Solaris University Challenge elnevezésű, egyetemisták részére meghirdetett versenyét. A kiírás értelmében a cég pénzzel és különböző díjakkal jutalmazza azokat a diákokat, akik a világ legfejlettebb operációs rendszerén, a Solaris 10-en, vagy a Sun nyílt forráskódú Solaris-közösségében, az OpenSolarisban fejlesztik projektjeiket. A Solaris 10 egyetemi versenyre a részt vevő országok akkreditált felsőoktatási intézményeinek hallgatói, végzős diákjai, valamint tanárai és informatikai munkatársai jelentkezhetnek.

A hazai érdeklődők a solaris_university@sun.com e-mail címen jelentkezhetnek konzultációra a versenyre vonatkozóan. A szakembek itt várják azoknak a kérdéseit, akik részt akarnak venni a nemzetközi programban.

A mobiltelefon elterjedésének regionális jellemzői Európában

ERDŐSI FERENC

Magyar Tudományos Akadémia, Regionális Kutatások Központja, Pécs
erdosi@rkk.hu

Kulcsszavak: NMT 450/900, GSM, kártyás mobil, penetráció, Európa, régiók, ellátási tényezők

A szerző e cikkben a mobiltelefonia európai elterjedésének időbeli folyamatával és térbeli mintáival foglalkozik a mindenkori technológiák (analóg, digitális) időszakaira vonatkozóan. A jellegzetességek csak komplex szemléletmóddal értelmezhetők, mert az egyes hatótényezők (a GDP, a természetföldrajzi adottságok, a népsűrűség, a lakosság kulturális színvonala, a távközlési politika) önmagukban nem definitívek. Számításba kell venni a nemzeti és más „soft” tulajdonságokat is.

Minden műszaki innováció élete a történelem során ciklikusan alakult. Korunk gazdaságának és társadalmának egyre inkább az infokommunikációs eszközök a bázisinnovációi. Beléptünk a posztindusztriális ötödik Kondratyev-ciklusba, annak a felfelé ívelő szakaszában vagyunk. Mint minden jelenségnek, a távközlésnek/infokommunikációnak is van egy időbeli dimenziója (eszközei, szolgáltatásai elterjedési folyamata), de térbeli jellegzetességei is izgalmasak. E cikk az internetnél az 1990-es években világviszonylatban 1,5-1,8-szor gyorsabban terjedő mobiltelefon földrészünkön végbement térbeli terjedésének sajátosságaival foglalkozik.

1. Az analóg cellás rendszerek lassú terjedése a mobiltelefonia „őkorában”

Az 1980-as években a még analóg elven működő és mérsékelt miniaturizált, első generációs (450 és 900

MHz-es) mobiltelefon-rendszerek eleinte nem hoztak átütő sikert, kivéve a ritkán lakott területekkel rendelkező Skandináviát, Angliát és Amerikát (1. táblázat). A vezetékes telefonhoz képest kétségtelen előnyeit csupán a belföldi távközlésben (NMT 450/900, TACS, Radiocom 2000 stb.) lehetett hasznosítani [1]. Az analóg rendszer különféle változatai az 1990-es évek elején lassan aztán a nyugat- és dél-európai országokban (Olaszország, Németország, Franciaország, Spanyolország stb.) is megjelentek.

Európa nyugati felének országai közül még a fejletlenebbek (például Spanyolország) is a volt szocialista országokhoz képest több éves (esetenként évtizedes) előnyre tettek szert az analóg mobilszolgáltatások bevezetésében.

1990-ben az Európai Közösség és az EFTA valamennyi országában működött már valamilyen analóg mobilszolgáltatás (Törökországban, Görögországban és

1. táblázat A cellás analóg mobiltelefon-rendszerek bevezetése Európában [2]

Év	NMT 450		NMT 900		TACS		Más rendszerek	
1981	Svédország	október						
	Norvégia	nov.						
1982	Dánia	január						
	Finnország	március						
	Spanyolország	június						
1983								
1984	Ausztria	nov.						
1985	Hollandia	január			Egyesült Kir. (x2)	január	Németo. (Netz-C)	szept.
	Luxemburg	június			Írország	dec.	Olaszo. (RTMS)	szept.
							Franciao. (R2000)	nov.
1986	Izland	július	Dánia	dec.				
			Finnország	dec.				
			Norvégia	dec.				
			Svédország	dec.				
1987	Belgium	április	Svájc	szept.				
1988			Ciprus	dec.				
1989	Feröer-szigetek	január	Hollandia	január			Portugália (Netz-C)	január
	Franciaország (SFR)	aug.						
1990	Andorra	július			Olaszország	április		
					Spanyolország	április		
					Ausztria	július		
					Málta	július		
1991	Magyarország							

Országcsoporthatár / ország	1990-ben			1995-ben			2000-ben			2002-ben			2004 aug.	Vezetékes kulmináció		A penetráció növekedésének mértéke (%)		
	vezeték fővonal	cellás mobil előfizetés	mobil / vezetékes, (%)	vezeték fővonal	cellás mobil előfizetés	mobil / vezetékes, (%)	vezeték fővonal	cellás mobil előfizetés	mobil / vezetékes, (%)	vezeték fővonal	cellás mobil előfizetés	mobil / vezetékes, (%)	cellás mobil előfizetés	éve	penetráció	fővonal 2002 / 1990	2002 / 1990	2004 / 1990
Nyugat-Európa													81,0					
Írország	29	0,7	2,4	36,3	4,4	12,1	48	68,0	141,7	50	76,3	152,6	87,8	→	—	72,1	10900,0	12542,9
Egyesült Királyság	43	1,1	2,6	50,2	9,8	19,5	59	72,2	122,4	59	84,1	142,5	91,3	2001	59,4	37,2	7545,5	8300,0
Hollandia	46	0,6	1,3	52,4	3,5	6,7	62	67,2	108,4	61	76,8	125,9	85,6	2001	62,1	32,6	12700,0	14266,7
Belgium	40	0,5	1,3	46,3	2,3	5,0	52	52,0	100,0	49	78,6	160,4	79,1	2000	51,7	22,5	15620,0	15820,0
Luxemburg	51	0,2	0,4	57,1	6,6	11,6	76	69,2	91,1	80	106,1	132,6	122,8	→	—	56,9	52950,0	61400,0
Franciaország	47	0,5	1,1	56,0	2,3	4,1	57	49,1	86,1	57	64,7	113,5	67,0	1998	58,3	21,3	12840,0	13400,0
Németország	40	0,4	1,0	51,3	4,6	9,0	61	30,6	50,2	65	72,8	112,0	77,6	→	—	62,5	18100,0	19400,0
Ausztria	41	0,8	2,0	48,0	4,8	10,0	47	77,0	163,8	48	83,6	174,2	92,2	2000	49,9	17,1	10350,0	11525,0
Svájc	58	2,0	3,4	63,7	6,4	10,1	73	64,3	88,1	74	78,9	106,6	86,3	→	—	27,6	3845,0	4315,0
Észak-Európa													96,3					
Izland	51	3,6	7,1	55,5	11,5	20,7	65	97,2	149,5	65	90,6	139,4	99,9	2000	68,3	27,5	2416,7	2775,0
Norvégia	51	4,2	8,2	56,7	22,5	39,7	74	75,1	101,5	73	84,4	115,6	84,1	→	—	43,1	1909,5	2002,4
Svédország	68	4,6	6,8	68,0	22,7	33,4	75	71,7	95,6	74	88,9	120,1	105,1	2000	75,8	8,8	1832,6	2284,8
Finnország	54	4,4	8,1	54,3	20,1	37,0	55	72,0	139,9	52	86,7	166,7	95,1	1997	55,6	-3,7	1870,5	21613,6
Dánia	57	2,1	3,7	61,1	15,7	25,7	72	63,1	87,6	69	83,3	120,7	92,0	2001	72,2	12,1	3866,3	4380,9
Dél-Európa													77,9					
Portugália	22	0,0	0,0	36,7	3,4	9,3	43	66,7	155,1	42	82,5	196,4	93,8	2000	43,0	90,9	•	•
Spanyolország	32	0,2	0,6	38,5	2,4	6,2	43	62,2	144,7	43	82,4	191,6	92,6	→	—	17,5	13633,3	463000
Olaszország	40	0,6	1,5	43,3	6,8	15,7	47	73,7	156,8	48	93,9	195,6	98,8	→	—	20,0	15650,0	16466,7
Málta	33	0,5	1,5	45,9	2,9	6,3	52	25,3	48,7	52	69,9	134,4	74,1	2001	53,0	57,6	13880,0	14820,0
Görögország	34	0,0	0,0	49,4	2,6	5,3	54	56,2	104,1	49	85,5	174,5	93,7	2000	53,6	44,1	•	•
Ciprus	44	0,7	1,6	55,2	7,1	12,9	65	32,2	49,5	69	58,4	84,6	87,8	→	—	56,8	8242,9	12542,9
Törökország	12	0,0	0,0	21,4	0,71	3,3	28	24,4	87,1	28	34,8	124,3	44,0	2001	28,1	133,3	•	•
Visegrádi és DNY-balkáni országok													69,8					
Lengyelország	9	0,0	0,0	14,8	0,19	1,3	28	17,5	62,5	30	36,0	120,0	50,7	→	—	233,3	—	—
Cseh Köztárs.	17	0,0	0,0	23,7	0,47	2,0	38	42,3	111,3	38	82,0	215,8	98,4	→	—	123,5	—	—
Szlovákia	15	—	—	20,8	0,23	1,1	31	20,6	66,4	27	54,7	202,6	72,8	2000	31,4	80,0	—	—
Magyarország	9	0,1	0,01	21,1	2,6	12,3	38	30,8	81,0	36	67,6	187,8	75,1	2000	38,0	300,0	—	—
Szlovénia	21	0,0	0,0	30,9	1,4	4,5	40	60,8	152,0	40	83,5	208,8	94,9	→	—	90,5	—	—
Horvátország	18	0,0	0,0	28,3	0,74	2,6	39	22,0	56,4	42	53,5	127,4	60,0	→	—	133,3	—	—
DK-Európa és Kelet-Balkán													36,8					
Bosznia-Herc.	14	—	—	—	—	—	10	2,4	24,0	24	19,6	81,7	32,4	→	—	71,4	—	—
Szerbia-Mont.	15	—	—	19,2	—	—	23	12,3	53,5	23	25,7	111,7	37,9	→	—	53,5	—	—
Macedónia	15	—	—	17,9	—	—	25	5,8	22,3	27	17,7	65,6	37,5	→	—	80,0	—	—
Bulgária	27	—	—	30,5	0,25	0,8	35	9,1	26,0	37	33,3	90,0	54,6	→	—	37,0	—	—
Románia	11	—	—	13,1	0,04	0,3	17	11,1	65,3	19	23,6	124,2	37,6	→	—	72,7	—	—
Albánia	1	—	—	1,3	—	—	4	0,8	20,0	7	27,6	39,4	31,4	→	—	600,0	—	—
Baltikum													66,7					
Észtország	21	0,1	0,5	27,7	2,05	7,4	36	38,7	107,5	35	65,0	185,7	80,1	2000	36,3	67,0	—	—
Lettország	23	—	—	27,9	0,59	2,1	30	16,7	55,6	30	39,4	131,3	60,5	2001	30,7	30,4	—	—
Litvánia	21	—	—	25,4	0,40	1,6	32	14,2	44,4	27	47,5	175,9	73,3	2000	32,2	28,6	—	—
Kelet-Európa													27,8					
Oroszország	14	—	—	16,9	0,06	0,4	22	2,2	10,0	24	12,0	50,0	33,5	→	—	71,4	—	—
Belarusz	15	—	—	19,3	—	—	27	0,4	1,5	30	4,7	15,7	15,2	→	—	100,0	—	—
Ukrajna	13	—	—	16,1	0,03	0,2	21	1,6	7,6	22	8,4	38,2	17,4	→	—	69,2	—	—
Moldávia	11	—	—	13,0	—	—	13	3,2	24,6	16	7,7	48,1	13,2	→	—	45,5	—	—

2. táblázat A vezetékes és mobil távbeszélő sűrűség (db/100 lakos) alakulása országcsopontonként
(Forrás: ITU évkönyvek, valamint a www.emc-database.com alapadataiból számított és összeállította a szerző)
Jelmagyarázat: → a vezetékes távbeszélő penetráció növekedése még 2002-ben sem kulminált

Portugáliában kísérleti stádiumban volt), de az előfizetők száma országonként még általában sehol nem haladta meg a tízezres-százezres nagyságrendet. Penetrációban kiemelkedően az észak-európai országok vezettek, élükön az NMT rendszerű készülékek – kezdetben monopolhelyzetet élvező – gyártója, az Ericsson által kedvezményezett Svédországgal (4,6%). A szűkebb értelemben vett Nyugat-Európán belül Svájc, Dél-Európában pedig Ciprus járt az élen a mobil elterjedésében.

A volt szocialista országok közül mindössze Észtországban kezdődött meg a mobilozás. A vezetékesekhez viszonyított mobil-előfizetői arány természetesen a mobilpenetrációhoz hasonlóan alakult, azzal a külön-

séggel, hogy Norvégia és Finnország megelőzte Svédországot (2. táblázat). Nem kétséges, hogy az analóg maroktelefon az általa kiszolgált térségek korlátozott-sága („államhatár tisztelete”) miatt csak egyfajta retardált diffúziós mintát képezett le.

A közel két évtizedes múlta visszatekintő analóg rendszerek elterjedettsége ma már marginális, de az európai országok felében még jelen van (3. táblázat). Megoszlásukhoz csak részben szolgál magyarázattal az a tulajdonságuk, hogy a ritkán lakott területek (Grönland, Izland, Svédország stb.) ellátására gazdaságosabbak, mivel celláik nagyok, tehát a tornyokat egymástól nagy távolságra lehet elhelyezni.

3. táblázat

A mobiltelefon műszaki jellemzőinek alakulása Európában

(Forrás: ITU és European Mobile Communications 2004 alapadatokból számította és összeállította a szerző)

Régió / ország	Az első mobil-szolgáltatás kezdete		A digitális szolgáltatások			A kártyás fizetés aránya	Mobil lefedettség a népesség százalékában
	év, hó		százalékos aránya				
	analóg	digitális	1995-ben	1999-ben	2002-ben	2004-ben	
Nyugat-Európa							
Írország	-	1993.06.	19,0	89,2	100,0	73,2	•
Egyesült Királyság	-	1991.02.	30,9	97,8	100,0	49,8	•
Hollandia	-	1994.07.	49,5	100,0	100,0	57,6	•
Belgium	-	1994.01.	80,4	100,0	100,0	63,9	99 (2002)
Luxemburg	-	1993.07.	99,4	100,0	100,0	31,2	98 (2002)
Franciaország	-	1992.07.	76,9	100,0	100,0	39,8	99 (2002)
Németország	-	1992.07.	84,0	99,8	100,0	49,8	99 (2002)
Ausztria	-	1992.02.	26,0	95,2	100,0	45,1	98 (2002)
Svájc	-	1993.03.	28,9	100,0	100,0	40,6	99 (2002)
Észak-Európa							
Izland	1986.07.	1994.08.	31,4	84,4	92,0	41,7	99 (2002)
Norvégia	1981.11.	1993.05.	50,3	92,1	99,2	44,9	97 (2002)
Svédország	1981.00.	1992.11.	50,2	94,4	98,6	58,4	•
Finnország	-	1992.01.	38,0	93,9	100,0	4,4	•
Dánia	-	1992.03.	62,8	95,9	100,0	26,5	83,3 (2002) 49,5 (1999)
Dél-Európa							
Portugália	-	1992.10.	94,4	100,0	100,0	79,9	•
Spanyolország	-	1995.07.	5,6	77,4	99,9	77,5	•
Olaszország	1990.05.	1995.04.	13,5	89,5	99,0	91,0	99,6 (2002)
Málta	-	1997.07.	—	32,3	100,0	92,9	99 (2002)
Görögország	-	1993.07.	100,0	100,0	100,0	71,5	99,6 (2002)
Ciprus	-	1995.04.	47,1	96,8	100,0	42,9	100 (2002)
Törökország	1986.10.	1994.03.	76,2	98,5	100,0	60,1	94 (2002)
Visegrádi és DNy-balkáni országok							
Lengyelország	1992.06.	1998.09.	—	96,7	99,9	55,2	97 (2002)
Cseh Köztársaság	1991.08.	1996.07.	—	96,5	99,6	74,0	•
Szlovákia	1991.08.	1886.01.	—	92,4	99,8	66,7	98 (2002)
Magyarország	1990.10.	1994.03.	73,6	94,0	100,0	75,7	96,2 (2002)
Szlovénia	1990.10.	1996.06.	—	93,3	97,7	47,5	99 (2002)
Horvátország	1991.11.	1996.03.	22,0	90,0	99,0	84,0	98 (2002)
Délkelet-Európa és Kelet-Balkán							
Bosznia-Hercegovina	-	1996.07.	—	—	—	74,5	—
Szerbia-Montenegró	1995.11. -	1996.10. 1996.07.	—	98,5	95,6	91,9 81,1	91,5 (2002)
Macedónia	-	1996.10.	—	98,0	100,0	82,6	92 (2002)
Bulgária	1993.12.	1995.09.	22,5	57,1	97,7	54,3	91,5 (2002)
Románia	1999.05.	1997.04.	—	99,0	100,0	61,7	98 (2002)
Albánia	-	1996.06.	—	100,0	100,0	97,2	90 (2002)
Baltikum							
Észtország	1991.01.	1993.09.	38,0	97,7	100,0	39,9	99 (2002)
Lettország	1991.10.	1995.01.	33,3	95,6	100,0	47,0	96,9 (2002)
Litvánia	1992.02.	1995.03.	46,6	94,6	100,0	42,2	100 (2002)
Kelet-Európa							
Oroszország	1991.12.	1994.07.	14,6	66,6	98,2	40,6	•
Belorusz	1993.04.	1999.04.	—	28,8	99,5	7,7	87 (2002)
Ukrajna	1993.12.	1996.12.	—	52,6	99,8	79,3	75 (2002)
Moldávia	-	1998.10.	—	94,4	100,0	83,9	77 (2002)

2. A digitális mobiltelefonía gyors terjedésének műszaki és politikai okai

Az „europaizálódási” folyamat, illetve a globalizáció következtében a távközlési igények egyre nagyobb hányada nemzetközi viszonylatú, amit az analóg rendszerek nem tudtak teljesíteni. Ezért már az 1980-as években elindult a nemzetközi szolgáltatásokra is alkalmas páneurópai mobilrendszer fejlesztési programja. Ennek eredményeként a második generációs digitális cellás GSM rendszer a külföldi kapcsolattartásra való alkalmasságával egyik technikai eszköze lehetett az egységes európai piac megteremtésének. A legtöbb európai országban 1991/92-ben megindult a GSM kereskedelmi szolgáltatása (3. táblázat). Bár a GSM tulajdonképpen európai használatra létrehozott rendszer (1987-ben 17 európai ország írta alá szabványáról az egyetértési emlékeztetőt), de ma a világ mobil-előfizetőinek túlnyomó többsége is ezt használja.

A GSM jól jött az éppen akkor egyesített és távközlési hálózatát rohamos gyorsasággal korszerűsítő Németországnak, a kontinens legnagyobb potenciális piacának. A volt NDK-nak nem volt analóg rendszerű kereskedelmi mobilszolgáltatása, ezért egyből a második generációs rendszerrel indult meg a mobilhálózat kiépítése. A digitális mobilszolgáltatás megindítására földrésünk legtöbb nagyrégiójában – Kelet-Európa kivételével – néhány éven belül sor került (lásd 3. táblázat). A digitális rendszer (alapvetően a GSM, egészen ritkán pedig más technológia) részaránya a mobiltelefonában 1995-re 5,6% (Spanyolország) és 99,4% (Luxemburg) között váltakozott és még csak Görögországban vált hegemónná (100%). 1999-re az országok egyötödében már kizárólag digitális szolgáltatás volt, míg a többiben a hányada 28,8% (Belorusz) és 99,0% (Románia) között változott.

2004 nyarán az egész Nyugat-Európában és a Baltikumban egyeduralkodóvá vált a digitális rendszer, továbbá túlnyomó lett Dél-Európában, de Észak- és Délkelet-Európa több országában is elérte a 100%-ot (lásd 2. táblázat).

