

A harmadik generációs mobil gerinchálózatok hívásszintű teljesítményelemzése*

Sándor Rácz¹, Miklós Telek², Gábor Fodor³

¹ Ericsson Research Hungary, Budapest, Sandor.Racz@eth.ericsson.se,

² Budapesti Műszaki Egyetem, Híradástechnikai Tanszék, Budapest, telek@hit.bme.hu

³Ericsson Research, Kista, Sweden, Gabor.Fodor@era-t.ericsson.se

Kivonat

Ebben a cikkben bemutatjuk az UMTS gerinchálózatok hívásszintű modelljét, ahol a négy UMTS szolgáltatási osztály valamelyikéhez tartozó hívások véletlenszerűen érkeznek be. A beérkező hívások számára garantált szolgáltatási szint a hívás szolgáltatási osztályától, a szükséges maximális és minimális sávszélességtől és a beérkezéskor rendelkezésre álló hálózati erőforrásoktól függ. Egy adott sávszélességű összeköttetést Markov-modellel modelleztük, hogy GoS (blokkolási valószínűség) és QoS (sávszélesség) jellemzőket határozassunk meg két ésszerű és műszakilag használható sávszélesség-megosztási módszer segítségével. Arra a következtetésre jutottunk, hogy e két módszer egyike képes garantálni a megfelelő GoS/QoS értékeket nagy mennyiségű, vegyes forgalom esetén. Megmutatjuk, hogy az eredmények felhasználhatók a teljesen IP-alapú, illetve az MPLS-alapú új UMTS architektúráknál.

Kulcsszavak: harmadik generációs mobilhálózatok, UMTS hálózatok, sávszélesség-megosztási célok, különböző sávszélességű veszteségi modellek, blokkolási valószínűség, Markov-modellek.

1 Bevezetés

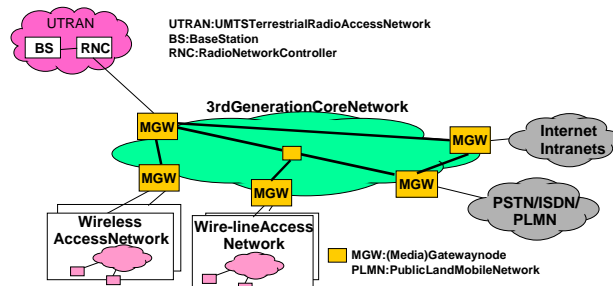
Architektúra szempontjából a harmadik generációs mobilhálózatok egyik kulcsfontosságú jellemzője a hozzáférési rész és a gerinchálózat elválasztása. Ez az elválasztás hozzájárul a különböző hozzáférési technológiák és a gerinchálózat által biztosított új szolgáltatások folyamatos fejlődéséhez [3], [4]. Ezért az UMTS/IMT-2000 hálózatok magja olyan többszolgáltatásos hálózat, amely GoS (blokkolási valószínűség) és QoS (sávszélesség) garanciát kínál a szolgáltatási osztályok következő négy típusa számára [16]:

- A társalgási osztály (*conversational class*) jó minőségű hozzáférést biztosít számos különböző szolgáltatáshoz, köztük nagy bitsebességű szolgáltatásokhoz is. Ez az osztály

*Ez a cikk a [9, 10, 11] publikációk eredményeinek magyar nyelvű összefoglalója.

olyan igényes felhasználók számára hasznos, akik az ATM CBR osztályához hasonló sávszélesség-garanciát kívánnak kapni.

- Az áramlási osztályt (*streaming class*) nagy sávszélességű, változó bitsebességű szolgáltatások, például közepes vagy igen jó minőségű video- vagy konferenciabeszélgetési szolgáltatások nyújtására fejlesztették ki. Az ehhez az osztályhoz tartozó hívások egyik közös vonása, hogy tartási idejük független a rendszerben tartózkodási idejük alatt kapott aktuális sávszélességtől.
- Az interaktív osztály (*interactive class*) kevésbé igényes, tipikusan a jelenlegi “best effort” IP hálózatok által támogatott szolgáltatásokat nyújt, például fájlátvitel, WEB-böngészés, illetve telefonos alkalmazások. Az UMTS hálózatoktól azt várják, hogy valamiféle sávszélesség garanciát nyújtsanak még az ilyen típusú szolgáltatások esetén is. Az interaktív osztályú hívások tartási ideje függ a sávszélességtől (kétszeres sávszélesség mellett például egy fájl továbbítása feleannyi időbe telik).
- A háttérosztály (*background class*) “best effort” típusú kiszolgálást kap, azaz a háttérhívások a magasabb prioritású szolgáltatási osztályokba tartozó hívások után megmaradó sávszélességet kapják meg. Ide tartozik az elektronikus levelezés és a gyengébb minőségű fájlátvitel. A tartási időt illetően ez az osztály hasonló az interaktív osztályhoz.



