híradástechnika

VOLUME LVII.

2002/9 Szeptember

Hívásengedélyezés

Titkosítás, biztonság

Gazdaságpolitika

Hírközlési és Informatikai Tudományos Egyesület folyóirata

Tartalom

Címlap: Régen jól elrendezett kövek hordozák a titkos üzenetet

Készült NHIT támogatással

Előszó a szeptemberi számhoz

A távközlési lehetőségek és alkalmazások választékának növekedése számos területen új kényelmi, gazdasági megoldásokat tett lehetővé. Az előnyök mellett azonban, mint minden más újdonság, vonzó azok számára is, akik az informatika újdonságaiból törvénytelen hasznot akarnak húzni. A távközlés és informatika elkerülhetetlen bűnözői használata mellett azonban ezen a területen megjelennek törvénytisztelő fiatalok, akik magukat kipróbálandó sportteljesítményként értékelik, ki tud több kódot feltörni, nehezebben hozzáférhető adatbázisba behatolni, és számukra talán értéktelen titkot megismerni. A távközlési és informatikai szakemberek elfogadják ezt a kihívást, és műszaki-matematikai módszerek kidolgozásával igyekeznek meggátolni, hogy illetéktelenek hozzáférjenek, megmásítsanak, felhasználjanak bizalmas információkat. Ha véglegesen nem is sikerül meggátolni az információ védelmét, de mindenesetre megnehezítik a jogtalan behatolást.

Ez utóbbi is nagy eredmény. Ha egy védelmi módszer megismerése és a jogtalan behatolás kidolgozásának ideje elegendően hosszú, akkor már számos értékes információ elavul. Sok esetben néhány óra, vagy nap elegendő ahhoz, hogy a beavatkozás elveszítse célját. Így a bűnözők és a titkosítók versenye, a támadók és a védők harca nem a tökéletes megoldásokat keresi, hanem elegendő, ha az információ jogtalan megszerzése hosszú időt igényel.

E verseny miatt sok értékes szakember ideje és sok intézmény pénze kárba vész. Mindkét oldal azonban közben pallérozza az agyát, amit esetleg máshol, más helyzetben a közösség javára tud hasznosítani. Érdemes ezért ezekről az eljárásokról írni és a szakmai közvéleményt az eredményekkel megismertetni, mert néhány gyakorlati kódolási, vagy kódfeltörési módszer mögött értékes matematikai újdonságok találhatók.

E havi számunk ezzel kapcsolatos négy cikke különböző esetekre, különböző szempontok szerint készült. Az első cikk az elliptikus görbék rendkívül érdekes tulajdonságait tárja az olvasók elé. Ennek számos hasznosítási területe lehet, de a cikkből azt az egyet ismerjük meg, mely nagyon nehezen, hosszú idő alatt feltörhető kódoláshoz vezet. A második cikk az elmélettől a gyakorlat felé vezető úton, a mobilhálózatok veszélyeivel és az elérhető biztonsággal foglalkozik. Napjaink aktuális problémája az adminisztráció digitalizálása, elektronizálása, mely elvezet az e-közigazgatásig és az e-kereskedelemig, melyek működésének és biztonságának előfeltétele az elektronikus aláírás. Bár ezt már a parlament is megtárgyalta, erre törvényt alkotott, bevezetése érdekében még számos teendő áll előttünk. A sort a szeptember 11-i terrortámadás titkosítási problémái és a rendkívül nehezen felismerhető, rejtett üzenetek kérdése zárja. Itt a terroristák terveinek megismerése a tisztességes emberek célja.

Az ezen témakör előtti első blokkban a hívásengedélyezés kérdését tárgyaljuk. Ennek bevezető cikke átfogóan ismerteti, hogy miért került ez a téma előtérbe, miért nem volt jelentősége a vonalkapcsolt, hierarchikus rendszerekben. Reméljük, hogy a bevezető cikket követő három módszer világosan megmutatja, hogy különböző helyzetekben előnyös lehet más-más hívásengedélyezési eljárás.

A szakma és a politika közös ügye a digitális szakadék, vagyis a hírközléssel, informatikával rendelkezők növekvő előnyének vizsgálata a leszakadók, a műszaki lehetőségekhez nem hozzáférőkkel szemben. Ezzel foglalkozik az utolsó csoportban lévő egyik cikk. Rendkívül tanulságos, hogy a szegények és a gazdagok különbsége minden műszaki újdonság megjelenésekor előtérbe kerül. Legyen az vasútépítés, autópálya-használat, gépkocsivásárlás, repülés, melynek lehetőségeit a gazdagok mindig előbb veszik igénybe és tovább növelik előnyüket. Valamennyi esetben azonban kiderül, hogyha a szegények a tudás megszerzéséhez hozzájutnak, és a gazdagok segítik, támogatják a tanulást, a szakmai képzést, akkor bármely szakadék áthidalásához kezükben lesz az eszköz.

Bár ebben a számban több olyan cikk van, amely az új eredmények kidolgozásához magas színvonalú matematikai apparátust használ és a működés bemutatása is ennek segítségével lehetséges, mégis úgy látszik, hogy egészében valamennyi cikk napi problémáinkhoz kapcsolódik. A távközlés és informatika széles körű használata nemcsak azoknak előny, akik ebből élnek, hanem azok számára is, akik felhasználásával eredményesebben dolgozhatnak, vagy segítségével magánéletüket tehetik szebbé.

Dr. Lajtha György
a szerkesztőbizottság elnöke

Hívásengedélyezés garantált minőségű hálózatokban
(áttekintés)

NÉMETH KRISZTIÁN

BME-TTT egyetemi hallgató
E-mail: nemeth_k@soha.ttt.bme.hu

Különböző rendszerű hírközlő hálózatokban közös igény, hogy a hálózat által felajánlott szolgáltatáshoz – az információátvitelhez – minőségi garanciákat is biztosítson a szolgáltató. Ezeket a kielégítő hálózati szolgáltatásokat nevezzük garantált minőségű szolgáltatásoknak (Quality of Service, QoS). Nem kapcsolt és vonalkapcsolt hálózatok esetében a minőség garantálása a technológiából fakadóan viszonylag egyszerű, csomagkapcsolt hálózatokban azonban ez sokkal nehezebb feladat. A szolgáltatások minőségét csomagkapcsolt hálózat esetén különböző paraméterekkel jellemezhetjük, melyek közül a leggyakrabban használtak az átvitel késleltetése, a késleltetés ingadozása (jitter) és a csomagok elvesztésének valószínűsége.

Különböző csomagkapcsolt hálózati technológiák a minőséget eltérő módszerekkel biztosítják, sőt egyazon hálózattípusra is vannak egymástól lényegileg különböző típusú, minőséget garantáló eljárások. Ez utóbbi kategóriába tartoznak az internetprotokollt (IP) használó hálózatok. Az IP technológia tervezésekor és fejlesztésekor sokáig a „minél többet, minél gyorsabban" elv érvényesült, de bármilyen nagy adatmennyiséget, bármilyen gyorsan képes is továbbítani ez a hálózattípus, garanciát semmire sem tud nyújtani, és az adatfolyamok egymás közötti prioritásának meghatározása sem lehetséges. Különös tekintettel arra, hogy az új csomagkapcsolt hálózati alkalmazások (például mozgókép és hang továbbítása) megkövetelik a garantált minőségű adatátvitelt, e szolgáltatások integrálása az IP hálózatokba – és ezáltal az internetre – egy nagyon fontos lépés e hálózatok fejlődésében. Ez a fejlesztési, szabványosítási munka jelenleg is folyamatban van. Különböző IP QoS architektúrajavaslatok születtek, például az Integrated Services (IntServ, integrált szolgáltatású hálózat), illetve a Differentiated Services (DiffServ, differenciált szolgáltatású hálózat). Bár e technológiák még kísérleti stádiumban vannak, valószínű, hogy a gyakorlatban ezeknek egy keveréke fog elterjedni, kiegészítve más megoldásokkal is. Ellentétben a hagyományos IP hálózatokkal, az ATM (Asynchronous Transfer Mode, aszinkron átviteli mód) hálózatokban a tervezés kezdetétől szempont volt a minőségi garanciák pontos rögzítése. Ennek ellenére itt is lehetnek különbségek az egyes részletekben, például az újság jelen számának egyik központi témája, a hívásengedélyezés is történhet különböző algoritmusok alapján. Az eddig elmondottakból is érződik, hogy a csomagkapcsolt hálózatokban a minőség garantálása meglehetősen összetett feladat. A teljes rendszer több építőelemből épül fel, melyek közt szerepelnek (a teljesség igénye nélkül) a különböző jelzésrendszerek, prioritásos sorban állási rendszerek, minőségi garanciát nyújtó forgalomirányítás (QoS routing), erőforrás-foglalás a forgalomirányítókban, és nem utolsósorban a hívásengedélyezés.

Vizsgáljuk meg tehát közelebbről, mi is az a hívásengedélyezés (Call Admission Control, CAC)! Ennek a modulnak a feladata eldönteni egy új igény megjelenésekor, hogy a hozzá tartozó adatfolyam beléphet-e a hálózatba. Abban az esetben lehet az új forgalmi igényt elfogadni, amennyiben két feltétel teljesül: egyrészt az új forgalom beengedésével a már elfogadott kapcsolatok minősége nem csökken az előzetesen megállapított szint alá, másrészt az új forgalom átvitelének a minősége is el fogja érni azt a szintet, amit az igény bejelentője elvár. A hívásengedélyezés folyamatában tehát éppen az a nehéz, hogy meg kell találni az egyensúlyt két egymásnak ellentmondó követelmény között: egyrészt garantálni kell a hálózatba már beengedett forgalmak integritását, másrészt maximalizálni kell a hálózat kihasználtságát. Ebből következik, hogy a hívásengedélyezés során kétfajta hiba fordulhat elő: első fajú hibának nevezzük azt, amikor elutasítunk egy kérést, holott az még befért volna a hálózatba; másodfajú hiba az az eset, amikor elfogadtunk egy igényt, pedig az már túlterheli a hálózatot. Hogy melyik fajta hiba mennyire súlyos, az a szolgáltató üzletpolitikai döntésétől függ. Az ebből a döntésből származó számszerű paraméterek akár bemenetei lehetnek az alkalmazott hívásengedélyezési algoritmusnak is.

A fentiekből látszik, hogy a hívásengedélyezés folyamata valójában egy párbeszéd a „felhasználó" (igénybejelentő) és a „hálózat" között. Megjegyzendő, hogy a „felhasználó" alatt itt azt értjük, aki az adott szolgáltatást igénybe akarja venni, ez lehet egy ember, de egy szoftver vagy eszköz is, sőt akár egy másik hálózat is.

Ezen túlmenően a „hálózat"-ban is különféleképpen helyezkedhet el az a modul, amely végül választ fog adni a felhasználó kérésére. Maga a párbeszéd is többféleképpen valósulhat meg. Kezdetként a felhasználó specifikálja az igényeit, amelyet egy szabványos formában, a forgalomleíró adatblokkban elküldi a hálózatnak. A forgalomleíró adatblokk szabványos, abban az értelemben, hogy egyazon hálózat mindig ugyanabban a formátumban várja az igényspecifikációt, azonban különböző hálózatoknál ez a formátum különböző lehet. Természetesen a törekvés az, hogy csak néhány leíró adatblokktípus terjedjen el, és azok egyikét használja minden hálózat. A felhasználó kérésére a hálózat válaszolhat egyszerű „igen/nem"-mel, jelezve a kérés elfogadását vagy elutasítását, elképzelhető azonban olyan eset is, ahol a válasz ennél bonyolultabb. Erre egy példa az olyan válasz, amelyben a hálózat a kérést elutasítja, de javaslatot tesz egy olyan forgalomleíróra, amelyet még el tudna fogadni. Az engedélyezési párbeszéd ez esetben tovább folytatódhat.

A hívásengedélyezési döntés helye is eltérő a különböző típusú hálózatokban. Az IntServ IP hálózatokban és az ATM hálózatokban például minden egyes forgalomirányító, illetve kapcsoló eldönti, hogy a saját belső erőforrásait és a kimeneti csatornáit figyelembe véve a felajánlott pluszforgalmat el tudja-e vezetni. Az igényt, majd ezeket a döntéseket egy erőforrás-foglaló protokoll továbbítja a hálózati csomópontok és a hívás kezdeményezője között. A hívást a hálózat részéről abban az esetben lehet elfogadni, ha minden egyes csomópont, amelyeken később az adatok át fognak folyni, elfogadta azt. A DiffServ architektúra nem tartalmaz hívásengedélyezést, arra azonban lehetőség van, hogy egy külső modul engedélyezze vagy tiltsa a DiffServ tartományba irányuló forgalmat. Ebben az esetben egyetlen döntés születik, amely egy adott hálózati tartomány egészére érvényes elfogadást vagy elutasítást jelent. E döntés meghozatalának helye is hálózatfüggő. A döntést meghozhatja az a csomópont, ahol a forgalom majd belép a tartományba (ingress router), de lehet ez egy erre külön kijelölt csomópont is, melyet sávszélességbrókernek (bandwidth broker) nevezünk. Vannak ezektől eltérő javaslatok is, azonban teljes képet már csak azért sem tudunk adni, mert mind a DiffServ, mind a köré épülő QoS architektúrák jelenleg is fejlesztési stádiumban vannak.

Csoportosíthatjuk a hívásengedélyező algoritmusokat aszerint is, hogy determinisztikus alapon döntenek, vagy pedig statisztikus alapon. Az első esetben az éppen felépítés alatt álló és a már elfogadott kapcsolatok forgalmi leíró blokkjaiból a hívásengedélyező algoritmus kiszámítja, hogy a hálózat képes lesz-e „a legrosszabb esetben" is az adatok továbbítására a megadott minőségben. Ez azt jelenti, hogy a hálózat akkor sem követhet el másodfajú hibát, ha minden forrás a neki megengedett maximális adatmennyiséget küldi minden időpillanatban (ún. mohó források). Ebben az esetben a döntési algoritmus viszonylag egyszerű: amennyiben az eddig elfogadott kapcsolatok és az új igény erőforrásigényének összege nem nagyobb, mint a hálózat vizsgált részének áteresztő képessége, akkor az új igény elfogadható. E módszer előnye, hogy egyszerű, továbbá sosem fordulhat elő másodfajú hiba. Hátránya, hogy nagyon konzervatív, és mivel nem veszi figyelembe a folyamok statisztikai paramétereit, ezért az esetek jelentős részében a hálózat kihasználtsága indokolatlanul alacsony lesz. A statisztikus alapú döntés esetén több forgalmi leíró paramétert is figyelembe lehet venni, így például az átlagos sávszélességértéket. Ebben az esetben lehetőség van a maximális erőforrásoknál kevesebbet allokálni egy folyam számára, így több igényt elfogadni, ezáltal gazdaságosabban üzemeltethető a hálózat. Ezzel a módszerrel akkor van esélyünk a jobb kihasználtságra, ha a források változó bitsebességgel bocsátják ki az adatokat (Variable Bitrate, VBR), azaz az általuk igényelt sávszélesség időben nem állandó (Constant Bitrate, CBR). E döntési típus előnye, hogy javul a hálózat kihasználtsága, hátránya azonban, hogy a döntés algoritmusa nagyon bonyolult lehet, továbbá, hogy megjelenhet a másodfajú hiba is.

A hívásengedélyezéshez szükséges forgalomleíró adatblokk is többféle lehet. Egyik formájában a felhasználó néhány paramétert ad meg az általa generált forgalomról, ilyen lehet például a maximális sávszélesség, az átlagos sávszélesség, a maximális löketméret. Egy másik lehetőség, hogy léteznek előre definiált forgalmi osztályok – például bronz, ezüst és arany osztály egyre növekvő minőségi garanciákkal –, és a felhasználó ezek egyikébe sorolhatja be az általa létrehozott adatforgalmat. Ebben az esetben, ha N osztályunk van, akkor minden egyes hálózati állapotot egy N dimenziós vektorral írhatunk le, melynek az i-dik eleme mondja, meg, hogy az i-dik osztályból hány darab adatfolyam van a hálózatban. Ekkor a hívásengedélyezés problémája megegyezik egy N dimenziós szeparáló hiperfelület létrehozásával, amely megmutatja, hogy mely állapotvektorok fogadhatók el, és melyek nem. Az engedélyezési algoritmus ezek után a következő: létrehozzuk azt a vektort, ami akkor keletkezne, ha az aktuális igényt elfogadnánk. Ha ez a vektor még elfogadható (a szeparáló hiperfelület elfogadó oldalán van), akkor a hívást beengedjük, ellenkező esetben pedig nem. E módszer előnye, hogy egyszerűbb az általános esetnél, ahol is mindenki tetszőleges forgalmat specifikálhat, hátránya pedig ennek következtében az, hogy kisebb a felhasználó választási szabadsága.

Létezik egy egészen más megközelítés is, a mérésalapú hívásengedélyezés. Bár ebben az esetben is el kell küldenie a felhasználónak a forgalomleíró blokkot, a döntés csak részben alapszik ezeken. Itt nem az történik, mint a hagyományos esetben, amikor is az eddig elfogadott forgalomleíró blokkok és az új igény leíróblokkjának összességéből számolnak ki egy értéket, amit összevetnek a rendelkezésre álló kapacitással. A mérésalapú esetben a hálózat jelenlegi terheltségére a forgalom méréséből következtetnek valamilyen módon, és azt vizsgálják, hogy emellé befér-e még a for-

galmi leíróblokkal adott új igény a kérdéses tartományba. A módszer előnye, hogy a forrásoknak kevesebb forgalmi leíró paramétert kell elküldeniük a kapcsolat felépítésekor (hiszen lehet, hogy ők maguk sem tudják előre ezeket), tipikusan az igényelt maximális sávszélesség (peak rate) elegendő. Hátránya viszont, hogy a rendszer bonyolultabb lesz, továbbá a másodfajú hiba megjelenésének tágabb teret ad a méréssel megállapított hálózati terheltségnek a bizonytalansága. Végezetül megemlítjük, hogy vannak olyan javaslatok is, amelyekben egyáltalán nem kell a forrásnak forgalmi leíró adatblokkot küldenie. Ebben az esetben egyszerűen elkezdi küldeni az adatait, és rövid idő múlva mérések segítségével a hálózat eldönti, hogy ez a pluszforgalom még befér-e a tartományba, vagy sem. Ez utóbbi esetben a forrásnak az adatok küldését meg kell szakítania.

A következő három cikk mélyebb betekintést nyújt a hívásengedélyezés témájába, felvázolva és értékelve egy-egy konkrét algoritmust.

Hírek

A korszerű gépkocsikban az indítómotort és a generátort felváltja az autóvillamossági „csomag", az ISG (integrált starter-generátor), melyet a motor és a sebességváltó közé szerelnek. Villamos motorként a belső égésű motort a hagyományos indítómotornál lényegesen gyorsabban indítja. Generátorként áramot generál minden fogyasztó táplálására. Az ISG a fejlesztett áram felesleges részével az akkumulátor, fékezéskor pedig segítségével a jármű mozgási energiája elektromos energiává alakítható és tárolható. Szükség esetén az ISG teljesítményt ad a tengelyekre, és „stop and go" menetben villanymotorként hajtja a járművet. Lámpánál a robbanómotor leáll, amíg a vezető a féken tartja a lábát, utána viszont az ISG a másodperc tört része alatt az üresjárati fordulatszámra pörgeti fel a főtengelyt, a motorvezérlés pedig aktiválja az üzemanyag-befecskendezést. Az ISG 15%-os fogyasztáscsökkenést eredményezhet.

❖ ❖ ❖

Háromféle új Ethernet-modul áll rendelkezésre a Siemens Sicomp IMC márkájú ipari mikroszámítógépeihez. CPCI felépítésű modulok 10 Mbit/s Ethernet és 100 Mbit/s Fast Ethernet átviteli sebességgel működnek. A három változat az RJ45 csatlakozótechnikával kínált interfészek számában (2-6) és fajtájában (hu funkcióval vagy anélkül) különbözik egymástól. A szoftver drivereket minden általánosan használt operációs rendszerhez – pl. Windows, Linux, VxWorks, QNX, RMOS53 – kifejlesztették.

❖ ❖ ❖

A szlovák távközlési törvény módosítása a befészkelt szolgáltatókat elfogadható feltételek mellett kötelezi az előfizetői vonalakhoz való szabad hozzáférés lehetővé tételére. A módosítást a szlovák kormány már jóváhagyta és a szlovák parlament elé terjesztette.

❖ ❖ ❖

A Hawaii-szigetcsoport egyik kis tagja, Kapolei céltudatosan épülő távközlési gócpont a Csendes-óceánon. Egyedülálló funkciója az ott végződő, együttesen több mint 7 Terabit/s kapacitású kábelek végberendezéseinek, 16 műholdas összeköttetés földi állomásának kiszolgálása és közöttük kapcsolatot létesítő eszközök működtetése. Legalább hat globális távközlési vállalat vetette meg a lábát a szigeten. A helyi idők különbségét kihasználják arra, hogy ugyanazon a napon Észak-Amerikával és Ázsiával is üzleti kapcsolatban legyenek, az e-business minden változatát gyakorolva.

Gyors és pontos hívásengedélyezés csomagkapcsolt hálózatokban

FANCSALI ALPÁR, VÁZSONYI MIKLÓS, DR. LEVENDOVSZKY JÁNOS

Budapesti Műszaki és Gazdaságtudományi Egyetem (BME), híradástechnikai tanszék
alpar@hit.hit.bme.hu

Reviewed

A statisztikus multiplexelésen alapuló hívásengedélyezési (Call Admission Control – CAC) algoritmusok valós időben elvégzik a CAC feladatát, miközben gyorsan igazodnak az időben változó forgalmi paraméterekhez. Így a csomagkapcsolt hálózatok kihasználtságát nagymértékben megnövelik. A cikk a CAC feladatát halmazszeparálásra vezeti vissza. A pontos szeparáló felület kiszámítása két dimenzióban, két forgalmii osztálynál nem okoz különösebb nehézséget, ezért a három forgalmi osztályos esetet visszavezeti több kétdimenziósra. A nem parametrikus esetben egy mérés alapján felvett mintahalmaz alapján kell a szeparáló felületet közelíteni. Erre a közelítésre újszerű neurális architektúrát javasolunk, amit gradiens módszerrel lehet tanítani.

1. Bevezetés

Az internet elterjedésével a csomagkapcsolt technológia alapvető kommunikációs eljárás lett, és a gerinchálózatok is IP alapúak. Egyre nagyobb igény mutatkozik a hálózat jobb kihasználására, miközben meghatározott minőségű (QoS) távközlési szolgáltatásokat kell biztosítani. A tervezés alapja, hogy az előfizetői hálózatok egy adott (vagy akár változó) kapacitású keresztmetszettel csatlakoznak a gerinchálózathoz. Ekkor az alhálózat fenntartójának az a célja, hogy a felhasználók statisztikus tulajdonságait figyelembe véve a lehető legjobban használja ki a szűk keresztmetszeteket, miközben adott minőségi paramétereket biztosít (pl. átlagos csomagkésleltetés, csomagvesztési arány).

Ha a felhasználók által generált összforgalom mellett sérülnek az iménti minőségi paraméterek, úgy az adott hívásigényt vissza kell utasítani. Ezt a feladatot a hívásengedélyezés (Call Admission Control – CAC) oldja meg. Ahogy azt az idevágó kutatások demonstrálták [1, 2], minél „változatosabb" a szállított forgalom, annál fontosabb szerepet játszik a forgalomirányítás.

E feltételek kielégítése során kompromisszumot kell kötnünk az algoritmikus bonyolultság és a valós idejű működés között. Eddig csupán olyan hozzávetőleges számításokon alapuló, de egyszerű algoritmusokat használtak, amelyek korlátozták a hálózat kihasználtságát [1, 8]. Ebben a cikkben a CAC feladatát egy feltételes optimalizációs problémaként fogalmazzuk meg, amellyel a lehető legnagyobb hálózati kihasználtságot érjük el, miközben teljesülnek az előírt minőségi (QoS) jellemzők.

A cikk a következők szerint épül fel:

- On/off forrásokat feltételezve megadjuk az egyes forgalmi szituációkhoz tartozó aggregált forgalom eloszlását.
- Mind puffer nélküli közelítésnél, mind G/D/1 sorállásmodell mellett kiszámoljuk a QoS paramétereket (az átlagos cellakésleltetés és a puffertúlcsordulás) adott forgalmi konfigurációnál.
- A QoS ellenőrzés során a forgalmi szituációkat két halmazba osztjuk, azaz halmazszeparálásként fogalmazzuk meg a problémát.
- Mivel a QoS ellenőrzés analitikusan bonyolult folyamat, ezért először csak két forgalmi osztálynál határozzuk meg a pontos szeparáló felületet.
- A kétdimenziós eredményeket kiterjesztjük három forgalmi osztályra.
- Ismeretlen források esetén mérések alapján végezzük a hívásengedélyezést egy újszerű neurális architektúrával.
- Végezetül sokoldalú valós paraméterek melletti szimulációk igazolják a kidolgozott módszerek hatékonyságát.

2. A probléma egzakt modellje

A probléma formális leírásához tekintsük a felhasználók által kibocsátott forgalmat egy véletlen folyamatnak. Ennek az összforgalomnak a tulajdonságai azért fontosak, mert el kell dönteni, hogy a rendelkezésre álló csatornakapacitás és pufferméret mellett a hálózatba beengedett forgalom megsérti-e az előírt QoS paramétereket. Ha a pufferben túlcsordul, akkor ez cellák elvesztéséhez vezet, ezért ennek a valószínűségét (Cell Loss Probability – CLP), valamint a cellák átlagos várakozási idejét (Mean Cell Delay – MCD) kell kiértékelni. Ezek alapján a CAC algoritmus szétválasztja a hálózatba beengedhető és visszautasítandó hívásokat.

2.1 A felhasználók modellezése

A pontos analitikus tárgyaláshoz a következő feltevésekkel éljünk:

- A felhasználókat a kibocsátott forgalmuk alapján (pl. adat, beszéd, videó) osztályokba soroljuk ($i=1,2,...,M$).
- Az egy forgalmi osztályba tartozó felhasználók azonos statisztikával rendelkeznek (átlagsebességüket m_i, csúcssebességüket h_i jelöli).
- Az on/off forrásmodellnek megfelelően a felhasználók forgalmát Bernoulli valószínűségi változók írják le, azaz pl. az i-dik forgalmi osztályból a j-dik felhasználó által kibocsátott forgalmat jelölje az $x_j^{(i)}$ valószínűségi változó, amely az alábbi eloszlással rendelkezik: $P(X_j^{(i)} = 0) = 1 - \rho_i$ és $P(X_j^{(i)} = h_i) = 1 - \rho_i$ minden j-re, ahol $\rho_i := m_i / h_i$.
- Vezessük be az $(n_1, ..., n_i, ..., n_M)$ forgalmi állapotvektort, amely a rendszer állapotát mutatja, és ahol n_i azt jelzi, hogy az i-dik forgalmi osztályból éppen hány felhasználó van jelen. Jelölje N az $\mathbf{n}$ vektorok halmazát.
- Az i-dik osztályból adó felhasználók $Y_i = Znij=1x_j^{(i)}$ összforgalma binomiális $Y_i = \sum_{j=1}^{n_i} X_j^{(i)}$ eloszlást követ:

$$P(Y_i = y_i h_i) = \binom{n_i}{y_i} \rho_i^{y_i} (1-\rho_i)^{n_i - y_i}. \tag{2-1}$$

- Ennek megfelelően az összterhelés eloszlása:

$$P(Y=l) = \sum_{y_1 h_1 + ... + y_M h_M = l} \prod_{i=1}^{M} \binom{n_i}{y_i} \rho_i^{y_i} (1-\rho_i)^{n_i - y_i}. \tag{2-2}$$

2.2 A kiszolgálási modell

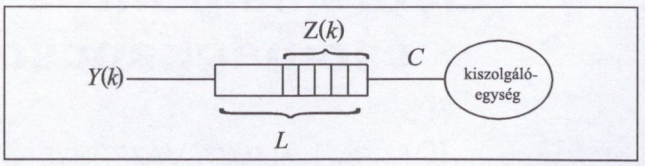
A gyakorlatban kétféle kiszolgálási modellt szoktak alkalmazni: a pufferes és a puffer nélküli közelítést. A felhasználók forgalmát mindig egy pufferben multiplexáljuk. A puffer nélküli közelítés folytonos modell, ahol feltesszük, hogy minden felhasználó egyszerre kerül kiszolgálásra. Ebben a modellben a kiszolgálás késleltetését nem lehet figyelembe venni. Azonban a késleltetésre is gyakran előírást kell tenni, ezért a cikk megemlít egy egyszerű sorállási kiszolgálási modellt.