3. A GSM elterjedésének términtái és regionális különbségei

Jóval gyorsabban terjedt el a GSM, mint az analóg rendszer. Általában 1-2 év elegendő volt ahhoz, hogy a sűrűbben lakott (>150 fő/km²) országokban az analógnál jóval kisebb cellából összetevődő hálózat 95-98%-os területi lefedettséget érjen el, míg a ritkábban lakott (<80 fő/km²) országokban ugyanilyen lefedettséghez 2-3 évre volt szükség. Az alacsony (<15 fő/km²) népsűrűségű országokban 3-4 év alatt értek el 40-60%-os GSM lefedettséget. Néhány ország szélsőségesen ritkán lakott részét gazdaságossági megfontolásból továbbra is a nagy cellákkal fedték le analóg rendszerekkel (például Lappföldet, Észak-Svédországot, Norvégia hegyvidéki területeinek többségét).

A GSM szolgáltatások diffúziójának területi mintája az alábbi volt:

- először a fővárosokban és egy-két nagyvárosban „pontoszerűen”;
- majd a közlekedési korridorokban, vasúti fővonalakon, de különösen az autópályák mellett „lineáris” megjelenésben;
- ezt követően a vidéki városokban és területi igazgatási központokban, valamint ezek környékén;
- végül a falusi térségekben „diszpergált” mintában teljesedett ki, vált az országok területét (a lakatlan hegységek kivételével) gyakorlatilag teljesen lefedő hálózattá.

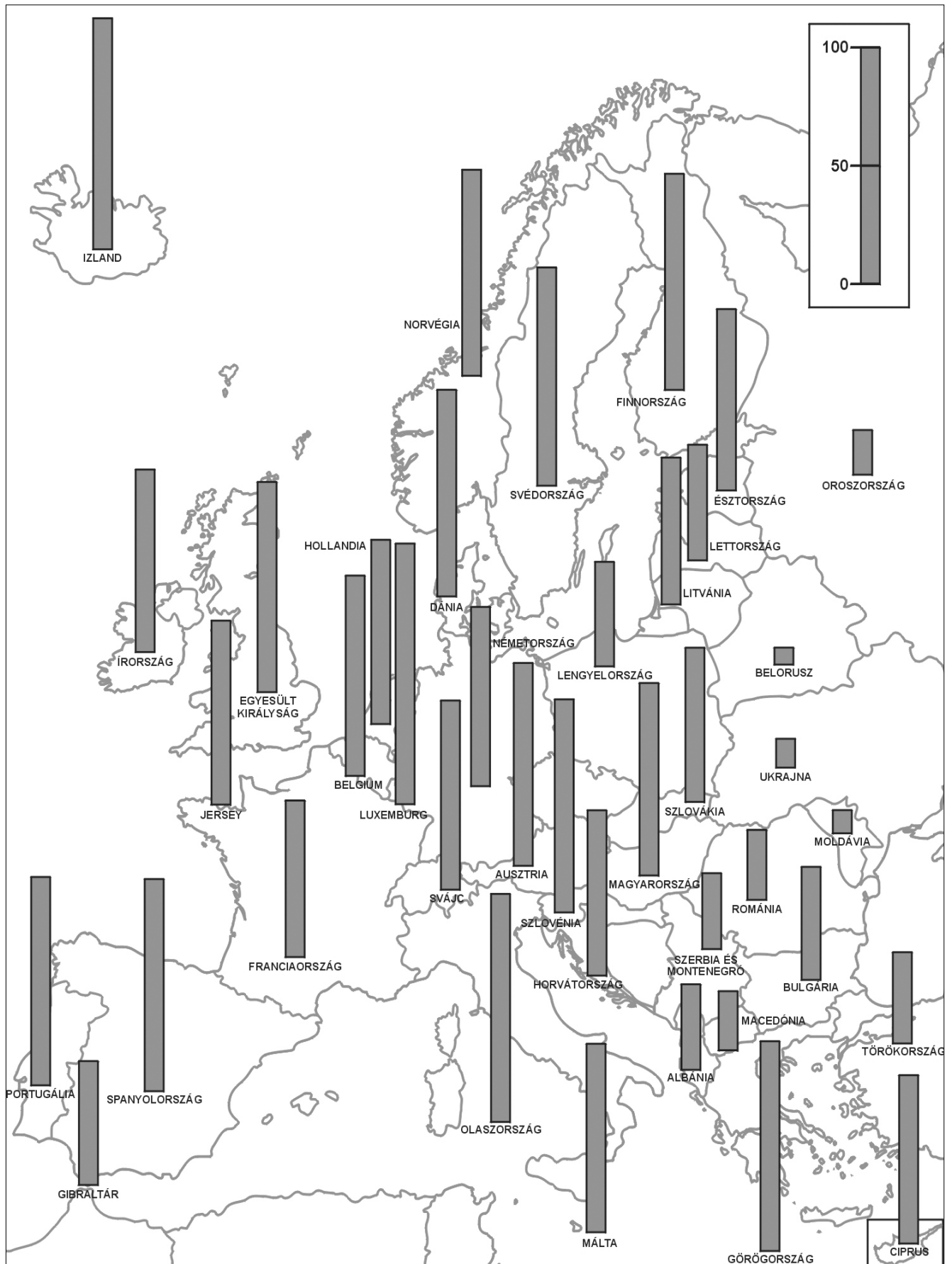
1995-re a 15 EU országban az összes mobil előfizetők száma tekintetében a különbség már „csak” 212-szeres volt az 1990. évi 1393-szorossal szemben. 1995-ben már az Egyesült Királyságé az első hely mind a mobil penetráció, mind a vezetékeshez mért százalékos arány tekintetében. Észak-Európa megőrizte erős fölényét még Nyugat-Európával szemben is, de a régión belül akkor a mobil penetrációban Izland lemaradt Skandináviától, míg a telefónián belüli mobilhányad tekintetében továbbra is Norvégia állt az élen. A nagyrégiók között (2004. évi adatok alapján) még mindig háromszoros a különbség (a 96,3-as penetrációjú Nyugat-Európa és a 27,8-as Kelet-Európa között). Az egyes régiókon belül is tekintélyes a szóródás az országok között (1. ábra).

Végül soron tehát Európa országai között több, mint kilencszeres (9,3-szeres) volt még 2004-ben is a különbség. A mobil penetráció terén az elmúlt közel másfél évtizedben végbement térbeli és intenzitási diffúzió eredményeképpen ugyan megállapítható az országok, nagyrégiók közötti közeledési tendencia, de a felzárkózás konkrét jegyei még ma is eléggé halványak. Európa országai közül 2005 januárjában Luxemburgon és Svédországon kívül már Izland, Olaszország és Csehország is bekerült az 100-as feletti penetrációjú országok előkelő körébe, amelybe a tengerentúli földrészek országai közül csak Hongkong, Izrael és Szingapúr tartozik.

Pontos adatok legkevesbé a mobilhálózatok által lefedett területeknek az adott ország területéből való százalékos arányáról állnak rendelkezésre. A MoU GSM Association értékelése alapján [4] alapvetően az országok alábbi típusai különíthetők el:

- A sűrűn és közepesen sűrűn lakott fejlett és közepesen fejlett országok, amelyekben az alig lakott hegységi területek kivételével szinte teljes lefedettségű a hálózat (például Németország, Svájc, Egyesült Királyság, Szlovénia, Csehország, Szlovákia, Románia);
- Az alacsony népsűrűségű, de magasan fejlett országok, ahol csak a nagyon ritkán lakott térségek maradtak a hálózaton kívül (Svédország, Norvégia, Finnország).
- Izland egyedi eset: magas fejlettsége ellenére csak a településeket hordozó, tengerpart közeli sávra korlátozódik a térerő.

1. ábra
A mobilpenetráció Európa országaiban 2004 nyarán
(Forrás: Az EMC World Cellular Database adataiból [3] szerkesztette a szerző)



- Kelet-Európa előbbre tartó országai, ahol a nagyvárosi agglomerációkon és közlekedési folyosókon kívül már egyes viszonylag sűrűn lakott kisvárosi-községi térségekre is kiterjedt a hálózat, de még sok kisebb apró „fehér folt” létezik a ritkábban lakott falusi térségekben.
- Kelet-Európa legkeletibb, legkevésbé fejlett és leg-ritkábban lakott térségei, ahol még mindig a pontszerű és lineáris elemek a meghatározók a mobilhálózatok területi térmentáiban.

A lefedettségnek a népességhányadban kifejezett másik mértékegységét elemezve kiderül, hogy 2002-ben a fejlett és sűrűn lakott országokban már a népesség 98-99,9%-ának volt módja megfelelő tértérít biztosító hálózat által a mobilszolgáltatásokra.

Érdekes kivétel Dánia a viszonylag alacsony (83,32 százalékos) értékével, mivel szigetei közel fekszenek egymáshoz és a távoli Grönlandot a statisztika nem foglalja be Dánia értékeibe. Lényegesen elmarad azonban Kelet-Európa, ahol (Oroszország általunk becsült

adatait is beszámítva) a népességnek 2002-ben csak körülbelül a 2/3-át, legújabbban talán már a 3/4-ét éri el a mobilhálózat.

4. A mobil és fix kapcsolatok versenye

Az 1980-90-es évek fordulóján az analóg mobilkészülékek még viszonylag nagyméretűek, és a fejlett országokban alapvetően az üzleti életet szolgálták Izland és Skandinávia északi, ritkán lakott és távbeszélővonalakkal gyengén behálózott régiói kivételével, ahol a lakosok vezetékes telefon helyettesítésére alkalmazták.

Európa – távközlésben még inkább – elmaradottabb keleti felében kétségtelen, hogy néhány évig a mobil szolgáltatásokat sokan vették igénybe azok közül, akik addig nem jutottak hozzá vezetékeshez (a várakozási idő még ebben az időszakban is a vidéki térségekben akár egy-két évet is kitett), de ahhoz elég pénzük volt, hogy a szinte hónapokon belül megjelentő mobil szolgáltatások anyagi feltételeit viselni tudják.

4. táblázat

Európa országai mobil penetrációja (2004), fajlagos GDP-je, továbbá a penetrációval esetleg összefüggésbe hozható egyéb gazdasági és kulturális mutatók a 2000-es évek elején (Forrás: [3],[8],[9])

a) nappali tagozaton felsőfokú oktatásban részesülők aránya,
b) százezer lakosra jutó, egy évben kiadott könyvek száma

Ország	Mobil / 100 lakos	GDP / lakos USD	Személygk. / 1000 lakos	Külker. forgalom / 1 lakos USD	Aktivitási ráta %	Szolgáltatásokban foglalkoztatottak, %	Munka- nélküliség %	Összes fogyasztásból az élelmiszer- fogyasztás, %	a) Felsőfokú oktatás %	b) Könyv- kiadás %	Napilap / 1000 lakos	Idegen- forgalom a GNP %-ában
Albánia	31,4	1.332	180	310	•	•	•	•	•	•	•	•
Ausztria	92,2	25.393	517	8.720	58,9	65,8	4,7	12,5	57	100	297	5,2
Belarusz	15,2	1.438	135	280	•	•	•	•	•	60	152	0,7
Belgium	79,1	23.681	460	17.419	51,4	73,4	7,7	14,7	58	•	161	2,9
Bosznia-Hercegovina	32,2	•	•	•	•	•	•	•	•	•	•	•
Bulgária	54,6	1.992	242	923	49,2	57,1	13,7	37,1	40	61	116	9,5
Ciprus	87,8	14.194	450	8.104	•	•	•	•	•	•	125	20,2
Csehország	98,4	6.852	344	3.559	59,4	•	7,8	21,3	34	122	254	5,7
Dánia	92,0	32.033	352	8.148	65,5	73,6	5,4	13,1	59	272	283	2,4
Egyesült Királyság	91,3	26.369	487	5.371	62,5	75,3	•	9,9	59	187	331	1,1
Észtország	80,1	4.732	414	5.430	58,4	61,3	10,0	27,3	60	237	174	9,6
Finnország	95,1	25.314	418	6.188	62,8	68,2	10,5	13,1	85	255	445	1,2
Franciaország	66,4	24.057	489	5.050	56,1	71,2	9,0	15,0	54	67	218	2,2
Görögország	93,7	12.084	325	2.987	48,6	61,7	8,9	21,3	61	39	130	3,3
Hollandia	85,6	25.866	419	12.152	64,7	77,7	3,6	11,8	55	95	300	1,7
Horvátország	60,0	5.125	294	•	•	53,5	14,3	•	•	167	115	16,4
Írország	87,8	31.041	382	13.360	59,6	65,9	4,5	15,0	47	•	149	4,0
Izland	99,9	26.617	550	•	•	69,7	3,3	•	•	641	•	•
Lengyelország	50,7	4.902	272	1.304	54,7	53,0	19,6	29,5	58	56	102	2,9
Litvánia	60,5	3.597	326	1.485	62,0	59,2	10,6	31,2	64	91	135	1,6
Litvánia	73,3	3.977	250	1.320	58,0	54,0	12,4	31,9	•	116	29	3,4
Luxemburg	122,8	47.255	620	22.653	54,8	78,2	3,7	•	•	•	328	•
Macedónia	37,5	1.705	250	520	•	•	•	•	•	36	53	0,7
Málta	74,1	9.839	430	6.800	•	•	•	•	25	•	•	17,3
Magyarország	75,1	6.486	269	3.400	49,8	61,2	5,9	30,5	44	110	192	8,0
Moldávia	13,2	337	110	240	•	•	•	•	•	27	60	•
Németország	77,6	24.122	539	5.902	57,3	66,2	9,8	12,1	48	95	305	0,9
Norvégia	84,1	42.149	421	7.135	72,8	74,9	4,5	15,6	70	112	589	1,3
Olaszország	99,8	21.024	578	4046	49,3	63,5	8,9	15,2	50	56	104	2,3
Oroszország	33,3	2.370	61	290	60,0	58,8	8,7	40,5	•	25	105	1,5
Portugália	93,8	11.800	362	3.783	62,1	54,4	6,1	23,1	50	22	75	5,0
Románia	37,6	2.107	144	694	54,8	34,5	7,0	35,9	27	35	120	0,9
Spanyolország	92,6	16.091	461	3.815	54,1	63,6	11,1	17,1	59	149	100	5,6
Svájc	86,3	36.738	510	10.662	67,7	73,4	4,1	15,9	•	256	373	2,7
Svédország	105,1	26.864	454	7.188	62,8	•	5,6	12,8	76	142	410	1,8
Szerbia-Montenegró	42,0	1.451	230	470	•	•	•	•	•	51	106	•
Szlovénia	94,9	11.020	443	5.097	56,6	•	•	20,4	61	174	169	5,1
Szlovákia	72,8	4.404	240	2.870	60,3	55,9	17,4	23,5	•	58	131	3,2
Törökország	44,0	2.722	71	603	48,7	43,0	10,5	29,1	25	4	110	5,3
Ukrajna	17,4	827	105	236	62,6	51,2	9,1	•	•	13	175	7,7

A GSM rendszer elterjedésével, a készülékek és a szolgáltatások együttes árának viszonylagos csökkenésével a mobil a rendszertulajdonságából adódó különleges használati értéke [5,6] folytán az 1990-es évek végétől az emberiség technikai történetében páratlanul gyorsan elterjedt, népszerű műszaki eszköz lett. Sőt, a mind „többet tudó” készülékek megjelenésével az „erkölcsi kopása” két-három év alatt olyan mértékűvé válik, hogy a fogyasztók napi életében már kikerült a „tartós fogyasztási cikkek” kategóriájából.

Bármennyire is viharos gyorsaságú volt a mobil diadalútja, ez mégis csak néhány országban ment végbe a vezetékes kárára, azaz készítette a fogyasztókat arra, hogy lemondják a vezetékes telefonjukat. Európa országai annak alapján, hogy a mobil terjedése befolyásolta-e és milyen mértékben, illetve mikortól a vezetékes ellátottságot, az alábbiak szerint csoportosíthatók:

- A mobilpenetráció 2000-ben még csak 15 országban (11 fejlettsen és 4 közepesen fejlettsen) haladta meg (vagy érte el) a vezetékest, de közülük 2000 előtt a vezetékes állomány elterjedése csak Franciaországban és Finnországban érte el kulminációs pontját. 2000-ben a vezetékes ellátottság még néhány magasan fejlett országban is megelőzte a mobilt; mint Luxemburg, Franciaország, Németország, Svájc, Svédország, Dánia, nem beszélve a gyengén fejlett országról (lásd 2. táblázat).

- 2002-ben már a nyugat-, észak- és dél-európai országokban (az egyetlen Ciprus kivételével) továbbá valamennyi visegrádi/nyugat-balkáni, valamint balti országban felülmúlta a mobilellátottság a vezetékest. Különlegesen nagy mobilfőlény jön létre Albániában, Európa legelmaradottabb országában (3,94-szeres), mégpedig alapvetően helyettesítő szerepet ellátva. Igen erős (1,8-2,2-szeres) mobilfőlényre tett szert Csehország, Szlovénia, Szlovákia, Ausztria, Svédország, Portugália, Magyarország, Észtország. Többségük hegyvidéki terület, ahol a cellás rendszer műszakilag általában könnyebben kiépíthető, mint a vezetékes és fenntartásuk, karbantartásuk is jóval olcsóbb. E természeti adottság azonban nem meghatározó, ahogyan a gazdasági fejlettség sem. 2004 óta már csak Kelet-Európában folytatódik érzékelhetően a vezetékes penetráció növekedése, annak ellenére, hogy ezekben az országokban az európai átlagnál még mindig jóval gyorsabban terjed a mobil (lásd 2. táblázat).

Egy spanyol szerzőpáros [7] szerint több jelentős nemzeti piac (Norvégia, Svédország, Finnország, Spanyolország) már 2001-ben közel került a telítettséghez. Kizárólag a 2G-re/2,5G-re vonatkoztatva e megállapítás elfogadható, azonban hogy hol van a penetráció határa, azt nehéz pontosan meghatározni. Formálisan a 100%-nál. A valóságban azonban több „mobilmániás” országban egy emberre már egynél több zsebtelefon jut – például azért, mivel a munkahelye is ellátja. A piac telítettségéről csak viszonylagos értelemben, egy meghatározott műszaki rendszerre, illetve annak szolgáltatásaira vonatkozóan lehet beszélni. Az egymást követő rendszerek (pl. 3G) újabb piacokat nyitnak.

Az European Mobile Communications Report különböző időpontokra vonatkozó adatait összehasonlítva kiderül, hogy *pár százalékkal csökkent a penetráció 2004-ben* Bulgáriában, Horvátországban, Franciaországban, Görögországban, Izlandon, Írországon, Lettországon, Portugáliában, Spanyolországban és az Egyesült Királyságban. Mivel a lakosság számának gyors növekedését a számításba jöhető hatótényezők közül törölnünk kell, vagy tényleg megkezdődött valamilyen speciális ok miatt az előfizetők számának csökkenése, vagy a statisztika adatelőállítás rendszerében volt néhány országban változás. De ez sem valószínű.

A mobil-fix verseny időbeli és térbeli folyamatának alakulását, a jelenlegi penetrációt és a forgalom mértékét a természeti viszonyok alig, a település-állomány és -szerkezet, valamint a képzettségi, műveltségi szint jobban, a jövedelmi viszonyok pedig még erősebben befolyásolják. (A 2004. évi penetráció és a GDP közötti korreláció: $r=0,87$.) Elvileg feltételezhető, hogy a távközlési szolgáltatások (így a mobil is) intenzitása a fogyasztói kosár részeként számos „puha” – kulturális/társadalmi/etikai/vallási/életmódbeli – tényező által befolyásoltan alakulnak (4. táblázat) még akkor is, ha egyenként vizsgálva, a korreláció alacsony értékeiben ($r=0,38-0,67$) ezek az összefüggések nem tükröződnek.

Számos országban a mobilelőfizetők gyakoriságát és különösen a forgalmat a jövedelmi viszonyoknál és minden más tényezőnél erősebben befolyásolja a távközlés-politika. A mobilszektorban a liberalizációs/deregulációs intézkedések nyomán a magántulajdon a meghatározó és a tényleges verseny feltételei a legtöbb országban teljesültek azzal, hogy több szolgáltató van jelen a piacon. 2004-ben inkább csak Kelet-Európa egyes országában igyekeztek megakadályozni a politikai erők a versenyt azzal, hogy kevés működési engedélyt adtak ki a nagyobb hányadban állami tulajdonú szolgáltatóknak.