1. ábra: UMTS gerinchálózat különböző hozzáférési hálózatokkal

Az 1. ábra ismerteti az UMTS gerinchálózat néhány jellemzőjét, nevezetesen a különböző hozzáférési hálózatok elválasztását a gerinchálózattól átjáró-csomópontokkal, amelyek a hívások beléptetésének felügyeletét végzik. Ami az alkalmazandó technológiákat illeti, a gerinchálózat valószínűleg valamilyen kapcsolatorientált erőforrás-tartalékolási mechanizmussal kiegészített gyors csomagkapcsolási technikákon alapul majd, mint például ATM vagy IP/MPLS [14, 17]. Az UMTS hálózatok méretezése és teljesítményelemzése megköveteli, hogy a blokkolási valószínűség és a sávszélesség közti kompromisszumot a dinamikusan változó környezetet (hívások véletlenszerűen érkezése és távozása) figyelembe vevő hívásszintű modellek alapján határozzuk meg. Ahogyan a 2. fejezetben részletezzük, az áramkör-kapcsolt [12] és az ATM [5] hálózatok esetén széles körben rendelkezésre álló hívásszintű modellek nem alkalmazhatók közvetlenül az UMTS környezetben, mivel nem veszik figyelembe az UMTS-specifikus szolgáltatási osztályok meghatározását. Jóllehet számos tanulmány foglalkozik a különböző hozzáférési hálózatokkal [13] - [15], ebben az írásban a gerinchálózatra koncentrálunk, ahol a fő cél a

sávszélesség-megosztási lehetőségek elemzése annak érdekében, hogy az UMTS szolgáltatási osztályok GoS/QoS követelményei teljesüljenek, és a sávszélesség-kihasználás nagy maradjon.

E tanulmány két dologgal is hozzájárul e cél eléréséhez. Egyrészt kiterjesztjük a klasszikus különböző sávszélességű szolgáltatás modelleket, amelyek így tartalmazni fogják az UMTS szolgáltatási osztályokat is. Másrészt az első három osztály sávszélességének biztosítása mellett, a háttérosztály teljesítményére koncentrálva, egyszerű sávszélesség-megosztási módszert javasolunk, amellyel nagyobb teljesítmény érhető el, mint a sávszélességre és a blokkolási valószínűségre vonatkozó tökéletes megosztással.

Tanulmányunk a következőképpen épül fel. A második fejezet a hálózati modellt ismerteti. A harmadik fejezet a teljesítmény jellemzőit és a hozzájuk tartozó sávszélesség-korlátokat határozza meg. A negyedik fejezet az UMTS gerinchálózat egyszerűen megvalósítható sávszélesség-megosztási módszereit tartalmazza. Az ötödik fejezetben a különböző vegyes forgalmakat feltételező számszerű eredményeket ismertetjük, a hatodik fejezetben pedig összefoglaljuk az eredményeket.

2 Az UMTS gerinchálózat modellezése

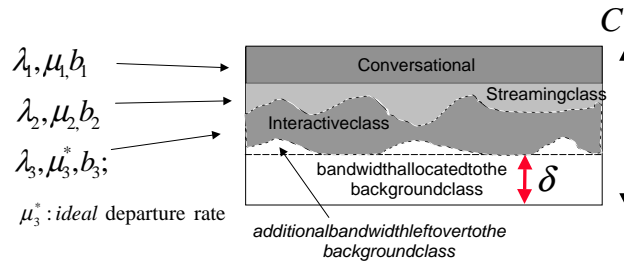
Mivel a gerinchálózat átviteli összeköttetésenként ellenőrzi a belépési feltételek teljesülését (CAC), ezért egyetlen átviteli összeköttetés Markov-modelljét készítettük el. Ezen az absztrakciós szinten modellünk egyaránt alkalmas ATM kapcsolat vagy MPLS címkekapcsolási útvonal modellezésére. Az UMTS szolgáltatási osztályok ismertetéséből világossá válik, hogy az UMTS szolgáltatási osztályok forgalmi viselkedése megfeleltethető például bizonyos ATM forgalmi osztályok viselkedésének. Az általunk alkalmazott modell ezért megegyező forgalmi viselkedés esetén nem csak az UMTS gerinchálózatok elemzésére alkalmazható. Ebben a cikkben elvégzett vizsgálat során modellezzük a sávszélesség garanciával rendelkező (azaz társalgási, áramlási és interaktív) szolgáltatási osztályokat, és kiszámítjuk a háttérosztály számára fennmaradó sávszélesség (közéértékét és eloszlását).

