

híradástechnika

1945 VOLUME LXII. 2007

hírközlés - informatika



Mobil infokommunikáció

Vertikális handover

3G mobil vagy WiFi/WiMAX?

Mobil központok villámvédelme

2007/5

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

Tartalom

<i>MOBIL INFOKOMMUNIKÁCIÓ</i>	1
Kovács Lóránt, Levendovszky János	
Minimális bithibaarány-stratégián alapuló közel optimális csatornakiegyenlítés statisztikai mintavételezéssel	2
Fülöp Attila, Zsíros Attila, Jeney Gábor	
Rétegek együttes optimalizálása UMTS-alapú videóátvitelben	8
Fülöp Péter, Szálka Tamás	
Vertikális handover – heterogén hálózatok együttműködése	15
Rácz András, Reider Norbert, Temesváry András	
Hívásátadási eljárás 3GPP-LTE újgenerációs hálózatokban	24
Huszák Árpád, Imre Sándor	
TFRC-alapú szelektív újraküldő algoritmus	30
Csaba Tamás	
Vezetéknélküli és mobil szélessávú hozzáférési technológiák alkalmazása	39
Babits László, Szedenik Norbert, Kiss István, Szűcs László, Lénárt Ferenc, Berta István	
Mobil kapcsolóközpontok primer és szekunder villámvédelme	43
Kovács László, Vidács Attila	
Dinamikus spektrumkiosztás: modell és árazás	49

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

Mobil infokommunikáció

imre@hit.bme.hu

A mobil és vezeték nélküli távközlés napjaink integrálódó infokommunikációs világában továbbra is meghatározó szerepet tölt be. A rádiós interfész sajátosságaiból és a mobilitásból fakadó műszaki nehézségek komoly kihívások elé állítják a mérnököket. A rádiós link többnyire szűk keresztmetszetet jelent a végpontok közötti kapcsolatokban, ráadásul a mozgási sebesség növekedésével a helyzet tovább romlik. A mobilitásból következő hálózati hozzáféréspont-váltás pedig nemcsak útvonalválasztási, hanem biztonsági és QoS problémákat is felvet.

Mostani számunkban a mobil rendszerek témaköréből járunk körül néhány elméleti és gyakorlati kérdést, illetve a rájuk adott válaszokat.

Kovács Lóránt és Levendovszky János a szűkös frekvenciaerőforrások hatékonyabb kihasználását lehetővé tevő új csatornaki egyenlítő algoritmust vezet be, a vezeték nélküli összeköttetésekben fellépő szelektív fading jelenségek kompenzálására. Az új algoritmus a kiegyenlítő szabad paramétereit közvetlenül a bithibaarány, mint költségfüggvény alapján optimalizálja, így a tradicionális eljárásoknál sokkal jobb hatásfok érhető el.

Fülöp Attila, Zsiros Attila és Jeney Gábor a PHOENIX EU projekt keretében kifejlesztett rendszert mutatják be, melynek célja egy olyan stratégia megvalósítása, ahol a forráskódolás, csatornakódolás és modulációs paraméterek együttesen kerülnek meghatározásra egy központi vezérlő intelligencia révén a hatékony multimédia átvitel érdekében. A cikk ismerteti a projektben alkalmazott szimulációs láncot, az UTRAN rendszerrészeket és a két szegmens összekapcsolását. Kitér az UTRAN optimalizálási megoldásra, valamint a szimulációs eredményekre is.

A vezetékes és vezeték nélküli, illetve mobil hálózatok együttműködése új szolgáltatások alapját teremti meg. A vertikális handover alkalmazásával a felhasználó szabadon mozoghat a különböző hálózatok között. Fülöp Péter és Szálka Tamás azt vizsgálják, miként nyújthat a szolgáltató integrált, hálózattól független szolgáltatásokat az optimális sebességű vagy költségű hozzáférési hálózat fellett. A szerzők áttekintik a vertikális handover fogalomkörét, annak tulajdonságait, alkalmazásának lehetőségeit és korlátait.

A 3rd Generation Partnership Project szabványosítási fórum már dolgozik a 3G hálózatok újabb generációjának szabványosításán, azzal a céllal, hogy a szabvány 2007 végére elkészüljön. A rendszer evolúciója érinti a rádiós hozzáférési hálózat (RAN) architektúráját, beleértve egy új, OFDM modulációra épülő rádiós interfészt, illetve a „core” hálózat (CN) architektúrájának megváltoztatását is. A rendszer RAN oldali evolúcióját Long Term Evolution

néven hivatkozzák a szabványosításban, míg a „core” hálózati részt System Architecture Evolution néven emlegetik. Rácz András, Reider Norbert és Temesváry András bemutatják az LTE RAN architektúra legfontosabb jellemzőit és részletesen elemzik az LTE hívásadási eljárást. Megmutatják, hogy a „hard handover” típusú hívásadás nem befolyásolja kedvezőtlenül a felhasználók által észlelt szolgáltatásminőséget és a rendszer teljesítményét.

Napjaink népszerű internetes protokolljainak hatásfoka jelentősen leromlik mobil környezetben a csomagvesztés eltérő okai miatt. Huszák Árpád és Imre Sándor egy olyan eljárást mutatnak be, amely az MPEG típusú videófolyam sérült, illetve elveszett csomagjait szelektíven újraküldi a hálózat pillanatnyi állapotától függően. Ezáltal a hálózat szabad kapacitását kihasználva tudunk a multimédiás tartalom minőségén javítani. A szabad kapacitást a forrás sebessége és a DCCP torlódáskezelő algoritmus (TCP Friendly Rate Control) által szolgáltatott paraméterek alapján határozzák meg.

A rádiós hozzáférési technológiák fontos fejlődési lépcsőhöz érkeztek, az általuk nyújtható adatátviteli sebesség jelentősen megnőtt. A két fő technológiai megközelítés, az IP-alapú hozzáférési módok (WLAN, WiMAX), és a mobil technológiák (GSM/(E)GPRS, UMTS/HSPA) között a számos hasonlóság mellett jelentős eltérések is mutatkoznak. Csaba Tamás cikkében ezen technológiák alkalmazását és alkalmazhatóságát veti össze a Pannon és annak anyavállalata, a Telenor tapasztalataira építve.

A Babits László, Szedenik Norbert, Kiss István, Szűcs László, Lénárt Ferenc és Berta István alkotta szerzői csapat egy nem hétköznapi mobilos témát mutat be, nevezetesen a mobil kapcsolóközpontok primer és szekunder villámvédelmének titkait. A mobil telefonhálózatok fontos részei a központi állomások (Mobile Switching Centres). Ezek adótornyait különösen veszélyeztetik a villámcsapások, a központok pedig – nagyszámú elektronikus berendezéseik miatt – különösen érzékenyek a túlfeszültségekre. A szerzők áttekintik az MSC központok villámvédelmének általános kérdéseit és az üzemeltetés során fellépő jellemző problémákat. Tárgyalják a villámvédelmi szempontból veszélyes tipikus helyzeteket, a védelem kialakításának elméleti hátterét és a kivitelezés során előforduló tipikus hibalehetőségeket.

Kovács László és Vidács Attila a szűkös frekvenciaspektrum hatékonyabb kihasználását lehetővé tevő új koncepciót, a dinamikus spektrumkiosztást mutatják be, ezen belül is az általuk javasolt modellt és árazási megoldást.

Imre Sándor
BME Híradástechn. Tanszék
vendég szerkesztő

Szabó Csaba Attila
főszerkesztő

Minimális bithibaarány-stratégián alapuló közel optimális csatornakegyenlítés statisztikai mintavételezéssel

KOVÁCS LÓRÁNT, LEVENDOVSKY JÁNOS

Budapesti Műszaki és Gazdaságtudományi Egyetem, Híradástechnikai Tanszék
{kovacs,levendov}@hit.bme.hu

Lektorált

Kulcsszavak: csatornakegyenlítés, adaptív algoritmusok, Monte-Carlo szimuláció, statisztikus mintavételezés

A cikk új csatornakegyenlítő algoritmust vezet be, a vezetéknélküli összeköttetésekben fellépő szelektív fading jelenségek kompenzálására. A keskenysávú átvitel során fellépő szelektív fading súlyos minőségromlást okozhat, amely jelentősen rontja a bithibaarányt. Az új algoritmus a kiegyenlítő szabad paramétereit közvetlenül a bithibaarány, mint költségfüggvény alapján optimalizálja, így a tradicionális eljárásoknál (amelyek csak szuboptimális célfüggvények alapján működnek, mint például négyzetes hiba, vagy csúcstorzítás) sokkal jobb hatásfok érhető el. Azonban a bithibaarány direkt minimalizálása exponenciális komplexitáshoz vezet, amelyet új statisztikai mintavételezésen alapuló algoritmusokkal lehet elkerülni, amelyekkel lehetővé válik a kiegyenlítő polinomiális komplexitásba történő optimalizálása. Így hatékony és egyben valós idejű kiegyenlítés biztosítható, valamint előírt minőségű szolgáltatás nyújtható a többutas terjedésből fakadó mélyfading esetén is. Ez hozzájárul a spektrális kihasználtság további növeléséhez, amely a jelenlegi vezetéknélküli kommunikációs technológiák egyik alapvető szűk keresztmetszete.

1. Bevezetés

Napjainkban ugrásszerűen megnőtt a szélessávú, vezetéknélküli mobil digitális adatátvitel iránti igény. Mivel a rendelkezésre álló fizikai sáv szélesség véges (és rendkívül drága), ezért a rendszerek spektrális kihasználtságának a növelése a cél, azaz az 1 Hz nominális sáv szélességen adott minőség mellett megvalósítható adatátviteli sebesség maximalizálása. A keskenysávú kommunikáció azonban nagyon érzékeny a többutas terjedésből fakadó fadingjelenségekre, amelyek súlyos lineáris torzításokat okozhatnak. Ennek csökkentése különös fontosságú a mobil rendszerekben, melyek napjainkban két irányban fejlődnek: egyrészt a jövő 3G-rendszerei CDMA-alapúak; másrészt a jelenlegi 2G (GSM, IS-136) rendszereket fejlesztik tovább oly módon, hogy a meglévő technológia segítségével nagyobb sáv szélességű szolgáltatásokat lehessen nyújtani [1].

Ez utóbbi megoldások alapja a 2G-rendszerekhez bevezetett új fizikai réteg, az EDGE (Enhanced Data rates for GSM Evolution). Az EDGE legfontosabb újítása a kétállapotú GMSK moduláció leváltása 8PSK-val, amely segítségével jelentősen javul a spektrális kihasználtság. A többállapotú moduláció bevezetése miatt irreálisan megnő a 2G rendszerekben detekcióra használt Viterbi-algoritmus komplexitása, amely a jelenlegi DSP technológia mellett nem kezelhető [1]. Ezért egyszerű MMSE (Minimum Mean Square Error) kiegyenlítő stratégiát alkalmaznak, amely azonban nem képes a bithibaarány jelentős csökkentésére a többállapotú moduláció esetén. Így továbbra is fontos kérdés marad a bithibaarány csökkentését hatékonyan javító, kis komplexitású kiegyenlítő algoritmusok kutatása.

Jelen cikkben a kiegyenlítő egy FIR szűrő, amelynek a súlyait közvetlenül a bithibaarány minimuma alapján

állítjuk be. A kiegyenlítőt egy küszöbdetektor követi. Az együttthatók optimalizálására vonatkozó tanulási algoritmus kis komplexitású gradiens kereséssel alapul. Az új módszerrel sokkal jobb teljesítőképesség érhető el keskenysávú átvitel esetén, mint a tradicionális ZF és MMSE algoritmusokkal. Ez tovább növeli a spektrális kihasználtságot.

A minimalizálás egyrészt a statisztikai megbízhatóság-analízisből ismert Li-Silvester határok, másrészt az átlagolt sztochasztikus approximáció segítségével történik.

A fenti eredményeket a cikk az alábbi szerkezet szerint tárgyalja: a 2. fejezetben a rendszermodell kerül bemutatásra; a következő fejezet témája a bithibaarány csökkentését egzakton minimalizáló, de exponenciális komplexitású algoritmus; majd a 4. fejezetben ezen algoritmus komplexitását csökkentjük determinisztikus, míg az 5. fejezetben sztochasztikus mintavételi módszerrel; végül pedig különböző terjedési viszonyokra vonatkozó szimulációs eredményeket mutatunk be.

2. Rendszermodell

Ebben a fejezetben az „egyfelhasználós” rendszerek széles osztályát leíró diszkrét idejű csatornamodell kerül bevezetésre, amely kiterjeszthető általánosabb esetre is. A modell-paraméterek pontos kapcsolata a folytonos átviteli rendszerrel számos cikkben megtalálható (pl. [5]), ezért itt erre bővebben nem térünk ki.

A csatorna diszkrét idejű modelljének szabad paramétereit jelölje $h_k, k = 0, \dots, M$ ahol M a lineáris torzítás tartója. A jelhez adódó zajt v_k jelöli, amelyről feltételezzük, hogy zérus várható értékű, N_0 spektrális sűrűségű fehér Gauss-zaj [5].

A vett sorozatot x_k jelöli, amely lineárisan torzított és zajos változata az elküldött y_k információs sorozatnak:

$$x_k = \sum_{j=0}^M h_j y_{k-j} + v_k \quad (1)$$

A döntőkészülék a kiegyenlítőből (FIR szűrő) és egy küszöbdetektorból áll. A FIR szűrő az

$$\tilde{y}_k = \sum_{j=0}^J w_j x_{k-j} \quad (2)$$

leképezést valósítja meg, ahol a $w_k, k = 0, \dots, J$ együtthatók jelölik a kiegyenlítő szabad paramétereit. A döntést előjeldetektorral képezzük: $y_k = \text{sgn}\{\tilde{y}_k\}$. A csatorna és a kiegyenlítő kaszkádjára külön jelölést vezetünk be:

$$q_k = \sum_{j=0}^M h_j w_{k-j} \quad (3)$$

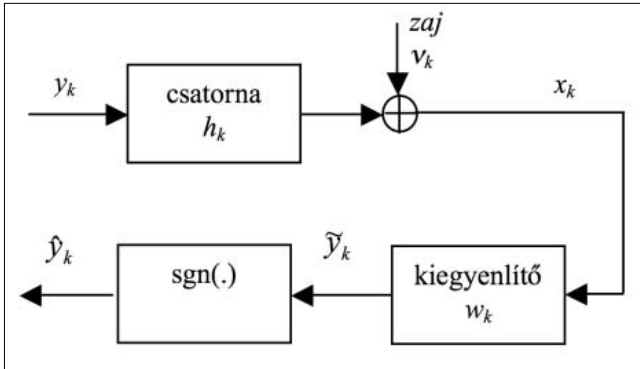
A hagyományos kiegyenlítők vagy a csúcstorzítást (PD – Peak Distortion):

$$\mathbf{w}_{opt} : \min \sum_{j=1}^L |q_j| \quad (4)$$

vagy a négyzetes középhibát (MSE – Mean Square Error):

$$\mathbf{w}_{opt} : \min E \left\{ \left(y_k - \sum_{j=0}^J w_j x_{k-j} \right)^2 \right\} \quad (5)$$

minimalizálják az egyszerű realizálhatóság érdekében [5]. Azonban ezen költségfüggvények nincsenek direkt kapcsolatban a bithibaarányval és a teljesítőképességük ennek megfelelően szerény [2,3]. A hatékonyság javítása érdekében közvetlenül a bithiba-valószínűséget minimalizáló algoritmusokra van szükség.



1. ábra A rendszermodell és az alkalmazott jelölések

3. A bithibavalószínűség minimalizálása egzakt gradiens-algoritmussal

A bithibavalószínűség (P_b) a kiegyenlítő együtthatóinak a függvényeként a korábbi irodalomból ismert [2]:

$$P_b(\mathbf{w}) = \frac{1}{2^L} \sum_{\mathbf{y} \in \Psi} \Phi \left(\frac{\sum_{n=0}^J w_n \sum_{l=n}^{M+n} h_{l-n} y_l}{\sqrt{N_0 \sum_{n=0}^J w_n^2}} \right) \quad (6)$$

ahol $L=J+M$ az eredő átvitel tartója, $\Phi(\cdot)$ a standard normális eloszlás eloszlásfüggvényét jelöli, továbbá

$$\Psi = \{\mathbf{y} = [-1, y_1, \dots, y_L], y_i \in \{-1, 1\}\} \text{ halmazt jelöli.}$$

A bithiba (6) szerinti kifejezését azon szűrőegyütthatók optimalizálják, amelyekre igaz, hogy kielégítik a következő egyenletet: $\mathbf{w}_{opt} : \text{grad}_{\mathbf{w}} P_b(\mathbf{w}) = 0$

A (6) gradiense analitikusan kifejezhető a következő formában:

$$\frac{\partial P_b(\mathbf{w})}{\partial w_i} = \frac{1}{2^L \sqrt{2\pi N_0} \left(\sum_{n=0}^J w_n^2 \right)^{3/2}} \sum_{\mathbf{y} \in \Psi} \exp \left(-\frac{\left(\sum_{n=0}^J w_n \sum_{l=n}^{M+n} h_{l-n} y_l \right)^2}{2 N_0 \sum_{n=0}^J w_n^2} \right) \cdot \left[\left(\sum_{n=0}^J w_n^2 \right) \left(\sum_{l=i}^{M+i} h_{l-i} y_l \right) - w_i \left(\sum_{n=0}^J w_n \sum_{l=n}^{M+n} h_{l-n} y_l \right) \right]$$

Így a bithibavalószínűség a gradiens módszerrel a következőképpen optimalizálható¹ (7):

$$\begin{aligned} w_i(k+1) &= w_i(k) - \Delta \frac{\partial P_b(\mathbf{w})}{\partial w_i} = \\ &= w_i(k) - \Delta \left\{ \frac{1}{2^L \sqrt{2\pi N_0} \left(\sum_{n=0}^J w_n^2 \right)^{3/2}} \sum_{\mathbf{y} \in \Psi} \exp \left(-\frac{\left(\sum_{n=0}^J w_n \sum_{l=n}^{M+n} h_{l-n} y_l \right)^2}{2 N_0 \sum_{n=0}^J w_n^2} \right) \cdot \right. \\ &\quad \left. \cdot \left[\left(\sum_{n=0}^J w_n^2 \right) \left(\sum_{l=i}^{M+i} h_{l-i} y_l \right) - w_i \left(\sum_{n=0}^J w_n \sum_{l=n}^{M+n} h_{l-n} y_l \right) \right] \right\} \end{aligned}$$

ahol $\mathbf{w}(k)$ a szűrőegyütthatók k -dik iterációbeli értékeit jelenti, illetve Δ a konvergenciát szabályozó lépéskönsz. Ez az algoritmus azonban a gradiensben szereplő exponenciálisan sok tagot tartalmazó összegzés miatt csak erősen korlátozott L értékekre² futtatható valós időben, hiszen minden lépésben $O(2^L)$ számítást kell végezni a szűrőegyütthatók adaptálásakor.

4. A bithiba-valószínűség minimalizálása a Li-Silvester módszerrel

Ebben a fejezetben az exponenciálisan növekvő számú összegzés elkerülésére statisztikai mintavételezést alkalmazunk.

Ennek alapja, hogy a bithiba-valószínűség (6) szerinti kifejezése egy várhatóértékként is értelmezhető, feltéve hogy a leadott információs sorozatok egyforma valószínűséggel fordulnak elő (ez az optimális forráskódolás jelenléte miatt általában fennáll). Így a $G(\mathbf{w}) = P_b(\mathbf{w})$ jelölést használva írhatjuk, hogy

$$\begin{aligned} \frac{\partial P_b(\mathbf{w})}{\partial w_i} &= \frac{1}{2^L} \sum_{\mathbf{y} \in \Psi} \Phi \left(\frac{\sum_{n=0}^J w_n \sum_{l=n}^{M+n} h_{l-n} y_l}{\sqrt{N_0 \sum_{n=0}^J w_n^2}} \right) = \\ &= \frac{1}{2^L} \sum_{\mathbf{y} \in \Psi} \Phi \left(\frac{\sum_{l=0}^L q_l y_l}{\sqrt{N_0 \sum_{n=0}^J w_n^2}} \right) = \\ &= \frac{1}{2^L} \sum_{\mathbf{y} \in \Psi} G(\mathbf{w}, \mathbf{y}) = E\{G(\mathbf{w}, \mathbf{y})\} \end{aligned} \quad (8)$$

¹ A gradienskeresés hátránya, hogy megakadhat lokális minimumokban, viszont a globális optimalizálást biztosító sztochasztikus keresési módszerekhez (pl. szimulált lehűtés) képest a gradiens gyorsabb konvergenciát eredményez.

² csatorna- + kiegyenlítőegyütthatók száma

továbbá vezessük be a következő jelöléseket:

$$g(\mathbf{w}) = \text{grad } G(\mathbf{w}) = \text{grad } E_y \{G(\mathbf{w}, \mathbf{y})\} = E_y \{g(\mathbf{w}, \mathbf{y})\}$$

A bithibavalószínűség (6) szerinti kifejezéséből látható, hogy a Ψ halmaz fölötti szummázás egyes tagjai a q_i együtthatók lineáris kombinációjának egy nemlineáris függvényeként állnak elő. A jel-zaj viszony növekedésével a standard normális eloszlásfüggvény mindinkább tart a $\text{sgn}(\cdot)$ függvényhez, azaz vagy 0-hoz közeli, vagy 1-hez közeli értékeket vesz föl, kivéve akkor, ha a függvényargumentum közelítőleg zérus. Ebből a tényből arra következtethetünk, hogy nagy jel-zaj viszont esetén a szummázás egyes tagjai között nagyságrendnyi eltérések adódhatnak, ami lehetővé teszi, hogy a teljes szummázást néhány domináns tag összegével közelítsük. Így egy kis komplexitásban kiszámolható, de éles alsóbecslést kaphatunk. (Ez a gondolat a statisztikai megbízhatóság-analízisből ismert Li-Silvester-módszer [7]).

Pontosabban, osszuk fel a Ψ teret két diszjunkt halmazzra, jelölje ezeket Ψ_1 és Ψ_2 . Az Ψ_1 halmaz számossága legyen K , méghozzá úgy, hogy azt a K db vektort tartalmazza, amelyre igaz, hogy

$$G(\mathbf{w}, \mathbf{y}_1) > G(\mathbf{w}, \mathbf{y}_2) > \dots > G(\mathbf{w}, \mathbf{y}_K) > G(\mathbf{w}, \mathbf{y}_i)$$

ahol $i = K+1, \dots, 2^L$. A $\Phi(\cdot)$ függvény tulajdonságaiból következően $0 \leq G(\mathbf{w}, \mathbf{y}_i) \leq 1$. Ezt kihasználva a bithiba alábbi korlátaikhoz jutunk:

$$\frac{1}{2^L} \sum_{\mathbf{y}_i \in \Psi_1} G(\mathbf{w}, \mathbf{y}_i) < P_b(\mathbf{w}) < \frac{1}{2^L} \left(\sum_{\mathbf{y}_i \in \Psi_1} G(\mathbf{w}, \mathbf{y}_i) + |\Psi_2| \right) \quad (9)$$

ahol $|\Psi_2|$ a halmaz számosságát jelenti. Esetünkben az $\mathbf{y}$ vektorok előfordulása egyenletes valószínűségű, így a (9) baloldalán szereplő alsó korlát igen szoros lehet, ugyanakkor a felső korlát nagyon laza, mivel a Ψ_2 -ben lévő G tagokat nagyon durván becsültük felülről.

A módszer alkalmazásának a kihívása a domináns tagok gyors megtalálása a (6) kifejezésben. Látható, hogy $P_b(\mathbf{w})$ invariáns $\mathbf{w}$ normájára, ezért – kihasználva az ebből adódó szabadságot – éljünk a $w_0 = 1/h_0$ választással, amiből következik, hogy $q_0 = 1$.

A (6) kifejezésből minden olyan tag elhagyható, amely nem tartalmazza $\mathbf{y}$ -t, hiszen $\mathbf{y}$ szerint szeretnénk maximalizálni. Ezek alapján a

$$\Phi\left(\sum_{i=0}^L q_i \mathbf{y}_i\right) = \Phi\left(-1 + \sum_{i=1}^L q_i \mathbf{y}_i\right)$$

kifejezést kell maximalizálni (hiszen a Ψ halmaz definíciójakor $y_0 = -1$ -t megkötöttük), amely a standard normális eloszlásfüggvény monotonicitása miatt az $\mathbf{y}$ argumentum maximalizálásával megtehető. A $\sum_{i=1}^L q_i \mathbf{y}_i$ maximuma éppen a csúcstorzítás (Peak Distortion – PD), ilyenkor

$$\mathbf{y}_1 = [-1, \text{sgn}\{q_1\}, \dots, \text{sgn}\{q_L\}.$$

Hasonló gondolatmenettel a legnagyobb tagot követő tagok megadhatók a következőképpen:

$$\mathbf{y}_k = [-1, \text{sgn}\{q_1\}, \dots, -\text{sgn}\{q_k\}, \dots, -\text{sgn}\{q_N\}, \dots, \text{sgn}\{q_L\}]$$

$$\text{ahol } \{i_1, \dots, i_N\} = C_k = \arg \min \left(|q_{i_1}| + \dots + |q_{i_N}| \right),$$

$$N = 1, 2, \dots, L \text{ és } C_k \neq C_{k-j}, j = 1, \dots, k-1.$$

Ezek után már csak a C_k halmazokat kell megkeresni. Amennyiben ragaszkodunk a tetszőleges számú első K darab legnagyobb tag megkereséséhez, az indexhalmazok megtalálása exponenciális komplexitású, hiszen az összes lehetséges $\mathbf{y}$ -ra ki kell számolni a $\mathbf{q}^T \mathbf{y}$ szorzatot, utána a szorzat értéke szerint sorbarendezeni, valamint a megfelelő $\mathbf{y}$ -okat kigyűjteni.

Azonban $K=4$ -re (amely a szimulációs eredmények szerint sok gyakorlati alkalmazásban teljesen elegendő), az első négy domináns tag az alábbi egyszerű algoritmussal megkapható.

$$\text{Legyen } \mathbf{q}' = [q_1, |q_2|, \dots, |q_L|]^T \text{ és } \mathbf{r} = S(\mathbf{q}')$$

ahol $S(\cdot)$ a csökkenő sorbarendezés operátora.

$$\text{Továbbá } \tilde{\mathbf{r}} = [q_0 | \mathbf{r}] = \tilde{S}(\mathbf{q})$$

ahol $\tilde{S}(\cdot)$ az első tagot nem érintő csökkenő sorbarendezés operátora és $s_i = \sum_{j=0}^i \tilde{r}_{L-j}$.

Az algoritmus lépései a következők:

1. Legyen $K=4$;
2. $\mathbf{y}_0 = [-1, 1, \dots, 1]$
3. $\mathbf{y}_1 = \tilde{S}^{-1}[-1, 1, \dots, 1, -1]$
4. $\mathbf{y}_2 = \tilde{S}^{-1}[-1, 1, \dots, -1, 1]$
5. Ha $s_1 + s_2 \leq s_3$
akkor $\mathbf{y}_3 = \tilde{S}^{-1}[-1, 1, \dots, -1, -1]$
másképpen $\mathbf{y}_3 = \tilde{S}^{-1}[-1, 1, \dots, -1, 1, 1]$
6. $N = \{\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3, \mathbf{y}_4\}$
7. A Li-Silvester-módszerrel kapott alsó becslőt jelölje $F(\mathbf{w}) \leq G(\mathbf{w})$.
Ennek becslő gradiense $f(\mathbf{w})$ nem más, mint $g(\mathbf{w})$, azzal a különbséggel, hogy a szummázás Ψ helyett N fölött történik.
8. Adaptáljuk a szűrő súlyokat a
 $w_i(k+1) = w_i(k) - \Delta f_i(\mathbf{w})$
algoritmus szerint, ahol $f_i(\mathbf{w})$
a gradiens i -dik komponensét jelöli.

4.1. Egyszerűsített mintavételezés

a BER meghatározására polinomiális komplexitásban

Ha lemondunk az éppen K legfontosabb tag megkereséséről, helyette K értékét maximalizáljuk és a ki-egyenlítés során ennél kisebb K -t is megengedünk, akkor egyszerű polinomiális komplexitású algoritmushoz jutunk. Vezessük be további jelöléseket: $K_{\max} = 2^p - 1$ jelöli a becslésben maximálisan használt $\mathbf{y}$ vektorok számát. $\mathbf{y}_0 = [-1, 1, \dots, 1]^T$ a PD-hez tartozó üzenetvektor. $\mathbf{y}_i = \mathbf{y}_0 \otimes \mathbf{z}_i$, ahol $\otimes$ az elemenkénti szorzás műveletét jelöli és $\mathbf{z}_i \in \{-1, 0\}^{L+1}$, valamint $\mathbf{e}_i$ az i -dik egységvektor.

Ebben az esetben a becslő feladata azon $\mathbf{y}$ vektorok halmazának megtalálása a (6)-beli szummázásban, amelyekhez az összegnek jelentős (de nem feltétlenül a legjelentősebb) tagjai tartoznak.

Az egyszerűsített becslő algoritmus lépései:

1. Legyen $p = \log_2(K_{\max} + 1) + 1$
2. Ha $r_p \geq s_{p-1}$
akkor $\mathbf{z}_i = [0, \dots, 0 | \mathbf{u}_i^T]^T$,
ahol $\mathbf{u}_i \in \{0, -1\}^{p-1}, i = 1, 2, \dots, 2^{p-1}$
 $\mathbf{y}_i = \tilde{S}^{-1} \{\mathbf{y}_0 \otimes \mathbf{z}_i\}$
másképpen $\arg \max r_j \geq s_j, j < p$
 $\mathbf{z}_i = [0, \dots, 0 | \mathbf{u}_i^T]^T$,
ahol $\mathbf{u}_i \in \{0, -1\}^{j-1}, i = 1, 2, \dots, 2^j$
 $\mathbf{z}_i = -\mathbf{e}_k, k = J+1, \dots, p-1, i = 2^j - j + k$
 $\mathbf{y}_i = \tilde{S}^{-1} \{\mathbf{y}_0 \otimes \mathbf{z}_i\}$
4. Az előző pontban adódó $\mathbf{y}$ vektorok halmazát jelölje M .
5. Az egyszerűsített Li-Silvester-módszerrel kapott alsó becslőt jelölje $F'(\mathbf{w}) \leq G(\mathbf{w})$. Ennek becslő gradiense $f'(\mathbf{w})$ nem más, mint $g(\mathbf{w})$, azzal a különbséggel, hogy a szummázás ψ helyett M fölött történik.
6. Adaptáljuk a szűrő súlyokat a
 $w_i(k+1) = w_i(k) - \Delta f'_i(\mathbf{w})$
algoritmus szerint.

5. A bithibavalószínűség minimalizálása sztochasztikus mintavételi módszerekkel

Az előzőekben determinisztikus mintavételi módszerrel választottuk ki a mintákat, amelyekkel a (8) egyenletben szereplő várható értéket – vagy annak gradiensét – közelítettük. Ebben a fejezetben a $G(\mathbf{w})$ függvény becslésére véletlenszerűen sorsolt mintákat használunk.

5.1. Sztochasztikus approximáció

Ebben az esetben a (7) algoritmusban a teljes Ψ tér fölötti szummázást egyetlen véletlenszerűen sorsolt $\mathbf{y}$ vektor alapján kiszámolt mintával közelítjük. A sorsoláshoz egyenletes eloszlást választunk, mert a forrás oldalán is egyenletes eloszlású sorozatokat tételeztünk fel. Az így adódó, úgynevezett sztochasztikus gradiens algoritmus a következő:

$$w_i(k+1) = w_i(k) - \Delta \hat{g}_i(\mathbf{w}(k), \mathbf{Y} = \mathbf{y})$$

ahol $\mathbf{Y} \in \Psi$ és $P\{\mathbf{Y}_i = -1\} = P\{\mathbf{Y}_i = 1\} = 0.5$, továbbá (10):

$$\begin{aligned} \hat{g}_i(\mathbf{w}(k), \mathbf{Y} = \mathbf{y}) &= \\ &= \frac{1}{2^L \sqrt{2\pi N_0} \left(\sum_{n=0}^J w_n^2 \right)^3} \exp \left(- \frac{\left(\sum_{n=0}^J w_n \sum_{l=n}^{M+n} h_{l-n} y_l \right)^2}{2 N_0 \sum_{n=0}^J w_n^2} \right) \cdot \\ &\cdot \left[\left(\sum_{n=0}^J w_n^2 \right) \left(\sum_{l=i}^{M+i} h_{l-i} y_l \right) - w_i \left(\sum_{n=0}^J w_n \sum_{l=n}^{M+n} h_{l-n} y_l \right) \right] \end{aligned}$$

Az algoritmus konvergenciájának vizsgálatát a Kushner-Clark [6] tétel segítségével végezzük. A tétel alkalmazásának feltétele, hogy a vizsgálni kívánt sztochasztikus approximációban szereplő mennyiségek eleget tegyenek bizonyos feltételeknek.

Ezek első csoportja a lépéskonstans megválasztásával kapcsolatos, amely a $\Delta = 1/k$ választással kielégíthető. A második feltételcsoport teljesüléséhez az szükséges, hogy a $\hat{g}_i(\mathbf{w}(k), \mathbf{Y} = \mathbf{y})$ függvény parciálisan differenciálható legyen y_i és w_i szerint, ami esetünkben teljesül. Ezek után a $\hat{g}_i(\mathbf{w}(k), \mathbf{Y} = \mathbf{y})$ függvény $\mathbf{y}$ szerinti várható értékét kell kiszámolnunk, s ennek segítségével írható föl az a determinisztikus differenciálegyenlet-rendszer, aminek a vizsgálata megadja a sztochasztikus differenciaegyenlet viselkedését is. (A Kushner-Clark tétel bizonyítása [6]-ban található).

Mivel

$$\begin{aligned} E\{\hat{g}_i(\mathbf{w}(k), \mathbf{Y} = \mathbf{y})\} &= \\ &= \frac{1}{2^L \sqrt{2\pi N_0} \left(\sum_{n=0}^J w_n^2 \right)^3} \sum_{\mathbf{y} \in \Psi} \exp \left(- \frac{\left(\sum_{n=0}^J w_n \sum_{l=n}^{M+n} h_{l-n} y_l \right)^2}{2 N_0 \sum_{n=0}^J w_n^2} \right) \cdot \\ &\cdot \left[\left(\sum_{n=0}^J w_n^2 \right) \left(\sum_{l=i}^{M+i} h_{l-i} y_l \right) - w_i \left(\sum_{n=0}^J w_n \sum_{l=n}^{M+n} h_{l-n} y_l \right) \right] \end{aligned}$$

ami nem más, mint a $P_b(\mathbf{w})$ függvény $\mathbf{w}$ szerinti deriváltja. A vizsgálandó differenciálegyenlet-rendszer:

$$\begin{aligned} \frac{dw_i(t)}{dt} &= \frac{1}{2^L \sqrt{2\pi N_0} \left(\sum_{n=0}^J w_n^2 \right)^3} \sum_{\mathbf{y} \in \Psi} \exp \left(- \frac{\left(\sum_{n=0}^J w_n \sum_{l=n}^{M+n} h_{l-n} y_l \right)^2}{2 N_0 \sum_{n=0}^J w_n^2} \right) \cdot \\ &\cdot \left[\left(\sum_{n=0}^J w_n^2 \right) \left(\sum_{l=i}^{M+i} h_{l-i} y_l \right) - w_i \left(\sum_{n=0}^J w_n \sum_{l=n}^{M+n} h_{l-n} y_l \right) \right] \end{aligned}$$

A fenti differenciálegyenlet numerikus megoldásával számos csatornamodellre bizonyítható a stabilitás és belátható, hogy a (10)-zel adott sztochasztikus approximáció stabil és konvergens, továbbá ugyanoda konvergal, mint a determinisztikus algoritmus.

5.2. Monte-Carlo módszer

Ebben az esetben iterációként nem egy, hanem több $\mathbf{y}$ vektort is kisorsolunk (az $\mathbf{y}$ eloszlása szerint, esetünkben egyenletes eloszlással, visszatevéssel), majd a mintákat ezekre átlagoljuk, a következőképpen:

$$\hat{g}_i^{MC}(\mathbf{w}(k)) = \frac{1}{N} \sum_{j=1}^N \hat{g}_i(\mathbf{w}(k), \mathbf{Y} = \mathbf{y}_j)$$

A súlyok beállítására szolgáló algoritmus:

$$w_i(k+1) = w_i(k) - \Delta \hat{g}_i(\mathbf{w}(k), \mathbf{Y} = \mathbf{y})$$

6. Szimulációs eredmények

A szimulációs eredmények az újonnan bevezetett algoritmusok két legfontosabb tulajdonságára,

- az egyes módszerekkel elérhető bithibaarányra;
- illetve az egyes módszerek konvergenciaidejére összpontosulnak.

Ezen vizsgálatokat három standard – különböző terjedési körülményeket figyelembe vevő – csatornamodell-

re végeztük el [4]: $\mathbf{h}^{(1)} = [10.6, -0.3]^T$, $\mathbf{h}^{(2)} = [1; 0.6; -0.45]^T$, $\mathbf{h}^{(3)} = [1; 0.12; 0.3; -0.8]^T$, (Megjegyezzük, hogy a $\mathbf{h}^{(2)}$ csatorna nem minimál-fázisú, amelynek a kiegyenlítésére a hagyományos algoritmusok alkalmatlanok.)

6.1. A bithiba a jel-zaj viszony függvényében

A szimulációk egyik célja a bithibaarány (BER) jelzaj viszonytól (SNR) való függésének a megadása volt, különböző kiegyenlítési algoritmus esetén. Az eredményeket a 2-4. ábrák mutatják.

Az eredményekből az látszik, hogy az új algoritmusok elsősorban jó jel-zaj viszony értékek (SNR > 20 dB) esetén nyújtanak a hagyományos megoldásokhoz képest lényegesen jobb teljesítményt: minimál-fázisú csatornák ($\mathbf{h}^{(1)}$, $\mathbf{h}^{(3)}$) esetén a bithibaarány az eredeti érték felére csökkent, míg nem minimál-fázisú esetben akár egy nagyságrendnyi javulás is tapasztalható. A legjobb teljesítményt minden esetben az egzakt gradiens algoritmus (TGS) szolgáltatja, azonban ettől csak egészen minimális mértékben maradnak el a véletlen mintavételezésen alapuló módszerek (sztochasztikus gradiens (STG), illetve Montecarlo (MC)³), ugyanakkor ez utóbbiak kis komplexitásúak.

A determinisztikus mintavételezésen alapuló egyszerűsített Li-Silvester módszer (LISI)⁴ is minden esetben jobb eredményt szolgáltat a hagyományos megoldásokhoz képest, azonban kevésbé elmarad a sztochasztikus módszerek teljesítőképességétől. (Ez utóbbi alól kivételt képeznek azok az esetek, ahol a domináns tagok száma kicsi az összes tag számához képest, és rossz a jel-zaj viszony, például a $\mathbf{h}^{(3)}$ csatorna esetében, 22 dB SNR alatt (lásd a 4. ábrát). Ebben az esetben 512 tag összegét 4 domináns tag felhasználásával becsültük.

A jel-zaj viszony növekedtével a tagok közötti eltérések egyre nagyobbak, hiszen a $\Phi(\cdot)$ függvény egyre inkább közelít a signum függvényhez, így nagyobb jel-zaj viszony értékek esetén jobb, míg gyenge jel-zaj viszony esetén rosszabb a Li-Silvester közelítés).

6.2. Konvergenciasebesség

A gyakorlati alkalmazások szempontjából mindenképpen fontos, hogy az újonnan bevezetett algoritmusokhoz csatornainformációra, azaz a h függvény ismeretére van szükség.

A szimulációk során feltételeztük, hogy egzakt csatorna-információ áll rendelkezésünkre. Ennek hiányában csatorna-identifikáló algoritmus segítségével becsülhetjük az ismeretlen koeficiensket.