5. A kártyás fizetési mód területi különbségei

A mobilszolgáltatások kártyás előfizetői (prepaid) módja általában az 1990-es évek második felében kezdett elterjedni.

2004-ben a kártyás fizetés százalékos aránya (3. táblázat) szerint az országok az alábbiak szerint kategorizálhatók:

- kiemelkedően magas (80% feletti) arányt ért el Albániában, Szerbiában, Montenegróban, Olaszországban, Máltán, Horvátországban, Macedóniában, Moldáviában;
- igen magas (70-80% közötti) arány jellemzi Írországot, Portugáliát, Spanyolországot, Görögországot, Csehországot, Magyarországot, Bosznia-Hercegovinát, Ukrajnát;
- magas (60-70% közötti) a kártyások aránya Belgiumban, Törökországban, Szlovákiában, Romániában;

- közepes (50-60% közötti)
Hollandiában, Svédországban, Lengyelországban és Bulgáriában;
- alacsony (40-50% közötti)
az Egyesült Királyságban, Németországban, Ausztriában, Svájcban, Izlandon, Norvégiában, Cipruson, Szlovákiában, Lettországon, Litvániában és Oroszországban;
- igen alacsony (20-40% közötti)
Luxemburgban, Franciaországban, Dániában és Észtországban;
- szélsőségesen alacsony
Finnországban és Beloruszban (lásd 3. táblázat).

Az egyes régiókon belül is erősen szóródnak az értékek, csupán a Baltikum tűnik ki viszonylagos kiegyensúlyozottságával. Nyugaton leginkább Máltán (93,8%), legkevésbé Dániában (31,5%) kedvelt e díjfizetési mód, míg a tágabb értelemben vett Kelet-Európán belül e kétféle szélsőség (a Baltikumtól eltekintve) Lengyelországra (33,9%) és Szerbia-Montenegróra (98,7%) jellemző.

E hányadok mögött semmi objektív (az ország karakteréhez köthető) rendezőelv nem fedezhető fel. Sem az életszínvonalra, a távközlés fejlettségére, sem a népességre jellemző életvitelbeli/gazdálkodási szokásokra nem vezethetők vissza az arányok. *A szolgáltató társaságok üzletpolitikája, azaz a szolgáltatási árstruktúra lehet az első számú hatótényező.* Végso soron a kártás fizetés elterjedtségét a belőle származó fogyasztói haszon, vagy más (például kényelmi természetű) előnyöknek a mértéke alakítja.

5. Összefoglaló

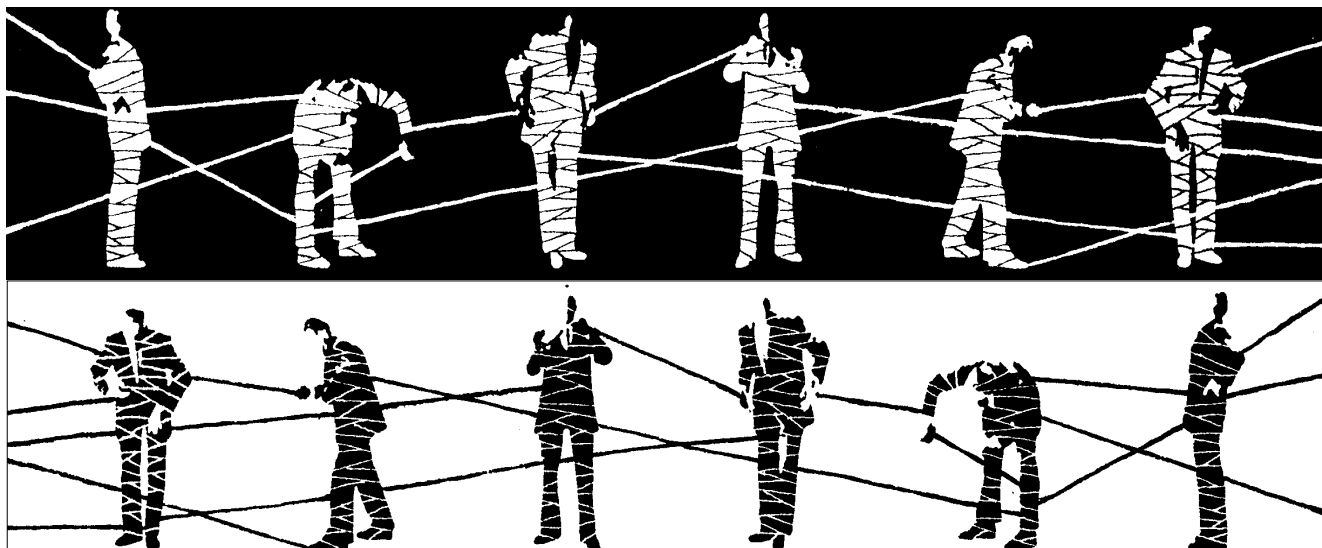
A mobil előfizetők száma korunkban még a képzetebb népességű Európában is meghaladja a világhálót használók (de különösen a hostok) számát, míg az elmáradott és alacsony képzettségű Harmadik Világ egyes régióiban a „netezőknél” sokszorosan többen „mobiloznak”.

A mobil szolgáltatások diffúzióját egy adott országban a nemzeti műszaki-innovációs háttér, a természeti adottságok és a gazdasági teljesítmény (GDP) mellett egyre inkább számos „puha tényező” is befolyásolja. Úgy gondoljuk, hogy a terjedés mikéntjét csak valamennyi tényező számbavételével lehet értelmezni.

Megállapításaink természetesen csak a legutóbbi időig markánsan szinguláris távközlési technológiára érvényesek. A mobiltelefon és az Internet/műsorvétel/fotószolgáltatás stb. már megvalósulóban levő konvergenciája olyan integrált technológiák létrejöttéhez vezet, melyek idő- és térbeli diffúziós mintái már nem elhanyagolható mértékben eltérnek a közelmúltban regisztráltaktól.

Irodalom

- [1] Standards in wireless telephone networks. (2004) Telecommunications Policy, 27. pp.325–332.
- [2] Garrard, G.A. (1998): Cellular Communications: Worldwide Market Development. Artech House, Boston-London.
- [3] EMC World Cellular Database (www.emc-database.com)
- [4] HIV 2003. évi Statisztikai Évkönyv
- [5] Madden, C.–Neal, G.C.–Dalzell, B. (2004): A dynamic model of mobile telephony subscription incorporating a network effect. Telecommunications Policy, 28. pp.133–144.
- [6] Steinbock, D. (2003): Globalization of wireless value systems from geographic to strategic advantages. Telecommunications Policy, pp.207–235.
- [7] Ramos, S.–Feijoo, C.–Perez, J.–Castejon, L.–Segura, I. (2003)
- [8] Nemzetközi Statisztikai Évkönyv, 2004.
- [7] Magyar Statisztikai Évkönyv, 2004.



MAIPAN – middleware személyi hálózatokban működő alkalmazások összekapcsolására

RÓNAI MIKLÓS AURÉL, FODOR KRISTÓF, BICZÓK GERGELY, TURÁNYI ZOLTÁN, VALKÓ ANDRÁS

Ericsson Magyarország Kft, Traffic Lab

{Miklos.Ronai, Kristof.Fodor, Gergely.Biczok, Zoltan.Turanyi, Andras.Valko}@ericsson.com

Kulcsszavak: PAN, mindenütt jelenlévő számítástechnika, hozzáférés-vezérlés, dinamikus kapcsolatrendszer-kezelés

Dinamikusan változó személyi hálózatok (Personal Area Network – PAN) létrehozására cikkünkben a MAIPAN elnevezésű middleware megoldást javasoljuk. A middleware eltakarja a PAN szétdaraboltságát, valamint az eszközök konfigurációját, és a hálózatot egységesen, egyetlen számítógépnek mutatja az alkalmazások felé. A megoldás lehetővé teszi a PAN-ban több eszközön működő elosztott alkalmazások egyszerű összekapcsolását, valamint a PAN konfiguráció változásakor az adatfolyamok transzparens átírányítását. A javasolt middleware a különböző eszközökön futó alkalmazások által nyújtott szolgáltatások ki és bemeneti csatlakozási pontjait virtuális csatornák létrehozásával köti össze. A csatornák szükség esetén újrafelrakhatók, ha a hálózat felépítése vagy a felhasználói igények megváltoznak. Az irodalomban található megközelítésekkel szemben egy olyan megoldást ajánlunk, ahol a szolgáltatások között kiépített kapcsolatrendszer dinamikusan menedzselhető, valamint szorosan egybe van építve egy újfajta jogosultság kezeléssel.

1. Bevezetés

A vezeték nélküli terminálok, mint például okostelefonok, digitális személyi asszisztensek (Personal Digital Assistant – PDA) és laptopok növekvő száma miatt egyre inkább szükség van arra, hogy személyi hálózatokat egyszerű módon lehessen felállítani, valamint be- és újraconfigurálni.

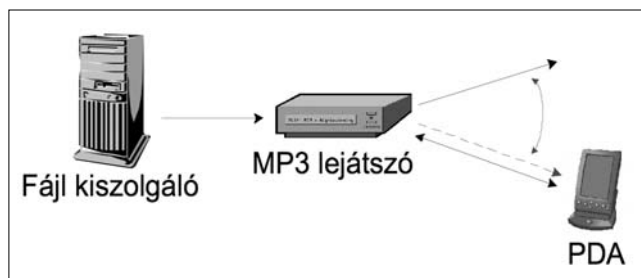
Cikkünkben azzal foglalkozunk, hogy hogyan lehet személyi hálózatokban egy dinamikusan változtatható, de egységes számítástechnikai környezetet létrehozni, mely kiterjed az egymással összekötött vezetékes és vezeték nélküli, valamint fix és mobil eszközökre. A MAIPAN middleware-t javasoljuk – ami az alkalmazások alatt és a hálózati réteg fölött helyezkedik el –, mely megoldás eltakarja a személyi hálózatban működő különálló eszközöket, és az ezeken futó alkalmazások képességeit úgy mutatja, mintha azok mind ugyanazon a számítógépen lennének elérhetőek.

Ezzel egy olyan egységes programozási platformot nyújt, amely lehetővé teszi a személyi hálózatok egyszerű felépítését és a PAN-ban futó alkalmazások dinamikus összekapcsolását, valamint szétválasztását. Alkalmazás programozók, akik a middleware által nyújtott egységes alkalmazás programozói interfészt (API) használják, úgy fejleszthetnek szoftvert, hogy nem kell foglalkozniuk a különböző PAN konfigurációkkal és dinamikákkal. Bizonyos képességek meglétét tételezhetik fel, de figyelmen kívül hagyhatják azt, hogy ezen képességeket egy bizonyos eszközök futó egyetlen alkalmazás vagy több eszközön futó több alkalmazás halmaz nyújtja. A programozóknak csak a middleware-be kell regisztrálniuk alkalmazásaik ki- és bemeneteit, és nem kell törődniük azzal, hogy milyen eszközök vagy alkalmazások fognak ezekhez a csatlakozási pontokhoz kapcsolódni, valamint a programjaikat használni.

A bemutatásra kerülő middleware jogosultság, valamint rugalmas és átruházható kapcsolatrendszer kezelést is tartalmaz. A middleware-ben van néhány intelligens funkció, mely hatékonyabbá teszi az ember számítógép kapcsolatot (Human Computer Interaction – HCI). Elméletben bármely szolgáltatás-felderítési, hálózati vagy adatkapcsolati megoldás használható a MAIPAN-nel.

Az 1. ábrán PAN alkalmazásra látható egy példa. Tegyük fel, hogy egy digitális személyi asszisztens felhasználó, akinek a készülékén a MAIPAN fut, belép egy szobába ahol hangszórók vannak a sarokban, melyek szintén MAIPAN-t futtatnak. Ha a felhasználó úgy dönt, hogy zenét szeretne hallgatni, de se mp3 fájllai, se mp3 lejátszója nincsen, akkor bekapcsolva a PDA-ját lekérdezheti, hogy milyen eszközök és szolgáltatások érhetők el a szobában. Ekkor először is – függetlenül MAIPAN-tól – a PDA hálózati rétege kapcsolatot hoz létre a szobában lévő eszközökkel és a szolgáltatás-felderítő protokollja pedig körbekérdez, hogy milyen szolgáltatások működnek a környéken. Ezt követően a felhasználó a MAIPAN vezérlővel bekonfigurálhat egy olyan személyi hálózatot, amely a PDA-t, a sarokban lévő hangszórókat, az mp3 lejátszót és a hálózaton elérhető fájlszerveret tartalmazza.

1. ábra Zenehallgatás



MAIPAN létrehozza a szükséges virtuális kapcsolatokat, azaz a fájlservert az mp3 lejátszó bemenetére, az mp3 lejátszó kimenetét a hangszórókra és az mp3 lejátszó vezérlő bemenetét a PDA-ra köti. Ettől kezdve a felhasználó a fájlserveren ki tudja választani azokat a zenéket, amiket hallgatni szeretne és a PDA-jával utasítani tudja az mp3 lejátszót, hogy a zenét a hangszórókon szólaltassa meg. Az ábrán a nyilak az adatáramlás irányát mutatják.

A cikk hátralévő részében először áttekintjük az irodalmat, majd a következő szakaszban bemutatjuk a MAIPAN middleware-t, végezetül pedig összefoglaljuk munkánkat.

2. Irodalom-áttekintés

Mark Weiser a mindenütt jelenlévő számítástechnika (ubiquitous computing) koncepcióját a 90-es évek elején alkotta meg [1,2]. Tíz évvel később, Satyanarayanan Weiser víziójának kihívásaival foglalkozott [3], és bevezette a mindent átható számítástechnika (pervasive computing) fogalmát, mely többek között az elosztott rendszerek, a mobil számítástechnika és az okos területek (smart spaces) keveréke.

A middleware-ek lényeges részei mind a mindenütt jelenlévő, mind a mobil számítástechnikának. Mascolo és társai [4] többek között azt feszegetik, hogy a hagyományos middleware rendszerek (mint például a CORBA [5]), miért nem annyira megfelelőek a mobil környezetben, valamint arról írnak hogyan is kell egy mobil számítástechnikai middleware-t tervezni. Másik kulcskérdése az ilyen mindenütt jelenlévő rendszereknek: a biztonság. Chandrasiri és társai [6] a személyi biztonsági területek használatát javasolják, mely terület a felhasználó személyes eszközeiből áll, valamint különféle biztonsági aspektusokat elemez a PAN-okkal kapcsolatban. Manapság számos projekt foglalkozik ezekkel a kutatási területekkel, ezek közül az alábbiakban kiemelünk néhányat. A területről és az alábbi projektek közül néhányról korábban már bővebben is írtunk [26,27].

Az AURA projekt [7] célja, hogy minden felhasználónak kialakítson egy láthatatlan számítástechnikai aurát. A GAIA [8] project egy middleware infrastruktúrát ajánl, hogy az okos területeket lehessen létrehozni. Az Oxygen project [9] nagyon intelligens, felhasználóbarát, könnyen használható mobil eszközök tervezésével

foglalkozik. A Portolano-projekt [10] adaptív felhasználói interfészek létrehozására fókuszál. A projekt keretében a one.world [11] architektúrát tervezték meg, mely mindenütt jelenlévő számítástechnikai alkalmazások készítését támogatja. Ugyanilyen alkalmazások készítésének elősegítésére az extrovert-Gadgets projekt [12] különféle architektúrákat vizsgált meg és a GAS-OS middleware használatát javasolja szenzorok és működtető szerkezetek (actuator) vezérlésére. A 2WEAR projekt [13] viselhető (wearable) személyi rendszereket analizált, amelyek dinamikusan rakhatók össze különböző eszközök összekapcsolásával. A Cortex projekt [14] olyan új típusú alkalmazásokkal foglalkozik, melyek emberi beavatkozás nélkül, önállóan is tudnak működni. Az EasyLiving project [15] célja intelligens otthoni és irodai környezet kialakítása. A Speakeasy megközelítés [16] az alkalmazások közötti minimális interfészek definiálására koncentrál mobil kód és mobil ügynökök használatával. A MobiDesk [17] megoldás egy virtuális számítástechnikai asztalt definiál, mely megoldással többek között a hálózati sebességre is lehet optimalizálni. A Virtual Device koncepció [18,19] a felhasználó környezetében minden autonóm eszközt egy nagy virtuális egységgé alakít.

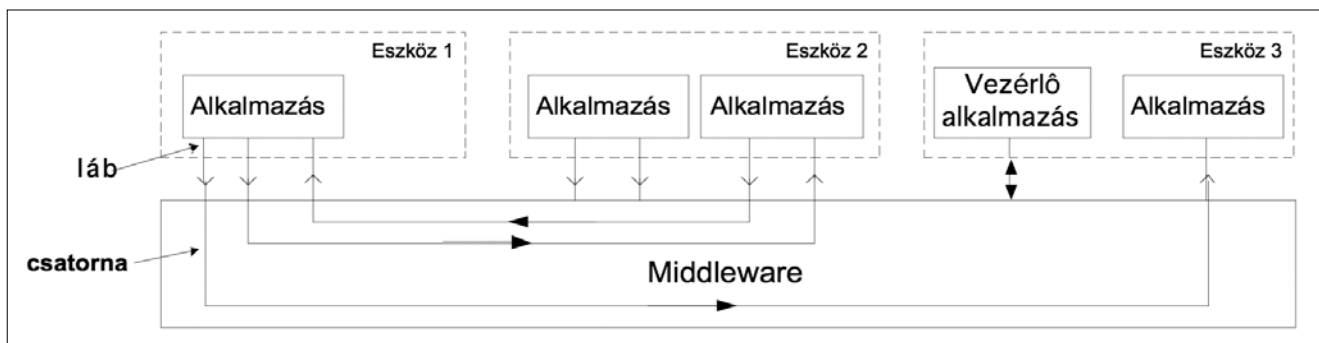
Hasonlóképpen néhány fent említett megoldáshoz MAIPAN a teljes személyi hálózatot az alkalmazások felé egyetlen eszköznek mutatja. A MAIPAN-nal a felhasználók egyszerűen hozhatnak létre és konfigurálhatnak újra személyi hálózatokat. A MAIPAN rendszer alapötleteit angolul a [20]-as cikkben publikáltuk.

3 A MAIPAN platform

3.1. Alapötletek és definíciók

A MAIPAN eszközöket, alkalmazásokat és szolgáltatásokat különböztet meg. Az „eszköz” szó a fizikai eszközt jelenti amin az „alkalmazás” fut, ami pedig nem más mint az a szoftver amelyik az adott „szolgáltatást” felajánlja. Ezeket az absztrakciókat például az mp3 lejátszó esetében alkalmazva azt mondhatjuk, hogy a „fizikai eszközön” fut az „mp3 lejátszó alkalmazás” ami az „mp3 lejátszó szolgáltatást” nyújtja. A különbségtétel az alkalmazás és a szolgáltatás között azért szükséges, mert szolgáltatást nem csak szoftver nyújthat, hanem hardver is (például éger).

2. ábra MAIPAN kapcsolatrendszer



A MAIPAN három koncepción alapszik (2. ábra): *láb, csatorna és kapcsolatrendszer*. Az alkalmazásoknak, melyek a szolgáltatásokat nyújtják, vannak bemenetek és kimenetek, amiket – kölcsönvéve az elnevezést az integrált áramköri világból – lábaknak nevezünk. A lábak az alkalmazások kapcsolódási pontjai a middleware felé, így a middleware az alkalmazásokat a PAN-ban be és kimeneti lábak csoportjainak látja. A lábaknak egy előre definiált típusuk lehet, amely a láb által generált, illetve elnyelt adat, azaz az adott alkalmazás által kezelt információ fajtáját mutatja meg (például egér mozgása, billentyű leütése). A szükségleteknek megfelelően bármikor új típusok definiálhatók. A PAN vezérlő alkalmazás a felelős azért, hogy csak a megfelelő lábakat lehessen összekötni.

A lábak közötti kommunikációhoz a middleware pont-pont kapcsolatokat, azaz csatornákat épít ki. Azt a csatorna halmazt, amely egy adott PAN szolgáltatás használatához szükséges, kapcsolatrendszernek nevezzük. Például, az 1. ábrán látható csatornák alkotják az mp3 lejátszó kapcsolatrendszert: van egy csatorna a fájlserver és az mp3 lejátszó, egy másik az mp3 lejátszó és a hangszórók, és egy vezérlő csatorna az mp3 lejátszó és a PDA között. A PDA ebben az esetben csak a vezérlő entitás szerepét játssza, ugyan ő a tulajdonosa, de nem tagja a kapcsolatrendszernek. Az alkalmazások nem tudnak a csatornákról, sem a kapcsolatrendszerekről, csak saját lábaikról van tudomásuk.