Puffer nélküli közelítés (zero buffer approximation – ZBA) modell esetén akkor van túlcsordulás, ha a felhasználók összforgalma meghaladja a C csatornakapacitást. Ekkor a cellavesztés valószínűsége:

$CLP = \sum_{l>C} P(Y=l)$, ahol a $P(Y=l)$ valószínűséget a (2-2) adja meg. Egy $\mathbf{n}$ forgalmi konfiguráció akkor elfogadható, ha a CLP nem halad meg egy előre megadott szintet, amit $e^{-\gamma}$ jelöl. Ekkor a CAC-t megvalósító ideális döntési függvény:

$$g(\mathbf{n}) = \begin{cases} 1, \text{ ha } CLP < e^{-\gamma} \\ -1 \text{ egyébként.} \end{cases} \tag{2-3}$$

G/D/1 sorállási modell esetében az elérési modult a következő egyszerű sorállási rendszerrel írjuk le:



2-1. ábra A G/D/1 sorállási modell

Feltesszük, hogy minden időrésben egy cellát vesz el a kiszolgálóegység (G/D/1 modell). A puffer túlcsordulásakor cellák vesznek el, ezért az egyik QoS kritérium most is a cellavesztési valószínűség (CLP), amit a $CLP<e^{-\gamma}$ alakban korlátozunk. Ugyanakkor a sorbaállás miatt számolni kell a késleltetési idővel is, ezért a másik QoS paraméter az átlagos cellakésleltetés, amire szintén előírást teszünk: $MCD<e^{-\beta}$. Egy $\mathbf{n}$ forgalmi konfiguráció akkor elfogadható, ha a $CLP<e^{-\gamma} \wedge MCD<e^{-\beta}$ feltétel teljesül. Ekkor a QoS kommunikáció megvalósul. A CAC feladata, hogy egy adott $\mathbf{n}$ esetén a QoS kritériumok alapján a hívás beengedéséről döntsünk.

A pontosabb leírás érdekében jelöljük a pufferben a k-dik időpillanatban várakozó cellák számát Z_k-val. A puffer állapotát a következő egyenlet írja le:

$$Z_{k+1} = \lceil Z_k - 1 \rceil^+ + Y_{k+1}, \text{ ahol } \lceil z \rceil^+ = \begin{cases} z, \text{ ha } z > 0 \\ 0 \text{ egyébként.} \end{cases} \tag{2-4}$$

Mivel Z_k csak az előző időpillanatban felvett értéktől függ, ezért L hosszúságú puffer esetén $L+1$ állapotú Markov-láncot alkot, amelynek állapotátmeneti valószínűségei a (2-4) felhasználásával:

$$P_{ij} = P(Z_1 = j | Z_0 = i) = P\left(\lceil Z_0 + Y_1 - 1 \rceil^+ = j | Z_0 = i\right) = P\left(\lceil i + Y_1 - 1 \rceil^+ = j\right) \tag{2-5}$$

Az Y aggregált forgalom eloszlását a (2-2) fejezi ki, ahol most h_i értelemszerűen 1 cellát jelöl i-től függetlenül (diszkrét on/off modell), ρ_i pedig annak a valószínűségét, hogy a forrás ad cellát. A (2-5) egyenletből a következő módon áll elő a puffer $\mathbf{P}$ állapotátmeneti mátrixa:

$$\mathbf{P} = \begin{bmatrix} p_0 + p_1 & p_2 & p_3 & \cdots & \cdots & \sum_{l=L+1}^{\infty} p_l \\ p_0 & p_1 & p_2 & \cdots & \cdots & \sum_{l=L}^{\infty} p_l \\ 0 & p_0 & p_1 & \cdots & \cdots & \sum_{l=L-1}^{\infty} p_l \\ 0 & 0 & p_0 & \ddots & \ddots & \sum_{l=L-2}^{\infty} p_l \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & p_0 & \sum_{l=1}^{\infty} p_l \end{bmatrix}, \tag{2-6}$$

ahol $p_i = P(Y = „i$ számú cella"). A puffer stacioner eloszlását jelölje a $\bar{\pi}$ sorvektor ($\mathbf{P}$ egységnyi sajátértékéhez tartozó sajátvektor), amit a $\bar{\pi}\mathbf{P} = \bar{\pi}$ egyenletből rekurzívan könnyen ki lehet fejezni (kihasználva, hogy $\mathbf{P}$ felső háromszög mátrixhoz hasonlít).

A cellavesztési valószínűség az elveszett cellák és a beérkező cellák várható értékének hányadosaként van definiálva, azaz:

$$CLP = \frac{\mathbf{E}(\text{elvesztett cellák})}{\mathbf{E}(Y)} = \frac{\sum_{k=1}^{\infty} k \sum_{l=0}^{L} \pi_l p_{L-l+k+1}}{\sum_{i=1}^{M} n_i m_i}. \quad (2\text{-}7)$$

Az átlagos cellakésleltetést egyszerűen a $\Delta \cdot \sum_{l=1}^{L} l\pi_l$ kifejezéssel lehet megkapni, ahol Δ jelöli a cellák kiszolgálásának időrését. Ezzel az ideális döntési függvény:

$$g(\mathbf{n}) = \begin{cases} 1, \text{ ha } CLP < e^{-\gamma} \text{ és } \sum_{l=1}^{L} l\pi_l < e^{-\beta} \\ -1 \text{ egyébként,} \end{cases} \quad (2\text{-}8)$$

ahol a CLP-t a (2-6) adja meg.

3. A CAC mint halmazszeparálás

A (2-4) kifejezés alapján látható, hogy a CAC a forgalmi állapotvektorok halmazát két részre osztja:

$$N^{(1)} := \{\mathbf{n}: g(\mathbf{n}) = 1\} \text{ és } N^{(-1)} := \{\mathbf{n}: g(\mathbf{n}) = -1\}. \quad (3\text{-}1)$$

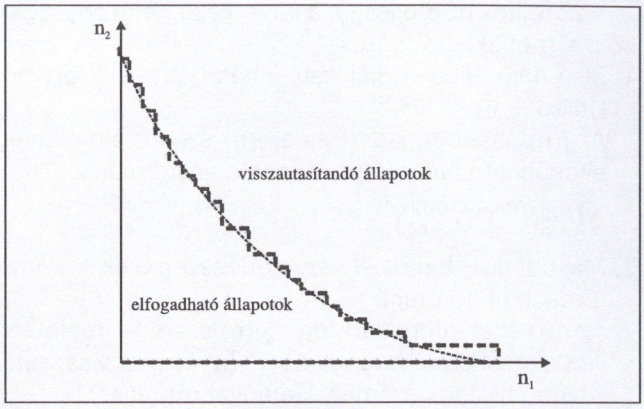
Sajnos $g(\mathbf{n})$ analitikus kiértékelése rendkívül hosszadalmas feladat, ezért ez alapján nem lehet valós idejű hívásengedélyezést végezni. Ehelyett csak K előre kiszámított minta alapján fogjuk elvégezni a döntési függvény közelítését.

3.1 A CAC két forgalmi osztály esetén

Ha kihasználjuk, hogy az $N^{(1)}$ és $N^{(-1)}$ halmazokat elválasztó hiperfelület konvex [3], viszonylag kevés forgalmi konfigurációban kiértékelve $g(\mathbf{n})$-t megkapható a pontos szeparáló felület. Ehhez definiáljunk egy $\mathbf{z}$ (egyelőre ismeretlen hosszú) vektort, ami a következők szerint legyen feltöltve:

1. $n_2 := 0, n_1 := 0, k := 0,$.
2. Növeljük n_1-et egyesével egészen addig, amíg $g(\mathbf{n}) = -1$ nem lesz. Az így kapott $n_1^* = n_1$ érték határozza meg $\mathbf{z}$ első elemét: $z_0 := n_1^* - 1$.
3. Növeljük 1-gyel n_2-t. Mivel ezzel a forgalmi állapot nagyobb forgalmat ír le, biztos, hogy továbbra is $g(\mathbf{n}) = -1$ lesz.
4. Csökkentsük egyesével n_1-et (z_{k-1}-ről) egészen addig, amíg $g(\mathbf{n}) = 1$ nem lesz újból. Az így kapott $n_1^* = n_1$ érték legyen $\mathbf{z}$ következő eleme: $z_k := n_1^*$.
5. Ha $z_k = 0$, akkor a teljes szeparáló felületet felderítettük és megállhatunk, máskülönben $k := k+1$ és ugorjunk a 3. ponthoz.

Az algoritmus során a következő utat járja be az $\mathbf{n}$ vektor:



3-1. ábra A kétdimenziós szeparáló felület feltérképezése

Ha $n_i^{(max)}$-szal jelöljük az egy forgalmi osztályból maximálisan beengedhető források számát, amikor a többi osztályból a források nem generálnak forgalmat, akkor látható, hogy összesen $K=2(n_1^{(max)}+1)n_2^{(max)}+1$ pontban kiértékelve $g(\mathbf{n})$-t, feltérképezzük a teljes szeparáló felületet. A $\mathbf{z}$ vektorral ekkor a következőképpen írható át $g(\mathbf{n})$:

$$g(n_1, n_2) = \begin{cases} 1 & \text{ha } n_1 \le z_{n_2} \\ -1 & \text{egyébként.} \end{cases} \quad (3\text{-}2)$$

A gyakorlatban általában 3 forgalmi osztály fordul elő (pl. voice over ADSL, internet access I, internet access II). Sajnos a fenti egyszerű módszerrel háromdimenziós szeparáló felület már nem deríthető fel. A következőkben egy közelítő eljárással úgy terjesztjük ki a módszert 3D-re, hogy a szeparáló felület meghatározásához szükséges pontszám ne növekedjen nagyságrendekkel. Ehhez először a közelítéssel szemben támasztott kritériumokat vizsgáljuk meg.

3.2 A közelítő CAC szempontjai

A CAC implementálása folyamán az $N^{(1)}$ és $N^{(-1)}$ halmazok $g(\mathbf{n})$ szerinti szeparálását az $f(\mathbf{n})$ osztályozó függvénnyel közelítjük:

$$f(\mathbf{n}) = \begin{cases} 1, & \text{ha } \mathbf{n} \in \tilde{N}^{(1)} \\ -1, & \text{ha } \mathbf{n} \in \tilde{N}^{(-1)}, \end{cases} \quad (3\text{-}3)$$

ahol a hullámjel a megoldás közelítő jellegére utal. Célunk, hogy az $f(\mathbf{n})$ által definiált $\tilde{N}^{(1)}$ elfogadási és az $\tilde{N}^{(-1)}$ visszautasítási halmazok minél jobban közelítsék az optimális $N^{(1)}$ és $N^{(-1)}$ részhalmazokat. Az optimalizálást úgy kell elvégezni, hogy minél több felhasználó legyen beengedve a hálózatba, minél nagyobb legyen a hálózat kihasználtsága (hiszen minél több felhasználó kommunikál, annál nagyobb a szolgáltató bevétele), miközben nem szabad elfogadni olyan hívást, amellyel megsérülnek a QoS paraméterek (az a szolgáltató szerződés megszegését jelentené, ami nem tolerálható). Ennek megfelelően a probléma az alábbi feltételes szélsőérték-keresési feladatként írható fel:

$$f_{opt} : \min_f \left| N^{(1)} - \tilde{N}^{(1)} \right| \text{ miközben } \tilde{N}^{(1)} \subseteq N^{(1)}. \quad (3\text{-}4)$$

A közelítés pontosságát a következő hatékonyságleírókkal mérjük:

1. Első fajú hiba – hálózati kihasználtság (Network Utilization)
   Az $f(\mathbf{n})$ által elfogadott és a $g(\mathbf{n})$ szerint elméletileg elfogadható hívások számának arányát méri:
   $$NU := |N^{(1)} \cap \tilde{N}^{(1)}| / |N^{(1)}| \ .$$
2. Másodfajú hiba – szerződésszegések aránya (Breach of contract)
   Az $f(\mathbf{n})$ által elfogadott, de a $g(\mathbf{n})$ szerint elméletileg visszautasítandó hívások és a $g(\mathbf{n})$ szerint visszautasítandó hívások számának arányát mutatja:
   $$BC := |\tilde{N}^{(1)} - N^{(1)}| / |N^{(-1)}| \ .$$

Ezek a definíciók értelemszerűen csak véges N állapottérnél érvényesek. Eddig elvben nem korlátoztuk az állapottér méretét, de nyilvánvaló, hogy ha az i-dik forgalmi osztályból összesen $n_i^{(max)}$ felhasználót engedhetnénk be önmagában, akkor az állapottér $n_i^{(max)}$ koordinátaértékeken kívüli részét teljesen értelmetlen vizsgálni, hiszen ahhoz mindenképpen –1-es (visszautasítás) döntés tartozna. Ennek megfelelően: $|N| = \prod_{i=1}^{M} n_i^{(\max)}$ ahol

$$n_i^{(\max)} = \max\left\{ n_i : g(\mathbf{n}) = 1, n_j = 0 \ \forall j \neq i \right\}. \qquad (3\text{-}5)$$

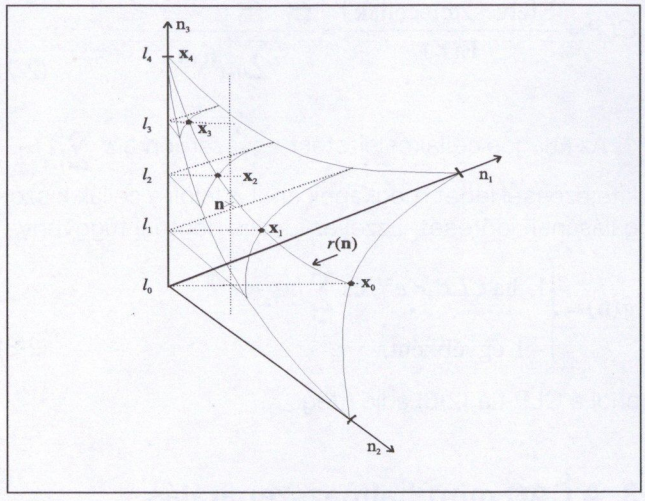
3.3 Kiterjesztés három forgalmi osztályra

A módszer alapötlete az, hogy az N tér harmadik koordinátáját felosztjuk rétegekre, és minden rétegen egy 2D-s ideális döntési függvénnyel képezzük a 3D-s szeparáló felület síkmetszeteit. A pontosabb leírás érdekében vezessük be a következő jelöléseket:

- Indexeljük úgy a 3 forgalmi osztályt, hogy $n_1^{(max)} \geq n_2^{(max)} \geq n_3^{(max)}$ legyen. Ekkor a 3. koordinátát fogjuk felosztani, hiszen ekkor kevesebb osztóponttal is pontos közelítés érhető el.
- A harmadik tengely osztópontjait a χ halmaz foglalja össze: $\chi = \{0, l_1, l_2, ..., l_H\}$. Ennek megfelelően $H+1$ réteget képeztünk.
- Minden egyes rétegen határozzuk meg az ideális $g_q(n_1,n_2) = g(\mathbf{n})\big|_{n_3=l_q}$ döntési függvényt a 3.1 alpontban adott módszerrel.

Ha a beérkező $\mathbf{n}$ állapotvektor egy adott rétegre esik, akkor a döntés egyértelmű. Ha két réteg között helyezkedik el, azaz $l_q < n_3 < l_{q+1}$, akkor a döntést képezzük interpolációval:

Metsszük el képzeletben a 3D-s szeparáló felületet egy a 3. koordinátán és az $\mathbf{n}$ vektor végpontján keresztülhaladó síkkal. Az így kapott szeparáló felület trajektóriát írja le a $\rho(\mathbf{n}) = 0$ egyenlet úgy, hogy a vele egy síkban fekvő alatta lévő pontokra pozitív, míg a felette lévő pontokra negatív legyen az értéke. Az egyes rétegeken a szeparáló felületeknek az $\mathbf{x}_0, \mathbf{x}_1, ..., \mathbf{x}_H$ első két koordinátájának iránya felé vett széleit jelöljék rendre az $\rho(\mathbf{n})$ vektorok, ahogy azt a 3-2. ábra szemlélteti. Feladatunk a $\rho(\mathbf{n})$ függvény közelítése interpolációval. Vagyis keressük azt a $\tilde{\rho}(\mathbf{n})$ függvényt, amelyre $\tilde{\rho}(\mathbf{x}_0) = 0$, $\tilde{\rho}(\mathbf{x}_1) = 0$, …, $\tilde{\rho}(\mathbf{x}_H) = 0$,.



3-2. ábra A 3D-s szeparáló felület közelítése rétegekkel

Minthogy az $\mathbf{x}_0, \mathbf{x}_1, ..., \mathbf{x}_H$ vektorok egy síkban vannak, így ez egy egyváltozós interpolációs feladat. Ennek megfelelően a 3D-s szeparáló felület közelítése a következő módon áll elő:

$$f_{3D}(\mathbf{n}) = \begin{cases} 1, & \text{ha} \quad \tilde{\rho}(\mathbf{n}) > 0 \\ -1, & \text{ha} \quad \tilde{\rho}(\mathbf{n}) \leq 0 \end{cases}. \qquad (3\text{-}6)$$

A zérusrendű interpoláció az $\mathbf{x}_0, \mathbf{x}_1, ..., \mathbf{x}_H$ pontok lépcsőzetes görbével való összekötését jelenti. Vagyis akkor fogadjuk el az $\mathbf{n}$ hívást, ha a fölötte lévő rétegen a döntés szintén „elfogadás":

$$f_{3D}(\mathbf{n}) = \begin{cases} 1, & \text{ha} \quad g_{q+1}(n_1, n_2) = 1 \\ -1 & \text{egyébként,} \end{cases} \qquad (3\text{-}7)$$

ahol $l_{q+1} > n_3 > l_q$.

A módszer előnye a döntés egyszerűsége mellett az is, hogy az interpolációs hiba csak az első fajú hibákat növeli. Ellenben a magas kihasználtsághoz nagyon sűrűn kellene felosztani a harmadik koordinátát, ezért célszerű magasabb rendű interpolációt alkalmazni.

Az elsőrendű interpolációnál az $\mathbf{x}_0, \mathbf{x}_1, ..., \mathbf{x}_H$ pontokat egyenes szakaszok kötik össze. A másodfajú hibák csökkentése végett legyenek ezek a határpontok úgy definiálva, hogy a hozzájuk tartozó $g(\mathbf{x}_r)$ $(r = 0,1,...,H)$ döntések mindegyike –1. Az elsőrendű interpoláció döntési függvénye részletezés nélkül:

$$f_{3D} = \begin{cases} 1, & \text{ha} \quad g_{q+1}(n_1, n_2) = 1 \\ -1, & \text{ha} \quad g_q(n_1, n_2) = -1 \\ \operatorname{sgn}\left(t_0 x_{q+1,3} + (1 - t_0) x_{q,3} - n_3\right) & \text{egyébként,} \end{cases} \qquad (3\text{-}8)$$

ahol

$$t_0 = \frac{|\mathbf{x}_q - \mathbf{n}|^2 - |\mathbf{n} - \mathbf{x}_{q+1}|^2 + |\mathbf{x}_{q+1} - \mathbf{x}_q|^2}{2|\mathbf{x}_{q+1} - \mathbf{x}_q|^2}. \qquad (3\text{-}9)$$

A részletes leírás megtalálható [4]-ben.

Habár ennél a módszernél a másodfajú hibákat is növeli az interpolációs hiba, a gyakorlatban olyan kicsi a szeparáló felület görbülete, hogy gyakorlatilag elhanyagolható romlás lép fel, miközben a kihasználtság számottevően megnő a lépcsőzetes közelítéshez képest.

Megjegyezzük, hogy a módszer hasonló elven elvileg kiterjeszthető háromnál magasabb dimenziószámra is. Ekkor azonban a számítási igény jelentősen megnő, ezért háromnál több forgalmi osztály esetén érdemesebb inkább véletlenszerűen sorsolt tanító minták alapján képezni a közelítő döntési függvényt, mint a következő részben tárgyalt nem parametrikus esetnél.

4. Nem parametrikus CAC

Ha ismeretlenek a források tulajdonságai és nincs túl szigorú cellavesztési arány előírva, akkor véletlenszerűen választott forgalmi konfigurációkat választva az elérési modult tesztüzemben működtetve a cellavesztési arány mérése alapján dönthetünk az adott konfiguráció elfogadásáról vagy visszautasításáról. Ebben az esetben nincs lehetőség a (3-4) feladat pontos megoldására, mert semmilyen információnk nincs a tényleges szeparáló felületről. Szorítkozzunk a hibás döntések számának minimalizálására, vagyis (3-4) helyett az

$$E(\mathbf{w}) = \frac{1}{2}\sum_{\mathbf{n}_k \in N} |f(\mathbf{n}_k) - d_k| p(\mathbf{n}_k) =$$

$$= \frac{1}{2}\sum_{\mathbf{n}_k \in N^{(1)}} (f(\mathbf{n}_k) - 1) p(\mathbf{n}_k) + \frac{1}{2}\sum_{\mathbf{n}_k \in N^{(-1)}} (1 - f(\mathbf{n}_k)) p(\mathbf{n}_k), \quad (4\text{-}1)$$

hibavalószínűséget minimalizáljuk, ahol $p(\mathbf{n})$ jelöli az $\mathbf{n}$ forgalmi állapot előfordulási valószínűségét:

$f_{opt} = \arg\min_f E(\mathbf{w})$. Az a feladat, hogy ezt az f_{opt}-ot mérések alapján közelítsük. A mérést K forgalmi állapotvektorra elvégezve egy τ mintahalmaz állítható össze a következő módon:

$$\tau = \{(\mathbf{n}_k, d_k), k = 1,2,...,K\}, \quad (4\text{-}2)$$

ahol d_k az $\mathbf{n}_k$ forgalmi konfigurációnál mért döntés. Elméletileg $d_k = g(\mathbf{n}_k)$ lenne, de a mérési pontatlanság miatt a $d_k = \mathrm{sgn}\{g(\mathbf{n}_k)+\nu(\mathbf{n}_k)\}$ modellt alkalmazhatjuk, ahol ν valamilyen zajt jelölő valószínűségi változó. Jelölje $f(\mathbf{n},\mathbf{w})$ a τ alapján előálló döntési függvényt, ahol a $\mathbf{w}$ vektorba összefoglalt paramétereket úgy kell megválasztani, hogy a (4-1) hibavalószínűséget minimalizáljuk. ($f(\mathbf{n},\mathbf{w})$ jelölheti bármilyen univerzális approximátor – pl. egy előrecsatolt neurális hálózat – átviteli függvényét. A megoldáshoz az [5]-ben ismertetett nem parametrikus döntéselméleti eredményeket alkalmazzuk.

4.1 Az optimális döntési függvény megválasztása

Célunk a (4-1) minimalizálása, az $f_{opt}(\mathbf{n}) = f(\mathbf{n},\mathbf{w}_{opt})$ osztályozó függvény megtalálása, ahol az optimális súlyvektor: $\mathbf{w}_{opt} = \arg\min_{\mathbf{w}} E(\mathbf{w})$. Mivel $E\mathbf{w}$ kiértékeléséhez ismerni kellene az $N^{(1)}$ és az $N^{(-1)}$ halmazokat, vagyis a pontos szeparáló felületet, így $\mathbf{w}_{opt}$-ot a τ mintahalmaz alapján közelítjük. A következő tétel ennek módjáról szól.

1. tétel: Legyen K számú $\mathbf{n}_k$ forgalmi állapotban megmérve a döntés. A mérési pontatlanságot zajként felfogva $d_k = \mathrm{sgn}\{g(\mathbf{n}_k)+\nu(\mathbf{n}_k)\}$, ahol ν ismeretlen, nulla várható értékű valószínűségi változó. Foglalja össze $\tau^{(K)}$ a megfelelő $(\mathbf{n}_k, d_k)$ párokat. Végezzük el a $\tau^{(K)}$ feletti

$$\tilde{E}(\mathbf{w},\tau^{(K)}) = \frac{1}{2K}\sum_{\{\mathbf{n}_k | d_k = 1\}} (1 - f(\mathbf{n}_k)) + \frac{1}{2K}\sum_{\{\mathbf{n}_k | d_k = -1\}} (1 + f(\mathbf{n}_k)) \quad (4\text{-}3)$$

empirikus hibafüggvény minimalizálását:

$\mathbf{w}^{(K)} = \arg\min_{\mathbf{w}} \tilde{E}(\mathbf{w},\tau^{(K)})$.

Ha

- az $\mathbf{n}_k$ mérési állapotok a $p(\mathbf{n}_k)$ eloszlással vannak sorsolva és
- $f(\mathbf{n},\mathbf{w})$ képes tetszőlegesen pontosan közelíteni bármilyen konvex $\gamma(\mathbf{n})$ osztályozó függvényt,

$$\exists \mathbf{w}^* : \left\| f(\mathbf{n},\mathbf{w}^*) - \gamma(\mathbf{n}) \right\| < \varepsilon \ (\forall \varepsilon > 0),$$

akkor $\lim_{K\to\infty} \mathbf{w}^{(K)} = \mathbf{w}_{opt}$ és az $f_{opt}(\mathbf{n}) = f(\mathbf{n},\mathbf{w}_{opt})$ osztályozó függvény a (4-1) hibavalószínűséget minimalizálja. Vagyis az $f(\mathbf{n},\mathbf{w}^{(K)})$ függvény aszimptotikusan optimális.