A tárgyalt rendszer egy C sávszélességű átviteli összeköttetés. C -t egész számnak tekintünk valamilyen használatos sávszélesség-egységben, mondjuk $Mbps$ -ban. A vizsgált összeköttetésre érkező hívások az alábbi három forgalmi osztály egyikéhez tartoznak:

- A társalgási szolgáltatásba tartozó hívásokat b_1 sávszélesség követelmény, λ_1 beérkezési intenzitás és μ_1 távozási intenzitás jellemzi;
- Az áramlási osztályra b_2 maximális sávszélesség, b_2^{min} minimális sávszélesség követelmény, λ_2 beérkezési intenzitás és μ_2 távozási intenzitás jellemző. Habár az áramlási hívások által elfoglalt sávszélesség a forgalomtól függően változhat, tényleges tartási idejüket nem befolyásolja a rendszerben tartózkodásuk teljes ideje alatt kapott sávszélesség. Ez az eset áll fenn például a videokodeknél, amely a sávszélesség romlásakor gyengíti a videokép minőségét, és így kisebb sávszélességet foglal, de a video program vetítési ideje nem változik.
- Az interaktív osztályú folyamokat b_3 maximális sávszélesség, b_3^{min} minimális sávszélesség követelmény, λ_3 érkezési intenzitás és μ_3 maximális sávszélesség melletti (ideális)

távozási intenzitás jellemzi. Az ideális távozási intenzitás csak a maximális sáv szélesség rendelkezésre állásánál tapasztalható. A valódi pillanatnyi távozási intenzitás a folyamatok sáv szélességétől függ.

A 2. osztályú (áramlási) és 3. osztályú (interaktív) folyam számára kiosztott tényleges sáv szélességet a rendszer egy adott állapotában b_2^r és b_3^r -al jelöljük, ahol b_2^r és b_3^r változik a különböző osztályú összeköttetések érkezésével és távozásával. Az áramlási és interaktív osztályok összenyomhatóságának határa $r_{min} := b_{min}/b$. Egy interaktív osztályú hívást úgy tekinthetünk, mint amely exponenciális eloszlású szolgáltatási követelményből sorsolt W továbbítandó adatmennyiséggel rendelkezik beérkezéskor. A továbbítandó adatmennyiség elosztása $G(x) = Pr(W \leq x) = 1 - e^{-\frac{b_3}{\mu_3}x}$, amely a hívás teljes ideje alatt b_3 maximális sáv szélesség rendelkezésre állása esetén exponenciális eloszlású kiszolgálási időt eredményez $1/\mu_3$ átlagértékkel. Mivel a vizsgált összeköttetés szabad sáv szélessége a szolgáltatásban részt vevő hívások pillanatnyi számától függően változik, az interaktív hívások számára rendelkezésre álló sáv szélesség a maximális sáv szélesség-követelmény értéke alá eshet, amikor is a hívás tartási ideje megnő.



2. ábra: Sáv szélesség megosztás a 4 UMTS szolgáltatási osztály között

Ahogy a 2. ábrán látható, a sáv szélesség garanciákkal rendelkező folyamatok első három típusa független Poisson folyamatok szerint érkeznek, a társalgási és az áramlási osztályba tartozó hívások tartási ideje exponenciális eloszlású, és az interaktív osztályba tartozó hívások (PH eloszlású) tartási idejét a vizsgált összeköttetés Markov-modelljével határozzuk meg.

A továbbiakban megvizsgáljuk két olyan egyszerű sáv szélesség-megosztási stratégia hatását, amelyek GoS/QoS korlátokat biztosítanak az első három szolgáltatási osztály számára, illetve sáv szélesség garanciát az alacsonyabb prioritású háttér osztály számára.

3 Teljesítményjellemzők és szolgáltatásminőségi korlátok

A klasszikus többsáv szélességű modellekkel [12] ellentétben, modellünk lehetővé teszi az elfoglalt sáv szélesség ingadozását, és így számítható teljesítményjellemzők a blokkolási valószínűség és az osztályokra vonatkozó sáv szélesség értéke is. A sáv szélesség-megosztás hatékonyságának megítéléséhez az osztályokra vonatkozó sáv szélesség értékek és a hozzájuk tartozó korlátok pontos meghatározására van szükség.

3.1 Sáv szélesség definíciója

A társalgási hívások sáv szélessége egyszerűen a kiosztott sáv szélesség, amely állandó érték (b_1). Látható, hogy az interaktív osztály folyamainak tartási ideje nemcsak az átvinni kívánt adatok mennyiségétől függ (amely véletlen változó), hanem a tartási idejük alatt kapott sáv szélességtől is. Hasonlóképpen, az áramlási osztályon keresztül átvitt adatmennyiség is a kapott sáv szélességtől függ. Ennek a viselkedésnek az elemzése céljából a következő mennyiségeket definiáljuk:

- $\theta_2(t)$ és $\theta_3(t)$ jelöli az áramlási, illetőleg interaktív folyamatok t időpontban kapott pillanatnyi sáv szélességét. (Pl., ha n_1, n_2, n_3 darab társalgási, áramlási, illetőleg interaktív folyamat van jelen a rendszerben t időpontban, a pillanatnyi sáv szélesség az áramlási és az interaktív osztályban $\min(b_2, (C - n_1 b_1 - n_3 r_3 b_3)/n_2)$ és $\min(b_3, (C - n_1 b_1 - n_2 r_2 b_2)/n_3)$). Figyeljük meg, hogy $\theta_2(t)$ és $\theta_3(t)$ diszkrét véletlen változók.
- $\tilde{\theta}_t = \frac{1}{t} \int_0^t \theta_2(\tau) d\tau$ jelöli a t tartási idejű áramlási folyamat sáv szélességét.
- $\tilde{\theta} = \int_0^\infty \tilde{\theta}_\tau dF(\tau) = \mu_2 \int_0^\infty \tilde{\theta}_\tau e^{-\mu_2 \tau} d\tau$ jelöli az áramlási folyamat sáv szélességét, ahol $F(t)$ jelöli az exponenciális eloszlású tartási idő eloszlásfüggvényét.
- $T_x = \inf\{t \mid \int_0^t \theta_3(\tau) d\tau \geq x\}$ (véletlen változó) jelöli azt az időt, amennyi egy interaktív folyamat x adatmennyiségének átviteléhez szükséges.
- $\hat{\theta}_x = x/T_x$ jelöli az interaktív folyamat sáv szélességét x adategység átvitele alatt. θ_x folytonos véletlen változó.
- $\hat{\theta} = \int_0^\infty \hat{\theta}_x dG(x) = \mu_3/b_3 \int_0^\infty \hat{\theta}_x e^{-x \mu_3/b_3} dx$ (véletlen változó) jelöli az interaktív folyamat sáv szélességét, ahol az átvitt adatmennyiség μ_3/b_3 paraméterű exponenciális eloszlású.

A következőkben feltételezzük, hogy az alkalmazott hívás engedélyezés (CAC) olyan, hogy biztosított a társalgási, áramlási és interaktív forgalmi osztályokhoz rendelt maximálisan elfogadott blokkolási valószínűség (B_1^{max} , B_2^{max} illetve B_3^{max}) és az áramlási és interaktív hívások minimálisan elfogadott sáv szélessége ($\hat{\theta}^{min}$, $\hat{\theta}^{min}$) [9]. Mivel a háttérosztály esetén nincsen minimálisan garantált sáv szélesség, ezért ezen osztálynál a következő két sáv szélesség korlátot vesszük figyelembe.

3.2 Sáv szélesség korlátok háttér forgalomra

Ha n_1, n_2, n_3 társalgási, áramlási, illetőleg interaktív folyamatok vannak jelen a rendszerben t időpontban, a (legalacsonyabb prioritási szintű) háttérosztály számára biztosított pillanatnyi sáv szélesség a következő:

$$\theta_4(t) = \max(0, C - n_1 b_1 - n_2 b_2 - n_3 b_3)$$

A háttérosztály egyensúlyi állapotbeli sáv szélességét θ_4 -el jelöljük. (θ_4 diszkrét véletlen változó.)

A háttérosztály számára fennmaradó sáv szélesség jellemzésére a következő két sáv szélesség korlátot határozzuk meg:

- átlagos sávszélesség korlát:

$$E(\theta_4) > \theta_4^{min},$$

ami azt követeli meg, hogy a háttérosztály számára fennmaradó sávszélesség átlaga nagyobb legyen, mint θ_4^{min} ,

- sávszélesség küszöb korlát:

$$\Pr(\theta_4 > \theta_4^{min}) > \epsilon.$$

ami azt követeli meg, hogy a háttérosztály számára fennmaradó sávszélesség értéke ϵ valószínűséggel nagyobb legyen, mint θ_4^{min} ,

A két korlát között az a gyakorlati szempontból lényeges különbség, hogy az első úgy is teljesíthető, hogy nagy ($\gg \epsilon$) valószínűséggel nem áll rendelkezésre szabad kapacitás a háttérosztály számára, míg a második esetben ez csak kis (ϵ) valószínűséggel fordulhat elő.

4 Sávszélesség-megosztási politikák

Régóta tanulmányozzák a sávszélesség megosztását a kapcsolt távközlő és az ATM hálózatokban, de a legtöbb modell nem vesz tudomást a forgalmi osztályok - például az áramlási és interaktív osztályok - rugalmas természetéről. A hívásszintű rugalmas forgalmat modellező tanulmányok például Altman és társai [1], Andersen és társai [2], illetve Nunez és társai [6] művei, de ezek kizárólag két forgalmi osztályra korlátozódnak, ráadásul inkább a rugalmas forgalom momentumaira, mintsem elosztásának vizsgálatára koncentrálnak.

A fenti sávszélesség és korlátdefiníciók felállításával a következő sávszélesség-megosztási szabályokat határozhatjuk meg. Az alábbiakban $C - \delta$ jelöli a sávszélességnek azon részét, amely az első három szolgáltatási osztályhoz van rendelve, és δ jelöli a kizárólagosan a háttérosztály számára fenntartott sávszélességet. A numerikus példákban a következő két politikát elemezzük: $\delta = 0$ (1. módszer: teljes megosztás) és $\delta > 0$ (2. módszer: teljes partícionálás). Sávszélesség-megosztási módszereinket a műszaki megvalósíthatóság és - ahogyan számszerű eredményekből látni fogjuk - az a megérzés indokolja, hogy mivel a magas prioritású rugalmas osztályok hajlamosak elfoglalni az összes rendelkezésre álló sávszélességet, ezért a háttérosztály számára szükséges lehet fenntartott sávszélesség biztosítása, hogy elfogadható szolgáltatás minőséget nyújtsen különböző forgalmi szituációkban.