3.2. Biztonság és hozzáférési jogosultság kezelés

A biztonsági és jogosultság kezelési funkciókat eszköz szinten definiáltuk. Ez azt jelenti, hogy a szolgáltatásokhoz (az adott alkalmazás lábaihoz) való hozzáférés az eszközöknek van engedélyezve. Így ha egy eszköz jogot kap egy szolgáltatás használatára, akkor minden alkalmazás, amely az adott eszközön fut jogosult lesz hozzáférni az adott szolgáltatáshoz. Ez az egyszerűsítés akkor engedhető meg, ha feltételezzük, hogy a PAN-ban kis eszközök vannak, melyek csak egy-két egyszerű szolgáltatást nyújtanak (például egér, mp3 lejátszó). Ebben az esetben főleg egy bonyolultabb eljárás alkalmazása, ahol a hozzáférést az eszközökön futó alkalmazásoknak egyesével kellene megadni.

A vezérlő alkalmazás, mely a vezérlő entitás szerepét játszó eszközön fut, tudja felépíteni és újrakonfigurálni a kapcsolatrendszereket. A vezérlő entitásnak kell ellenőriznie a hozzáférési jogosultságok meglétét. Ha ezek hiányoznak, akkor a szolgáltatást nyújtó eszköztől meg kell kérdeznie, hogy a felhasználó használhatja-e az adott szolgáltatást. Központi vezérlő entitás lehet például egy PDA, amelynek elegendő számítástechnikai kapacitása van ahhoz, hogy menedzselni tudjon egy teljes személyi hálózatot. Minden más eszköz, amely a PAN-ban van, *résztevőknek* nevezzük. Speciális esetekben a résztvevők delegálhatják a hozzáférés engedélyezés döntési jogkörét más eszközöknek, amelyeket *menedzsereknek* hívunk.

Itt most nem megyünk részletekbe a hitelesítés és engedélyezéssel kapcsolatban, viszont a [6]-ban talál-

ható megoldások a MAIPAN-nal is alkalmazhatóak a biztonságos kommunikáció megvalósítására.

3.3. Kapcsolatrendszerek áthelyezése

A PAN-ban legalább egy eszköznek kell lennie, mely a vezérlő alkalmazás szerepét játssza. Abban az esetben, ha ez az eszköz eltűnik (mert kikapcsolják, vagy a felhasználó távozik), akkor az általa vezérelt összes kapcsolatrendszer automatikusan lebontódik. Ahhoz, hogy a kapcsolatrendszereket a vezérlő eszköz eltűnése után is fenntarthassuk, a MAIPAN lehetőséget ad a működő kapcsolatrendszerek egyik vezérlő alkalmazásról a másikra való átvitelére.

Például, hogy zenét tudjanak hallgatni egy tárgyalóteremben, az egyik felhasználó egy mp3 lejátszó kapcsolatrendszert konfigurál be. Így az adott felhasználó eszköze lesz a vezérlő entitása a kapcsolatrendszernek. Ha a felhasználó el akarja hagyni a szobát, akkor a MAIPAN-nel közli, hogy a kapcsolatrendszert vigye át egy másik vezérlő entitásra, ami például egy másik felhasználó PDA-ja lehet.

3.4. A kapcsolatrendszerek újrakonfigurálása

Abban az esetben, ha egy a PAN-ban résztvevő eszköz eltűnik (például a felhasználó elhagyja a termet, vagy az eszköz akkumulátora lemerül), akkor azok a csatornák, melyekben az eszköz részt vett automatikusan lebontódnak és a kapcsolatrendszert újra kell konfigurálni, ha a továbbiakban is használni akarják.

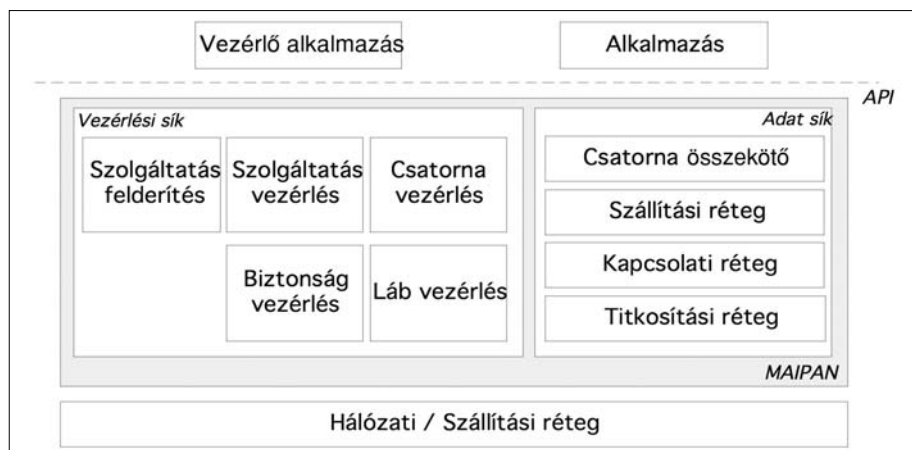
Az első lépésben a MAIPAN értesíti az érintett vezérlő alkalmazás(oka)t az eltűnés tényéről. A második lépésben a vezérlő alkalmazás eldöntheti, hogy mely más szolgáltatás(oka)t kívánja használni a kieső(k) helyett. Több választási lehetőség esetén a vezérlő alkalmazás kérheti a felhasználót, hogy döntse el ő, mely más szolgáltatásokat szeretné használni, vagy az alkalmazás önmagában is dönthet, ha egyetlen lehetőség van, illetve ha – ismerve a felhasználó preferenciáit – ki tudja választani a legmegfelelőbb szolgáltatás(oka)t az eltűnt(ek) pótlására. A harmadik lépésben a vezérlő vagy felépíti az új csatornákat vagy lebontja az érintett kapcsolatrendszert.

3.5. Architektúra

A fenti ötleteken alapulva alkottuk meg a MAIPAN middleware-t, melynek architektúrája a 3. ábrán (a következő oldalon) látható. Az alkalmazások a middleware fölött futnak, míg a middleware alatt lévő rétegek végponttól végpontig történő adatkommunikációt biztosítanak. Ahogy az ábrán látható, a protokoll verem függőlegesen két részre van osztva. Az egyik rész az *adat sík*, melynek feladata az alkalmazások közötti hatékony és biztonságos adatátvitel. A *vezérlési sík* a lábak, csatornák és kapcsolatrendszerek menedzseléséért, valamint a kulcs- és jogosultság kezelésért felel.

Adat sík

Az alkalmazások a lábaikon keresztül küldik a middleware felé az adatot, amit a *csatorna összekötő réteg* a megfelelő csatornába irányít. A *szállítási réteg* létre-



3. ábra A MAIPAN architektúrája

hozza a csomagokat és olyan szolgáltatásokat nyújt, mint folyam szabályozás, sorrendezés, automatikus új-rakúldás, QoS stb. Csak azok a funkciók kerülnek bekapcsolásra az adott csatornákon, amelyek az adott szolgáltatáshoz szükségesek és hiányoznak a middleware alatti rétegekből. A *kapcsolati réteg* olyan információkat ad a csomagokhoz, melyek a célba éréshez szükségesek: a forrás és cél eszköz címét, valamint a csatorna azonosítóját. Végezetül a *titkosítási réteg* a csomag integritásának ellenőrzéséhez paritás biteket számol és szükség esetén titkosítja a csomagot.

Vezérlési sík

A vezérlési sík tartalmazza azokat a vezérlési funkciókat, melyek a PAN menedzseléséhez szükségesek. A *szolgáltatás-vezérlő* rész a helyi szolgáltatások regisztrálását és hozzáférési jogosultságokat kezeli, valamint kommunikál a szolgáltatás-felderítő protokollal. Elméletben bármilyen szolgáltatás-felderítési protokoll hozzákapszolható a MAIPAN-hez, úgy mint az SLP, az UPnP vagy a Salutation [21,22]. A vezérlő alkalmazás utasításainak megfelelően a *csatorna-vezérlő* hozza létre és szükség esetén konfigurálja újra a kapcsolatrendszereket. A vezérlő alkalmazás igényeinek megfelelően, a csatorna-vezérlő kéri meg a PAN-ban résztvevő eszközök *láb-vezérlőit*, hogy kívánt lábak között építsék ki a csatornákat. A láb-vezérlő utasítja a csatorna összekötő réteget, hogy hozza létre a szükséges csatornát, a szállítási rétegben aktiválja az adott csatornához szükséges szállítási funkciókat és beállítja a csatorna végpontját a kapcsolati rétegben. A *biztonsági vezérlő* kezdeményezi és koordinálja az eszközök közötti azonosítási eljárást, kezeli a szolgáltatás hozzáférési jogosultságokat és tárolja a biztonságos kommunikációhoz szükséges információkat (például biztonsági kulcsokat).

3.6. Implementáció

A MAIPAN egy korábbi verzióját Linux operációs rendszerre C-ben implementáltuk [23]. Készítettünk néhány MAIPAN alkalmazást (pl. mp3 lejátszó, fájlserver) és egy vezérlőt is, hogy vizsgálhassuk a middleware viselkedését [24,25]. Az implementációból tanultakat is beépítettük a middleware most bemutatott verziójába.

4. Összefoglalás

Cikkünkben a MAIPAN platformot mutattuk be, ami egy olyan middleware, mely személyi hálózatokban működő alkalmazások összekapcsolását teszi lehetővé. E middleware lényege, hogy egy olyan PAN programozási platformot nyújtson, ahol a hardveres és szoftveres erőforrások összeköthetők egymással és a személyi hálózat elosztottsága el van takarva a szolgáltatások elől.

A javasolt architektúrát használva, az elosztott PAN alkalmazásokat készítő programfejlesztőknek nem kell foglalkozniuk a PAN konfigurálással és dinamikákkal (például eszközök eltűnése, illetve megjelenése), és a rendszer által nyújtott egységes alkalmazás programozási interfészt használva egyszerűen készíthetnek PAN szoftvereket.

A MAIPAN a biztonságos jogosultság kezelési mechanizmusával és a központi vezérlő entitásával a területen egy új megközelítést reprezentál. A MAIPAN jogosultság kezelése

- 1) biztosítja a felhasználó eszközei közötti együttműködést,
- 2) biztonságot nyújt a felhasználó eszközeinek más felhasználó eszközeivel szemben és
- 3) emellett megengedi a különböző felhasználók eszközei közötti kontrollált kommunikációt.

A megoldás ezen kívül kényelmes jogosultság kezelést és PAN konfigurációt biztosít a központi vezérlő alkalmazás segítségével, valamint abban is egyedülálló, hogy megengedi a központi entitás cseréjét, azaz a kapcsolatrendszer vezérlési jogai eszközök között szabadon átvihetők.

A MAIPAN jogosultság kezelési sémája szerint az adott szolgáltatás használati jogát eszközök kaphatják meg, így ha egy eszköz jogosultságot szerez egy adott szolgáltatás használatára, akkor az adott eszközön futó összes alkalmazás használhatja az alkalmazást. A dinamikus kapcsolatrendszer menedzsment azt a szituációt kezeli, amikor egy PAN szolgáltatásban résztvevő eszköz eltűnik. Ebben az esetben MAIPAN értesíti a vezérlő alkalmazást, ami ezek után megpróbálja új résztvevők bevonásával újrakonfigurálni a kapcsolatrendszert, vagy ha több lehetőség adódik és a felhasználó preferenciái nem ismertek, akkor a felhasználó beavatkozását kérheti.

A MAIPAN kapcsolatrendszer áthelyezést is támogat, ami abban az esetben hasznos, amikor a kapcsolatrendszer birtokló eszköz el akarja hagyni a PAN-t. Ebben az esetben, ha a felhasználó a távozása után továbbra is fenn akarja tartani a kapcsolatrendszert, akkor a MAIPAN-t a vezérlő alkalmazás segítségével utasíthatja, hogy helyezze át a kapcsolatrendszer vezérlését egy másik eszközre.

Irodalom

- [1] Mark Weiser:
„The Computer for the 21st Century”,
Scientific American, September 1991.
- [2] Mark Weiser:
„Some Computer Science Issues in
Ubiquitous Computing”,
Communications of the ACM, July 1993.
- [3] M. Satyanarayanan:
„Pervasive Computing: Vision and Challenges”,
IEEE Personal Communications, August 2001.
- [4] C. Mascolo, L. Capra, W. Emmerich:
„Middleware for Mobile Computing (A Survey)”,
In Advanced Lectures in Networking.
Editors: E. Gregori, G. Anastasi, S. Basagni.
Springer, LNCS 2497, 2002.
- [5] A. Pope:
„The Corba Reference Guide: Understanding the
Common Object Request Broker Architecture”,
Addison-Wesley, Januar 1998.
- [6] P. Chandrasiri, O. Gurleyen, Y. Shahabi,
C. Gehrmann, A. Jonsson, M. Naslund:
„Personal Security Domains”,
Contribution to the 10th WWRF Meeting,
New York, October 27-28, 2003.
- [7] D. Garlan, D. Siewiorek, A. Smailagic, P. Steenkiste:
„Aura: Toward Distraction-Free Pervasive Computing”,
IEEE Pervasive Computing, 2002.
<http://www-2.cs.cmu.edu/aura/>
- [8] Gaia Project: „Active Spaces for
Ubiquitous Computing”;
<http://gaia.cs.uiuc.edu/index.html>
- [9] „MIT Project Oxygen”, Online Documentation,
<http://oxygen.lcs.mit.edu/publications/Oxygen.pdf>
- [10] M. Esler, J. Hightower, T. Anderson, G. Borriello:
Next Century Challenges:
„Data-Centric Networking for Invisible Computing:
The Portolano Project at the Univ. of Washington”,
(MobiCom'99),
<http://portolano.cs.washington.edu/proposal/>
- [11] R. Grimm, J. Davis, E. Lemar, A. MacBeth,
S. Swanson, T. Anderson, B. Bershad, G. Borriello,
S. Gribble, D. Wetherall:
„System support for pervasive applications”,
ACM Transactions on Computer Systems,
22(4):421–486, November 2004.
- [12] A. Kameas, S. Bellis, I. Mavrommati, K. Delaney,
M. Colley, A. Pounds-Cornish:
„An Architecture that Treats Everyday Objects as
Communicating Tangible Components”,
in Proc. of the 1st IEEE Intern. Conf. on Pervasive
Computing and Communications (PerCom'03),
Fort Worth, Texas, USA, March 23-26, 2003., p.115.
- [13] 2WEAR project: „A Runtime for Adaptive and
Extensible Wireless Wearables”;
<http://2wear.ics.forth.gr>
- [14] CORTEX Project:
„CO-operating Real-time sentient objects:
architecture and EXperimental evaluation”;
<http://cortex.di.fc.ul.pt/index.htm>
- [15] B. Brumitt, B. Meyers, J. Krumm, A. Kern, S. Shafer:
„EasyLiving: Technologies for Intelligent Environments”
in Proc. of Handheld and Ubiquitous Computing
Symposium, (Bristol, England), 2000.
- [16] W. K. Edwards, M.W. Newman, J. Sedivy, T. Smith:
„Challenge: Recombinant Computing and
the Speakeasy Approach” (MobiCom'02),
September 23-28, 2002., Atlanta, Georgia, USA.
- [17] R. Baratto, S. Potter, Gong Su, J. Nieh:
„MobiDesk: Mobile Virtual Desktop Computing”,
Proc. of the 10th Annual ACM Intern. Conf. on
Mobile Computing and Networking (MobiCom'04),
Philadelphia, PA, September 26-October 1, 2004.
- [18] Jonvik, T.E., Engelstad, P.E., Thanh, D.V.:
„Building a Virtual Device on Personal Area Network”,
Proc. of 2003 Intern. Conf. on Software,
Telecom. and Computer Networks (SoftCom'03),
Dubrovnik (Croatia) / Ancona, Venice (Italy),
October 7-10, 2003.
- [19] Jonvik, T.E., Engelstad, P.E., Thanh, D.V.:
„Dynamic PANBased Virtual Device”,
Proc. of 2nd IASTED Int. Conf. on Communications,
Internet and Information Technology (CIIT'2003),
November 17-19, 2003.
- [20] M. A. Rónai, K. Fodor, G. Biczók, Z. Turányi, A. Valkó:
„MAIPAN: Middleware for Application Interconnection
in Personal Area Networks”,
Poster at Mobiquitous 2005 Conference,
San Diego, CA, USA, July 17-21, 2005.
- [21] Reakesh John:
„UPnP, Jini and Salutation – A look at some popular
coordination frameworks for future networked devices”,
California Software Laboratories Inc.,
Technical Report, June 17, 1999.
- [22] F. Zhu, M. Mutka, L. Ni:
„Classification of Service Discovery in Pervasive
Computing Environments” MSU-CSE-02-24,
Michigan State University, EastLansing, 2002.
- [23] Kristóf Fodor:
„Implementation of a Protocol Stack for
Personal Area Networks”,
diplomamunka, 2003. június
- [24] Fodor K., Kovács B.:
„A Blown-up rendszer megvalósítása”,
HTE-BME diákkonferencia, Budapest, 2003. május
- [25] Balázs Kovács:
„Design and Implementation of Distributed
Applications in Ad Hoc Network Environment”,
diplomamunka, 2003. május
- [26] Biczók G., Fodor K., Kovács B., Szabó Á.:
Pervasive computing – rejtett számítástechnika,
Híradástechnika, 2003. március
- [27] Biczók G., Fodor K., Kovács B., Szabó Á.:
„Blown-up rendszer tervezése és megvalósítása”,
első helyezett TDK/OTDK dolgozat,
Győr, 2002. november/2003. április

Időben változó sorbanállási rendszer vizsgálata mátrix-geometrikus eljárások segítségével

ÉLTETŐ TAMÁS, VADERNA PÉTER

{tamas.elteto, peter.vaderna}@ericsson.com

Kulcsszavak: sorbanállási rendszer, Markov-érkezési folyamat, mátrix-geometriai módszer

Ebben a cikkben egy végtelen sok kiszolgálóval rendelkező sorbanállási rendszert vizsgálunk időfüggő esetben, azaz nem a stacionárius megoldását keressük, hanem a rendszerben lévő igények számát követjük nyomon az idő múlásának függvényében. A sorbanállási rendszerbe igények érkeznek egy nem stacionárius pontfolyamat szerint, és az igények kiszolgálása is véletlen ideig tart. Ezt a nem stacionárius érkezési folyamatot egy Markov érkezési folyamatként adjuk meg. Az időben változó sorbanállási rendszer egy inhomogén lineáris differenciálegyenlet-rendszerre vezet, melynek megoldását kiszámítjuk. A számítási módszert egy gyakorlati példára – egy távközlési szolgáltatás vizsgálatára – alkalmazzuk.

1. Bevezetés

Végtelen kiszolgálóval rendelkező sorbanállási modellek először a hagyományos telefonközpontok méretezése során vetődtek fel. A modellben Poisson folyamat szerint érkező telefonhívásokat fogadott és továbbított a telefonközpont, ahol csak véges sok vonal állt rendelkezésre a hívások továbbítására. Ha ismerjük a telefonközponton egy időben átmenő hívások számának stacionárius eloszlását, akkor meg tudjuk tervezni a ki- menő vonalak számát úgy, hogy csak nagyon kicsi valószínűsége legyen az érkező hívások kapacitáshiány miatti visszautasításának.

A távközlési hálózatok fejlődésével, a szolgáltatások sokaságának megjelenésével az eredeti feladat általánosítása is szükségessé vált. Az Internet forgalom statisztikai elemzése során kimutatták, hogy a Poisson folyamat sok esetben nem elég jó modell adathívások, adatátviteli kapcsolatok érkezési folyamatának a leírására [3]. Másrészt, mint az ismertetésre kerülő példa is mutatja, vannak esetek, mikor a stacionárius állapot leírása nem elegendő kérdéseink megválaszolásához.