Bizonyítás:

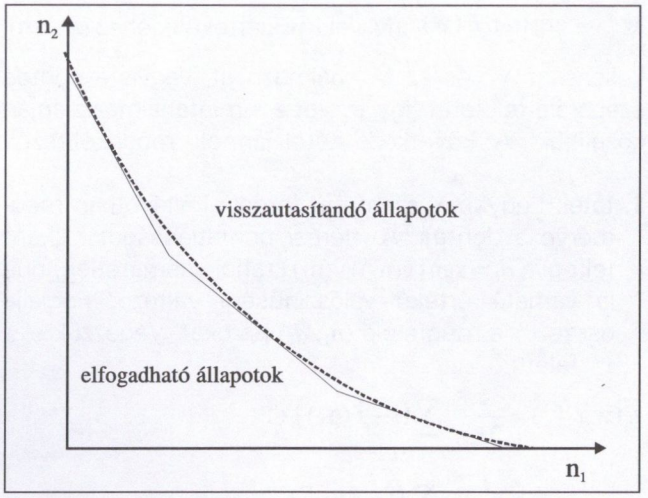
Tudjuk, hogy ha a hibavalószínűséget a relatív gyakorisággal becsüljük a mérés során, akkor $\mathrm{sgn}(\mathbf{E}\{d_k\})+g(\mathbf{n}_k)$. Ekkor belátható, hogy ha a mérési állapotok a $p(\mathbf{n}_k)$ eloszlással vannak sorsolva, akkor az ergodikus hipotézis és a Vapnik–Chervonenkis-tétel [6] értelmében igaz a következő határérték:

$$\lim_{K\to\infty} \min_{\mathbf{w}} \tilde{E}(\mathbf{w},\tau^{(K)}) = \min_{\mathbf{w}} E(\mathbf{w}), \quad (4\text{-}4)$$

ami maga után vonja, hogy $f(\mathbf{n},\mathbf{w}^{(K)}) \to f(\mathbf{n},\mathbf{w}_{opt})$. Mivel $f(\mathbf{n},\mathbf{w})$ képes tetszőleges konvex osztályozó függvény közelítésére: $f_{opt}(\mathbf{n}) = f(\mathbf{n},\mathbf{w}_{opt})$.

4.2. A döntési architektúra: a „golyós" szeparátor

Tudjuk, hogy a $g(\mathbf{n})$ osztályozó függvény egy konvex szeparáló felületet definiál az N állapottérben, amely poligoniálisan hipersíkokkal közelíthető úgy, hogy az $\tilde{N}^{(1)} \subseteq N^{(1)}$ kritérium teljesüljön (ld. 4-1. ábra).



4-1. ábra A szeparáló felület poligoniális közelítésének 2D-s szemléltetése

Az i-dik hipersíkot az $\mathbf{x}^T\mathbf{w}^{(i)}+c=0$ egyenlettel lehet leírni, ahol $c>0$ konstans. Ennek megfelelően M hipersík esetén a következőképpen áll elő a döntési függvény:

$$f(\mathbf{n},\mathbf{w}) = \bigvee_{i=1}^{M} \operatorname{sgn}(\mathbf{n}^T\mathbf{w}^{(i)} + c), \tag{4-5}$$

ahol V a logikai „vagy" kapcsolatot jelenti –1-es „L" szint és +1-es „H" szint mellett.

1. lemma: Létezik olyan M síkszám, hogy

$\min_{\mathbf{w}} \|f(\mathbf{n},\mathbf{w}) - g(\mathbf{n})\| < \varepsilon$ tetszőleges $\varepsilon > 0$-ra,

vagyis a (4-5) osztályozó függvény képes az ideális döntési függvény tetszőlegesen pontos közelítésére.

Az állítás triviális, hiszen az ideális szeparáló felületről tudjuk, hogy konvex, és konvex felületek tetszőleges pontossággal közelíthetők poligoniálisan a hipersíkok számának növelésével.

A (4-5)-öt (4-3)-ba helyettesítve rögtön látszik, hogy az $\tilde{E}(\mathbf{w},\tau)$ nem folytonos függvénye $\mathbf{w}$-nek, gradiens típusú módszerrel nem minimalizálható. A probléma feloldásához alakítsuk át (4-5)-öt a következő alakúra:

$$f(\mathbf{n},\mathbf{w}) = \bigvee_{i=1}^{M} \operatorname{sgn}\left(\left|\mathbf{n}-\mathbf{w}^{(-i)}\right|^2 - \left|\mathbf{n}-\mathbf{w}^{(+i)}\right|^2\right). \tag{4-6}$$

Az átalakítás lényege, hogy a $\mathbf{w}^{(+i)}$ és a $\mathbf{w}^{(-i)}$ vektorok szintén egy hipersíkot definiálnak, mert a végpontjuktól egyenlő távolságra lévő pontok halmaza hipersík. Figyeljük meg, hogy teljesül a következő állítás:

2. lemma: Ha a szeparálófelület az $N^{(1)}$ térrészből nézve konvex, akkor az

$$f(\mathbf{n},\mathbf{w}) = \bigvee_{i=1}^{M} \operatorname{sgn}\left(\left|\mathbf{n}-\mathbf{b}\right|^2 - \left|\mathbf{n}-\mathbf{w}^{(+i)}\right|^2\right). \tag{4-7}$$

osztályozó függvénnyel tetszőleges poligoniális közelítés megvalósítható, ahol $\mathbf{b} \in N^{(1)}$ konstans vektor és $\mathbf{w}$ a $\mathbf{w}^{(+i)}$ súlyvektorokat összefogó vektor.

Bizonyítás

Mivel a $\mathbf{b}$ vektor az $N^{(-1)}$ térrészben helyezkedik el, az i-dik hipersíkra vett tükörképük az $N^{(1)}$-ben van. Ha a pontokat ezeknek a tükörképeknek feleltetjük meg, akkor a $(\mathbf{b}, \mathbf{w}^{(+i)})$ pontpároktól egyenlő távolságra lévő pontok halmaza pontosan az i-dik síkot adja. Mivel ez a tükrözés tetszőleges poligoniális közelítésnél végrehajtható, az állítás igaz.

A továbbiakban a $\mathbf{w}^{(+i)}$-re, mint az i-dik hipersík bázispontjára fogunk hivatkozni. Figyeljük meg, hogy egy adott beérkező $\mathbf{n}$ esetén a döntést a hozzá legközelebb eső bázispont fogja meghatározni. Ezzel (4-7) a következőképpen írható át:

$$f(\mathbf{n},\mathbf{w}) = \operatorname{sgn}\left(\left|\mathbf{n}-\mathbf{b}\right|^2 - \left|\mathbf{n}-\mathbf{w}^{(+i)}\right|^2\right)\Big|_{i=\arg\min_{j=1,\ldots,M}\left|\mathbf{n}-\mathbf{w}^{(+j)}\right|} \tag{4-8}$$

Az $\tilde{E}(\mathbf{w},\tau)$ folytonossá tételéhez jó ötletnek tűnik, hogy a (4-8)-ban szereplő „kemény" nemlinearitást „lággyá" tegyük pl. egy szigmoid jellegű görbével:

$$\hat{f}(\mathbf{n},\mathbf{w},\alpha) := \varphi^{(\alpha)}\left(\left|\mathbf{n}-\mathbf{b}\right| - \left|\mathbf{n}-\mathbf{w}^{(+i)}\right|\right),$$

ahol

$$\varphi^{(\alpha)}(z) = \frac{2}{1+e^{-\alpha z}} - 1. \tag{4-9}$$

Nyilvánvaló, hogy ha $\alpha\to\infty$, akkor $\hat{f}(\mathbf{n},\mathbf{w},\alpha)\to f(\mathbf{n},\mathbf{w})$. Az viszont korántsem egyértelmű, hogy $\hat{f}(\mathbf{n},\mathbf{w},\alpha)$-át (4-3)-ba helyettesítve folytonos függvényt kapunk.

2. tétel: Az

$$\hat{E}(\mathbf{w},\tau^{(K)},\alpha) = \frac{1}{2K}\sum_{\{\mathbf{n}_k|d_k=1\}}\left(1-\hat{f}(\mathbf{n}_k,\mathbf{w},\alpha)\right) + \frac{1}{2K}\sum_{\{\mathbf{n}_k|d_k=-1\}}\left(1+\hat{f}(\mathbf{n}_k,\mathbf{w},\alpha)\right)$$

hibafüggvény folytonos és

$$\mathbf{w}_{\text{opt}} = \lim_{K\to\infty}\lim_{\alpha\to\infty}\arg\min_{\mathbf{w}}\hat{E}(\mathbf{w},\tau^{(K)},\alpha). \tag{4-10}$$

Bizonyítás

A tétel második állítása közvetlenül adódik a (4-9) szigmoidfüggvény határérték-tulajdonságából és az 1. tételből (ahol az 1. és a 2. lemma biztosítja, hogy (4-7) megfelelő osztályozó függvény). Az első állítás bizonyításához bontsuk szét $\tau^{(K)}$-t diszjunkt halmazokra a következő rekurzióval:

$\tau_0 := \varnothing$ és

$$\tau_i := \left\{(\mathbf{n}_k,d_k)\in\tau^{(K)}\setminus\bigcup_{l=0}^{i-1}\tau_l : \left|\mathbf{n}_k-\mathbf{w}^{(+i)}\right|^2 \le \left|\mathbf{n}_k-\mathbf{w}^{(+j)}\right|^2 \forall j\right\}, \tag{4-11}$$

A $i=1,2,\ldots,M$ részhalmazok a τ_i bázisponthoz legközelebb eső mintákat tartalmazzák. Ezzel (4-10) az alábbi alakra hozható:

$$\hat{E}(\mathbf{w},\tau^{(K)}) = \sum_{i=1}^{M}\sum_{\mathbf{n}_k\in\tau_i(\mathbf{w})}\psi_i(\mathbf{n}_k,\mathbf{w}), \tag{4-12}$$

ahol

$$\psi_i(\mathbf{n}_k,\mathbf{w}) := \begin{cases} \frac{1}{2K}\left(1-\varphi\left(\left|\mathbf{n}_k-\mathbf{b}\right|^2-\left|\mathbf{n}_k-\mathbf{w}^{(+i)}\right|^2\right)\right), & \text{ha } d_k=1 \\ \frac{1}{2K}\left(1+\varphi\left(\left|\mathbf{n}_k-\mathbf{b}\right|^2-\left|\mathbf{n}_k-\mathbf{w}^{(+i)}\right|^2\right)\right), & \text{ha } d_k=-1. \end{cases} \quad (4\text{-}13)$$

Nyilvánvaló, hogy $\mathbf{w}$-nek abban a tartományában, amelyben a $\tau_i(\mathbf{w})$-k állandók, (4-12) folytonos. Most tegyük fel, hogy a $\mathbf{w} = (\mathbf{w}^{(+1)}\ \mathbf{w}^{(2)}\ \ldots\ \mathbf{w}^{(+M)})$ súlyvektor $\Delta\mathbf{w}$-vel történő elmozdításával a $\tau^{(K)}$ halmaz q-adik eleme átkerül τ_i-ből τ_j-be, miközben a többi halmaz változatlan marad, azaz

$$\begin{aligned} &\tau_i(\mathbf{w}+\Delta\mathbf{w}) = \tau_i(\mathbf{w}) \setminus \{(\mathbf{n}_q, d_q)\} \\ &\tau_j(\mathbf{w}+\Delta\mathbf{w}) = \tau_j(\mathbf{w}) \cup \{(\mathbf{n}_q, d_q)\},\ j \neq i \\ &\tau_l(\mathbf{w}+\Delta\mathbf{w}) = \tau_l(\mathbf{w})\ \ \forall l \neq i, j. \end{aligned} \quad (4\text{-}14)$$

Ekkor léteznie kell egy olyan $0<c<1$ konstansnak, hogy a $\hat{\mathbf{w}}=\mathbf{w}+c\Delta\mathbf{w}$ helyen $\left|\mathbf{n}_q-\hat{\mathbf{w}}^{(+i)}\right|^2=\left|\mathbf{n}_q-\hat{\mathbf{w}}^{(+j)}\right|^2$, ami maga után vonja, hogy $\psi_i(\mathbf{n}_q, \hat{\mathbf{w}}) = \psi_j(\mathbf{n}_q, \hat{\mathbf{w}})$, vagyis a $\tau_i(\mathbf{w})$ halmazok módosulása az $\hat{E}(\mathbf{w},\tau^{(K)})$ hibafüggvényben nem okoz ugrást; az átmenet folytonos marad.

A gyakorlatban $\hat{E}(\mathbf{w},\tau^{(K)},\alpha)$ minimalizálására a

$$\mathbf{w}[k+1] = \mathbf{w}[k] - \delta \cdot \operatorname{grad} \hat{E}(\mathbf{w},\tau^{(K)}) \quad (4\text{-}15)$$

iterációt alkalmazzuk, ahol $\delta>0$ konstans és $\hat{E}(\mathbf{w},\tau^{(K)})$ analitikusan kifejezhető. Sajnos az iménti bizonyításból kiderül, hogy $\operatorname{grad} \hat{E}(\mathbf{w},\tau^{(K)})$ nem létezik a $\tau_i(\mathbf{w})$ halmazok módosulási helyeinél, mert a bal és jobb oldali parciális deriváltak nem azonosak. Mivel egy minta (4-11) miatt csak egy τ_i-ben lehet, így az egyik oldali parciális deriváltak mindig léteznek és ezekkel (4-14) működőképes marad. A minimum közelében fellépő oszcilláció a δ tanulási tényező fokozatos csökkentésével csillapítható.

A (4-12) hibafüggvény gradiens módszerrel történő minimalizálásakor problémát jelenthet a lokális minimumokban való megakadás. Az eddigi tapasztalatok azt mutatták, hogy kicsi α értékeknél ($\alpha \approx 5...50\text{K}/|N|$) több véletlenszerűen sorsolt kezdeti súlyvektorból indítva a minimalizálást, közel azonos minőségű megoldásokat kapunk, ami arra utal, hogy kicsi α értékeknél csökken a lokális minimumok száma. (A jelenség pontos okának kiderítése további kutatásokat igényel.) A pontos optimalizáláshoz viszont minél nagyobb α értékre van szükség. Az iménti ellentmondás úgy oldható fel, hogy δ csökkentésével α-t növeljük az iterációs lépések során egészen addig, amíg

$$\left|\hat{E}(\mathbf{w}[k],\tau^{(K)},\alpha) - \tilde{E}(\mathbf{w}[k],\tau^{(K)})\right| < 1/K$$

nem lesz. Ekkor α növelésével $\tilde{E}$ nem csökkenthető tovább, minthogy $\tilde{E}$ legkisebb változása $1/K$ lehet csak.

Nagy előnye a (4-7) osztályozó függvénynek, hogy a $\mathbf{w}^{(i+)}$ bázispontok az állapottérben helyezkednek el, ezért könnyebb a kezdeti értékeket beállítani. A (4-14) súlybeállítási, tanulási folyamat során úgy is értelmezhetők ezek a bázispontok, mintha egy-egy golyó koordinátáját írnák le, amik „begurulnak" az optimumba. Ezért adtam a golyós szeparátor elnevezést az döntési architektúrának.

A (4-3) hibafüggvény minimalizálásakor a másodfajú hibák nincsenek megkülönböztetve. A numerikus eredményekből viszont kiderül, hogy így is elfogadható megoldásokat kapunk. Már számos módszer született a másodfajú hibák „büntetésére" (pl. büntetőfüggvényes tanulás [7], átlagos veszteségminimalizálás [5]), amelyek könnyen átültethetők a golyós szeparátoros optimalizálásra is.

5. Numerikus eredmények és konklúziók

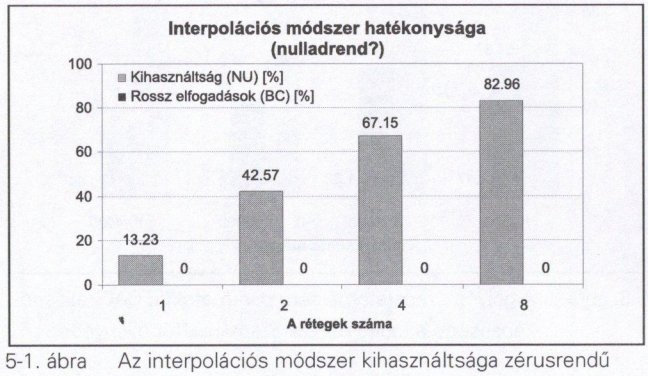
A kidolgozott módszerek hatékonyságát a következő 3 forgalmi osztály esetén vizsgáltuk:

	m_i [kbit/s]	h_i [kbit/s]	$n_i^{(\max)}$
Voice over DSL	24	40	404
Internet access I	20	384	226
Internet access II	20	2048	34

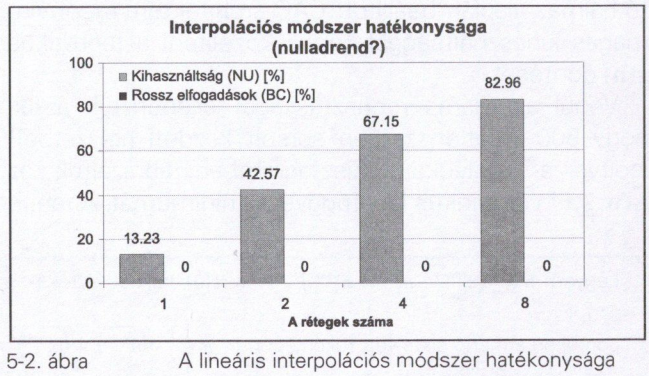
A linkkapacitás $C=11.52\text{Mbit/s}$, míg a CLP korlát 10^{-6} volt. A táblázat utolsó oszlopa az egy forgalmi osztályból önmagában beengedhető felhasználók számát mutatja. Mivel a hatékonyságadatokban nem mutatkozott számottevő eltérés a pufferes és a ZBA modell között, így a gyorsabban kiértékelhető ZBA (puffer nélküli) modellt használtam a szimulációk során.

5.1. Háromdimenziós eset: interpolációs módszer

Az 5-1. és 5-2. ábrák a zérusrendű és az elsőrendű interpolációval kapott hatékonyságértékeket mutatják az állapottér harmadik dimenzió felosztásának függvényében.



5-1. ábra Az interpolációs módszer kihasználtsága zérusrendű esetben a rétegszám függvényében

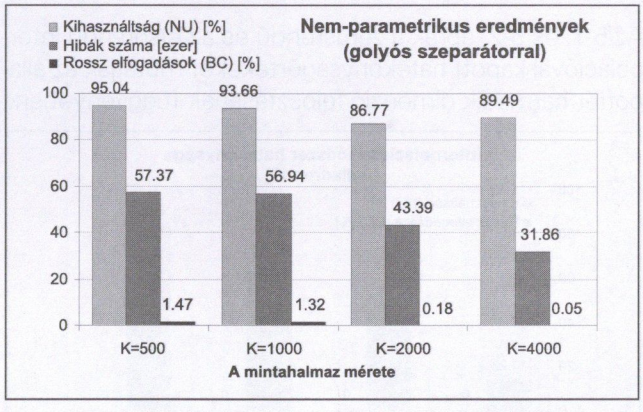


5-2. ábra A lineáris interpolációs módszer hatékonysága

Megfigyelhető, hogy a zérusrendű interpolációnál valóban nem lépnek fel másodfajú hibák, viszont a magas kihasználtsághoz sok réteget kell alkalmazni, ami jelentősen megnöveli a szeparáló felület közelítéséhez szükséges időt. Elsőrendű interpolációnál a másodfajú hibák fellépése jelenti a legfőbb gondot, mivel a kihasználtság a rétegszámtól függetlenül magas. A rétegszám növelésével a rossz elfogadások aránya is jelentősen csökken.

5.2. A golyós szeparátor teljesítőképessége

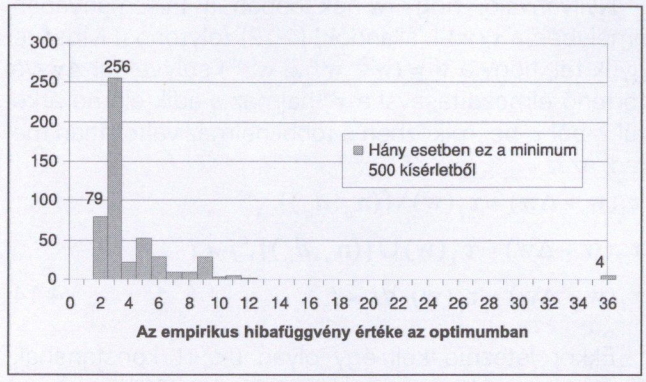
Ebben az esetben a nem parametrikus esetnek megfelelően egy tanító mintahalmazt sorsoltunk egyenletes eloszlással az állapottérből. A döntés mérésének szimulációjához feltettük, hogy 10^6 cella átvitelét figyeljük. Mivel a CLP korlát 10^{-6}, ezért akkor legyen az aktuálisan mért forgalmi konfiguráció elfogadható, ha a 10^6 cellából egyetlenegy sem veszik el. Ennek a valószínűsége $\mu = (1-\mathrm{CPL})^{10^6} \approx e^{-\mathrm{CPL}\cdot 10^6}$. On/off modell mellett a CLP kiszámítható (ld. 2.2.1. alpont), így az $\mathbf{n}_k$ forgalmi állapotvektorhoz sorsoljuk d_k-t a következő eloszlás szerint: $P(d_k = -1) = \mu$ és $P(d_k = 1) = 1-\mu$. Az így felvett mintahalmaz jól szimulálja a mérési pontatlanságot. A numerikus analízis során először azt vizsgáltuk, hogy a mintahalmaz méretének függvényében hogyan javul az eljárás teljesítőképessége. Ahogy az 5-3. ábrán látható, a hibák száma egyértelműen csökkent a mintaszám növelésével, míg a másodfajú hibák száma hamar elhanyagolhatóvá vált.



5-3. ábra A golyós szeparátoros nem parametrikus CAC teljesítőképessége a tanító minták számának függvényében

Megfigyelhető, hogy csupán a méréssel felvett tanító halmaz alapján beállított CAC architektúra mennyire magas kihasználtsággal képes közelíteni a teoretikus $g(\mathbf{n})$ döntését.

Végül az eljárás robusztusságát ellenőriztük azzal, hogy 500 véletlenszerűen sorsolt kezdeti helyzetből[1] indítva a „golyókat" hisztogramban ábrázoltuk az $\tilde{E}(\mathbf{w}_{\mathrm{opt}}, \boldsymbol{\tau}^{(K)})$ empirikus hibafüggvény minimumát. A tanító halmaz mérete $K = 500$ volt. Az 5-4. ábrából kiderül, hogy alig játszik szerepet a kiindulási érték, ami arra utal, hogy a lokális minimumok hatását a kis α értékről való indítással jelentősen le lehetett csökkenteni. Az esetek 82%-ában csak 5, vagy annál kevesebb hiba lépett fel a tanító halmazon.



5-4. ábra Az empirikus hibafüggvény minimalizálásával kapott optimuma hisztogramban ábrázolva 500 különböző kezdeti érték esetén

A hagyományos előrecsatolt neurális hálózat jóval nagyobb érzékenységet mutat a kezdeti értékekre, így alkalmazásuk nehézkesnek tűnik. A golyós szeparátornál ez a probléma nem vetődik fel, a tanulási időt kézben lehet tartani. A jövőben a jelenség elméleti magyarázatát kell megtalálni. A bemutatott szimulációs eredmények azt mutatják, hogy ezt a speciális halmazszeparálási problémát kimagaslóan jól oldja meg a „golyós" szeparátor. Ha a forrásparaméterek változnak, akkor új mintahalmazt kell felvenni. Ha a CAC készenlétbe állására a mintahalmaz felvétele után kevesebb idő áll rendelkezésre, mint a golyós szeparátor tanulási ideje, akkor a halmazszeparálás gyors analóg számítógépekkel (celluláris neurális hálózattal) is megoldható [4] szerint.

6. Összefoglalás és konklúziók

A cikk a [3,7,8] cikkekhez hasonlóan halmazszeparálási problémára vezette vissza a CAC feladatát. Parametrikus esetben két forgalmi osztályesetén gyorsan meghatározható az ideális elválasztófelület, ezért a gyakorlatban sokkal fontosabb három forgalmi osztályos esetet interpolációs technikával visszavezettük több két dimenziós halmazszeparálásra, amivel viszonylag egyszerű számításigénnyel magas hálózati kihasználtságot értünk el.

Nem-parametikus esetben a [7,8] cikkek előrecsatolt neurális hálózattal közelítették az elválasztófelületet. Azonban ezek hosszadalmas tanítási folyamata megakadályozza a gyors adaptáció lehetőségét és vál-

[1] Legyen $\mathbf{b} = \frac{1}{3}\left(n_1^{(\max)} \quad n_2^{(\max)} \quad n_3^{(\max)}\right)$. A $\mathbf{w}^{(+i)}[0]$ kezdeti értékek a $\mathbf{b}$ vektornak az $(\mathrm{n}_1^{(\max)},0,0)$, $(0,\mathrm{n}_2^{(\max)},0)$ és az $(0,\mathrm{n}_3^{(\max)},0)$ koordinátájú pontokon átmenő síkra vett tükörképének $R = \frac{1}{6}\sqrt{\left(n_1^{(\max)}\right)^2 + \left(n_2^{(\max)}\right)^2 + \left(n_3^{(\max)}\right)^2}$ sugarú környezetében lettek egyenletesen sorsolva.

tozó forgalomhoz. Éppen ezért kidolgoztunk egy egyszerűbb döntési architektúrát (a „golyós" szeparátort), ami úgyképes magas hálózati kihasználtságot biztosítani, hogy jelentősen lerövidül a tanítási folyamat, mert a golyós szeparátor energiafüggvényében a lokálisminimumok hatása számottevően lecsökken. Ezzel a tanulási időt kézben lehet tartani. A jövőben a jelenség elméleti magyarázatát kell megtalálni. A bemutatott szimulációs eredmények azt mutatják, hogy ezt a speciális halmazszeparálási problémát kimagaslóan jól oldja meg a „golyós" szeparátor.

Ha a forrásparaméterek változnak, akkor új mintahalmazt kell fölvenni. Ha a CAC üzembehelyezésére a mintahalmaz felvétele után kevesebb idő áll rendelkezésre, min a golyós szeparátor tanulási ideje, akkor a halmazszeparálás gyors analóg számítógépekkel (celluláris neurális hálózattal) is megoldható [4] a hálózati kihasználtság kis romlása árán.

Irodalom

[1] Levendovszky, J.: „Validation of novel CAC algorithms", ICAM- IEEE 1999, pp. 195–211.

[2] Levendovszky, J.: „Call admission control of ATM networks based on modulated Markov chains", Journal on Communication dedicated to ATM Networks, VOL. XLVII, pp. 19–24, March 1995.

[3] Hiramatsu, A.: ATM Traffic Control Using Neural Networks, Neural Networks in Telecommunications, edited by Yuhas, B. and Ansari, N., Kluwer Academic Publishers, 1994.