- Ha elegendő sávszélesség áll rendelkezésre az összes folyam számára, hogy megfeleljenek maximális sávszélességi követelményeiknek, akkor a 2. és 3. osztályú folyamatok b_2 , illetve b_3 sávszélességet foglalnak le.
- Ha a sávszélesség csökkentésre van szükség, azaz $n_1 \cdot b_1 + n_2 \cdot b_2 + n_3 \cdot b_3 > C - \delta$, akkor a folyamatok sávszélességét úgy kell csökkenteni, hogy $r_2 = r_3$, ahol $r_2 = b_2^r/b_2$ és $r_3 = b_3^r/b_3$, amíg mindkét osztálynál teljesül a minimális sávszélességkorlát, (azaz $b_2^{min}/b_2 \leq r_2 \leq 1$ és $b_3^{min}/b_3 \leq r_3 \leq 1$).
- Ha további sávszélesség csökkentésre van szükség, de a két osztály valamelyike új folyam érkezésekor már nem csökkentheti tovább sávszélességét, mert r_i már b_i^{min}/b_i , valamelyik

$i = 2$ vagy $i = 3$ osztályra, akkor a tovább csökkenthető szolgáltatási osztály összeköttetései egyenlő mértékben csökkentik a sávszélességüket mindaddig, amíg a minimális sávszélességük még biztosított.

Érdemes megemlíteni a fenti szabályokra (és a modellre) vonatkozó három alapvető feltételt. Először is feltételezzük, hogy mind az áramlási, mind az interaktív folyamatok *falánkak*, azaz mindig elfoglalják a kapcsolat maximálisan lehetséges sávszélességét. (A maximális sávszélességet b_2 , illetve b_3 , valamint a nem összenyomható folyamatok által (a kapcsolatkiosztási eljárástól függően) meghagyott sávszélesség határozza meg.) Másodszor feltételezzük, hogy az összes aktív áramlási és interaktív folyamat egyenlően részesül a rendelkezésre álló sávszélességből (azaz az r_i értékek egyenlőek), tehát az újonnan érkezett folyamat és a már zajló folyamatok is ugyanazt az r_i értéket kapják. Ha egy újonnan érkezett folyamat b_2^{min} , illetve b_3^{min} alá csökkentené a sávszélességet (azaz mind az áramlási, mind az interaktív osztályok a minimális sávszélesség értékre lettek csökkentve), akkor az új folyamat nem kerül be a rendszerbe, hanem blokkolódik és elveszik. Fontos, hogy az összes beérkező folyamat képes a szolgáltatásban részt vevő áramlási és interaktív folyamatok sávszélességének csökkentésére mindaddig, amíg a szükséges minimális sávszélesség minden osztályban rendelkezésre áll. Harmadszor, a modell feltételezi, hogy a folyamatban lévő áramlási és interaktív összeköttetések sávszélességvezérlése ideális, abban az értelemben, hogy bármilyen rendszerállapot-változást (pl. kapcsolat belépése vagy távozása) követő végtelenül kicsiny idő alatt a forgalomforrások újraállítják aktuális sávszélességüket. Bár ez nyilvánvalóan ideális feltételezés, feltesszük, hogy az IP csomagrétegen lévő pufferek elég nagyok ahhoz, hogy elnyeljék az IP csomagokat addig, amíg pl. a TCP vagy bármelyik másik felső rétegi protokoll csökkenti a kapcsolatokon az adatküldési sebességet.

5 Alkalmazási példák

5.1 A rendszer Markov modelljének felépítése és megoldása

A közös erőforrást használó különböző osztályú kommunikációs folyamatok QoS/GoS paramétereit már elemezték Markov-modellek segítségével [6, 1, 9]. E megközelítés alkalmazásának két fő nehézsége van: az (általában nagy) Markov-modell automatikus generálása és megoldása. Mi az Interneten hozzáférhető MRMSolve elnevezésű Markov hozam modell számító programmal végeztük számításainkat, amely a modell magas szintű leírása alapján végzi el automatikusan mindkét lépést [8].

5.2 A vizsgált átviteli rendszer ismertetése

A társalgási, az áramlási és az interaktív forgalmi osztályok *felajánlott* forgalmát és a teljes felajánlott forgalmat, a következőképpen definiáljuk:

$$\rho_i = \frac{\lambda_i}{\mu_i} b_i, \quad i = 1, 2, 3; \quad \rho = \sum_{i=1}^3 \rho_i,$$

A vizsgált összeköttetésen kialakuló forgalom különbözik a felkínált forgalomtól a hívásblokkolás és a sávszélesség-csökkenés miatt. Ha a folyamatban lévő áramlási és interaktív hívások

maximális tömörítése után sincsen elég szabad sáv szélesség, akkor az újonnan beérkező hívások elvesznek. Az így elutasított hívások következtében elvesző terhelést korlátozzák a maximálisan elfogadható blokkolási valószínűségekre vonatkozó értékek (B_1^{max} , B_2^{max} , illetve B_3^{max}).