A cikkben a Poisson folyamat általánosításaként az úgynevezett Markov érkezési folyamatot [1] fogjuk alkalmazni. Az érkezéseket egy Markov lánc állapotai, illetve állapot-átmenetei befolyásolják. Mivel célunk az volt, hogy a végtelen kiszolgálós sorbanállási rendszerek tranziens viselkedésének lehető legáltalánosabb analitikus leírását adjuk, így a kiszolgálási időnek – a hívások hosszának – az eloszlására is kellőképp gazdag eloszláscsaládot választottunk. A cikkben fázis típusú eloszlásokat használunk, melyek elnyelő állapottal rendelkező Markov láncok abszorpcióig eltelő idejét írják le. Ez az eloszláscsalád a gyakorlatban is jól használható, mivel ismert, hogy kellően nagy állapotterű fázis típusú eloszlással bármely eloszlást meg tudunk közelíteni pontosan [4].

A Markov érkezési folyamatok és fázis típusú tartási idők segítségével a feladat, bár nagyon általános, tel-

jes mértékben markovi maradt. Ez lehetővé teszi a rendszer analitikus vizsgálatát nem csak stacionárius esetben, hanem követni lehet a rendszer időbeli változását is, ami nagy előnyt jelent más – inkább a stacionárius állapot leírására törekvő – módszerekkel szemben. A rendszerben tartózkodó hívások számának momentumaira megadott időfüggő differenciálegyenlet-rendszert megoldjuk, és megmutatjuk, hogy a megoldás paraméterei hogyan függenek a bemeneti paraméterektől. Megjegyezzük, hogy a vizsgálatunk tárgyát képező rendszer-nél [5]-ben egy általánosabb esetet elemeznek, mi egy eltérő, alkalmazás-központúbb megközelítést mutatunk.

A számítási módszer segítségével egy gyakorlatban felmerülő problémát oldunk meg. A feladat egy internetes hírszolgáltató működéséhez kapcsolódik. Azt feltételezzük, hogy a hírszolgáltató a különböző operációs rendszerekkel és sokféle internetes böngészővel rendelkező felhasználóknak egységes külalakkal kívánja a híreket megjeleníteni. Ezt úgy éri el, hogy amikor egy felhasználó egy hírt kér a kiszolgálótól, akkor a kiszolgáló érzékeli a felhasználó beállítását. A rendelkezésre álló információk alapján egy konvertáló segítségével elkészíti, majd elküldi a hírek egy olyan változatát, amelyet a beállításoknak megfelelően formázott. A hírek konvertálásának folyamatát írjuk le az ismertetett számítások segítségével.

A cikk következő fejezetében definiáljuk a használt matematikai eszközöket, majd ismertetjük a számítási eljárás lényeges lépéseit. Ezek után bemutatjuk a gyakorlati problémát, végül összefoglaljuk a cikket. Az írásunk végén található függelékben pedig elméleti állításaink bizonyításai olvashatók.

2. Matematikai háttér

Az alábbiakban röviden összefoglaljuk a Markov érkezési folyamatot (angol betűszóval: MAP). A téma részletesebb tárgyalása [1]-ben található angol nyelven.

A Markov érkezési folyamat egy olyan egyszerű pont-folyamat, amelynek érkezéseit egy folytonos idejű Markov folyamat befolyásolja. Tekintsünk egy sztochasztikus folyamatot:

$$\{(N(t), J(t)) : t \geq 0\},$$

ahol $N(t)$ a $(0, t)$ intervallumban történő érkezések számát, míg $J(t)$ a t időpontban a folyamat „fázisát” jelöli. Itt $J(t)$ egy folytonos idejű Markov folyamat egy véges, Melemű állapotterén, $N(t)$ pedig egy $J(t)$ -től függő sztochasztikus folyamat. Az $(N(\cdot), J(\cdot))$ együttes sztochasztikus folyamat egy folytonos idejű Markov folyamat lesz az $\{(n, j) : n \geq 0, 1 \leq j \leq M\}$ állapotterén, az alábbi definíció szerint.

Osszuk fel az $\{(n, j) : n \geq 0, 1 \leq j \leq M\}$ állapotteret rész-halmazokra:

$$I(k) = \{(k, 1), (k, 2), \dots, (k, M)\},$$

ahol k nemnegatív egész. Ekkor legyen Q mátrix az $(N(\cdot), J(\cdot))$ folyamat infinitezimális generátor mátrixa az alábbi definícióval:

$$Q = \begin{bmatrix} D_0 & D_1 & 0 & 0 & \dots \\ 0 & D_0 & D_1 & 0 & \dots \\ 0 & 0 & D_0 & D_1 & \dots \\ 0 & 0 & 0 & D_0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix},$$

ahol

- D_0 és D_1 $M \times M$ -es mátrixok,
 - $D_1 \geq 0$,
 - $[D_0]_{i,j} \geq 0, 1 \leq i \neq j \leq M$, és $[D_0]_{i,i} \leq 0, 1 \leq i \leq M$,
 - $D = D_0 + D_1$ mátrix sztochasztikus,
- vagyis minden sorösszege 0.

Szemléletesen arról van szó, hogy a $J(\cdot)$ folytonos idejű Markov folyamat infinitezimális generátor mátrixa D . Ez a mátrix bomlik fel két mátrix összegére. Ezek közül a D_0 mátrix tartalmazza azokat az állapotátmeneteket, ahol nem történik érkezés, míg D_1 mátrix írja le azokat az állapotátmeneteket, ahol történik érkezés, illetve szintén D_1 mátrix adja meg a feltételes Poisson folyamatok érkezési intenzitását, amikor is nem vált állapotot a $J(\cdot)$ folyamat, de érkezés történik. (A Markov érkezési folyamatokat általánosíthatjuk nem egyszerű pontfolyamatokra is, azaz megengedhetünk egyszerre több érkezésre is. Ekkor további D_n mátrixok írják le a csoportos érkezéseket.)

A Markov érkezési folyamatok egy speciális osztályát Markov-modulált Poisson folyamatoknak nevezzük. A Markov-modulált Poisson folyamatok esetén a D_1 mátrixnak csak a diagonálisában vannak nem nulla elemek, azaz érkezés nem fordulhat elő, amikor a $J(\cdot)$ Markov folyamat állapotot vált. A Markov-modulált Poisson folyamatok részletes összefoglalása [2]-ben található.

Egy másik speciális osztály az úgynevezett PH felújítási folyamatok osztálya. Ezek olyan felújítási folyamatok, melyek felújítási időközének eloszlása PH (fázis típusú) eloszlás. A PH eloszlások tulajdonképpen különböző paraméterű Erlang eloszlások keverékei, de inkább

egy elnyelő állapottal rendelkező Markov folyamatnak az indítás és az elnyelődés között eltelt idejeként definiálják őket.

Ezen második definíciónak megfelelően, a PH eloszlásokat egy (α, T) párral reprezentálhatjuk. Az α vektor adja meg a tranziens állapotok eloszlását a Markov lánc indításakor. A T mátrix pedig egy szubsztokhasztikus mátrix, azaz

- T invertálható mátrix,
- $[T]_{i,j} \geq 0, 1 \leq i \neq j \leq M$, és $[T]_{i,i} \leq 0, 1 \leq i \leq M$,
- T mátrix sorösszege nem pozitív.

A T mátrix a lényeges része az elnyelő Markov folyamatunk generátor mátrixának, ami a következőképpen írható le:

$$\begin{bmatrix} T & \tau \\ 0 & 0 \end{bmatrix},$$

ahol

$$\tau = -T \mathbb{I}, \quad \mathbb{I} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}.$$

A PH eloszlás eloszlásfüggvényét a következő módon fejezhetjük ki az (α, T) pár segítségével:

$$F(t) = 1 - \alpha e^{Tt} \mathbb{I},$$

ahol az α vektorral fejezzük ki az elnyelő Markov lánc kezdeti eloszlását.

Ekkor $D_0 = T$ és $D_1 = \tau \alpha$, vagyis a $J(\cdot)$ Markov folyamat állapottere megegyezik az elnyelő Markov folyamat tranziens állapotaival. Amikor az elnyelődés következik, akkor a Markov érkezési folyamatban érkezés történik, és a $J(\cdot)$ folyamat újraindul α kezdeti eloszlással.

3. Egy korlátlan számú kiszolgálóval rendelkező sorbanállási rendszer momentumai

Tekintsünk egy olyan sorbanállási rendszert, melybe az igények MAP folyamat szerint érkeznek, és végtelen sok kiszolgáló egymástól függetlenül, párhuzamosan szolgálja ki az igényeket PH eloszlású kiszolgálási idővel. A cél a rendszerben egyidejűleg kiszolgálás alatt álló igények számának meghatározása.

Jelölje $X(t)$ a kiszolgálás alatt lévő igények számát t időpontban és $J(t)$ a MAP érkezési folyamat fázisát t -időpontban! Jelölje továbbá $\mu^{(K)}(t)$ azt az M dimenziós vektort, melynek i -ik koordinátája $\mu_i^{(K)}(t)$, ahol

$$\begin{aligned} \mu_i^{(K)}(t) &= E[X^{(K)}(t) | X(0) = 0, J(0) = i], \\ K &\geq 1, 1 \leq i \leq M, \end{aligned} \quad (1)$$

$$X^{(K)}(t) = X(t)[X(t) - 1] \cdots [X(t) - K + 1]$$

Azaz, itt $\mu_i^{(K)}(t)$ a kiszolgálás alatt lévő igények számának K -dik faktoriális momentuma, ha a rendszert az i -dik fázisból indítjuk.

A cél az, hogy számítsuk ki a fenti sorbanállási rendszer momentumainak időbeli változásait leíró függvé-

nyeket. Ezeket a függvényeket egy differenciálegyenlet-rendszer határozza meg. A momentum függvények kiszámításának módját a Függelékben ismertetett (10)-(16) egyenletekből olvashatjuk le. A momentum függvények lényegében

$$t^j e^{-\gamma_{i,k} t}, \text{ illetve } t^j e^{\gamma_{m} t}$$

alakú függvények lineáris kombinációi. Minél magasabb momentum függvényt kívánunk kiszámítani, annál több tagja van a függvényt megadó összegnek. Az exponensek és a lineáris kombináció együtthatóinak a kiszámítását a Függelékben ismertetjük.

4. Gyakorlati példa

Ebben a fejezetben egy példa segítségével vizsgáljuk a momentumok gyakorlati kiszámításának lehetőségeit. A példa egy hírszolgáltató működését vizsgálja. A hírszolgáltatás egy honlapon keresztül működik. A hírszolgáltató szervere érzékeli a felhasználók böngészőinek típusát és az operációs rendszereket. A cél az, hogy a lehető legtöbb típusú operációs rendszeren futó különféle böngészőkkel a lehető legjobb minőségben jelenjenek meg a cikkek. (Tehát a képek mérete, felbontása, a szöveg tördelése, a honlapon esetleg előforduló animációk hasonlóképpen mutassanak a különféle típusú számítógépeken.)

A szolgáltatás működését az 1. ábrán illusztráljuk.

Célunk a konvertálási funkcionalitás vizsgálata egy új hír megjelenésétől egészen addig, amíg már az összes lehetséges formátumban létezik a cikk. A felhasználók gyors kiszolgálása érdekében egyszerre több konvertálás is előfordulhat. Ha a konvertálók száma nem elég nagy, akkor előfordulhat az az eset, amikor az összes konvertáló foglalt, és ezért a kiszolgáló nem képes egy új beállításokkal rendelkező felhasználó kérését teljesíteni. A konvertálók száma viszont nem lehet nagyon nagy sem, például azért, mert egy ilyen berendezés drága. A feladat az, hogy megállapítsunk a foglalt konvertálók számára egy olyan felső korlátot, melynél csak nagyon kicsi valószínűséggel lesz egyszerre több foglalt.

Az egyszerűség kedvéért feltételezzük, hogy a felhasználók kérései Poisson folyamat szerint érkeznek, bár ennél általánosabb eseteket is tudnánk kezelni. A példában használt adatok az illusztrációt szolgálják, a gyakorlatban nem vizsgáltuk, hogy hányféle böngészővel olvasnak cikkeket a felhasználók, illetve azt sem, hogy mennyi ideig tarthat egy konvertálás elvégzése egy valós számítógépen. Feltesszük tehát, hogy összesen tízféle böngészőre van konvertáló eljárás, és átlagosan másodpercenként 8 kérés érkezik. A konvertáláshoz szükséges idő exponenciális eloszlású, és egy konvertálás átlagosan 5 másodpercig tart.

$$H(t) = 1 - e^{-\frac{t}{5}}.$$

A tízféle böngészőtípus előfordulási gyakoriságait és a másodpercenként előforduló kérések számát (érkezési intenzitás) az 1. táblázat foglalja össze:

Típus	Gyakoriság	Intenzitás
1. típus	21%	1.68
2. típus	20%	1.6
3. típus	13%	1.04
4. típus	13%	1.04
5. típus	13%	1.04
6. típus	4%	0.32
7. típus	4%	0.32
8. típus	4%	0.32
9. típus	4%	0.32
10. típus	4%	0.32

1. táblázat

A 2. táblázatban látható, hogy a Markov érkezési folyamat állapotterének leírása számnégyesek segítségével lehetséges. Ezekkel a számnégyesekkel körülbelül azt írjuk le, hogy melyik típusú felhasználó számára van már a tárolóban megfelelő változat.

Gyakoriság	Típusok	Intenzitás
21%	1.	1.68
20%	2.	1.6
13%	3.,4.,5.	1.04
4%	6.,7.,8.,9.,10.	0.32

2. táblázat

1. ábra A hírszolgáltatás működésének vázlata



A számnégyes tagjainak a jelentései sorrendben:

- 1) Ha még nem történt konvertálás az 1. típusú felhasználók részére, akkor értéke 0, különben 1.
- 2) Ha még nem történt konvertálás a 2. típusú felhasználók részére, akkor értéke 0, különben 1.
- 3) Itt számoljuk a 3.-5. típusú felhasználók számára elvégzett konvertálásokat. Ez a tag 0 és 3 közötti értékeket vehet fel.
Itt fontos megjegyezni, hogy nem számít az, hogy milyen sorrendben érkeztek a 3.-5. típusú felhasználók, mivel gyakoriságaik egyformák.
- 4) Itt számoljuk a 6.-10. típusú felhasználók számára elvégzett konvertálásokat.
Ez a tag 0 és 5 közötti értékeket vehet fel.
Itt szintén azért elegendő egy számláló, mivel a 6.-10. típusú felhasználók gyakoriságai megegyeznek, és ezért itt sem számít, hogy milyen sorrendben érkeztek.

Az állapotok száma $2 \cdot 2 \cdot 4 \cdot 6 = 96$.

Az állapot átmenet csak olyan számnégyesek között lehetséges, ahol csak egyetlen tag különbözik eggyel. A kiinduló állapot a $(0,0,0,0)$, vagyis ekkor még a tároló üres. Az állapotátmeneteket a $\mathbf{D}_1$ 96x96-os mátrixban adjuk meg. A $\mathbf{D}_1$ mátrix elemeit az alábbi módon adjuk meg.

Ha a mátrixelemnek megfelelő állapotátmenet az 1. típusú felhasználók részére történő konvertálást ír le, akkor a mátrixelem értéke $8 \cdot 0.21 = 1.68$, ahogyan azt az 1. táblázatban is láthatjuk az „Intenzitás” oszlopban. A 2. típusú konvertálási igények beérkezéséhez tartozó állapotátmenetek intenzitása 1.6.

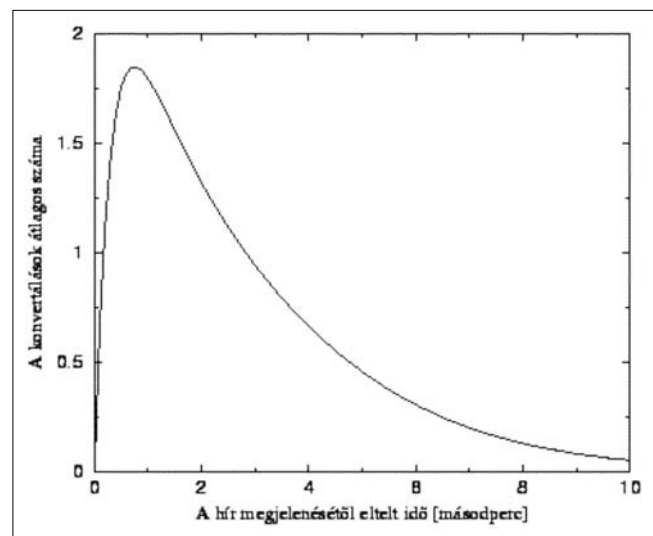
A 3.-5. típusú felhasználók esete egy kicsit különbözik az előzőektől. Ezek érkezési intenzitása egyforma, ezért az olyan átmeneteknél, ahol az állapotleíró számnégyes 3. tagja változik, elegendő csupán annyit számon tartani, hogy eddig hány konvertálást végeztünk el a 3.-5. típusok számára. Az induláskor ugyanis $3 \cdot 1.04 = 3.12$ intenzitással jöhet kérés a 3.-5. típusok valamelyikétől. Az első kérés után azonban már csak a megmaradó két típustól jöhet kérés, tehát az intenzitás lecsökken $2 \cdot 1.04 = 2.08$ -ra, míg ha már kettő típust kiszolgáltunk, akkor a megmaradó 3.-5. típusú kérés érkezési intenzitása 1.04 lesz.

A 6.-10. típusú kéréseket leíró állapot átmenetekre ugyanazok vonatkoznak, mint a 3.-5. típusúak kéréseire azzal a különbséggel, hogy a kezdeti intenzitás $5 \cdot 0.32 = 1.6$ lesz, és ez csökken le 0.32-re, amikor már csak 1 hiányzó típus marad.

Ekkor a hír megjelenésétől kezdve folyamatosan érkeznek a kérések a kiszolgálóhoz. Az első másodpercekben sorra érkeznek be olyan böngészőkről kérések, amelyekre a hírek még nem létezik konvertált változata. Ekkor van a konvertálóknak a legtöbb dolguk. Amikor a leggyakrabban előforduló beállításoknak megfelelő változatok már elkészültek, akkor egyre kevesebb konvertálásra van szükség. Ha 10 konvertáló áll rendelkezésre, akkor azonnal ki tudunk szolgálni minden igényt. Azonban látni fogjuk, hogy ennél kevesebb konvertáló is elegendő a feladat elvégzéséhez.

Miután megkonstruáltuk a $\mathbf{D}_0$ és $\mathbf{D}_1$ mátrixokat, kiszámítjuk a (6) egyenlethez szükséges $\mathbf{V}$ mátrixot, és ezzel megkapjuk a differenciálegyenlet-rendszer homogén részének a megoldását is. Megjegyezzük, hogy a $\mathbf{D} = \mathbf{D}_0 + \mathbf{D}_1$ mátrix minden sajátértéke valós és egyszeres. A partikuláris megoldás felírásához először kiszámítjuk a $H(t)$ függvényből kiindulva a (7) egyenletben szereplő $f_k(t)$ függvényeket, melyek (8)-as alakúak. Ezek után felírhatjuk a $\hat{c}_m(t)$ függvényeket (10) alapján. A jelen példában a $\hat{c}_m(t)$ függvények felírására nincs szükség. Végül pedig a momentum függvények a homogén rész és a partikuláris megoldás összegeként állnak elő (15)-ben.

A 2. ábrán az egyszerre foglalt konvertálók átlagos számának időbeli változását láthatjuk. Az is látható az ábrán, hogy a legnagyobb foglaltság a hír megjelenése után körülbelül 0.75 másodperccel következik be.



2. ábra

A foglalt konvertálók átlagos számának alakulása röviddel egy új hír megjelenése után

Annak érdekében, hogy egy olyan korlátot állapíthassunk meg, melynél több foglalt konvertáló előfordulásának a valószínűsége kicsi, kiértékeljük a párhuzamosan futó konvertálások számának momentumait az új hír megjelenése után 0.75 másodperccel. Ekkor azt feltételezzük, hogy tetszőleges számú konvertáló áll a rendelkezésünkre. A 3. táblázat tartalmazza az eredményeket:

	Faktorális momentum	Momentum
1.	1.848	1.848
2.	3.063	4.911
3.	4.499	15.537
4.	5.763	56.048
5.	6.307	224.208
6.	5.734	976.641
7.	4.157	4573.540
8.	2.253	22805.388
9.	0.812	120181.397
10.	0.146	665327.954

3. táblázat

Mivel legfeljebb 10 párhuzamos konvertálás fordulhat elő, így az első 10 momentum segítségével egyszerűen kiszámítható a párhuzamos konvertálások számának az eloszlása egy Vandermonde típusú mátrix segítségével [10]. A 4. táblázatban azt láthatjuk, hogy milyen valószínűséggel lép túl a párhuzamos konvertálások száma (N) különböző korlátokat, amelyeket n -el jelöltünk.