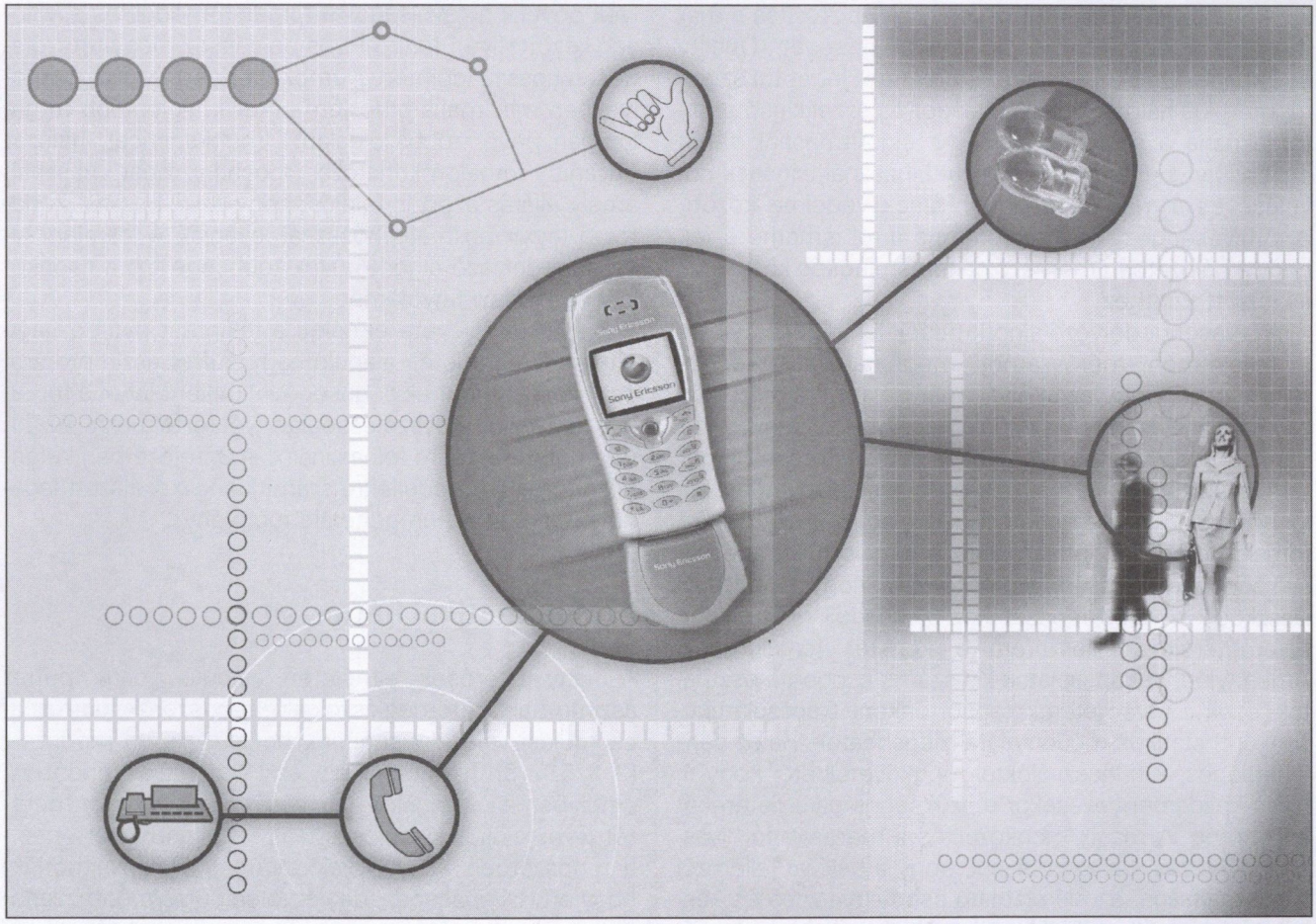
[4] J. Levendovszky, A. Fancsali: „Real-time call admission control for packet switched networking by cellular neural networks", Submitted to IEEE/CAS, April 2002.

[5] Fancsali A.: „Back propagation típusú neurális hálók alkalmazása átlagos veszteség minimalizálására", BME Végzős konferencia, 1999. április 28.

[6] V. Vapnik: The Nature of Statistical Learning Theory, Springer, 1995.

[7] Levendovszky J., Mészáros Á.: „Neuron based penalty function classifiers", Proceedings, 19th WIC Conference, Veldhoven, Holland, May, 1998.

[8] Levendovszky, J., Imre, S., Pap, L., E.C. van der Meulen, Varga, B.: Comparative Analysis of Call Admission Control Algorithms for ATM Networks, Annual Scientific Progress Report, COPERNICUS C579, August, 1995.



Egy nagy eltérések elméletén alapuló hívásengedélyezési algoritmus analitikus megközelítése

JAKABFY TAMÁS

Eötvös Loránd Tudományegyetem, Természettudományi Kar,

SZLÁVIK ÁRPÁD, SERES GERGELY

Ericsson Magyarország Kommunikációs Rendszerek Kft.
{Tamas.Jakabfy, Arpad.Szlavik, Gergely.Seres}@eth.ericsson.se

Reviewed

A cikk célja olyan forgalommodellek analitikus vizsgálata, amelyek figyelembe veszik a korlátozottságot. Ilyen típusú modellek vizsgálata annak érdekében történik, hogy olyan új hívásengedélyezési szabályokat lehessen konstruálni, amelyekben a hálózat kihasználtsága növelhető a szolgáltatásminőség megtartása mellett.

1. Bevezetés

A hívásengedélyező (Connection Admission Control, CAC) algoritmus a legfontosabb erőforrás-kezelő eljárások egyike. A szerepe hálózat felé továbbított forgalomkorlátozás annak érdekében, hogy biztosítsa a már meglévő kapcsolatok szolgáltatásminőségét (Quality of Service, QoS). Ez az algoritmus nem lehet túl szigorú, mert kis hálózatkihasználtságot és ezzel kieső profitot okozna a szolgáltatónak. Az algoritmusnak tehát egyensúlyt kell találnia a hálózat kihasználtságának növelése és a már elfogadott hívások védelme között. A jó hívásengedélyező algoritmus arról ismerhető fel, hogy a hálózatban mért szolgáltatásminőség közel esik a kitűzött szinthez.

A hívásengedélyező algoritmusok lényegét a hálózati teljesítmény-mérőszámok meghatározása képezi, feltételezve, hogy az új hívás már bekerült a rendszerbe. Ez a felhasznált kiszolgáló kapacitástól vagy a várható csomagvesztési aránytól és a késleltetéstől függ.

A hívásengedélyezési feladatnak két alapvetően különböző megközelítése létezik. Az első módszer szerint forgalommodelleket használunk a forgalom általános viselkedésének leírására, ami feltételezi, hogy a modell jól illeszkedik a valós rendszerhez. A modell mint sablon paraméterekhez illeszthető a konkrét forgalomhoz. Amennyiben a kapcsolatok kis számú csoportra oszthatók, ezek a paraméterek csoportonként (kapcsolattípusonként) statikus módon előre megadhatók. Ha ez nem lehetséges, a felhasználóktól megkövetelhető, hogy új hívás kezdeményezésekor a szükséges paramétereket adják meg. A másik megközelítés a mérésalapú hívásengedélyezés. Ez azon alapul, hogy a hálózat jellemzői (kihasználtsága, az előrelátható csomagvesztés és késleltetés) becsülhetők a forgalom statisztikai tulajdonságainak mérésével. Ebben az esetben nincs szükség sem forgalommodellre, sem paramétermegfeleltetésre.

Az aszinkron átviteli mód (Asynchronous Transfer Mode, ATM) típusú hálózatok hívásengedélyezési algoritmusainál gyakran használják a forgalommodell-alapú megközelítést. Sokszor azonban olyan feltételezésekkel éltek az eddigi megoldások [4, 8] kidolgozói, amelyek növelik a távolságot a forgalommodell és a valóság között. Ilyen feltételezés például az, hogy tetszőlegesen hosszú időintervallumokat engednek meg, amelyekben maximális bitfolyam érkezik. A legtöbb forgalomban levő többszolgálatú (multi-service) hálózat ugyanis forgalomkorlátozó mechanizmust használ, amely eljárás azért is felelős, hogy ezek hosszát korlátozza (egy maximális hossznál levágja). A jelenlegi hívásengedélyező algoritmusok többsége ezt a csonkolást nem veszi figyelembe.

A cikk 2. fejezete elméleti áttekintést nyújt a nagy eltérések elméletén alapuló nagy puffer aszimptotáról és arról, hogy ennek segítségével miként lehet a forgalom hatékony sávszélességigényét értelmezni és hívásengedélyezésre felhasználni. A 3. fejezetben különböző forgalommodellekre számítjuk ki a skálázott logaritmikus momentumgeneráló függvényt.

2. Matematikai háttér

A nagy eltérések elméletén alapuló „nagy puffer aszimptotika" tömegkiszolgáló rendszerek puffertúlcsordulási valószínűségének becslésére ad módszert [2, 3, 5, 7, 9]. Tekintsünk egy egy kiszolgálós rendszert, amelyben a kiszolgáló kapacitását c-vel, a puffer méretét b-vel jelöljük. A túlcsordulás valószínűségét ebben a rendszerben azzal a valószínűséggel közelíthetjük, hogy egy c kapacitású, de végtelen pufferméretű rendszer sorhossza (Q) meghaladja a b értéket. Az elméletileg a túlcsordulás valószínűségének logaritmusa aszimptotikusan lineáris b-ben, azaz

$$\lim_{b\to\infty}\frac{1}{b}\ln P(Q>b) = -\delta(c), \qquad 1.$$

ahol $\delta(c) = \sup\{s : SCGF(s) \le cs\}$. Itt pedig az *SCGF* az érkezési folyamat skálázott logaritmikus momentumgeneráló függvénye (scaled cumulant generating function, SCGF):

$$SCGF(s) := \lim_{t\to\infty}\frac{1}{t}\ln E\left[e^{sX[0,t)}\right], \qquad 2.$$

amelyben $X[0,t)$ a $[0,t)$ időintervallumban beérkezett forgalommennyiséget jelenti.

Ez azt jelenti, hogy nagy b esetén a túlcsordulási valószínűségre a becslés (1) alapján használható a következő kifejezés

$$P(Q>b) \approx e^{-b\delta(c)}.$$

Az érkezési folyamat sávszélességigényén (ekvivalens kapacitásán) azt a minimális c_{equ} sávszélességet értjük, amellyel a forgalom az előre megadott ϵ csomagvesztési valószínűség mellett kiszolgálható. Ezt a csomagvesztési valószínűséget a fentiek szerint becsülve az ekvivalens kapacitás tehát:

$$c_{equ} := \inf\left\{c : e^{-\delta(c)b} \le \epsilon\right\}.$$

A $\delta(c)$ függvény c-ben monoton növekvő, így megfigyelhető, hogy c_{equ} kiszolgáló kapacitás mellett a várható csomagvesztés éppen megegyezik az előírttal:

$$e^{-\delta(c_{equ})b} = \epsilon, \quad \text{azaz} \quad \delta(c_{equ}) = \frac{-\ln\epsilon}{b}.$$

Ezt összevetve azzal a megfigyeléssel, hogy az *SCGF* függvény konvexitása miatt $SCGF(\delta(c)) = c\delta(c)$, az ekvivalens kapacitásra a következő kifejezést kaphatjuk:

$$c_{equ} = \frac{SCGF(\delta(c_{equ}))}{\delta(c_{equ})} = \frac{SCGF\left(\frac{-\ln\epsilon}{b}\right)}{\frac{-\ln}{b}}. \qquad 3.$$

A nagy eltérések elméletének nagy puffer aszimptotikáján alapuló hívásengedélyezési algoritmus az ekvivalens kapacitást (3) használhatja a hívásengedélyezési döntéshez: ha egy új hívás érkezik, ellenőrzi, hogy ennek az értéknek (c_{equ}) és az újonnan érkező adatfolyam maximális sebességének (p) összege túllépi-e a kimenő kapacitást (c). Ha nem $(c_{equ}+p \le c)$, engedélyezi a hívást, ellenkező esetben elutasítja. Látható (3)-ból, hogy a sávszélességigény meghatározásához elegendő az érkező forgalom skálázott logaritmikus momentumgeneráló függvényét (2) meghatározni, mert ϵ és b előre megadott mennyiségek. Ez a függvény becsülhető forgalmi mérésekből, vagy analitikusan számolható egyes valóságot közelítő forgalommodellek esetén.

3. A skálázott logaritmikus momentumgeneráló függvény kiszámítása ON-OFF forgalmakra

Célunk a forgalom minél reálisabb skálázott logaritmikus momentumgeneráló függvényének kiszámítása. A matematikai kezelhetőség érdekében ON-OFF modelleket fogunk vizsgálni (más típusú modellek vizsgálata pl. a [4, 1] munkákban található). Egy forrást ON-OFF forrásnak nevezünk, ha élettartama során a forgalomban egymást váltó aktív (ON) és inaktív (OFF) periódusokat különböztethetünk meg: az ON periódus alatt a forrás maximális sávszélességgel generál adatot, az OFF periódus alatt pedig nem forgalmaz. Az ON-OFF források a tömörített hangátvitelt jól modellezik, adat- és képátvitel esetén pedig megfelelő beállításokkal elfogadható közelítést adnak. Előnyük az is, hogy a forrás leírói könnyen illeszthetők a bevezetőben már említett forgalomkorlátozás során kapott forgalomleírókkal (a maximális és az átlagos cellasebességgel, valamint a maximális ON periódusmérettel).

Az irodalomban gyakran Markov-típusú ON-OFF modellt használnak a forgalom leírására: exponenciális eloszlású hosszú ON periódus után esetleg más paraméterű exponenciális eloszlású hosszúságú OFF periódus következik. Felhasználva az exponenciális eloszlás tulajdonságait, könnyű kezelhetőségét, erre a modellre általános analitikus eredmény adódik a skálázott logaritmikus momentumgeneráló függvényre [8]. A pozitív eséllyel előforduló, tetszőlegesen hosszú ON periódus azonban nem reális feltételezés a gyakorlatban a forgalomkorlátozás miatt. Ezért a továbbiakban korlátos ON-periódusú ON-OFF források skálázott logaritmikus momentumgeneráló függvényének kiszámítására törekszünk.

A skálázott logaritmikus momentumgeneráló függvény definíciójában (2) szereplő várható értéket a beérkezett adatmennyiség, $X[0,t)$ eloszlásának segítségével fogjuk számolni.

Független kapcsolatok multiplexálása esetén a forgalmak skálázott logaritmikus momentumgeneráló függvényei összeadódnak. A továbbiakban így – azonos típusú kapcsolatok esetén – elegendő csak egy-egy hívás (forrás) skálázott logaritmikus momentumgeneráló függvényét kiszámítani.

3. 1. Markov ON-OFF modellek

A következőkben Markov-típusú ON-OFF forrásokat vizsgálunk. Ez azt jelenti, hogy a forrást folytonos idejű kétállapotú Markov-lánccal modellezzük, amelyben az ON és az OFF periódus (állapot) hossza is exponenciális eloszlású μ és η paraméterekkel, és az ON állapotban p intenzitással generálódik adat, az OFF állapotban pedig nem keletkezik forgalom. Ezen forrás skálázott logaritmikus momentumgeneráló függvénye [8] alapján

$$SCGF(s)=\lim_{t\to\infty}\frac{1}{t}\ln\left(\left[\frac{\mu}{\eta+\mu},\frac{\mu}{\eta+\mu}\right]\times\right.$$

$$\left.\times\exp\left[\begin{pmatrix}-\mu+ps & \mu\\ \eta & \eta\end{pmatrix}t\right]\times\begin{pmatrix}1\\1\end{pmatrix}\right]=$$

$$=\frac{1}{2}\left(ps-\mu-\eta+\sqrt{(ps-\mu-\eta)^2+4\eta\mu}\right).$$

A folytonos idejű eset diszkrét megfelelője az, hogy az exponenciális eloszlás helyett geometriai eloszlást tételezünk fel a forrást modellező kétállapotú Markovláncban: egy ilyen diszkrét idejű ON-OFF modellt kapunk, ha feltesszük, hogy minden diszkrét időpillanatban q valószínűséggel p forgalom generálódik, $(1-q)$ valószínűséggel pedig nulla. Erre a forrásra

$P(X[0,n]=kp)=\binom{n}{k}q^k(1-q)^{n-k}$, és így a skálázott logaritmikus momentumgeneráló függvénye

$$SCGF(s)=\lim_{n\to\infty}\frac{1}{n}\ln\sum_{k=0}^{n}\binom{n}{k}q^k(1-q)^{n-k}e^{kps}=$$

$$=\lim_{n\to\infty}\frac{1}{n}\ln\sum_{k=0}^{n}\binom{n}{k}\left(qe^{ps}\right)^k(1-q)^{n-k}=$$

$$=\lim_{n\to\infty}\frac{1}{n}\ln\left(qe^{ps}+1-q\right)^n=\ln\left(qe^{ps}+1-q\right).$$

3.2. A folytonos idejű ON-OFF modellek

Ebben a fejezetben általános ON-OFF forrásokkal foglalkozunk: általános eloszlású ON és OFF periódusok váltakoznak egymással. A továbbiakban egy ON és egy OFF periódust együtt blokknak fogunk nevezni, az i. ON periódus tartásidejét X_i-vel, az i. OFF periódus tartásidejét Y_i-vel, valamint az ebből a kettőből összeálló blokk hosszát Bi-vel jelöljük. Minden i-re X_i azonos, X-szel megegyező eloszlású valószínűségi változó, az Y_i-k eloszlásai is azonosak (Y) és $B=X+Y$. Ezekről az X_i és Y_i valószínűségi változókról feltesszük, hogy függetlenek egymástól. Egy adott t időponthoz definiáljuk az

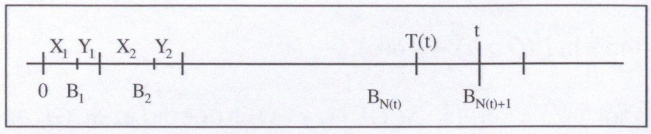
$$N(t):=\sup\left(k:\sum_{i=1}^{k}B_i\le t\right)\quad és\ a \tag{4.}$$

$$T(t):=\sum_{i=1}^{N(t)}B_i \tag{5.}$$

számokat!

Az $N(t)$ valószínűségi változó azt adja meg, hogy hány teljes blokk fér el a t időpontig, míg a $T(t)$ az utolsó ilyen blokk végét adja meg (lásd az 1. ábrát).

A fenti definíciók felhasználásával kifejezhető $X[0,t)$ eloszlásfüggvénye. A kérdéses eloszlás az X_i $N(t)$-szeres konvolúciójának és az utolsó $t-T(t)$ töredékintervallumban érkezett forgalommennyiség $(R(t))$ összegeként áll elő:



1. ábra

$$P\left(X[0,t)<x\right)=P\left(X^{*N(t)}+R(t)<x\right). \tag{6.}$$

$N(t)$ eloszlásának ismeretében $X[0,t)$ eloszlása egyszerűen becsülhető:

$$\sum_{k=1}^{\infty}P\left(X^{*N(t)}>x\middle|N(t)=k\right)P(N(t)=k)\le$$

$$\le P\left(X[0,t)>x\right)\le$$

$$\le\sum_{k=1}^{\infty}P\left(X^{*(N(t)+1)}\ge x\middle|N(t)=k\right)P(N(t)=k). \tag{7.}$$

Megfigyelhetjük $N(t)$ eloszlása és B konvolúcióhatványai közötti kapcsolatot (1. ábra és definíciók):

$$P(N(t)\ge k)=P\left(\sum_{i=1}^{k}B_i<t\right)=P(B^{*k}<t).$$

Tehát ha X és Y minden konvolúcióhatványát ki tudjuk számítani $(B^{*k}=X^{*k}*Y^{*k})$, akkor $N(t)$ eloszlása is ismert:

$$P(N(t)=k)=F_{B^{*k}}(t)-F_{B^{*(k+1)}}(t). \tag{8.}$$

Az $X[0,t)$ eloszlásának (7) szerinti becsléséhez tehát egyrészt az ON periódushosszak konvolúcióira, másrészt a blokkhosszak konvolúcióira (8) van szükségünk.

3.3. Korlátos ON tartásidő-eloszlású modellek

Ebben a fejezetben az egyszerűség kedvéért olyan forgalommodelleket vizsgálunk, amelyekben egy ON periódus tartásideje egy korlátos valószínűségi változó, miközben az OFF periódus hosszát konstansnak választjuk (l_{OFF}). Mint láthattuk (8)-ban, $N(t)$ eloszlásának meghatározásához B konvolúcióhatványainak kiszámítására van szükségünk, amelyek a konstans OFF periódushossz miatt könnyen számíthatók X konvolúcióhatványaiból

$$F_{B^{*k}}(t)=P(B^{*k}<t)=P(X^{*k}<t-kl_{OFF}). \tag{9.}$$

Determinisztikus OFF periódushosszak esetén tehát csupán az ON periódushossz konvolúcióinak ismerete elegendő a (7)-ben szereplő becslés használatához.

Egyenletes eloszlású ON periódus

Ha a konstans OFF periódus mellett az ON periódusok hossza egyenletes eloszlású a [0,1] intervallumon (a módszer illusztrációja végett), akkor $F_{B^{*k}}(t)$ kiszámítható:

$$F_{B^{*k}}(t)=\frac{1}{k!}\sum_{i=0}^{[t-kl_{OFF}]-1}(-1)^i\binom{k}{i}(t-kl_{OFF}-i)^k,$$

ahol felhasználtuk egy tetszőleges [0,1]-en egyenletes eloszlású U valószínűségi változó k. konvolúcióhatványának sűrűségfüggvényét [6]: minden $i \in \{1,2,\ldots,k-1\}$ és $x \in [i,i+1]$ esetén

$$f_{U^{*k}}(x)=\frac{1}{(k-1)!}\sum_{j=0}^{i}(-1)^j\binom{k}{j}(x-j)^{k-1}. \qquad 10.$$

Azt várjuk, hogy az ebből kapott alsó és felső korlátok (7) határértékei ($t \to \infty$) zárt formulával kifejezhetők, azaz analitikus becslést adnak a skálázott logaritmikus momentumgeneráló függvényre. Az itt még szükséges mennyiség, $P(X^{*N(t)}>x/N(t)=k) = P(X^{*k}>x)$, közvetlenül kifejezhető a (10) egyenletből $U=X$ helyettesítéssel.

Csonkolt exponenciális eloszlású ON periódus

Ebben a fejezetben a determinisztikus OFF periódus mellett, az ON periódusról azt tesszük fel, hogy csonkolt exponenciális eloszlású, l_{ON} korláttal és λ paraméterrel. A csonkolt exponenciális eloszlás az exponenciális eloszlás egy korlátos változata, sűrűségfüggvénye:

$$f_X(x)=\begin{cases}\frac{\lambda e^{\lambda x}}{1-e^{\lambda l_{ON}}} & ha,\ x \in [0,l_{ON}) \\ 0 & különben\end{cases}$$

Bár az ON periódusunk csonkolt exponenciális eloszlású, a (9) egyenletben használt k. konvolúcióhatvány sűrűségfüggvénye előáll (10) segítségével a következő alakban:

$$f_{X^{*k}}(x)=\frac{\lambda^k e^{-\lambda x}}{(1-e)^{-\lambda l_{ON}}}f_{U^{*k}}(x/l_{ON}).$$

A $P(X^{*N(t)}>x/N(t)=k) = P(X^{*k}>x)$ mennyiség kiszámítása ebből a formulából történhet. Hasonlóan az egyenletes eloszlás esetéhez, itt is azt várjuk, hogy (7) felhasználásával a skálázott logaritmikus momentumgeneráló függvényre zárt képletben kifejezett alsó és felső korlátot kapunk.

Véletlen indítású determinisztikus ON-OFF modell

Az $X[0,t)$ eloszlása a (7) becslés szerint még az ON periódusok (vagy blokkok) konvolúcióhatványainak ismerete esetén sem mindig állítható elő zárt formában. Amint az a fenti két példa alapján is látható, a becslés alsó es felső korlátainak explicit meghatározása is további munkát igényel. A skálázott logaritmikus momentumgeneráló függvény kiszámítása ezekben az esetekben például numerikusan lehetséges.

A következő rendkívül egyszerű modellben demonstráljuk, hogyan lehet segítségünkre a blokkos megközelítés, (4) és (5), $X[0,t)$ eloszlásának (6) pontos kiszámításához és ezáltal a skálázott logaritmikus momentumgeneráló függvény zárt alakban való előállításához.

Ez a modell konstans l_{OFF} hosszú OFF tartásidőkkel és konstans l_{ON} hosszú ON tartásidőkkel dolgozik. Az egyetlen véletlen elem az, hogy az első periódust sorsoljuk, p valószínűséggel ON-nal, $(1-p)$ valószínűséggel OFF-al indítunk. Az $N(t)$ és $T(t)$ értékünk determinisztikus lesz:

$$N(t)=\left[\frac{t}{l_{OFF}+l_{ON}}\right], T(t)=N(t)(l_{OFF}+l_{ON})$$

Az $X[0,t)$ eloszlása csak a kezdéstől függ és a determinisztikus $[l_{ON}N(t), l_{ON}(N(t)+1)]$ intervallumra koncentrálódik. Ha feltesszük, hogy $x \in (0, l_{ON})$, akkor ezen az intervallumon az eloszlásfüggvény számítható, és

$$P(X[0,T)>N(t)l_{ON}+x)=$$
$$=p\frac{l_{OFF}+l_{ON}-x}{l_{OFF}+l_{ON}}+(1-p)\frac{l_{ON}-x}{l_{OFF}+l_{ON}}.$$

A sűrűségfüggvény \ ebből deriválással kapható:

$$\frac{\partial\left(1-P(X[0,t)>N(t)l_{ON}+x)\right)}{\partial x}=\frac{1}{l_{OFF}+l_{ON}}.$$

Ez az egyenletes eloszlású rész azt jelenti, hogy a t időpont egy ON periódusra esik. Amiatt, hogy ez egy OFF periódusba is eshet, $X[0,t)$ eloszlásában van egy-egy az intervallum két végpontjára koncentrált rész:

$$P(X[0,t)=(N(t)l_{ON})=(1-p)\frac{l_{OFF}}{l_{OFF}+l_{ON}},\ illetve$$
$$P(X[0,t)=(N(t)+1)l_{ON})=p\frac{l_{OFF}}{l_{OFF}+l_{ON}}.$$

Ebből a skálázott logaritmikus momentumgeneráló függvény kiszámítható:

$$SCGF(s)=\frac{sl_{ON}}{l_{ON}+l_{OFF}}.$$

4. Összefoglalás

A hívásengedélyező algoritmus fontos építőeleme, hogy megjósolja a rendszer viselkedését arra az esetre, ha egy újonnan érkező hívást beengednénk a rendszerbe. Ez például a már a rendszerben levő forgalom által felhasznált kimenő kapacitásnak (ekvivalens kapacitás) becslésével tehető meg. Az ekvivalens kapacitás meghatározásának egyik lehetséges megközelítése a forgalom folyamatos mérése és a kapott idősoron statisztikai próbák elvégzése. Ennek a megközelítésnek pozitívuma, hogy semmilyen előzetes információt nem kell feltételezni a forgalomról. Másrészről a másik, az analitikus modelleken alapuló megközelítésnek megvan az az előnye, hogy a forgalomról előzetesen rendelkezésre álló adatokat felhasználva zárt formulákat

adnak meg, melyek könnyen átvihetők a gyakorlatba. Alkalmazásuk a hívásengedélyezésben tehát általában egyszerűbb, és a hívásengedélyezési döntés is rövidebb időt vesz igénybe, mint a mérésalapú esetben.

A legtöbb létező analitikus eredmény a nagy eltérések elméletén alapuló hívásengedélyező eljárásokra azzal a gyakorlatban nem reális feltételezéssel él, hogy a források tetszőlegesen hosszú ideig generálhatnak forgalmat. Mivel a legtöbb hálózatban jelen vannak forgalomkorlátozó mechanizmusok, mi olyan forrásokat igyekeztünk megvizsgálni, amelyeknél a forgalomgenerálás csak korlátos ideig tarthat.

Az általunk választott hívásengedélyezés a sávszélességigény becslésén alapul, amelyet a nagy eltérés-elmélet nagy puffer aszimptotájával határoztunk meg. A sávszélességigényt (vagy ekvivalens kapacitást) a beérkező forgalom skálázott logaritmikus momentumgeneráló függvényének segítségével lehet kiszámolni.