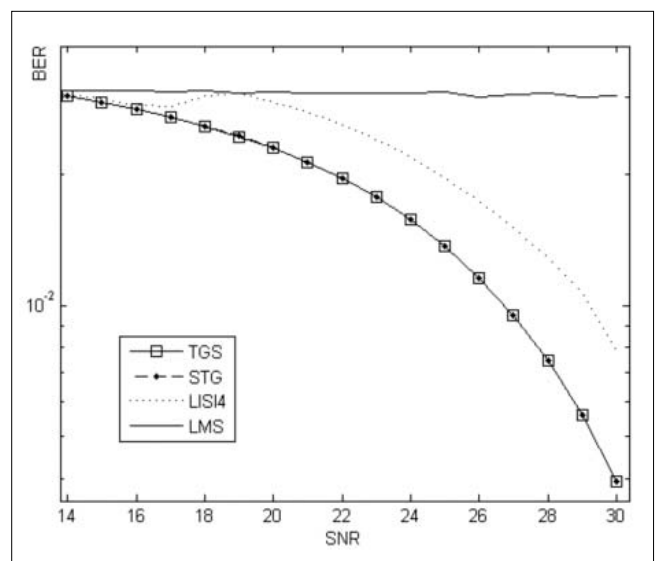
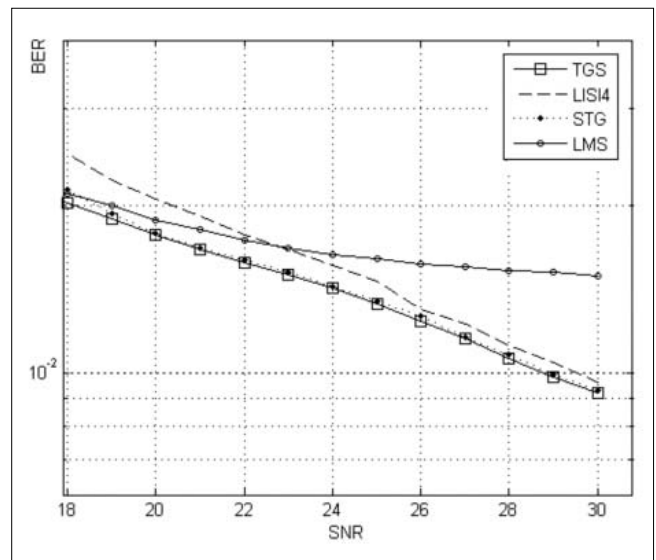
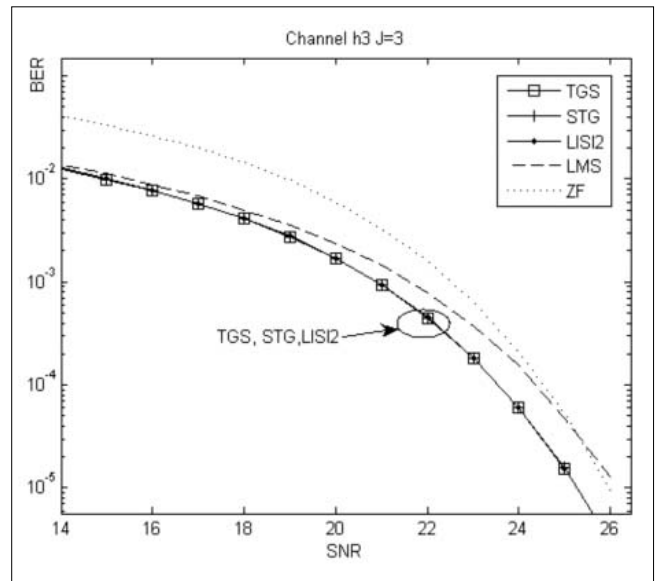
Mivel a csatornaki egyenlítés valós időben megoldandó feladat, így fontos kérdés az algoritmusok komplexitása illetve iteratív algoritmusoknál a konvergencia sebessége. Az előbbi meghatározza, hogy milyen szimbólumsebességű forrást tudunk kiszolgálni adott számítás

Bithibaarány a jel-zaj viszony függvényében

2. ábra a $\mathbf{h}^{(1)}$ csatorna és 3 kiegyenlítő-együttható esetére

3. ábra a $\mathbf{h}^{(2)}$ csatorna és 6 kiegyenlítő-együttható esetére

4. ábra a $\mathbf{h}^{(3)}$ csatorna és 6 kiegyenlítő-együttható esetére

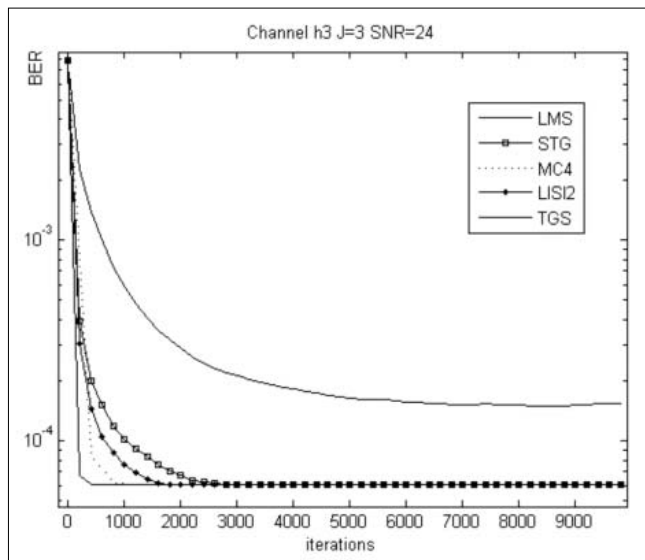


³ A grafikonokon az MC rövidítés után álló szám arra utal, hogy iterációként hány véletlen vektor sorsolásával futtatuk az algoritmust.

⁴ A LISI rövidítés után álló szám arra utal a grafikonokon, hogy adott esetben mekkora volt a maximálisan felhasznált domináns tagok száma az egyszerűsített Li-Silvester becslő futtatásakor

kapacitás esetén. Az utóbbi arra vonatkozóan ad útmutatást, hogy a csatorna milyen változási sebessége mellett képes az adaptív algoritmus követni az optimumot.

Egy adott jel-zaj viszony értékre vonatkozó konvergencia-görbe (bithibaarány az iterációk függvényében) látható az 5. ábrán.



5. ábra

Bithibaarány az iterációk függvényében a $h^{(2)}$ csatorna, 3 kiegyenlítő-együttható és 24 dB SNR esetén

Leggyorsabban minden esetben az egzakt gradiens algoritmus (TGS) konvergál. A sztochasztikus mintavételre alapuló módszerek konvergenciaideje körülbelül annyszorosára növekszik ehhez képest, amennyivel kevesebb tagot használunk az összegzésben. Vagyis ezekkel a módszerekkel eredményesen csökkenthető az 1 iteráció során elvégzendő műveletek száma, de nem csökkenthető jelentősen a teljes konvergencia során elvégzendő műveletek száma. Az egyszerűsített Li-Silvester módszer konvergenciája a sztochasztikus módszerekkel nagyjából megegyezik. Ugyanakkor bármelyik új módszer lényegesen gyorsabban konvergál az LMS algoritmusnál, a különbség – csatornától függően – 1-2 nagyságrendnyi is lehet (lásd 5. ábra). Továbbá figyelemreméltó, hogy az új módszerek alkalmazása esetén (főleg TGS és MC) a kezdeti lépésekben rendkívül meredeken csökken a bithibavalószínűség.

7. Összefoglalás

A cikkben a bithibavalószínűséget, mint a digitális összeköttetések legfontosabb minőségjellemzőjét alkalmaztuk kiegyenlítési stratégiaként. Ezáltal megnövekedett teljesítőképességű algoritmusok adódnak. Az egyre olcsóbbá váló és folyamatosan növekvő számítási kapacitás kihasználásával azonban ezek a kismértékben bonyolultabb algoritmusok is gazdaságosan megvalósíthatóak, ezáltal kisebb bithiba-valószínűséget és jobb spektrális kihasználtságot eredményezve.

Irodalom

- [1] W. Gerstacker, R. Schober: „Equalization Concepts for EDGE”, IEEE Trans. Wireless Comm., Januar 2002, Vol. 1, No.1., pp.190–199.
- [2] O. Shimbo, M. Celebiler: „The probability of error due to intersymbol interference and gaussian noise in digital communication systems”, IEEE Trans. on Communication Technology, 1971. Vol. COM-19, pp.113–119.
- [3] C. Yeh, J.R. Barry: „Adaptive minimum bit-error rate equalization for binary signaling”, IEEE Trans. Comm. Vol. 48 pp. 1226–1235. Jul. 2000.
- [4] R. Steele and L. Hanzo (editors): „Mobile Radio Communications” Wiley, 1999.
- [5] J. G. Proakis: „Digital Communications” McGrawHill, 1995.
- [6] H.J. Kushner, D.S. Clark: „Stochastic approximation methods for constrained and unconstrained systems” Springer Verlag, 1978.
- [7] O. K. Li, J. A. Silvester: „Performance Analysis of Networks with Unreliable Components”, IEEE Trans. Comm., 1984. Vol. COM-32, pp.1105–1110.

Rétegek együttes optimalizálása UMTS-alapú videóátvitelben

FÜLÖP ATTILA, ZSIROS ATTILA, JENEY GÁBOR

Budapesti Műszaki és Gazdaságtudományi Egyetem, Híradástechnikai Tanszék
{fulopa, zsirosa, jeneyg}@hit.bme.hu

Kulcsszavak: együttes forrás és csatornakódolás, multimédia átvitel, IPv6, UMTS rádiós elérési hálózat, erőforrás vezérlés

Ebben a dokumentumban egy olyan szoftverplatform kerül ismertetésre, amely demonstrációs és teljesítményelemzési céllal lett kifejlesztve. A platform alkalmazási és egyéb (pl. fizikai) rétegek együttes optimalizálásán alapuló vezeték nélküli videó átvitelt szimulál. A fejlesztés az IST-PHOENIX projekt keretein belül valósult meg és a projekt által kifejlesztett rendszer optimalizálási modellen nyugszik. A kifejlesztett platform egy lényeges elemét, a vezeték nélküli hálózati szegmensét azonban egy részletes, szabványos UTRAN hálózati szimulációra cseréltük. A cikk tartalmazza a projekt szimulációs láncának rövid ismertetését, illetve az UTRAN rendszerrészek ismertetését, a két szegmens összekapcsolását. Kitérünk az UTRAN optimalizálásban való szerepére, a konfigurációjára, vezérlésére, majd néhány szimulációs eredmény kerül bemutatásra.

1. Bevezetés

A vezeték nélküli telekommunikációs hálózatok fejlődése egy egyre integráltabb és globálisabb rendszer irányába mutat. A hagyományos felfogás (ISO/OSI) szerint rétegekre bontjuk a kommunikációban résztvevő álmások funkcióit, hogy azokat könnyen megfogható módon, egyénileg alakíthassuk tökéletessé. Manapság azonban a felhasználói és ipari igények kielégítése csak a különféle technológiák konvergenciájával válik lehetővé. Ez egyben a különálló rétegek összemosását, a funkciók átlapolódását, keveredését jelenti. Ebben a cikkben egy olyan projekt és szoftveres demonstrációs platform bemutatására kerül sor, amelynek célja az utóbbi felfogás támogatása. A szoftver lehetővé teszi a rádiós videóátvitel alkalmazási és egyéb rétegek közötti együttes optimalizálásának teljesítményelemzését. A platform az IST PHOENIX projekt keretein belül készült (<http://www.ist-phoenix.org>) és egy, a projekt által fejlesztett modellen alapul.

A rendszer felépítése követi a hagyományos ISO/OSI modellt, ugyanakkor célja egy olyan stratégia megvalósítása, ahol a forráskódolás, csatornakódolás és modulációs paraméterek együttesen kerülnek meghatározásra egy központi vezérlő intelligencia által. Ezt nevezzük közös forrás- és csatornakódolás/dekódolásnak (Joint Source and Channel Coding/Decoding – JSCC/D). A tradicionális esetben a forráskódolás (tömörítés) és a csatornakódolás (hibajavító kódolás) egymástól elválasztva kerül megvalósításra Shannon jól ismert szeparációs tétele alapján [1].

Ez az elv komplex, ám jól teljesítő rendszereket eredményez, de nem különösebben hatékony az olyan mai népszerű alkalmazások esetében, mint például az audió-videó jelfolyam átvitel [2]. A mai modern alkalmazásoknak gyakran olyan, a szeparációs tétel ideális feltevéseit nem kielégítő jellemzőik vannak, amiket réteges megközelítésben nem lehet kielégíteni. Ilyenek többek

közt a valós idejű átviteli követelmény, vagy a változó hibaérzékenységű forrásadatok változó hibavédelme.

A cikkben közölt architektúra különböző hozzáférési technikákkal alkalmazható, például OFDM vagy WCDMA. Mind a H.264/AVC [3], mind az MPEG-4 videó kódolás támogatott az alkalmazási rétegben. Különböző szállítási rétegbeli protokollok (UDP, UDP-Lite, DCCP) használhatóak az átvitelhez. A projekt által fejlesztett modell és szimulációs eszköz lehetőséget biztosít számunkra, hogy egy valós rendszert, az UMTS hálózatát használva is kipróbáljuk az együttes optimalizálás (JSCC/D) elvét. A cikkben az UTRAN WCDMA rendszerre ültetett rendszermodell bemutatására koncentrálunk, az általunk fejlesztett, részletes, szabványhoz igazodó UMTS szimuláción keresztül [4]. A modell részeinek funkcionális bemutatása után az optimalizálási lehetőségek elemzése következik, részletesebben kitérve az UTRAN által támogatott átviteli konfigurációkra. Különböző futtatási beállítások mellett mutatjuk be az optimalizálás hatását az átvitt videó folyam minőségére.

A 2. fejezetben a szimulációs rendszer felépítését ismerhetjük meg, ezt követően pedig a jelzési mechanizmusokat részletezzük. A protokoll-hierarchiát két részre, alkalmazási és adattovábbítási modulokra bontva a 4. és 5. fejezet tartalmazza az egyes modulok rövid, az optimalizálás szempontjából releváns működésének ismertetését.

2. A szimulációs rendszer architektúrája

Az 1. ábrán látható a rendszermodell UTRAN WCDMA hálózatra ültetve. Az ábra egy videófolyam átvitelét mutatja be. Az egymáshoz kapcsolódó elemek vezetékes (IPv6), és vezeték nélküli (jelen esetben UMTS) hálózatokat reprezentálnak, a vezetékes közeg IP hálózat felhőben szerepel, míg a vezeték nélküli közeg a csatornablokk. A rendszer adattovábbítást végző blokkjai

ISO/OSI rétegmodell szerint rendeződnek. A rendszer JSCC/D vezérlőinformációkat használ, amik például az UEP (Unequal Error Protection) védelmet, vagy a JSCC adaptált csatornakódolást vezérlik. Külön vezérlőréteg van definiálva az alkalmazási és a fizikai rétegekhez, ez utóbbi az UTRAN-ban használt rádiós erőforrásvezérlési réteg (RRC). Az alkalmazási rétegbeli visszacsatolt vezérlő (JSCC/D) a forráskódolóhoz, az alkalmazási műveletekhez (titkosítás, UEP) valamint a streaming, transzport és IPv6 csomagképző protollokhoz kapcsolódnak. Ezen utolsó három elem biztosítja egy valós idejű multimédia folyamtnál megkövetelhető QoS-t, sorrendhelyes továbbítást stb. Ezen elemek funkcióját a 4/C szakasz ismerteti bővebben. Az egyes elemek információt nyújtanak a felhasználói igényekről, és a hálózat, illetve a csatorna állapotáról. Az információ a vezérlőelemekhez továbbítódik, amely ezek alapján döntéseket tud hozni. Ilyen módon képes az adaptációt vezérelni, megadja például a videó kódoló tömörítési rátáját, vagy a csatornakódolás szükséges védelmi szintjét.

Az ábrán a folytonos vonal jelöli a hasznos adat, azaz a payload (videó) és a hozzá kapcsolódó információk áramlását. A videóadat a forráskódolás és egyéb alkalmazási műveletek (pl. UEP) után streaming és transzport protollok alkalmazásával egy IPv6 hálózaton halad keresztül. A hasznos adathoz, az optimalizáláshoz elengedhetetlen információk (SSI, SRI) és protokoll felek kapcsolódnak. A videóadathoz nem szinkronizált információk külön jelfolyamban továbbítottak (lásd szaggatott vonal). A visszacsatolt vezérlő megfelelő műkö-

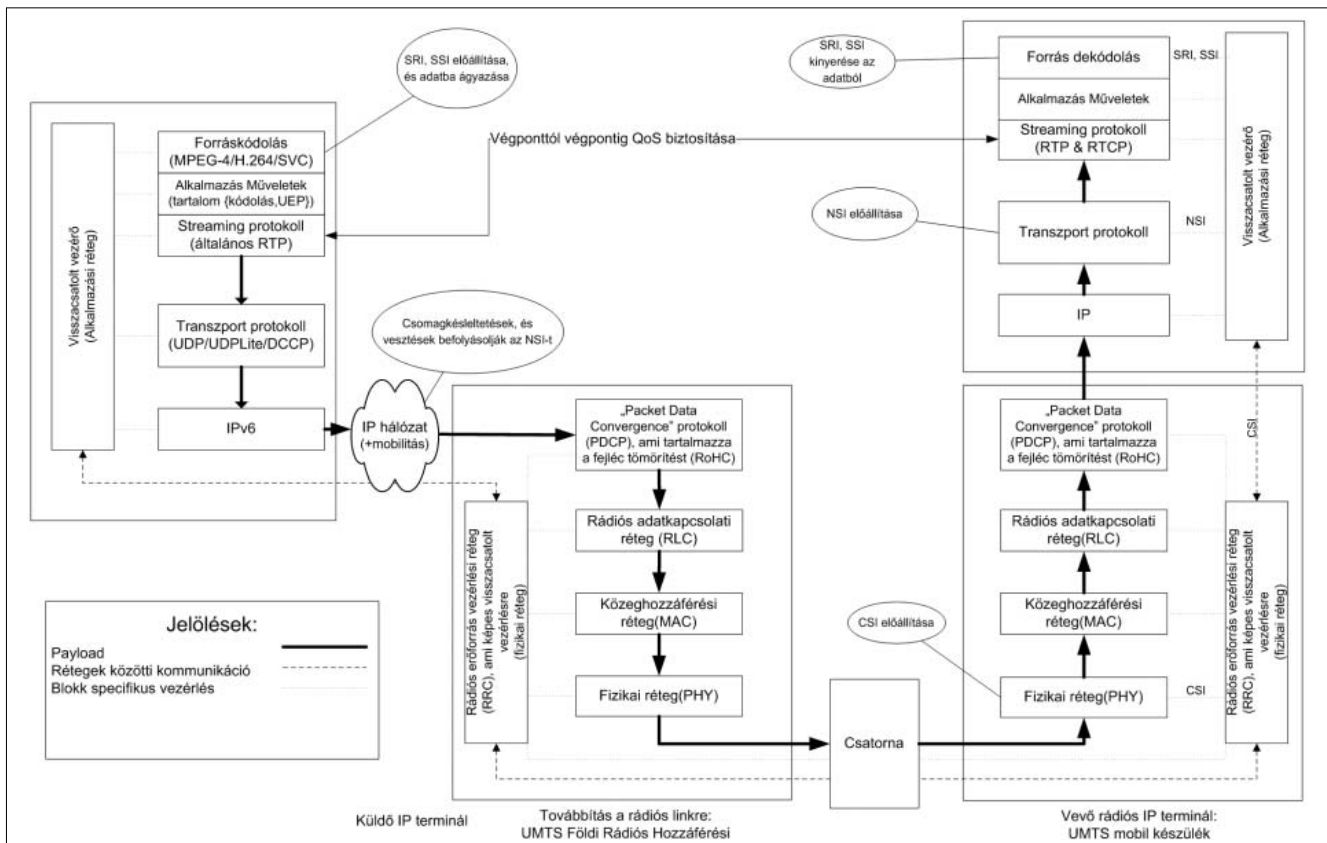
déséhez szükséges volt kialakítani az egyes blokkokkal való interfészeket, melyeket az ábra pontozott vonallal jelöl. Ugyancsak így jelöltük az általunk használt UTRAN modulban a rádiós erőforrás vezérlési réteg és az egyes jelfeldolgozásban résztvevő rétegek (PDCP, RLC, MAC, PHY) közötti interfészeket. Ezek a szolgáltatás hozzáférési pontok (Service Access Points) a szabványokban [8] definiáltak, azonban bővítenünk kellett funkcionalitásukat, hogy lehetővé tegyük az együttes optimalizálást.

3. Együttes optimalizálás vezérlése

Vezetéknélküli videó átviteli rendszerekben alkalmazott rétegek közötti optimalizálásra több megoldás is létezik. Ezek közül egy a közeghozzáférési rétegbeli (MAC) automatikus ismétléskérés (ARQ), és fizikai rétegbeli előremutató hibajavító kódolás (FEC) adaptív kiválasztására alapuló módszer [5]. Ezzel szemben a Phoenix projektben követett megoldás (JSCC/D) a teljes protokollvermen keresztül alakít ki kommunikációt, mind az alkalmazási, mind a fizikai és közeghozzáférési rétegek viselkedését próbálja összehangolni a tényleges átviteli körülményektől függően.

Az alábbi releváns információk cseréjére kerül sor [6] az optimalizálás során: SSI (source significant information), azaz a forrás csatornahibákra mutatott érzékenysége; SRI (source a-priori information), azaz a forrás statisztikai információja az egyes adatelemek gyakorisá-

1. ábra Teljes rendszermodell és szimulációs összeállítás



gáról; valamint a DRI (decision reliability information), a csatornadekódoló megbízhatósági információja (ha szoft értékeket tartalmazó kimenete van); a SAI (source a-posteriori information), a forráskódolás statisztikai kimenete és végül a visszacsatolt információk, a CSI (channel state information), például jel-zaj arány; NSI (network state information), ami késleltetési, jitter és csomagvesztési értékeket tartalmaz többek között és a videó minőség.

Az egyes vezérlő információk jelentésbeli és funkcionális jellegük miatt eltérnek frissítési gyakoriságukban, generálási módjukban, adatmennyiségükben és bitsebességükben, így különböző módon kell azokat kezelni adattovábbítás szempontjából. Az SSI-t a felhasználói adattal kell szinkronizálni, mivel az adatforrás csatornahibákra mutatott érzékenységét írja le, amit az UEP modulnak kell kezelni. Ezzel szemben például az NSI információ a hálózat állapotát írja le és visszacsatolás útján jut el az együttes, rétegek közötti vezérlőbe (JSCC/D), azaz nem kell szinkronizálni az adatcsomagokkal.

A szinkronizálásra megjelölt információkat a felhasználói adatfolyamban továbbítjuk, melyre megoldás lehet egy extra RTP fejléc (2. ábra) vagy az úgynevezett IPv6 kiterjesztett fejléc (extension header). A nem szinkronizált visszacsatolt adatok továbbítására pl. ICMPv6 (Internet Control Message Protocol v.6) használható.

Az ábrán a hasznos videó-adat (Payload) és az egyes protokoll fejlécek (header) mellett az extra információ (extra information) kap helyet, mint specifikus RTP fejléc, tartalmazva a titkosítási (CIPHER), SRI, SSI, a ritkítás paraméterét UEP modulban (PUNC), illetve a szoft információhoz köthető adatokat. Ez utóbbi feltételezi, hogy a csatornadekódoló kimenetén 4 biten kvantált szoft információt is rendelkezésre áll. A szoft információt a rendszer a rádiós csatorna kimenetétől a video dekódolóig továbbítja. A szoft információk átvitelre az UTRAN szimuláció nem lett felkészítve, így ez a lehetőség jelenleg nem áll rendelkezésre.

4. Az alkalmazási rész

Ebben a fejezetben kerülnek részletezésre az együttes rétegek közötti vezérlő (JSCC/D) által közvetlenül vezérelt egységek, melyek klasszikus értelemben a szállítási réteg és az alkalmazási réteg közötti funkciók szerepét töltik be. A JSCC/D vezérlő az IP hálózat után a vezeték nélküli hozzáférési hálózatot a fizikai rétegbeli alegység-

vel tudja befolyásolni (lásd 1. ábra). Ezekre a modulokra koncentrálunk az ezt követő fejezetben.

A. Forráskódoló

Az 1. ábrán látható forráskódoló és dekódoló modulok legfelül helyezkednek el. A forráskódolásnak gyakorlatilag a használt videókódek felel meg, feladata a nyers videofolyam hatékony tömörítése. A vevőoldalon a videodekódolónak a kódolt, ám hibás adatfolyamot a lehető legkisebb észlelhető torzítással kell visszaadnia. A szimuláció fejlesztése során 3 kódoló került vizsgálatra: az MPEG-4, H.264/AVC és H.264/SVC kodekek.

B. UEP, tikosítás

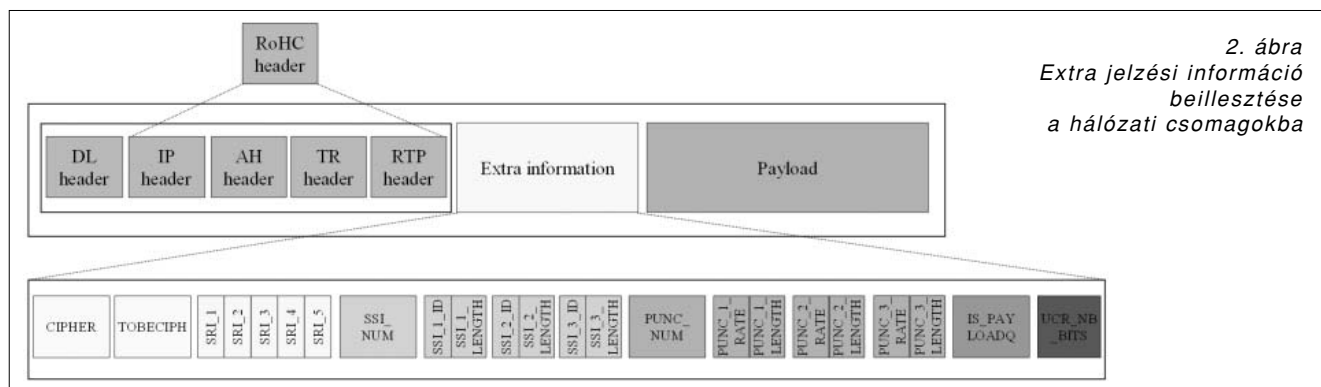
Egy tömörített videofolyam elemei különböző érzékenységet mutatnak a hibák (bithibák és csomaghibák) iránt. Ennek megfelelően különböző védelmet kell biztosítani a különböző érzékenységi osztályoknak. A feladatot az eltérő hibavédelem (Unequal Error Protection) blokk oldja meg. Az adatok védelme az UEP modulban ritkított (punctured) konvolúciós kódolóval van megvalósítva.

C. Streaming modul

A streaming modul feladata adóoldalon az adatfolyam IP-s csomagokba ültetése, a vevőoldalon pedig a fogadott IP-s csomagok alapján az adatfolyam visszaállítása a felsőbb rétegek részére. Az IP-s csomag azonban nem csak az IP protokollt rejti: a szállítási rétegbeli funkciók kívül ide tartoznak még az RTP és az RTCP is.

A PHOENIX szimulációs láncban mind az RTP, mind az RTCP protokollok implementálva vannak. Ezek a protokollok végponttól végpontig terjedő QoS-t biztosítanak a valós idejű multimédia folyamának időbéllyeggel, és sorszámmal címkézik fel a multimédia csomagokat.

A szimulációs láncban UDP, UDP-Lite és DCCP transzport protokollok vannak implementálva. Az UDP protokoll összeköttetés-mentes, best-effort szolgáltatást nyújt, amelynél nincsen sorszámozás, és duplikátumok is előfordulhatnak. Az UDP-Lite protokoll az UDP protokoll továbbfejlesztése, amely annyiban változott az eredeti UDP-hez képest, hogy egy részleges ellenőrző összeget (CRC) vezettek be. A részleges ellenőrző összeg használatával a teljes csomag helyett csak a fejrész és a payload egy részére képződik ellenőrző összeg. Amennyiben a csomagnak abban a részében ke-



lethezik hiba, amelyet az ellenőrző összeg lefed, a csomag eldobásra kerül.

Ha az UDP ellenőrző összege az egész csomagot lefed, az UDP-Lite működése megegyezik a hagyományos UDP működésével. Ha viszont csak a fejrészt védjük ellenőrző összeggel, akkor jelentősen javul a hatások. Ezzel az eljárással ugyanis körülbelül 40%-kal csökken az eldobások száma. Egy másik alkalmazott szállítási protokoll a DCCP, ami egy nem megbízható, ütközésvezérelt adatfolyam-szolgáltatást nyújt visszajelzéssel a helyesen megérkezett adatokról. A megvalósított DCCP protokoll nem tartalmaz újraküldési lehetőségeket a datagrammokra és az UDP-Lite-hoz hasonlóan tartalmaz egy részleges ellenőrző összeget. A DCCP ellenőrző összeg minden esetben lefed az egész fejléct, és a payload $n \cdot 4$ byte-os részét. Megjegyzendő, hogy a IPv4/IPv6 csomaggenerálás megköveteli, hogy a szállítási rétegbeli protokollok ellenőrzőösszeget alkalmazzanak, habár ez lehet részleges, mint ahogy az UDP-Lite és DCCP esetében említettük.

A szállítási protokoll fejrészt tartalmazó csomagok végül IP-s csomagba lesznek ültetve. A szimulációs lánc kizárólag az IPv6, azaz 6-os verziójú Internet Protokollt támogatja, mivel ez a jövő Internetének protokollja. Az Európai Bizottság minden általa támogatott információtechnológiai projektben az IP 6-os verzióját követeli meg.

A mai hatékony audió/video dekóderek sérült, hibás csomagok feldolgozására is képesek, növelve a képminőséget. Ennek következtében kihasználható a fentebb említett protokollok által nyújtott részleges hibaellenőrző összeg használata, aminek következtében a csomagvesztés csökken és a videó dekódoló által feldolgozott csomagok száma nő.

A RTP/RTCP protokollok jól használhatóak az UDP/UDP-Lite protokollokkal, viszont DCCP használata esetén nélkülözhetők, mivel a DCCP ellátja azokat a szük-

séges funkciókat, amik az RTP/RTCP-t szükségessé tennék. A szimulációs lánc jelenlegi verziója azonban gyakorlati okokból mindenképpen alkalmazza az RTP/RTCP protokollokat.

5. Az adattovábbítási rész

A. IPv6 hálózat

A szimulációban modellezett IP hálózat egy IP routerek ismeretlen halmazának tekinthető. A virtuális hálózat átviteli kapacitása, és buffer mérete szoftveresen tetszőleges értékre beállítható. Minden router okozhat késleltetést, illetve csomagvesztést. A végponttól végpontig terjedő késleltetés Gamma-eloszlású, míg a csomagvesztés egyenletes eloszlású minden csomópontban. A routerek száma, és a hozzájuk tartozó paraméterek szintén beállíthatóak. Ilyen módon minden hálózat modellezhetővé válik. A vezetékes közeg megbízhatósága magas. Emiatt az általunk szimulált 20 másodperces átvitel ideje alatt a csomagvesztés valószínűsége elhanyagolhatóan kicsi. Az IPv6 hálózat jellemzően a csomagok késleltetését változtatja meg, aminek következtében azok sorrendje változhat, illetve borsztössé válhat. Szimulációink során az UTRAN vezetéknélküli hálózati szegmensre koncentráltunk, így az IP hálózat torlódásból adódó csomagvesztését minimálisra (0-ra) állítottuk. Ezáltal a 20 másodperces videó szekvenciák ideje alatt nem volt az IPv6 hálózatban csomagvesztés.

B. UTRAN

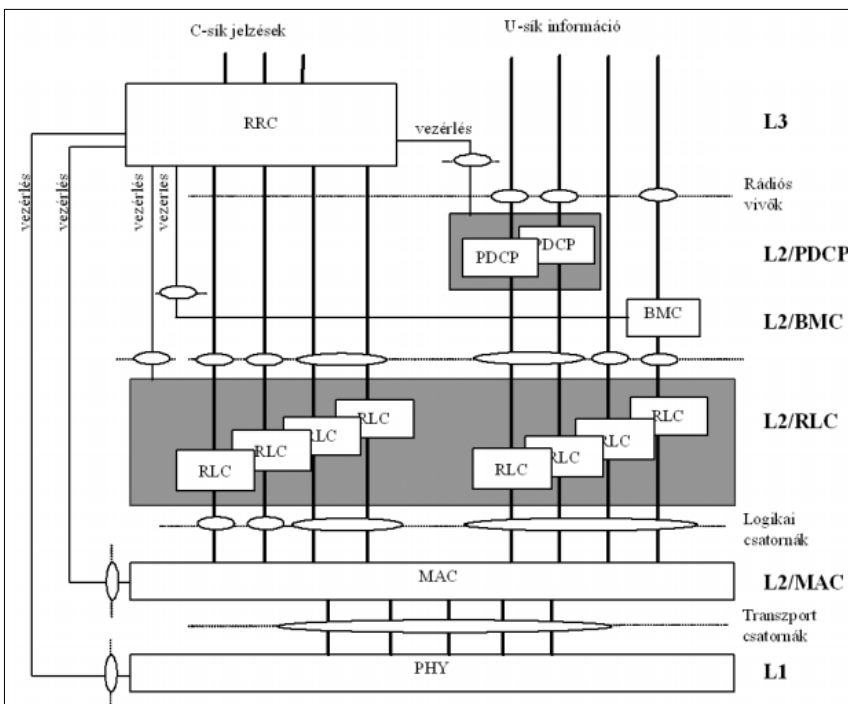
1) Felépítés és funkcionalitás

Az általunk implementált UTRAN modulok tervezési szempontjai a szabványosság, a teljesség és részletesség, illetve a hatékonyság voltak. A szabványosságot a [7-10] ETSI technikai specifikációk követése biztosította, míg a funkcionalitás teljességét egy korlátozott működési mód teljes körű implementálásával valósítottuk meg. Az UMTS-ből csak a dedikált adatátvitelt, illetve a hozzá kapcsolódó járulékos kontroll információk átvitelét implementáltuk. Célunk szempontjából a járulékos funkciók, mint a hálózati hívás felépítés nem fontosak. Vizsgálataink során feltételeztük a kapcsolat kiépített voltát, és csak az adat továbbítására koncentráltunk.

A 3. ábrán látható a UTRAN protokoll architektúrája, melyben az RRC réteg került módosításra a szabványhoz képest, hogy képes legyen fogadni a JSCC/D vezérlő utasításait és a PDCP, RLC, MAC, illetve fizikai rétegeket annak megfelelően konfigurálni.

A 3. ábrán látható a UTRAN protokoll architektúrája, melyben az RRC réteg került módosításra a szabványhoz képest, hogy képes legyen fogadni a JSCC/D vezérlő utasításait és a PDCP, RLC, MAC, illetve fizikai rétegeket annak megfelelően konfigurálni.

3. ábra
UTRAN protokoll-architektúra



2) Adatfolyam továbbítás

Az UMTS vízszintes rétegei függőleges irányban két részre oszthatóak, C-síkra, és U-síkra. A C-sík felel meg a vezérlésnek, az U-sík pedig az adatfolyam továbbításáért felelős. A legfelső szinten külön rétegek felelősek a vezérlésért, és adattovábbításért, az alsóbb szinteken pedig ugyanaz a réteg valósítja meg mindkét funkciót.

A harmadik (L3) rétegbeli rádiós erőforrást vezérlő réteg (RRC) irányítja az UMTS többi rétegét. Vezérli az RLC, MAC és fizikai rétegeket és ez a réteg valósítja meg a felsőbb rétegek felé a kapcsolatok jelzését, ezzel biztosítva az adatfolyam megfelelő továbbítását az UMTS rendszeren keresztül. Az adatcsomag konvergencia protokoll réteg (PDCP) az RRC réteg mellett helyezkedik el. Ez a réteg kapja felsőbb rétegektől az IP adatcsomagokat, amit fejléctömörítés után SDU csomagokban továbbít az RLC rétegnek.

A rádiós link vezérlő réteg (RLC) szabványai három támogatott módot írnak le, AM (Acknowledged Mode) nyugtázott mód, UM (Unacknowledged Mode) nyugtázatlan mód és TM (Transparent Mode) transzparens mód. A nyugtázott mód használata ellentmondana az együttes optimalizálás elvének, mivel az JSCC/D kontroller által elért adaptáció nem adhatna jobb eredményt a hagyományos átvitelnél, ha mindig újraküldjük a meghibásodott adatcsomagokat (legalábbis csak nagyon szélsőséges esetben, a rádiós csatorna tartósan rossz minősége esetén, amikor az újraküldés sem vezet jobb átvitelre). Az UM és TM módok használata célravezető számunkra. Az RLC réteg szegmentálja a felülről kapott adatcsomagokat (eltérően UM és TM módok esetén), sorszámozza azokat (csak UM esetén), majd PDU egységekben továbbítja azokat a MAC rétegnek. UM, illetve TM mód esetén hibás csomagoknál nincsen újraküldés. A közeghozzáférési réteg (MAC) a kapott PDU-kat transzportcsatornákra képezi le esetleges paddinggel kiegészítve, majd a kiválasztott transzportformátummal együtt továbbítja azokat a fizikai réteg felé.

A fizikai réteg sok más funkció mellett CRC-t fűz az adatfolyamokhoz, csatornakódolás után pedig leképezi fizikai csatornára azokat. A fizikai csatornákon levő adat QPSK moduláció után kerül továbbításra a rádiós interfészen keresztül. A komplex alapsávi jelként kiadott adat egy többutas fadinges rádiós átviteli közeget szimuláló modulba kerül, mely AWGN zajt ad a jelhez, mielőtt az a vevőhöz jutna. A vevő egy koherens gereblye-vevő (rake receiver), mely optimálisan becsli a csatorna jelútjainak (6 fő-terjedési, 20 alút) komplex csillapításait. A vevő után a vételi oldalon az egyes rétegek (fizikai, MAC, RLC, PDCP) inverz jelfeldolgozási folyamatai mennek végbe.

3) A modul vezérlése

Az UMTS rétegek vezérlését a rádiós erőforrás vezérlési réteg (RRC) valósítja meg. A Phoenix projekt szimulációs lánc együttes optimalizálást vezérlő paraméterei ezen a rádiós erőforrás vezérlési rétegen keresztül hatnak az UMTS rétegeire. Az UTRAN-ban az adat átvitelének módját az aktuálisan érvényes transzportformátum határozza meg, mely összefogja az RLC, a

MAC és fizikai rétegek beállításait. Az aktuálisan érvényes transzportformátumot a MAC réteg választja az RRC által beállított transzportformátum-halmazból. A kiválasztás algoritmusá figyelembe veszi az egyes adatok prioritását és az RLC bufferében lévő adat mennyiségét [4]. Ez az elsődlegesnek is nevezhető konfigurálás, melyet a MAC réteg végez viszonylag gyorsan nevezhető, mivel 10-80 ms időközökkel történik. Egy transzportformátum tartalmazza

- a csatornakódolás módját (turbó, konvolúciós, nincs kódolás);
- a csatornakódolás kódsebességét;
- az alapvető időtartomány alatt átvitt adatmennyiség (transzportblokk-méret, transzportblokkok száma);
- az alapvető időtartomány nagyságát (10, 20, 40 vagy 80 ms);
- az alkalmazott CRC méretét (0, 8, 12, 16, 24 bit);
- sebességillesztési paramétert (rate matching, puncturing) stb.

A RRC réteg azonban változtathatja a transzportformátumok azon halmazát, melyből a MAC választ. Ezt nevezhetjük másodlagos konfigurációnak, ami az elsődlegeshez képest lassan mehet csak végbe, mivel az adónak és a vevő egységnek szinkronizálnia kell ezt a halmazt egymással. A szimulációs láncunkban az JSCC/D által vezérelt adaptáció egy másodperces időközönként megy végbe, ami időben összemérhető az RRC által meghatározott formátumhalmaz kialakítással.

Az RRC réteg a JSCC/D-től kapott vezérlési információk alapján [7] választ egyet az előre elkészített konfigurációs beállítások közül. A beállításhalmaz a [11] alapján lettek meghatározva, többek között tartalmazza a transzportformátumok halmazát, az RLC réteg által használt módot, a csatorna típusát, a payload és fejlec méreteket bitben, valamint a maximális átviteli sebességet stb. Egy szimulációs futtatás során a transzportformátumok halmazán kívül más nem változik. Ez természetesen mind feltöltési (uplink) és mind letöltési (downlink) irányokban szabályozható.

A szimulációban használt konfigurációs beállítások a szabványban leírt értékeknek felelnek meg [11], letöltési irányban 8-2048 kbps, feltöltési irányban pedig 8-384 kbps átviteli sebességet biztosítva a rendszernek.

A JSCC/D kontroller a teljes torzítás (DS+C) figyelembevételével választja ki az átvitt adat mennyiségének, a sebességillesztési paraméternek illetve csatornakódolási rátának a tartományát. Ezen adatok alapján az RRC választ az előre elkészített konfigurációs beállításokból, azon szempont szerint, hogy melyik beállítás felel meg legjobban a paramétertartományoknak.