4. táblázat
A párhuzamosan futó konvertálások számának eloszlása a hír megjelenése után 0.75 másodperccel

n	$P(N > n)$
0.	0.871
1.	0.577
2.	0.276
3.	0.095
4.	0.023
5.	0.004
6.	$5 \cdot 10^{-4}$
7.	$4 \cdot 10^{-5}$
8.	$2 \cdot 10^{-6}$
9.	$3 \cdot 10^{-8}$
10.	0

Ha a tervezés során 1 ezrelék alá szeretnénk csökkenteni annak a valószínűségét, hogy a foglalt konvertálók, illetve párhuzamosan futó programszálak száma meghaladja a rendelkezésre álló kapacitásokat, a táblázatból megállapíthatjuk, hogy 10 konvertáló helyett elegendő csupán 5 konvertáló, illetve ennek megfelelő processzorleltésmény használata, mivel annak valószínűsége, hogy több, mint 5-re lenne szükség $5 \cdot 10^{-4}$.

5. Összefoglalás

Jelen cikkben egy végtelen számú kiszolgálóval rendelkező (Kendall jelölésével MAP/PH/ ∞) sorbanállási rendszer tranziens viselkedését vizsgáltuk analitikus eszközök segítségével. A sorbanállási rendszer momentumaira egy differenciálegyenlet-rendszert adtunk meg, melynek megtaláltuk az időfüggő megoldását. A sorban tartózkodó igények számának átlagát, szórásnégyzetét illetve magasabb momentumait felhasználva megkaptuk a sorban tartózkodó igények számának az eloszlását is, szintén az eltelt idő függvényében.

A számítási módszer hasznosságát egy gyakorlati példán keresztül mutattuk be. A példa egy internetes hírszolgáltatás lehetséges működéséről szól. A hírszolgáltató minden hírt testre szab a különböző beállításokkal rendelkező előfizetői csoportok számára. Ez a folyamat azonban időigényes, és előfordulhat, hogy egy időben több előfizetőnek is konvertálni kell a hírt. Az analitikus modell segítségével kiszámítjuk az egy időben szükséges konvertálások számát, és egy felső korlátot adunk erre, amelyet a rendszer nagy valószínűséggel nem lép túl. A korlát pedig megadja a szolgáltatás zavartalan működéséhez szükséges számítási kapacitás nagyságát.

Mind az elméleti, mind a gyakorlati példához kapcsolódóan vannak további kutatási lehetőségek.

A momentumok kiszámításakor feltételeztük, hogy a rendszer az indulás pillanatában nem szolgál ki egyetlen igényt sem. Ezt a (2) illetve (3) egyenletekkel leírt differenciálegyenlet-rendszer kezdeti feltételeinél rögzítettük. Feltételezhetjük azonban azt is, hogy induláskor a rendszerben véletlen számú igény van. Az igények

számának eloszlása ismeretében meghatározhatjuk a faktoriális momentumokat, és ezek a momentumok lesznek a differenciálegyenlet-rendszer kezdeti feltételei.

A gyakorlati példában felmerül az a probléma, hogy nagyszámú és sokféle lehetséges beállítási típusnál a Markov érkezési folyamat állapottere igen nagy lesz, és emiatt a számítás futási ideje jelentős mértékben megnő. A példában megjelenő Markov érkezési folyamat azonban tulajdonképpen hasonló struktúrájú független Markov érkezési folyamatok szuperponáltja, így a számítások egy elemi esetre egyszerűsíthetők le, mely szintén analitikusan kezelhető, és ezzel a számítási igény csökkenthető.

További problémát kaphatunk a gyakorlati példa módosításával. Tegyük fel, hogy egyetlen konvertáló eszközünk van, és a konvertálási kéréseket egy FIFO sorban tároljuk, amíg kiszolgálásra nem kerülnek. A probléma ekkor egy véges számú kiszolgálóval rendelkező sorbanállási rendszerre vezethető vissza. Ebben az esetben más matematikai eszközöket felhasználó analitikus vizsgálatokra van lehetőségünk.

6. Függelék

[6] alapján a következő differenciálegyenlet-rendszert adhatjuk meg a $\mu^{(K)}(t)$ vektorokra:

$$\frac{d}{dt} \mu^{(1)}(t) = \mathbf{D} \mu^{(1)}(t) + \{1 - H(t)\} \mathbf{D}_1 \mathbf{1}, \quad \mu^{(1)}(0) = 0, \quad (2)$$

$$\text{és } K \geq 2 \text{ esetén} \quad (3)$$

$$\frac{d}{dt} \mu^{(K)}(t) = \mathbf{D} \mu^{(K)}(t) + K \{1 - H(t)\} \mathbf{D}_1 \mu^{(K-1)}(t), \quad \mu^{(K)}(0) = 0.$$

Itt a kiszolgálási idő eloszlásfüggvénye $H(t)$, valamint $\mathbf{1}$ az 1-ekből álló M dimenziós oszlopvektor. A bevezetőben megjegyeztük, hogy a vizsgálatunk tárgyát képező rendszerrel [5]-ben egy általánosabb esetet tárgyalnak. [5]-ben a rendszer időfüggő momentum-generáló függvényét közvetlenül adják meg, míg mi a differenciálegyenlet-rendszert megoldva kapjuk meg momentumok függvényeit az időfüggő generátorfüggvény felhasználása nélkül.

Látható, hogy a (2) egyenlet megoldásából a magasabb momentumok iteratív módon számolhatóak a (3) egyenlet felhasználásával.

Bár nem tartozik ezen cikk szűken vett témakörébe, megjegyezzük, hogy faktoriális momentumok aszimptotikus viselkedésére a következő tételt mondják ki [6]-ban:

1. Tétel: Ha az igények kiszolgálási idejének az átlaga véges ($\mu = \int_0^\infty (1 - H(t)) dt < \infty$), akkor a $\mu^{(K)}(t)$ vektor egy $c_K e = (c_K, \dots, c_K)$ vektorhoz konvergál $K \geq 1, t \rightarrow \infty$ esetén.

A tételből látható tehát, hogy véges várható értékű kiszolgálási idő esetén a faktoriális momentumok $t \rightarrow \infty$ határértéke független a kezdeti állapottól.

[6] cikk szerzői a (2) és (3) differenciálegyenletekre numerikus megoldást javasoltak. Megkaphatjuk azonban az egzakt megoldást is, ha feltételezzük, hogy az

igények kiszolgálási ideje PH eloszlású, vagyis felírható különböző paraméterű Erlang eloszlások kevertjeként. Azaz feltételezzük, hogy $H(t)$ a következő alakban írható fel:

$$H(t) = 1 - \sum_{i=1}^I e^{-\beta_i t} \sum_{j=0}^{J_i} \alpha_{i,j} t^j, \quad (4)$$

ahol $I \geq 1$, $J_i \geq 0$, $\beta_i \geq 0$.

A megoldás az elsőrendű inhomogén lineáris differenciálegyenletek szokásos megoldási menetét követi (részletek: [7,8]). Az (1) egyenlettel definiált faktoriális momentumok felírhatóak az egyenlet homogén részének általános megoldása és az inhomogén rész partikuláris megoldása összegeként. Mivel a $\mathbf{D}_0$ és $\mathbf{D}_1$ mátrixok konstansok, a homogén rész megoldása az összes momentum esetében ugyanaz. A K -ik momentumra vonatkozó differenciálegyenlet inhomogén részének megoldása viszont a $(K-1)$ -ik momentumtól függ ($K \geq 2$). Először kiszámítjuk a megoldást $\mu^{(1)}(t)$ -re, majd a bemutatott módszert alkalmazva kaphatjuk meg a megoldást a magasabb momentumokra.

A (2) egyenlet megoldásának első lépése a homogén rész megoldása:

$$\frac{d}{dt} \mu^{(1)(H)}(t) = \mathbf{D} \mu^{(1)(H)}(t). \quad (5)$$

A megoldás exponenciális függvények lineáris kombinációjaként adható meg:

$$\mu_n^{(1)(H)}(t) = \sum_{m=1}^M [\mathbf{V}]_{n,m} c_m e^{r_m t}. \quad (6)$$

Itt az $M \times M$ -es $\mathbf{V}$ mátrix a $\mathbf{D}$ mátrixot Jordan alakúra hozó hasonlósági transzformáció mátrixa (oszlopai az $\mathbf{D}$ mátrix általánosított jobboldali sajátvektorai). Ha a $\mathbf{D}$ mátrix minden sajátértéke egyszeres, akkor a $\mathbf{V}$ mátrix oszlopai a $\mathbf{D}$ mátrix sajátvektorai, az r_m számok a sajátvektorokhoz tartozó sajátértékek. A c_m -ek pedig meghatározandó ismeretlenek.

A $\mathbf{D}$ mátrix sztochasztikusságából azonnal következik az alábbi tétel:

2. Tétel: A $\mathbf{D}$ mátrixnak a 0 az egyik sajátértéke és az összes sajátértékének a valós része nem pozitív.

Mivel a 0 is egy sajátértéke $\mathbf{D}$ -nek, így a homogén rész megoldásában van egy tag, amelyik nem függ t -től. Ez a tag marad meg $t \rightarrow \infty$ esetén. Az általánosság megtartása mellett tegyük fel, hogy $r_1 = 0$.

A $\mathbf{D}$ valós együttthatós mátrix komplex sajátértékei konjugált párokat alkotnak. Emiatt azonban a továbbiakban megjelenő egyenletek komplex együttthatói is konjugált párokba rendezhetőek, így a végeredményben a képzetes részek kiegyenlítik egymást. Ennek részletes ellenőrzésével ebben a cikkben nem foglalkozunk.

Abban az esetben, ha valamelyik sajátérték többszörös, akkor a $\mu_n^{(1)(H)}(t)$ függvény felírásában az $e^{r_m t}$ együttthatói között t szerinti polinom függvények is megjelennek,

$$\mu_n^{(1)(H)}(t) = \sum_m [\mathbf{V}]_{n,m} \sum_j c_{m,j} t^j e^{r_m t},$$

de a megoldási módszer lényegében változatlan marad.

A differenciálegyenlet-rendszer általános megoldásának a felírásához szükség van egy partikuláris megoldásra is. Ezt az állandók variálásának módszerével kaphatjuk meg, azaz tekintsük a homogén megoldás c_m együttthatóit t szerinti függvényeknek.

$$\sum_{m=1}^M [\mathbf{V}]_{n,m} \frac{d}{dt} c_m(t) e^{r_m t} = f_k(t), \quad (7)$$

ahol

$$f_k(t) = (1 - H(t)) \sum_{i=1}^M [\mathbf{D}_1]_{k,i}.$$

Az $f_k(t)$ függvény általános formája a következő:

$$f_k(t) = \sum_{i=1}^{I^{(K)}} e^{-\gamma_{i,k} t} \sum_{j=0}^{J_i^{(K)}} \alpha_{i,j}^{(K)} t^j. \quad (8)$$

Ha a (2) egyenlet partikuláris megoldását keressük, akkor az $f_k(t)$ (8)-beli kifejezésében $\gamma_{i,k} = \beta_i$ minden k -ra. Megjegyezzük, hogy a magasabb momentum függvények egyre több tagból állnak. Amint azt látni lehet a (3) egyenlet inhomogén részén, a $\gamma_{i,k}$ exponensek és a β_j exponensek összegei adják meg a $\gamma_{l,k+1}$ exponenseket.

A $c_m(t)$ ($1 \leq m \leq M$) függvények deriváltjait kifejezhetjük az $f_k(t)$ függvényekkel és a $\mathbf{V}$ mátrix inverzével. Jelölje $\mathbf{A} = \mathbf{V}^{-1}$ a $\mathbf{V}$ mátrix inverzét! Ekkor a (7) egyenletet megoldva $\frac{d}{dt} c_m(t)$ -re kapjuk a következő kifejezést:

$$\begin{aligned} \frac{d}{dt} c_m(t) &= \sum_{k=1}^M [\mathbf{A}]_{m,k} f_k(t) e^{-r_m t} = \\ &= \sum_{k=1}^M [\mathbf{A}]_{m,k} \sum_{i=1}^{I^{(K)}} e^{-(\gamma_{i,k} + r_m)t} \sum_{j=0}^{J_i^{(K)}} \alpha_{i,j}^{(K)} t^j, \quad 1 \leq m \leq M \end{aligned} \quad (9)$$

A $c_m(t)$ függvényt a (9) egyenlet integrálásával kaphatjuk meg. Itt az integrálás során az összeg különböző t -től függő tagjait a következő séma szerint kaphatjuk meg:

$$\int e^{\delta t} t^j dt = -\frac{j!}{\delta^{j+1}} \sum_{l=0}^j \frac{e^{-\delta t} \delta^l t^l}{l!} + C, \quad \text{ha } \delta \neq 0,$$

és

$$\int t^j dt = \frac{t^{j+1}}{j+1} + C, \quad \text{ha } \delta = 0.$$

Egy partikuláris megoldásnál elegendő az integrálás során kapott konstans 0-nak venni, a $c_m(t)$ függvényeket a következő alakban írhatjuk fel (10):

$$\hat{c}_m(t) e^{r_m t} = \sum_{k=1}^M \sum_{i=1}^{I^{(K)}} \sum_{j=0}^{J_i^{(K)}} [\mathbf{A}]_{m,k} \alpha_{i,j}^{(K)} \frac{j!}{(\delta_{i,k}^{(m)})^{j+1}} \frac{e^{-\gamma_{i,k} t} (\delta_{i,k}^{(m)})^j}{l!}, \quad 1 \leq m \leq M,$$

itt $\delta_{i,k}^{(m)} = \gamma_{i,k} + r_m$. Ha valamely i és k indexekre $\delta_{i,k}^{(m)} = 0$, akkor az ehhez tartozó $c_m(t)$ függvények a következő alakúak:

$$\bar{c}_m(t) e^{r_m t} = \sum_{k=1}^M \sum_{i=1}^{I^{(K)}} \sum_{j=0}^{J_i^{(K)}} [\mathbf{A}]_{m,k} \alpha_{i,j}^{(K)} \frac{t^{j+1}}{j+1} e^{r_m t}. \quad (11)$$

Ekkor a partikuláris megoldás n -ik komponensét $\hat{c}_m(t)$ és $\bar{c}_m(t)$ függvényekből kaphatjuk meg:

$$\hat{\mu}_n^{(1)(P)}(t) = \sum_{m=1}^M [\mathbf{V}]_{n,m} \hat{c}_m(t) e^{r_m t}, \quad (12)$$

$$\bar{\mu}_n^{(1)(P)}(t) = \sum_{m=1}^M [\mathbf{V}]_{n,m} \bar{c}_m(t) e^{r_m t}, \quad (13)$$

és a teljes partikuláris megoldás pedig

$$\mu_n^{(1)(P)}(t) = \hat{\mu}_n^{(1)(P)}(t) + \bar{\mu}_n^{(1)(P)}(t), \quad (14)$$

ahol $1 \leq n \leq M$.

Megjegyezzük, hogy mind $\hat{\mu}_n^{(1)(P)}(t)$, mind $\bar{\mu}_n^{(1)(P)}(t)$, 0-hoz konvergál $t \rightarrow \infty$ esetén.

Miután megkaptuk az első momentumra vonatkozó differenciálegyenlet homogén részének a megoldását ($\mu_n^{(1)(H)}(t)$) és egy partikuláris megoldását ($\mu_n^{(1)(P)}(t)$) is, így az általános megoldást szintén megkaptuk.

$$\mu_n^{(1)}(t) = \mu_n^{(1)(H)}(t) + \mu_n^{(1)(P)}(t). \quad (15)$$

A (6) egyenletben található c_m együtthatók meghatározásához a (2) egyenlet kezdeti feltételét használhatjuk, azaz

$$\mu_n^{(1)}(0) = \mu_n^{(1)(P)}(0) + \sum_{m=1}^M [\mathbf{V}]_{n,m} c_m = 0.$$

Jelölje $c = [c_1, c_2, \dots, c_m]$ az együtthatók vektorát. Ekkor ismételtén az $\mathbf{A} = \mathbf{V}^{-1}$ jelölést használva a konstansokat a

$$c = -\mathbf{A} \cdot \mu^{(1)(P)}(0) \quad (16)$$

egyenletből kaphatjuk meg.

Mivel az egyenletrendszer általános alakjában az $(n+2)$ -dik momentumra vonatkozó egyenletet ugyanúgy vezettük vissza az $(n+1)$ -dik momentum függvényére, mint ahogyan az $(n+1)$ -dik momentumra vonatkozó egyenletet az n -dik momentum függvényére, a momentum függvényeket iteratív úton kaphatjuk meg a (3) egyenlet jobb oldalába való behelyettesítéssel.

Megjegyezzük, hogy a momentum függvények tulajdonképpen a faktoriális momentumokat adják meg. Az eredeti momentumokat a másodfajú Stirling számok ($\Delta_{i,n}$) segítségével kaphatjuk meg:

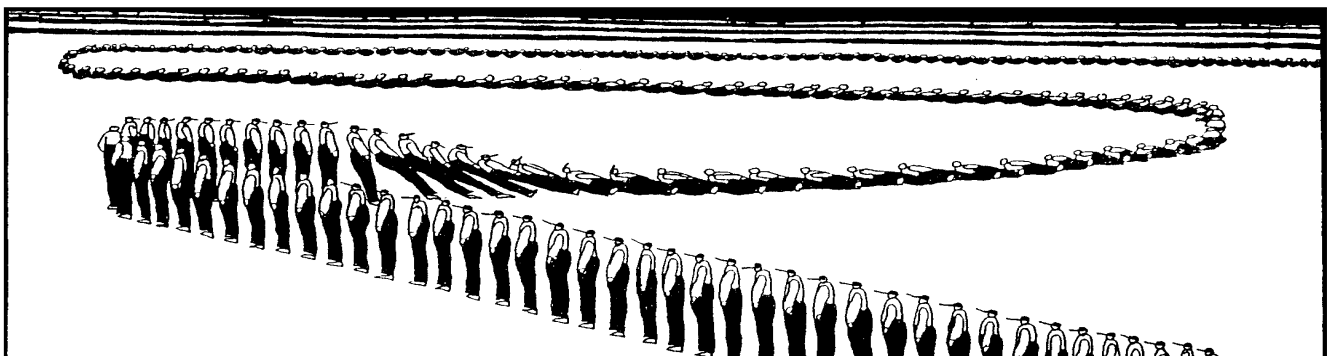
$$k^n = \sum_{i=1}^n \Delta_{i,n} k(k-1)\dots(k-i), k \geq n.$$

$\Delta_{i,n}$ meghatározására pedig találhatunk egy rekurzív kifejezést [9] 5.2 fejezetében.

A gyakorlatban az iterációs lépések során csak szorzást és összeadást végzünk. Ekkor az $(n+1)$ -ik momentum függvényben szereplő $\gamma_{i,k}$ és $\alpha_{i,j}^{(k)}$ konstansokat állítjuk elő az n momentum függvény ismeretében. Sajnos tetszőlegesen nagy momentum függvényeket nem tudunk meghatározni, mivel az iterációs lépések során a $\gamma_{i,k}$ és $\alpha_{i,j}^{(k)}$ konstansok száma nem korlátos.

Irodalom

- [1] G. Latouche, V. Ramaswami: „Introduction to Matrix Analytic Methods in Stochastic Modeling”, ASA-SIAM Series on Statistics and Applied Probability, 1999.
- [2] W. Fischer, K. Meier-Hellstern: „The Markov-Modulated Poisson process (MMPP) cookbook”, Performance Evaluation 18, pp.149–171., 1993.
- [3] N. Hohn, D. Veitch, P. Abry: „Does fractal scaling at the IP level depend on TCP flow arrival processes?”, ACM SIGCOMM Internet Measurement Workshop (IMW-2002), Marseille, France, 2002.
- [4] A. Horváth, M. Telek: „Approximating heavy tailed behaviour with phase type distributions”, Advances in algorithmic methods for stochastic models, MAM3 (G. Latouche and P. Taylor, eds.), pp.191–214., Notable Publications Inc., 2000.
- [5] Hiroyuki Masuyama, Tetsuya Takine: „Analysis of an Infinite-Server Queue with Batch Markovian Arrival Streams”, Queueing Systems, vol. 42, No.3, pp.269–296., 2002.
- [6] V. Ramaswami, M. F. Neuts: „Some explicit formulas and computational methods for infinite-server queues with phase-type arrival”, Journal of Applied Probability 17, pp.498–514., 1980.
- [7] Ordinary Differential Equation System with Constant Coefficients, MathWorld, <http://mathworld.wolfram.com/>
- [8] I. N. Bronstein, K. A. Semendjajew, G. Musion, H. Mühlig: Taschenbuch der Mathematik, Verlag Harri Deutsch, Frankfurt am Main, 1999.
- [9] D. J. Daley, D. Vere-Jones: „An Introduction to the Theory of Point Processes, Volume I: Elementary Theory and Methods”, Second Edition, Springer, 2003.
- [10] Rózsa Pál: „Lineáris algebra és alkalmazásai”, Műszaki Könyvkiadó, 1974.