Egy általános valószínűségelméleti megközelítést adtunk ON-OFF típusú forgalmakra, amelynek segítségével képlet vezethető le a skálázott logaritmikus momentumgeneráló függvényre. A legegyszerűbb forgalmakra ezt alkalmazni is tudtuk, mivel azonban kicsit bonyolultabb, de még mindig egyszerű modellekre már nehézségekbe ütköztünk, egy olyan alsó és felső korlátot adtunk a skálázott logaritmikus momentumgeneráló függvényre, amelyektől azt várjuk, hogy zárt alakra hozhatók.

A jövőben bonyolultabb ON-OFF típusú forgalmakra szeretnénk kiterjeszteni az új megközelítést, hogy a rendkívül elterjedt modell helyett, amelyben az ON és az OFF periódusok hosszát exponenciális eloszlásúnak feltételezzük, a korlátos, tehát csonkolt exponenciális eloszlású ON periódussal rendelkező, sokkal reálisabb modellt vizsgáljuk.

Hivatkozások

1. Jakabfy Tamás: A nagy eltérések elméletén alapuló hívásengedélyezési algoritmusok viselkedése néhány forgalomtípus esetére, TDK dolgozat, Eötvös Loránd Tudományegyetem, 2002
2. John T. Lewis, Raymond Russell: An Introduction to Large Deviations for Teletraffic Engineers, Technical Report, Dublin Institute for Advanced Studies - Applied Probability Group, 1996
3. Simon Crosby, Ian McGurk, John T. Lewis, Raymond Russell, Fergal Toomey: Statistical Properties of a Near-Optimal Measurement-Based CAC Algorithm, Proc. IEEE ATM '97, Lisbon, Portugal, 1997
4. Frank Kelly: Notes on Effective Bandwidths, Stochastic Networks: Theory and Applications, vol. 4, pp. 141–168, Oxford University Press, 1996
5. Nick G. Duffield: Economies of scale in queues with sources having power-law large deviation scalings, J. Appl. Prob., no. , pp. 840–857, 1996
6. Prékopa András: Valószínűségelmélet műszaki alkalmazásokkal, Műszaki Könyvkiadó, Budapest, 1962
7. Peter W. Glynn, Ward Whitt: Logarithmic asymptotics for steady-state tail probabilities in a single-server queue, Studies in Applied Probability, 1993
8. George Kesidis, Jean Walrand, and Cheng-Shang Chang, Effective bandwidths for multiclass Markov fluids and other ATM sources, IEEE/ACM Trans. Networking, vol. 1, no. 4 pp. 424–428, 1993
9. Dembo-Zeitouni: Large Deviations Techniques and Applications, Springer, 1998

UMTS hozzáférési hálózatok teljesítőképesség-vizsgálata

MALOMSOKY SZABOLCS, NÁDAS SZILVESZTER, SONKOLY BALÁZS

Traffic Analysis and Network Performance Laboratory, Ericsson Research
E-mail: Szabolcs.Malomsoky@eth.ericsson.se

Reviewed

Az UMTS hozzáférési hálózatok (UTRAN) transzport linkjein, de főleg az Iub interfészen, amely a bázisállomásokat a rádiós hálózatvezérlő központokkal köti össze, az erőforrás-foglalás nehéz a csomagkésleltetési és vesztési követelmények szigorúsága és a rendelkezésre álló átviteli erőforrások szűkössége miatt. Ebben a cikkben az UTRAN-ban használható, linkszintű kapcsolat engedélyezési (LAC – link admission control) algoritmusokat javaslunk, melyek FIFO és prioritásos ütemező alkalmazásakor is működnek. A hálózatszintű engedélyezési döntések (CAC – connection admission control) a kapcsolási csomópontokban történő, független LAC döntésekből tevődnek össze. Ezért megvizsgáljuk azt, hogy a hálózatra megfogalmazott minőségi követelményeket mennyire képes teljesíteni a rendszer. Elsődleges célunk az ATM/AAL2 alapú UTRAN vizsgálata, de modellünk IP alapú hálózati infrastruktúrák vizsgálatára is alkalmas.

1. Bevezetés

Az UTRAN első verzióiban a kapcsolási és multiplexálási technológiák főleg az Aszinkron Transzfer Módra (ATM) és az AAL2 (ATM Adaptation Layer type 2) ATM adaptációs rétegre épülnek [4] [7]. A későbbi kiadású rendszerek az IP (Internet Protocol) alapú átvitelt is támogatni fogják [2]. A kulcskérdés a hozzáférési hálózat hatékony erőforrás-menedzsmentje, a sok bázisállomás miatt. Ezzel a szakirodalomban viszonylag kevés cikk foglalkozik. Az ATM/AAL2 technológia használatát az UTRAN-hoz mutatják be a [7] irodalomban, továbbá foglalkoznak azok teljesítőképességével és minőségi (QoS – quality of service) problémáival. Egyszerű, a Chernoff-korláton alapuló, linkszintű engedélyezési algoritmusok (LAC) találhatók a [8]-ban. Egy linkméretező módszert mutatnak be beszédforgalomra [9]-ben. A [10] és [11] az AAL2 hálózatok sávszélesség-allokációjával foglalkozik, illetve az AAL2 csomagok elhelyezésével az ATM cellákba. Ezen vizsgálatok alapján a [12]-ben az UBR (Unspecified Bit Rate – nem specifikált erőforrás-allokációjú) VC (Virtual Channel – ATM virtuális csatorna) kapcsolást javasolják az UMTS-hez.

Munkánk motivációi a következők:

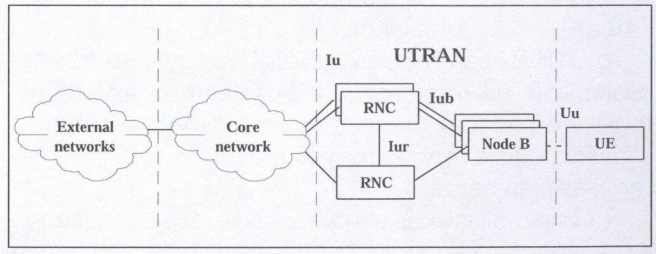
- Az UTRAN transzport hálózat megfelelő modellezése (a [7]-hez és [8]-hoz képest újdonság a forgalom Iub-specifikus modellezése),
- LAC algoritmusok fejlesztése FIFO (first-in-first-out) és prioritásos ütemező esetére, az algoritmusok jóságának igazolása szimulációval és matematikai analízissel a [14], [16], és [18] alapján,
- Hálózati szintű QoS értékelése, és
- Különböző kapcsolási alternatívák hatékonyságának vizsgálata új módszerrel.

A cikk felépítése a következő. A 2. rész a rendszer felépítését és működését mutatja be. A 3. és 4. részben az *egylinkes* esetet vizsgáljuk. A 3. rész egy link sorban állási modelljét tárgyalja. A 4. részben a LAC algoritmusokat közöljük, és számpéldákkal bizonyítjuk ezek alkalmazhatóságát. Az 5. részben a *többlinkes* (hálózati) esetet elemezzük; a hálózati szintű QoS-t analizáljuk, és különböző kapcsolási (forgalomkoncentrálási) alternatívák hatékonyságát értékeljük. A cikk konklúziója a 6. részben található.

2. A rendszer felépítése és működése

Az UTRAN transzporthálózata

Az UMTS hálózat alapelemei a következők: felhasználói végberendezés (UE – user equipment), a hozzáférési hálózat (UTRAN), melynek fő elemei a bázisállomások (Node B) és a rádiós hálózatvezérlők (RNC), valamint a maghálózat (core network) (1. ábra).

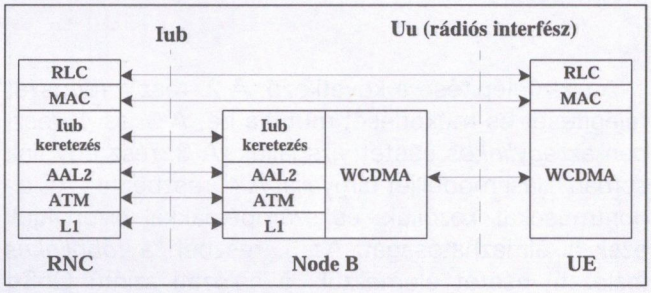


1. ábra UMTS hálózati architektúra

Az UTRAN végzi a rádiós erőforrás-menedzsmentet, a hívásátadás (handover) vezérlését stb. A maghálózat alapvető feladata, hogy a hozzáférési hálózatot a külső

hálózatokkal (external networks), például a telefonhálózattal vagy az internettel összekapcsolja. A mobil (UE – user equipment) a bázisállomásokhoz (Node B) a WCDMA [6] (széles sávú CDMA) rádiós interfészen (Uu) keresztül kapcsolódik. Egy UE egyszerre több bázisállomással is kapcsolatot tarthat „soft handover" esetén.

Az ATM/AAL2 alapú UTRAN protokoll felépítése a 2. ábrán látható. A rádiós link kontroll (RLC) protokoll újraküldési mechanizmusa biztosítja a veszteségmentes átvitelt a rádiós interfészen. A MAC (medium access control) protokoll állítja össze a rádiós csomagokat és ezeket a rádiós interfész időzítési követelményei szerint periodikusan ütemezi. A periódust TTI-nek hívják (transmission time interval), hossza 10 ms többszöröse. A rádiókapcsolatok (RAB – radio access bearer) bitsebességei tipikusan 8 és 384 kbps között vannak. Ha a legegyszerűbb esetet tekintjük, amikor egy felhasználóhoz egy RAB van rendelve, minden aktív UE-hez egy RLC és egy MAC entitást hoz létre a rendszer az RNC-ben (illetve ezek párját a UE-ben).



2. ábra Az ATM/AAL2 alapú UTRAN protokolljai

A keretezés után a MAC csomagokból Iub csomagok jönnek létre. Az Iub csomagok a szegmentálás és az AAL2 fejléc hozzáadása után AAL2 CPS (common part sublayer) csomagokba kerülnek, melyeket az AAL2 multiplexer ATM cellákba csomagol. Az AAL2 csomagok változó méretűek (fejléc nélkül max. 45 byte), az AAL2 fejléc 3 byte. Az ATM cellák 53 byte hosszúak, amiből 5 byte a fejléc. Az AAL2 multiplexálás során különböző kapcsolathoz tartozó AAL2 csomagok kerülhetnek egy ATM cellába (tipikusan átlapolódás útján 2 vagy 3 csomag). Az AAL2 fejléc CID (connection identifier) mezeje azonosítja az AAL2 kapcsolatokat egy ATM VC-ben. Egy VC-ben maximum 248 AAL2 kapcsolat lehet.

Az UTRAN-ban minden új RAB-hoz egy új AAL2 kapcsolat épül fel. A transzporthálózatban az AAL2 CAC foglal erőforrást az új kapcsolatoknak, mely döntéseit a kapcsolatokhoz rendelt forgalomleírók és QoS paraméterek alapján hozza.

A sorban állási késleltetés csökkentése érdekében az egyes kapcsolatok kezdőfázisai véletlenszerűen (egyenletes eloszlással) vannak elosztva a TTI fölött. A felhasználói, illetve applikációs szintű forgalmi jellemzők úgy jelennek meg, hogy a transzport hálózatban az egyes kapcsolatok forgalma nem egy folyamatos periodikus csomagfolyamként modellezhető, hanem aktív és inaktív intervallumok váltakoznak (csomagküldés minden TTI-ben, illetve nincs csomagküldés). Ezeket az intervallumokat ON és OFF periódusoknak hívjuk. Például beszédátvitel esetén az ON és OFF periódusok jellemzőit a beszédfolyamat- és a kódolóban található beszédaktivitás-detektor interakciója határozza meg. A *forgalomleírókat* a kapcsolatok felépítését megelőző jelzésekben küldik a csomópontoknak. Az UTRAN-ban a forgalomleírók a következők: csomagméret (AAL2 szegmentálás előtt, minden felsőbb protokoll fejlécét beleszámolva), TTI (a csomagok érkezésének periódusideje), és az ún. „aktivitási faktor". Az aktivitási faktor 0 és 1 közötti szám, definíciója szerint: az átlagos ON periódushossz elosztva az átlagos ON és OFF periódus hosszak összegével. (Ez azt jelenti, hogy az ON és OFF periódusok hosszának eloszlásáról a LAC-nak nincs információja.) Az UMTS-ben a következő TTI értékek lehetségesek: 10, 20, 40 és 80 ms. Például egy beszédszolgáltatás tipikus leírói lehetnek a következők: 40 byte hosszú csomagok, 20 ms TTI, aktivitási faktor 0,6. Fontos megjegyezni, hogy az aktivitási faktor sok kapcsolat felett számított átlagérték, melyet az operátor állíthat be, ha az ebből eredő statisztikus multiplexálási nyereséget ki szeretné használni.

A *QoS követelményeket* tekintve, számos okból, a transzporthálózatban a csomagkésleltetés a legfontosabb követelmény. Például beszédforgalom esetén a telefontól telefonig (vagy telefontól gatewayig) tartó késleltetésre vonatkozó követelmény határozza meg az elviselhető sorban állási késleltetés mértékét, ami 5-7 ms [1]. Egyéb forgalomtípusokra is szigorúak a követelmények. Ennek egyik oka az, hogy soft handover esetén az egy felhasználónak szánt MAC csomagokat minden handoverben részt vevő bázisállomástól egyszerre (t_{out}) kell kiküldeni a UE felé. Ezért a hálózati elemek szinkronizálva vannak, és a hálózatnak biztosítania kell, hogy minden Iub csomag megérkezzen a bázisállomásokhoz t_{out} előtt. Ez – még best-effort (pl. internet-) forgalom esetén is – szigorú késleltetéskövetelményt eredményez. További fontos tényező, hogy az RLC csomagok körbefordulási idejét (round-trip time) célszerű minimalizálni a best-effort forgalom áteresztő képesség maximalizálása érdekében [13]. A szigorú késleltetés-követelmények miatt kisméretű tárolókat alkalmaznak a rendszerben, melyek főleg a rövid időskálán bekövetkező forgalomfluktuációkat képesek kiegyenlíteni. Mi azt feltételezzük, hogy egy kapcsolat vég-vég (Node B – RNC) késleltetés követelménye kisebb a kapcsolatra jellemző TTI-nél.

A forgalomleírók alapján a kapcsolatokat forgalmi osztályokba sorolhatjuk. Ha minden osztály csomagjai ugyanabban a sorban várakoznak, és a csomagok kiszolgálása érkezési sorrendben történik (first-in-first-out, FIFO), akkor a legszigorúbb követelménynek kell teljesülnie. Ezt el lehet kerülni *QoS differenciálással,* amikor külön sorokban állhatnak különböző követelményű csomagok. Az AAL2 protokoll első kiadása (Capability Set 1) nem támogatja a QoS differenciálást. A második kiadás (Capability Set 2) lehetővé teszi, hogy egy kapcsolathoz

a hordozó ATM VC-t QoS követelmény alapján válasszuk [4]. A későbbi verziók valószínűleg lehetővé fogják tenni az AAL2 szintű differenciálást is.

ATM/AAL2 kapcsolási alternatívák

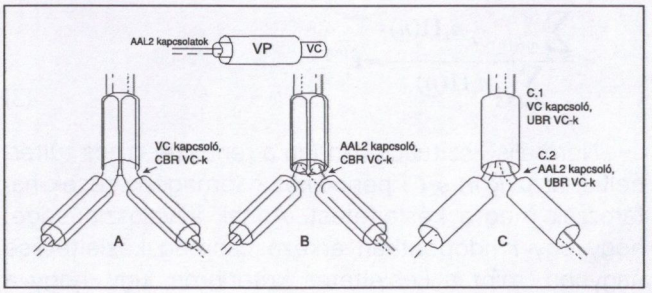
Ebben a cikkben a következő ATM/AAL2 kapcsolási alternatívákkal foglalkozunk (3. ábra):

A *ATM VC kapcsolók és CBR VC-k használata:* Ebben a változatban vég-vég CBR (constant bit rate - konstans bitsebességű) VC-ket használnak. A VC kapacitásokat a CID-korlát (max. 248 kapcsolat egy VC-n) figyelembevételével kell meghatározni. Ha túl nagy kapacitásokat választunk, a CID-korlát esetleg nem fogja engedni a teljes kapacitás kihasználását (sok kis erőforrás-igényű kapcsolat esetén). Ha túl kicsi kapacitásokat választunk, a csomagszintű statisztikus nyereséget csökkentjük. A teljes erőforrás túlzott elaprózása egyéb okokból sem praktikus.

B *Kapcsolás az AAL2-ben és CBR VC-k használata:* Ez az alternatíva hasonló az **A**-hoz, de itt nincsenek vég-vég VC-k. A hálózatban lévő koncentrációs csomópontok az AAL2 csomagokat demultiplexálják és egy másik VC-be multiplexálják. Az ilyen eszközöket mostantól AAL2 kapcsolóknak fogjuk nevezni. Minden AAL2 kapcsolóban szükséges a LAC algoritmust implementálni.

C.1 *ATM VC kapcsolók és UBR VC-k használata (CBR VP-kben):* Ez az eset hasonló az A alternatívához, de CBR helyett UBR (unspecified bit rate – nincs erőforrás-foglalás) VC-ket használunk. A [12]-ben ezt az alternatívát javasolják az UMTS-hez. Ebben az esetben a CID-korlát nem jelent erőforrás-foglalással kapcsolatos problémát, viszont az AAL2 jelzéseket az ATM kapcsolóknak is kezelniük kell, mert az AAL2 LAC-nak minden koncentrációs csomópontban a VP erőforrásait kell allokálnia (további részletek a [12]-ben találhatók). A megoldás teljesítőképessége hasonló a **C.2**-höz. Az általunk használt modell **C.1**-re és **C.2**-re azonos.

C.2 *Kapcsolás az AAL2-ben és UBR VC-k használata (CBR VP-kben):* Ez az alternatíva a **B**-hez hasonlít, de itt UBR VC-ket használunk. A **C.1** esethez képest fontos különbség, hogy az AAL2 kapcsolókban az AAL2 LAC képes a VP erőforrásait allokálni. Várhatóan ez az alternatíva eredményezi a legnagyobb statisztikus multiplexálási nyereséget.



3. ábra ATM/AAL2 kapcsolási (switching) alternatívák UTRAN-ban

A fenti alternatívákban CBR VC-k helyett rt-VBR (real time-Variable Bit Rate – valós idejű változó bitsebességű) VC-ket is használhatnánk. Ezzel kapcsolatban a következő megjegyzéseink vannak. Az erőforrás-foglalás rt-VBR VC-k fölött lényegesen bonyolultabb, mint CBR VC-k fölött. Az elsődleges probléma az, hogy a forgalomleírókban a börsztösség jellemzése nem szerepel, pedig ez szükséges lenne, ha rt-VBR VC-ket használnánk. Ha a börsztösségről feltételezéseket teszünk, vagy méréseket végzünk, természetesen használhatunk rt-VBR VC-ket, de akkor a fenti opciókkal az összehasonlítás nehézkes. AAL2 kapcsolatok rt-VBR VC-k fölötti erőforrás foglalásával foglalkozik [21].

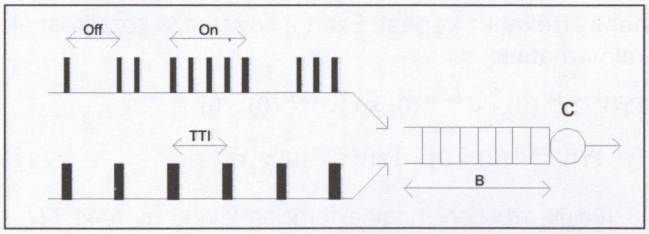
3. Sorban állási modell

Ebben a részben egy sorban állási modellt mutatunk be az egylinkes esetre, FIFO, illetve prioritásos ütemezést feltételezve.

FIFO ütemezés

A 4. ábra egy példát mutat két kapcsolattal, melyek különböző forgalmi osztályokba tartoznak. A csomagkésleltetési kritérium megsértésének valószínűségét szeretnénk kiértékelni: $\Pr\{D_i > \tilde{D}_i\}$, ahol D_i valószínűségi változó írja le az i osztályba tartozó csomagok késleltetését, és $\tilde{D}_i$ az i osztályú csomagok késleltetés kritériuma (vagy maximálisan megengedhető késleltetése). A tárolóban elszenvedett késleltetés két részre osztható: $D_i = Q_i + S_i$, ahol S_i a kiszolgálási idő és Q_i a sorban történő várakozási idő. A csomagvesztéseket végtelen késleltetéseknek fogjuk fel. Összefoglalva, modellünkben a következő okokból származó késleltetésekkel foglalkozunk:

- az ON-OFF viselkedés, mely időleges rendszer túlterheléseket okoz, és
- a periodikus csomagküldés az ON állapotokban, ahol a kapcsolatok csomagküldési fázisai véletlenszerűek, ami a tárolókban csomagtorlódásokhoz vezethet.



4. ábra Példa két különböző forgalomleíróval meghatározott kapcsolattal

Jelöljük a forgalmi osztályok számát K-val. Az i osztályból a rendszerben lévő kapcsolatok száma N_i. Az azonos osztályba tartozó kapcsolatok forgalomleírói azonosak: α_i az aktivitásfaktor, a csomagméret b_i, és a csomagok érkezési periódusa TTI_i. A kiszolgáló kapaci-

tása (a link kapacitása) C. Az i osztály késleltetési kritériumának megsértésének megengedett valószínűsége $\tilde{\varepsilon}_i$, ami azt jelenti, hogy a LAC-nak a következőt kell biztosítania: $\Pr\{D_i > \tilde{D}_i\} > \tilde{\varepsilon}_i$ ahol $i=1,\ldots,K$.

Börszt- és csomag szintű hatások szétválasztása

Jelöljük $A(t)$-vel, $t \geq 0$ a rendszerbe érkezett munka mennyiségét a $[-t,0)$ intervallumban. Definiáljuk a többletmunkát a $[-t,0)$ intervallumban: $W(t) = A(t) - C\,t$, $t \geq 0$. A bejövő forgalom átlagos rátája a $[-t - TTI^{max}, -t)$ intervallumban $R(t) = (A(t + TTI^{max}) - A(t)) / TTI^{max}$, ahol $(TTI^{max} = \max_i TTI_i)$. Legyen az akkumulált többletmunka, $W_{acc}(t)$, a $W(t)$ azon komponense, mely az ON-OFF viselkedés (börszt szintű fluktuáció) következménye:

$$W_{acc}(t) = \int_0^t R(u)\,du - C\,t\,.$$

A $W_{acc}(t)$ alakulása az ON és OFF periódusok eloszlásától, a források közti összefüggéstől, és még sok egyéb tényezőtől függ. Csak akkor növekedhet, ha az átlagos ráta $R(t)$ nagyobb, mint C, a kiszolgálási sebesség.

Egy FIFO sorban a várakozási idő a *munkaterheléssel* (vagy virtuális várakozási idővel) közelíthető [16]. A rendszer stacionárius, tehát $t = 0$ egy véletlenszerű időpontot reprezentál. A munkaterhelés (a továbbiakban terhelés) a következőképpen számítható: $V(0) = sup_{t\geq 0} W(t)$. A terhelés „börszt komponense", melyet csak $W_{acc}(t)$ figyelembevételével számítunk, a következő: $V^{burst}(0) = sup_{t\geq 0} W_{acc}(t)$. Ezután a „csomag komponens" egyszerűen a terhelés és a terhelés börszt komponensének különbségeként határozható meg: $V^{packet}(0) = V(0) - V^{burst}(0)$.

Minket a terhelés eloszlása érdekel: $Q(x) = \Pr\{V(0) > x\}$.

A sorban állási folyamat a következőképp dekomponálható:

$$Q(x) = \Pr\{V^{packet}(0) + V^{burst}(0) > x \mid V^{burst}(0) > 0\} \cdot \Pr\{V^{burst}(0) > 0\} + \Pr\{V^{packet}(0) > x \mid V^{burst}(0) = 0\} \cdot \Pr\{V^{burst}(0) = 0\} \quad (1)$$

A [14]-ben megmutatják, hogy ha a terhelés börszt komponense pozitív, akkor általában elmondható, hogy a csomag komponens várható értéke, $\boldsymbol{E}\{V^{packet}(0) \mid V^{burst}(0) > 0\}$, és szórásnégyzete, $\boldsymbol{Var}\{V^{packet}(0) \mid V^{burst}(0) > 0\}$, elég kicsi a börsztkomponens értékéhez képest. Ezért a következő közelítést alkalmazhatjuk:

$$\Pr\{V^{packet}(0) + V^{burst}(0) > x \mid V^{burst}(0) > 0\} \cdot \Pr\{V^{burst}(0) > 0\} \approx \Pr\{V^{burst}(0) > x\} \quad (2)$$

Tudjuk azonban, hogy a forgalomleírók, α_i, b_i és TTI_i, nem írják le az ON és OFF periódusok hosszát. Ez azt jelenti, hogy a börsztkomponens eloszlását, $\Pr\{V^{burst}(0) > x\}$, nem tudjuk kiértékelni.

Mivel rendszerünkben a késleltetéskövetelmények szigorúak $(\tilde{D}_i > TTI_i)$, a várakozási idő egy túlterhelési szituációban (amikor egy bizonyos ideig $R(t) > C$) hamar eléri a késleltetés kritériumot. Más szavakkal, ha a túlterhelési szituációk nem elég rövidek, a tároló nem tudja kisimítani hatékonyan az időleges túlterheléseket, még akkor sem, ha a tároló mérete végtelen. Ezért azt a konzervatív feltételezést tesszük, hogy, hogy minden túlterheléskor érkező csomag késleltetése nagyobb, mint a késleltetési kritérium. Használva ezt a feltételezést, a börsztkomponens eloszlása helyett annak a valószínűségét fogjuk kiértékelni, hogy egy csomag túlterheléskor érkezik, és ezzel a valószínűséggel fogjuk közelíteni a csomagvesztés valószínűségét.

Ha a börsztkomponens nulla, $V^{burst}(0) = 0$, a tároló periodikusan kiürül (TTI^{max} periódussal), és a terhelés eloszlását a vizsgált $t = 0$ időpont előtti rövid intervallumban érkező csomagok határozzák meg.