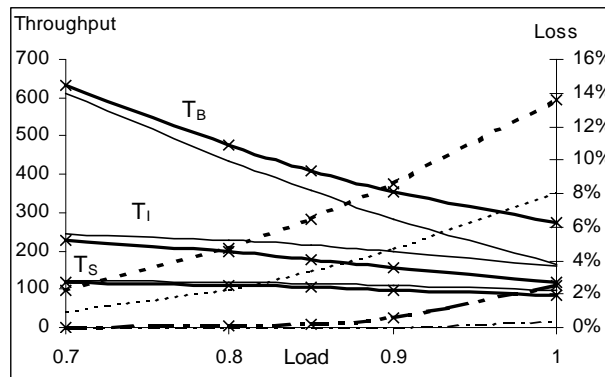
Az áramlási és interaktív folyamatok sáv szélessége a vizsgált összeköttetés forgalmi terhelésétől függően ingadozik. Az interaktív folyamatok sáv szélesség-ingadozása nem eredményez semmilyen terheléscsökkenést, mivel az interaktív osztályú összeköttetésen átvitt adatmennyiség független a rendelkezésre álló sáv szélességtől (az alacsonyabb sáv szélesség hosszabb szolgáltatási időt okoz). Ugyanakkor az áramlási folyamatok ingadozása további terheléscsökkenést idéz elő, hiszen az áramlási osztályú összeköttetésen átvitt adatmennyiség az összeköttetés alatt rendelkezésre álló sáv szélességgel arányos.

A vázolt jelenség mennyiségi jellemzésére két veszteségre vonatkozó jellemzőt vezetünk be:

$$B_{call} = B_1 \frac{\rho_1}{\rho} + B_2 \frac{\rho_2}{\rho} + B_3 \frac{\rho_3}{\rho};$$

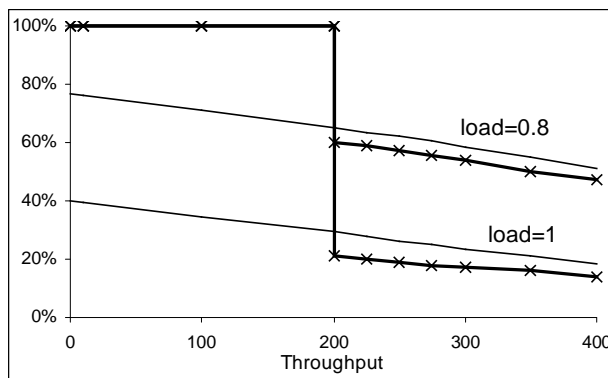
$$B_{bw} = B_1 \frac{\rho_1}{\rho} + \left(B_2 + (1 - B_2) \frac{b_2 - E(\theta_2)}{b_2} \right) \frac{\rho_2}{\rho} + B_3 \frac{\rho_3}{\rho}.$$

B_{call} csak a hívásblokkolás okozta veszteségekre vonatkozik, míg B_{bw} a hívásblokkolás és a sáv szélesség-csökkenés miatt bekövetkező veszteségeket együttesen jellemzi.



3. ábra: UMTS forgalmi osztályok vesztesége és sáv szélessége egyenletes forgalmi terhelés esetén

A 3. ábrán láthatók a teljes felkínált forgalom függvényében a sáv szélesség- és veszteség-jellemzők. A vizsgált forgalmi eset situáció a következő: $C = 2000$, $b_1 = 12$, $b_2 = 128$, $b_2^{min} = 64$, $b_3 = 256$, $b_3^{min} = 51.2$, $\mu_1 = 1/180$, $\mu_2 = 1/300$, $\mu_3 = 1/30$, a nagy prioritású osztályok felajánlott forgalma megegyezik, $\rho_1/\rho = \rho_2/\rho = \rho_3/\rho = 1/3$. (A hívások beérkezési intenzitása ennek megfelelő.) Két módszert hasonlítottunk össze. Az első esetben a teljes sáv szélességet használhatja az összes szolgáltatási osztály (teljes megosztás), $d=0$, míg a második esetben fenntartott sáv szélességrészt kap a háttérforgalom, $d=200$. Az első esethez tartozó eredmények vékony-, míg a második esethez tartozó eredmények vastag vonallal vannak ábrázolva. A sáv szélesség, a B_{call} és a B_{bw} görbék folyamatos -, szaggatott -, illetve pontvonallal vannak jelölve. T_I és T_S az interaktív és az áramlási osztály összeköttetések átlagos sáv szélességét, míg T_B a háttér forgalom számára rendelkezésre álló összes sáv szélesség átlag értékét jelöli.