A CRC használata is hathat a rendszer teljesítményére. Mint már említettük, a fizikai rétegben történik a CRC megfelelő adategységekhez (transzport blokkokhoz) való csatolása az adóoldalon, majd a vételi oldalon a hibás (nem megfelelő CRC ellenőrzőösszegű) csomagokat az RLC nem továbbítja a felsőbb PDCP rétegnek. Azonban, ha nem használunk CRC-t, akkor több hibás adat jut el a video dekódolóig, mivel kevesebb kerül eldobásra az UMTS-ben. A csomagok persze el-

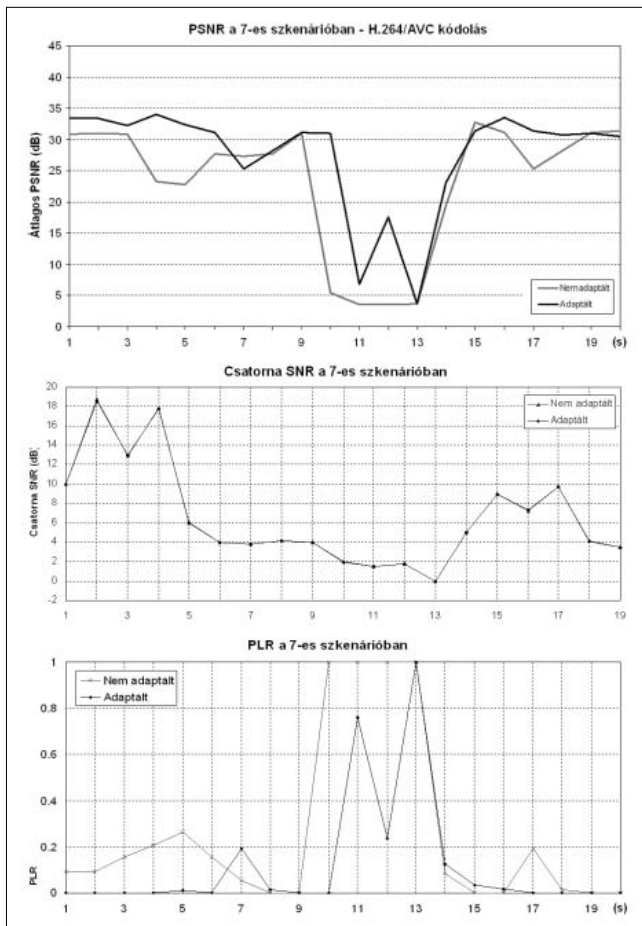
veszhetnek akkor is, ha nem használunk CRC ellenőrző összeget, mivel értelmezhetetlen vezérlő információk (szekvenciaszám, adathossz információk) esetén az RLC rétegben és a fejléc kitömörítéskor a PDCP rétegben is eldobódhatnak. A szimulációban az optimalizált eset teljesítményét javítottuk a CRC mellőzésével. A csomagvesztés ilyen módon való csökkentése természetesen teljesítménynövekedést okoz, azonban nem kizárólag ez felelős az adaptációval elérhető javulásért. A CRC használat hatásaira a [12] tartalmazza az általunk generált részletesebb szimulációs eredményeket.

6. Szimulációs eredmények

A. Szimulációs futtatások, tesztbeállítások

Ebben a fejezetben egy futtatási példát emeltünk ki, bemutatva a szimulációs lánc teljesítményét. Az egyes modulok beállításai a projekt által leírt 7-es scenárió alapján történtek, a részletes beállítások megtalálhatók a [13] dokumentumban. Ez a scenárió az úgynevezett sürgős videóátvitelt (pushed video information) próbálja utánozni, mint például egy élő hírműsor, melynek jellemzője az alacsony késleltetés, többes adás és streaming mód, illetve mobil felhasználók jelenléte. Természetesen szimulációnk egy mobil állomás, mint vevő és a „híreket” adó másik végpont között szimulál átvitelt.

4. ábra Szimulációs eredmények a 7-es scenárióban mérve 20 sec videószekvencia esetén: átlagos PSNR, rádiós csatorna SNR és csomagvesztés (PLR)



A nyers videószekvencia a „foreman” és az „akiyo” QCIF formátumú YUV fájlok voltak, melyeket H.264/AVC forráskódolóval tömörítünk a szimulációs lánc alkalmazási rétegében. A szimulációs futtatás RTP, UDP-Lite szállítási protokollokkal történt, mely után IPv6 csomaggenerálás következett. Szimulált IPv6 hálózaton vezetjük keresztül a médiafolyamot, amelyet az UTRAN hálózati szegmensében lévő hálózati terminál vesz. Az RLC réteg által használt módot viszont nyugtázatlan UM módra állítottuk, amely a csomagok sorszámozása mellett nem tartalmaz újraküldési lehetőséget a hibás csomagok számára. Mint már említettük, ez az adaptáció miatt szükséges, mivel az adathibákat nem újraküldéssel, hanem együttes optimalizálással, az egész rendszer újrakonfigurálásával akarjuk csökkenteni, újraküldéssel az RLC réteg viszont elrejtene a hibás csomagokat a felsőbb rétegek előtt, ahol a döntés születik az együttes optimalizálásról. Rádiós csatorna szintjén a [14] alapján implementált mikrocellás belvárosi környezetet szimuláltunk. A vezeték nélküli UTRAN szakasz letöltési (downlink) irányban vitte át a videószekvenciát az 1. ábrának megfelelően.

Az együttes rétegek közötti vezérlővel (JSCC/D) optimalizált futtatást jelöltük adaptált esetnek. A nyers videóadat forráskódolása másodpercenként történik, amivel szinkronban, azaz másodpercenként történik az egyes modulok átkonfigurálása adaptált esetben. A nem adaptált esetben a JSCC/D kontroller nem működött és 384 kbit/s-os átvitelnek megfelelő transzport formátumot használt a MAC réteg [11]. Nem adaptált esetben 8 bites CRC-t használtunk, illetve adaptációnál kerültük a CRC használatát az UTRAN rendszerbeli átvitel során.

Adaptív és nem adaptív átvitelünk csatornakódolás szempontjából ugyanazon szabványos turbó csatornakódolást használnak. Jelentős különbség adódna a két eset között, ha nem azonos erősségű hibajavítást használnánk, azaz például nem adaptív esetben konvolúciós kódolást. Az adaptivitást így a kódoló típusára nem terjesztettük ki, mivel a viszonylag rossz minőségű rádiós környezet (fading hatás, alacsony jel-zaj viszony) megköveteli az erős csatornakódolás alkalmazását.

Az átvitt adat mennyisége azonban változik az engedélyezett transzport formátumok alapján, itt a sebességillesztési paraméter által meghatározott bitkiszűrés (puncturing), illetve az átviteli időintervallum (TTI) alatt átvihető bitek száma stb. A forráskódoló szintjén természetesen ugyancsak jelentősen változik az adaptáció hatására, azaz a tömörítés mértékére az adott másodpercben átvitt videó adat mennyisége.

B. Futtatási eredmények

A 4. ábra jobb felső diagramja az átvitel egyes másodperceire vonatkozó átlagos jel-zaj viszonyt ábrázolja szimulált fadinges csatornán. A csomagvesztési arány (PLR – Packet Loss Ratio), mely az ábra jobb alsó részén látható, a két esetben eltér a csatorna ingadozása illetve az eltérő adatmennyiség miatt. A nem adaptált esetben a forráskódoló a rádiós csatornától függetlenül működve, mindig ugyanolyan minőségben kódol.

Adaptált esetben azonban az egyes másodpercekben különbözhet a videójel kódolása, illetve a fizikai rétegbeli konfigurációtól is függ a rádiós csatornára érkező adat mennyisége. Az átvitt videófolyam PSNR görbéje (az ábra bal oldali grafikonja) átlagosan 3-4 dB-lel jobb az adaptált esetben [GJ1].

5. ábra

Példa
a szimuláció
által elért
eredményekre:
nem adaptált
(bal) kontra
adaptált (jobb)
eset



Az 5. ábrán láthatóak a „foreman” QCIF és „akiyo” QCIF videó szekvenciák átviteléből pillanatképek ugyanazon képkockára vonatkoztatva adaptált és nem adaptált beállításoknál. A vizuális példán is látható, hogy bizonyos másodpercekben mennyivel jobb minőséget produkálhat a JSCC/D-vel vezérelt eset. Más futtatási környezetekre, beállításokra vonatkozó eredményeinket a projekt [13] dokumentuma tartalmazza.

7. Összefoglalás

A túlzott részletekre nem kiterjedő, de a szimulációs modell átfogó működését ismertető leírás után körvonalaztuk az UTRAN hálózati szegmens funkcionalitását és beillesztését a modellbe. Az egyes rendszerblokkok között áramló adat struktúráját is felvázoltuk, melyre a végponttól végpontig terjedő optimalizálás alapul. A valósághoz még jobban közelítettük rendszerünket egy általunk implementált, szabványokhoz nagymértékben igazodó, vezeték nélküli UMTS hálózati szegmens használatával. A UTRAN modul konfigurálhatóságát úgy alakítottuk, hogy megfelelően illeszkedjen a JSCC/D kontroller vezérlő interfészéhez.

Kapott eredményeink is mutatják az alkalmazási rétegek szintjén (forráskódoló) és fizikai réteg szintjén (csatornakódolás, moduláció stb.) használt együttes adaptáció nyereségét egy sokkal realisztikusabb környezetben, mint ami eddigi tudományos kutatásokban szerepelt. Remélhetőleg eredményeink a szabványosítók és a gyártók figyelmét is felkeltik majd, és a jövő rendszerei, berendezései alkalmazni fogják az együttes forrás- és csatorna(de)kódolás (JSCC/D) alapelvét.

Köszönetnyilvánítás

A szerzők köszönetüket fejezik ki a PHOENIX projekt minden résztvevőjének, akik értékes munkájukkal hozzájárultak a projekt szimulációs rendszerének kifejlesztéséhez, lehetővé téve jelen munka megszületését is.

Irodalom

- [1] C. E. Shannon, „A mathematical theory of communication”, Bell System Technical Journal, July-October 1948. Vol. 27., pp.379–423; 623–656.
- [2] S. B. Zahir Azami, P. Duhamel, O. Rioul, „Joint source-channel coding: panorama of methods”, in Proc. of CNES workshop on data compression, Toulouse, France, November 1996.
- [3] Final Draft Int. Standard of Joint Video Specification (ITU-T Rec. H.264, ISO/IEC 14496-10 AVC), Doc JVT-G050r1, Geneva, Switzerland, May 2003.
- [4] A. Zsiros, A. Fülöp, G. Jeney, „Easily Configurable Environment of UTRAN Physical Layer”, in Proc. 5th EURASIP Conf. Speech and Image Processing, Multimedia Communications and Services, Smolenice, Slovakia, 2005, pp.93–98.
- [5] M. van der Schaar, S. Krishnamachari, S. Choi, X. Xu, „Adaptive Cross-Layer Protection Strategies for Robust Video Transmission Over 802.11 WLANs”, IEEE Journal on Select. Areas Comm. Vol. 21., Dec. 2003, pp.1752–1763.
- [6] C. Lamy-Bergot, J. Huusko, M. G. Martini, P. Amon, C. Bergeron, P. Hammes, G. Jeney, S. X. Ng, G. Panza, J. Peltola, F. Sidoti, „Joint optimisation of multimedia transmission over an IP wired/wireless link”, European Symposium on Mobile Media Delivery, September 2006.
- [7] ETSI Technical Standards, TS 125.212 V6.1.0 (2004-03), UMTS „Multiplexing and Channel Coding (FDD)”
- [8] ETSI Technical Standards, TS 125.301 V6.0.0 (2003-12), UMTS „Radio Interface Protocol Architecture”
- [9] ETSI Technical Standards, TS 125.211 V6.0.0 (2003-12), UMTS „Physical Channels and Mapping of Transport Channels onto Physical Channels (FDD)”
- [10] ETSI Technical Standards, TS 125.213 V6.0.0 (2003-12), UMTS „Spreading and Modulation (FDD)”
- [11] ETSI Technical Standards, TS 134.108 V6.0.0 (2005-10), UMTS „Common test environments for User Equipment (UE) conformance testing”
- [12] M. G. Martini, M. Mazotti, C. Lamy-Bergot, P. Amon, G. Panza, J. Huusko, J. Peltola, G. Jeney, G. Feher, S. X. Ng, „A Demonstration Platform for Network Aware Joint Optimization of Wireless Video Transmission”, IST Mobile Summit, Juny 2006.
- [13] „Deliverable 4.4b: Experiments and test results”, PHOENIX project deliverable, December 2006.
- [14] ETSI Technical Report, TS 125.996 V6.1.0 (2003-09), UMTS „Spatial Channel Model for Multiple Input Multiple Output (MIMO) Simulations”

Vertikális handover – heterogén hálózatok együttműködése

FÜLÖP PÉTER, SZÁLKA TAMÁS

Budapesti Műszaki és Gazdaságtudományi Egyetem, Híradástechnikai Tanszék
{fulopp, szalkat}@hit.bme.hu

Kulcsszavak: vertikális handover, heterogén hálózatok, mobilitás, hálózathoz tartozó szolgáltatások, hívásátadás

A vezetékes, vezeték nélküli illetve mobil hálózatok együttműködése új szolgáltatások alapját teremti meg. A vertikális handover alkalmazásával a felhasználó szabadon mozoghat a különböző hálózatok között. Integrált, hálózathoz tartozó szolgáltatásokat nyújthatnak számára a szolgáltatók, az optimális sebességű vagy költségű hozzáférési hálózat felett. A cikkben a vertikális handover fogalmakörét, tulajdonságait, alkalmazásának lehetőségeit és korlátait tekintjük át.

1. Bevezetés

Napjainkban a mobil cellás hálózatok második generációja mellett a harmadik generációja is mindennapunk részévé kezd válni (3G UMTS). Emellett újabb és újabb fix telepítésű vezeték nélküli hozzáférési hálózatok (FBWA – Fixed Broadband Wireless Access) jelennek meg, melyek egyelőre alig támogatják a mobilitást, de a korábbiaknál lényegesen nagyobb sáv szélességű és áruk gyakran pótolja ezt a hiányosságot (WLAN, HIPERLAN).

A mobil és vezeték nélküli hálózatok közötti, korábban élesen megjelenő határ a hálózati technológiák fejlődése nyomán elmosódik. A korlátlan (gyakorlatilag globális) mobilitást támogató mobilhálózatok jellemzően alacsony sáv szélességet nyújtottak drágán. A vezeték nélküli, de mozgást nem támogató (erősen korlátozott lefedettségű) hálózatok sebessége és ára is lényegesen kedvezőbb volt. A hálózat sebessége szempontjából a két kategória képviselői az utóbbi egy-két évben közelítettek egymáshoz, a mobil cellás rendszerek sebessége nőtt drasztikusan (HSDPA esetében elméletileg 2 Mbit/s-ra), míg a fix vezeték nélküli hálózatok esetében új szabványok jelennek meg a mobilitás támogatására. A hozzáférések díján is a sebesség kapcsán lezajló folyamat figyelhető meg, jóllehet ez a konvergencia lassúbb.

A közeljövőben várható, hogy a mobilitást támogató cellás rendszerek vezetékes terminológiában is szélessávú hozzáférést adnak majd (például a 802.16 elméletileg 70 Mbit/s-ot biztosít). Jónéhányuk a szabad, illetve nem lezárt frekvenciasávokban is üzemeltethető, amely a mobil szolgáltatók számára is komoly versenyhelyzetet teremt, hatása az árakon és a szolgáltatások palettáján fog jólékonnyan jelentkezni.

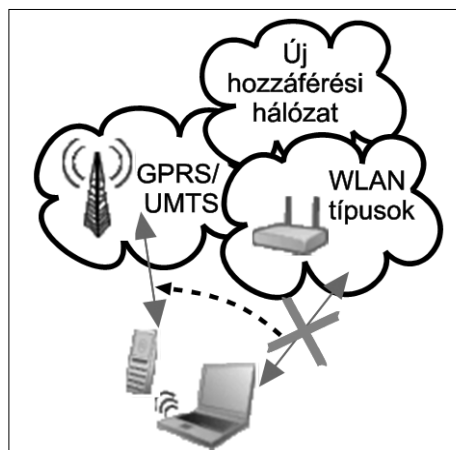
A piacon elérhető és hamarosan megjelenő hozzáférési technológiák (GSM/GPRS, UMTS, EDGE, 802.11 WIFI, 802.16 WiMAX, 802.20, HIPERLAN, HIPERLAN2) óriási választékot nyújtanak a felhasználók számára [1-3]. Az igény, hogy ezt a választékot szabadon használhassák, hamarosan megkerülhetetlenné válhat a kü-

lönöző hálózati szolgáltatókkal szemben támasztott elvárások sorában. A felhasználói igények mellett a szolgáltatók szempontjából is hasznos a különböző hálózatok erőforrásainak és tulajdonságainak kihasználása, valamint optimális összehangolása.

A vertikális handover (VH) fogalma magában foglalja mindazokat a képességeket és elvárásokat, melyek a különböző hálózatok együttes használatát, együttműködésüket valósítják meg.

2. Vertikális handover

A vertikális handover fogalmának értelmezéséhez a cellás hálózatokból ismert handover-t hívhatjuk segítségül. A mozgó felhasználók a mozgás során elhagynak cellákat és új cellák területére érnek. A cellahatáron a folyamatban levő hívás folytonossága miatt a korábbi cellából logikailag át kell adni a hívást annak a cellának, amelybe megérkezett a felhasználó. Ez a hívásátadási eljárás a handover, vagy hívásátadás. Ezt a típusú handover-t horizontális handovernek is szokták nevezni, mivel logikailag horizontális irányban, két szomszédos cella között történik a hívásátadás. Ezzel szemben a vertikális handover elnevezés a következő handover eljárást jelöli.



1. ábra
Vertikális
handover

Egy adott földrajzi ponton többféle hozzáférési hálózat érhető el. Ismert példával élve, a ferihegyi repülőtér termináljain elérhető egy (vagy több) WLAN hálózat, illetve a Budapest nagy részét lefedő 3G UMTS hálózat. Ekkor a repülőtér várójában egy alkalmas eszközzel – PDA-val, okostelefonnal – mindkét hozzáférési technológia használható lenne. A vertikális handover az általános értelemben vett hívás – vagyis hang- vagy adatkapcsolat – átadása egyik hálózatról a másikra (UMTS-ről WLAN-ra, vagy fordítva). A vertikális hívásátadás egymást lefedő hálózatok között, vagyis függőleges irányban történik, innen ered az elnevezés (1. ábra).

A legfőbb különbség a cellás hálózatok cellái között alkalmazott horizontális hívásátadás és a vertikális handover között az, hogy a hívás információinak, leírójának átadása egy szolgáltató egy technológiát használó hálózatán belül vagy különböző szolgáltatók különböző technológiát alkalmazó hálózatai között történik. Az alábbi esetek képzelhetők el:

1. egy szolgáltató egyfajta technológiát használó pontjai között – horizontális handover
2. egy szolgáltató különböző technológiájú hozzáférései között – például mobiltelefon-szolgáltató által üzemeltetett WLAN hálózatra történő átlépés a GPRS-ről
3. különböző szolgáltatók egy technológiát használó hálózatai között – például mobilszolgáltatók közötti váltás (ennek praktikusán nincs nagy jelentősége)
4. különböző szolgáltatók különböző technológiát alkalmazó hálózatai között – mint a fenti példákban

Az első eset a horizontális handover esete, míg a többi pont esetén vertikális hívásátadásról beszélünk. Ebből a megközelítésből a horizontális handover hálózaton belüli (intranetwork), míg a vertikális handover hálózatok közötti (internetwork) hívásátadást jelöl.

Az említett megkülönböztetésből adódóan egy vertikális hívásátadás sokkal kiterjedtebb feltételrendszerrel és komplexebb következményekkel jár, hiszen hálózatok közötti jelzést és menedzsmentet kell megvalósítani a kezeléséhez.

A vertikális handover hasonló funkciót lát el a felhasználók szempontjából, mint a horizontális handover. Egy

GSM hálózatban hozzászokhattunk, hogy a cellák közötti barangolás közben a lebonyolított hívásban semmilyen fennakadás nem érzékelhető. A vertikális handover szintén ezeket a váltásokat rejtje el a felhasználó előtt, így a hívás – általános esetben az alkalmazás – megszakadás nélkül és ideális esetben háborítatlanul folytatható. A hívásra illetve a hálózati alkalmazásra nézve az alábbi kategóriákba sorolhatjuk a hívásátadás különböző változatait (1. táblázat):

- *Hard handover*: az eszköz képtelen egyszerre két hálózattal kommunikálni, ezért a váltás előtt lebontja a régi kapcsolatot, utána építi fel az újat.
- *Soft handover*: az eszköz a handover idejére fenntartja mindkét kapcsolatot [4].
- *Seamless handover*: észrevétlen a váltás, nincs adatvesztés, gyors hívásátadás.

A fenti kategóriák közül a seamless handover rendelkezik azokkal a tulajdonságokkal, amelyeket a horizontális handover esetén megszokhattunk. A folyamatban levő hívás, illetve a futó alkalmazás számára átlátó módon történik meg a hálózatok közötti váltás.

A vertikális handover esetében sok esetben kivitelezhetetlen a horizontális handoverhez hasonló észrevétlen hívásátadás. A felhasznált hozzáférési technológiák eltérő átviteli sebessége kihat a kapcsolatra, egy lassúbb hálózat használata esetén a hálózati forgalom elkerülhetetlenül lelassul. Ez a hatás különböző, felsőbb rétegbeli intézkedésekkel csökkenthető, de teljesen nem szüntethető meg.

A hívásátadások típusai csoportosíthatóak aszerint, hogy a döntési intelligencia a hálózati oldalon vagy a mobil eszközben található meg. A horizontális handover esetében a döntés a hálózati oldalon születik meg, a bázisállomások adják át egymásnak a mobil eszközök kezelésének jogát (GSM, 3G UMTS). Ekkor hálózat által irányított hívásátadásról (Network Controlled Handover – NCHO) beszélünk. Abban az esetben, ha a mobil eszköz autonóm módon, a rendelkezésre álló információk (hálózati paraméterek, szolgáltatási díjak) alapján dönt a váltásról, akkor mobil terminál által vezérelt hívásátadás (Mobile Controlled Handover – MCHO) történik (WLAN hálózat hozzáférési pontjai között). A vertikális handover esetében az utóbbi módszer jellem-

1. táblázat
Vertikális handover
típusai
és tulajdonságaik

	Hard	Soft	Seamless
A kommunikációs partnerrel létesített kapcsolat biztosan megszakad			
Kapcsolat megszakadása bizonyos protokollok esetén lehetséges			
Nincs érzékelhető kapcsolatszakadás			
Biztos csomagvesztés a handoverből kifolyólag			
Nincs csomagvesztés a handover miatt			
Egyszerre csak egy élő link érhető el			
Párhuzamos linkcsatlakozás			
A kapcsolat tovább él (megakadás lehetséges)			

ző. A hálózatok közötti váltás esetén egyik hálózattól sem elvárható, hogy hívásátadáshoz szükséges vezérlési és jelzési információk birtokában legyen idegen hálózatok esetében is. A mobil eszköz felelőssége, hogy a megfelelő intézkedéseket tegye és átregisztrálja magát a másik hálózatba, annak aktív közreműködése nélkül.

3. Alkalmazási lehetőségek

A vertikális handover egy általános megoldás, amelynek segítségével a felhasználók mobilitást nem támogató hálózatok használata során is szabadon mozoghatnak. A feladata a mobilitás kezelése és elrejtése az azt nem támogató hálózatok előtt. Illusztrációként ebben a részben felvázoljuk, hogyan alkalmazható néhány területen a vertikális hívásátadás a mobilitást nem támogató hálózatok felett.

Magánhálózatok távoli elérése

A céges privát hálózatok szerepe folyamatosan nő az üzleti folyamatok lebonyolításában. A hálózaton céges adatok, erőforrások és más, a munkához elengedhetetlen kellékek találhatók meg, amelyekre a munkatársaknak a telephelytől távol is szükségük lehet.

A vertikális handover segítségével könnyedén kialakítható egy biztonságos protokoll, amely lehetővé teszi, hogy a felhasználó mozgás közben bármely távoli pontról bejelentkezzen a céges hálózatba. A kapcsolat tetszőleges ideig fenntartható, függetlenül attól, hogy a felhasználó közben mozog, és barangol a különböző hálózatok között.

Mozgó felhasználók szinkronizálása

Számos olyan szervezet és vállalat működik, melyek több tucat vagy több száz mozgó felhasználó valamilyen követését és kontrollját igényli. A vertikális hívásátadás segítségével a felhasználók állandó hálózati kapcsolattal rendelkezhetnek, miközben utaznak és vándorolnak a hálózatok között. Ha jobb minőségű vagy olcsóbb hálózati lefedettséget találnak, akkor válhatnak, ezzel pénzt és időt takarítva meg. A jobb paraméterekkel rendelkező hálózatokat ki lehet használni többlet információk átvitelére is. Gondoljunk például fuvarozócégek fuvarosait, ügynököket, taxitársaságok alkalmazottjait összekötő állandó hálózati kapcsolatra.

Egy taxi fedélzeti számítógépe folyamatos GPRS kapcsolattal rendelkezik a diszpécserközpont felé. Hálózaton keresztül kapja meg a megrendeléseket, küld helyzetinformációt vagy autó-diagnosztikát a központnak, ha az átvitel nem tűr késleltetést. Ha a taxitársaság által üzemeltetett WLAN hotspothoz ér, akkor az ingyenes és gyors kapcsolaton keresztül letölthet reklámfilmeket, híreket, információkat, amelyek a sofőr vagy az utas számára lehetnek hasznosak. Így fontos információk késleltetés nélkül juthatnak el a címzetthez a GPRS kapcsolaton, az egyéb adatok átviteléért pedig nem kell drága forgalmi díjat fizetni.

4. Mobilitás támogatásához szükséges hálózati követelmények

A különböző vezeték nélküli és mobil hálózatok szabványai rendelkeznek azokról a funkciókról, amelyek az adott hálózaton belül a felhasználók mozgásának kezelését végzik (GSM hálózat bázisállomásai, egy WLAN hálózat hozzáférési pontjai között). A vertikális handover feladata a mobilitást támogató funkciók kiterjesztése a hozzáférési hálózatok feletti szintre.

Az adott hálózaton belüli barangolást figyelmen kívül hagyhatjuk, mert azt az adott hálózat átlátszó módon kezeli. Így bármely hálózathoz csatlakozunk, azon folyamatos kommunikáció zajlik. Röviden áttekintjük, milyen problémák vetődnek fel a hálózatok közötti mozgások esetén.

4.1. Kapcsolatleíró

A mobil hálózatokon folyó kommunikációs folyamatok túlnyomó többsége kapcsolat-orientált. A partnerek az adatcserét megelőzően felépítenek egy kapcsolatot, amely logikai csatornát biztosít a végpontok között (hanghívás GSM hálózaton, TCP kapcsolat WLAN hálózaton). A kapcsolat leírója tartalmazza azokat az információkat, amelyek alapján a végpontok és a közbenső csomópontok az átvitelt vezérlik (GSM VLR/MSC, TCP/IP terminológiában a címek és portok párosa). Horizontális handover esetén a hálózat vezérlő funkciói biztosítják a kapcsolat leírójának átadását a cellák között. A vertikális handover során szintén át kell adni a kapcsolat-leíró a hálózatok között ahhoz, hogy a horizontális handoverhez hasonló, kapcsolatmegszakadással nem járó hívásátadás történjen. Ekkor azonban eltérő szabvány által definiált, részben (fizikai és adatkapcsolati rétegben) inkompatibilis hálózatok között történne a hívásátadás, ami újabb kérdéseket vet fel.

A kapcsolatleíró különböző hálózatok esetén különböző szabványokat követhet. A transzformáció sok esetben nem lehetséges, amely eleve megghiúsítja a vertikális handover lehetőségét (áramkörkapcsolt GSM hívás átadása más hálózatra nem lehetséges). A handover akkor valósítható meg, ha a kapcsolat leírója olyan szabványt követ, amely több hálózaton is használható. Jó példa erre az Internet Protocol, amely a mai hálózatok nagy részében megtalálható és a jövő hálózatainak közös alapját adja. IP alapú kommunikáció esetén a kapcsolatleíró átadható minden olyan hálózat között, amely ismeri az IP protokollt.

4.2. Sebességkülönbség

A különböző hálózatok rádiós és adatkapcsolati tulajdonságai, illetve az alkalmazott rádiós technológia miatt több nagyságrendnyi a sebességkülönbség. A skála a GPRS néhány 10 kbit/sec sebességétől kezdve a WiMAX elméleti 70 Mbit/sec sebességéig, vagy még tovább terjed.

A horizontális handover hálózaton belüli hívásátadást végez, így az átadás előtt és után is azonos marad az átviteli sebesség. Vertikális handover esetén az

előbb felvázolt sebesség-skálán is történik ugrás, nagyobb vagy alacsonyabb sebességű hálózatra vált a mobil eszköz.

A különböző sávszélességekből adódó hatások CBR (Constant Bit Rate – állandó bitsebességű) forgalom esetén (például hanghívás) esetén nem jelentkeznek, ha a hálózatok rendelkeznek kellő kapacitással az átvitelhez. VBR (Variable Bit Rate – változó bitsebességű) forgalom esetén az adott idő alatt átvitt információ mennyisége időben változó, gyakran a legnagyobb elérhető sebességgel forgalmaznak a végpontok (például best-effort TCP kapcsolatok). A változó sávszélességű hálózatok közötti váltás esetén az aktuális átviteli sebesség jelentősen meghaladhatja az elérhető maximumot, vagy éppen sokkal kisebb nála. Amennyiben a handover során nő a sávszélesség, az nincs negatív hatással az átvitel paramétereire, hiszen nagyobb sávszélesség áll rendelkezésre. Ellenkező esetben a nagy bitsebességű adatfolyam átviteléről kell gondoskodnia egy keskenyebb sávú hálózatnak, amelyen nem garantál megfelelő sebességet. Az átvitel a fizikai lehetőségek miatt elkerülhetetlenül lelassul, a felhasználó számára is észrevehető módon.

A sebességkülönbség áthidalására és hatásának csökkentésére több lehetőség adódik.

A TCP forgalomszabályozásra és torlódásmenedzsmentre alkalmas funkciói segítségével automatikusan adaptálódik a megváltozott sávszélességhez. Az adaptáció közvetett beavatkozással gyorsítható a handover megelőző néhány tizedmásodpercben alkalmazott mesterséges csomagkésleltetéssel vagy csomagdobással. A TCP a csomag eldobása után beindítja a torlódásvezérlési algoritmust (exponential backoff – exponenciális visszaesés), így a handover pillanatában lényegesen kisebb átviteli sebességgel forgalmaz az állomás, amely nagyobb eséllyel vihető át a kisebb sávszélességgel rendelkező hálózaton. A proaktív bitsebesség-csökkentés kisimítja a hívásátadás során a sávszélesség visszaesésének hatását, mely nélkül a csomagvesztések miatt hosszú időre feltartaná a folyamatos átvitelt.

Hang- és videofolyamok (stream-ek) esetén a végponti kódolás hangolása segíthet a sebességkülönbség elrejtésén. A kodek a sávszélesség függvényében adaptív módon kódolja a folyamat, vagyis nagy sávszélesség esetén kis tömörítéssel jobb minőségben küldi át az adatokat, míg kis sebesség esetén erősebb tömörítést alkalmazva rosszabb minőségben. A minőség ugyan romlik, de a valós idejű folyamatok esetében fontosabb késleltetés és jitter nem növekszik jelentősen. Az adaptív kodek nemcsak a kapcsolat végpontjaiban működhet, a közbelső állomások és csomópontok rendelkezhetnek stream-átkódolási képességekkel.

4.3. Handover késleltetés

A hívásátadás – horizontális és vertikális – a hálózatban jelzésüzenetekkel jár, az átregisztrálási procedura időbe telik. A követelmények sorában hangsúlyos elemként szerepel az, hogy az átregisztrálási késleltetés az elérhető legkisebb legyen.

Horizontális handover esetén az átregisztrálás néhány milliszekundum, a hálózati intelligencia a felhasználó által nem érzékelhető késleltetéssel elvégzi feladatát. A vertikális handover a korábban említett tulajdonságok miatt nagyobb késleltetéssel jár. A hálózatok közötti váltást a mobil eszköz vezérli, amely eleve lassúbb, mint a hálózat belső jelzésforgalma.

A késleltetés akkor nő meg és válik zavaróvá a felhasználó számára is, amikor a váltásról nem értesül az elhagyott hálózat, illetve a mobil kommunikációs partnere (Correspondent Node – CN). A hívásátadás után a CN az elhagyott hálózaton keresi a mobil eszközt, nem ismerve annak új helyzetét. Az elhagyott hálózat sem tud információval szolgálni, így mindaddig megszakad a kommunikáció, amíg a mobil eszköz újból be nem jelentkezik a CN-nél. Közvetve csökkenti a késleltetést, ha az elhagyott hálózatot értesíti a mobil eszköz a hívásátadásról és új helyzetéről, az elhagyott hálózat a mobil után küldheti az annak érkező adatokat. A CN értesítésével lekerülhet, hogy a mobil régi pozíciójába küldjön csomagokat, mindig az aktuális pozíciót lehet címezni. A felvázolt módszert a MobileIP követi, minden handover esetén értesítést küld a CN-nak a mobil eszköz új helyzetéről (Binding Update üzenet) [5,6].

4.4. Lefedettség

A handover tulajdonságait és jellegét jelentősen befolyásolják a rádiós lefedettségi viszonyok. Hívásátadás nemcsak az optimális hozzáférés használatához szükséges, hanem a lefedettség megszűnése esetén is. Ha a felhasználó elhagyja a lefedettségi zónát, akkor a váltást időben kell kezdeményezni, hogy a kapcsolat ne szakadjon meg, a mobil eszköz ne maradjon hálózati hozzáférés nélkül.

Az automatikus vertikális handover érdekében a mobil eszköz felruházható olyan funkciókkal, melyek a rádiós paraméterek vizsgálatát végzik. A lefedettségi zóna elhagyása a rádiós jellemzők romlásával – a jelerősség, a jel-zaj viszony csökkenésével – jár, amely felismerhető a mobil eszköz fizikai interfészén. A paraméterek alapján döntési függvény definiálható, amely bizonyos határértékek alatt inicializálja a handover-t. Felhasználható az országos lefedettségű GPRS (esetleg UMTS) kapcsolat, amely nagy valószínűséggel bárhol elérhető. A lefedettségi zónák elhagyása esetén általában gyorsan kell dönteni a handoverről, gyors mozgás esetén túl nagy lehet a handover késleltetés és a forgalom megakadását okozhatja.

Egy hálózat lefedettségi zónájába érkezve a rádiós paraméterek alapján működő döntési függvény kezdeményezheti a hívásátadást, ha forgalmi, minőségi vagy díjazási szempontból előnyösebb hozzáférést biztosít.

5. Mobilitás heterogén hálózatok között

A különböző hálózatok közötti vertikális handover működése az OSI modell rétegeinek definíciója miatt szigorúan nem tartozik egyik szinthez sem, több helyen is

megvalósítható. A megoldásnak bizonyos követelményeket ki kell elégíteni:

- az átlátszó átvitel, azaz a hálózatok közötti váltás lehetőleg ne okozzon adatvesztést;
- alacsony legyen a késleltetés;
- a mobil eszköz hívásátadásoktól függetlenül egy statikus azonosító segítségével folyamatosan elérhető legyen;
- a hálózat szempontjából a vertikális handover megvalósítás a lehető legkevesebb módosítással járjon.

Ennek megfelelően röviden tekintsük át a következő alfejezetekben hogy az egyes OSI rétegek mennyire alkalmasak egy, minden követelményt kielégítő vertikális handover megvalósítására.

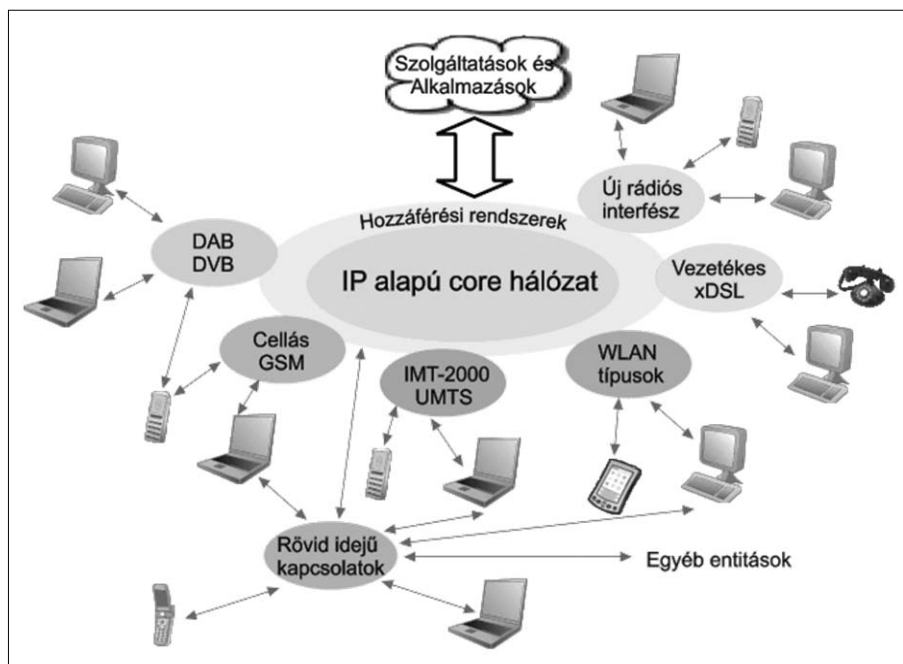
Adatkapcsolati réteg

A fizikai réteg feletti mobilitáskezelés csupán lokális megoldásként szerepelhet, egy adott alhálózaton belüli szabad mozgás valósítható meg a segítségével. Hozzáférési hálózatok közötti átjárásra önmagában teljesen alkalmatlan. Más, magasabb rétegbeli megoldással karöltve azonban gyors és hatékony megoldások születhetnek, erre jó példa a Mobile IP és CIP (Cellular IP) együttműködése [7]. A Mobile IP-CIP alkalmazásának hátránya, hogy a globális megvalósítása nehézkes, a hálózat állomásainak nagy részét fel kell készíteni az új protokollok kezelésére.

Hálózati réteg

A hálózatokban egyre erősebb az IP dominanciája. A jövőbeni fejlődés az integrált NGN (Next Generation Networks) IP-alapú csomagkapcsolt hálózatok irányába mutat [8]. A „homokóra alakú” protokoll-modell alapján az IP egyeduralmukodóvá válik a hálózati átviteli protokollok között (2. ábra).

2. ábra „All-IP core” hálózati struktúra



Mindezek azt jelentik, hogy a hozzáférési technológiák sokasága egységesen IP-alapon működik és fog működni, ebből következően az IP-alapon megvalósított handover megoldásoknak nagy jelentősége lesz, vagyis érdemes IP-alapon keresni a megoldást az átjárhatóság kérdésére.

A vertikális handover szempontjából a cél úgy megoldani a hálózatsváltást és az ezzel járó IP-cím-változást, hogy arról a felhasználó és az általa futtatott alkalmazás ne értesüljön. Amennyiben mégis értesül, az ne zavarja meg. A mobil eszközön a cím-váltást elrejtetni csak belső erőforrást igényel, egy szoftvermodult, amely a váltást kezelve a felhasználói alkalmazás felé állandó IP-címet, egy új, virtuális réteget mutat.

Nagyobb körületekintést igényel a távoli kommunikációs partner kezelése. Ebből a szempontból a megoldásokat két csoportra oszthatjuk. Első megközelítés, hogy a mobil eszközön futó megoldás funkcionális tükörképét alkalmazzuk. Ez nehezebben megvalósítható, hiszen minden kiszolgáló szoftverét ki kell egészíteni, ugyanakkor nem minden esetben oldja meg a kapcsolat azonosíthatóságát. A kapcsolat azonosítására azért van szükség, hogy a kiszolgáló tudja, hogy melyik csomag, melyik mobil eszköztől, milyen céllal érkezett. A TCP/IP szokványos kapcsolatazonosítójának része az IP-cím. Ha tehát a handover miatt változik az IP-cím, a kiszolgáló nem tudhatja, hogy még mindig ugyanazzal a mobil eszközzel kommunikál. Erre a csoportra jó példa a Host Identity Protocol.

Másik lehetőség, hogy egy állandó címmel rendelkező harmadik entitást (nevezzük vertikális handover szervernek) vonunk be a kommunikációba, amit vagy az otthoni hálózatunkba, vagy a teljes hálózat egy optimális pontjára helyezünk el. A mozgó készülék minden pillanatban egy védett folyamattal informálja a szervert az aktuális elérhetőségéről. A mobil terminállal kapcsolatba lépő CN a szervernek küldi az adatokat, amely továbbítja a mobil eszköz utolsó ismert címére. Fordított irányban a kommunikáció szintén a szerveren keresztül zajlik.

A megoldásnak két alapvető hátránya van:

- a csomagok nem optimális utat tesznek meg, mivel mindkét irányban is átmennek a szerveren;
- a szerver meghibásodási pontot jelent (single point of failure).

Ez utóbbi megoldás esetén az alkalmazások egy virtuális IP címen keresztül kommunikálnak a távoli kommunikációs partnerrel. Ahogy előbb említettük, a harmadik entitás segítségével megoldható, hogy egy egységes címen keresztül folyamatosan elérhető legyen a mobil állomás.

Szállítási réteg

A felhasználó mozgásáról érdemes értesíteni a szállítási réteget, mivel tradicionálisan ennek a rétegnek a feladata a torlódásvédelem. Ahhoz, hogy a torlódásvédelem hatékony legyen, adatok kelljenek a két végpont közötti útról. (A mobilitás akármelyik rétegben legyen megvalósítva, a torlódásvédelemmel ellátott átviteli protollokban mindenképpen szükséges a változtatás). Jellemző példa a szállítási rétegbeli mobilitásra az mSCTP (Mobile Stream Controll Transmission Protocol). Az eredeti SCTP protokollt arra tervezték, hogy TCP-t és esetleg még az UDP-t is leváltsa. Hasonlít a TCP-re, de képes például multistreaming és multihoming támogatására. A multihoming az az új tulajdonság, ami miatt az SCTP alkalmas lehet mobilitás kezelésére [9].

Egy állomás akkor „multihomed”, ha több hálózati rétegbeli címmel rendelkezik. Egy transzport protokoll akkor támogatja a multihomingot, ha egy transzport végponthoz több hálózati rétegbeli címet lehet hozzárendelni. A vertikális handover esetében lehetőség nyílik arra, hogy a végpont megváltoztassa az IP címét miközben a transzport végpont-végpont kapcsolat nem szakad meg.