A javasolt modell

Feltételezéseinket alkalmazva a következő modellt állítjuk fel. A késleltetés kritérium kétféle megsértését különböztetjük meg: (1) csomagok elvesznek, ha a tároló tele van, és (2) csomagok késleltetést szenvednek, de nem vesznek el. Ennek megfelelően, két mennyiséget definiálunk:

$$\hat{\varepsilon}_i^{lost} = \frac{\text{elveszett csomagok száma}}{\text{összes csomag száma}} \quad \text{és} \quad \hat{\varepsilon}_i^{delayed} = \frac{\text{késő csomagok száma}}{\text{összes csomag száma}}\,.$$

Jelöljük az i osztályba tartozó aktív kapcsolatok számát (kapcsolatok ON periódusban) a t időpontban $N_i^{act}(t)$-vel, és az aktív kapcsolatok vektorát $\underline{N}^{act}(t) = [N_1^{act}(t), N_2^{act}(t), \ldots, N_K^{act}(t)]$-vel. Azt mondjuk, hogy egy rögzített t_0 időpontban a rendszer az $\underline{n}$ állapotban van, ha a véletlen vektor, $\underline{N}^{act}(t_0)$, az $\underline{n}$ értéket veszi fel (tehát, $N_i^{act}(t_0) = n_i$; $i = 1, 2, \ldots, K$). Az i osztályú aktív kapcsolatok száma binomiális eloszlású, jelölése $\Pi_i(n_i) = \Pr\{N_i^{act}(t_0) = n_i\}$, az állapoteloszlás, $\Pi(\underline{n})$, többdimenziós binomiális eloszlású:

$$\Pi(\underline{n}) = \prod_{i=1}^{K} \Pi_i(n_i) = \prod_{i=1}^{K} \binom{N_i}{n_i} \alpha_i^{n_i} (1-\alpha_i)^{N_i - n_i}$$

Egy aktív kapcsolat „súlyát" a csomagméret és a periódusidő hányadosaként definiáljuk: $\rho_i = b_i / TTI_i$. A bejövő ráta az $\underline{n}$ állapotban:

$$R(\underline{n}) = \sum_{i=1}^{K} n_i \rho_i$$

Annak a valószínűsége, hogy egy i osztályú csomag túlterhelési szituációban érkezik (amikor $R(n) > C$), a következő:

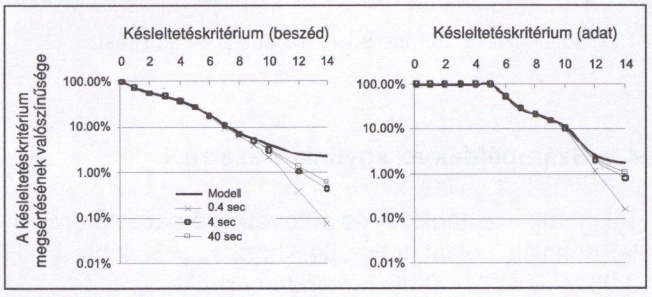
$$\varepsilon_i^{lost} = \Pr\{\text{egy csomag túlterhelési szituációban érkezik}\} = \frac{\sum_{\underline{n}: R(\underline{n}) > C} n_i \Pi(\underline{n})}{\sum_{\forall \underline{n}} n_i \Pi(\underline{n})} \approx \hat{\varepsilon}_i^{lost} \quad (3)$$

„Normális" szituációban (ha a rendszer nincs túlterhelve, azaz $R(n) \leq C$) periodikus csomagérkezések határozzák meg a késleltetést. Annak a valószínűsége, hogy egy t_0 időpontban érkező csomag késleltetése nagyobb, mint a késleltetés kritériuma, úgy, hogy a csomag nem vész el:

$$\varepsilon_i^{delayed} = \frac{\sum_{\underline{n}:R(\underline{n})\le C} n_i \Pi(\underline{n}) \cdot \Pr\{D_i > \tilde{D}_i \mid \underline{N}^{act} = \underline{n}\}}{\sum_{\forall \underline{n}} n_i \Pi(\underline{n})} \approx \hat{\varepsilon}_i^{delayed}. \quad (4)$$

Végül, hasonlóan az (1) egyenlet esetén látott dekompozícióhoz, a késleltetési kritérium megsértésének valószínűsége két valószínűség összegeként írható: $\varepsilon_i = \varepsilon_i^{lost} + \varepsilon_i^{delayed}$.

A javasolt modell helyességének ellenőrzése céljából különböző ON-OFF forrásokkal szimuláltunk. Az 5. ábrán Markov modulált források csomagkésleltetéseit hasonlítjuk össze a javasolt modell által számított eloszlással. A Markov modulált források ON periódus hosszának átlagai 0.4 sec, 4 sec és 40 sec. Egészen 10 ms késleltetéskritériumig a késleltetések alig függnek az ON periódusok hosszától, a kritérium megsértésének valószínűségét a rövid távú, periodikus viselkedés dominálja. Nagyobb kritériumok esetén a javasolt modell konzervatív, az eloszlás alakulását a túlterhelési szituációk dominálják.



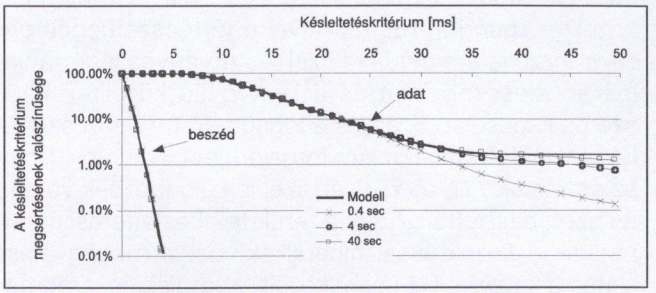
5. ábra Szimulációs eredmények C = 520 kbps link és 15 ms hosszú tároló esetén, 30 beszédkapcsolattal (TTI_1 = 20 ms, b_1 = 40 bytes, α_1 = 0.6) és 2 adatkapcsolattal (TTI_2 = 40 ms, b_2 = 360 bytes, α_2 = 1)

Prioritásos ütemező

A prioritásos ütemezés azt jelenti, hogy egy alacsonyabb prioritású sorban álló csomag csak akkor kerülhet kiszolgálásra, ha minden nála magasabb prioritású sor üres. Ha egy nagyméretű csomag egy alacsony prioritású sorból a kiszolgálóba kerül, a kiszolgálása alatt érkező magasabb prioritású csomagok nagy késleltetést szenvedhetnek miatta. Ezért szegmentálást alkalmaznak, a szegmensméret s. Ekkor Q_i és S_i az egész csomag helyett csak utolsó szegmensre vonatkoznak. Fontos különbség a FIFO-hoz képest, hogy a Q_i eloszlását nem lehet a terhelésfüggvény segítségével számolni, mert magasabb prioritású csomagok megelőzhetik az alacsonyabb prioritásúakat. Az ε_i^{lost} és az $\varepsilon_i^{delayed}$ számítása a FIFO-tól annyiban különbözik, hogy az i osztály szempontjából a rendszer csak akkor van túlterhelve, ha a magasabb (vagy ugyanolyan) prioritású forgalom rátája nagyobb, mint a C kapacitás.

Hasonlóan a FIFO esethez, a késleltetéskövetelményt megsértő csomagokat kétfelé bontjuk: elveszett és csak késleltetett csomagokra. A 6. ábrán a modell ellenőrzésére készített szimulációs eredmény látható, prioritásos ütemezőt feltételezve.

Megállapítható, hogy a modell (kis tárolókkal) ebben az esetben is konzervatív. A prioritásos ütemező miatt a beszédcsomagok sokkal kisebb késleltetést szenvednek, mint a FIFO esetében.



6. ábra Szimulációs eredmények, a modell paraméterei ugyanazok, mint az 5. ábrán, a szegmensméret s = 47 byte, a beszéd magasabb prioritású, mint az adat

4. LAC algoritmus FIFO és prioritásos ütemezővel

A LAC algoritmus feladata, hogy egy új kapcsolat érkezésekor ellenőrizze, hogy minden, a rendszerben lévő kapcsolatra teljesülnek-e a QoS követelmények. A kapcsolatok azon halmazát, melyre a követelmények teljesülnek, *elfogadható területnek* (admissible region) nevezzük. A modellnek megfelelően az $\tilde{\varepsilon}_i$ követelményt két részre osztjuk: $\tilde{\varepsilon}_i = \tilde{\varepsilon}_i^{lost} + \tilde{\varepsilon}_i^{delayed}$.

Ebben a cikkben a (3) egyenletet használjuk a csomagvesztés ($\tilde{\varepsilon}_i^{lost}$) kiértékelésére. Nagy kapacitású linkekre lassú lehet a számítás (komplexitása $O(N^K)$). Gyors közelítések találhatók az irodalomban, például [19]-ben és korábbi munkánkban, [22]-ben. A következőkben az $\varepsilon_i^{delayed}$ valószínűségszámolásával foglalkozunk.

Legyen $Q(x)$ a terhelés eloszlásfüggvénye egy FIFO sorban. Az érkezési folyamat növekményei stacionáriusak, ezért $Q(x)$ a következő általános sorbanállás elméleti eredmény segítségével határozható meg [16]:

$$Q(x) = \Pr\left\{V^{packet}(0) > Cx\right\} = \Pr\left\{\sup_{\tau \ge 0}\left(A(\tau) - C\tau\right) > Cx\right\}. \quad (5)$$

Mivel $Q(x)$ meghatározásakor csak nem túlterhelt eseteket vizsgálunk ($R(\underline{n}) \le C$), amikor a sor periodikusan kiürül, csak a $\tau \in [0, TTI^{\max}]$ intervallumot tekintjük. Ha az i osztályból sok független, α_i=1 aktivitásfaktorú, periodikus forrást összegzünk, az így nyert érkezési folyamat közelíthető egy Brown-híddal [18]. A Brown-híd közelítés alkalmazása lehetővé teszi az (5) egyenlet zárt alakú megoldását [17], [18]:

$$Q_i(x) = \exp\left\{-\frac{2Cx}{TTI_i N_i \rho_i^2}\left(\frac{Cx}{TTI_i} + C - N_i \rho_i\right)\right\}, \quad (6)$$

Ha több különböző forgalmi osztályunk van, de mindegyiknek azonos a csomagérkezési periódusa

($TTI_i=TTI_i, i=1,...,K,$), az elfogadható terület hipersíkkal közelíthető (N_i-ben):

$$Q_i(x) \le \varepsilon \Leftrightarrow \sum_i N_i \left(\rho_i + \frac{\rho_i^2}{C} \frac{\gamma\, TTI}{2x} \right) \le C + \frac{C\,x}{TTI}, \tag{7}$$

ahol $\gamma=-\ln(\varepsilon)$, és minden i-re $\tilde{\varepsilon}_i^{delayed}=\varepsilon$. A Brown-híd közelítés igen jó [16], és mivel a (6) összefüggés egzakt, igazolja a hipersík- közelítés érvényességét a valós érkezési folyamat és a Brown-híd közelítés közti szoros kapcsolat. Számos elfogadható területet szimuláltunk UTRAN-specifikus forgalomleírókkal, úgy, hogy az osztályok TTI_i és $\alpha_i \le 1$ értékei is különbözőek voltak, és azt tapasztaltuk, hogy a területek késleltetéslimitált határai jó közelítéssel hipersíkok. Ezt a megfigyelést alátámasztják a [15]-ben közölt szimulációs eredmények, valamint részletes analízisünk a [22]-ben.

Mivel a hipersík-közelítést (inhomogén forgalomra is) jónak találtuk, a LAC algoritmusban csak a hipersíkok tengelymetszeteit számoljuk. Legyen TN_{ij} az i osztályú kapcsolatok maximális száma, feltéve, hogy egy, a j. osztályból érkező csomag teljesíti a j osztály késleltetés követelményét ($\Pr\{D_j > \tilde{D}_j\} \le \tilde{\varepsilon}_i^{delayed}$). TN_{ij}-vel közelítjük az i osztályú kapcsolatok maximális számát, ha még egy j osztályú kapcsolat is a rendszerben van. (További részletek a [22]-ben találhatók.) A TN mátrix alkalmazásával az $(N_1,N_2,...,N_K)$ kapcsolatok elfogadásának szükséges feltétele a következő:

$$\sum_{i=1}^{K} \frac{TN_{jj}}{TN_{ij}} \cdot N_i \le TN_{jj} + 1 \qquad j=1,2,...,K;\ N_j > 0. \tag{8}$$

A TN_{ij} értékek meghatározásánál három esetet különböztetünk meg az i osztály és a j osztály prioritásszintjétől függően.

Ha az *i és j osztályok azonos prioritásúak* (illetve FIFO ütemező esetén), a TN_{ij} értékeit a (6) összefüggés segítségével határozhatjuk meg: $(\Pr\{D_j > \tilde{D}_j\} = Q_i(\tilde{D}_j - b_j / C)$. Ezt a (4) egyenlettel kombinálva számítjuk TN_{ij} értékeit. (Alternatív formulákat a [22]-ben közöltünk.)

Ha *i osztály magasabb prioritású j osztálynál*, más módszert használunk. Az esemény, hogy a j osztályú csomag utolsó, s_{last} méretű szegmense $\tilde{D}_j$ időn belül nem kap kiszolgálást, ekvivalens azzal, hogy az összes szegmens nem szolgálódik ki $D'=\tilde{D}_j - s_{last}/C$ idő előtt. Jelölje $B(t)$ a kiszolgáló j osztályú csomag által látott rendelkezésre állási idejét a $[0,t]$ periódusban, ha a j osztályú csomag a 0 időpontban érkezett. Ekkor:

$$\Pr\{D_j > \tilde{D}_j\} = \Pr\left\{ B\left(0, \tilde{D}_j - \frac{s_{last}}{C}\right) < \frac{b_j - s_{last}}{C} \right\}. \tag{9}$$

A kiszolgáló rendelkezésre állási idejére a következő (konzervatív) becslést használjuk: $B(0,t) \approx t - A_i(0,t) / C$, ahol $A_i(t)$ a magasabb prioritású (i osztályú) forgalom a $[0,t]$ intervallumban. Az érkezési folyamatra folytonos, gaussi közelítést (azaz a Brown-híd közelítést) alkalmazva a következő eredményre jutunk:

$$\Pr\{D_j > \tilde{D}_j\} \approx \Phi\{b_j - s_{last};$$
$$(C - N_i\rho_i)D',\ N_i\,\rho_i\,D'(TTI_i - D')\}, \tag{10}$$

ahol $\Phi\{x;\mu,\sigma^2\}$ a normális eloszlás μ várható értékkel és σ^2 szórásnégyzettel. Az 1. táblázat a (10) közelítéssel számolt TN_{ij} értékeket hasonlít össze egzakt értékekkel. A példa paraméterei: C=920 kbps, b_i=320 bit, s_{last}=320 bit, TTI_i=20 ms, $\tilde{D}_j$=10 ms, $\tilde{\varepsilon}_i^{delayed}$= 0.1%.

Ha *i osztály alacsonyabb prioritású j osztálynál*, elhanyagoljuk az esetlegesen a kiszolgálóban levő i osztályú csomagszegmens j osztályú csomagoknak okozott késleltetését. (Ez gyakorlatilag jó közelítés. Az általunk elhanyagolt hatást figyelembe vevő precíz megoldást mutatnak be [20]-ban.) Ez azt jelenti, hogy TN_{ij} értékeket nem számolunk erre az esetre ($TN_{ij}=\infty$ a (8)-ban).

b_j [bit]	320	640	960	1920	2880	3840
TN_{ij}^{exact}	41	38	36	30	25	20
TN_{ij}^{approx}	37	36	34	29	24	20

1. táblázat A (10) formula pontosságát érzékeltető példa

4.1. Számpéldák az egylinkes esetre

Tekintsük a 2. táblázat és a következő paraméterek által definiált példát: C=1504 kbps, s_{last}=384 bit, $\tilde{\varepsilon}_i^{lost}$= 0.05 %, $\tilde{\varepsilon}_i^{delayed}$= 0.05 % minden i-re.

RAB típus	TTI [ms]	b [bit]	α	prioritásszint
beszéd	20	336	0.55	magas
64k RAB	20	1480	1	alacsony
384k RAB	10	4360	1	alacsony
DCCH	40	216	0.2	magas
PCH	10	480	0.5	alacsony
FACH1	10	432	0.5	alacsony
FACH2	10	456	0.5	alacsony

2. táblázat A példa forgalmi osztályai

A jelzőcsatornákról azt feltételezzük, hogy a rendszerben állandóan van 3 PCH, 3 FACH1 és 3 FACH2 kapcsolat, míg a DCCH-k száma megegyezik a beszéd, 64k RAB és 384k RAB kapcsolatok számának összegével. (A jelzőcsatornák szerepéről lásd a [3].)

	0 of 384k RAB	**1** of 384k RAB	**2** of 384k RAB
0 of 64k RAB	90, 73, 83, 90	43, 39, 53, 53	--, 8, 16, 16
1 of 64k RAB	84, 65, 76, 84	33, 33, 46, 46	--, 1, 10, 10
2 of 64k RAB	77, 60, 71, 77	24, 28, 40, 40	--, --, 4, 4
3 of 64k RAB	71, 56, 67, 71	14, 24, 34, 34	
4 of 64k RAB	64, 51, 63, 64	4, 20, 27, 27	
5 of 64k RAB	58, 47, 58, 58	--, 15, 21, 21	
6 of 64k RAB	52, 43, 52, 52	--, 11, 15, 15	
7 of 64k RAB	45, 38, 45, 45	--, 6, 9, 9	
8 of 64k RAB	39, 34, 39, 39	--, 2, 3, 3	
9 of 64k RAB	30, 29, 33, 33		
10 of 64k RAB	21, 25, 26, 26		
11 of 64k RAB	11, 20, 20, 20		
12 of 64k RAB	1, 14, 14, 14		
13 of 64k RAB	--, 8, 8, 8		
14 of 64k RAB	--, 2, 2, 2		

3. táblázat A beengedett beszédkapcsolatok száma ***FIFO, PRIO-7.5, PRIO-10*** és ***PRIO-15*** esetekben

A következő FIFO és prioritásos rendszereket vizsgáljuk: ***FIFO*** (késleltetéskövetelmény: 5 ms), ***PRIO-7.5*** (késleltetéskövetelmények: magas: 5 ms, alacsony: 7.5 ms), ***PRIO-10*** (késleltetéskövetelmények: magas: 5 ms, alacsony: 10 ms), ***PRIO-15*** (késleltetés követelmények: magas: 5 ms,: alacsony: 15 ms). Az elfogadható tartományokat a 3. táblázat tartalmazza. Az eredmények az előző részben ismertetett algoritmusokkal lettek kiszámolva. Az elfogadható tartományokat szimulálva megállapítottuk, hogy a LAC algoritmusok konzervatívak, azaz a QoS követelmények minden forgalmi helyzetben teljesülnek.

Az eredményekből látszik, hogy a prioritásos ütemező használatakor, ahogy az alacsony prioritású forgalomkésleltetés követelményén lazítunk, az elfogadható alacsony prioritású kapcsolatok száma nő. Az elfogadható magas prioritású kapcsolatok száma csökkenhet, ha az alacsony prioritású forgalom késleltetés követelménye közelíti a magas prioritásúét. Ebben a példában a prioritásos ütemezés 7.5 ms alacsony prioritású késleltetéskövetelménynél még nem, de 10 ms-nál már előnyösebb a FIFO-nál.

5. Többlinkes eset

A következőkben a 2. részben bemutatott kapcsolási alternatívák használatakor elérhető statisztikus multiplexálási nyereséget vizsgáljuk, illetve számpéldákkal alátámasztva keressük a választ arra a kérdésre, hogy AAL2 kapcsolók használata esetén a vég-vég QoS követelményeket lehet-e a LAC algoritmusokkal megfelelően teljesíteni.

Kapcsolási alternatívák hatékonysága

Ha egy új AAL2 kapcsolat érkezik, elfogadásáról LAC algoritmusok döntenek minden csomópontban, ahol AAL2 multiplexálás vagy kapcsolás történik az AAL2 kapcsolatnak kijelölt útvonal mentén. Habár a $\tilde{D}_i$ késleltetéskövetelmény végpontok között van előírva, *minden LAC-ban ezt az értéket állítjuk be.* A LAC-ban beállított C kapacitás egy CBR VC vagy (UBR VC-k használata esetén) egy CBR VP kapacitását reprezentálja. Egy AAL2 kapcsolat beengedésre kerül, ha $\mathrm{PR}\{D_i > \tilde{D}_i\} < \varepsilon_i$ minden osztályra, minden LAC-nál.

A rendszerben a kapcsolások maximális blokkolási valószínűségeire is előírások vannak. A különböző osztályú kapcsolatok számainak valószínűségei (és így a blokkolási valószínűségek is) függnek a kapcsolatok érkezési folyamatától. A blokkolások természetesen attól is függnek, hogy milyen CAC algoritmust alkalmazunk. Ha több VC (vagy VP) is rendelkezésre áll két csomópont között, a kapcsolatok ezen erőforrások feletti elosztásának stratégiája is befolyásolja a blokkolási valószínűségeket.

A blokkolási valószínűségek számításához egyszerű Markov-láncok (mint pl. az Erlang-B formula kiterjesztései) kevéssé használhatók, mert: (1) egy LAC által számolt elfogadható tartomány tipikusan nem egy hipersík [22], (2) a VC-k (vagy VP-k) feletti kapcsolat elosztási stratégiák figyelembevétele nehézkes, (3) a kapcsolatok érkezési folyamata általában nem Poisson, valamint (4) a blokkolási valószínűségek a különböző hierarchiaszinten lévő linkeken nem függetlenek.

Mivel a LAC algoritmus megfelelően gyors (on-line működik az UTRAN-ban), jól alkalmazható egy kapcsolat szintű szimulátorban is. Ezért készítettünk egy ilyen szimulátort, melyben a csomag szintű forgalomleírók a generált kapcsolatok attribútumai. A forgalomkoncentrációs pontokban LAC algoritmusok futnak, és az általuk blokkolt hívások arányát mérjük.

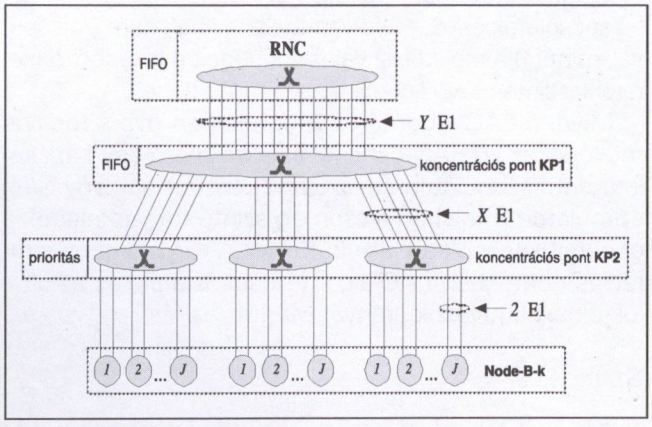
Számpélda

A példahálózat a 7. ábrán látható. A „downlink" irányú forgalmat vizsgáljuk, ami azt jelenti, hogy a kapcsolatok az RNC-től a bázisállomások felé generálódnak. A hierarchiaszinteket nevezzük L1, L2 és L3-nak (L1: Node B-k és KP2 között, L2: KP2 és KP1 között, L3: KP1 és RNC között). A forgalom egyenletesen oszlik meg a bázisállomások között.

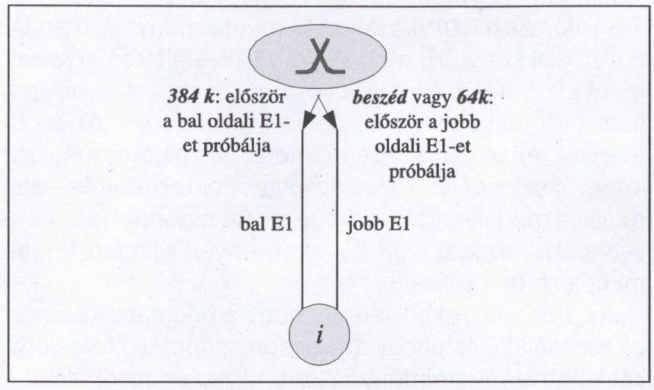
Az L1 szinten 2 E1-et használunk bázisállomásonként. A fragmentáció csökkentése érdekében a 8. ábrán látható kapcsolat elosztási mechanizmust, a hatékonyság növelése érdekében prioritásos ütemezőt alkalmazunk. Az L2 és L3 szinteken nagy sebességű fizikai linkeket (pl. 155 Mbps) használunk, és FIFO ütemezőt alkalmazunk. Az egyszerűség kedvéért azt feltételezzük, hogy a VP és VC kapacitások az E1 kapacitás többszörösei lehetnek. Így az L2 és L3 szinteken X, illetve Y „E1"-et használunk. Egy E1 bruttó kapacitása 1920 kbps. Ebben a példában csak felhasználói (user-plane) forgalmat generálunk, és azt feltételezzük, hogy az üzemeltetési és fenntartási (O&M, operation and maintenance) forgalom 212 kbps-t, a közös jelzésrendszer jelzései (FACH és PCH) 234 kbps-t foglalnak bázisállomásonként. Ha például CBR VC-ket használunk az L1-en, ezeknek a nettó kapacitása ATM fejléc nélkül $47/53 \cdot (2 \cdot 1920 - (212+234))/2 = 1504.8$ kbps. Azt feltételezzük, hogy ez a kapacitás (*1504.8* kbps) áll rendelkezésre minden „E1"-nél, minden szinten, és hogy megkapjuk a korrekt bruttó kapacitást, a jelzés és O&M többletet a számítás végén adjuk hozzá. Például, ha Y=10 és J=3 (lásd a 7. ábrát), akkor a bruttó kapacitás az L3-on $53/47 \cdot Y \cdot 1504.8 + 3 \cdot J \cdot (212+234) = 20983$ kbps.

A forgalomleírókat, QoS és blokkolási követelményeket a 4. táblázat tartalmazza. Adatforgalom esetén 64 kbps és 384 kbps sebességű RAB-okat feltételezünk. Nehéz feltételezéseket tenni arra vonatkozóan, hogy ezek a RAB-ok mikor és milyen hosszan kerülnek használatra, mert ez nagyban függ a rendszer általunk nem vizsgált egyéb tulajdonságaitól és az applikációk viselkedésétől. Ezért ebben a cikkben hipotetikus kapcsolat szintű forgalmi paramétereket használunk, a kapcsolatokat Poisson folyamat szerint generáljuk. Azt feltételezzük továbbá, hogy ha egy kapcsolat blok-

kolódik, nincs újrapróbálkozás (pl., ha egy 384 kbps igény blokkolódik, nincs újrapróbálkozás 64 kbps sebességen). Ezek a feltételezések természetesen nem szükségesek a szimulátor használatakor.



7. ábra A példahálózat



8. ábra Kapcsolatelosztás az erőforrás L1 fragmentációjának csökkentésére

Először *X*-et és Y-t úgy állítjuk be, hogy ne legyen blokkolás az L2 és L3 szinteken, és a RAB használók számát addig növeljük, míg a blokkolás valószínűsége el nem éri a blokkolási követelményt L1 szinten ($B_{i,L1}$-et). Ezután X-et csökkentjük addig, amíg $B_{i,L2}$-t el nem éri a blokkolás L2-n. Végül *Y*-t csökkentjük, amíg $B_{i,L3}$-at el nem érjük. Homogén forgalom esetére, a **B** kapcsolási alternatívát használva, az eredmények a 5. táblázatban találhatók. A táblázat szerint tisztán beszédforgalom esetén kevés nyereséget lehet elérni (pl. 18 helyett 15 VC L3 szinten, és nincs nyereség L2-n), míg nagyobb sebességű kapcsolatok esetén nagyok a nyereségek (pl. 18 helyett 9 VC L3-on tisztán 384 kbps forgalomra).

	Csomagszint			Kapcsolatszint	QoS köv.		Max. blokkolás		
	TTI_i [ms]	b_i [bit]	α_i	Felajánlott forg. / RAB használó [mErlang]	D_i [ms]	ε_i [%]	$B_{i,L1}$ [%]	$B_{i,L2}$ [%]	$B_{i,L3}$ [%]
beszéd	*20*	*336*	*0.55*	*15*	*5*	*0.1*	*0.1*	*0.2*	*0.3*
64k	*20*	*1480*	*1*	*2.775*	*15*	*0.1*	*1*	*2*	*3*
384k	*10*	*4360*	*1*	*0.925*	*15*	*0.1*	*1*	*2*	*3*

4. táblázat A példa paraméterei

Heterogén forgalom esetén (ahol a RAB használók aránya: 80% beszéd, 18% 64 kbps és 2% 384 kbps) tételezzük fel először azt, hogy ha a **B** vagy a **C.2** alternatívát használjuk KP1-ben és az RNC-ben, akkor **B**-t használjuk KP2-ben. Így a RAB használók száma bázisállomásonként 8700-ra adódik (104.4 Erlang beszéd, 4.34 Erlang 64 kbps és 0.16 Erlang 384 kbps). Az eredményeket a 6. táblázat mutatja. A nyereségeket a következőképpen számoltuk: *bruttó kapacitás **A**-val mínusz bruttó kapacitás **B**-vel (vagy **C.2**-vel), az egész osztva bruttó kapacitás **A**-val.*

Ezután tételezzük fel azt, hogy ha C.2-t használjuk KP1-ben és az RNC-ben, akkor szintén C.2-t használunk KP2-ben. Így a RAB használók száma bázisállomásonként 9600-ra nő. Ha J = 3, akkor X = 5 és Y = 11, és a nyereségek L2 és L3 szinteken 15% és 34%.