Dokumentumkategorizálás távközlési alkalmazásban

BENKŐ BORBÁLA KATALIN, PÁNDI ZSOLT

{bbenko, pandi}@hit.bme.hu

Kulcsszavak: hibatűró rendszerek, trouble ticket, dokumentumkategorizálás, szövegfeldolgozás

Cikkünk kiindulópontját egy konkrét probléma, egyes trouble ticketing rendszerek tökéletlensége adta. A vizsgált trouble ticketing rendszerek – noha igen hasznosak lehetnének a hálózat üzemeltetőjének és a hibatűró hálózatokkal kapcsolatos kutatások szempontjából –, nem használhatók elég jól, mert a ticketek kategóriamezője csak kevés információt tartalmaz. Célunk az automatikus kategória detekció a ticketben található szöveges leírás alapján. Ehhez a dokumentumkategorizálás széles eszköztárát vonultatjuk fel. A konkrét probléma megoldásán túl áttekintő képet adunk a dokumentumkategorizálás jelenlegi helyzetéről, a szokásos módszerekről, a rendelkezésre álló eszközökről is. Ez különösen hasznos lehet olyan olvasóinknak, akik szintén szövegfeldolgozási feladatokkal (például szöveges üzenetek, ügyfélszolgálat levelek stb) szembesülnek.

1. Bevezetés

A hibatűró hálózatokkal kapcsolatos kutatások egyik alapvető problémája, hogy a valós rendszerek rendelkezésre állását és megbízhatóságát jellemző paraméterekhez igen nehéz, gyakran lehetetlen hozzáférni. A berendezéseket és kábeleket gyártó cégeknek nem áll érdekükben az ilyen, verseny szempontjából érzékeny adatok nyilvánosságra hozatala, a szolgáltatók pedig vagy nem rendelkeznek az erre vonatkozó megfelelő statisztikai adatokkal, vagy ugyancsak üzleti titokként kezelik ezt az információt.

Távközlő hálózatok üzemeltetése során úgynevezett trouble ticket rendszerekben rögzítik az előforduló hálózati hibákat, azok (valószínűsíthető) okát, a javításhoz szükséges időt, és további kapcsolódó információkat. Néhány nyilvános kutatási célú nemzeti hálózat az Interneten is hozzáférhetővé teszi hálózati hibabejelentő, más néven trouble ticketing rendszerét. Bár garantáltan teljes képet nem, valós rendszerek üzemeltetéséből származó értékes adatokat mindenképpen szolgáltathat az ilyen jellegű adatbázisok szisztematikus feldolgozása.

A hálózati adatok statisztikai vizsgálatának kérdése a hálózatüzemeltetők szempontjából is felvetődik. Lényeges különbség a nyilvános adatbázisok és a távközlési szolgáltatók éles adatbázisai között tulajdonképp nincs, csupán annyi, hogy az üzemeltetőnek érdeke az adatok pontossága és teljessége.

Sajnos a jelenleg alkalmazott, publikusan is hozzáférhető adatokat tartalmazó rendszerek [1,2] néhány hiányossága miatt a statisztikai elemzés az ember által írt folyószöveg értelmezését teszi szükségessé. Valószínűleg az üzemeltetők által alkalmazott rendszerek hiányosságai is hasonlóak ahhoz, amelyek az Interneten hozzáférhetőek.

Ennek a nagy kihívást jelentő feladatnak a megoldásához, azaz az ember alkotta szövegek elemzésé-

hez ad segítséget a dokumentumkategorizálás, amely nyelvészeti és matematikai alapokra épülő módszereket használ. Cikkünkben bemutatjuk a trouble ticket rendszerekben fennálló problémát; áttekintő képet adunk a dokumentumkategorizálás eszközeiről; végül javaslatot teszünk, hogy a rendelkezésre álló eszközökből hogyan érdemes rendszert építeni a trouble ticketek elemzéséhez.

2. Trouble ticket

Egy trouble ticket rendszer bejegyzései tipikusan az 1. ábrán látható releváns információkat tartalmazzák.

A ticketek tartalmának értelmezéséhez célszerű vázlatosan áttekinteni egy ticket életét, amely lehetőséget ad arra is, hogy az adathalmazból levonható következtetéseket megalapozzuk. Egy ticket létrehozása vagy egy észlelt hibaeseménynek, vagy pedig egy jövőbeli, várhatóan a hálózatra hatással levő (például kockázatonövelő) eseményről kapott információnak köszönhető. A megnyitás időpontja értelmezéstől függően megegyezhet az adott esemény észlelésének vagy aktuálisá válásának idejével, de lehet a jövőbeli információ rögzítésének ideje is. A lezárás ideje a hibaesemény el-

1. ábra Trouble ticket tipikus tartalma

Trouble ticket
▪ Ticket azonosító ▪ Megnyitás időpontja ▪ Státusz (nyitott/zárt) ▪ Kategória / Rövid jellemzés ▪ Lezárás időpontja ▪ Részletes szöveges leírás (a tickethez kapcsolódó események leírása kötetlen formában)

hárítási időpontja, a bejegyzés aktualitásának elvesztése, vagy a nyitva felejtett ticketek ellenőrzésének időpontja lehet tipikusan.

Egy ticket létrehozásakor nyitott állapotba kerül, jelezve azt, hogy az üzemeltető részéről valamilyen reakciót igényel (beavatkozást vagy akár fokozott készséget). A ticket lezárása idején a ticket zárt állapotba kerül. Ennek az attribútumnak a jelentősége, hogy az üzemeltető a rendszerben a nyitott ticketekre szűrve gyorsan tud információt szerezni az aktuális problémákról.

A ticket kategóriája általában a feldolgozást megkönnyítendő, előre definiált hibaosztályok egyikébe történő manuális besorolás eredménye, míg a rövid jellemzés egy lényegretörő, átfogó listában is megjeleníthető emlékeztető. E két mező funkciója nem fedi egymást teljesen, továbbá megemlíthetjük azt is, hogy a kategóriák megfelelő meghatározása kulcsfontosságú lehet a hálózatról generálandó statisztikák számára.

A tickethez kapcsolódó események leírása adja a legtöbb és a legpontosabb információt, ám sajnos csak az ember számára. Gépiesített kategorizálásnál igen komoly feladatot jelenthet a folyószöveg bejegyzések értelmezése, főleg, ha formai szabályoktól mentesek a bejegyzések, és a használt nyelv sem mindig ugyanaz.

Problémánk, hogy a ticketekhez kapcsolódó eseményekről a kategória mező tartalma alapján az [1] hálózat esetében csak *outage* vagy *scheduled maintenance* besorolás áll rendelkezésre, a [2] hálózathál pedig a kategória mező helyett rövid szöveges jellemzés szerepel. Ahhoz, hogy ennél pontosabb kategória besorolást kapjunk, dokumentumkategorizálásra van szükség.

3. Dokumentumkategorizálás

A dokumentumkategorizálás az 1960-as évek óta kutatott, de még mindig aktuális tudományterület [3], mely alapvetően két feladattal foglalkozik:

- **Dokumentumklaszterezés:** a dokumentumhalmazt hasonlósági alapon klaszterekbe soroljuk. A klaszterek automatikusan – és a dokumentumoktól függően dinamikusan – alakulnak ki. Bemenő adat a klaszter sugara (vágási határ) vagy a kért klaszterek száma.

- **Dokumentumkategorizálás:** a dokumentumhalmaz elemeit előre definiált kategóriákhoz soroljuk. (Zavaró, de sajnos ennek a feladatnak a neve megegyezik a tématerület nevével.)

Ipari alkalmazásoknál tipikusan kategorizálási feladatról van szó, ám a jól működő kategorizálás alapfeltétele a kategóriák – nem csak jó, hanem – tökéletes kialakítása. Ezért gyakori, hogy a kategóriák kialakítása előtt egy reprezentatív részhalmazon klaszterezést végeznek, és az itt tapasztaltak alapján alakítják ki – esetleg hierarchikusan – a kategóriákat. Így elkerülhető, hogy túl általános vagy túl specifikus kategóriák alakuljanak ki, hisz nyilván nem hordozna túl sok információt, ha a dokumentumok 80%-a ugyanabba a kategóriába tartozna.

Jelenleg is aktívan kutatott terület a *szcenárió-hozzárendelés*, ahol a dokumentumot valamilyen előre definiált forgatókönyvhöz soroljuk (például kábelhiba lépett fel, majd ki lett javítva).

Mivel még a dokumentumok teljeskörű automatikus megértésétől (szemantika stb.) nagy távol állunk, a dokumentumfeldolgozók egyszerűsített modellt használnak. A két fő *dokumentummodell* a következő:

- **Szavak halmaza** (set of words) **modell**. A dokumentumot szavak halmazának tekinti, a szavak közti relációt nem veszi figyelembe. Egyszerű, gyors és meglepően hatékony [4].

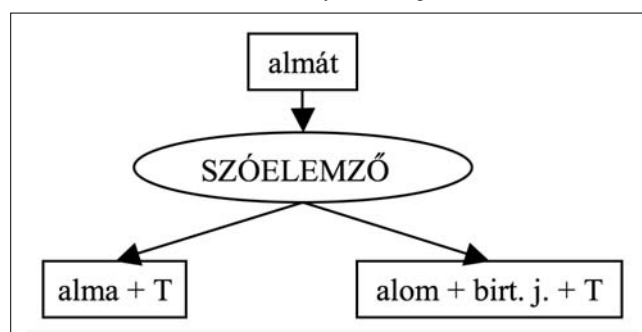
- **Nyelvészeti indíttatású** (linguistically motivated) **modellek**. Figyelembe veszi a szavak közti nyelvtani relációkat (alany-állítmány, tagadás, távoli kapcsolatok stb). Alapfeltétele egy jó nyelvtani elemző. Főleg szcenárió-hozzárendelésnél alkalmazzák, hisz itt nagy szükség van a viszonyok pontos azonosítására (megjavították vagy nem javították meg; ki okozta a kárt és ki hozta helyre) [5-7].

A gépi feldolgozás szempontjából alapvető különbség van a szeparáló és az agglutináló/flektáló nyelvek között. A szeparáló nyelvek (például angol) esetén a szóalak nem (illetve ritkán) változik, a nyelvtani reláció a szavak sorrendjében van kódolva. A dokumentum szavakra bontása közvetlen előállítja a szavak listáját. Az agglutináló vagy flektáló nyelveknél a felszíni szóalak és a szótó között akár nagyon nagy különbség is lehet. Ezért a dokumentum szótójának előállításához szótóvesítésre van szükség (morfológia). A szótóvesítés sajnos újabb bizonytalanságot visz a rendszerbe, hisz a szótó igen gyakran nem egyértelmű (2. ábra). A modern szóelemzők közlik a szótó variánsok becsült valószínűségét is.

Egyes alkalmazásokban célszerű *hierarchikus kategorizálást* használni, azaz a dokumentumot előbb egy főkategóriába sorolni, majd azon belül finomítani. Ez nem csak a kategorizálás sebességét növeli meg (például 5 fő- és 10-10 alkategória esetén nem 50, hanem csak 15 illeszkedési vizsgálatot kell elvégezni), hanem a felhasználhatóságot is, hiszen így lehetőség lesz mind általános (főkategória alapján), mind specifikus (alkategória alapján való) szűrő kérdések megfogalmazására.

A kategorizálás szemantikai támogatása még gyerekipőben jár, noha igen lényeges lenne. Néhány fontosabb kezdeményezés [8] és [9].

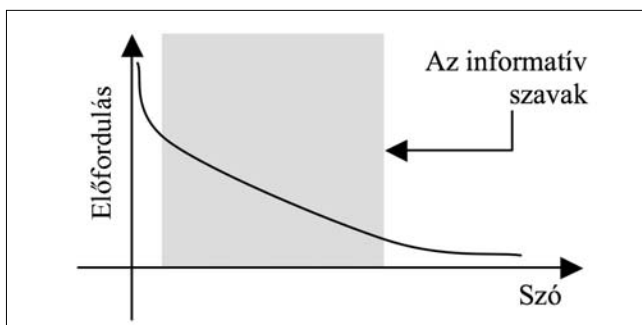
2. ábra
A szótóvesítés bizonytalanságot visz a rendszerbe



3.1. „Szavak halmaza” modell

A „szavak halmaza” modell arra a feltételezésre épül, hogy a kategorizálási feladat jól megoldható csupán a dokumentumban előforduló (illetve a hiányzó) szavak alapján. A „szó” fogalmát szokás rugalmasan kezelni, például tulajdonnév vagy tagadás esetén (egy szónak számít: „Harry Potter”, „doesn't repair”, „no problem”).

A Luhn-megfigyelés, vagy Luhn-szabály (3. ábra) szerint [10], ha a dokumentumban előforduló szavakat az előfordulások száma szerint sorbarendezzük, majd elhagyjuk a leggyakoribb és a legkevésbé gyakori szavakat, pont az informatív szóhalmaz marad meg. Heurisztikus magyarázat erre, hogy a nagyon gyakori szavak (névelők, kötőszavak) nem informatívak, a nagyon ritkák pedig valószínűleg nem kapcsolódnak szorosan a tárgyhoz.



3. ábra A Luhn-megfigyelés

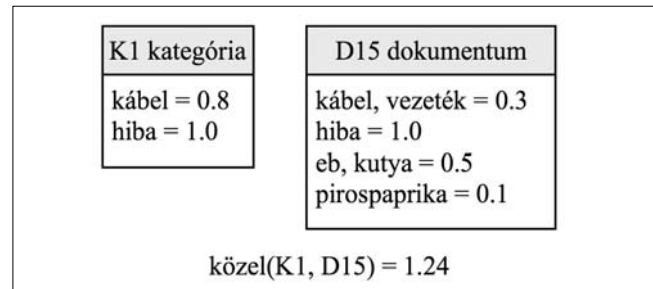
A Luhn-megfigyelés ebben a formájában nem használható a gyakorlatban (például egy 20 szavas dokumentum esetén nincs értelme gyakori és ritka szavakról beszélni). Ezért a szógyakoriságot egy nagyobb, reprezentatív korpuszon (szöveghalmazon) mérjük meg. Ez alapján készül el egy tiltólista (a vágandó szavak), illetve referencia szógyakoriságok. Fontos, hogy a mérési korpusz reprezentatív legyen (a feldolgozandó dokumentumokhoz hasonló elemekből álljon), hisz nyilván más mérési érték adódik általános szövegre, műszaki leírásra, dadaista versgyűjteményre stb.

A dokumentum és a kategória összerendelésére több különböző lehetőség van:

- **Bool-i összerendelés.** A dokumentumot akkor soroljuk a kategóriába, ha a kategória összes kulcsszavát tartalmazza. Ez több problémát is felvet, például mi történik, ha egy dokumentum az x és y kategória kulcsszavait egyaránt tartalmazza. Nem szokás használni.

- **Vektortér alapú összerendelés:** Egy vektort használ a dokumentum, illetve a kategória információtartalmának leírására. A vektor minden eleme egy szó súlyát jelenti (például az első elem a „kábelhiba” szóét). A kategória leíró vektorában a v_i elem a szó fontosságát jelenti; a dokumentum leíró vektorában pedig a szó jellemzőségét a dokumentumra (a szó előfordulási számán és a Luhn-szabályon alapuló normált érték). A dokumentum és a kategória közelségét egyszerű vektoriális szorzattal kapjuk meg (esetleg normálva). Lehetőség van tiltó szavak felvételére is (negatív súly a kategória-vektorban). A dokumentumot a hozzá legközelebb eső

kategóriába soroljuk. A vektortér-modellbe könnyen beépíthető *rejtett szemantika* is, hisz csak annyit kell tenni, hogy a vektor i . eleme nem egy szót, hanem egy szinonima halmazt jelöl. A gyakorlatban természetesen a vektor helyett ésszerűbb adatstruktúrákat használunk.



4. ábra Vektortér alapú összerendelés

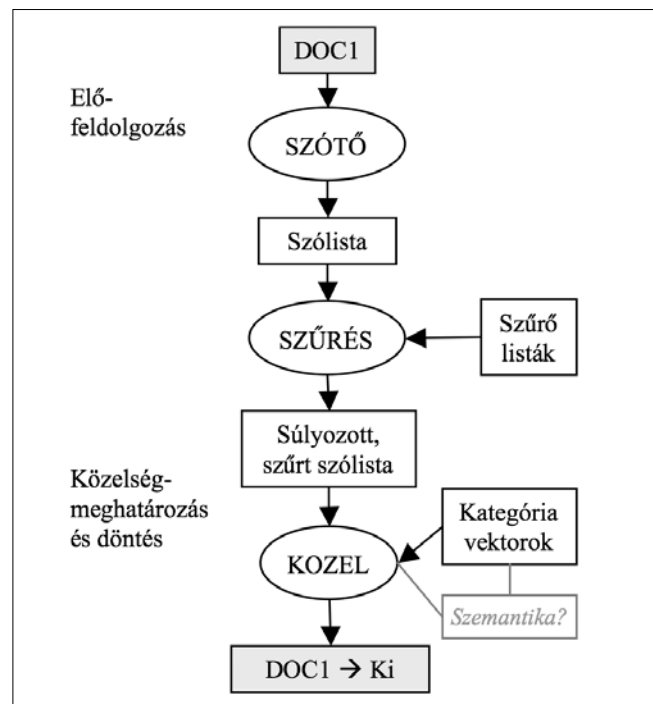
A 4. ábrán egy normálás nélküli, rejtett szemantikus összerendelés látható. A közelség a mindkét helyen előforduló szavak súlyának összege.

Egy szó alapú kategorizáló tipikus architektúrája a 5. ábrán látható. A dokumentumból először elő kell állítani a szavak listáját (általában szótövesítéssel), majd ezt a Luhn-szabály és/vagy szűrőlisták alapján redukálni és súlyozni. Ezután meghatározzuk a kategóriánkénti közelséget (lehetőleg valamilyen szemantikai támogatással), majd a dokumentumot a legközelebbi kategóriához rendeljük.

3.2. Nyelvészeti indíttatású modellek

A nyelvészeti indíttatású modellek alapötlete, hogy a kategorizáláshoz ne pusztán a szavakat, hanem a köztük levő relációt is vegyük figyelembe. A relációk azonosításához mondattani elemzésre van szükség, mely

5. ábra A szó alapú kategorizálás architektúrája



lehet mély elemzés (a teljes mondat elemzési fáját elő-állítjuk) vagy sekély elemzés (csak részlegesen elemzünk, például csak a névszói szerkezeteket).

A mondatelemzés alapján lehetőség van a mondat fő tartalmának beazonosítására. Például az esetek döntő többségében az alany, állítmány (és ha van, a tárgy) hordozza a fő információt. Ha elhagyjuk az egyéb mondatrészeket és a felesleges jelzőket és határozókat, pont a fő mondanivaló áll elő. Ezért gyakori, hogy a kategorizálás az alany-állítmány-tárgy hármas alapján történik, illetve kiegészítésként figyelembe veszik a hely és idő információt (ezek azonosítása sem mindig triviális, például „az előadás alatt” nem helyre, hanem időre utal).

Elvileg lehetséges lenne, hogy a kinyert információból egy háttérbeli szemantikai tudás (*domain knowledge*) segítségével további következtetéseket vonjunk le, de ez ma még nem realitás (a probléma elsősorban nem a műszaki oldallal, hanem a háttértudás leírásával, illetve a következtetések értelmezésével van). A jelenlegi (kísérleti!) rendszereknél a leginkább szinonima relációt használnak.

Nemrég zárult le egy érdekes kutatási projekt [11], ahol gazdasági hírek feldolgozása volt a cél. A rendszer az adás-vétel eseményt képes felismerni akár eladásról, akár vételről szól a hír.