Viszonyréteg

A viszonyrétegben megvalósított vertikális handover hasonló előnyökkel rendelkezik, mint a szállítási rétegbeli mobilitás, azonban kevésbé bonyolult.

A címváltozás esetén létrehoz egy új transzport kapcsolatot, a régit megszünteti. Működő megoldás nem sok létezik, az egyik példa a Pennsylvaniai Egyetemen kidolgozott DHARMA (Distributed Home Agent for Robust Mobile Access), mely a MobileIP megoldás és a viszonyréteg előnyeit ötvözi [10].

Alkalmazási réteg

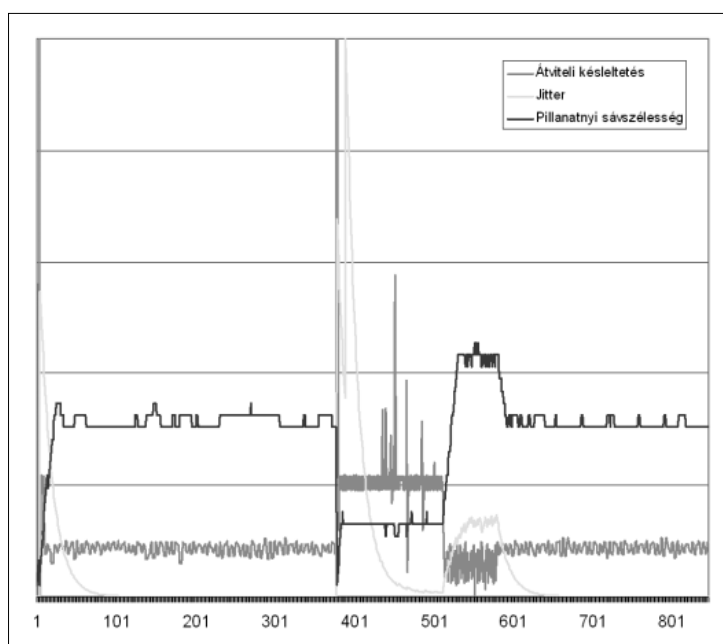
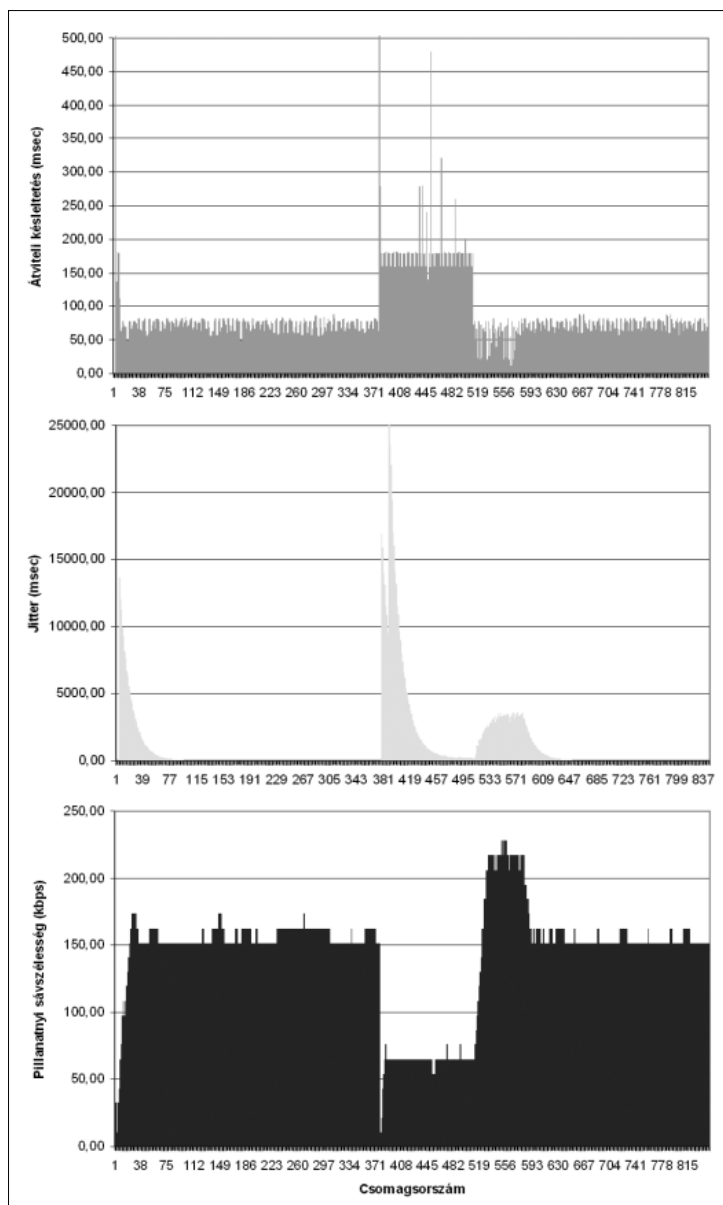
Az alkalmazás-szintű megoldások lényege, hogy a konkrét alkalmazás kezeli a vertikális handovert.

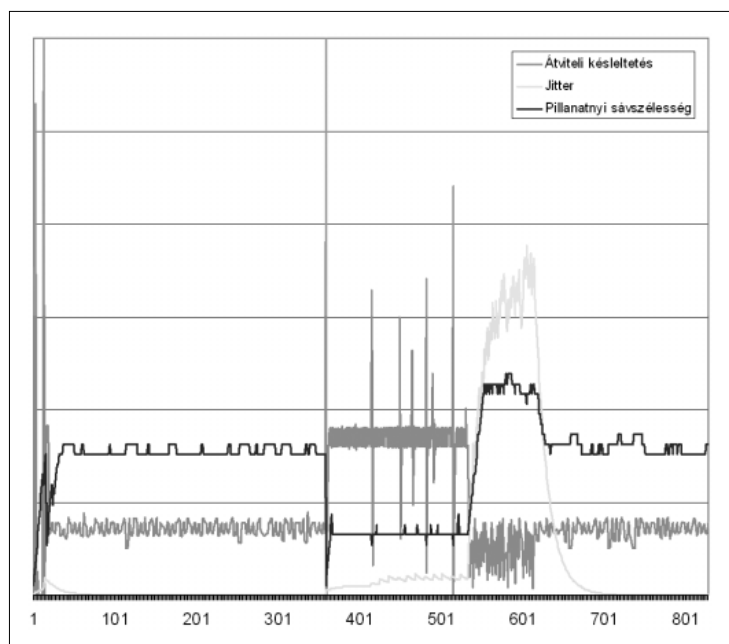
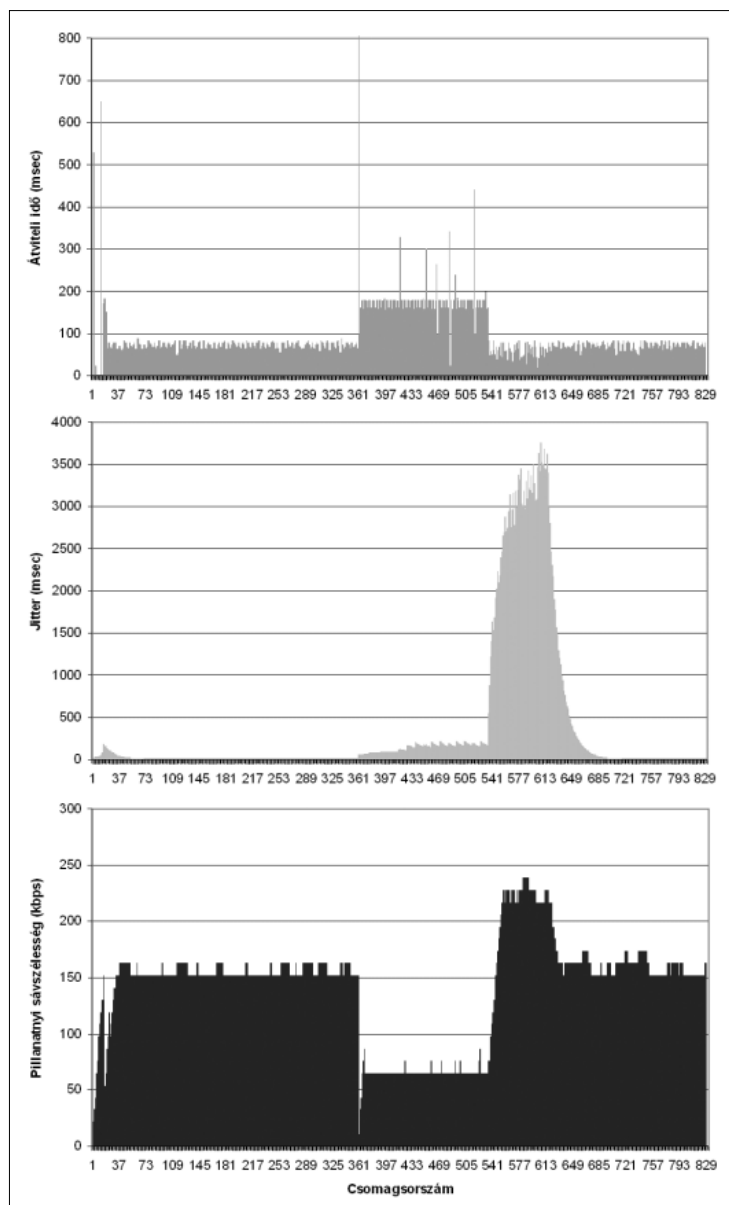
Ismertek a Session Initiation Protocol (SIP) protokollt használó multimédiás alkalmazások [11], vagy például a H323 stream vezérlő szabvány [12]. Előnyük az egyszerű mobilitás-kezelés, hátránya viszont a korlátozott használhatóság.

Első mérés

3. ábra
Átviteli idő, jitter és pillanatnyi sávszélesség értékek az első mérés során

4. ábra
A handover során tapasztaltak időbeli korrelációja az első mérés során





6. Az alkalmazott vertikális handover vizsgálata

Ebben a részben bemutatjuk a vertikális handover egy alkalmazási esetét. A vertikális handover a korábban felvázolt szempontok alapján a mobil eszközön az IP-rétegben működő szoftver vezérli. A vezérlés soft handoveret valósít meg, a hívásátadás idejére párhuzamosan két hálózati kapcsolatot tart fenn. A párhuzamosan fenntartott kapcsolattal a rendszerben a csomagvesztés esélye minimálisra csökken, a vertikális handover okozta késleltetés növekedése azonban lehetséges.

A mérési környezetben GPRS és 802.11b szabványú WLAN hozzáférést alkalmazunk. A folyamatos hálózati forgalmat egy 16 kbit/sec bitsebességű video- és egy ugyanekkora sebességű egysávos hangfolyam letöltésével generáljuk a mobil eszközön. A folyamatok a kísérletek ideje alatt változatlan kódolással érkeznek, vagyis nem történik adaptív átkódolás. Elméleti sávszélességet tekintve (GPRS 40 kbit/sec, WLAN 11 Mbit/sec) mindkét hálózat alkalmas a folyamatok együttes átvitelére, a gyakorlatban azonban a valóban rendelkezésre álló effektív sávszélesség mindkét hálózati hozzáférés esetében kisebb. A GPRS hálózaton ezért várható némi csomagvesztés a szűk kapacitásból adódóan.

A kísérletek során a folyamatok letöltését a WLAN hálózaton kezdjük, majd vertikális handoveret hajtunk végre. Ezután a GPRS hálózaton érkezik a stream mindaddig, míg egy ismételt vertikális handover történik, és ismét a WLAN hálózat lesz aktív.

Az 3-6. ábrákon két mérési alkalom eredményei láthatóak. Az ábrákon sorrendben az átviteli késleltetés (az átvitelhez szükséges idő), a mért jitter és a pillanatnyi bitráta látható. Mindhárom paramétert ábrázoltuk közös diagramban is, amelyen az időbeli változásuk közötti kapcsolat követhető nyomon.

Értékelés

Mindkét mérési esetben jól látható a két vertikális handover pillanata. Az átviteli idők megnőnek a GPRS esetében, hiszen nagyobb utat járnak be a csomagok, többszöri protokollkonverzió esnek át, illetve kisebb bitsebességgel haladnak.

Második mérés

5. ábra
Átviteli idő, jitter és pillanatnyi sávszélesség értékek a második mérés során

6. ábra
A handover során tapasztalt időbeli korrelációja a második mérés során

A pillanatnyi bitsebesség értékein láthatjuk, hogy a GPRS kisebb átviteli teljesítményt tud biztosítani. A kis sáv szélesség a video- és a hangfolyam paraméterein is érzékelhető volt, megakadást, szakadozást tapasztaltunk a stream vizsgálatokor.

A jitter mérési eredményeiben közös, hogy a GPRS-WLAN hívásátadás után megnöttek az értékek. A lassú GPRS hálózatról a gyors WLAN hálózatra váltva az átviteli késleltetés drasztikusan lecsökken, aminek eredménye a késleltetés ingadozás, vagyis a jitter növekedése.

7. Összefoglalás

A vertikális handover egy eszköz a szolgáltatók és a szabványosító testületek kezében, amellyel új szolgáltatások alapját teremthetik meg. Segítségével ötvözhetőek a különböző hálózatok előnyei, integrált hálózati szolgáltatások alapját teremtve meg.

Áttekintettük a vertikális handover által felvetett problémákat, amelyek megoldásáról gondoskodni kell a használható szolgáltatások kialakításához. A megvalósítás több módon is elképzelhető, különböző rétegekben működő protokollok révén. Az előnyök és hátrányok alapján a leghatékonyabb megoldás az IP rétegben működhet.

Irodalom

- [1] Alan Sicher, Randall Heaton,
„GPRS Technology overview”,
White Paper February 2002.
- [2] www.4gforum.org.
- [3] www.wimaxforum.org.
- [4] G. Cunningham, P. Perry, L. Murphy,
„Soft, Vertical Handover of Streamed Video”,
Dept. Of Computer Science,
University College Dublin, Ireland, 2004.
- [5] C.E. Perkins,
„Mobile IP”,
IEEE Communications Magazine, 1997.
- [6] Chung-Kuo Chang,
„A mobile-IP based mobility system for wireless metropolitan area networks”,
ICPP 2005 Workshops, 14-17. June 2005.
- [7] T. Campbell, J. Gomez, A. G. Valkó,
„An Overview of Cellular IP”,
IEEE 1999., pp.29–34.
- [8] Cui Hongyan, Cai Yunlong,
Wang Ying, Zhang Ping,
„Design and implementation of all IP architecture for beyond 3G system”,
Personal, Indoor and Mobile Radio Communications, PIMRC 2004.
15th IEEE Int. Symposium, 5-8. September 2004.
- [9] Li Ma, Fei Yu, Leung, V.C.M., Randhawa, T.,
„A new method to support UMTS/WLAN vertical handover using SCTP”,
Wireless Communications, IEEE Vol. 11., Issue 4, August 2004, pp.44–51.
- [10] Yun Mao, Bjorn Knutsson,
Honghui Lu, J. M. Smith,
„DHARMA: Distributed Home Agent for Robust Mobile Access”,
IEEE INFOCOM 2005, Miami, Florida, March 2005.
- [11] Schulzrinne, H., Rosenberg, J.,
„The Session Initiation Protocol: Internet-Centric Signaling”
IEEE Communications Magazine, 2000.
- [12] Wanjiun Liao,
„Mobile Internet telephony: mobility extension to H.323”,
Vehicular Technology, IEEE Transactions, Vol. 50., Issue 6, November 2001, pp.1403–1414.

Tájékoztató a Híradástechnika szerzőinek

A Híradástechnika szerkesztőbizottsága szeretné, ha egyre több szerzője lenne különböző területekről, így tovább bővülne az újságban megjelenő témák köre, és változatosabbá válna az eltérő szemléletű szerzők gondolatvilágától. Leendő szerzőink számára a cikkírással kapcsolatban szeretnénk néhány tájékoztató gondolatot közölni:

• **Témák:** A lap profilja egyfelől felöleli a távközlés „klasszikus” műszaki témaköreit, továbbá az informatika távközléshez, kommunikációhoz kapcsolódó vonatkozásait, a média-technológiák és média-kommunikáció kérdéseit, ezzel is elősegítve a távközlés-informatika-média konvergenciájának folyamatát. Másfelől helyet adunk a távközléshez és média-kommunikációhoz kapcsolódó gazdasági, szabályozási, marketing, menedzsment témáknak és a távközlés-informatika-média társadalmi vonatkozásainak is.

• **Terjedelem:** A szakmai cikkek az újságban általában 4-8 oldal terjedelemben jelennek meg. Ennél rövidebbek inkább csak a hírek vagy beszámolók lehetnek. 8-10 oldalnál hosszabban pedig csak olyan alapvető újdonságok írhatók le, ahol a megértéshez az elméleti alapok és a gyakorlati megvalósítás egyaránt szükséges. Ez azt jelenti, hogy ábrák nélkül 12-20 ezer karakter lehet egy cikk szövege. Nyomtatott oldalanként max. 1-3 ábra elhelyezése teszi az olvasó számára áttekinthetővé, vonzóvá az ismertetést.

• **Forma:** Sem betűtípus, sem rajzkivitel nem köti a szerzőket. Az újság egységessége kedvéért ugyanis az elektronikus érkező szövegeket a layoutban használt betűtípusban dolgozzuk fel. A cikkeket minden esetben elektronikus formában is kérjük, tehát e-mailen, vagy lemezen. A szövegeket *word formátumban* kérjük elkészíteni. Az ábrák megrajzolásánál egyetlen köztétel, hogy az újság *fekete-fehér kivitelben* jelenik meg, így a színes ábrák is szürkeárnyalatos képként lesznek láthatók az oldalakon. Ennek megfelelően kérjük a szerzőket, hogy lényeges dolgokra ne hivatkozzanak úgy, hogy a piros vonal, vagy a kék alapterületű rész, ehelyett szaggatott, pontozott, vastag és vékony vonalak legyenek megkülönböztethetők, illetve a területnél sraffozással lehet különbséget tenni. Fotóillusztrációk esetén lehetőség szerint nyágfelbontású, külön képfájlokat is kérünk.

• **Szerkezeti elvárások:** A cikk kötelező részei a Bevezetés (első fejezet) és az Összefoglalás (utolsó fejezet). A bevezetésben a szerzők röviden ismertessék a téma hátterét, a cikk fő mondanivalóját és azt, hogy a további részekben miről lesz szó. A cikkhez csatolni kell egy rövid, néhány mondatos tartalmi összefoglalót magyar és angol nyelven, továbbá meg kell adni néhány jellemző kulcsszót is, szintén magyarul és angolul. A cikk

végén kérjük a kapcsolatos, vagy előzményként felhasznált publikációkat megadni. A hivatkozásokat szögletes zárójelben számozzuk, amely után következik a szerző, majd a cikk vagy a könyv címe, a megjelenés helye és időpontja.

• **Lektorálás:** A cikkek különböző minősítési folyamatoknál értékes pontokat jelenthetnek. Az új eredményeket tartalmazó cikkeket a szerkesztőség bíraltatja. A bírálók véleménye alapján a cikket visszaadhatjuk a szerzőnek javításra, esetleg átdolgozásra. Minden félév végén az azt megelőző öt számból kiválogatjuk azokat a cikkeket, melyek a külföldi, nem magyar anyanyelvű olvasók számára is érdekesek lehetnek. Ezeket angolra fordítva az 1. és 7. számban „Selected Papers” címen jelentetjük meg, ami idegen nyelvű publikációnak számít.

• **Megjelenés:** A folyóirat minden hónap végén jelenik meg. A pontos időpont függ az ünnepektől és a hétvégek helyzetétől. Mindig az előző hónap utolsó napjáig végleges változatban beérkezett cikkeket vesszük számításba. Tematikus megfontolásokból előfordulhat, hogy egy későbbi számban előnyösebbnek látszik az adott téma tárgyalása. Általában a beküldést követő negyedévben helyet kap a munka az újságban. Kérés esetén az átnézés vagy lektorálás után a beküldéstől számított két héten belül a szerző visszaigazolást kaphat a cikk elfogadásáról.

• **Szerzői adatok:** Annak érdekében, hogy az olvasók problémáikkal, véleményükkel közvetlenül kapcsolatba léphessenek a szerzőkkel, a cikk előtt lévő szürke részben (a cím alatt) szerepel nevük, munkahelyük és e-mail címük. Célszerű tehát, hogy a cikkeket úgy küldjék be, hogy a felsorolt adatokat, valamint a szerzők telefonos elérhetőségét tartalmazzák. Ez utóbbi a szerkesztés, illetve a lektorálás közbeni esetleges kérdések tisztázásához elengedhetetlen.

• A beküldés módja:

A cikkek eljuttathatók a főszerkesztőhöz:

Szabó Csaba Attila

(BME, Híradástechnikai Tanszék, szabo@hit.bme.hu)

vagy a HTE titkárságára:

Hírközlési és Informatikai Tudományos Egyesület

(Bp., Kossuth L. tér 6-8. IV.emelet, info@hte.hu).

Reméljük, hogy ezen ismeretek segítik kollégáinkat, hogy gondolataikat, új eredményeiket, műszaki megoldásaikat, számítási módszereiket közkinccsá tegyék. Várjuk tehát a cikkeket az oktatási intézményekből, fejlesztőhelyekről, gyártóktól, üzemeltetőktől, tanulóktól, szakértőktől, oktatóktól és mindenkitől, akinek mondanivalója van a közösség számára.

A Szerkesztőbizottság

Hívásátadási eljárás 3GPP-LTE újgenerációs hálózatokban

RÁCZ ANDRÁS

Ericsson Kft. Kutatás-Fejlesztési Igazgatóság, andras.racz@ericsson.com

REIDER NORBERT, TEMESVÁRY ANDRÁS

Budapest Műszaki és Gazdaságtudományi Egyetem
andras.temesvary@gmail.com, rn488@hszk.bme.hu

Lektorált

Kulcsszavak: 3GPP Long Term Evolution, hívásátadás, teljesítmény analízis, csomagtovábbítás, csomagfelicserélődés

A 3GPP (3rd Generation Partnership Project) szabványosítási fórum már dolgozik a 3G Long Term Evolution (LTE) hálózatok szabványosításán, azzal a céllal, hogy a szabvány 2007 év végére elkészüljön. A cikkben bemutatjuk az LTE RAN architektúra legfontosabb jellemzőit és részletesen elemezzük szimulációk segítségével az LTE hívásátadási eljárást. Megvizsgálunk olyan szempontokat mint például a csomagtovábbítás szükségessége a helyes TCP működés érdekében, vizsgáljuk a csomagfelicserélődési problémát illetve a HARQ/ARQ állapotinformációk elvesztéséből adódó hatásokat. Megmutatjuk, hogy a hard-handover típusú hívásátadás nem befolyásolja kedvezőtlenül a felhasználók által észlelt szolgáltatás minőséget és a rendszer teljesítményét.

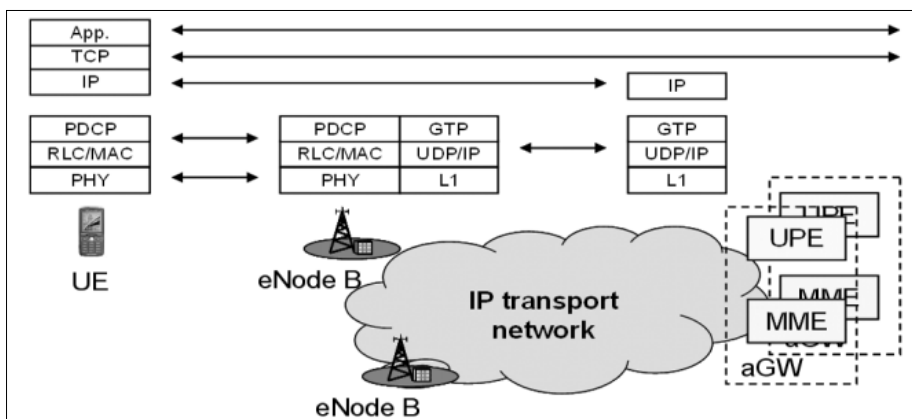
A 3G hálózatok rendszer-evolúciója érinti a rádiós hozzáférési hálózat (RAN) architektúráját [1], beleértve egy új, OFDM modulációra épülő rádiós interfészt, illetve a „core” hálózat (CN) architektúrájának megváltoztatását is [2]. A rendszer RAN oldali evolúcióját Long Term Evolution (LTE) néven hivatkozzák a szabványosításban, míg a „core” hálózati részt System Architecture Evolution (SAE) néven emlegetik. A CN oldali hálózati változtatásokat elsősorban azon tényezők motiválják, miszerint a szolgáltatások egyre inkább az IP technológiára épülnek, és emellett fontos tényező a multi-access technológiák integrálása, illetve a fix és mobil hálózatok konvergenciája. A RAN oldali változtatások elsősorban az új, OFDM alapú rádiós interfész következményei, melyeket a nagyobb sáv szélesség, a kisebb késleltetésértékek és a minél inkább a csomagkapcsolt forgalomhoz való alkalmazkodás igényei vezérelnék.

Az LTE RAN architektúra legfőbb jellemzője az elosztottság, vagyis a RAN funkciók jelentős része a bázisállomásokban elosztott módon kerülnek megvalósításra, melynek következménye, hogy a rádiós vezérlő funkciók, mint például a hívásengedélyezés, hívásátadás-vezérlés szintén elosztott módon működnek. Mindez jelentősen különbözik a 3G RAN hálózatok filozófiájától, ahol az RNC (Radio Network Controller), mint központi vezérlő végezte ezeket a feladatokat. A hívásátadási eljárást illetően fontos különbség a 3G WCDMA rendszerekhez képest, hogy LTE-ben nem használjuk az úgynevezett soft-handover eljárást, minden hívásátadás hard-handover típusú.

1. LTE hálózati architektúra

Mielőtt a hívásátadási folyamat részleteit ismertetnénk, röviden áttekintést adunk az LTE hálózatok architektúrájáról és a használt protokollokról. A hálózati architektúra és az adatkapcsolati sík protokollverme az 1. ábrán látható. A hálózatban alapvetően kétféle csomóponti eszköz különböztethető meg: a bázisállomás, melyet eNodeB-nek neveznek és az átjáró, más néven access-Gateway (aGW) csomópont. Az átjáró egység felbontható egy adatkapcsolati (user plane) és egy vezérlő síkbeli (control plane) eszközre, melyet UPE-nek (User Plane Entity) és MME-nek (Mobility Management Entity) hívunk.

A szabvány definiálja az MME és UPE közötti interfészt, ezáltal megadva a gyártóknak azt a lehetőséget, hogy külön fizikai eszközként valósítsák meg az adatkapcsolati, illetve vezérlő funkciókat. A szabvány támogatja továbbá az MME/UPE eszközök úgynevezett pool-okba szervezését, mely lehetővé teszi, hogy egy adott adminisztratív területhez tartozó bázisállomásokat bármely, ugyanazon pool-ba tartozó MME/UPE kiszolgáljon. Mindez azt jelenti, hogy egy adott adminisztratív régióba tartozó bázisállomások és az ugyanazon pool-



1. ábra
LTE hálózati architektúra
az adatkapcsolati síkbeli
protokollveremmel

ba tartozó MME/UPE eszközök között több-több típusú kapcsolat létezik, ami a megbízhatóságot növeli, vagyis egy MME/UPE eszköz meghibásodása esetén nem válik hozzáférhetetlenné a hálózat egy teljes régióban.

Ahogy az a protokollverem-ábrából is látszik, az UPE eszköz alapvetően az IP csomagok megfelelő továbbítását végzi, nagyon hasonlóan egy IP útvonalválasztóhoz (router). Az UPE és a megfelelő eNodeB között egy-egy GTP tunnel (GPRS Tunneling Protocol) [3] van kiépítve minden egyes QoS folyamra és terminálra nézve. Az UPE felelős a lefele irányba érkező csomagok megfelelő tunnelre való leképzéséért és a tunnel átkapcsolásáért egyik bázisállomástól a másikra hívásátadások esetén.

Az UPE eszköz főbb funkciói

- Csomagszűrés és leképzés tunnelekre, QoS folyamok (QoS bearer) alapján.
- Tunnel kapcsolás a terminálok bázisállomás váltásainak alapján.

Az MME eszköz valósítja meg a vezérlő protokollokat a terminál, a bázisállomás és az UPE eszköz felé és egyfajta koordinátor, vezérlő szerepet tölt be a hálózatban. Ide tartozik például a terminál azonosítása (authenticáció), a bázisállomás, illetve az UPE eszköz felkonfigurálása (QoS folyamok, csomagszűrők stb.) a felhasználó számára engedélyezett szolgáltatásminőségi osztályok alapján stb.

Az MME eszköz főbb funkciói

- Vezérlő protokollok megvalósítása a terminál (pl. authenticáció), a bázisállomás (pl. terminál QoS folyamok felépítésének vezérlése) és az UPE irányába (pl. csomagszűrők beállítása).
- Az IDLE módban lévő terminálok kezelése, elhelyezkedési információjuk tárolása (paging területekre lebontva), a terminálhoz tartozó kontextus tárolása, ideértve például a terminál QoS folyamainak konfigurációját, a titkosítási kulcsokat stb.

Mindezen MME illetve UPE funkciókból az is látszik, hogy ezek az eszközök nem végeznek semmilyen rádiós jellegű feladatot, a rádiós interfészhez kötődő funkciók teljes egészében a bázisállomásban végződnek. Az eNodeB valósítja meg természetesen a rádiós fizikai réteg feladatait, mint amilyen a kódolás, a moduláció. Az LTE rádiós interfész az OFDM modulációs technikát használja [4].

A rádiós erőforrásokat úgynevezett „chunk”-ok formájában osztja meg az ütemezési algoritmus a felhasználók között. A „chunk” egy blokk az idő-frekvencia síkban, ahol a frekvencia az OFDM vivőket jelenti. Az eNodeB minden egyes TTI (Transmission Time Interval = 1 ms; rádió-keretek ideje) időközönként ütemezési döntést hoz és kijelöli, mely felhasználók mely chunk-okat használhatják a következő időintervallumra.

Az idő-frekvencia chunk-ok használatát szemlélteti a 2. ábra, különböző tónusokkal jelölve a különböző felhasználókhoz rendelt blokkokat egy adott kiosztás esetén.

Az RLC/MAC protokoll valósítja meg az olyan funkciókat, mint például a csomagok feldarabolása az éppen aktuális rádió-keret méretének megfelelően, a hibamentes átvitelhez szükséges újraküldési mechanizmus megvalósítása (ARQ=Automatic Repeat Request). Hasonlóan a WCDMA HSDPA [5] esetén alkalmazottakhoz, az újraküldés első szintjén a Hybrid ARQ (HARQ) megoldást használják, ami N darab (N=6) egymással párhuzamosan futó 1 bites STOP-and-WAIT ARQ folyamból áll. Továbbá a HARQ módszer lehetőséget ad az inkrementális újraküldésre is, ami alapvetően arra utal, hogy a hibásan megkapott csomagot és az újraküldött csomagot együttesen használja a vevő ahhoz, hogy az eredeti adategységet dekódolja, ezáltal hasznosítva a hibásan vett adategység információtartalmát is. Azonban a HARQ újraküldés nem garantál tökéletes megbízhatóságot (a megmaradó csomagvesztés körülbelül 10^{-3}), ezért alkalmaznak fölötte még egy második szintű, hagyományos ablakozó eljárást használó ARQ protokollt is.

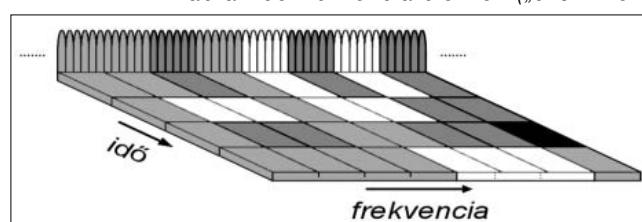
A PDCP protokoll felel az IP-fejléc tömörítésért és az adatcsomagok titkosításáért.

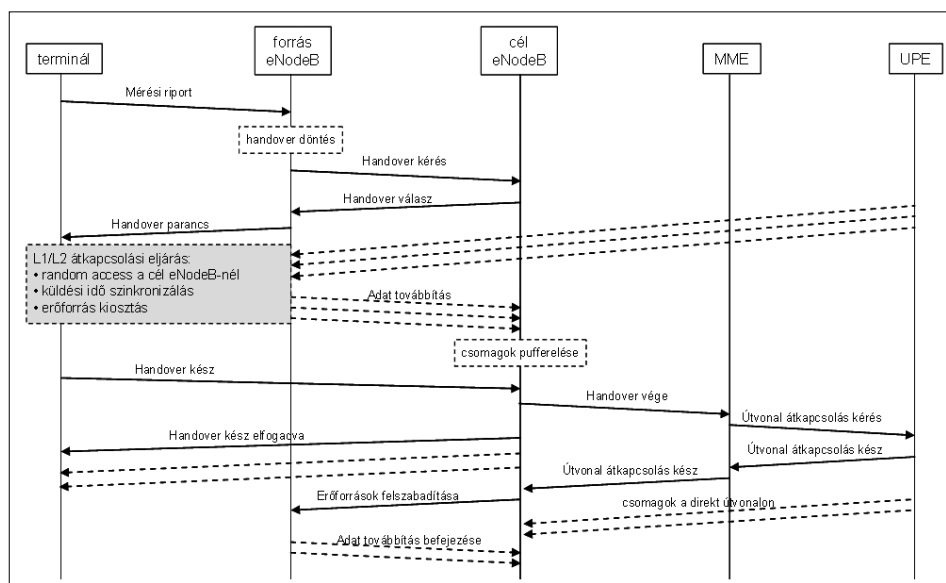
2. LTE hívásátadási eljárás

A rendszer általános ismertetését követően bemutatjuk részletesebben a hívásátadási eljárás folyamatát. A hívásátadás során a felhasználó hálózathoz való csatlakozási pontja átkerül az egyik bázisállomásról a másikra. Ez azt is jelenti, hogy azon protokollok végpontjai, amelyek a bázisállomásban végződnek (RLC/MAC/PDCP), áthelyeződnek a forrás eNodeB-ről a cél eNodeB-re. Mindez felveti annak kérdését, hogy mely protokollállapotokat vigyük át és mely állapotokat töröljük hívásátadás esetén.

Például kérdésként merülhet fel, hogy az ARQ állapotinformációt (ARQ ablak állapota stb.) átvigyük-e vagy eldobjuk. Az eldobás azt eredményezheti, hogy egyszer már helyesen átküldött csomagokat dobunk el és küldünk majd újra a cél-bázisállomástól. Ugyanígy kérdéses lehet az is, hogy az IP-fejlécet tömörítő algoritmus állapotát átvigyük-e vagy inkább inicializáljuk újra az algoritmust. Ez utóbbi eset azt jelentheti, hogy a hívásátadás után néhány IP-csomagot tömörítetlenül kell átvinni. Mindezekből látszik, hogy a várható nyereség és a többlet-komplexitás körütekintő mérlegelésére van szükség annak eldöntésére, hogy mely állapotinformációkat érdemes átvinni az eNodeB-k között egy hívásátadás során.

2. ábra Idő-frekvencia blokkok („chunk”-ok)





3. ábra
Hívásátadási folyamat

a cél-bázisállomástól kapott. Ezzel egyidőben felfüggeszti az adást/vételt és elkezdi a pufferében maradt csomagok, illetve a későbbiekben még beérkező csomagok továbbítását a cél bázisállomás felé. Minderre azért van szükség, hogy veszteségmentes legyen a hívásátadás. Később látjuk majd, hogy ennek nagy jelentősége van a TCP hatékony működésére.

Ezek után a terminál elkezdi a rádiós csatorna átkapcsolását, amely a következő főbb lépésekből áll: idő- és frekvencia-szinkronizáció az új bázisállomás adásához, majd egy random-access hozzáférés végrehajtása a cél-bázisállomás RACH (Random Access Channel) csatornáján. A random access küldés alapján a bázisállomás megbecsüli a szükséges Timing Advance (TA) értéket, amely megadja, hogy mennyivel korábban/később kell a terminál a felfele irányú adását megkezdje, annak érdekében, hogy az szinkron módon érkezzon a bázisállomáshoz. Ezáltal kompenzálható a jelterjedési késleltetés és annak különbözősége a különböző terminálok között. Egy adott rádiós keretidőben több terminál is adhat ugyanazon cellában, más-más OFDM vivőket használva, ezért szükséges annak biztosítása, hogy az egyes terminálok jelei szinkronban érkezzenek meg, máskülönben az OFDM moduláció ortogonalitása vesz el.

Ezután az ütemező dedikált erőforrásokat (chunk-ot) oszt a terminál számára, melyen az elküldheti a „Handover kész” üzenetet. Ezek után megkezdődhet az adás-vétel az új cellában.

A cél eNodeB értesíti az MME és UPE eszközöket a hívásátadás bekövetkeztéről, melynek hatására az UPE átkapcsolja a csomagtovábbítást a forrás-tunnelről a cél eNodeB felé irányuló tunnelre. Ezután a csomagok már a direkt útvonalon érkeznek a cél-bázisállomáshoz. Ekkor azonban még lehetnek továbbított csomagok is útközben a hálózatban (UPE–forrás eNodeB–cél eNodeB útvonalon), ami csomagcserélődéshez vezethet, mely problémát a későbbiekben még részletesebben vizsgáljuk.

A 3. ábra szemlélteti a hívásátadási folyamatot az üzenetváltási diagram segítségével. A terminál folyamatosan méri a szomszédos bázisállomások „pilot” csatornájának jelerősségét és mérési riportokat küld a saját bázisállomásának. A méréseket és a riport küldésének feltételeit a bázisállomás határozza meg. Lehetőség van például periodikus illetve valamely esemény bekövetkeztéhez kötött riportküldési szabályok bekonfigurálására a terminálban. A mérési riport alapján a forrás eNodeB dönt a hívásátadás szükségességéről és kezdeményezi annak előkészítését a cél-bázisállomásnál. Ennek során átadja az adott terminálhoz tartozó kontextust a cél-bázisállomásnak, mely tartalmazza például a terminál QoS vivőinek leíróit, azok rádiós konfigurációit, a titkosításhoz használatos kulcsokat, az UPE felé használatos tunnel azonosítókat stb. A cél bázisállomás hívásengedélyezést hajt végre a terminál egyes QoS vivőire és lefoglalja az azokhoz szükséges rádiós erőforrásokat.

A cél eNodeB emellett lefoglal egy-egy tunnel-végpontot az ideiglenes csomagtovábbítás céljára a terminál minden egyes QoS vivője számára. Ezen információkat visszaküldi a forrás bázisállomáshoz, beleértve mindazon rádiós paramétereket is, amelyek a terminál számára szükségesek ahhoz, hogy kapcsolódhasson az új cellában (például random access csatorna pozíciója, egyedi terminál azonosító az új cellában stb.).

A cél-bázisállomás által nyugtázott hívásátadás előkészítés után a forrás-bázisállomás utasítja a terminált a hívásátadás végrehajtására és a parancsüzenetben transzparens módon átadja mindazt az információt, amit

1. táblázat
A szimuláció főbb paraméterei

Paraméter	Érték
Mobil terminál sebessége	10 m/s
Cellasugár	200 m
Handover hiszterézis	0.5 dB
Adatforrás	File letöltés
TCP maximális torlódási ablakmérete	128 Kbyte

3. A hívásátadási eljárás analízise szimuláció segítségével

Ebben a fejezetben szimulációs eredmények segítségével elemezzük a fentebb ismertetett hívásátadási eljárást, összpontosítva a felhasználó által kapott teljesítményre. A szimulátor részletesen modellezi a rádiós csatornát (például fading, interferencia stb.) és megvalósítja a magasabb szintű protokollrétegeket is (például RLC/MAC, TCP). A szimulált hálózat főbb paramétereit az 1. táblázat tartalmazza.

3.1. A csomagtovábbítás hatása a TCP átvitelre

Az LTE hívásátadás szimulációs vizsgálata során elsőként a csomagtovábbítás kérdésével foglalkoztunk. Mint már korábban utaltunk rá, csomagtovábbításról akkor beszélünk, ha a hívásátadási folyamat során a mobil terminál mind a pufferében lévő, mind a még hozzá

érkező csomagokat a transzportálózaton a cél-bázis-állomás felé továbbítja. Két különböző esetet vizsgáltunk, melyek rámutattak a csomagtovábbítás fontosságára.

Az első esetben a rádiós link sebességét 2 Mbit/s-ra, míg a második esetben 20 Mbit/s-ra korlátoztuk. A TCP torlódási ablak mérete az első esetben 64 kbyte volt, mely érték teljes link-kihasználtságot biztosított. A körülfordulási idő körülbelül 80 ms volt, mely magában foglalta az Internet-késleltetését is.