RAB típus	RAB használók száma / Node B	X	Y
beszéd	*14634*	*6*	*15*
64k	*10527*	*5*	*14*
384k	*2105*	*4*	*9*

5. táblázat Eredmények homogén forgalom esetén (J = 3)

	kapcsolási alternatíva	X	Y	L2 bruttó kapacitás	L3 bruttó kapacitás	nyereség *A*-hoz képest L2-n	nyereség *A*-hoz képest L3-n
3	***B***	*5*	*12*	*29447 bit/s*	*24359 bit/s*	*15 %*	*29 %*
	C.2	*4*	*11*	*24359 bit/s*	*22663 bit/s*	*29 %*	*34 %*
6	***B***	*9*	*23*	*53805 bit/s*	*47021 bit/s*	*22 %*	*32 %*
	C.2	*8*	*19*	*48718 bit/s*	*40238 bit/s*	*29 %*	*42 %*
9	***B***	*13*	*32*	*78165 bit/s*	*66293 bit/s*	*25 %*	*36 %*
	C.2	*11*	*27*	*67988 bit/s*	*57813 bit/s*	*34 %*	*44 %*

6. táblázat Eredmények inhomogén forgalom esetén, a **B** és **C.2** alternatívák statisztikus multiplexálási nyereségei **A**-hoz képest

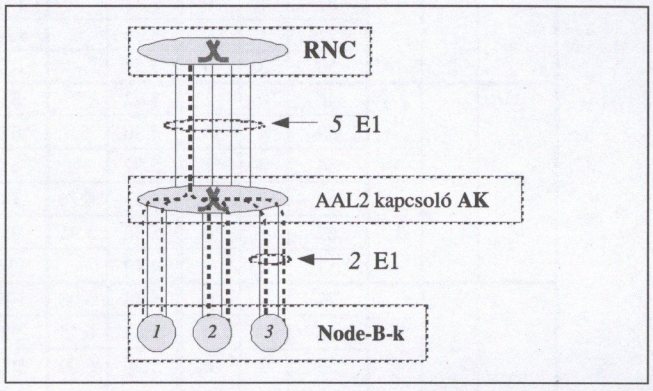
A 6. táblázat eredményei azt mutatják, hogy AAL2 kapcsolók használatával jelentős nyereséget lehet elérni, még akkor is, ha a beszédkapcsolatok dominálnak a forgalomban. A leghatékonyabb megoldás UBR VC-k használatával érhető el (**C.2** alternatíva). Ennél a megoldásnál figyelni kell azonban arra, hogy csak a LAC által figyelt UBR VC-k osztozzanak a VP erőforrásain. Például, ha egy operátor VC-ket bérel, **C.2** nem használható. Az eredményekből megfigyelhetjük azt is („nyereség **A**-hoz képest L2-n" a 6. táblázatban), hogy még akkor is jelentős nyereségeket lehet elérni, ha nincs koncentrálás L3-on (az RNC és KP2 a VC-k végpontjai, és $Y = 3 \cdot X$).

Vég-vég csomagkésleltetés

Az UTRAN-ban nincs lehetőség arra, hogy egy olyan CAC algoritmust futtassunk, amely a teljes hálózat állapotát ismeri. Ehelyett minden csomópontban, ahol AAL2 multiplexálás történik, LAC algoritmusokat futtatnak. Ezért a vég-vég késleltetéskövetelményt linkekre vonatkozó követelményekbe kellene transzformálni. Mivel az AAL2 kapcsolóknak nincs információjuk a hálózati topológiáról és a hálózati erőforrásokról, ezt a transzformációt nem tudják végrehajtani. Másrészt, a késleltetéskövetelmények valószínűségi megfogalmazása és a különböző linkeken a késleltetések összefüggősége miatt a vég-vég késleltetés követelményeket nem lehet egyszerűen elosztani a linkek között. Mi azt feltételezzük, hogy a legtöbb forgalmi szituáció esetén a vég-vég késleltetésben egy szűk keresztmetszetű link hatása dominál, ezért minden LAC-ba a vég-vég késleltetés követelményt konfiguráljuk. Ebben a részben megmutatjuk, hogy ez a közelítés a gyakorlatban megfelelő.

Az UTRAN-ban az AAL2 kapcsolókkal felszerelt forgalomkoncentrálási szintek száma tipikusan 2-3. Mi egy egyszerű kétszintű hálózati elrendezést használtunk (9. ábra) a vég-vég késleltetések ellenőrzéséhez. A két szint a következő: L1: Node Bs - AK, L2: AK - RNC. Mindkét szinten FIFO ütemezést használunk. A VC-k kapacitása 1504 kbps. A forgalomleírók és QoS követelmények ugyanazok, mint a 4. táblázatban. AK-ban a **B** kapcsolási alternatívát használjuk.

Amikor egy beszédkapcsolat egy kiválasztott L2 VC-n blokkolódik, a kapcsolat szintű szimulátor rögzíti a rendszer állapotát (kapcsolatok száma az egyes linkeken stb.). Ezt az információt képes felhasználni egy csomag szintű szimulátor, mely a vég-vég csomagkésleltetéseket méri azokon a kapcsolatokon, melyek a kiválasztott L2 VC-t használják. A 9. ábra példáján balról a második VC-t választottuk L2 szinten, és a szaggatott vonalak mentén haladó csomagok késleltetését mértük. A kiválasztott L2 VC-n a LAC akkor blokkol egy beszédkapcsolatot, ha ez a VC „tele van", azaz, az éppen érkező kapcsolat már nem fér be az elfogadható területbe. De ebben az esetben a L1 VC-k nem feltétlen vannak tele. A kapcsolat szintű szimulátor által rögzített eseteken kívül extrém eseteket is szerettünk volna megvizsgálni, amikor minkét szint szűk keresztmetszet. Ezért a rögzített esetekben addig növeltük a kapcsolatok számát az L1 szinten, amíg az L1 szint összes linkje tele nem lett. Ezeket a hozzáadott kapcsolatokat nem tudná a rendszer kiszolgálni, hiszen L2 szinten kevesebb VC van, mint L1 szinten! A szimulátorban ez nem okoz problémát, mert az L2 szinten csak a kiválasztott VC forgalmát szimuláljuk.



9. ábra A vizsgált hálózati elrendezés

A 7. táblázat három szimulációs példa eredményeit mutatja. Mindegyik példa esetén megvizsgáltuk (1) a „normál esetet", amikor a kapcsolatszintű szimulátorból kinyert állapotot szimuláltuk csomagszinten, (2) azt az esetet, amikor az L1 szintű VC-ket beszédkapcsolatokkal telítettük, (3) végül azt az esetet, amikor az L1 szintű VC-ken először a 384 kbps kapcsolatok, majd a 64 kbps, és végül a beszédkapcsolatok számát növeltük addig, amíg lehetett. A vég-vég késleltetések 0.1 % kvantilis értékeit tartalmazza a táblázat.

Amint azt a 4. részben bemutattuk, a LAC által meghatározott elfogadható terület kétféle terület metszetéből adódik, a késleltetéslimitált és a csomagvesztés-limitált területekéből. A 7. táblázat példáit úgy választottuk, hogy a három példában az L2 szinten blokkolt beszédkapcsolat különböző okok miatt került blokkolásra. Az első esetben csak a csomagvesztés ($\tilde{\varepsilon}_i^{lost}$), a másodikban a csomagvesztés és a késleltetés ($\tilde{\varepsilon}_i^{delayed}$), míg a harmadikban csak a késleltetés korlátja limitált.

Az eredmények azt mutatják, hogy még ezekben a ritka, illetve extrém esetekben is mindig kisebb volt a maximum vég-vég késleltetés, mint 7 ms, ami a gyakorlatban elfogadható. A 64 kbps kapcsolatok késleltetései jóval a követelmény (15 ms) alatt voltak, mert FIFO ütemezőt használtunk. Prioritásos ütemezőt használva, illetve L2 szinten nagyobb sebességű VC-kkel a beszédcsomagok késleltetései még kedvezőbbek lennének. A vég-vég késleltetéseket úgy is csökkenthetjük a gyakorlatban, hogy a LAC-okban a valós követelménynél kissé szigorúbbat állítunk be.

6. Konklúzió

Munkánk során először bemutattuk az UTRAN rendszer sorban állási modelljét, majd erre építve kapcsolat-

a kapcsolatok száma a kiválasztott L2 VC-n			normál eset					a beszédkapcsolatok számát növeltük					A 384k, majd a 64k, végül a beszédkapcsolatok számát növ.				
			kapcsolatok száma az L1 VC-ken			vég-vég késleltetés [ms]		kapcsolatok száma az L1 VC-ken			vég-vég késleltetés [ms]		kapcsolatok száma az L1 VC-ken			vég-vég késleltetés [ms]	
beszéd	64k	384k	beszéd	64k	384k	beszéd	64k	beszéd	64k	384k	beszéd	64k	beszéd	64k	384k	beszéd	64k
128	0	0	50	1	0	3.21	-	121	1	0	4.21	-	50	1	1	5.22	-
			79	3	0	4	-	107	3	0	4.75	-	86	6	0	5.54	-
			67	1	0	3.35	-	121	1	0	4.33	-	67	8	0	5.72	-
			67	1	0	3.35	-	121	1	0	4.34	-	67	8	0	5.74	-
			71	1	0	3.36	-	121	1	0	4.29	-	78	7	0	5.87	-
			50	3	0	3.57	-	107	3	0	4.81	-	56	9	0	5.82	-
86	6	0	46	0	1	5.92	-	61	0	1	6.03	-	50	1	1	6.06	-
			59	8	0	6.36	6.9	67	8	0	6.55	7.05	67	8	0	6.51	7.07
			54	0	1	5.92	-	61	0	1	5.95	-	61	0	1	5.88	-
			63	7	0	6.31	6.74	78	7	0	6.92	7.2	67	8	0	6.64	7.09
			67	5	0	5.38	5.91	93	5	0	6.18	6.48	67	8	0	6.31	6.81
			57	2	0	4.49	-	114	2	0	5.72	-	67	8	0	6.21	-
23	12	0	77	3	0	5.43	6.08	107	3	0	6.02	6.56	78	7	0	6.57	7.16
			68	5	0	5.63	6.45	93	5	0	6.29	6.89	78	7	0	6.66	7.24
			55	3	0	5.36	6.09	107	3	0	6.25	6.72	56	9	0	6.71	7.36
			77	5	0	5.7	6.35	93	5	0	6.2	6.73	78	7	0	6.5	7.03
			57	0	1	6.11	-	61	0	1	6.15	-	61	0	1	6.13	-
			73	7	0	6.64	7.25	78	7	0	6.85	7.33	78	7	0	6.85	7.37

7. táblázat Csomag szintű szimulációs eredmények

engedélyezési algoritmusokat fejlesztettünk FIFO és prioritásos ütemezőkre. Feltételezve, hogy a javasolt engedélyezési algoritmusokat használja a rendszer, különböző kapcsolási (forgalomkoncentrálási) alternatívákat értékeltük, és azt találtuk, hogy az AAL2 kapcsolók használata jelentős hatékonyságnövekedést eredményezhet. Ha AAL2 kapcsolók vannak a rendszerben, minden AAL2 kapcsolóban kell engedélyezési algoritmusnak futnia, és egy kapcsolat csak akkor kerül elfogadásra, ha minden általa használt AAL2 kapcsolóban elfogadásra került. Magyarázatot adtunk arra, hogy miért lehet ebben az esetben is megfelelő vég-vég csomagkésleltetést biztosítani.

Hivatkozások

1. 3GPP, Delay Budget within the Access Stratum, TR 25.853, 2001.
2. 3GPP, IP Transport in UTRAN, TR 25.933, 2001.
3. 3GPP, Multiplexing and channel coding (FDD), TS 25.212, 2001.
4. ITU-T, AAL Type 2 Signalling Protocol (Capability Set 1 and 2), New ITU-T Rec. Q.2630.1 and Q.2630.2
5. The ATM Forum Technical Committee, Traffic Management Specification V. 4.0, 1996.
6. T. Ojanpera and R. Prasad, editors, Wideband CDMA for Third Generation Mobile Communications, Artech House, 1998.
7. G. Eneroth, et al, Applying ATM/AAL2 as a Switching Technology in 3G Mobile Networks, IEEE Comm. Mag., 37(6): 112–122, 1999.
8. G. Fodor, et al, Comparison of Call Admission Control Algorithms in ATM/AAL2 Based 3rd Generation Mobile Access Networks, Proc. IEEE WCNC, 1999.
9. K. M. F. Elsayed, N. Gerlich and P. Tran-Gia, Efficient Design of Voice Carrying Fixed-Network Links in CDMA Mobile Communication Systems, Telecom. Systems, 17(1,2): 9–29, 2001.
10. H. Saito, Bandwidth Management for AAL2 Traffic, IEEE Tr. on Vehicular Tech., 49(4): 1364-1377, 2000.
11. H. Saito, Performance Evaluation and Dimensioning for AAL2 CLAD, Proc. IEEE INFOCOM, pp. 153–160, 1999.
12. H. Saito, Effectiveness of UBR VC Approach in AAL2 Networks and Its Application to IMT-2000, IEICE Transactions on Communications, E83-B, 11, 2000.
13. J. Peisa and M. Meyer, Analytical Model for TCP File Transfers over UMTS, Proc. 3G Wireless 2001, 2001.
14. I. Norros, et al, The Superposition of Variable Bit Rate Sources in an ATM Multiplexer, IEEE JSAC, 9(3): 378–387, 1991.
15. L. He and A. Wong, Connection Admission Control Design for GlobeView-2000 ATM Core Switches, Bell Labs Technical Journal, pp. 94–110, Jan.-March, 1998.
16. Methods for the performance evaluation and design of broadband multiservice networks, Part III, Traffic models and queuing analysis, The COST 242 Final Report, 1996.

17. B. Hayek, A Queue with Periodic Arrivals and Constant Service Rate, in Kelly, F.P. (ed.) Probability, Statistics and Optimisation, a Tribute to Peter Whittle., Wiley, 147–157, 1994.
18. F. P. Kelly, Notes on Effective Bandwidths, Stochastic Networks: Theory and App., Vol. 4., Oxford University Press, 141–168, 1996.
19. A. M. Makowski, Bounding On-Off sources - Variability ordering and majorization to the rescue, ISR TR 2001-13, http://www.isr.umd.edu/TechReports/ISR/2001/TR_2001-13/TR_2001-13.pdf
20. K. Iida, et al, Delay Analysis for CBR Traffic under Static-Priority Scheduling, IEEE/ACM Transactions on Networking, 9(2), April, 2001.
21. A. Rácz, N. Fias, P. Rácz, Effective Bandwidth and Associated CAC Procedure for Traffic Streams Multiplexed over a VBR Link, Proc. SPECTS, pp. 72-79., Vancouver, 2000.
22. Sz. Malomsoky, S. Rácz and Sz. Nádas, Connection Admission Control in UMTS Radio Access Networks, submitted to Computer Communications, Elsevier Science, 2002.
23. Sz. Nádas, S. Rácz and Sz. Malomsoky, On Quality of Service Differentiation in UMTS Radio Access Networks, submitted to IEEE GLOBECOM, Taipei, Taiwan, 2002

Hírek

Új, vezeték nélküli internetszolgáltatás telefondíj nélkül
Az EnterNet széles sávú szolgáltatása a Cisco Systems vezeték nélküli technológiájára épül

Az EnterNet internetszolgáltató a Cisco Systems vezeték nélküli technológiájára épülő szolgáltatást indít, amely telefondíj nélkül kínál internetkapcsolatot. A technológia multipont rendszerű. Az új szolgáltatás bekapcsolásának átfutási ideje a megrendelést követően előreláthatóan két hét, nagymértékű túljelentkezés esetén ez az idő legfeljebb egy hónapra emelkedhet.

Az vezeték nélküli internetelérés augusztus 10-től a következő városokban lesz elérhető: Baja, Balatonboglár, Balatonfüred, Kalocsa, Kaposvár, Keszthely, Pápa, Salgótarján, Szentes. A fenti kilenc várost követően már augusztus folyamán további 15 város bekötését tervezik, majd szeptember és október hónapokban további 30 települést kívánnak a szolgáltatási körbe bevonni. Jelenleg az EnterNet által kiépített infrastruktúra 4 km-es sugarú körben biztosít internet-hozzáférést, de igény esetén, az eszközök láncba fűzésével a szolgáltatási terület tovább növelhető. A szolgáltatásba bevont, vagy bevonandó városokat és településeket az alternatív megoldások iránt mutatkozó piaci igények alapján választotta ki az EnterNet.

A különböző csomagok árának kialakításánál nagy hangsúlyt kapott, hogy a vidéki kis- és középvállalkozások kedvező áron – az eddigi 60-70 ezer forintos költség helyett már 20 ezer forintért – tudjanak a bérelt vonali hozzáférés minőségének megfelelő internetkapcsolathoz jutni, ezért az EnterNet a magyar internetpenetráción belül az üzleti szegmens súlyának 20%-os növekedését prognosztizálja.

Ennek a Magyarországon áttörő szolgáltatásnak a bevezetésére és technikai hátterének kiépítésére az EnterNet eszközgyártó partnert keresett. A Cisco Systems az a partner, amellyel az EnterNet belevágott az új, mikrohullámú internetszolgáltatás megvalósításába.

A vezeték nélküli internetes hozzáférést a Cisco Aironet 1200-as bázisállomás (access point) biztosítja a hozzá tartozó klienskártyákkal. A bázisállomásról küldött jeleket a felhasználói oldalon a klienskártyákba csatlakozó antennák fogadják, így biztosítva a számítógépek (asztali PC-k vagy hordozható számítógépek) internetes csatlakozását. A bázisállomás a vezeték nélküli hálózatokra kidolgozott, nemzetközileg legelterjedtebb 802.11b szabványának megfelelően a 2,4 GHz-es szabad frekvencián működik, és jelenleg 11 Mbs adatátviteli sebességet biztosít. A Cisco Aironet 1200-as rendszere a hamarosan piacra kerülő bővítőmodul segítségével képes lesz az 5 GHz-es frekvenciasávban akár 54 Mbs sebességgel is működni. A technológia telepítése a gyakorlatban egy vagy több kisugárzó antenna felszerelését jelenti az adott településen, illetve az előfizetői oldalon egy egyszerű telepítési eljárást kíván. A mikrohullámú technika lehetővé teszi, hogy az előfizetők alacsony költség mellett, telefondíj nélkül internetezhessenek.

Elliptikus görbén alapuló titkosírás

CSIRMAZ LÁSZLÓ

CEU

Reviewed

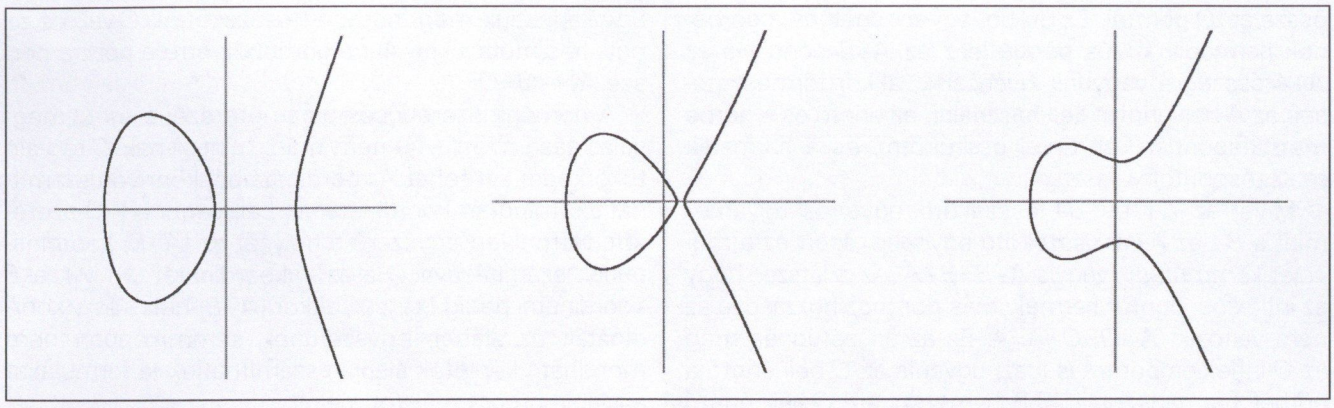
A nyilvános kulcsú titkosírás (általában: titkosírás) hátterében egy-egy matematikai probléma áll, melyről azt tételezzük fel, hogy nehéz megoldani. A *Ron L. Rivest, Adi Shamir* és *Leon Adleman* nevéhez fűződő RSA módszer azt használja, hogy míg viszonylag egyszerű eldönteni, hogy egy több száz jegyű szám összetett-e vagy sem, addig prímtényezőire bontani már reménytelenül nehéz feladat. *Whitheld Diffie* és *M. E. Hellmann* nevét viselő módszer hátterében az úgynevezett *diszkrét logaritmus* probléma áll: adott az alap és a hatványozás eredménye, keressük meg a kitevőt. A hatványozás természetesen ismételt szorzás, viszont a szorzást nem a szokásos, általános iskolában tanult módon kell elvégezni, hanem maradékosan, valamilyen előre meghatározott prímszám modulussal. A problémát általánosan is meg lehet fogalmazni, a következőképpen: definiáljunk egy szorzásnak nevezett műveletet sok (tipikusan 10^{100}) elem között. Választunk egy *g* alapot, egy n kitevőt, és kiszámítjuk a $y=g^n$ hatványt. A feladat *y* és *g* ismeretében megkeresni az *n* kitevőt. Mivel az *n* kitevő is 100 jegyű szám, az *y* hatvány kiszámítása sem lehetséges ismételt szorzással. Ezért még feltesszük, hogy a definiált művelet az adott elemeken *csoportot* alkot. Ilyenkor a hatványozást jóval gyorsabban el tudjuk végezni: a *g* generátorelemet ismételten négyzetre emeljük (ehhez egy-egy szorzást használva), amivel megkapjuk *g*-nek a 2^0, 2^1, 2^2, 2^3 stb. kitevős hatványait. Az *n* kitevőt kettes számrendszerben írjuk fel, és a jegyeinek megfelelő hatványokat szorozzuk össze. Ezzel a szorzások számát még száz jegyű kitevő esetén is 1000 alatt tudjuk tartani.

Diffie és Hellmann ismerték fel, hogy véges csoportokon a diszkrét logaritmus nehéz probléma, vagyis a kitevőt az alap és a hatvány értékéből nem lehet gyorsan visszakeresni. (Természetesen nem ez a helyzet a valós számok esetében, ekkor a kitevőt – vagyis *y*-nak g alapú logaritmusát – gyorsan ki tudjuk számítani.) Különböző speciális csoportok esetén más és más, igen bonyolult és mély matematikai eszközöket használó algoritmusokat fejlesztettek ki. Tetszőleges csoportra működő leggyorsabb ismert algoritmus futásideje a csoport elemszámának négyzetgyökével arányos. Vagyis 10^{100} elemű csoportnál az algoritmus futása 10^{50} körüli műveletet igényel, amit összevetve a világegyetem korára becsült becsült 10^{22} évvel, elég meggyőző érv a probléma nehézségére. Speciális csoportokon – így például egy *p* prímszámmal való maradékos szorzás esetén is – gyorsabb algoritmusok is ismeretesek. Abban az esetben, mikor a csoport éppen a modulo *p* *összeadás*, *g* pedig az 1 szám, igazán nincs is feladat, hiszen az eredmény és a kitevő (jelen esetben a szorzótényező, hiszen most a művelet-összeadás) egy és ugyanaz. Ez is mutatja, hogy kriptorendszer biztonsága szempontjából igen fontos a felhasznált csoport megfelelő megválasztása.

Kriptográfiai rendszerekben két típusú csoportot szokás használni. Az egyik a az 1, 2, ..., *p*-1 számokon a modulo *p* maradékos szorzás alkotta csoport. Ezekre a csoportokra ismeretes a fent említett $\sqrt{p}$ idejű általános algoritmusnál gyorsabb eljárás is. Ha a *p* prímszám ezenkívül még speciális alakú is, például *p*-1-nek (vagy *p*+1-nek) mindegyik prímosztója legfeljebb 10 jegyű, akkor a diszkrét logaritmus problémát emberi idő alatt (néhány hónap vagy év) is meg lehet oldani. A csoportoknak egy másik gazdag, és egyre többet használt forrása az *elliptikus görbék*. Általános elliptikus görbékből adódó csoportok esetén a diszkrét logaritmus problémára nem ismeretes a csoport méretének gyökénél kevesebb műveletet igénylő algoritmus. Természetesen vannak kifejezetten rossz, illetve „nem ajánlott" görbék, viszont elliptikus görbékből jóval több van, mint természetes számokból, nagyobb a választási lehetőség, kisebb esélyünk van arra, hogy egy rossz csoportba botlunk.

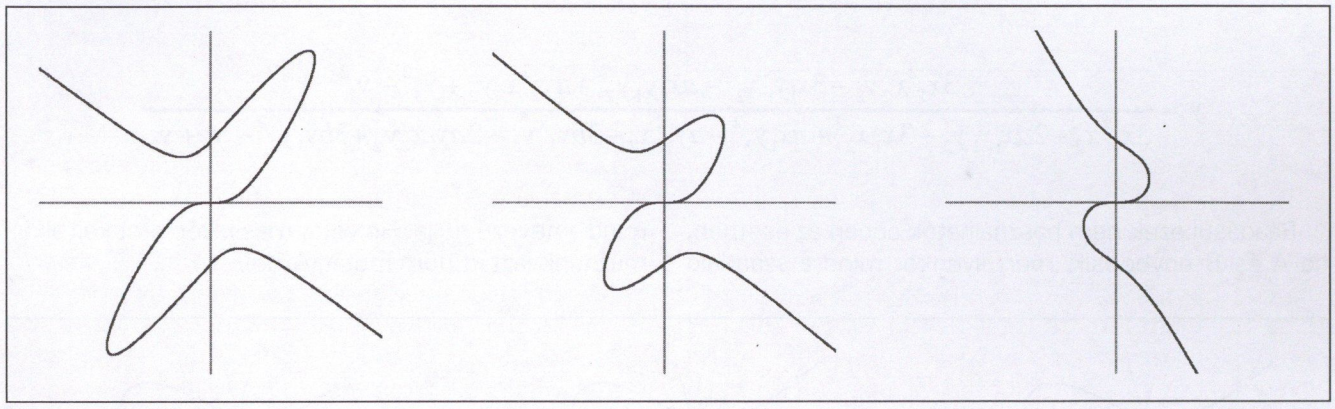
Az elliptikus görbék a középiskolából jól ismert kör, ellipszis, parabola, hiperbola általánosításai. Míg például az egységkör a síknak azokból az *(x, y)* koordinátájú pontjaiból áll, amikre $x^2+y^2=1$, addig például az $y^2=x^3+ax+b$ egyenletet kielégítő pontok egy harmadrendű algebrai görbét határoznak meg.