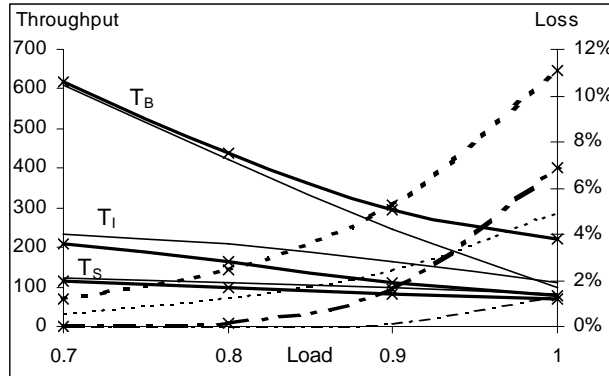
4. ábra: A háttér forgalom számára elérhető sáv szélesség eloszlása egyenletes forgalmi terhelés esetén

Látható, hogy gyakorlatban tipikus forgalmi terhelésnél, $\rho < 0.9$, a csökkethető sáv szélességű osztályoknál némi sáv szélesség-csökkenést és alacsony hívásblokkolási valószínűséget $B_{call} < 2\%$ tapasztalunk, míg a kombinált veszteség mértéke, B_{bw} , sokkal magasabb értéket mutat. Nagy terhelésnél, $\rho > 0.9$, folytatódnak ezek a trendek. Érdekes módon nem jelentkezik a háttérforgalom számára dedikált sáv szélesség fenntartásának, $d=200$, jelentős kedvező hatása a háttérforgalom átlagos sáv szélességére ($E(\theta_4)$). Azonban az átlagos sáv szélesség a háttérforgalom számára rendelkezésre álló sáv szélességnek nem elég pontos jellemzője. A sáv szélesség-fenntartás valódi előnye a 4. ábrán válik láthatóvá, amely θ_4 elosztását mutatja (azaz $\Pr(\theta_4 \geq x)$). Annak ellenére hogy a háttérforgalom átlagos sáv szélessége nem különbözik jelentősen, ha $d=200$; θ_4 eloszlása a teljesen más ebben az esetben. A háttérforgalmi szolgáltatások minőségére nézve a legjelentősebb különbség az, hogy annak valószínűsége, hogy nincs rendelkezésre álló sáv szélesség a háttérforgalom számára, $\Pr(\theta_4 = 0)$, lényegesen kisebb. Ha $\delta = 0$, akkor nagy (20% – 60%) a valószínűsége, hogy a nagyobb prioritású szolgáltatások teljesen kihasználják a rendelkezésre álló sáv szélességet, és a háttérszolgáltatások számára egyáltalán nem marad sáv szélesség. Ugyanakkor ha $\delta = 200$, akkor ez a minimális sáv szélesség mindig rendelkezésre áll a háttérosztályba tartozó szolgáltatások számára.

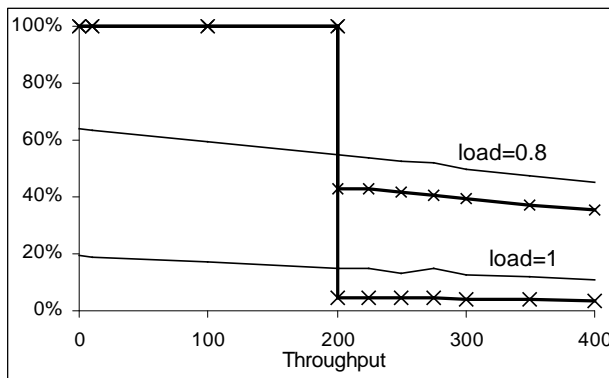
Az áramlási és interaktív osztály adaptív természeté az oka annak, hogy a magas prioritású szolgáltatások nagy valószínűséggel teljes mértékben kihasználják a rendelkezésre álló sáv szélességet. Ezek a szolgáltatások úgy alakítják a sáv szélességüket, hogy az összeköttetés teljes sáv szélességét a lehető legjobban kihasználják. Az 5. és 6. ábra megmutatja, hogyan növekszik a sáv szélesség teljes kihasználásának valószínűsége a forgalom összetételének változásával és az adaptív forgalmi osztályok arányának növekedésével. Ezeken az ábrákon ugyanazokat az teljesítmény jellemzőket ábrázoltuk, megváltozott forgalom-összetétel mellett: $\rho_1/\rho = 0.1, \rho_2/\rho = 0.1, \rho_3/\rho = 0.8$. A 6. ábrán látható, hogy ennél a forgalom-összetételnél sokkal nagyobb a valószínűsége, hogy nem marad sáv szélesség a háttérforgalom számára.