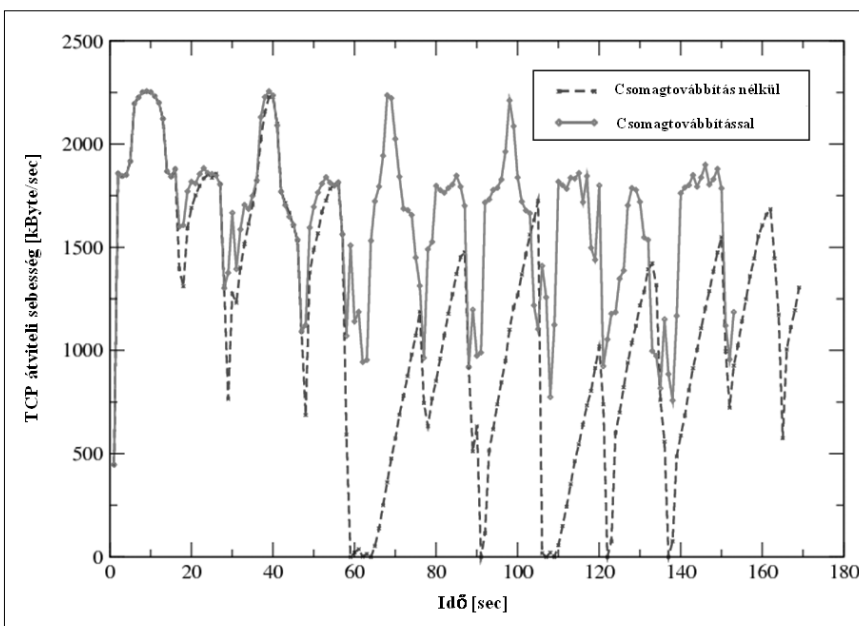
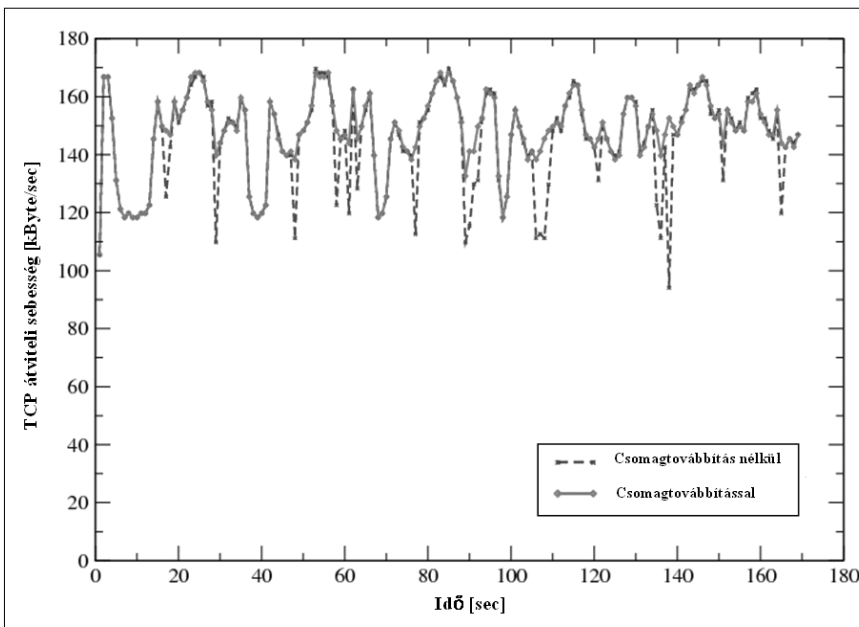
A 4. ábrán, a 2 Mbit/s-ra korlátozott rádiólink esetében kapott TCP átviteli sebességet figyelhetjük meg az idő függvényében, csomagtovábbítás nélküli és csomagtovábbítást alkalmazó esetekben, ahol az egyes mérési eredmények az egy másodpercre vetített TCP átviteli értékeket ábrázolják. Ebben az esetben a két átviteli görbe közti eltérés ugyan érzékelhető, de nem túl jelentős, csupán hívásátadások esetén figyelhető meg kisebb előny a csomagtovábbítást alkalmazó működési módban.

Más a helyzet azonban, ha a 20 Mbit/s sebességre korlátozott rádiós linket használó hálózatot tekintjük. Ebben az esetben a TCP torlódási ablak méretét 256 kbyte-ra növeltük, hogy a megnövekedett sávszélességű linket ismét maximálisan kihasználjuk. Az 5. ábrát megfigyelve jelentős különbséget láthatunk a TCP átviteli görbékben, a csomagtovábbítást alkalmazó eset előnyére. Ez a meghatározó különbség abból adódik, hogy a csomagtovábbítás nélküli esetben bekövetkezett nagy mennyiségű csomagvesztés hatására TCP timeout történik. Megjegyezzük, hogy a TCP átviteli értékek csomagtovábbítást használó esetben is megfigyelhető jelentős ingadozása a rádiós link minőségének változásának köszönhető.

Megállapítható tehát, hogy az LTE hard-handover típusú hívásátadása során érdemes csomagtovábbítást használni, hiszen ekkor nagy sebességű linkek esetén a sorozatos csomagvesztéseket elkerülve a TCP timeout mechanizmus is elkerülhető, így maximális link kihasználtság érhető el a teljes adatátviteli folyamat során.

4. ábra
TCP átvitelek 2 Mbit/s-os link esetén

5. ábra
TCP átvitelek 20 Mbit/s-os link esetén



3.2. A csomagsorrend felcserélődése

Mint bemutattuk, nagy sebességek esetén a csomagok átirányítása a hívásátadás során szükséges ahhoz, hogy az adatátviteli sebesség ne essen vissza jelentős mértékben.

Sajnálatos módon azonban a csomagtovábbításnak nem kívánt mellékhatásai is lehetnek, amiket kezelni kell. A probléma abból adódik, hogy amikor az aGW átkapcsolja a felhasználóhoz tartozó adatfolyamot a forrás-bázisállomás felé vezető tunnel helyett a cél-bázisállomás felé vezető tunnelre, akkor a hálózaton még lehetnek átirányított csomagok, melyek felcserélődhetnek az immár direkt útvonalon haladó csomagokkal. Ez a jelenség a TCP működésére is hatással lehet, tekintve, hogy már kettőnél több egymás után rossz sorrendben érkező csomag a TCP torlódási ablak feleződését okozza.

A csomagfelcserélődési probléma vizsgálata során fontos kérdés volt az utolsó, átirányított útvonalon közlekedő csomag, illetve az első új útvonalon közlekedő csomag megérkezési ideje a cél eNodeB-hez. Amennyiben az új útvonalon érkező első csomag korábban érkezik meg, mint az utolsó átirányított csomag, úgy a csomagok sorrendje felborul. Az is látható, hogy minél nagyobb a két időpont közti különbség, annál több csomagot érint a probléma, és annál kritikusabb a jelenség a TCP átviteli folyamatra.

A probléma megoldására egy, a cél eNodeB-kben használt egyszerű prioritásos ütemező alkalmazását javasoltuk és vizsgáltuk meg alkalmazását szimuláció segítségével. Az egyszerű prioritásos ütemező két pufferrel dolgozik, egy magasabb prioritású pufferbe az átirányított csomagok, míg egy kisebb prioritásúba a direkt úton érkező csomagok kerülnek. Amennyiben a hívásátadási folyamat során az átirányított csomagokat tartalmazó puffer nem ürül ki addig, míg az utolsó átirányított csomag is megérkezik, úgy a felcserélődési probléma megoldódik, hiszen az ütemező addig csak átirányít-

tott csomagokat fog a MAC rétegnek szolgáltatni, és a felborult sorrend ismét helyreáll. Elmondható, hogy az esetek többségében, amikor a csomagok áramlása folyamatos, ez a helyzet áll fenn és ekkor a prioritásos ütemezés sikeresen megoldja a problémát.

A szimulációs vizsgálatok során a 6. ábrán látható TCP átviteli görbét kaptuk, melyet a késleltetési paraméterek egy olyan beállítása okozott, ahol a transzport hálózati linkek késleltetése viszonylag magas volt (20 ms), ami lehet például hálózati torlódás következménye is. Látható, hogy az adatátviteli folyamat során lezajlott hívásátadások jelentősen visszavetik a TCP átvitelt a csomagok felcserélődését nem kezelő esetben.

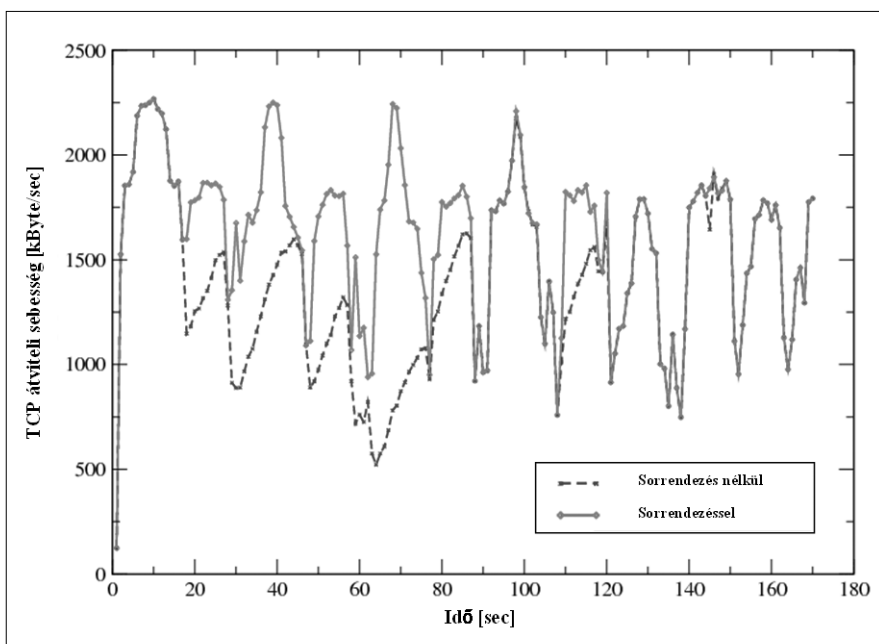
Részletesebben megvizsgálva a jelenséget, a 7. ábrán láthatjuk, hogy mi is történik valójában a TCP szegmensek sorrendjével. Látható, hogy négy egymás után rossz sorrendben érkező TCP szegmens újraküldést okoz az adóoldalon és a torlódási ablak felezését okozza. Amint a 8. ábrán látható, az általunk javasolt módszer megoldotta a problémát, a csomagok sorrendje teljesen helyreállt.

Következtetésképpen elmondható, hogy érdemes felkészíteni az eNodeB-eket a csomagok átirányítására, ugyanakkor a csomagok keveredésének reális veszélye miatt gondoskodni kell a csomagok újra sorrendezéséről is.

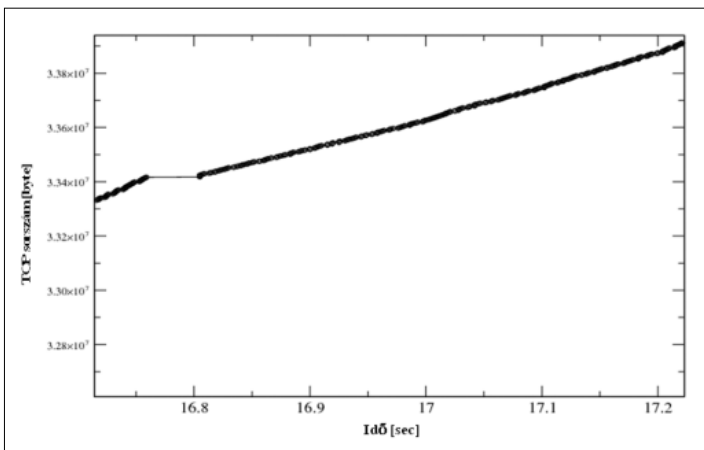
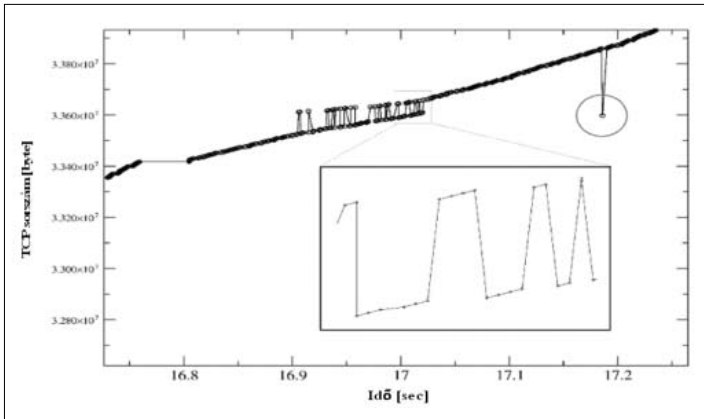
3.3. A HARQ/ARQ állapotinformáció megtartása

Ahogy már korábban említettük, hívásátadásakor a forrás bázisállomásban végződő protokollok végpontjait át kell helyezni a cél bázisállomáshoz. Ennek következtében új MAC kapcsolat jön létre a cél eNodeB és a mobil terminál között, vagyis a HARQ/ARQ állapot nem kerül átvitelre. Ehelyett mindazon MAC SDU-k továbbításra kerülnek a cél eNodeB-hez, amelyekre nem érkezett pozitív nyugta a Handover parancs kiadásáig. Mivel a MAC réteg szükség szerint szegmentálást is végez, vagyis egy SDU-t MAC PDU-kra darabol fel, ezért

amennyiben egy SDU-ból akárcsak egy darab PDU is hiányzik a vevőnél, a teljes SDU újraküldésre fog kerülni a cél-bázisállomástól. Mindez bizonyos pazarlást jelent a rádiós interfészen. Szimuláció segítségével megvizsgáltuk, hogy a HARQ/ARQ állapot elvetése miatt átlagosan hány sikeresen megérkezett MAC PDU-t dobunk el a hívásátadás következtében.



6. ábra
TCP átvitel
sorrendezéssel és nélküle



7. és 8. ábra

TCP sorszámok sorrendezés nélkül és sorrendezéssel

A szimulációk során statisztikát készítettünk arról, hány helyesen vett MAC PDU van az adási ablakban a hívásátadás megkezdésének pillanatában. A statisztikákból közelítőleg megállapítható, hogy egy hívásátadás bekövetkeztekor hány sikeresen átküldött PDU-t kell eldobnia a vevőnek. A 9. ábra tartalmazza a kapott eredményeket 1, illetve 10 Mbit/s-os rádiós link kapacitás esetén.

Az ábrán látható, hogy az esetek 50-70 százalékában egyetlen helyesen vett MAC PDU sincs az adási ablakban, ami azt jelenti, hogy ekkor a PDU-k sorrendhelyesen érkeztek meg a vevőhöz. Az esetek többi részében gyakorlatilag egy-két PDU található csak az adási ablakban, tehát csak egy-két olyan MAC PDU van, amelyet helyesen detektált a vevő, de még van előtte nyugtázatlan MAC PDU. Ebből azt a következtetést lehet levonni, hogy a HARQ/ARQ állapotinformáció elhagyásából adódó rádiós erőforrás-pazarlás elhanyagolható. Ez alapján az is elmondható, hogy annak a valószínűsége, hogy egy sikeresen átküldött teljes SDU az adási ablakban van, meglehetősen kicsi. A szimulációk eredményei szintén ezt támasztják alá, mivel egyszer sem fordult elő, hogy hívásátadáskor egy teljes SDU előbb került átküldésre, mint egy sorrendben előtte álló.

4. Összefoglalás

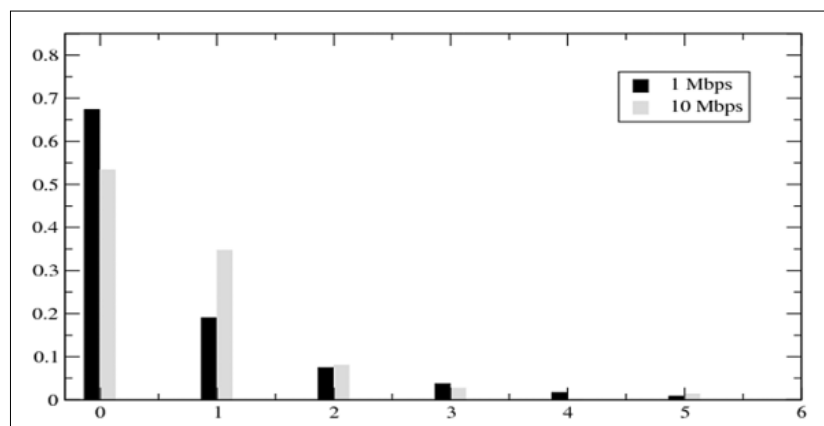
Ebben a cikkben a 3GPP-LTE hálózatban használt hívásátadási eljárást mutattuk be és elemeztük teljesítőképességét. Megmutattuk, hogy csomagtovábbítás szükséges a nagysebességű adatátvitel minőségének megtartásához, rávilágítottunk az ebből eredő csomagsorrend felcserélődési problémára, melyre egy megoldást is javasoltunk.

Következésképpen elmondható, hogy az általunk vizsgált hard-handover típusú hívásátadás nem befolyásolja kedvezőtlenül a felhasználó által észlelt szolgáltatás minőségét és a rendszer teljesítményét.

Irodalom

- [1] 3GPP TS 36.300, „Evolved UTRA and evolved UTRAN, overall description”, ftp.3gpp.org/Specs/archive/36_series/36.300
- [2] 3GPP TS 23.401, „GPRS enhancements for LTE access”, ftp.3gpp.org/Specs/archive/23_series/23.401
- [3] 3GPP TS 29.060, „GPRS tunneling protocol across the Gn and Gp interface (Rel 7)”, ftp.3gpp.org/Specsarchive/29_series/26.060
- [4] R. van Nee, R. Prasad, „OFDM for Multimedia Wireless Communications”, Artech House, 2000.
- [5] S. Parkvall, et al., „Evolving 3G mobile systems – broadband and broadcast services in WCDMA”, IEEE Communications Magazine, February, 2006.

9. ábra
Egy bázisállomás HARQ/ARQ adási ablakában lévő helyesen vett PDU-k számának hisztogramja hívásátadáskor



TFRC-alapú szelektív újraküldő algoritmus

HUSZÁK ÁRPÁD, IMRE SÁNDOR

Budapesti Műszaki és Gazdaságtudományi Egyetem, Híradástechnikai Tanszék
Mobil Távközlési és Informatikai Laboratórium (MC2L)
{huszak, imre}@hit.bme.hu

Kulcsszavak: multimédia, adaptív újraküldés, video streaming

A cikkben egy olyan általunk kidolgozott eljárást mutatunk be, amely az MPEG típusú videó folyam sérült, illetve elveszett csomagjait szelektíven újraküldi a hálózat pillanatnyi állapotától függően. Ezáltal a hálózat szabad kapacitását kihasználva tudunk a multimédiás tartalom minőségén javítani. A szabad kapacitást a forrás sebessége és a DCCP torlódáskezelő algoritmus (TFRC – TCP Friendly Rate Control) által szolgáltatott paraméterek alapján tudjuk meghatározni.

Az IP alapú hálózatokban egyre nagyobb mértékben megjelenő korszerű multimédiás alkalmazások, sok új és izgalmas problémát vetnek fel. Ezek az alkalmazások többnyire az UDP-t használják szállítási rétegbeli protokollként, de újabb szabványok is megjelentek már, amelyek hatékonyabbnak bizonyulnak audio/videó folyamatok átvitele esetén (például UDPLite, SCTP, DCCP stb.). A DCCP nagyon előnyös tulajdonsága, hogy torlódáskezelő algoritmust (TFRC, TCPLike) is használ a hálózat túlterhelésének elkerülése érdekében.

A multimédiás folyamatban nem minden csomag egyformán értékes, hiszen az MPEG szabvány három különböző kerettípust használ (I-, P- és B-keretet). Az I-kép egy teljes video-kép információját tartalmazza, míg a P- és B-keretek csupán a különbségeket tárolják, ezért az I-keret hibája hatással van a többi keretre is. Az általunk kidolgozott eljárás tehát a hálózat szabad kapacitását az elveszett csomagok prioritása alapján tölti fel.

A torlódást elkerülő és fontossági alapon szelektíven újraküldő eljárást, Ns2 szimulációs környezetben vizsgáltuk. Az Ns2 hálózati szimulátor kimenetét feldolgozva vizsgáltuk meg az algoritmus hatékonyságát. Az eredmények alátámasztották várakozásainkat.

1. Bevezetés

Az új vezeték nélküli technológiák ma már lehetővé teszik a multimédiás alkalmazások használatát, hiszen ezek már rendelkeznek azokkal a jellemzőkkel, melyek lehetővé teszik az ilyen típusú alkalmazások kiszolgálását. Ez az oka annak, hogy ezek az audio/ videó-szolgáltatások robbanásszerű fejlődésen mennek át és sok új és izgalmas problémát vetnek fel. Az újonnan megjelenő hozzáférési hálózatok már képesek kielégíteni az új igényeket.

Ilyen technológia például a nemrégiben megjelent WiMAX (IEEE 802.16), amely akár ötven kilométer sugarú körben, 70 Mbit/s sebességen kínál vezeték nélküli internetes hozzáférést, vagy az UMTS. Ennek okán min-

denképpen nagymértékű növekedésre kell számítani az IP-alapú audio/videó-alkalmazások terén.

Hasonló fejlődési tendencia mutatható ki a végberendezések fejlődésében is, hiszen ma már nem jelent technológiai problémát kép- és hanganyag lejátszása a mobil terminálokban. Ugyanakkor ezek az eszközök már nagymértékben IP-alapú kommunikáción alapulnak. Jelenleg előrehaladott fejlesztések folynak adatszolgáltatások biztosítására vezeték nélküli hálózatokban, a végső cél: a mobil kommunikáció és az Internet teljes konvergenciája, a mobil Internet kifejlesztésével. 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 egy garantált minőségű adatfolyam átvitele, vagyis QoS (Quality of Service) biztosítása. Ez az adatfolyam lehet szigorúan érzékeny a hibákra, mint például az FTP, WWW alkalmazásokhoz tartozó adatátvitel, de lehet a hibákra érzéketlenebb, mint a hang-, kép-, illetve videóátvitel esetén. Ennek oka, hogy a kisebb hibákat az emberi érzékszervek nehezebben 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.

Ezek a multimédiás alkalmazások többnyire az UDP-t [1] használják szállítási rétegbeli protokollként, de újabb szabványok is megjelentek már, amelyek hatékonyabbnak bizonyulnak audio/videó-folyamok átvitele esetén (pl. UDPLite [2], SCTP [3], DCCP [4] stb.). A DCCP nagyon előnyös tulajdonsága, hogy torlódáskezelő algoritmust (TCPLike [5], TFRC [6]) is használ a hálózat túlterhelésének elkerülése érdekében.

A cikkben egy olyan általunk kidolgozott eljárást mutatunk be, amely az MPEG típusú videófolyam sérült illetve elveszett csomagjait szelektíven újraküldi a hálózat pillanatnyi állapotától függően. Ezáltal a hálózat szabad kapacitását kihasználva tudunk a multimédiás tartalom minőségén javítani. A szabad kapacitást a forrás sebessége és a DCCP torlódáskezelő algoritmus (TFRC – TCP Friendly Rate Control) által szolgáltatott paraméterek alapján tudjuk meghatározni.

A cikk következő szakaszában megismerhetjük a jelentősebb szállítási rétegbeli protokollokat, amelyek alkalmasak lehetnek az audió/videó-folyam továbbítására. Ezt követi az általunk kidolgozott algoritmust ismertetése, mely egy új transzport protokoll (DCCP) alkalmazásával javítani képes az MPEG típusú multimédiás adatfolyam minőségén a TFRC torlódáskezelő algoritmuson alapuló szelektív újraküldési technika segítségével. A negyedik részben az algoritmus hatékonyságának vizsgálatát foglaljuk össze. Majd végül az ötödik szakaszban levonjuk a következtetéseinket és ismertetjük a továbblépési lehetőségeket.

2. A szállítási réteg protokolljai multimédiás alkalmazásokhoz

Az ISO/OSI rétegstruktúrában a hálózati réteg (network layer) felett elhelyezkedő szállítási réteg feladata a sorrendhelyes, duplikációmentes és megbízható átvitel biztosítása a felhasználói alkalmazástól függően. Számos transzport-protokoll áll rendelkezésünkre, melyek közül az alkalmazás típusától függően választjuk ki a legalkalmasabbat. A szállítási réteg protokolljait alapvetően két osztályba sorolhatjuk: megbízható és megbízhatatlan.

A megbízható transzport protokoll legfőbb jellemzője, hogy minden csomagot eljuttat a vevőhöz. Ezt természetesen csak a megsérült vagy elveszett csomagok újraküldésével lehetséges. Az újraküldésnek hátrányai is vannak, mégpedig az, hogy a sérült illetve hiányzó csomagok újraküldését és a vevőhöz való megérkezését meg kell várni, és a többi csomagot csak ezt követően tudjuk a hálózatra küldeni. Ez időnként jelentős késleltetés okozhat. A megbízható protokollok legismertebb képviselője a TCP (Transmission Control Protocol) [7], de a nemrégiben kifejlesztett SCTP (Stream Control Transport Protocol) is ehhez a csoporthoz tartozik.

A megbízhatatlan protokollok, mit például az UDP (User Datagram Protocol) vagy a DCCP (Datagram Congestion Control Protocol) nem küldi újra a sérült adatcsomagokat, emiatt nem fordulhat elő az sem, hogy egy újraküldött csomag megérkezésére kelljen várni. Ezeket a típusú protokollokat ezért olyan alkalmazások esetén használják, ahol a késleltetés alacsony szinten tartása a legfontosabb szempont és bizonyos számú csomag elvesztése elfogadható a felhasználó számára. Általában megbízhatatlan protokollokat használnak a multimédiás alkalmazások esetén, az ismertetett tulajdonságok alapján.

A szállítási réteg két régi nagy protokollja a TCP és az UDP. Ezek a protokollok a mai napig meghatározói a számítógépek valamint számítógép-hálózatok közötti adattovábbításnak, pedig immár 25 éves szabványok, hiszen a TCP-t és az UDP-t is az 1980-as évek elején fejlesztették ki. Ez az oka annak, hogy mindkét szabványt viszonylag kis hibavalószínűségű, vezetékes hálózatra dolgozták ki, azonban a ma egyre szélesebb körben használt vezeték nélküli hálózatok jellemzői jelentősen különböznek a vezetékes hálózatok adatátviteli tulajdonságaitól.

Az eltelt évek alatt a vezetékes számítógépes hálózatok is nagyon sokat fejlődtek. Emiatt a régi protokollokat felül kell vizsgálni, és az új igényeknek megfelelően módosítani kell azokat, illetve újabb szabványok kifejlesztésére van szükség. Így került sor a régi protokollok újabb változatainak megjelenésére, TCP esetén a TCP Reno, TCP New-Reno, TCP Vegas, TCP SACK, míg az UDP módosított változata az UDP Lite (Lightweight User Datagram Protocol). A közelmúltban több új transzport-protokollt fejlesztettek ki, melyek megpróbálják kiküszöbölni a régebbi protokollok hibáit, mint például a multimédiás átvitelre szánt megbízható SCTP és a megbízhatatlan DCCP.

A továbbiakban a DCCP működését ismertetjük, mivel a TFRC torlódáskerülő mechanizmuson alapuló szelektív újraküldés ezt a protokollt használja számos előnye miatt.

A DCCP egy megbízhatatlan transzport protokoll, amely torlódásszabályozási algoritmus használatára, valamint sorrendhelyes csomagtovábbításra is alkalmas a TCP-hez hasonlóan. UDP esetén a torlódás elleni védekezést az alkalmazásoknak kellett megoldaniuk, míg DCCP esetén ez már a protokoll szerves része. A torlódásszabályozó algoritmus feladata, hogy elkerülje a hálózat túlterhelését.

Amennyiben a hálózat útvonalválasztóinak bufferai megtelnek a még nem továbbított csomagokkal, a továbbra is érkező csomagok elvesznek. A forrásnak tehát értesülnie kell arról, hogy a hálózatban torlódás alakult ki, ezért a küldési sebességet csökkentenie kell. A DCCP-t úgy próbálták alakítani, hogy a TCP és az UDP előnyeit egy protokollként valósítsák meg. A DCCP fejlődésben így ráismerhetünk az előbbi protokollokból ismert fejlődésekre.

A fejléc hossza minimálisan 12 byte, maximálisan pedig a 1024 byte-ot 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. Ez lehetőséget ad arra, hogy azok az alkalmazások, amelyek képesek kezelni a sérült adatokat, hatékonyabban működjenek.

A kapcsolorientált DCCP a kapcsolat felépítése során megbízható protokollként működik. A torlódásszabályozással kapcsolatos üzenetek szintén megbízható adatfolyamként kerül továbbításra. A torlódásszabályo-

zási 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.

A torlódásszabályozó algoritmusokat a fejléc CCID (Congestion Control Identifier) mezőjében különböztetjük meg:

- TCP-like Congestion Control [CCID 2]
- TFRC (TCP-Friendly Rate Control) Congestion Control [CCID 3]

A TCP-Like algoritmus a TCP torlódáselkerülési technikáját alkalmazza, amire jellemző a hirtelen sebességcsökkenés csomagvesztés esetén. A TFRC az (1) képlet alapján határozza meg a küldési sebességet, aminek köszönhetően nincsenek drasztikus sebességváltozások.

$$T = \frac{s}{R \sqrt{\frac{2p}{3} + 4R \left(3 \sqrt{\frac{3p}{8}} \cdot p \cdot (1 + 32p^2) \right)}} \quad (1)$$

A képletben szereplő változók: T – küldési sebesség; s – csomagméret; R – körbefordulási idő (RTT) becsült értéke; p – csomagvesztési arány.

A körbefordulási idő becslésére a következő képlet szolgál:

$$R = qR + (1-q)RTT \quad (2)$$

A képletben RTT az aktuális körbefordulási idő, míg az R a becslés, $q=0.9$ pedig állandó. A becslés célja, hogy a pillanatnyi ingadozás ne legyen hatással küldési sebességre. Ugyanilyen módon kerül meghatározásra a csomagvesztési arány is hasonló módon számítható. A p számításához, a korábbi időintervallumokban mért csomagvesztési arány értékét súlyozva használják. Ennek az ad értelmet, hogy az egyszeri hibák, amelyek lehetnek a torlódás következményei vagy a csatorna hibái, csak kisebb mértékben befolyásolják a küldési sebességet. A becslések teszik lehetővé, hogy a küldési sebesség ne változzon drasztikusan, hiszen pont ez a TFRC módszer lényege.

Az adatfolyam ugyan megbízhatatlan, de az adóoldalt 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. 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-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.

3. TFRC alapú szelektív újraküldő algoritmus

A mobil eszközök teljesítménye ma 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. Emellett pedig a megnövekedett forgalom okozhatja a hálózat túlterhelését, ami csomagvesztéshez vezet.

Az MPEG típusú multimédiás adatfolyamok javítására már korábban kidolgoztunk egy eljárást, ami a csomagok fontosságának függvényében újraküldi a hibás vagy sérült csomagokat. Mivel az MPEG videó három különböző képtípus használ, melyeknek prioritása nem egyforma, a különböző típusú keretekhez tartozó információ elvesztése különböző hatással van a mozgókép minőségére:

I-képek: Intra frame coded – csak képkockán belül kódolt, ami azt jelenti, hogy a képkockákon belül JPEG tömörítést alkalmazva kódolja a teljes képet. Ez a kerettípus a legfontosabb, hiszen ez szolgál alapul a többi képtípushoz.

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 P-képek tömörítési aránya nagyobb, mint az I-kereté.

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.

Az általunk kidolgozott szelektív újraküldés átmene- tet képez a megbízható és a megbízhatatlan alappro- tokollok között, hiszen az elvesztett csomagok csak egy részét küldi újra. A korábbi munkákban [8] a csomag tartalma alapján döntöttünk a csomagok újraküldéséről, azonban ezt már szempontok alapján is megtehetjük. A csomag újraküldéséről tehát több feltétel alapján dönt- hetünk:

- a csomag tartalma alapján,
- a hálózati terhelés alapján,
- az újraküldött csomag várható vételi ideje alapján.

A szelektív újraküldés alapötlete az, 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. Az egyes csomagok újraküldése már a mi feladatunk. Ebből a szempontból a DCCP összesíti a többi transzport-protokoll előnyét, ezért választottuk ezt a protokollt a probléma megoldásához.

Ebben a dolgozatban a hálózati terhelés alapján határozzuk a csomagok újraküldéséről. Ehhez azonban ismernünk kell a hálózat aktuális állapotát. Tudnunk kell, mikor van lehetőség újabb csomagok beküldésére a hálózatba, anélkül hogy torlódást okoznánk, vagy a már túlterhelt hálózat terheltségét tovább fokozzuk. Annak megállapítására, hogy mikor tekinthető a hálózat túlterheltnek, vagy túlterheléshez közeli állapotban lévőnek, a DCCP TFRC (CCID3) torlódáskezelő algoritmus ad választ. Az TFRC algoritmus a hálózat jellemzőit használja fel, a torlódás elkerülése érdekében, melyeket az (1) képletben alkalmaz a küldési sebesség megállapítására. Abban az esetben, amikor a hálózat túlterheltség miatt csomagot dob el, a küldő oldalra nem érkezik nyugta az adott csomagról. Ennek hatására a p csomagvesztési arány nő, ami a küldési sebesség csökkenéséhez vezet. Hasonló hatással van a küldési sebességre a körbefordulási idő (RTT) növekedése is. Ha ez az érték növekszik, abból arra következtethetünk, hogy az elküldött csomagok egyre tovább várakoznak a routerek bufferjeiben, mire továbbküldésre kerülnek, vagyis a hálózat közeledik a túlterhelt állapothoz.

A hálózatban tehát két módon kerülhet sor csomagvesztésre vagy csomagsérülésre. Az egyik lehetőség, hogy az egyik a hálózati router bufferjének túlszordul, a másik pedig, hogy a csatorna hibázása miatt sérül a csomag. Az első esetben egy ideig nem küldhetjük újra a csomagot, hiszen az további terhelést okoz a hálózatnak, míg ha a routereknek vannak szabad kapacitásai, csupán a rossz minőségű csatorna miatt sérült a csomag, a csomagot újra lehet küldeni. A feladat tehát, hogy ezt a két eseményt szétválasszuk. Ez azonban nem egyértelmű, hiszen a vevő, illetve a küldő oldal csupán a csomagvesztés tényét és körbefordulási időt tudja észlelni.

Abban az esetben, ha ismerjük a forrásunk sebességét, ami jelen esetben egy MPEG videó folyam, el tudjuk dönteni, hogy egy adott pillanatban a TFRC által meghatározott küldési sebesség esetén, van-e még szabad kapacitás a hálózatban. Tehát abban az esetben, ha:

$$X_{TFRC}(t) > S_{MPEG}(t) \quad (3)$$

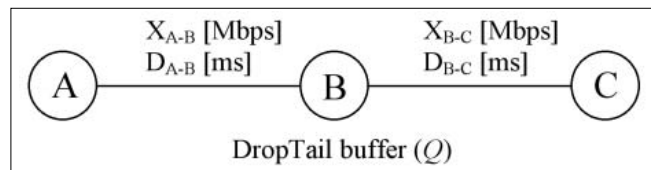
újraküldhetjük az elveszett csomagot, anélkül hogy túlterhelnénk a hálózatot. A képletben $X_{TFRC}(t)$ a TFRC torlódáselkerülő algoritmus által meghatározott pillanatnyi küldési sebességet jelenti, míg $S_{MPEG}(t)$ az MPEG folyam sávszélesség igényét jelenti.

A TFRC alapú szelektív újraküldő algoritmus olyan szituációkban bizonyul hatékonynak, amikor a TFRC küldési sebesség az MPEG folyam sebesség körül ingadozik. Ilyen torlódás közeli állapotban sikeresen meg tudjuk védeni a hálózatot az összeomlástól. A vizsgálatok során is az ilyen állapotú hálózatokat vizsgáltuk.

Abban az esetben, amikor a hálózat kapacitása sokkal nagyobb, mint a videófolyam sebessége, a TFRC algoritmus is nagy küldési sebességet engedélyez a küldő terminálnak: $X_{TFRC}(t) \gg S_{MPEG}(t)$. Ekkor az újraküldött csomag biztosan nem fog torlódást előidézni a hálózatban.

Előfordulhat, hogy a videófolyam sebessége túl nagy és emiatt a hálózat képtelen azt továbbítani, a szűkös erőforrásai miatt. Ilyenkor már maga az MPEG folyam is torlódást okoz, tehát nincs lehetőség a minőség javítására az újraküldés alkalmazásával. A Wakamiya és társai által kidolgozott [9] TFRC alapú MPEG kódolás megoldást nyújt arra, hogy mindig a felhasználható szabad hálózati kapacitáshoz alkalmazkodva kódoljuk a videó anyagot. Ez a megoldás azt eredményezi, hogy a TFRC által meghatározott küldési sebesség és az MPEG folyam sebessége közel azonos. Ilyen helyzetekben pedig jól alkalmazható az általunk kidolgozott újraküldési mechanizmus.

A TFRC alapú szelektív újraküldést Ns2 szimulációs környezetben vizsgáltuk, valamint a szimulátor kimeneti adatai alapján kiegészítő méréseket végeztünk. A hálózatot egyszerűsített formában implementáltuk, ami azt jelenti, hogy három csomópontot helyeztünk el a következő módon:



1. ábra Topológia

Az 1. ábrán az A csomópont jelöli az MPEG streaming szervert, amely a B routeren keresztül továbbítja a csomagokat a C felhasználóhoz. A szimulációban a B router reprezentálja a többi elemét is, így amikor a B router Q elemű bufferje túlszordul, a hálózat túlterheltté válik. A B router DropTail típusú buffert alkalmaz, melynek tulajdonsága, hogy azok a csomagok melyek már nem férnek el a tárolóban, elvesznek. A mérésekben $Q=10$. Az A-B link sávszélessége (X_{A-B}) és késleltetés (D_{A-B}) a mérés során nem változik, míg a B-C link sávszélessége (X_{B-C}) a hálózat terhelését szimulálva ingadozik.

Az MPEG folyam sebessége is állandó és mindegyik mérésben kisebb, mint az A-B link kapacitása. Ezen a szakaszon tehát torlódás nem történhet. Azokban a mérési elrendezésekben, ahol a csatorna veszteségét is vizsgálni akarjuk, ezen a linken fogunk hibákat generálni.

4. Szimulációs eredmények

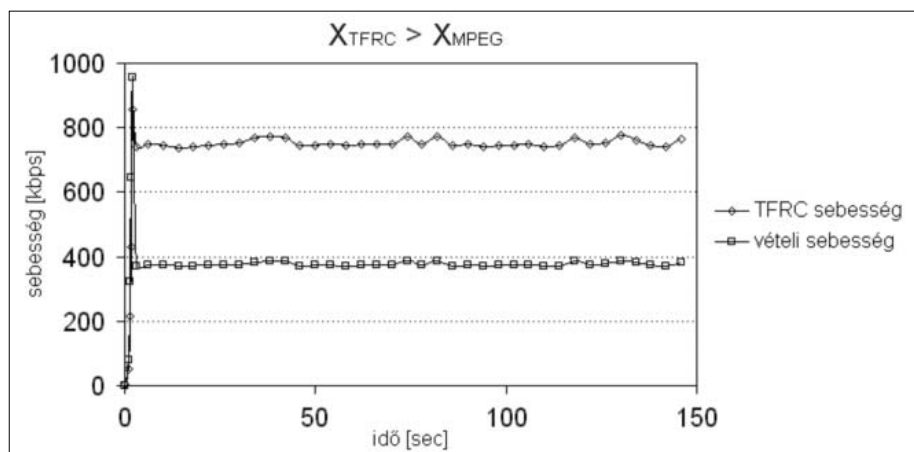
Az újraküldött csomagok, amennyiben időben érkeznek a vevő oldalra, hatékonyan javítják a videó folyam minőségét. A cél pedig, hogy feleslegesen ne küldjük olyan újraküldendő csomagokat a hálózatba, amelyek nagy valószínűséggel nem érkeznek meg, vagy más csoma-

gok elvesztéséhez vezetnek. A hálózat terheltségét jól mutatja az 1. ábrán B -vel jelölt router bufferjének aktuális szintje.

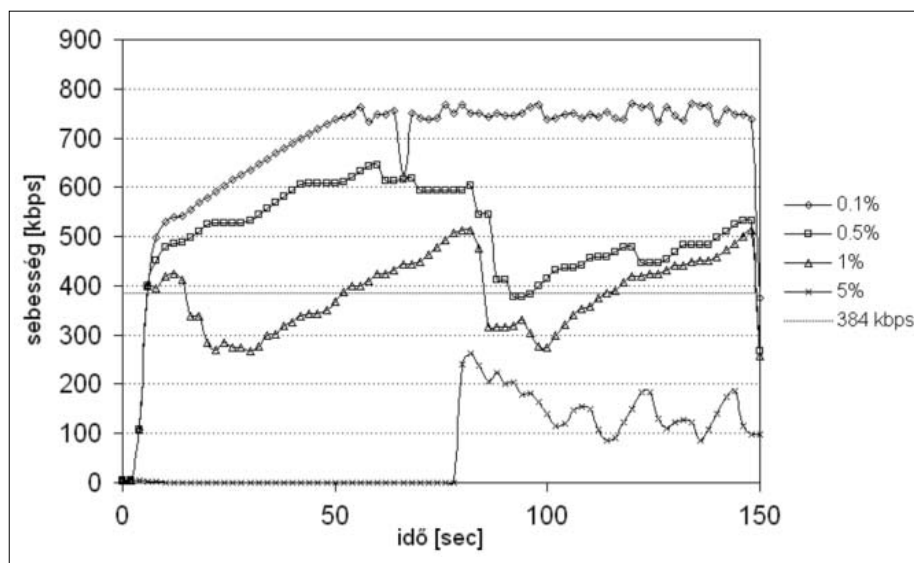
Abban az esetben, amikor az A - B és B - C link kapacitása is meghaladja az MPEG videó sebességét, az újraküldést a hálózati erőforrások szűkössége nem korlátozza. A 2. ábrán a TFRC által meghatározott küldési sebesség és a vételi sebesség látható az idő függvényében, 384 kbit/s kódolási sebességű MPEG videó esetén. A mérési elrendezésben a csomagvesztési valószínűség nulla, D_{A-B} és D_{B-C} egyaránt 30 ms.