A nyelvészeti indíttatású dokumentummodell komplexitása, mondattani elemző problémái, szemantikai háttérrendszer fejletlensége stb. miatt kategorizálásra olyan esetekben szokás használni, amikor a szó alapú modell nem hatékony. Ez elsősorban két esetet jelent:

- Homogén dokumentumhalmazt esetén. Itt a dokumentumok szókészlete kvázi-állandó (ugyanazok a szavak vannak az összes dokumentumban), ezért csupán a szavak alapján lehetetlen lenne kategorizálni.

- Ha a dokumentum szerkezete ezt indokolja (például sok tagadás).

4. Kategorizálás trouble ticketing rendszerekben

4.1. Milyen feladatot oldjunk meg?

Nézzük, milyen kategorizálási feladatnak van értelme trouble ticketing rendszerekben.

- **Klaszterezés.** A dokumentumhalmaz előzetes analíziséhez tökéletes eszköz, de éles rendszerben értelmetlen lenne használni: nem rendel nevet a klaszterekhez, így a humán operátorok számára nehezen kezelhető; valamint új dokumentumok érkezésekor újra kéne klaszterezni, és lehetséges, hogy az új körben alapvetően más csoportok alakulnak ki.

- **Kategorizálás.** Egy jól megalkotott, (lehetőleg hierarchikus) kategóriarendszer nagyban megkönnyítené az operátorok munkáját. Egyaránt hasznos lenne az aktuális beavatkozások felügyeleténél (például a kategória alapján megállapítható a hiba súlyossága, ezáltal a beavatkozás sürgőssége), és a hálózati statisztikák készítésénél.

- **Szcenárióhozzárendelés.** A trouble ticket tartalma időben változik, nyilvánvalóan egy folyamatot ír le, azaz kézenfekvő folyamat-lefutási sablonhoz (scenárióhoz) rendelni.

Bármely feladatot választjuk is, az biztos, hogy nem lehet teljes egészében az eredeti modellre hagyatkoznunk; figyelembe kell venni a trouble ticketek *nem időinvariáns* voltát (azaz hogy a ticketek tartalma és ezáltal a kategóriája az idő előrehaladtával változhat). A ticketeket minden új bejegyzés után újra kell kategorizálni.

4.2. Milyen modellt használunk?

Tekintve, hogy a trouble ticketing rendszer alapvetően angol nyelven áll rendelkezésre, és az angol első sorban szeparáló nyelv (ráadásul elég lokális természetű szerkezetekkel), elegendő a szóhalmaz modell. A tagadásokat és néhány egyéb fontos szófordulatot egy szónak tekintve elég jó modell adódik. A mondattani elemzés nagyon elbonyolítaná a rendszert, és minden valószínűség szerint kevesebb haszonnal járna, mint amennyi bizonytalanságot eredményez.

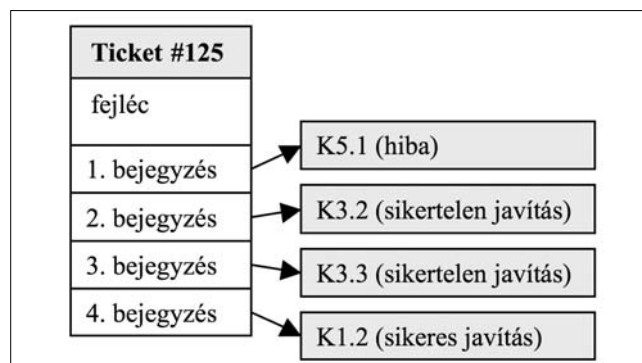
A másik kérdés, hogy pontosan mit akarunk kategorizálni, a trouble ticket bejegyzéseit, vagy az egész ticketet egyben. Az egyes bejegyzések kategorizálása mindenképp értelmesebb feladatnak tűnik. Később, egy újabb körben a bejegyzések kategóriái alapján már az egész ticketet is kategorizálhatjuk.

4.3. Rendszer vázlat

Első lépésként meg kell oldani a dokumentum (értsd: bejegyzések) szavakra bontását, majd elő kell állítani a szűrőlistákat, és a referencia szósúlyokat. A szómodell megalkotásához többféle segédeszközt is használhatunk (szűrők, named entity detector stb). Egy néhány ezer elemű dokumentumhalmaz alapján már nagy biztonsággal tudunk használható szűrőlistát készíteni és a referencia súlyokat számolni. Tekintve, hogy szakszövegről van szó, nem szerencsés a nyilvánosan hozzáférhető szósúlygyűjtemények használata.

A következő lépés a kategóriák megalkotása. Érdekes néhány próba klaszterezést futtatni (különböző vágási határokkal), majd az előálló klasztereket al-klaszterekre bontani. A klasztereket és alklasztereket áttanulmányozva képet kaphatunk a dokumentumok tartalmi eloszlásáról (vagy rossz esetben a szómodellünk, szű-

6. ábra Egy ticket bejegyzéseinek kategorizálása



rőlistáink hibáiról). A kategória leírások első változatát akár közvetlen is kinyerhetjük a klaszterekből (néhány egyszerű vektorművelet). A kinyert leírást emberek számára értelmezhető címkékkel kell ellátni, illetve szükség esetén tovább finomítani.

Ha a klaszterek nem az általunk kívánt elven alakulnak ki, bevethetünk néhány trükköt. Közismert, hogy a klaszterezés érzékeny az első néhány elemre; ezt a dokumentumok permutációjával kivédhetjük. A másik eszköz, hogy kiválasztunk néhány tipikus dokumentumot – minden tervezett kategóriából egyet-egyet – majd a kiválasztott dokumentumot megsokszorozzuk (például 800 példányban lemásoljuk). Ezzel tudjuk „kényszeríteni” a klaszterezőt, hogy a többi dokumentumot – ha lehetséges – ezen 800 elemű gócpontok köré rendezze.

Mivel már minden előfeltétel adott (szűrőlista, kategórialeírások), megkezdhetjük a dokumentumok (ticket bejegyzések) kategorizálását. A 6. ábrán egy ticket bejegyzéseinek kategorizálása látható. Példánkban a kategóriák 2 szintű hierarchiát alkotnak (például a K3 a sikertelen javítást jelenti, a K3.2 és a K3.3 pedig ennek két külön alosztala).

A kategorizált bejegyzések alapján előáll a ticket életút-leírása. Erre alapozva akár az egész tickethez rendelhetünk egy szcenárió mintát. Itt azonban az egyes bejegyzések sorrendje is lényeges.

Gyors és egyszerű egy véges automatát megalkotni, mely a bejegyzések kategóriája alapján rendel szcenáriókat a tickethez. Például a 7. ábrán látható véges automata részlet a hibabejelentések és sikeres/sikertelen javítások alapján vált állapotot. A bolyongás végállapota alapján kapjuk meg a szcenáriókat. Az ábrán láthatónál bonyolultabb automatával lehetséges lenne olyan szcenáriók felismerése is, mint „többszörös javítási kísérlet után sem működőképes”, „többszörösen javított, most működőképes” stb.

A véges automata modell különösen szerencsés választás, mert hosszú (sok bejegyzésből álló) ticketek esetén sem bonyolódik el.

Ha a bejegyzéseket meta-információval látjuk el (pl. idő, státusz), az operátor kényelmesen formálhat lekérdezéseket akár bejegyzés, akár ticket szinten. Sőt, mi-

vel formalizált rendszert hoztunk létre, lehetőség van az adatok automatizált feldolgozására (például automatikus heti statisztika, átlagos hibajavítási idő számítása a kábelhibákra stb).

5. Összefoglalás

Cikkünkben áttekintettük, hogy a modern dokumentumkategorizálási módszerek hogyan használhatók fel egy aktuális telekommunikációs probléma megoldására.

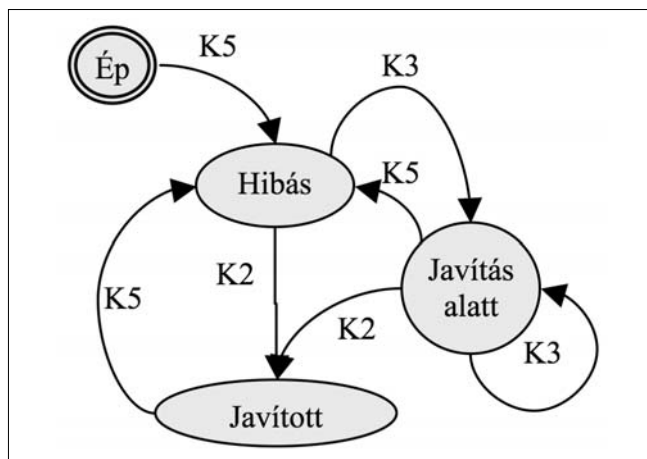
Amellett, hogy a konkrét problémára vázoltuk a megoldás útját, cikkünk átfogó képet igyekezett adni a jelenleg rendelkezésre álló dokumentumkategorizálási eszközökről és a dokumentumkategorizálás jelenlegi helyzetéről (state-of-the-art).

A dokumentumkategorizálás alkalmazásával a trouble ticketing rendszerek használhatósága, ezáltal értéke megnő. Reményeink szerint ebből minden érintett fél profitál, és idővel jobb hálózat rendelkezésre állás érhető el.

Irodalom

- [1] <http://dooka.canet4.net>
- [2] <http://www.switch.ch>
- [3] G. Salton, M. J. McGill, Introduction to Modern Information Retrieval, McGraw-Hill, Auckland, 1983.
- [4] G. G. Chowdhury, 'Searching and retrieval', Introduction to Modern Information Retrieval, Library Association Publishing, London, 1999., pp.158–178.
- [5] J. M. Ponte, Language modeling approach to information retrieval, In Proceedings of SIGIR'98, pp.275–281.
- [6] D. Hiemstra, A Linguistically Motivated Probabilistic Model of Information Retrieval, 1998. In C. Nicolaou and C. Stephanidis (eds.) Proc. of the 2nd European Conference on Research and Advanced Technology for Digital Libraries, ECDL-2, pp.569–584.
- [7] B.K. Benkő, T. Katona, Information retrieval in homogeneous document sets using syntactical parse information. In Proceedings of the 9th International Symposium for Social Communication, Santiago de Cuba, 2005. ISBN 959-7174-05-7
- [8] WordNet – <http://wordnet.princeton.edu/>
- [9] OWL – <http://www.w3.org/2004/OWL/>
- [10] H. P. Luhn, The Automatic Creation of Literature Abstracts, IBM Journal of Research and Development 2., 1958, pp.159–165.
- [11] NewsML project – <http://www.morphologic.hu>

7. ábra Véges automata a bejegyzések kategorizálására



Felhívás cikkek írására

a Híradástechnika „Web technikák és technológiák” cílszámába

A Híradástechnika 2006. januári számát web technikákkal és technológiákkal foglalkozó cílszámként kívánjuk megjelentetni. A cílszámhoz főként a terület egyes részterületeit áttekintő, tutorial jellegű cikkeket várunk, de helyet kívánunk biztosítani kutatók saját eredményeit tartalmazó rövid beszámolóknak is.

Javasolt tématerületek:

- web-orientált kommunikációs protokollok
- webtvé, webrádió, webtelefon, közvetítések
- web-alapú szolgáltatások
- speciális web alkalmazások (mobil telefónia, folyamatirányítás, mérés-technika)
- média konvergencia a technológiák szintjén
- információmodellezés, struktúra modellezés és analízis
- minőségi paraméterek értelmezése, definiálása, mérése
- minőségbiztosítás
- web-alapú tartalom menedzselés
- web-alapú tudás menedzselés
- elosztott virtuális terek
- tartalomszétosztás
- elosztott rendszerek
- keresőrendszerek
- web-alapú információgyűjtés
- web-biztonság
(protokollok, sérülékenységi, beágyazott rosszindulatú programok, védekezés)
- cache-elés, hatékonyságvizsgálat
- kiszolgálási modellek
- céleszközök
- rendszermodellek
- felhasználás (log) analízis
- méretezési problémák
- üzleti megoldások (szolgáltatások, fizetés)
- szabványok, szabványosítás
- fejlődési irányok, korlátok

Rövid beszámolókat várunk a web technológiákkal kapcsolatos európai kutatási projektekből való részvevőkről, eredményekről, konferenciákról. Recenzió közzétételére is van lehetőség értékesnek tartott web technológiákkal foglalkozó frissen megjelent könyvekről.

A beküldött cikkeket világos, érthető stílusban, a tématerülettel nem specialista szinten foglalkozó híradástechnikai szakemberek számára is jól érthető stílusban, magyar nyelven kell megírni. A cikk hossza maximum 5000 szó, az ábrák száma maximum tíz lehet.

A Híradástechnikában megjelenő cikkek formai követelményei a folyóirat 2005/8. számában találhatók.

Határidők:

A kézirat beküldése: 2005. december 10.

Visszajelzés az elfogadásáról és javaslatok a véglegesítéshez: 2005. december 15.

A végleges kézirat beküldése: 2006. december 31.

A cikkjavaslatokat a főszerkesztő és/vagy a vendégszerkesztő címére kérjük elküldeni.

Szabó Csaba főszerkesztő
szabo@hit.bme.hu

Vonderviszt Lajos vendégszerkesztő
vonderviszt.lajos@nhh.hu

Felhívás cikkek írására

a Híradástechnika „Optikai távközlés” célszámába

A Híradástechnika 2006. februári számát ismét optikai távközléssel foglalkozó célszámként kívánjuk megjelentetni. A célszámhoz főként a terület egyes részterületeit áttekintő, tutorial jellegű cikkeket várunk, de helyet kívánunk biztosítani kutatók saját eredményeiket tartalmazó rövid beszámolóknak is. Az eddig megjelent célszámok a 2004/2-es és a 2005/2-es voltak.

Javasolt tématerületek:

- optikai szálak, jellemzők, PMD
- nemlineáris optikai jelenségek, eszközök
- fotonikus kristályok és alkalmazásai
- fényterjedés szabad térben, FSO rendszerek
- optikai kábelek, technológiák, építési technológiák
- optikai aktív és passzív eszközök
- optikai erősítők (EDFA, Raman)
- MEMS-ek, optikai kapcsolók
- optikai adók, vevők, moduláció, kódolás, zajok, PMD kompenzálás
- optikai berendezések, rendszerek long-haul DWDM, Metro CWDM, OXC
- optikai hálózatok: core- és access hálózati technológiák, tervezési módszerek
- többszörös optikai hozzáférési technikák, pl. optikai CDMA, WDM/TDMA
- passzív optikai hálózatok, GPON, EPON, WPON,
- FTTx, FTTH rendszerek
- újgenerációs SDH, OTN rendszerek, MSPP berendezések
- Ethernet over SDH, GFP, VCAT, LCAS
- Ethernet, GbE, 10GbE rendszerek hálózatok, Ethernet O&M,
- IP, NGN optikai vonatkozásai, QoS
- többretegű, többtartományú optikai hálózatok tervezése (védelem, útvonalirányítás)
- ASTN/ASON/GMPLS control plane , UNI, protokollok
- optikai csomagkapcsolás, OBS, OPS hálózatok
- optikai műszerek, mérés technika
- optikai rendszerek modellezése
- optikai szabványok, szabványosítási fórumok legújabb eredményei
- beszámoló kiemelkedő nemzetközi optikai konferenciákról
- új termékek, termékismertető, alkalmazási példák, méretezés

Rövid beszámolókat várunk az optikai hírközléssel kapcsolatos európai kutatási projektekben való részvételéről, eredményekről, konferenciákról. Recenzió közlésére is van lehetőség értékesnek tartott optikai hírközléssel foglalkozó frissen megjelent könyvekről.

A beküldött cikkeket világos, érthető stílusban, a tématerülettel nem specialista szinten foglalkozó híradástechnikai szakemberek számára is jól érthető stílusban, magyar nyelven kell megírni. A cikk hossza maximum 5000 szó, az ábrák száma maximum tíz lehet.

A Híradástechnikában megjelenő cikkek formai követelményei a folyóirat 2005/8. számában található.

Határidők:

A cikkek címe és rövid, max 10 soros összefoglalója: 2005. december 10.

Visszajelzés a cikkjavaslat elfogadásáról: 2005. december 20.

A végleges kézirat beküldése: 2006. január 15.

A cikkjavaslatokat a főszerkesztő és/vagy a vendégszerkesztő címére kérjük elküldeni.

Szabó Csaba főszerkesztő
szabo@hit.bme.hu

Paksy Géza vendégszerkesztő
paksy@tmit.bme.hu

Evaluation of IPv6 services in mobile WiFi environment

Keywords: Internet2, IPv6, WiFi, L2 and L3 roaming, TCP Slow Start algorithm, TCP Windowing algorithm

How deeply does the mobility impact the TCPv6 and UDPv6 services? In our paper, we demonstrate the effects of the processes occurred during the roaming phase of the WiFi system on the IPv4 and IPv6 connections. The TCP connections are significantly affected by the interaction between the relative speed of mobile clients to the APs and the execution of roaming, whilst it has minor effect to the UDP transfer. We got a realistic view about the behavior of the IPv4 and IPv6 in mobile environment by using statistical methods to gain results from the comparative measurements. Furthermore we obtained details about how the IPv6 protocol really provides a higher quality on mobile environment over wireless data-link compared to its predecessor, the IPv4.

Trends in next generation mobile networks

Keywords: next generation mobile systems, mobile multimedia, QoS

Mobile communications and networks are developing at an astounding speed, with evidences of significant growth in the areas of mobile subscribers, mobile access networks, and mobile services and applications. Current times are just the beginning for deploying 3G mobile communication systems, while research on the next generation of mobile communications, 4G networks begin to pave the way for the future. 4G is simply an initiative by academic R&D labs to move beyond the limitations and problems of 3G which is having trouble getting deployed and meeting its promised performance and throughput.

Selective Retransmission of MPEG Video Streams over IP Networks

Keywords: selective retransmission, DCCP, MPEG streaming

In this paper a selective retransmission algorithm was presented for video streaming over IP networks where the video is streamed using DCCP transport protocol. The goal was to minimize the number of the corrupted bits, with acceptable transmission delay. The performance of the algorithm was confirmed by analytical examinations. The result shows a significant increase the quality of MPEG streams compared with the currently used methods.

Regional indicators of mobile penetration in Europe

Keywords: NMT, GSM, penetrations, supply factors

In this article the communication-geographer author is dealing with the timing and spatial patterns of the spread of mobile phone industry in the different periods (analogue, digital) of applied technologies. The features revealed in the article may be interpreted in a complex way only, as the individual factors (GDP, physical geo-

graphical factors, population density, cultural level, telecommunication policy) are not definitive alone. National and other 'soft' factors should also be taken into account.

MAIPAN – a middleware for interconnecting applications in personal networks

Keywords: pervasive/ubiquitous computing, access control, dynamic session management

This paper proposes the Middleware for Application Interconnection in Personal Area networks (MAIPAN), a middleware that provides a uniform computing environment for creating dynamically changing personal area networks (PANs). The middleware hides the physical scatteredness and device configuration of the PAN and presents its capabilities as a single computer to the applications. The solution provides easy set-up of PAN-wide applications utilizing multiple devices and allows transparent redirection of ongoing data flows, when the configuration of the PAN changes. In contrast to the approaches found in the literature, this paper presents a solution where session transfer, dynamic session management are tightly integrated with strong and intuitive access control security.

Analysis of a time-variant queueing system using matrix-geometric methods

Keywords: queueing system, Markov-arrival process, matrix-geometric method

The transient analysis of an infinite server queueing system is presented in this paper. The goal is to calculate the time dependent number of customers staying parallel in the system. The customers arrive according to a non-stationary point process and they have random service time. The non-stationary point process is given by a Markov Arrival Process. A differential equation system is presented describing the queueing system together with its solution. The solution is applied to a practical scenario where a telecommunication service is analysed.

Document categorization in a telecommunication context

Keywords: fault-tolerant systems, trouble ticket, document categorization, text processing

The motivation of this paper is to solve some deficiencies in today's trouble ticketing systems. The problem of the examined trouble ticketing systems is that the header-namely the category field-of the tickets doesn't contain enough information so it is not as helpful as it could be for analysis and research of reliable networks. Our goal was to achieve automatic category detection based on the free text description of the tickets. We used a wide range of tools of modern document categorization. Besides the solution of this concrete problem we give an overview about the available techniques and the state-of-the art of document processing.