6 Összefoglalás

E tanulmányban kiterjesztettük a széles körben használt különböző sáv szélességű szolgáltatások veszteségi modelljét az UMTS forgalmi osztályokra. Azaz egységes modellt készítettünk,



5. ábra: Az UMTS osztályok vesztesége és sávszélessége domináns összenyomható forgalom esetén



6. ábra: A háttér forgalom számára rendelkezésre álló sávszélesség eloszlása domináns összenyomható forgalom esetén

amelybe belevettük a rugalmas osztályok két fő, a harmadik generációs mobilhálózatok (UMTS) számára szabványosított típusát, az áramlási és az interaktív osztályt. E két osztály viselkedése között nagyon fontos különbség (bár ezt gyakran figyelmen kívül hagyják), hogy tartási idejük és ténylegesen átvitt forgalmuk eltérő módon függ a rendszer terhelésétől.

A kialakított modell segítségével a blokkolási valószínűségekre és a bevezetett sávszélesség korlátokra vonatkozó teljesítményértékeket határoztunk meg két egyszerű sávszélesség-megosztási módszer felhasználásával. Arra a következtetésre jutottunk, hogy a sávszélesség egy részének az alacsonyabb prioritású háttérosztályhoz rendelésével ennek az osztálynak a szolgáltatás minőségét hatékonyan javíthatjuk anélkül, hogy számottevő blokkolást vagy sávszélesség-csökkenést okoznánk a magasabb prioritású osztályoknak. Ez az intuitív elvárásoknak ellentmondó eredmény az áramlási és az interaktív osztály rugalmas természetének következménye, mert ezek az osztályok - az vonalkapcsolt rendszerekkel ellentétben - hajlamosak az összes rendelkezésre álló sávszélesség felhasználására.

A kidolgozott modell és a segítségével számítható eredmények általánosan hasznosíthatók a jövő integrált szolgáltatású hálózatainak kialakításában.

Hivatkozások

- [1] Eitan Altman, Damien Artiges and Karim Traore, "On the Integration of Best-Effort and Guaranteed Performance Services", *INRIA Research Report No. 3222*, July, 1997.
- [2] Allan T. Andersen, Sören Blaabjerg, Gábor Fodor and Miklós Telek, "A Partially Blocking-Queueing System with CBR/VBR and ABR/UBR Arrival Streams", *5th IFIP International Conference on Telecommunications System*, Nashville, TN, USA, March 1997.
- [3] Emano Berruto, Giovanni Colombo, Pantelis Monogioudis, Antonella Napolitano and Kyriacos Sabatakis, "Architectural Aspects for the Evolution of Mobile Communications Toward UMTS", *IEEE Journal of Selected Areas in Communications*, Vol. 15, No. 8, pp. 1477-1487, October 1997.
- [4] G. Fodor, G. Malicsko, S. Malomsoky, "A Joint Radio- and IP Resource Reservation Scheme in All-IP 3rd Generation Networks", accepted to the *IEEE Wireless Communications and Networking Conference*, Chicago, IL, USA, September 2000.
- [5] D. Mitra, J. A. Morrison and K. G. Ramakrishnan, "ATM Network Design and Optimization: A Multirate Loss Network Framework", *IEEE/ACM Transactions on Networking*, Vol. 4, No. 4, pp. 531-543, August, 1996.
- [6] R. Nunez Queija, J. L. van den Berg, M. R. H. Mandjes, "Performance Evaluation of Strategies for Integration of Elastic and Stream Traffic", *International Teletraffic Congress*, UK, 1999.
- [7] T. Ojanpera and R. Prasad, eds., "Wideband CDMA for Third Generation Mobile Communications", *Artech House*, 1998.
- [8] S. Rácz, B. P. Tóth, and M. Telek, "MRMSolve: Numerical analysis of large markov reward models," in *Tools 2000*, pp. 337–340, Springer, LNCS 1786, 2000.
- [9] S. Rácz, M. Telek, and G. Fodor, "Link capacity sharing between guaranteed- and best effort services on an atm transmission link under GoS constraints," in *System Performance Evaluation: Methodologies and Applications* (E. Gelenbe, ed.), pp. 69–79, CRC Press, 2000. Ch. 5.
- [10] S. Rácz, M. Telek, and G. Fodor, "Link capacity sharing between guaranteed- and best effort services on an atm transmission link under GoS constraints." to appear in *Telecommunication System Journal*, 2001.
- [11] S. Rácz, M. Telek, and G. Fodor, "Call level performance analysis of 3rd generation mobile core network." ICC 2001, Helsinki, Finland, June, 2001.
- [12] K. W. Ross, "Multi-service Loss Models for Broadband Telecommunication Networks", *Springer Verlag London Limited*, ISBN 3-540-19918-7, 1995.
- [13] "Third Generation Mobile Systems in Europe", *IEEE Personal Communications Magazine*, April 1999.

- [14] "The Evolution of TDMA to 3G", *IEEE Personal Communications Magazine*, June 1999.
- [15] "Multiple Access for Broadband Wireless Networks", *IEEE Communications Magazine*, July 2000.
- [16] 3GPP TS 23.107, "Services and System Aspects: QoS Concept and Architecture", ver. 3.2.0, 2000.
- [17] 3GPP TR 23.922, "Architecture for an All-IP Network", ver. 1.0.0, 1999.