A TFRC küldési sebességét jelentősen befolyásolja a csomagvesztés alakulása. Az adóoldal nem tudja meghatározni, hogy mi a csomagvesztés oka, ezért egyformán reagál a torlódás miatti és a rossz minőségű csatorna okozta csomagvesztésre. Ha a csomag nem érkezik meg a vevő oldalra és a videó sebessége a link szabad kapacitása alatt van, a csomagot újra lehet küldeni, ezzel javítva a videó minőségét.

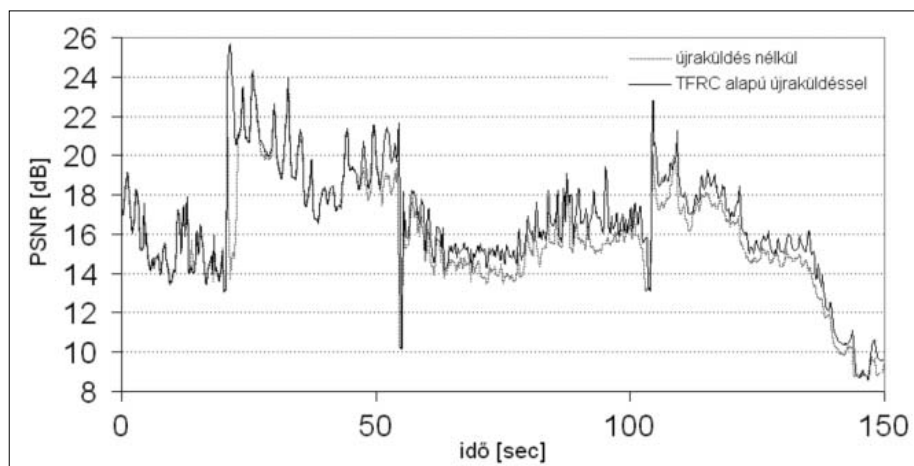
Megfigyelhető, hogy torlódás esetén az RTT értéke jelentősen megnő, mivel a csomagok a routerek tárolói-ban várakoznak, míg a csatorna hibázása esetén az



2. ábra
A TFRC küldési sebesség és a vételi sebesség az idő függvényében



3. ábra
TFRC sebesség a csomagvesztés függvényében



4. ábra
MPEG videó minősége (PSNR)

RTT értéke nem változik jelentősen. A TFRC algoritmus ezért nem a pillanatnyi RTT értékkel számol, hanem annak simított értékével, ahogy azt korábban bemutatuk.

A 3. ábrán a TFRC által meghatározott küldési sebesség alakulását ábrázoltuk, különböző csomagvesztési valószínűségek mellett. A linkek kapacitása (1 Mbit/s) továbbra is nagyobb, mint a videó sebessége (384 kbit/s), tehát van szabad kapacitása a csomagok újraküldésére.

5. ábra MPEG videók összehasonlítása



(a) eredeti videó



(b) TFRC alapú újraküldéssel (PSNR=17.18dB)



(c) újraküldés nélkül (PSNR=16.36dB)

A TFRC működésére az RTT értéke is hatással van, ami az *A-B* és a *B-C* link esetén is 30 ms. A csomagvesztés 0.1% és 0.5% értéke esetén a torlódáskerülő algoritmus által meghatározott küldési sebesség folyamatosan a videó sebessége felett van. Hangsúlyozni kell azonban, hogy nagyobb RTT esetén ez már nem állna fenn. Amikor a csomagvesztés 1%, a TFRC sebessége többször a videófolyam sebessége alá csökken. A TFRC alapú újraküldési algoritmus alapján, az adott mintahálózaton, 0,5%-nál kedvezőbb csomagvesztési arány esetén a sérült csomag újraküldése lehetséges, míg 5%-os csomagvesztés esetén az újraküldés nem engedélyezett.

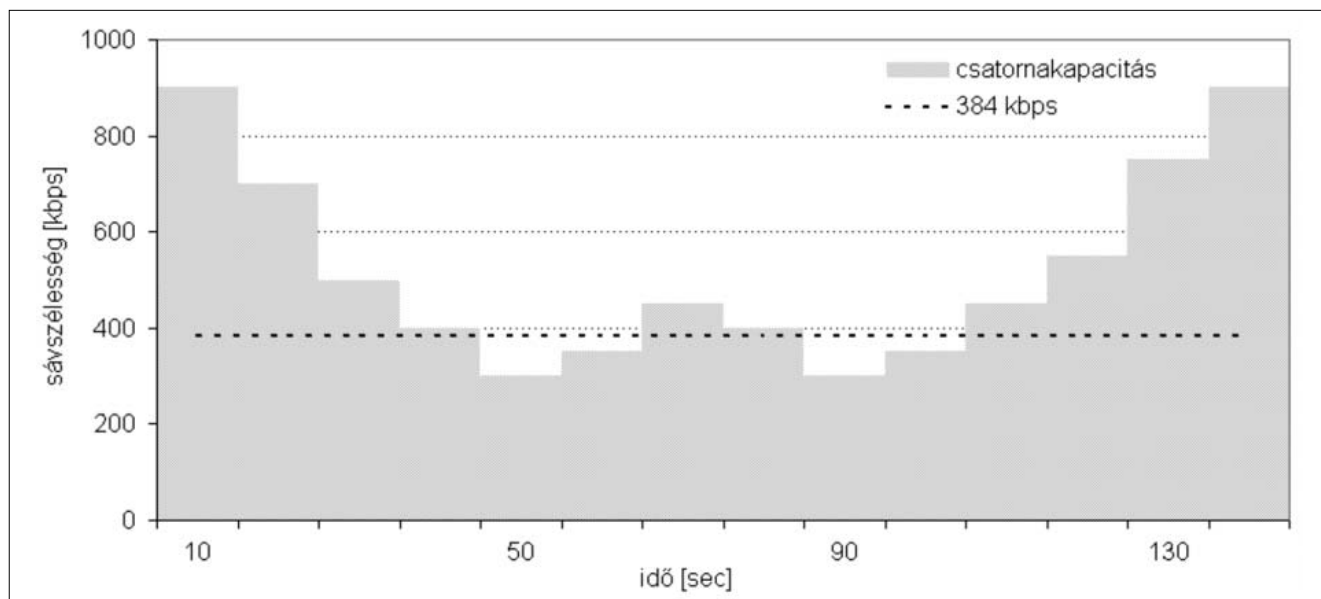
Látható, hogy ez az arány már a kapcsolat felépítésénél is gondot okoz, hiszen csak a 70. másodpercben sikerült azt létrehozni. A TFRC nagyon nagy csomagvesztési arány mellett nem tudja jól megadni a szabad csatornkapacitás mértékét, emiatt tapasztalhatjuk azt, hogy annak ellenére, hogy van szabad kapacitás, a TFRC már jelentősen visszafogja a küldési sebességet.

Az újraküldött csomagok jótékony hatással lesznek a videófolyam minőségére. Ha egy I-kerethez tartozó információt sikerül sikeresen eljuttatni a vevőhöz, a javulás jelentős lesz, hiszen erre a kerettípusra épül a többi. A videó minőségének mérésére a PSNR (Peak to Peak Signal to Noise Ratio) eljárást használjuk, ami egy objektív mérőszám. A TFRC alapú szelektív újraküldés hatását a videó minőségére a 4. ábrán mutatjuk be.

A TFRC alapú újraküldés alkalmazásával a jel-zaj arány növekedett, úgy, hogy közben a forrás sebessége nem haladta meg a TFRC küldési sebességét. Újraküldés nélkül az átlagos jel-zaj arány 15,86 dB, míg az általunk kidolgozott módszerrel 16,6 dB. A PSNR módszer az eredeti MPEG videóhoz hasonlítja a sérült videófolyamot, emiatt az időbeni elcsúszást is természetesen hibának érzékeli. Érdemes tehát más eljárásokkal együtt vizsgálni a videók minőségét, mivel egy adott pillanatban a kép minősége sokkal jobb is lehet, azonban az időbeni késés miatt a PSNR módszer csak kis különbséget érzékel. Az 5. ábra a 450. képkeretet mutatja, ahol PSNR csak kis különbséget mutat a (b) ábra javára.

A következő vizsgálatban változó szabad kapacitás áll rendelkezésre, amely időnként 384 kbit/s alá csökken, miközben a forrás pontosan ekkora sebességgel küldi a csomagokat a hálózatra. Ezekben a periódusokban tehát a *B* csomópont bufferje túlcsoordulhat, vagyis torlódás van a hálózatban. A TFRC pedig ennek megfelelően csökkenti a küldési sebességet. A vizsgálatban a *B-C* link kapacitása a 6. ábra szerint alakul (lásd a következő oldalon).

A csatorna hibavalószínűsége 0,1%, azonban a csomagvesztések nagy száma a buffer túlcsoordulása miatt történnek. Egy vizsgálat 150 másodpercig tart, amely alatt körülbelül 4800 felhasználói adatot tartalmazó csomag kerül továbbításra. A csatorna 0,1%-os hibavalószínűsége miatt közel 5 csomag veszik el, míg a buffer túlcsoordulása miatt több mint 50.



6. ábra A csatorna kapacitása

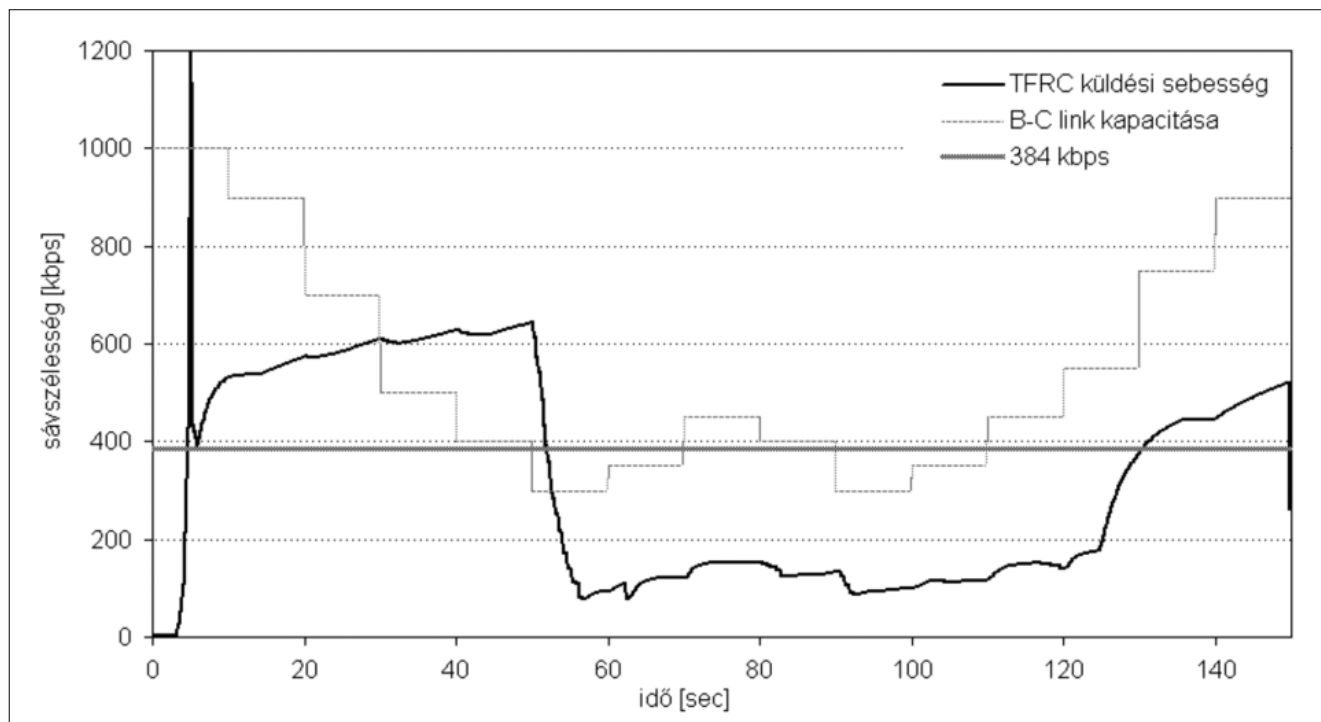
A TFRC küldési sebesség nem tér el jelentősen a 7. ábrán bemutatott eredményektől akkor sem, ha nincs csomagvesztés a csatornán. Szembetűnő, hogy a TFRC szinte azonnal a 384 kbit/s-os szint alá csökkentette a küldési sebességet, ahogy a szabad sávszélesség kisebb lett ennél a szintnél. A sebességnövelő szakaszon viszont sokkal visszafogottabban nő a küldési sebesség. Ennek az oka az, hogy a becsült körbefordulási idő értéke csak lassan csökken, részben a számítási módja miatt, másrészt mivel a router bufferjei lassan ürülnek ki.

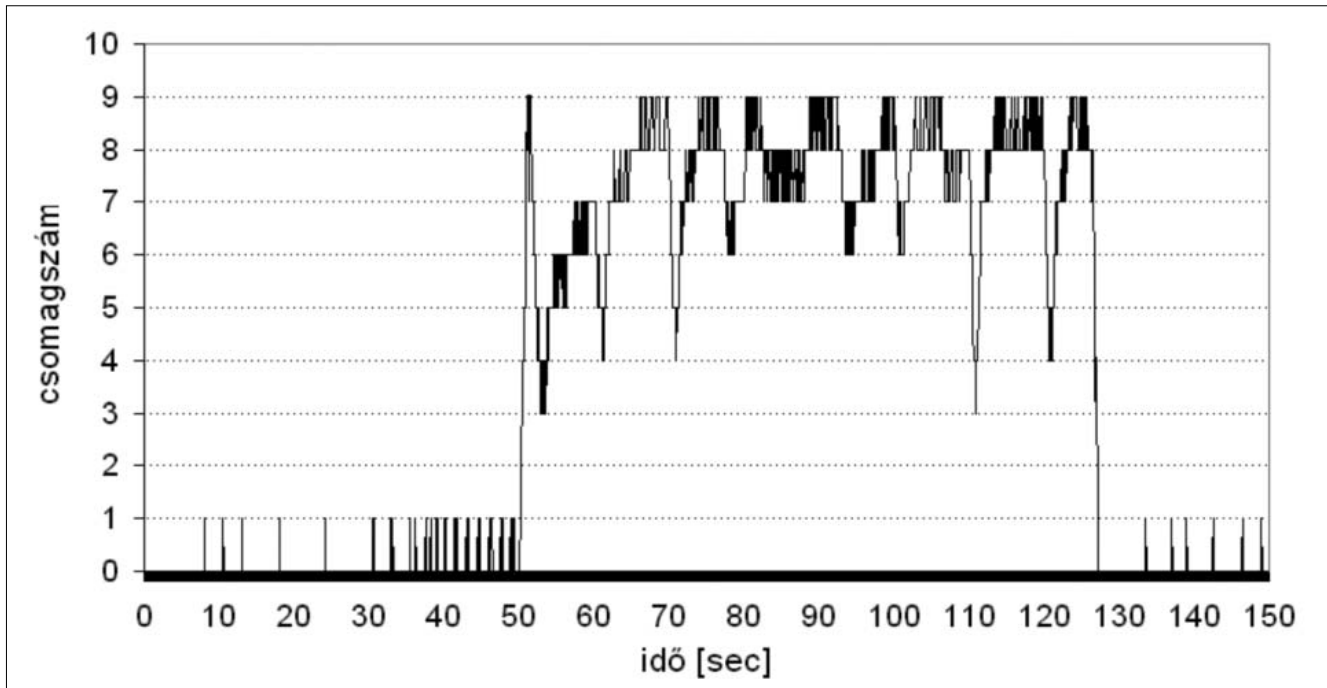
A következő ábra a B csomópont telítettségét mutatja, mely a 10. csomagot már nem tudja tárolni, így az

elveszik. A 8. és 9. ábrák jól mutatják, hogy a buffer telítettségi szintje és az RTT értéke szorosan összefügg.

A TFRC alapú újraküldés algoritmus az ábrák alapján az 50-130 sec intervallum kivételével lehetőséget ad a hiányzó csomagok újraküldésére. Megfigyelhető, hogy az RTT értéke is az 50-130 sec intervallumban tér el jelentősen a kezdeti értékektől. Ebben az intervallumban hiába küldenénk újra a csomagokat, azok csak tovább növelnék a hálózat terheltségét. Elképzelhető, hogy az újraküldött csomag kis valószínűséggel megérkezik a vevő oldalra, azonban nagy valószínűséggel másik csomag elvesztését okozza, így valójában nem javít a szolgáltatás minőségén.

7. ábra Szabad sávszélesség és TFRC sebesség





8. ábra A B csomópont bufferjének telítettsége

A vizsgált változó kapacitású hálózatban ismét javítani tudtunk a videó minőségén. A PSNR módszerrel történt összehasonlítás eredménye a következő oldali, 10. ábrán látható.

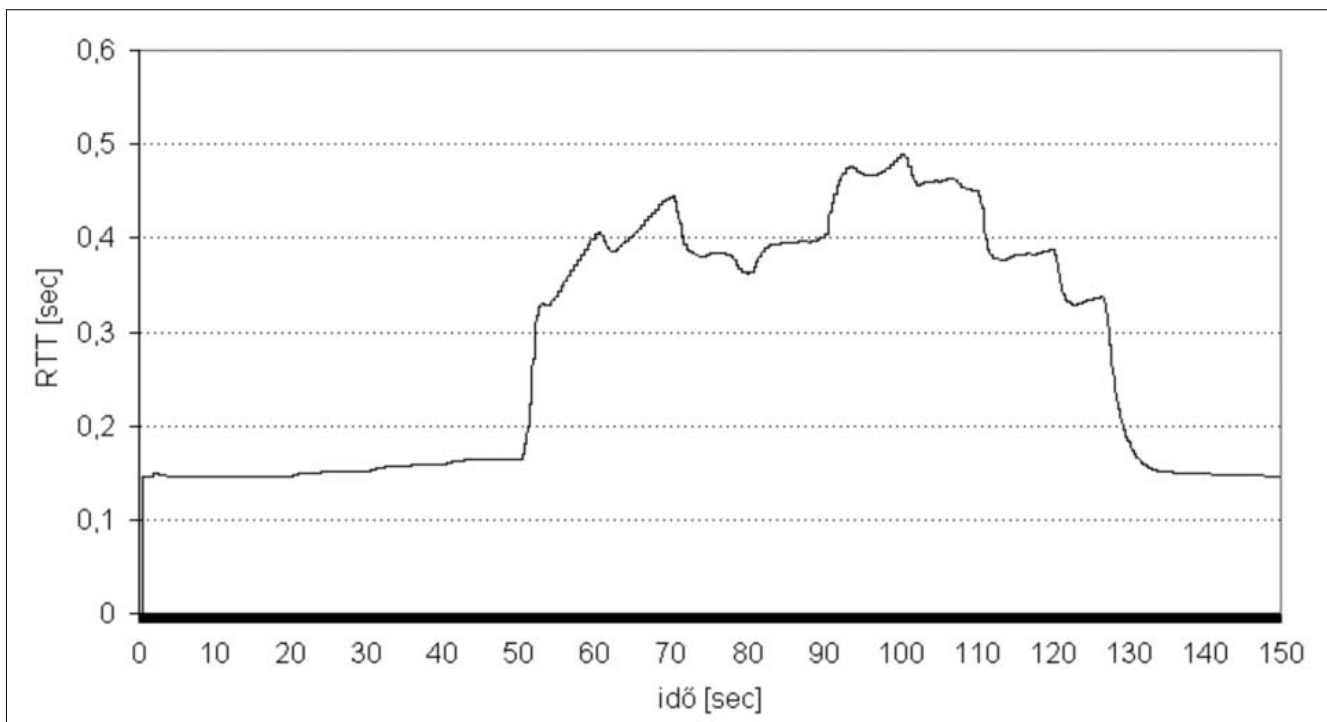
A minőség javítása az 50. másodperc előtti szakaszon nagyon hatékony, hiszen itt minden hibás csomag újraküldhető, ezáltal időbeni csúszásra sem kerül sor. A 130. másodperc után volt ismét lehetőség újraküldésre, azonban a PSNR ezt nem tudta kimutatni. Szemmel azonban jól érzékelhető a különbség. A TFRC alapú új-

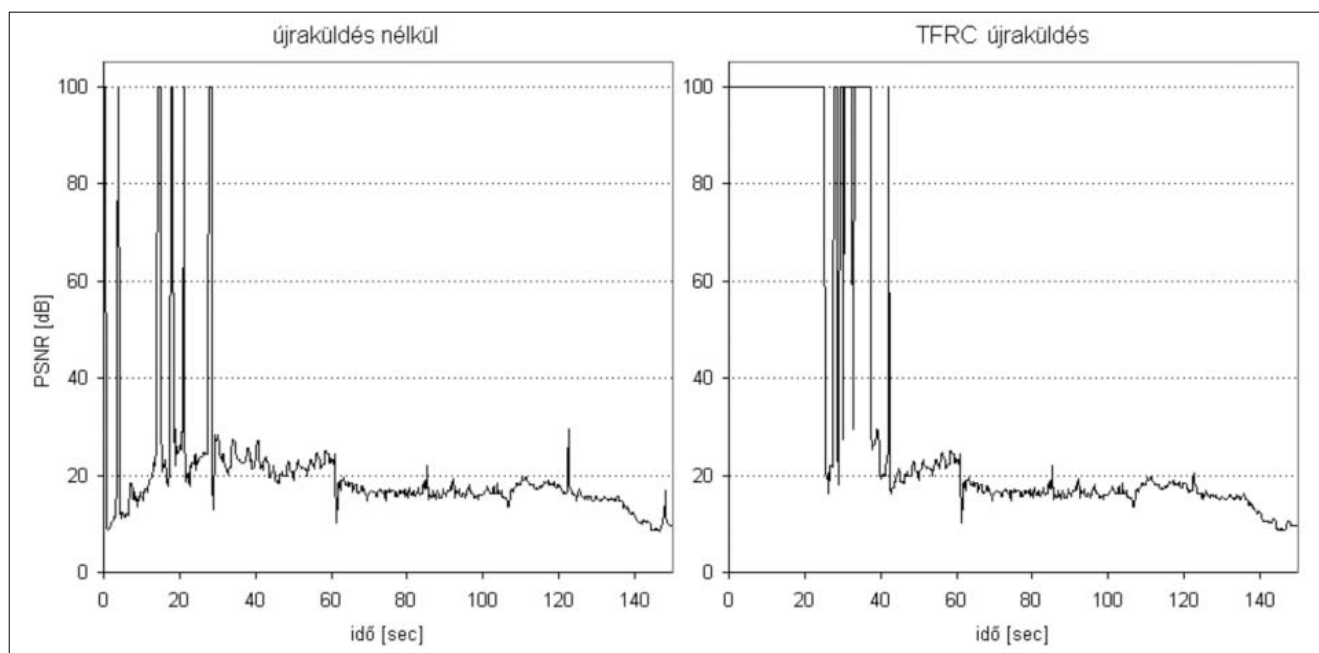
raküldéssel az átlagos PSNR értéke 36 dB, újraküldés nélkül pedig 19,9 dB. A nagy különbség az 50. másodperc előtti szakasz hibátlan átvitelének köszönhető.

5. Összefoglalás

A TFRC alapú szelektív újraküldés lehetővé teszi jobb minőségű szolgáltatások nyújtását, olyan alkalmazások esetén, amelyek nem igényelnek megbízhatóságot, vagy-

9. ábra Körbefordulási idő (RTT) alakulása





10. ábra A videó minőségének összehasonlítása

is a csomagvesztést kezelni tudják. Vizsgálatainkat állandó sebességű MPEG típusú videófolyam esetére végeztük el, azonban az eljárás használható változó sebességű egyéb forrás esetén is.

Az eredmények megmutatták, hogy jelentős javulást tudunk elérni a videó minőségében, a hálózati szabad kapacitás figyelembevételével. Ezáltal úgy javítunk a végleges csomagvesztési arányon, hogy a megosztott csatorna többi felhasználóinak kommunikációját ez nem zavarja.

A jövőbeni tervek között szerepel az idő alapú szelektivitás vizsgálata. Valós idejű multimédiás szolgáltatás esetén fontos tényező a csomagok késleltetése, ezért az újraküldés csak akkor lesz eredményes, ha a csomag még időben megérkezik a vevőhöz.

Irodalom

- [1] J. Postel:
„User Datagram Protocol”,
RFC-768, August 1980.
- [2] Larzon, Degermark, Pink:
„The Lightweight User Datagram Protocol”,
RFC-3828, July 2004.
- [3] R. Stewart:
„Stream Control Transmission Protocol”,
RFC-2960, October 2000.
- [4] Kohler, Handley, Floyd:
„Datagram Congestion Control Protocol”,
RFC-, March 2006.
- [5] S. Floyd, E. Kohler:
„Profile for DCCP Congestion Control ID 2:
TCP-like Congestion Control”, 10 March 2005.
draft-ietf-dccp-ccid2-10.txt
- [6] S. Floyd, E. Kohler, J. Padhye:
„Profile for DCCP Congestion Control ID 3:
TFRC Congestion Control”, 10 March 2005.
draft-ietf-dccp-ccid3-10.txt
- [7] J. Postel:
„Transmission Control Protocol”,
RFC-793, September 1981.
- [8] Huszák Á., Imre S.:
„Selective retransmission of MPEG video streams
over IP networks”, 5th Int. Symposium on Comm.
System Networks and Digital Signal Processing,
CSNDSP 2006, Patras (GR), 2006. július 18-23.,
ISBN 960-89282-0-6.
- [9] N. Wakamiya, M. Miyabayashi, M. Murata, H. Miyahara:
„MPEG-4 video transfer with TCP-friendly rate control”,
LNCS 2216, p.29.
- [10] NS2 – Network Simulator,
<http://www.isi.edu/nsnam/ns/index.html>

Vezetéknélküli és mobil szélessávú hozzáférési technológiák alkalmazása

CSABA TAMÁS

Pannon GSM Távközlési Zrt.

tcsaba@pannon.hu

Kulcsszavak: UMTS/HSPA, WiMAX, beruházás igény

A rádiós hozzáférési technológiák fontos fejlődési lépcsőhöz érkeztek, az általuk nyújtható adatátviteli sebesség jelentősen megnőtt. A két fő technológiai megközelítés az IP alapú hozzáférési módok (WLAN, WiMAX), és a mobil technológiák (GSM/(E)GPRS, UMTS/HSPA) között a számos hasonlóság mellett jelentős eltérések is mutatkoznak. A különbségek részben a technológiában, részben azok helyi alkalmazhatóságában rejlenek. Az itt leírtak nagyban támaszkodnak a Pannon, illetve annak anyavállalata, a Telenor tapasztalataira.

1. Bevezetés

A rádiós hozzáférési technológiák jelentős fejlődési lépcsőhöz érkeztek. Az üzletileg is sikeres beszédátvitel mellett megjelentek a szélesebb illetve szélessávú hozzáférési módok, melyek már versenytársaivá válhatnak az elterjedt vezetékes xDSL technológiáknak. A kettős konvergenciahatás következtében részben a mobil technológiák által nyújtott felhasználószintű sáv szélességek nőttek meg, részben az IP alapú már eleve szélessávú hozzáférési módok váltak vezetéknélkülivé. Ennek eredményeképpen számos technológiai megoldás született, melyek viszonyai és alkalmazási lehetőségei kerülnek összefoglalásra. A figyelembe vett aspektusok a következők:

- felhasználói igények,
- a technológia képességei, alkalmazási lehetőségei,
- rentabilitás, megvalósíthatóság.

2. Felhasználói igények

A megfelelő szélessávú technológia kiválasztását nagyban befolyásolja a már ismert és a várható szolgáltatási igények hálózattal szemben támasztott követelményei. Ezek a jellemzők a szolgáltatás sáv szélességigénye, késleltetésre való érzékenysége, illetve a felhasználás módja és környezete.

A szolgáltatások többsége aszimmetrikus, a hálózattól az előfizető irányában többszörös adatátviteli sebességet igényel. Az elvárt csúcssebesség a TV műsorszórástól, illetve filmek gyors letöltésétől eltekintve, jellemzően nem haladja meg az 1 Mbit/s-ot.

A hálózattal szembeni elvárások alapján a következőképpen kategorizálhatjuk a felhasználási módokat.

- *Helyhez kötött vagy fix (F)* használat, ahol egy adott pontban vagy annak szűk környezetében jelentkezik az igény. A hálózattervezés szempontból az előfizetők nagy területen egyenletesen oszlanak el.

- *Nomád (N)* a használat, ha az igény területileg elszórva több helyszínen is jelentkezhet, de azok között nem. Az előfizetők koncentráltan jelennek meg az egyes hozzáférési pontokon.
- *Mobil (M)* használat, ami részben a mozgás közbeni használatot, részben a szolgáltatás egy tetszőleges helyről történő elérését jelenti. Az elvárt sáv szélességigénye ennek a módnak jellemzően kisebb az előbbieknél. Hasonlóan a fix használathoz a felhasználók egyenletesen oszlanak el az ellátott területen, de használat közben helyet is változtathatnak.

A mobil hozzáférési hálózat a területi hozzáférést vagy a lefedettséget tekintve alkalmas fix használat kiszolgálására is, egyedül az igényelt kapacitássűrűség lehet korlátja a tényleges alkalmazhatóságnak.

3. Rádiós hozzáférési technológiák

3.1. Jellemzők

Két fő fejlődési irány ismert. Az egyik az IP alapú vezetéknélküli technológiák képességeit egészíti ki költségkímélőbb, nagyobb területi lefedés lehetőségével és a mobilitás támogatásával. A WLAN mellett 2006-ban vált elérhetővé a WiMAX (Worldwide Interoperability of Microwave Access) rendszer fix változata. Ebben az anyagban a WiMAX-ot a felhasználók szempontjából egy hálózati hozzáférést nyújtó megoldásnak tekintjük. A másik fő irányvonal technológiai a 3GPP szabvány családban tartoznak. Ezek az EDGE (Enhanced Data rate for Global Evolution), az UMTS (Universal Mobile Telecommunication System) és annak továbbfejlesztett változata a HSPA (High Speed Packet Access). A technológiák főbb jellemzőit foglalja össze a következő oldalon az 1. táblázat.

A felsorolt jellemzők alapján, hosszabb távon a HSPA és a WiMAX alkalmas nagyterületű hozzáférést biztosítására és a felsorolt felhasználási módok mindegyikének lefedésére. Természetesen nem szükségszerű az

összes felhasználói mód kiszolgálása, de ez a tulajdonság növeli a technológia piaci lehetőségeit.

A lényegesebb eltérések a hozzáférés módja, a támogatott modulációk fajtái, a támogatott működési frekvenciák, illetve a technológiák jelenlegi érettségében, bevezethetőségében mutatkozik meg.

A WiMAX jelenleg magasabbrendű modulációs módok támogatása miatt nagyobb hálózati kapacitásra és felhasználó szintű adatsebességre képes. Azonban napjainkban csak a fix változat tekinthető érettnak, alkalmasnak a piaci bevezetésre. Ezen felül a technológiai korlátok következtében a nagyvárosi használata is limitált. A mobil változat a mobilitás támogatása mellett, olyan megoldásokkal rendelkezik majd – az UMTS/HSPA-hoz hasonlóan és a fix változattal szemben – amelyekkel képes lesz nagyvárosi területek spektrumhatékony lefedésére is. A mobil WiMAX szabványt is támogató berendezések piaci megjelenése azonban késik, a valós hálózati körülmények között szerzett tapasztalatok gyűjtése idén is folyik.

Az UMTS/HSPA előnye a technológia érettsége, alkalmazhatósága. Bár jelenleg kevésbé gyors átvitelt támogat, annak mértéke nagy valószínűséggel kielégíti majd az összes szolgáltatás által igényelt szintet.

A táblázatban lévő adatátviteli sebességek a jelenre vonatkoznak, és elsősorban a különbségeket érzékelte. A valós hálózati környezetben tapasztalható minőség függ a szolgáltató állomás távolságától, a hálózat terheltségétől, a zaj és az interferencia szintjétől, az alkalmazott terminál képességeitől egyaránt.

A Pannon EDGE, UMTS, illetve HSDPA technológiai tapasztalatokkal rendelkezik. Az előbbi kettő esetén a táblázat értékei megegyeznek a hálózatban mért adatsebességekkel. A maximális alkalmazási szinten mért letöltési sebesség 16QAM moduláció használatával elérte a 3 Mbit/s-et, míg a különböző rádiós környezetekben végzett mérések átlaga QPSK modulációval is meghaladta az 1 Mbit/s-ot, nomád felhasználás esetén az 1,2 Mbit/s-ot.

A technológiák alkalmassága, a bennük rejlő üzleti lehetőségek a piaci kereslet mellett a különböző felhasználási módok költséghatékony támogatásától függ. Rádiós hozzáférés esetén a rendelkezésre álló sávszélesség, a nagy területek költséghatékony ellátása, valamint a mobilitás támogatása a legfőbb ismérvek. A technológiai adottságok alapján az EDGE és az UMTS inkább a mobil típusú, a WLAN inkább a nomád típusú, míg a WiMAX és az UMTS/HSPA mindhárom módra egyaránt alkalmas. Fontos megjegyezni, hogy a felhasználók várhatóan kevésbé fogják mozgás közben igényelni a szélessávú szolgáltatásokat, viszont azokhoz számos helyről szeretnének majd hozzáférni.

A táblázatból az is kiolvasható, hogy mindkét technológiai irányvonal egyaránt alkalmas lehet az ADSL kiváltására, azzal megegyező szolgáltatási minőség nyújtására.

3.2. Előretékinítés

A fejlődés következő fázisai a spektrumhatékonyság, a hálózati kapacitás, a felhasználói szintű adatátviteli

1. táblázat Technológiák jellemzői

Jellemző	EDGE /3GPP/	UMTS - Rel. 99' /3GPP/	UMTS / HSPA /3GPP/	WiMAX (fix: 802.16.a/g, mobile: 802.16.e) /IP/	WLAN (802.11.a, b) /IP/
Hozzáférési mód	TDMA – FDD	WCDMA – FDD	WCDMA – FDD	OFDM(A) – TDD (mobil), FDD	DSSS, OFDM
Moduláció	GMSK / 8PSK	QPSK	QPSK, 16QAM	QPSK, 16QAM, 64QAM	BPSK, QPSK, 16QAM, 64QAM
Max. felhasználói adatsebesség le / fel (indikatív értékek)	200 / 100 kbit/s	384 / 64 kbit/s	3.0 (10) Mbit/s / 384 kbit/s (5 Mbit/s)	20 / 5 Mbit/s	6, 3 / 6,3 Mbit/s
Működési frekvenciák [MHz] (tervezett)	900, 1800	2000, (900, 1800, 2500)	2000 , (900, 1800, 2500)	Mobile: 2300, 2500, Fix: 3500 , (5800)	2400, 5400-5800
Frekvenciasáv engedélykötelessége (Magyarországon)	Igen	Igen	Igen	Mobil: nem áll rendelkezésre 3500 MHz: igen	Nem
Mobilitás támogatása	Igen	Igen	Igen	Mobil verzió 2007/2008 -tól	Korlátozott
Nagy területek lefedésére alkalmas	Igen	Igen	Igen	Igen, de városi környezetben inkább a mobil változat	Nem
Jövőállóság (1-3)	1	1	3	3	3
Érettség (1-3)	3	3	2	1	3
Támogatott felhasználói módok	M	M	F, N, M	F, N, M	F, N

sebességek további növelését célozták meg mindkét szabványcsaládon belül. A MIMO (Multiple-Input Multiple-Output) technika bevezetése és újabb modulációs módok támogatása várható a következő években.

Az EU szabályozás által szorgalmazott technológia független frekvenciasáv használat elősegíti majd a rádiós szélessávú hálózatok további növekedését.

4. Rádióhálózat méretezése

Egy szolgáltató számára fontos szempont, hogy az arra alkalmas technológiával minél költséghatékonyabban tudja a szolgáltatást nyújtani, mekkora lesz annak beruházásigénye, a beruházásnak várhatóan milyen lesz az időbeli eloszlása. A mobil hálózatokról általánosságban elmondható, hogy nagy beruházásigényűek, amelynek döntő többsége az első pár évben jelentkezik és legnagyobb hányadát a rádióhálózat képezi. Így a szükséges bázisállomások száma van a legnagyobb hatással az üzleti terv eredményére. A hálózat méretezését a szükséges elemszámok meghatározását két tényező, az ellátandó terület nagysága és az ott felmerülő kapacitásigény befolyásolja. A következőkben magyarországi feltételek figyelembe vételével a WiMAX és az UMTS/HSPA kerül összehasonlításra azonos területi ellátottságot megcélózva.

4.1. Lefedettség

A két technológia viszonylatában a cellaméreték közötti különbséget alapvetően az állomás és a felhasználó közötti átviteli út elemei, azok frekvenciafüggése, a maximális kisugározható teljesítmény, a célterület jellemzői illetve a technológia sajátosságai határozzák meg.

A magyarországi szabályozás jelenleg az UMTS számára a 2 GHz körüli, míg a WiMAX számára a 3,5 GHz körüli frekvencián ad lehetőséget szolgáltatásnyújtásra. A szolgáltatók rendelkezésére álló sáv szélességek közel azonosak, 2x15 MHz illetve 2x14 MHz. Továbbá a 3,5 GHz-es sávban kisugározható maximális teljesítményt 14 dBW/MHz-re korlátozták.

A hullámterjedés frekvenciafüggő részei a szabadtéri csillapítás, a diffrakciós csillapítás, a növényzet által okozott csillapítás, az épületcsillapítás illetve a fading nagysága. A szabadtéri szakaszcsillapítás, a diffrakciós csillapítás [1] és a kábelcsillapítás okozta teljesítményvesztés a frekvenciával arányosan növekszik. Az épületek csillapítása függ az épület anyagától és szerkezetétől. Míg a vasbeton falak és a fémfelületek vesztesége a frekvenciával növekszik, addig a rövidebb hullámhossz miatt a magasabb frekvenciás jel kevésbé csillapodik az ablakon való áthaladásakor.

Mindent összetéve nagyvárosi környezetben az épületek csillapítása nagyobb a magasabb frekvenciákon. A hullámterjedési modellek a szakaszcsillapítás várható értékét jelzik előre. A mért lokális átlagok normális eloszlással szóródnak a várható érték körül. A lassú vagy log-normál fading szórása szintén nő a frekvencia növelésével [2], a beépítettség mértékével egyaránt. Eb-

ből következik, hogy egy adott lefedettségi valószínűség eléréséhez nagyobb fadingtartalék szükséges a magasabb frekvenciákon.

Vizsgáljuk meg magyarországi környezetben az egyes beépítettségi kategóriák lefedéséhez szükséges állomásszámokat. A számítások a *Cost 231-Hata*, és a 3,5 GHz-es sávra optimistább *Erceg-Greenstein* [3] modellel történtek. További fontosabb tervezési paraméterek: cellahatáron 90%-os szolgáltatáselérési valószínűség, a log-normál fading és annak frekvenciafüggése, egyetlen fal csillapítása. A kiindulási adatokat és az eredményeket a 2-4. táblázatok foglalják össze.

Paraméterek	UMTS / HSPA	WiMAX-802.16.a
Működési frekvencia [MHz]	2000	3500
1 vivő sáv szélessége [MHz]	5	3.5
Vivők száma	3	4
Frekvencia újrafelhasználás sector alapon	1	2
Max. Vivőszám / sector	3	2
Szektorok száma	3	4
Adó teljesítmény EIRP [dBm]	51	49.4
RX sensitivity level [dBm]	-101	-87

2. táblázat Technológia-specifikus kiindulási adatok

3. táblázat
Számított cellaátmérő
[km]

Terület típus	UMTS / HSPA	WiMAX-802.16.a
Városi	0.4	0.2
Külvárosi	1.2	0.7
Vidéki	4.5	3
LOS	>30	22

4. táblázat
Állomásszám 100 km²/
típus lefedéséhez

Terület típus	UMTS / HSPA	WiMAX-802.16.a
Városi	321	797
Külvárosi	36	103
Vidéki	3	6

Az eredményekből látható, hogy a HSDPA állomás-szükséglete az alacsonyabb működési frekvenciának köszönhetően sokkal kedvezőbb. Érdeemes megjegyezni, hogy tipikus városi környezetben, ahol a felhasználói terminálokat az épületek eltakarják, nem akadálymentes az átviteli szakasz. A felhasználó oldalára telepített szűk nyálábú antennával, az egymással szemben sugárzó cellák nem szeparálhatók. Ezért a mobil technológiákkal szemben a WiMAX fix változatnak használata ebben a környezetben kérdéses.

4.2. Kapacitás

Amennyiben az ellátottsági igények kielégítéséből adódó állomásszámok nem elegendőek a tervezett forgalom lekezelésére, az állomásszámot a kapacitás igény határozza meg.

A számításokat folytatva a lefedettséghez szükséges állomásszámból és a rendszerekre jellemző cellakapacitásokból meghatározható területtípusonként az a kapacitássűrűség-érték, amittől kezdve a WiMAX hálózat fejlesztése válik kedvezőbbé, igényel kevesebb állomást.

Kapacitás	UMTS / HSPA	WiMAX-802.16.a
Vivő kapacitás állomás - terminál [Mbps]	3.2	9
Site kapacitás állomás - terminál [Mbps]	28.8	72

5. táblázat Kapacitásjellemzők állomás-terminál irányban

6. táblázat
Az adódó kapacitás-
sűrűség [Mbps/km²]

Terület típus	UMTS / HSPA	WiMAX-802.16.a
Városi	92	574
Külvárosi	10	74
Vidéki	1	4

Terület típus	Küszöbérték
Városi	230
Külvárosi	30
Vidéki	2

7. táblázat

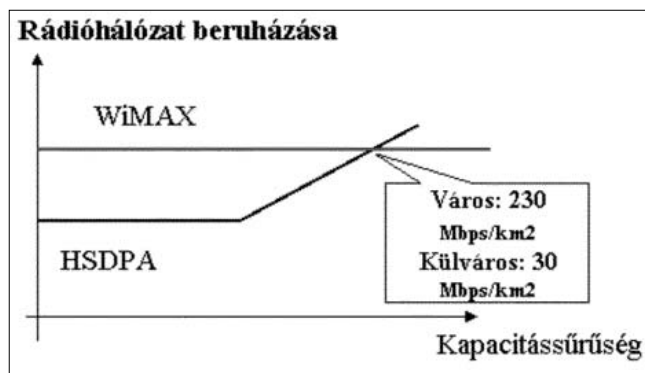
Kapacitássűrűség-értékek ahonnan kezdve a WiMAX kedvezőbbé válik [Mbit/s/km²]

A nagy állomásszám miatt a WiMAX hálózat induló kapacitása jelentősen meghaladja a várható igényeket.

4.3. Beruházásigény

Egy állomás beruházási költsége leegyszerűsítve a telekommunikációs eszközökre, az állomás civil infrastruktúrájára (torony, konténer, áramellátás) és ezek telepítésére fordított költségekből tevődik össze. Míg az utóbbi tíz évben a telekommunikációs eszközök ára jelentősen csökkent a növekvő világpiaci kereslet következtében, addig a többi tényező ára legfeljebb enyhe csökkenést mutatott az alapanyagok, és a hazai munkaerő költségek alakulása miatt. A technológiák különbözősége és a piaci igények különbsége okán a berendezések árában lehetnek eltérések, de azok jelentősége az említettek miatt csökken.

A számítások következményét mutatja az 1. ábra.



1. ábra

Technológiák beruházásszükséglete a kapacitássűrűség függvényében

Azonos területi lefedettség biztosítására több induló állomást, magasabb induló beruházást igényel a WiMAX rendszer. Előnye a nagy kapacitássűrűség a nagy forgalomsűrűség megjelenésével jelentkezhet. A gyártóktól kapott információk alapján az UMTS/HSPA cella-kapacitása az elkövetkező években folyamatosan fog nőni. A növekedést elsősorban újabb rendszerszoftverek telepítésével lehet elérni. Fontos szerepe van tehát az időnek, avagy mikor fog megjelenni a WiMAX képességeit is kihasználó forgalom.

5. Összefoglalás

Az UMTS/HSPA és a WiMAX egyaránt alkalmas a mobil és vezeték nélküli szélessávú Internet-elérés biztosítására. A HSPA-ról szerzett ismereteink megalapozottabbak, a WiMAX hálózatról kapott információk főként tanulmányokból származnak. Ezek alapján megállapítható, hogy az üzemelési frekvencia miatt nagy kiterjedésű, összefüggő UMTS/HSPA hálózat gazdaságosabban, kisebb beruházási lépésekkel fejleszthető. A WiMAX technológia szerepe nagy területen, nagy kapacitássűrűség esetén nő meg, illetve a működési frekvenciasávok kiterjesztésével értékelődhet fel.

A WiMAX megjelenésének további akadályai, hogy a fejlesztések és ezzel együtt a piaci bevezetés a korábbi tervekhez képest késik, elmarad a konkurens technológiától.

Irodalom

- [1] Kazimierz Siwiak (1995): Radiowave Propagation and Antennas for Personal Communication, pp.98–99.
- [2] Recommendation ITU-R P.1406 (1999)
- [3] Channel Models for Fix Wireless Application (2001), IEEE 802.16.3c-01/29r4
- [4] Harri Holma and Antti Toskala (2006): HSDPA/HSUPA for UMTS, pp.130–131.
- [5] IEEE 802.16a-2003 Standard

Mobil kapcsolóközpontok primer és szekunder villámvédelme

BABITS LÁSZLÓ

Pannon GSM Távközlési Zrt., lbabits@pannon.hu

SZEDENIK NORBERT, KISS ISTVÁN, SZÜCS LÁSZLÓ, LÉNÁRT FERENC, BERTA ISTVÁN

Budapesti Műszaki és Gazdaságtudományi Egyetem

{szedenik, kiss}@ntb.bme.hu, {szucs, lenart}@mht.bme.hu, berta@shock.hu

Kulcsszavak: elektronikus berendezések és rendszerek elektromágneses impulzusok elleni védelme, villámfigyelő rendszerek

A mobiltelefon hálózatok fontos részei a központi állomások (MSC – Mobile Switching Centres). Ezeket – a környezetükből jelentősen kiemelkedő adótornyuk miatt – különösen veszélyeztetik a villámcsapások. A központok pedig – a bennük üzemelő nagyszámú elektronikus berendezés okán – érzékenyek a villámcsapás okozta túlfeszültségekre. A szerzők áttekintik az MSC központok villámvédelmének általános kérdéseit, és az üzemeltetés során fellépő jellemző problémákat. Tárgyalják a villámvédelmi szempontból veszélyes tipikus helyzeteket, a védelem kialakításának elméleti hátterét és azokat a tipikus hibalehetőségeket, amelyek a kivitelezés során előfordulhatnak.

1. Bevezetés

A Budapesti Műszaki és Gazdaságtudományi Egyetemen, a Villamosmérnöki és Informatikai Karon működő villámvédelmi iskola évtizedek óta részt vesz a hazai és nemzetközi villám- és villámvédelmi kutatásokban. Kutatómunkájuk lényeges része a különlegesen érzékeny elektronikus eszközök és berendezések fokozott villámvédelmének tervezése és megvalósítása, e kérdés elméleti hátterének és gyakorlati problémáinak vizsgálata.

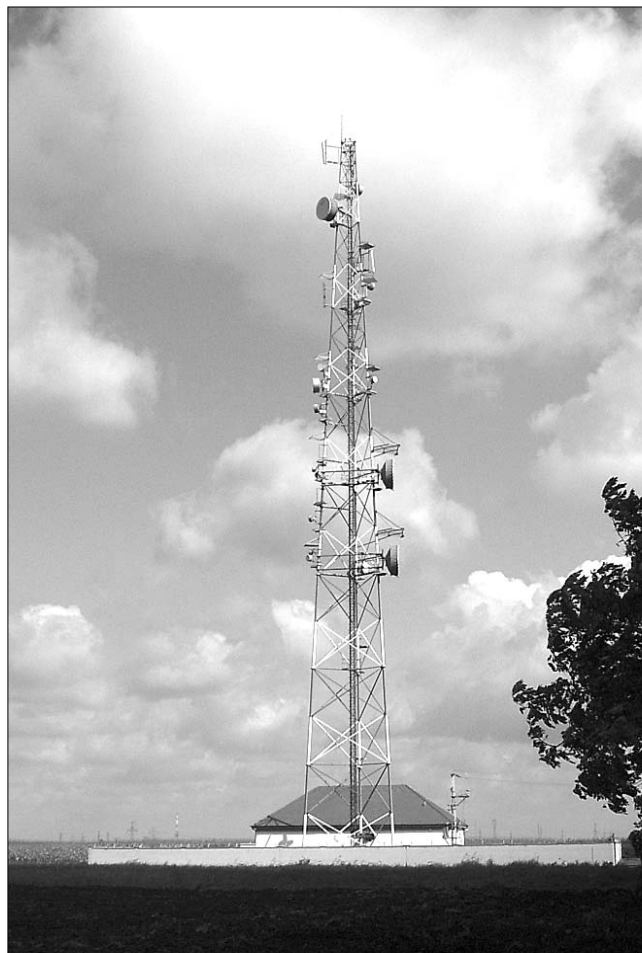
Napjainkra az ilyen eszközöket tartalmazó összetett rendszerek külső elektromágneses zavarok miatti sérülékenysége jelentősen megnőtt. Az elektromágneses zavarok egyik leggyakoribb forrása a villámcsapás, ezért az elektronikus berendezések, illetve rendszerek elsődleges és másodlagos villámvédelme kiemelkedő fontosságú. Különösen igaz ez olyan rendszerek esetén, amelyeknél az esetleges meghibásodásoknak kiterjedt következményei vannak (például különböző vállalatok, intézmények, közszolgáltatók számítógépes és telekommunikációs rendszerében alkalmazott információtechnológia).

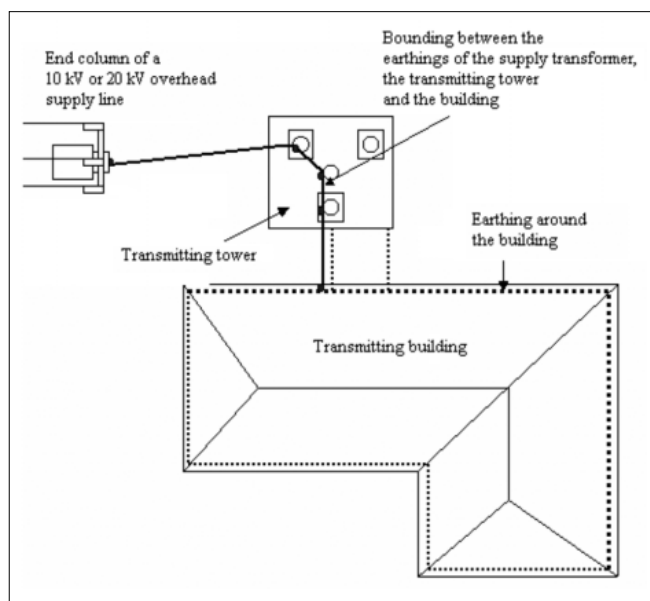
2. Az MSC állomásokon villámcsapás esetén fellépő problémák

A mobil kapcsolóközpontok (MSC – Mobile Switching Center) különösen fontos szerepet játszanak a mobil telefonhálózatban. Gyakorlatilag a telefonközpontokhoz hasonlóan összeköttetést képeznek a mobil telefonhálózat és egy másik telefonhálózat között. Az MSC kezeli a hívási folyamatokat, az átadásokat, az előfizetői szolgáltatásokat stb. Az MSC általában különálló egység, és többnyire vidékre telepítik. Két fő része az adóépület és az adótorony. Ezek a tornyok rendkívül magasak, általában 100 m körüliek. Minden MSC lega-

lább egy külső galvanikus kapcsolattal rendelkezik (energia betáplálás), de további vezetékes kapcsolatok is lehet, mint például telefonkapcsolat, vagy más központokkal kialakított vezetékes kapcsolat. Egy tipikus MSC képét láthatjuk az 1. ábrán, míg a központ sematikus felépítését a 2. ábra mutatja.

1. ábra Tipikus kapcsolóközpont képe





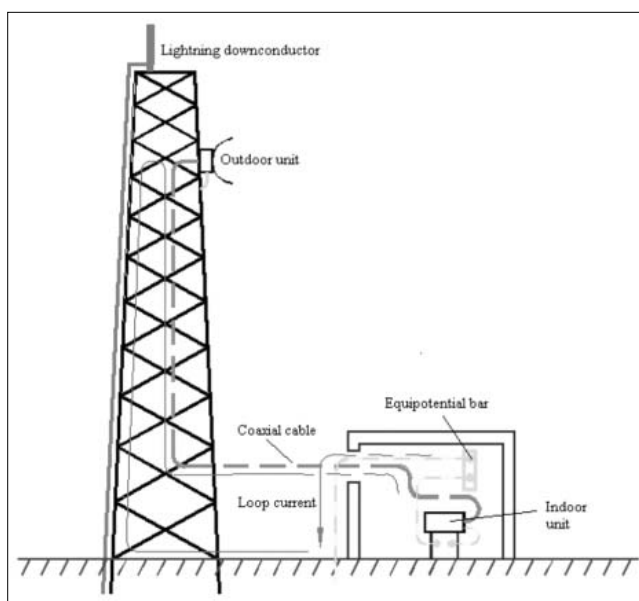
2. ábra Kapcsolóközpont tipikus felépítése

Az adóépület belsejében lévő berendezések és a tornyon elhelyezett kültéri egységek, valamint antennák között számos belső galvanikus összekötés létezik. Ez a felépítés az állomást rendkívül érzékennyé teszi a villám káros hatásaival szemben. A tornyot érő közvetlen villámcsapás bekövetkezési valószínűsége egy átlagos épülettel összehasonlítva viszonylag nagy. Közvetlen villámcsapás esetén a földbe levezetett villámáram a torony és minden hozzá kapcsolódó földelt rész potenciálját megemeli. Ha néhány berendezés „távoli földpotenciálhoz” is csatlakozik, a kialakuló áramok tönkreteszhetik az elektronikát (3. ábra). A villámáram bizonyos hányada a vezetékvezésen keresztül a tornytól közvetlenül a berendezésekhez jut. Ez az áram károsíthatja a berendezések bemeneti részeit.

A meghibásodások másik forrása a villám másodlagos hatása. Mind a közvetlen villámcsapás, mind pedig a közelben bekövetkező villámcsapás esetén nagy indukált feszültség keletkezik a nagyméretű vezető hurokban. Ha egy ilyen hurok érzékeny berendezéshez csatlakozik, az indukált feszültség, amely a 30-40 kV-ot is elérheti, meghibásodásokat okoz.

3. A kapcsolóközpontok túlfeszültségvédelme

A túlfeszültségek okozta igénybevételekkel szembeni védelem hiányában az MSC berendezései károsodhatnak. A meghibásodások két csoportra oszthatók. Egyrészt azokra a túligénybevételekre, amelyek a berendezések fizikai károsodását okozzák, másrészt olyanokra, amelyek hibás működést eredményeznek. A káresemény költsége szintén két részre osztható, a fizikai károsodás költségére és a szolgáltatás kiesése miatti költségre. Bizonyos esetekben az utóbbi sokkal nagyobb lehet, mint a károsodást szenvedett berendezés javításának, vagy cseréjének költsége.



3. ábra Potenciálkülönbségeket okozó tipikus áramutak

A korszerű védelemnek valamennyi fajta túlfeszültséggel szemben hatásosnak kell lennie. Egyik legtöbb gondot okozó hatás a központot a környezettel összekötő valamely vezetón érkező, úgynevezett vezetett túlfeszültség. A vezetett túlfeszültség elleni védelmet valamennyi, a központba befutó vezetékre (vezető anyagból készült összekötésre) ki kell építeni. A túlfeszültségvédelmi eszközöket (SPD – Surge Protective Device) az úgynevezett villámvédelmi zónák határára kell elhelyezni. A zónák definícióját a vonatkozó szabványban találjuk [1]. Az SPD-ket mind az energiaellátás, mind a kommunikációs rendszer vezetékein el kell helyezni. A túlfeszültségvédelmi eszközökből a villámvédelmi zónák szerint kell koordinált védelmi rendszert kialakítani és különös figyelmet kell fordítani a túlfeszültségimpulzusok várható energiaszintjére. Az általános egyenpotenciálra hozás alapvető fontosságú, különösen a villámvédelmi zónák határán, valamint az egymáshoz közel lévő földelések összekötésénél.

A megelőző intézkedések másik részének célja az indukált feszültségek elleni védelem. Annak érdekében, hogy az utóbbi típusú túlfeszültségek csökkenthetők legyenek, a lehető legkisebbre kell korlátozni a lehetséges hurokakat. Ez részben megfelelő nyomvonalak (elrendezések) kialakításával, részben a veszélyeztetett berendezések bemenetén megfelelő finomvédelmi eszközök (D osztályú) alkalmazásával érhető el.

4. Esettanulmány egy feltételezett villámcsapás vizsgálatára

Az előbbieken tárgyalt jelenségek alapján belátható, hogy villámcsapás esetén a megfelelő védelemmel nem rendelkező berendezések károsodhatnak. Ennek szemléltetésére tekintsünk át egy lehetséges forgatókönyvet, amelynek eseményeit az 1. táblázat alapján követhetjük nyomon. Ez olyan információkat tartalmaz, ame-

lyeket egy zivatart követően különböző forrásokból gyűjthetünk össze. A legfontosabb információforrás maga a GSM szolgáltató, amelynek rendszere rögzíti a villámcsapást megelőző és az azt követő eseményeket, valamint a kapcsolóközpontban szolgálatot teljesítő őrt, aki szem- és fültanúja lehet a zivatarnak, illetve a villámcsapásnak. A másik forrás az áramszolgáltató, melynek nyilvántartásában szerepelnek az energiaellátás zavaraihoz kapcsolatos események. A harmadik forrás a meteorológiai szolgálat villámfigyelő rendszere által gyűjtött adatsor. Magyarországon a SAFIR nevű villámfigyelő rendszer működik, az Országos Meteorológiai Szolgálat üzemeltetésében [8]. A SAFIR rendszer által gyűjtött adatok lehetővé teszik a villámok különböző paramétereinek, becsapási helyének és időpontjának meghatározását. A helymeghatározás pontossága néhány kilométer. Szintén rendelkezésre állnak a zivatarfelhő haladásáról készített radarfelvételek.

1. táblázat

Egy zivatar során feltételezett eseménysor idődiagramja

Idő	MSC biztonsági őr	MSC	Országos Meteorológiai Szolgálat	Áramszolgáltató	Ok
02:45		A zivatar miatti leállások az állomástól 100 km távolságra	Zivatarfelhők a környéken		Villámlás, szakaszcsillapítás növekedés az eső miatt
04:25:48	Nagy dörrenés	A térvilágítás kialudt, a kaputelefon tönkrement	Radarfelvételen zivatarfelhő az MSC felett	Közele 20 kV-os transzformátor-nál földzárlat, magától megszűnt	Valószínű villámcsapás a 20 kV-os szabadvezetéken
04:31:12	Nagy dörrenés, ég a 20 kV-os túlfeszültség-levezető	A kültéri, beltéri és a felügyeleti eszközök tönkrementek	Radarfelvételen zivatarfelhő az MSC felett		Valószínű villámcsapás a toronyba
04:39	Kisebb dörrenés		Radarfelvételen zivatarfelhő a környéken	Földzárlat	Valószínű villámcsapás a 20 kV-os távvezetéken
05:44		A dízelgenerátor az áramkimaradás érzékelését követően indul		Kézi lekapcsolás	A rendszer többi részén bekövetkezett hibák elhárítása miatt
05:45				Visszakapcsolás	
05:51-07:39		A dízelgenerátor az áramszünet érzékelése miatt többször indul			Az elszakadt fázisvezető időnként hozzáér a földelt tartóoszlophoz
07:39		A dízelgenerátor indul		A transzformátor előtti oszlopkapcsoló kézi kikapcsolása	Hibaelhárítás
08:24		A dízelgenerátor leáll		Visszakapcsolás	
09:19		Hibaelhárítás kezdete			
03:58		Hibaelhárítás befejezése			

5. Magyarázat

A különböző jelenségek okai a bekövetkezett események sorrendjének tanulmányozásával azonosíthatók. Feltetésünk szerint a villámcsapások egyike akkor történt, amikor a zivatarfront délnyugati irányból érkezve elérte a 20 kV-os körvezetékét 04:25:48-kor az MSC közelében. Tételezzük fel, hogy ennek a villámnak az árama 98 kA. (Ez az érték a SAFIR adataiból állapítható meg.) Ezt a villámáramot az MSC betáplálási pontjánál található 20 kV-os távvezeték-oszlopon elhelyezett levezetők szállították a föld felé. Erre utal az áramszolgáltató jelentésében regisztrált zárlat, amelyet automatikusan elhárítottak. A másodlagos hatások, mint például az indukált feszültség, károsították az udvar világítását, valamint a kaputelefont, amely feltételezésünk szerint nem rendelkezett D osztályú finomvédelemmel. Ekkor az MSC-ben más meghibásodás nem történt, a védelmek tökéletesen működtek, az érzékeny alkatrészeket megvédték a túlfeszültségekkel szemben.

A feltételezett esettanulmány szerint rövidesen egy újabb villámcsapás történt 04:31:12-kor. Tegyük fel, hogy ennek a villámcsapásnak az árama egy rendkívül nagy érték, a SAFIR adatai alapján 107 kA. Feltetésünk szerint ez a villám közvetlenül az adótoronyba csapott. Ennek eredményeként számos kültéri és beltéri berendezés hibásodott meg. A berendezések nagyobbik része azonnal tönkrement, de néhány közülük csak pár órával később. A túligénybevételt jelentősen korlátozta az érzékeny alkatrészek szigetelésének villamos szilárdsága. Teljes átütés ugyan nem következett be, de a túligénybevételeket elszenvedő alkatrészek a továbbiakban nem működtek megfelelően és röviddel a villámcsapás után javíthatatlanokká váltak. A rádiótelefon-forgalom a reggeli órák során növekedett meg annyira, hogy a redundancia hiányában ezek a meghibásodások kiderültek.

A 20 kV-os távvezeték-oszlop tetején elhelyezkedő levezetők szintén károsodtak (4-5. ábra). A beltéri 0.4 kV-os hálózati berendezések nem sérültek, mert ezek nem voltak a toronnyal galvanikus kapcsolatban.

A meghibásodások magyarázata az, hogy a toronyba csapó villám a torony potenciálját jelentősen mege-meli, hiszen a villámáram a földbe a torony levezetési impedanciáján keresztül jut el. Ez azt eredményezi, hogy minden, a toronnyal összekötött földelésnek is megemelkedik a potenciálja, ezáltal jelentős feszültségkülönbségek alakulnak ki a bemeneti csatlakozók és a készülékek távoli földpotenciálra kapcsolódó részei között.

Azoknál a berendezéseknél, ahol ezek a feszültségkülönbségek meghaladták a szigetelések számára elviselhető értéket, meghibásodások léptek fel. Jellemzően a tápegységek és a több berendezéssel is kapcsolatban lévő távfelügyeleti egységek vonalmeghajtó áramkörei hibásodtak meg, hiszen itt alakulhattak ki a legnagyobb potenciálkülönbségek. A kültéri egységnek is a tápegysége hibásodott meg a túlfeszültségre jellemző módon kormos átívelési nyomokat hagyva a nyomtatott áramköri lemezen. A távfelügyeleti berendezésekben



4. és 5. ábra
Meghibásodott 20 kV-os túlfeszültség-levezetők

az integrált áramkörök meghibásodása (szétrobbanás) szintén a túlfeszültségekre jellemző tünet.

Különösen fontos valamennyi vezeték árnyékolásának földelése az épületbe való belépés pontján (villámvédelmi zónahatár!), hiszen a csak a toronyhoz földelt árnyékolás a tornyot ért villámcsapás esetén korlátozás nélkül beviszi a veszélyes zavaráramokat a helyiségben lévő egységekbe. A kialakuló potenciálkülönbségek nagysága csak a levezetett villámáram nagyságától függ, így egy, a fentihez hasonló rendkívül nagy áramú villámcsapás esetén nagy valószínűséggel sok egység hibásodik meg. A gyakrabban jelentkező kisebb villámok esetén nem lépnek fel a berendezések meghibásodását okozó feszültségkülönbségek.

Külön kell szólnunk a 20 kV-os túlfeszültséglevezető meghibásodásáról. Példánkban ezt a meghibásodást is a földelt oszlop potenciálemelkedése okozta. A túlfeszültséglevezetők elektródjai között túlságosan nagy feszültségkülönbség alakulhatott ki. Ez az – előző villámcsapásban már működött, valószínűleg már mechanikai sérülésekkel rendelkező – egyik túlfeszültséglevezető átíveléséhez vezetett. A meggyengült szigetelésű túlfeszültséglevezetőn a 20 kV-os hálózati feszültség tartotta fent az ívet, a villámcsapást követően még órákon keresztül is. Az íváram ugyanis kisebb volt, mint a hálózati védelmek beállítási értékei, így azok nem kapcsoltak le. A tartósan lobogó ív azonban roncsolta a szigetelőtestet, sőt az alumínium sodronyt is megolvasz-

totta. Az elszabadult fázisvezetőt a szél időnként a földelt oszlophoz fújta és ez már sorozatos rövididejű lekapcsolással járt. Ez a magyarázata annak, hogy a dízelgenerátor időnként próbált bekapcsolni.

A meteorológia szolgálat adatai szintén segíthetnek megérteni a jelenségeket. A 2. táblázatban a SAFIR rendszer által egy példaként bemutatott zivatar során gyűjtött adatok találhatók. A táblázat a villámcsapások legfontosabb jellemzőit tartalmazza, így a csúcsáramot, a di/dt értékét, az energiát és a töltést.

A SAFIR rendszer adatait a villámcsapások helyének meghatározásához is fel lehet használni. Egy jellemző diagramot mutat a 6. ábra, amelyen a feltételezett adótorony helyét is bejelöltük.

6. Esettanulmány egy feltételezett villámcsapás vizsgálatára

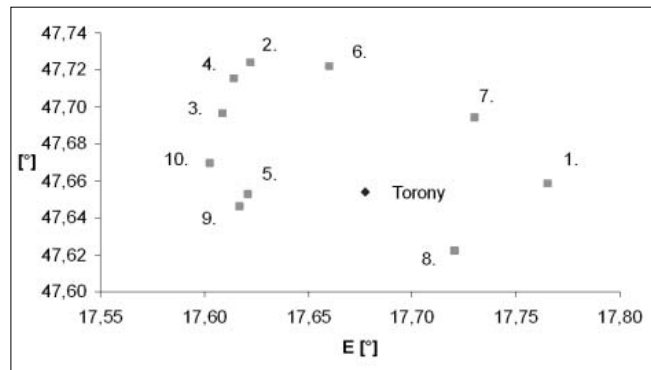
Egy feltételezett villámcsapás során bekövetkező meghibásodások a következők:

- mikrohullámú berendezések meghibásodása (a tornyon és/vagy az átviteltechnikai helyiségben);
- meghibásodások az épületfelügyeleti berendezés (EMS) központi PLC-jében az RS 232-es csatlakozások áramkörei (dízel, DC áramellátó);
- kaputelefon tönkremenetele.

Első lépésként az Országos Meteorológiai Szolgálattól kell az adatokat begyűjteni. A feltételezett esetben a kapott adatok szerint a villámok közvetlenül a tornyot érték, becsaptak az állomást tápláló 20 kV-os légvezetékbe, illetve ezeket elkerülve, de az állomás néhány km-es környezetében csaptak le.

Nézzük először azt a feltételezést, hogy a villám a tornyot érte. Ebben az esetben a mikrohullámú berendezések meghibásodását a toronyból érkező koaxiális kábelben megjelenő feszültségemelkedés idézte elő. Ez akkor valószínű, ha csak azok a berendezések mentek tönkre, amelyek legutoljára lettek telepítve, és a túlfeszültségvédelmük még nem volt kiépítve. A kábelek ár-

6. ábra A torony közelében lecsapó villámok helye



2. táblázat Egy példaként bemutatott zivatar során detektált villámok jellemzői

	Időpont		Ic	di/dt	W	Q
	<i>h/m/s</i>	<i>Ms</i>	<i>kA</i>	<i>kA/μs</i>	<i>A²s</i>	<i>As</i>
1	04:30:41	599,6	-22,1	-4,1	4636	-0,31
		731,6	-21,9	-8,4	5078	-0,35
		824,3	-25,7	-3,1	7142	-0,42
2	04:30:19	101,7	-41,3	-11,8	19647	-0,71
		258,5	-33,8	-4,2	11866	-0,53
3	04:32:08	504,4	-13,9	-2,9	1459	-0,16
4	04:32:16	349,2	-25,0	-7,1	5296	-0,32
	04:32:16	937,6	-29,3	-5,9	7445	-0,38
5	04:32:21	007,1	-22,1	-4,7	4195	-0,29
		108,0	-34,0	-9,7	15243	-0,67
		321,9	-25,8	-2,5	8340	-0,48
		597,0	-24,6	-8,2	5045	-0,31
		779,8	-15,9	-3,0	6080	-0,57
6	04:33:09	78,9	-48,6	-6,2	52732	-1,63
7	04:33:40	745,8	-98,4	-12,1	106964	-1,63
8	04:34:02	589,3	+15,2	+3,8	2092	+0,21
9	04:34:38	955,2	-107,0	-11,1	204464	-2,87
10	04:34:42	558,3	-13,2	-3,3	1211	-0,14
		938,7	-21,4	-8,6	3585	-0,25

nyékolásait a falon való átlépésük előtt le kell földelni, a mikrohullámú berendezéshez való csatlakozásuknál pedig túlfeszültségvédelmi eszközt kell elhelyezni. Mivel feltételezésünk szerint ezek a berendezések nem voltak a torony felől érkező túlfeszültségvédelemmel ellátva, így a tornyot ért villámcsapás esetén gyakorlatilag teljesen védtelenek voltak. Ez megmagyarázza a meghibásodásukat.

Az épületfelügyeleti berendezés (EMS) RS 232-es áramkörei hosszú vezetékvezéssel csatlakoznak a központi PLC áramköreihez, ahol az egyik ér földelt. Ezért a közelben lecsapó villám hatására jelentős indukált feszültség keletkezik a kápcok között, amely tönkretelheti az érzékeny elemeket.

A kaputelefon meghibásodásának oka az épület és a kapu között kialakított kedvezőtlen nyomvonalvezetés lehet. Ilyen esetben minden, a közelben lecsapó villám olyan mértékű túlfeszültséget indukál a vezetékvezetésben, amely meghibásodást okoz. Ez az eset jó példa arra, hogy csak műszaki és gazdasági szempontok együttes figyelembevételével célszerű villámvédelmet kiépíteni. Jelen esetben a kaputelefon megvédésének költsége lényegesen nagyobb, mint az esetleg fellépő kár.

A második feltételezés szerint a villámcsapás a 20 kV-os hálózatot érte az állomás közelében. Feltételezzük, hogy az állomás területén álló oszlop tetején az egyik fázison a 20 kV-os túlfeszültségvezető a villámcsapás során tönkrement. Ha ennek a helyreállítása nem történt meg, akkor újabb villámcsapás esetén az egyik fázis védtelen volt. Így módon a betápvezetéken a túlfeszültségvédelmet csak a B, illetve C osztályú védelmek látták el, amelyek önmagukban nem tudják megfelelően korlátozni a túlfeszültséget. A helyzetet súlyosbíthatja, ha az állomáson a 20/0,4 kV-os oszloptranzformátor és az állomás földelése nincs összekötve.

A közelben becsapó, de sem a tornyot, sem a 20 kV-os vezetékét nem érő villám esetén a meghibásodásokat az indukált feszültség okozza. Ekkor csak azok a berendezések mennek tönkre, amelyekhez nagy kiterjedésű, a zónahatárokon szabadon – azaz túlfeszültségvédelmi intézkedés nélkül – áthaladó vezetékek csatlakoztak.

7. Következtetések

A fenti feltételezett – de tipikus gyakorlati helyzeteket bemutató – esettanulmányokból jól látható, hogy átgondolt, következetes túlfeszültségvédelem hiányában az érzékeny elektronikus berendezések komoly veszélynek vannak kitéve. A hasonló üzemzavarok elkerülése érdekében rendkívül fontos a korszerű túlfeszültségvédelem kiépítése, különös tekintettel az alábbi intézkedésekre:

- valamennyi földelés (épület, torony, 20 kV-os oszlop földelések) összekötése a földben (lásd a 2. ábrát);
- többlépcsős, koordinált túlfeszültségvédelem az energiaellátás felől;

- az épület falát (zónahatárt) átlépő koaxiális kábelek árnyékolását a falon kívül össze kell kötni a földelősinnel;
- az épület falát (zónahatárt) átlépő kábelekhez csatlakozó berendezések megfelelő szintű túlfeszültségvédelmének kiépítése.

Köszönetnyilvánítás

Köszönetet mondunk a Pannon GSM Távközlési Zrt-nek a kutató-fejlesztő munka támogatásáért, valamint néhány tipikus helyszínről és néhány jellemző, túlfeszültség okozta eszközserüléstről készített felvétel rendelkezésünkre bocsátásáért.

Irodalom

- [1] MSZ IEC 61312-1:1997,
Az elektromágneses villámimpulzus elleni védelem.
1. rész: Általános alapelvek
- [2] MSZ EN 62305:2006,
Villámvédelem
- [3] Gulyás, A., Németh, B., Szonda, S., Berta, I.:
„Application of preventive measures in lightning protection”, Proc. of 28th Int. Conference on Lightning Protection, Kanazawa, Japan, 2006.
pp.1211–1216.
- [4] R. Markowska, A. W. Sowa, L. K. Augustyniak:
„Lightning effects on external and internal cabling in telecommunication centres with towers”.
Proc. of 26th Int. Conference on Lightning Protection, Cracow, Poland, 2002. pp.568–573.
- [5] Szedenik, N., Kiss, I., Babits, L., Berta, I.:
„Lightning Protection of Mobile Services Switching Centres”, Proc. of 28th Int. Conference on Lightning Protection, Kanazawa, Japan, 2006.
pp.1316–1321.
- [6] Horváth, T.:
Computation of Lightning Protection,
Research Studies Press, Tauton, 1991.
- [7] Horváth, T.:
Understanding Lightning and Lightning protection.
Whiley, RSP, 2006.
- [8] Dombai, F.:
Magyarországi villámlás-lokalizációs hálózat 2005-2006.
Magyar Elektrotechnikai Egyesület,
XVII. Villámvédelmi Konferencia, 2006. november 7.

Dinamikus spektrumkiosztás: modell és árazás

KOVÁCS LÁSZLÓ, VIDÁCS ATTILA

Budapesti Műszaki és Gazdaságtudományi Egyetem, Távközlési és Médiainformatikai Tanszék
{kovacs, vidacs}@tmit.bme.hu

Kulcsszavak: dinamikus spektrumkiosztás, vezeték nélküli hálózatok, aukció, árazás, interferencia-tolerancia

A rádiós spektrum véges, és ezért rendkívül drága erőforrás. Az erőforrások hatékony kihasználása fontos lenne, ugyanakkor a jelenlegi menedzsment-rendszerek nem támogatják azt. A merev spektrumkiosztás miatt az erőforrások jelentős része „elveszik” az igények térbeli és időbeli változása miatt. Ezenfelül a spektrumhasználati jogokat nem mindig azok birtokolják, akiknek a legnagyobb szükségük van rá, akik a legjobb szolgáltatást tudnák nyújtani rajta. Erre egy lehetséges megoldás a dinamikus spektrumkiosztás (DSA) és ezzel együtt a frekvenciagazdálkodás piaci alapokra helyezése. A cikkben javaslatot teszünk egy központosított spektrumárverési eljárásra, mely az általunk javasolt DSA keretrendszerbe illeszkedik, és alkalmas a fix infrastruktúrájú vezeték nélküli hálózatok hatékony spektrumkiosztásának a menedzselésére. Mivel a dinamikus spektrumkiosztású hálózatok esetében az interferencia tolerancia fontos kérdés, az általunk javasolt megoldás biztosítja, hogy azok a szolgáltatók, amelyek nem tűrik az interferenciát, vagy nagy zavaró hatást fejtenek ki szomszédaikra, többet fizessenek a lefoglalt erőforrásért. A javasolt DSA modell lehetővé teszi, hogy a rendelkezésre álló szűkös erőforráson a szolgáltatók jobb minőségű vagy többféle szolgáltatást biztosítsanak. A javasolt árazási mechanizmus tovább növeli a szolgáltatások minőségét, mivel célja a „közjólét” maximalizálása.

1. Bevezetés

A jelenlegi frekvenciagazdálkodás és -kiosztás következtében a rendelkezésre álló szabad frekvenciák száma egyre kevesebb. Másrészt viszont amellet, hogy a spektrum nagy része már ki van osztva, annak jelentős része kihasználatlan és ez a kihasználatlanság térben és időben is nagyon változó. Ez a motivációja egy hatékonyabb kiosztási technikának, a dinamikus spektrumallokációnak (DSA – Dynamic Spectrum Allocation). A spektrum kiosztásának dinamikussá tételével a szolgáltatások minősége tovább javulhat és a felszabaduló frekvenciák miatt lehetőség nyílik új szolgáltatások bevezetésére is.

A dinamikus spektrumkiosztás maga után vonja a spektrumgazdálkodás piaci alapokra helyezését is. Két fontos tényezőt kell figyelembe vennünk a piaci alapon működő spektrumgazdálkodás esetén. A frekvenciák „kereskedelme” lehetővé teszi a spektrumhasználati jogok átruházását rövid időskálákra is, a „liberalizáció” pedig biztosítja a szolgáltatás- és technológia-semlegességet. A két tényező együttes megvalósítása vezet a hatékony spektrumhasználatához [1].

Szükséges továbbá különböző jogok és kötelezettségek definiálása is a biztonságos és hatékony működéshez. A használati jogok négy fontos paraméterhez köthetők. Ezek a földrajzi terület (egy ország, régió, vagy egy bázisállomás környéke), az időtartam, a spektrumblokk (például a biztosított frekvenciatartomány) és a védelem az interferenciától. A kötelezettségek pedig az okozott interferencia mértékére vonatkoznak. Ugyanakkor a használati jogok tulajdonosa szabadon használhatja a frekvenciát bármire (szolgáltatás-semlegesség),

és bármilyen technológiával (technológia-semlegesség), ha az interferenciára vállalt kötelezettségeket betartja [1].

A következő fejezetben röviden ismertetjük a dinamikus spektrumallokáció és a spektrumgazdálkodás piaci alapokra helyezéséről szóló fontosabb munkákat. A 3. fejezetben megadunk egy tér- és időbeli DSA-keretrendszert, ami a 4. fejezetben bemutatott allokációs és árazási módszer alapját képezi. Ezek után ismertetjük a minimális interferenciájú allokáció meghatározásának módját, majd a 6. fejezetben egy példán keresztül bemutatjuk a javasolt módszer működését. Végül összegezzük az elmondottakat.

2. Kapcsolódó munkák

A felügyelt DSA problémával először az IST-DRiVE és az IST-OverDRiVE projektek foglalkoztak. A cél az volt, hogy kifejlesszenek egy olyan eljárást, amely lehetővé teszi a dinamikus spektrumkiosztást egyazon frekvenciasávban különböző technológiájú hálózatok számára [2-4]. DSA régiókat definiáltak, amin belül az igények csak időben változhattak, térben állandóak voltak. Buddhikot már részletes leírást adott egy tervezett rendszer architektúrájáról [5]. A DSA hálózatokkal kapcsolatos kutatásokról további részletes összefoglalók találhatók [5-7]-ben.

Rodriguez és társai egy olyan DSA rendszert javasoltak [8,9], ahol egy spektrummenedzser periódikusan elárverezi a rövid időtartamra szóló használati jogokat. A javasolt aukció másodfokú volt, a különböző területeket úgy tekintették, mintha szigetek lennének, így nem foglalkoztak a régiók egymásra hatásával.

Amennyiben egy tetszőleges mértékben osztható erőforrást akarunk kiosztani, az egyik alkalmas módszer a progresszív másodáras aukció. Ennek használatára már többen tettek javaslatot vezetékes hálózatokban történő erőforrás kiosztására [10-12].

A fentiekkel összehasonlítva a mi megoldásunkat, a különbségek a következők. A korábbi megoldások az általunk javasolttal ellentétben nem veszik figyelembe a régiók között fellépő interferenciát. A dinamikus spektrumallokáció sajátosságai miatt a vezetékes hálózatok sáv szélesség kiosztására vonatkozó javaslatok nem használhatók a mi esetünkben, de az árak meghatározásának ötletét figyelembe tudtuk venni az általunk kidolgozott megoldás során.

3. Tér- és időbeli DSA modell

Egyik előző publikációnkban [13] egy olyan DSA modellt javasoltunk, amely a régiók között fellépő interferenciát képes rugalmasan kezelni. Mivel a javasolt árazási módszer ezen a modellen alapul, ebben a fejezetben röviden összefoglaljuk a modell alapjait.

A teljes területet kisebb régiókra osztjuk és felteszünk, hogy egy-egy régió belül a spektrumigények térbeli eloszlása állandó, csak az időbeli változás megengedett. A spektrumot egy központi entitás, a spektrumbróker felügyeli és osztja ki a szolgáltatóknak. A használati jogok rövid időre szólnak, továbbá a szolgáltatók a kapott spektrumszeletet bármire használhatják (technológia semlegesség).

A tér- és időbeli dinamikus spektrumallokációt a következőképpen értelmezhetjük. Minden régió belül időbeli DSA-t valósítunk meg. Azaz a régióban működő szolgáltatók periodikus időközönként beküldik a spektrumigényeiket és a spektrumbróker az igényhez hozzárendel egy folytonos spektrumblokkot, amennyiben ez lehetséges. A blokkok mérete változhat az időben, ezáltal követve az igények változását. A térbeli DSA feladata az egy időben, de különböző régiókban fellépő igények összehangolása. A cél az, hogy a spektrumblokkokat úgy osszuk ki, hogy a lehető legkisebbre csökkentsük a régiók között fellépő interferenciát.

Fontosnak tartjuk megjegyezni, hogy nem célunk a dinamikus spektrumallokáció cellás szinten történő megvalósítása. A régiók ennél nagyobb területek (például egy-egy városrész), amelyeken belül több bázisállomás is található. Hasonlóan az időskálát is olyan nagyságrendűnek képzeljük (például óránkénti újraallokálás), amely képes követni az igények napi váltakozását, de nem célunk az igények hívásonkénti kezelése.

3.1. Spektrumallokáció

A teljes területet K darab nem átlapolódó régióra (R_k) osztjuk. Az adott régió belül M szolgáltató versenyez az erőforrásért.

A rendelkezésre álló spektrum, más néven a koordinált hozzáférésű sáv [6] (Coordinated Access Band, CAB), az $(s, \hat{s})$ frekvenciatartomány.

A spektrumbróker feladata a szolgáltatók igényeinek és a rendelkezésre álló spektrumnak az összerendelése. Az m . szolgáltatónak a k . régióban kiosztott spektrumblokk egy adott t időpontban:

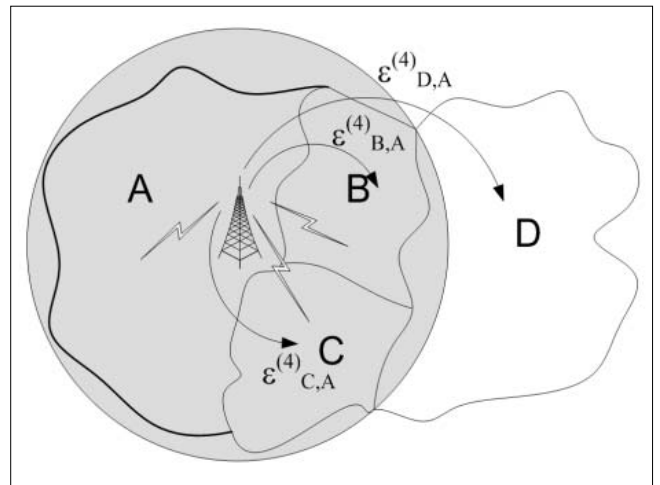
$$S_{m,k}(t) = (\hat{s}_{m,k}(t), \hat{s}_{m,k}(t)). \quad (1)$$

A fenti jelölés kifejezi, hogy a spektrumallokáció mind térben, mind időben dinamikus, minden szolgáltató különböző régióban és különböző időpontokban eltérő igényekkel léphet fel. (A jelölések egyszerűsítésére az időtől való függést a továbbiakban elhagyjuk.)

Jelölje továbbá $|S_{m,k}|$ a kiosztott spektrumblokk méretét, azaz $|S_{m,k}| = \hat{s}_{m,k} - s_{m,k}$.

3.2. Interferencia

A javasolt modellben az interferencia a spektrum kihasználtságának a csökkenését okozza.



1. ábra Geometriai csatolási tényező

A kihasználtság csökken, ha különböző szolgáltatók ugyanazt a frekvenciasávot használják a szomszédos régiókban és a határterületeken interferencia lép fel. Az interferencia szintje függ a földrajzi helyzettől, a régiók méretétől és a használt rádiós technológiától is. Ez a fajta interferencia jellemezhető egy földrajzi csatolási tényezővel (ϵ). Jelölje $0 \leq \epsilon_{l,k}^{(m)} \leq 1$ a „zajt”, amit az R_k régióban működő m . operátor okoz az R_l régióban. A fenti paraméter értéke nulla, ha egyáltalán nincs áthallás a két régió között, és egy, ha teljes mértékben áthallatszik a másik régióba is. Azaz minél kisebb a földrajzi csatolási tényező értéke, az annál jobb az interferencia szempontjából.

A különböző szolgáltatók által használt rádiós technológiák egymást zavaró hatását a rádiótechnológiai csatolási tényező írja le (η). Jelölje $0 \leq \eta_{m,n} \leq 1$ a csatolást az m . és az n . szolgáltató által használt rádiós technológia között.

Ismét a két szélsőséget nézve, ha két szolgáltató ugyanazt a spektrumszeletet használja ugyanabban a régióban és $\eta_{m,n}$ értéke 0, akkor a két szolgáltató gyakorlatilag nem zavarja egymást, míg ha $\eta_{m,n}$ értéke 1, akkor egyik szolgáltató sem tudja használni a spektrumszeletet. A két csatolási tényező együttes hatása a két

paraméter szorzata. Azaz, az R_k régióban működő m . szolgáltatónak az R_l régióban működő, vele azonos spektrumszeletet használó n . szolgáltató által okozott zavaró hatás $\varepsilon_{l,k}^{(n)} \cdot \eta_{m,n}$.

3.3. Spektrum kihasználhatóság

Miután megvannak az interferenciát leíró modellünk paraméterei, jelölje $\xi(S_{m,k})$ az $S_{m,k}$ spektrumblokk kihasználhatóságát, melyet a következőképpen számolhatunk:

$$\xi(S_{m,k}) = \frac{1}{|S_{m,k}|} \int_{S_{m,k}} \xi_{m,k}(\lambda) d\lambda, \quad (2)$$

ahol $\xi_{m,k}(\lambda)$ a λ frekvencia kihasználhatósága az R_k régióban működő m . szolgáltató számára, azaz

$$\xi_{m,k}(\lambda) = \prod_{i=1}^M \prod_{j=1}^K \left(1 - \varepsilon_{j,k}^{(i)} \cdot \eta_{m,i} \cdot I_{\{\lambda \in S_{i,j}\}} \right) \quad (3)$$

Ahol $I_{\{\lambda \in S_{i,j}\}}$ azt jelöli, hogy λ ki van-e osztva az R_j régióban az i . szolgáltatónak. A kihasználhatóság értéke 1, ha egyáltalán nem lép fel interferencia, és kisebb mint 1, ha a szomszédos régiókban működő szolgáltatók interferenciát okoznak.

3.4. Megvalósítható spektrumallokáció

Az $S = \{S_1, \dots, S_{M,K}\}$ spektrumallokációt megvalósíthatónak nevezzük, ha a szolgáltatóknak kiosztott spektrumblokkok teljesítik a következő feltételeket:

$$\xi(S_{m,k}) \geq \beta_m, \quad \forall m, k, \quad (4)$$

$$\min_{\lambda \in S_{m,k}} \xi_{m,k}(\lambda) \geq \alpha_m, \quad \forall m, k. \quad (5)$$

A fő kérdés a megvalósíthatóság mögött az, hogy a szolgáltatás mennyire érzékeny a zavaró hatásokra, milyen mértékben romlik az adott szolgáltatás minősége a kihasználhatóság csökkenésével. Ezért a fenti két feltétel célja annak biztosítása, hogy a kiosztott spektrumblokk „elég jó” minőségű legyen az adott szolgáltatás számára.

Az α és β paraméterekre egyfajta tolerancia küszöbként tekinthetünk, amelyek azt fejezik ki, hogy az adott szolgáltatás mennyire érzékeny az interferenciára. Azt feltételeztük, hogy ha a kihasználhatóság az adott küszöb felett van, a szolgáltatás zavar nélkül biztosítható. A β_m paraméter azt a minimális spektrum-kihasználhatóságot adja meg, amit átlagos szinten el kell érünk (lásd (4)), míg az α_m paraméter egy olyan minimális kihasználhatóságot ír elő, amit minden frekvencián biztosítani kell a kiosztott spektrumblokkban (lásd (5)).

A megvalósíthatóság vizsgálata alapkövetelmény. Ugyanakkor a megvalósíthatóság nem mond semmit az optimális allokációról. A spektrumot kiosztathatjuk jól és rosszul is. Az, hogy mit nevezünk jónak, attól függ, hogyan definiáljuk a kiosztás nyereségét. [14]-ben definiáltunk különböző nyereségeket mind a szolgáltató, mind pedig a hatóság szemszögéből, amelyek egyaránt maximalizálandó célok lehetnek.

A szolgáltatók valós toleranciaszintjének ismerete nagy segítség a dinamikus spektrumallokáció számára. Ugyanakkor a szolgáltatók szemszögéből sokkal egyszerűbb azt mondani, hogy nem toleráljuk az interfe-

renciát, „tisztá” spektrumblokk szükséges a megfelelő minőségű szolgáltatáshoz. Ez viszont nagymértékben lecsökkenti a dinamikus spektrumallokáció hatékonyságát. Ezért szükséges olyan mechanizmusok megvalósítása, amelyek „rákényszerítik” a szolgáltatókat, hogy tolerálják az interferenciát olyan mértékben, ahogy az lehetséges. Így nagy fontosságú egy olyan árazási modell, amely magas árat kér az intoleráns szolgáltatóktól, és azoktól, akik nagy zavaró hatást fejtenek ki a környezetükben lévő többi szolgáltatóra.

4. Aukció és árazás

Ebben a fejezetben egy olyan allokációs és árazási modellt javasolunk, mely figyelembe veszi a dinamikus spektrumkiosztás sajátosságait. A progresszív másodfokos aukció [10] jó megoldás tetszőlegesen osztható erőforrás kiosztására, esetünkben azonban a szétosztható spektrum mérete nem határozható meg az interferencia miatt. Előfordulhat ugyanis, hogy egy régió belül lehetnek olyan frekvenciasávok, melyet senkinek nem tudunk kiosztani a szomszédos területek zavaró hatása miatt. Ugyanakkor bizonyos frekvenciákat akár több szolgáltatónak is odaadhatunk, amennyiben azok nem zavarják egymást. Ezenfelül az árazási modellnek teljesítenie kell azt a követelményt is, hogy azon szolgáltatók, akik nem tolerálják az interferenciát, vagy nagy interferenciát okoznak a szomszédjaiknak, többet fizeszenek a lefoglalt erőforrásért, mint a toleráns, másokat nem zavaró szolgáltatók.

4.1. Licitek, allokáció és árazás

A spektrumot adott időintervallumonként osztjuk újra. Az intervallum kezdete előtt a szolgáltatók beküldik a lehetséges licitjeiket egy központi egységnek, az úgynevezett spektrumbrókernek. Ezt követően a spektrumbróker meghatározza azt az optimális kiosztást, mely a „közjót” maximalizálja, továbbá meghatározza az egyes szolgáltatók által fizetendő összegeket is.

4.1.1. Egy lépéses, több licites aukció

Az iteratív aukciós algoritmusok konvergenciaideje bizonyos esetekben nagyon hosszú lehet és jelentős jelzéstöbblettel járhat. Továbbá spektrumkiosztás esetén az „értelmes” licitek száma véges, így ezeket könnyen megadhatjuk egyetlen licithalmazzal. Ezért a javasolt aukciós modell egy lépéses, továbbá a szolgáltatók az összes lehetséges licitjüket beküldik a spektrumbrókernek. (A továbbiakban a játékelmélet fogalmait használjuk az aukciós eljárás intuitívabb leírásához, amelyben a frekvenciasávokért „versengő” szolgáltatókat mint „játékosokat” vesszük figyelembe. A „játzsma” nem-kooperatív, a játékosok „önzők”, azaz a résztvevők nem működnek együtt, minden játékos a saját érdekeit tartja csak szem előtt. A játékosokat racionális stratégiaalkotóknak tekintjük.)

Jelölje $I = \{1, \dots, I\}$ a játékosok halmazát. Mivel különböző régiókban ugyanazon szolgáltató igényei eltér-

rőek lehetnek, ezért a szolgáltatókat régióként külön kezeljük, azaz $I = M \cdot K$.

i . játékos $N^{(i)}$ darab kétdimenziós licittel rendelkezik:

$$B_i = \{b_{i,1}, \dots, b_{i,N^{(i)}}\}, \quad (6)$$

ahol

$$b_{i,n} = (q_{i,n}, p_i(q_{i,n})), \quad n = 1, \dots, N^{(i)}, \quad (7)$$

q jelöli az igényelt erőforrás mennyiségét, $p(q)$ pedig az érte ajánlott árat.

A bróker összegyűjti a licitek halmazát és létrehozza a licithalmaz-leíró, az árazási modell bemenetét:

$$B = (B_1, \dots, B_I). \quad (8)$$

4.1.2. Allokációs szabály és árazás

A fenti leíró alapján a bróker meghatározza minden $i \in I$ játékos számára a kiosztott erőforrás mennyiséget (a_i) és a hozzá tartozó árat (c_i) az A alokációs szabály és a C árazási szabály alapján.

Az A alokációs szabály egy alokációs vektort határoz meg:

$$A(B) = (a_1, \dots, a_I), \quad (9)$$

$$\text{ahol } a_i \in \{0, b_{i,1}, \dots, b_{i,N^{(i)}}\}, \quad i = 1, \dots, I, \quad (10)$$

azaz az i . játékos számára kiosztott spektrumblokk mérete. A C árazási szabály:

$$C(A(B)) = C(\mathbf{a}) = (c_1, \dots, c_I), \quad i = 1, \dots, I, \quad (11)$$

ahol $c_i \leq p_i(a_i)$ az i . játékos által az a_i mennyiségű erőforrásért fizetendő összeg.

4.2. Hatékony és megvalósítható alokáció

Az alokációs és árazási szabályokat úgy kell megalkotnunk, hogy a kialakult helyzet elfogadható (megvalósítható) legyen minden játékos számára és valamilyen szempontból optimális is. Célunk egy olyan alokációs szabály megvalósítása, amely a teljes spektrumkihasználtságot maximalizálja. (Megjegyezzük, hogy ez a mérték megfelel a széles körben használt „közjóllét” mértékének – lásd [10]).

4.2.1. Megvalósítható alokáció

Definiáljuk az A alokációs szabályt egy adott B licithalmaz-leíróra az alábbiak szerint:

$$A(B) = \arg \max_{\mathbf{a} \in Q^f} \sum_{i=1}^I p_i(a_i), \quad (12)$$

ahol Q^f a megvalósítható alokációk halmaza, azaz minden $\mathbf{a} \in Q^f$ vektorra létezik egy $S(\mathbf{a}) = \{S_{1,1}, \dots, S_{M,K}\}$ alokáció, ahol $|S_{m,k}| = a_{(m-1)K+k}$, és az alokáció teljesíti a megvalósíthatósági feltételeket (lásd (4) és (5)).

4.2.2. Hatékony alokáció

Egy alokáció $\Theta(\mathbf{a})$ hatékonyságát definiáljuk a következő módon:

$$\Theta(\mathbf{a}) = \sum_{i=1}^I \Theta_i(a_i), \quad (13)$$

ahol $\Theta_i(a_i)$ az adott alokáció (piaci) értéke az i . játékos számára. Az $\tilde{\mathbf{a}}$ alokáció optimális, ha ez a leghatékonyabb megvalósítható alokáció, azaz:

$$\tilde{\mathbf{a}} = (\tilde{a}_1, \dots, \tilde{a}_I) = \arg \max_{\mathbf{a} \in Q^f} \Theta(\mathbf{a}). \quad (14)$$

(12) és (14) alapján a javasolt alokáció akkor optimális, ha a játékosok licitjei $p_i(q) = \theta_i(q)$ alakúak. Ezt egy olyan árazási szabállyal érhetjük el, amely arra „készteti” a játékosokat, hogy megmondják a licitjükben a kívánt spektrumblokk tényleges értékét. Ezt a másodáras aukciók elvét használva érhetjük el.

4.3. Másodáras árazási szabály

A másodáras aukciók hátterében az a gondolat áll, hogy a játékos az általa fizetett árat, ami veszteséget a piacra lépésével a többieknek okoz. Belátható, hogy ezt az elvet követve a játékosok számára az az optimális, ha az igazat mondják, azaz a licitjükben az árat akkora állítják, amennyit az igényelt spektrum blokk ér számukra [10]. Ezért a továbbiakban azt feltételezzük, hogy az egyes játékosok licitjei $(q_i, \theta_i(q_i))$ alakúak.

Az eredeti licithalmaz-leíróban az i . játékos licitjeit törölve az „ellenfél-leíró” kapjuk:

$$B^{(-i)} = (B_1, \dots, B_{i-1}, 0, B_{i+1}, \dots, B_I). \quad (15)$$

Hasonlóan (14)-hez, a $B^{(-i)}$ leíróra is meghatározhatjuk az optimális alokációt:

$$\tilde{\mathbf{a}}^{(-i)} = (\tilde{a}_1^{(-i)}, \dots, \tilde{a}_I^{(-i)}) = \arg \max_{\mathbf{a}^{(-i)} \in Q^{f(-i)}} \Theta(\mathbf{a}^{(-i)}) \quad (16)$$

ahol

$$\mathbf{a}^{(-i)} = A(B^{(-i)}). \quad (17)$$

Ahogy már említettük, az i . játékos költsége:

$$c_i(A(B)) = \Theta_i(\tilde{a}_i) - [\Theta(\tilde{\mathbf{a}}) - \Theta(\tilde{\mathbf{a}}^{(-i)})] \quad (18)$$

$$= \sum_{\substack{j=1 \\ j \neq i}}^I [\Theta_j(\tilde{a}_j^{(-i)}) - \Theta_j(\tilde{a}_j)]. \quad (19)$$

Azaz az i . játékosnak pontosan annyit kell fizetni, amennyit azoknak a játékosoknak ért volna az adott spektrumszelet, akiket kizárt a jelenlétével.

5. Minimális interferenciájú alokáció

A 3. fejezetben ismertetett szabályok alapján meghatározhatóak a kiosztott spektrumszeletek ($\tilde{a}_i$) és a hozzájuk tartozó árak (c_i) minden egyes szolgáltatóra, minden régióban, úgy, hogy a hozzájuk tartozó $S = \{S_{1,1}, \dots, S_{M,K}\}$ alokáció megvalósítható. Azonban számos különböző megvalósítható alokáció létezik, amelyek teljesítik a megvalósíthatósági feltételeket ((4) és (5)), és ugyanaz a maximális hatékonyság tartozik hozzájuk. Az elrendezésekben fellépő összes interferencia azonban nagyon különböző lehet.

Célunk az optimális alokációs vektor

$$\tilde{\mathbf{a}} \equiv (|S_{1,1}|, \dots, |S_{1,K}|, \dots, |S_{M,1}|, \dots, |S_{M,K}|) \quad (20)$$

ismeretében a minimális interferenciájú spektrumallokáció megtalálása, azaz azt az $\tilde{S}(\tilde{\mathbf{a}})$ spektrumallokációt keressük, amelyik maximalizálja a $\xi(S)$ kihasználhatóságot, ahol

$$\xi(S) = \left(\sum_{m,k} |S_{m,k}| \right)^{-1} \sum_{m,k} \xi(S_{m,k}) |S_{m,k}|, \quad (21)$$

$$\text{így} \quad \tilde{S} = \arg \max_{\{S_{1,1}, \dots, S_{M,K}\}} \left\{ \sum_{m,k} |S_{m,k}| \xi(S_{m,k}) \right\}. \quad (22)$$

Ennek a meghatározásához egy szimulált lehűtésen alapuló heurisztikát használtunk.

5.1. Szimulált lehűtés

A szimulált lehűtés széles körben alkalmazott heurisztikus megközelítés globális optimalizációs problémák megoldásához [15]. A szimulált lehűtés esetében a keresési tér minden $\mathbf{s}$ pontját egy rendszer állapotaként értelmezzük és az adott pontokhoz tartozó $E(\mathbf{s})$ energiaértékek által meghatározott függvény minimalizálása az algoritmus célja.

Esetünkben az állapotvektort az alábbi formában definiálhatjuk:

$$\mathbf{s} = (\tilde{s}_{1,1}, \dots, \tilde{s}_{1,K}, \dots, \tilde{s}_{M,1}, \dots, \tilde{s}_{M,K}). \quad (23)$$

Észrevehetjük, hogy egy S spektrumallokáció egyértelműen megadható az $\mathbf{s}$ állapotvektorral és az $\tilde{\mathbf{a}}$ allokációs vektorral, mivel

$$S = \{(\tilde{s}_{m,k}, \hat{s}_{m,k})\} \text{ ahol } \hat{s}_{m,k} = \tilde{s}_{m,k} + \tilde{a}_{(m-1)K+k}$$

A rendszer energiája az $\mathbf{s}$ állapotban:

$$E(\mathbf{s}) = -\xi(S) + P_f, \quad (24)$$

ahol a P_f büntető függvény értéke

$$P_f = \sum_{m=1}^M \sum_{k=1}^K \left(I_{\{\xi(S_{m,k}) < \beta_m\}} + I_{\{\min_{\lambda \in S_{m,k}} \xi_{S_{m,k}}(\lambda) < \alpha_m\}} \right). \quad (25)$$

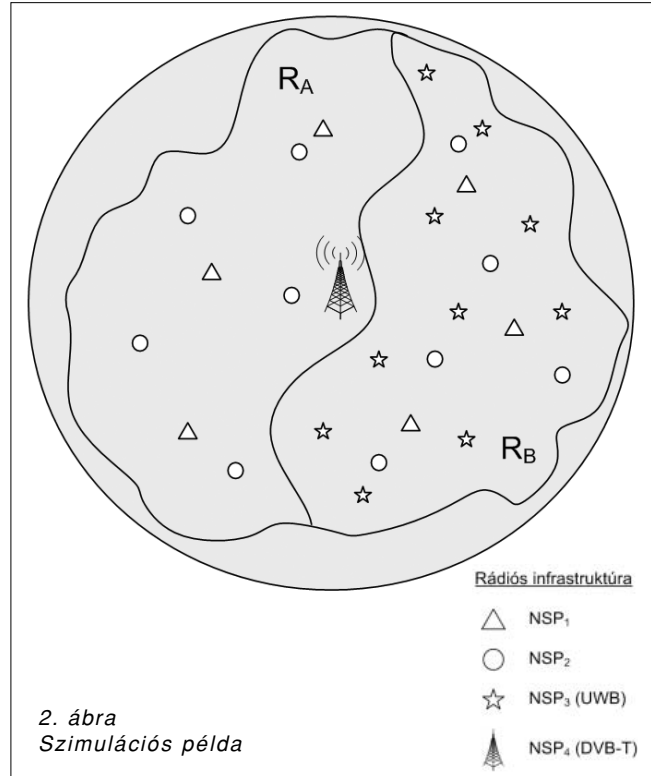
A büntetőfüggvény alakja olyan, hogy értéke 0, ha a megvalósíthatósági feltételek teljesülnek minden spektrumblokkra, egyébként pedig legalább 1. Ez azt eredményezi, hogy az E energiafüggvény mindig pozitív, ha az allokáció nem megvalósítható, és negatív, ha megvalósítható. Az energiafüggvényt minimalizálva pedig a legnagyobb kihasználhatóságú allokációt kapjuk meg.

Az optimalizálás során minden lépésben megnézzük az aktuális $\mathbf{s}$ állapot néhány $\mathbf{s}'$ szomszédját és véletlenszerűen döntünk arról, hogy az aktuális $e = E(\mathbf{s})$ energiájú állapotban maradunk-e, vagy átmegyünk az $e' = E(\mathbf{s}')$ energiájú állapotba.

A helyi minimumok elkerülése érdekében a P átlépési valószínűséget úgy kell megválasztani, hogy értéke ne legyen nulla $e' > e$ esetén sem, azaz a rendszer képes legyen „felmászni” az új állapotba akkor is, ha az rosszabb, mint a jelenlegi. Ugyanakkor az átlépési valószínűséget úgy kell meghatározni, hogy a futás során egyre inkább előnyben részesítsük az alacsonyabb energiájú állapotokba való átmenetet. Ezért az átlépési valószínűség függvénye nemcsak az aktuális és a következő állapothoz tartozó energiaértékektől függ, hanem egy időben változó T paramétertől (hőmérséklet) is. Az átlépési valószínűség-függvény széles körben használt alakja:

$$P(e, e', T) = e^{-\frac{e' - e}{T}}. \quad (26)$$

A T paraméter időbeli változása határozza meg azt, hogy milyen mértékű „felfelé” mozgás engedélyezett az adott időpontban. A szimulációk során a széles körben használt $T_{k+1} = \alpha \cdot T_k$ hőmérséklet-csökkentő függvényt használtuk, $T_0 = 1$ és $\alpha = 0.98$ kezdőértékekkel.



2. ábra
Szimulációs példa

6. Példa

Nézzünk egy egyszerű példát mindössze két régióval, ahogyan az a 2. ábrán látható. Két szolgáltató mindkét régióban jelen van és versenyez a rendelkezésre álló erőforrásért. Az A régióban jelen van egy DVB-T szolgáltató is, amely mindkét régiót lefedi. Továbbá a B régióban működik egy UWB jellegű szolgáltató, amely gyakorlatilag nem okoz interferenciát a többieknek és tolerálja az őt ért interferenciát.

A DSA modell jellemezhető az η (1. táblázat) és ε (2. táblázat) mátrixokkal, továbbá az α és β (3. táblázat) vektorokkal.

η	NSP-1	NSP-2	DVB-T	UWB
NSP-1	0	0.5	0.6	0.05
NSP-2	0.4	0	0.7	0.1
DVB-T	0.5	0.6	0	0.05
UWB	0.05	0.05	0.05	0

1. táblázat: Rádiotechnológiai csatolási tényezők

$\varepsilon_{AB}^{(i)}$	NSP-1	NSP-2	DVB-T	UWB
	0.2	0.3	1	0

2. táblázat: Földrajzi csatolási tényezők

	NSP-1	NSP-2	DVB-T	UWB
α	0.5	0.7	0.95	0.5
β	0.9	0.95	0.975	0.8

3. táblázat: Interferencia tolerancia paraméterek

Az η mátrix a rádiótechnológiai csatolási tényezők értékét tartalmazza, ahol az m . és az n . szolgáltató által használt rádiós technológiák egymásra gyakorolt zavaró hatását $\eta_{m,n}$ írja le. Minél kisebb ez az érték, annál kedvezőbb a helyzet az interferencia szempontjából. Az η mátrix elemeit megnézve láthatjuk, hogy az UWB szolgáltató által használt technológia gyakorlatilag nem zavarja a többi szolgáltatót és azt sem zavarják más szolgáltatók ($\eta < 1$).

A 2. táblázat a földrajzi csatolási tényezőket tartalmazza. Itt $\varepsilon_{i,k}^{(m)}$ azt a „zajt” jellemzi, amit az R_k régióban működő m . operátor okoz az R_i régióban. Ha megnézzük $\varepsilon^{(3)}$ értékét (DVB-T), egy erős csatolást ($\varepsilon=1$) láthatunk a két régió között, ezzel biztosítva, hogy a DVB-T szolgáltató lefedje mindkét régiót. Ezzel ellentétben az UWB szolgáltató egyáltalán nem zavarja a szomszéd régiót ($\varepsilon^{(4)}=0$).

A 3. táblázatban az átlagos (β) és a minimális (α) kihasználhatóságra vonatkozó paramétereket láthatjuk. Minél kisebbek ezek az értékek, a szolgáltató annál toleránsabb az interferenciával szemben. Láthatjuk, hogy a DVB-T szolgáltató tolerancia szintje alacsony ($\alpha \approx 1$ és $\beta \approx 1$), ami azt jelenti, hogy a spektrumblokkot kizárólagos használatra igényli.

A szolgáltatók licitjeit a 4. táblázatban láthatjuk, az 5. táblázat pedig az optimális allokációt és a szolgáltatók költségeit tartalmazza.

	R_A	R_B
NSP-1	$B_1 = \{(25, 75), (30, 85)\}$	$B_2 = \{(20, 40), (30, 50)\}$
NSP-2	$B_3 = \{(10, 30), (25, 45)\}$	$B_4 = \{(15, 45), (20, 55)\}$
DVB-T	$B_5 = \{(5, 25), (10, 35)\}$	$B_6 = \emptyset$
UWB	$B_7 = \emptyset$	$B_8 = \{(20, 20)\}$

4. táblázat: Liciték

	R_A		R_B	
	(q, c)	c/q	(q, c)	c/q
NSP-1	(25, 15)	0.6	(20, 10)	0.5
NSP-2	(10, 10)	1	(15, 10)	0.6
DVB-T	(5, 20)	4	-	-
UWB	-	-	(20, 0)	0

5. táblázat: Optimális allokáció és költségek

Az 5. táblázat c/q oszlopa az egyes szolgáltatók által fizetendő egységárat tartalmazza. Az eredmények azt mutatják, hogy a DVB-T szolgáltató fizeti a legtöbbet, mivel a toleranciaszintje alacsony, és mindkét régiót lefedi.

Ezzel ellentétben az UWB szolgáltató nem fizet semmit (ezt értelmezhetjük úgy, hogy csak a spektrum alapárát kell megfizetnie), mivel gyakorlatilag nem fejt ki zavaró hatást a többi szolgáltatóra. Láthatjuk továbbá, hogy az 1. szolgáltató földrajzi csatolási tényezője kisebb, toleranciája pedig jobb, mint a 2. szolgáltatóé, ennek megfelelően az általa fizetett egységár is kevesebb lesz.

7. Összefoglalás

Javaslatot tettünk egy árazási algoritmusra, mely segítségével hatékonyan szétoszthatók a spektrumhasználati jogok az egymással versengő szolgáltatók között. A javasolt megoldás egy egylépcsős többlicites aukció, amely az általunk javasolt [13] DSA keretrendszert használja alapként.

Ez a DSA keretrendszer alkalmas az interferencia és a különböző szolgáltatók interferencia tűrésének rugalmas modellezésére. A javasolt árazási mechanizmus pedig biztosítja, hogy azon szolgáltatók, amelyek nem tolerálják az interferenciát, vagy nagy zavaró hatást fejtenek ki a szomszédokra, többet fizessenek az erőforrásért.

Irodalom

- [1] „Study on conditions and options in introducing secondary trading of radio spectrum in the European Community”, – tanulmány az Európai Bizottság számára, készítők: Analysys, DotEcon és Hogan and Hartson, LLP, 2004. május (a jelentés kivonata).
- [2] „IST-DRiVE project”, <http://www.ist-drive.org>
- [3] „IST OverDRiVE project”, <http://www.ist-overdrive.org>
- [4] P. Leaves, J. Hushke, R. Tafazolli, „A summary of dynamic spectrum allocation results from DRiVE” – Proc., IST Mobile and Wireless Telecommunications Summit, Thessaloniki, Görögország, 2002. június 16-19., pp.245–250.
- [5] M. Buddhikot, P. Kolodzy, S. Miller, K. Ryan, J. Evans, „DIMSUNet: New directions in wireless networking using coordinated dynamic spectrum access”, Position Paper in IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks (IEEE WoWMoM 2005), Taormina/Giardini Naxos, Olaszország, 2005. június.
- [6] M. Buddhikot, K. Ryan, „Spectrum management in coordinated dynamic spectrum access based cellular networks”, – Proc., 1st IEEE International Symposium on New Directions in Dynamic Spectrum Access Networks, Baltimore, MD, 2005. november 8-11.
- [7] Kovács L., Tapolcai J., Vidács A., „Spatio-temporal dynamic spectrum allocation with interference handling”, – Proc., IEEE Int. Conf. on Communications (ICC 2007), Glasgow, Skócia, UK, 2007. június 24-28.
- [8] V. Rodriguez, K. Moessner, R. Tafazolli, „Market-driven dynamic spectrum allocation: Optimal end-user pricing and admission control for CDMA”, – Proc., 14th European Inform. Society Technologies (IST) Mobile and Wireless Communications Summit, Drezda, Németország, 2005. június.

- [9] V. Rodriguez, K. Moessner, R. Tafazolli,
„Auction driven dynamic spectrum allocation over
space and time: DVB-T and multi-rate, multi-class
CDMA over a two-island geography”,
– Proc., 15th European Inform. Society Technologies
(IST) Mobile and Wireless Communications Summit,
Myconos, Görögország, 2006. június 4-8.
- [10] A. Lazar, N. Semret,
„Design and analysis of the progressive second
price auction for network bandwidth sharing”,
Telecommunication Systems – Special issue on
Network Economics, 1999.
- [11] Tuffin,
„Revisited progressive second price auction
for charging telecommunication networks”,
Telecommunication Systems,
20(3):255–263, 2002.
- [12] P. Maille, B. Tuffin,
„Multi-bid auctions for bandwidth allocation
in communication networks”,
– Proc., 23rd Conf. of the IEEE Comm. Society
(INFOCOM 2004),
Hong Kong, 2004. március 7-11.
- [13] Kovács L., Vidács A.,
„Interference-tolerant spatio-temporal dynamic
spectrum allocation,”
– Proc., 2nd IEEE International Symposium on New
Directions in Dynamic Spectrum Access Networks,
Dublin, Írország, 2007, április.
- [14] Kovács L., Vidács A.,
„Spatio-temporal spectrum management model
for dynamic spectrum access networks,”
– Proc., 1st International Workshop on Technology
and Policy for Accessing Spectrum (TAPAS),
Boston, USA, 2006.
- [15] S. Kirkpatrick, C. D. Gelatt, M. P. Vecchi,
„Optimization by simulated annealing,”
Science, Nr.4598, 13 May 1983.,
pp.671–680.

Hírek

A Microsoft Magyarország Fejlesztési Platform Üzlet-ága és a BME Automatizálási és Alkalmazott Informatikai Tanszékén működő MSDN Kompetencia Központ idén először rendezte meg **„Win the Web” elnevezésű programozói versenyét**, azzal a céllal, hogy a magas szintű webfejlesztői ismereteket az egyetemen és iparági közegben is népszerűsítse, továbbá megteremtse a lehetőséget egy hazai webfejlesztői közösség kialakulására. A verseny egy 24 órás problémamegoldás keretében kizárólag webes fejlesztésekre koncentrált, ahol a háromfős csapatoknak Microsoft technológiával kellett megoldani egy nagy- és három kisfeladatot. A versenyen 21 csapat indult, amelyek közül az Evosoft Giczi György, Lipták Olivér és Virág Imre által képviselt csapata vitte el a trófeát. Második helyezést ért el az mEKF (Miskolci Egyetem és Eszterházy Károly főiskola), amelyet Túrócy Attila és Csirmaz Péter képviselt. A harmadik helyen a BME.NET csapata végzett, tagjai: Waszlavik Márton, Ritzinger Péter és Melkó Tamás voltak.

A **Magyar Telekom és a Cisco** bejelentette, hogy a Magyar Telekom elsőként építi ki új generációs szolgáltatói IP-hálózatot Magyarországon. A vállalat az IP-gerinchálózatának megújítása keretében két Cisco CRS-1 típusú terabitrouter helyez üzembe, amelyek lehetővé teszik a hálózat megbízhatóságának jelentős növelését és a jövőben jelentkező kapacitásigények kiszolgálását. A Cisco IP NGN (Cisco IP Next Generation Network, új generációs IP-hálózat) architektúrájának bevezetésével lehetőség nyílik az interneten túlmutató alkalmazások kiszolgálására és a távközlési infrastruktúra megújítására, az alkalmazások és előfizetők egyedi kezelésén alapuló szolgáltatások fejlesztésére és bevezetésére.

A **Novell** bejelentette Novell SUSE Linux Enterprise 10 magas rendelkezésre állású tárolóeszköz-felügyeleti megoldásának elérhetőségét az SAP NetWeaver és a mySAP Business Suite rendszerekhez. A Novell és az SAP LinuxLab által közösen kifejlesztett nyílt forráskódú elemekből felépülő terméke az adatok védelme mellett biztosítja a vállalatok zökkenőmentes működését, ezáltal csökkenti a költségeket és leegyszerűsíti az adattároló-kezelést. A SUSE Linux Enterprise magas rendelkezésre állású tárolóeszköz-felügyeleti megoldása egy integrált clusterszoftvert is tartalmaz, amely folyamatosan figyeli a vállalati rendszer különböző részeinek, többek közt az SAP-alkalmazásoknak az állapotát. Esetleges probléma, például hardvermeghibásodás esetén elindítja a feladatok átvitelét a redundáns kiszolgálókra. Emellett koordinálja a clusterszintű tevékenységeket a megosztott lemezalapú clusterok tároló-erőforrásainak megbízható felügyelete érdekében.

Near optimal channel equalizer based on minimum bit error rate strategy

Keywords: channel equalization, adaptive algorithms, Monte-Carlo simulations, statistical sampling

In this paper novel channel equalizer algorithms are introduced for wireless communication systems to combat channel distortions resulting from multipath propagation yielding severe performance degradation. This paper aims at developing small complexity channel equalizer algorithms by directly minimizing the bit error rate (BER) instead of minimizing the mean square error (MSE) or the peak distortion (PD). Unfortunately, the direct minimization of BER with respect to the equalizer coefficients is of exponential complexity, therefore we develop algorithms based on statistical sampling techniques minimizing the bit error rate and guarantee better performance than the traditional Zero Forcing (ZF) or Minimum Mean Square Error (MMSE) algorithms. The performance of the new algorithms are tested by extensive simulations on standard mobile channels.

Joint optimization of layers in UMTS-based video-transmission

Keywords: joint source channel coding and decoding, cross-layer design, IPv6, UMTS Radio Access Network

In this document we introduce a software demonstration platform. This platform has been developed for performance evaluation of network aware joint optimization of wireless video transmission. The platform has been developed in the framework of IST-PHOENIX project and is based on the system optimization model of the project. The essential component, the wireless network module is replaced by the standard compliant, highly detailed UTRAN simulation which was developed at Budapest University of Technology and Economics (BME). The paper gives a brief overview of the project simulation chain, the main specification of the UTRAN network and the combination of both systems. In our aspect, the role of the UTRAN network in optimization is very important along with its configuration and control. Some simulation results are also included.

Vertical handover: interworking of heterogenous networks

Keywords: mobile networks, mobility management, network-independent services, seamless handoff

The interoperation of wired, wireless and mobile networks can establish the fundamentals of new telecommunication services. The application of vertical handover enables the user to move freely between different access networks. Service providers can offer integrated, network-independent services using the network with optimal bandwidth or cost. In this article we observe the properties and both the potential capabilities and drawbacks of vertical handover.

A call forwarding procedure in 3GPP-LTE next generation networks

Keywords: Long Term Evolution, hard handover, performance evaluation, packet forwarding and reordering

The specification of the Long Term Evolution (LTE) of 3G systems is currently ongoing in 3GPP with a target date of ready specification at the end of 2007. The evolved Radio Access Network (RAN) involves a new radio interface based on OFDM technology and a radically different RAN architecture, where radio functionality is distributed into the base stations. In this paper we give an overview of the LTE intra-access handover procedure and evaluate its performance focusing on the user perceived performance aspects of it.

TFRC-based selective retransmission algorithm

Keywords: multimedia, adaptive retransmission, Ns2, video streaming

In our work we present a new selective retransmission algorithm, which retransmits the lost or damaged packets of the MPEG video stream according to the current state of the network. The packets will be retransmitted only if the additional load will not overload the network. With our method the free capacity of the network will be used to increase the quality of the multimedia stream. The free capacity is defined by the source sending rate and parameters of the DCCP congestion control algorithm (TCP Friendly Rate Control). Our method will fill up the free capacity of the link according to the priority of the packets. We used Ns2 (Network Simulator) to prove the efficiency of the selective retransmission algorithm based on congestion control and packet differentiation. The test results show that this method has significant effect on the improvement of the quality.

Wireless and mobile broadband access technologies from a service provider's point of view

Keywords: UMTS/HSPA, WiMAX, investment needs

The evolution of radio access technologies has reached an important milestone due to increased and high data rate capability. There are two main streams of radio access technologies the IP based WLAN and WiMAX, and the evolution of traditional mobile technologies like (E)GPRS and UMTS/HSPA. Besides the similarities there are also some notable differences which are originated in the distinctness between technologies and their applicability in various environments. The results and the conclusions of this paper are based on the experiences of Pannon and Telenor.

Primary and secondary lightning protection of Mobile Services Switching Centers

Keywords: electromagnetic protection, lightning detection

The Core Network Sites (e.g. Mobile Switching Centres) are very important parts of the mobile telephone network. They are specially endangered by lightning strokes, because of their tall transmitting tower. They are full of electronic equipment, which are very sensitive against the overvoltages caused by lightning. The authors give a review concerning both the general aspects of the lightning protection of such Core Network Sites (CNS) and typical problems of the practical operation. Basically dangerous situations, theoretical principles of the protection and possible typical faults of the installation are discussed in the paper.

Dynamic spectrum allocation: modelling and pricing

Keywords: dynamic spectrum allocation, wireless networks, auction, pricing, interference-tolerant

The radio spectrum is a scarce, valuable and thus expensive resource. An efficient use of frequencies is primordial, but existing management systems fail to achieve this. The possibility of on-demand spectrum allocation with Dynamic Spectrum Access methods, together with the pressure of a liberalized spectrum market would lead to a much more efficient spectrum usage. In this paper we propose a centralized spectrum auction method that fits to our proposed DSA framework to distribute spectrum for fixed infrastructure wireless networks according to the needs of providers. Since interference tolerance in a DSA scenario is much rewarded, our pricing scheme charges providers who do not tolerate others and interfere to a larger extent than necessary.