Az 1. ábrán különböző *a* és *b* paraméterek mellett mutatjuk be magát a görbét. A görbe maga egy vagy két részből áll, a jobb oldali ágai egyre meredekebben tartanak a végtelenbe. Azokat, melyek elmetszik saját magukat, vagy csúcsban végződnek, mint az $a=b=0$ esetben, *szinguláris görbéknek* nevezzük. A függőleges egyeneseket kivéve minden más egyenes vagy egy vagy három pontban metszi a görbét, ezért a függőle-



1. ábra $y^2=x^3+ax+b$ egyenletű elliptikus görbék különböző paraméterekkel

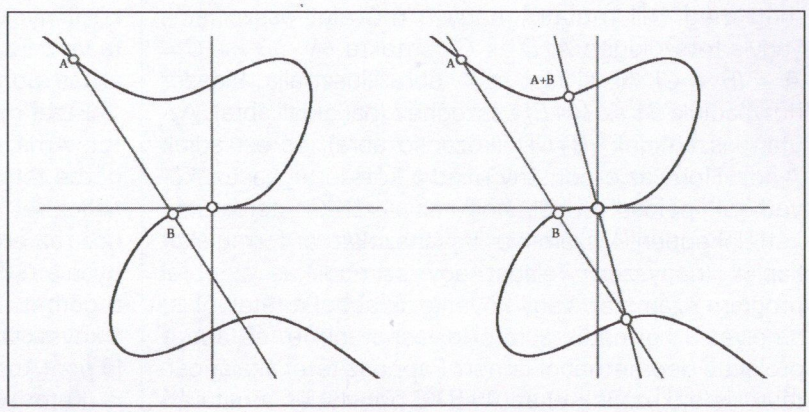
ges egyenesek ideális (végtelen távoli) pontját is hozzá szokás venni a görbéhez, hogy azok se legyenek kivételek. A görbének ezt a pontját egy projektív transzformációval az origóba vive a görbe formája is megváltozik:



2. ábra $x^3=y+axy^2+by^3$ egyenletű duális görbék, az inflexiós pont az origó

Egyetlen olyan pont van, ahol az érintő harmadrendben simul a görbéhez, ezt a görbe *inflexiós pontjának* nevezzük. Az $y^2=x^3+ax+b$ egyenletű görbéknél ez a függőleges egyenesek ideális pontja, az $x^3=y+axy^2+by^3$ görbéknél viszont az origó. Az ilyen harmadrendű kifejezések által meghatározott görbékkel, és általában egy kétváltozós polinomot kielégítő *(x, y)* pontok tulajdonságaival foglalkozik a matematika egyik legszebb, de ugyanakkor legnehezebb ága, az algebrai geometria. *Elliptikus görbék* a harmadrendű kifejezés által definiált nem szinguláris alakzatok. Ezek számtalan érdekes tulajdonságai közül azt használjuk fel, hogy pontjain definiálható egy művelet, amivel az elliptikus görbe pontjai csoportot alkotnak. Ezt a csoportot hívjuk *elliptikus csoportnak*. A műveletet szokás szerint összeadásként értelmezzük, vagyis a görbe pontjait *összeadhatjuk*. Nullelemnek, vagyis amihez a görbe bármely más elemét adva saját magát kapjuk vissza, a formulák egyszerűsítése érdekében a görbe inflexiós pontját választjuk.

A 3. ábrán mutatjuk be, hogyan számítjuk ki a görbe *A* és *B* pontjainak az összegét. Legyen *O* a görbe inflexiós pontja, ez az ábrán a koordinátarendszer középpontja. Tetszőleges egyenes a görbét egy vagy három pontban metszi, ez következik abból, hogy egy harmadfokú egyenletnek vagy egy, vagy három valós gyöke van. Ezért ha egy egyenesnek és a görbének van két közös pontja, akkor van egy harmadik is. (Speciálisan ez egybeeshet valamelyik ponttal is, ha éppen érintőről van szó.) Legyen tehát *A* és *B* a görbe két pontja. Kössük össze *A*-t és *B*-t egy egyenessel. Ez az egyenes a görbét még egy pontban metszi, ezt a metszéspontot kössük



3. ábra A görbe *A* és *B* pontjainak összege

össze az *O* ponttal. Ez utóbbi egyenesnek és a görbének harmadik közös pontja lesz az *A+B* pont. Ha az *A+A* összegre vagyunk kíváncsiak, akkor természetesen az *A*-beli érintőt kell használni, az érintő és a görbe metszéspontját kell *O*-val összekötni, és a harmadik metszéspontot választani.

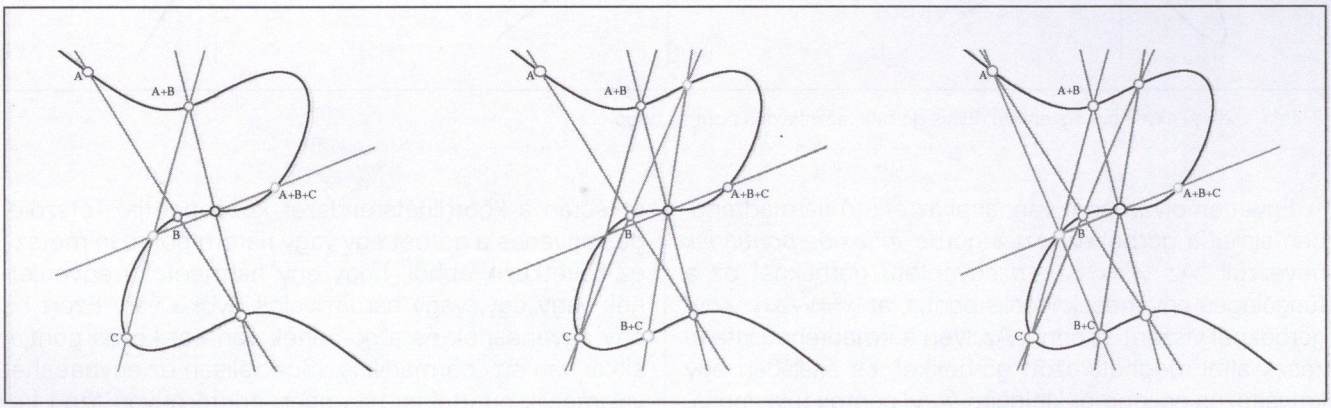
Mivel az *A*-t és *B*-t összekötő egyenes ugyanaz, mint a *B*-t az *A*-val összekötő egyenes, azért ez a művelet *kommutatív*, vagyis *A+B=B+A*. Az is látszik, hogy az inflexiós pontot bármely más ponthoz hozzáadva az nem változik: *A+O=O+A=A*. Ez az összefüggés még az *O* inflexiós pontra is igaz, ugyanis az *O*-beli érintő a görbét harmadszor is *O*-ban metszi (az *O*-beli érintő harmadrendben érint). Az *A* ellentettjét, vagyis a *(–A)*-t úgy kaphatjuk meg, hogy *A*-t összekötjük *O*-val, ez az egyenes metszi ki *(–A)*-t a görbéből; erre a pontra persze *A+(–A)=O*.

A görbénk szerencsés paraméterezése és *O* megválasztása miatt *(–A)* nem más, mint *A*-nak *O*-ra való tükörképe, ezt tehát *A* koordinátáiból könnyen számíthatjuk: mindkét koordinátának kell venni a (–1)-szeresét. Nem ilyen egyszerű a helyzet az *(A+B)* koordinátáinak számításával. Ha az *A* koordinátái (x_1, y_1), a *B* koordinátái pedig (x_2, y_2), akkor *(A+B)*-nek *x* és *y* koordinátáit az alábbi, egyszerűnek semmiképpen nem mondható képletek alapján számíthatjuk (a formulákat a Maple program állította elő):

$$x := \frac{ax_1^2y_2^2 + 3bx_1y_1y_1^2 - ay_1^2x_2^2 - 3by_1^2x_2y_2 + x_1y_1 + 2x_1y_2 - 2y_1x_2 - x_2y_2}{3x_1^2x_2 - 2ax_1y_1y_2 - 3x_1x_2^2 + ax_1y_2^2 - ay_1^2x_2 - 3by_1^2y_2 + 2ay_1x_2y_2 + 3by_1y_2^2 - y_1 + y_2}$$

$$y := \frac{3x_1^2x_2y_2 - 3x_1y_1x_2^2 - ax_1y_1y_2^2 + ay_1^2x_2y_2 + y_1^2 - y_2^2}{3x_1^2x_2 - 2ax_1y_1y_2 - 3x_1x_2^2 + ax_1y_2^2 - ay_1^2x_2 - 3by_1^2y_2 + 2ay_1x_2y_2 + 3by_1y_2^2 - y_1 + y_2}$$

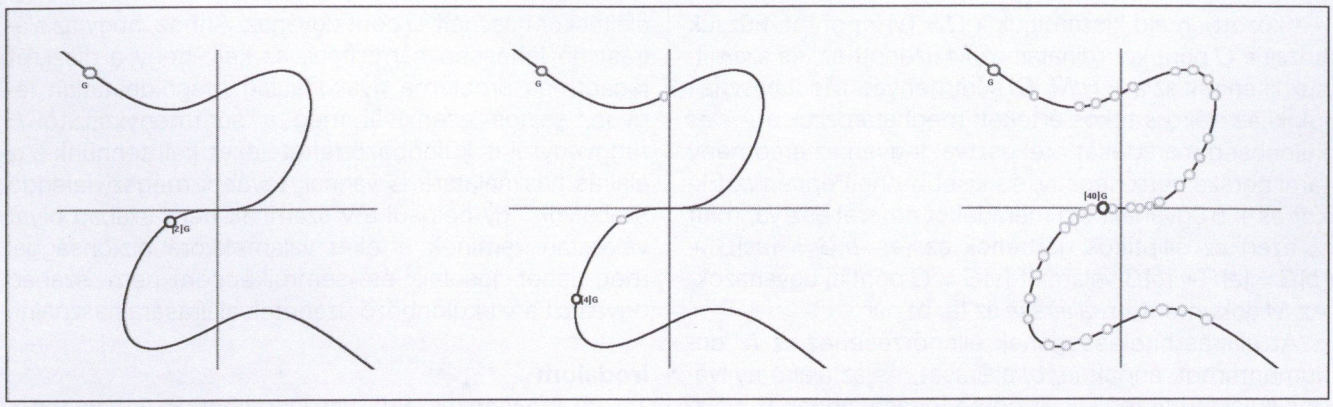
Ráadásul ezek nem használhatók abban az esetben, ha A és B egybeesik, mert ilyenkor mind a számláló mind a nevező nulla. Helyette másik képletet kell alkalmaznunk (azt itt nem mutatjuk be).



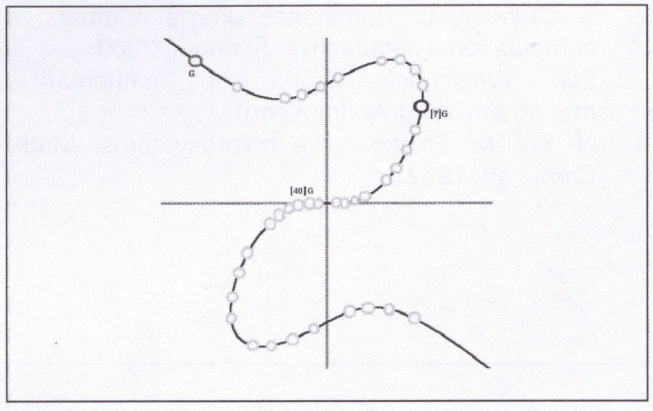
4. ábra Az összeadás asszociatív: *(A + B) + C = A + (B + C)*

Ahhoz, hogy ily módon tényleg csoportot kapjuk, még meg kell mutatni, hogy a művelet *asszociatív*, vagyis tetszőleges *A, B* és *C* pontokra *(A + B) + C = A + (B + C)*. Az állítást az 4. ábra illusztrálja. Először hozzáadjuk *C*-t az *(A+B)* összeghez (bal oldali ábra). Azután kiszámítjuk *(B+C)*-t (középső ábra), és ezt adjuk *A*-hoz. Hogy az eredmény mind a kétszer ugyanaz, következik például abból, hogy az *A+B+C* koordinátáit kétféleképpen kiszámítva ugyanazokat a formulákat kapjuk (nagyszerű feladat egy szimbolikus algebrai program számára). Vagy következik abból a tételből is, amelyet a harmadik ábráról olvashatunk le, és ami a projektív geometriából ismert Papposz-tétel általánosítása. Felejtkezzünk el az *A+B+C* pontról és a rajta átmenő egyenesről. Az ábrán marad hat egyenes, melyek kilenc pontban metszik egymást. Ha e kilenc pont közül nyolc rajta van a görbén, akkor a kilencedik is rajta van. Ez a *kilenc asszociált pont tétele* néven ismert állítás (pontosabban annak is csak egy speciális esete).

Ezzel megkaptuk a titkosításhoz szükséges csoportot. Amit még választanunk kell, egy *G* generátor (a görbe tetszőleges pontja), aminek többszöröseit használjuk fel. Szokás szerint ha *G*-t n-szer adjuk önmagához, az eredményt *[n]G* jelzi. Az ábrán láthatjuk, ahogyan a *G*, *[2]G*, *[3]G* stb. értékek egyre jobban beteríti a görbét. Esetünkben a *diszkrét logaritmus probléma* a következő: adott a görbe paramétereivel, a *G* generáló pont koordinátáival, továbbá a görbének egy *Q* pontja. Keressük meg azt az n egész számot (ami lehet negatív is), amire *Q=[n]G*. Az elliptikus görbéken alapuló



5. ábra A *G* generáló elem és többszörösei



6. ábra

nyilvános kulcsú titkosírás, szokásos angol rövidítéssel *ECC*, éppen ennek a problémának a nehézségét használja ki.

Természetesen a gyakorlati megvalósításban a harmadfokú görbét nem a valós számok fölött használjuk, hanem valamilyen véges testet választunk erre a célra. Mivel egy véges test elemszáma mindig valamilyen prímszámhatvány, gyakorlatilag kétféle test jöhet szóba. Az egyiknél a test elemszáma egy $2^n - 1$ alakú prímszám, a másiknál pedig kettőhatvány. A kitevőt mindkét esetben 200 körül szokás választani, így a test egy elemének felírásához ennyi bitre van szükség, a görbe egy pontját pedig két koordináta határozza meg, tehát ahhoz kétszer ennyi bitet kell tárolni.

További problémát okoz, hogy ha az alaptest elemszáma kettőhatványa, akkor az elliptikus görbe algebrai alakját nem lehet a szokásos $y^2 = x^3 + ax + b$ formára hozni, hanem csak az úgynevezett *Weierstrass* formára: $y^2 + ay = x^3 + bx^2 + cxy + dx + e$. A minél egyszerűbb alakra azért van szükség, hogy egy-egy elliptikus művelet végrehajtása minél gyorsabb legyen. Ahhoz, hogy két pont összegét meghatározzuk, jó néhány szorzást, összeadást, és ráadásul két osztást is el kell végeznünk az alaptest számain. Ráadásul egy kettőhatvány elemű testben két elem összege nem más, mint ezek bitenkénti mod 2 összege (vagyis a két bitsorozat XOR-ja), vagyis gyorsan számítható, addig két elem *szorzatát* már csak bonyolult módon (tipikusan polinomok szorzatát kell maradékosan osztani) lehet számítani. Ezért bár az ECC „csak" 200 bites számokat használ, a kódolás/dekódolás ideje nem kevesebb, mint az ötször több bitet használó RSA vagy Diffie-Hellmann módszereknek. A műveleti igény csökkentésére több trükköt vetettek be, például a számlálót és a nevezőt külön-külön tárolják, és csak a számítás legvégén osztják el egymással. Ha a görbe egy pontjának ismerjük az *x* koordinátáját, akkor ehhez két lehetséges *y* koordináta tartozik, melyek egymás negáltjai. A számításokat el lehet úgy is végezni, hogy csak az *x* koordinátákkal foglalkozunk, és soha nem mondjuk meg, hogy a két lehetőség közül melyik is az igazi.

További probléma, hogy nagyon sok algoritmusban szükség van a generált csoport *rendjére*, vagyis arra a legkisebb *r* pozitív egészre, amivel $[r]G = O$. Ennek értékét a modulo *p* szorzás esetén viszonylag egyszerű megkapni, míg elliptikus görbék esetében ennek meghatározására nem ismeretes polinominális algoritmus. A legfrissebb kutatások viszonylag gyors algoritmusokat eredményeztek, amivel egy-egy görbe rendjét már kevesebb, mint egy óra alatt meg lehet kapni.

Végül álljon itt a diszkrét logaritmus probléma egy konkrét alkalmazása. Az alábbi digitális aláírás *Claus Schnorr* 1991-ből származó módszerének egy változata. Elsőként az aláíró választ egy elliptikus görbét, azon egy *G* pontot. Meghatározza a *G* pont rendjét, vagyis azt a legkisebb pozitív *r* számot, amire $[r]G = O$. Választ még 0 és r–1 között egy s titkos számot, és kiszámítja a görbének a $P = [s]G$ pontját. Az aláíró nyilvános kulcsa a következő információkat tartalmazza: a görbét (vagyis annak a paramétereit, előállítási módját), a *G* generátor pontot, valamint a *P*-t (illetve ennek koordinátáit).

Az *M* dokumentum aláírásához szükség van még egy mindenki által ismert *sűrítménykészítő* algoritmusra is. Angol kifejezéssel az ilyen leképezéseket hash függvényeknek is nevezik. Léteznek minden szempontból kielégítő tulajdonságú hash függvények, egyszerűen feltesszük hogy H(.) egy mindenki megelégedésére kiválasztott sűrítménykészítő eljárás, ami ráadásul mindig a generátorpont rendjénél kisebb értéket állít elő.

Az *M* dokumentum aláírása a következőképpen történik. Először is választunk egy *v* véletlen számot 0 és

r–1 között, majd kiszámítjuk a $Q = [v]G$ pontot. Fűzzük hozzá a Q pont koordinátáit az M üzenethez, és számítsuk ki ennek az $a = H(M, Q)$ sűrítményét. Most használjuk ki az aláíró s titkos értékét: meghatározzuk a $v - as$ különbség maradékát r-rel osztva, legyen az eredmény (ami persze nem negatív, és kisebb r-nél) éppen b. Ekkor $as + b$ ugyanakkora maradékot ad r-rel osztva, mint v, azért az elliptikus görbének az $[as+b]G = [as]G + [b]G = [a]P + [b]G$ valamint $[v]G = Q$ pontjai ugyanazok. Az M dokumentum *aláírása* az (a, b) pár.

Az aláírás hitelességének ellenőrzéséhez az M dokumentumot, annak *(a, b)* aláírását, és az aláíró nyilvános kulcsát használjuk. A görbe leírása, annak G, valamint P pontjai ismeretében kiszámítható a görbe $Q' = [a]P + [b]G$ pontja. Számítsuk ki ezek után az $a' = H(M, Q')$ sűrítményt. Az aláírást hitelesnek fogadjuk el, ha az így kapott a' és az aláírásban szereplő a szám megegyezik.

Könnyű látni, hogy egy helyesen aláírt dokumentumot mindig hitelesnek fogunk elfogadni. Ez azon múlik, hogy ilyenkor az ellenőrzéskor számított Q' és az aláíráskor használt Q pont ugyanaz. Ahhoz, hogy az aláírást ne lehessen hamisítani, az kell, hogy a diszkrét logaritmus probléma gyakorlatilag megoldhatatlan legyen, sajnos ezenkívül még a sűrítménykészítő H függvényről is különböző feltételeket kell tennünk. Az aláírás használatára is vannak további megszívlelendő szabályok. Így például a v számnak nem szabad olyat választani, aminek értékét valamekkora biztonsággal meg lehet jósolni, és semmiképpen nem szabad ugyanazt a v-t különböző üzenetek aláírására használni.

Irodalom

1. A. J. Menezes, P. C. van Oorschot, S. A. Vanstone: Handbook of applied cryptography, CRC Press, 1997
2. R. Crandell, C. Pomerance: Prime numbers, a computational perspective, Springer, 2000
3. Eric Weisstein's world of mathematics, http://mathworld.wolfram.com
4. N. Koblitz, Elliptic curve cryptosystems, Math. Comp, 48, 1987

Hírek

A fordulat éve jöhet el az internetes könyvpiacon, és két-háromszorosára nőhet az interneten könyvet rendelők száma ez év decemberére. A további növekedés legnagyobb gátja az alacsony internetprezentáció. A lakosság kevesebb mint 14 százaléka használja egyáltalán az internetet Magyarországon. A Studentnet felméréséből kiderült, hogy az interneten vásárlók 65 százaléka a naponta a világhálót használók köréből kerül ki. Arra is fény derült, hogy volumenét tekintve a könyvvásárlás áll az első helyen az internetes áruházak körében, megelőzve a számítástechnikai cikkeket is.

❖ ❖ ❖

Az Ericsson és a Mobtel szerbiai hálózatüzemeltető a Mobtel GSM 900/1800 hálózatának bővítésére és felújítására vonatkozó szerződést írt alá.

A telepítést azonnal megkezdik, és a 12 hónapig tartó munkálatok befejeztével a Mobtel már az új MMS multimédiás üzenetkezelési szolgáltatást kínálhatja előfizetőinek. A hálózat kapacitása a jelenlegi egymilliójáról közel a kétszeresére növekszik.

Mobil ad hoc hálózatok biztonsága

GÉMESI ROLAND[1]–IVÁDY BALÁZS[1]–ZÖMBIK LÁSZLÓ[2]

gemesiro@sch.bme.hu–ivadyb@sch.bme.hu–laszlo.zombik@eth.ericsson.se

BME, Távközlési és Telematikai Tanszék[1], Ericsson Magyarország Kft.[2]

Az ad hoc hálózatok nem igényelnek kiépített és fix infrastruktúrát, minden résztvevő egyenrangú. A hálózat működésének biztosítása együttes erővel történik, ám teljesen elosztott és önszervező hálózatról lévén szó, a kommunikáció számos veszélyforrásnak van kitéve. Cikkünk célja, hogy bemutassa az ad hoc hálózatok biztonsági kérdéseit, rávilágítva néhány felmerült megoldási lehetőségre.

1. Bevezetés

A hagyományos hálózatokkal szemben az ad hoc hálózatok nem igényelnek kiépített infrastruktúrát, a feladatokat az egyenrangú résztvevők elosztva, önszervező módon végzik. Egy ilyen rendszerben nincsenek kitüntetett szerepű eszközök, melyek központilag irányíthatnák vagy ellenőrizhetnék a folyamatokat.

A mobil ad hoc hálózatok különféle speciális tulajdonságokkal rendelkeznek. Dinamikus topológiájuk miatt a felépülő kapcsolatok, így az utak is csak korlátozott ideig érvényesek. A résztvevők mozgása vagy eltűnése nem befolyásolhatja a hálózatműködést. Az egységek rendszerint kisméretű, kézi, hordozható készülékek, melyek korlátozott CPU-, memória- és telepkapacitással rendelkeznek. A vezeték nélküli kapcsolatok gyakran kisebb sávszélességűek, és e csatornák gyakran osztottak és limitáltak. Mind a kapcsolatok, mind a készülékek sokkal sebezhetőbbek, mint vezetékes hálózatok esetében.

Egy ad hoc hálózat elosztottságánál fogva számos olyan veszélyforrásnak van kitéve, melyek a korábbiakban még csak fel sem merültek. A továbbiakban biztonságtechnikai oldalról vesszük szemügyre e hálózatokat, biztonságossá tételükhöz eddig felmerült gondolatok ismertetésére törekszünk.

A 2. fejezetben bemutatjuk az ad hoc útválasztó mechanizmusok típusait, alapelveit, majd két, biztonságtechnikailag különböző konkrét protokollt is ismertetünk. A 3. fejezet a biztonság kérdéskörét részletezi, rámutat az elérendő célokra, valamint bemutat jó néhány fenyegetést jelentő tényezőt. Majd a problémák megoldásához rendelkezésünkre álló biztonsági mechanizmusokat ismertetjük. Ezután az 5. fejezetben rámutatunk arra, miként vezet a biztonságosság kérdése az útválasztás témakörébe, és lehetőségeket mutatunk a biztonságos megoldás eléréséhez.

2. Ad hoc hálózatok útvonal-választási mechanizmusai

A kommunikációhoz megfelelő útvonalon továbbított csomagok szükségesek. Hagyományos hálózatokban léteznek olyan kitüntetett pontok (gateway, szerver…), amelyek a hálózat felépítéséről információt hordoznak, így tudják a helyes továbbítási irányt. Ezzel szemben az ad hoc hálózatokban nincsen fix infrastruktúra, nincsenek központosított szerepek, minden egységnek részt kell vállalnia az útvonalválasztásban és csomagtovábbításban.

Útvonalválasztás protokolljai a mobil kommunikáció területén két nagy csoportba sorolhatók: *proaktív és reaktív* protokollok.

Egy *proaktív* protokoll állandóan figyelemmel kíséri a hálózat változásait és a lehetséges útvonalakat. Hátránya, hogy a résztvevőknek sok adatot kell tárolniuk, valamint a hálózat rendszeres felderítése nagy többletterhelést eredményez. Ezzel ellentétben állnak a *reaktív* vagy más néven igény szerinti protokollok, melyek csak akkor keresnek útvonalat, amikor arra szükség van. Ez jobb megoldás, de késedelmet jelenthet a kommunikáció kezdetekor, hiszen az útvonalat ekkor kell kialakítani. Léteznek ezen túl még e két típust ötvöző *hibrid* útvonalválasztó protokollok is. A következőkben két tisztán reaktív protokollt tekintünk át.

A ***Dynamic Source Routing (DSR)*** protokoll forrás útvonal-választási (source routing) algoritmust használ, vagyis mindig a küldő határozza meg a csomag teljes útját, melyet a csomag fejrészébe épít. A közbülső csomópontok számára ebből egyértelműen kiderül a továbbítás iránya.

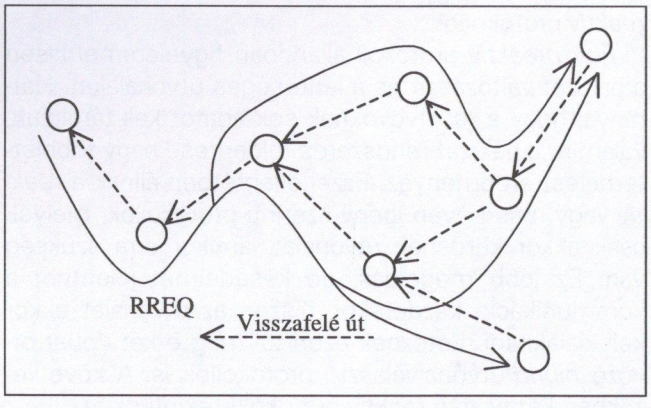
A résztvevők útvonal-gyorsítótárat (route cache) tartanak fenn, melyben bizonyos ideig tárolódnak a használt útvonalak. Megfelelő bejegyzés hiányában egy út-

vonalkérés csomag (route request) indul. Ez az üzenet szétterjed a hálózatban (broadcast message), minden továbbító résztvevő beleírja saját címét. Idővel a kérés eljuthat a célállomáshoz is, aki a fordított úton nyugtázhatja az esélyes útvonalat. Ha nem található útvonal a célállomásig, akkor hibaüzenet keletkezik, később a küldő ismét indíthat útvonalkeresést.

Kis mobilitású hálózatban ez a protokoll meglehetősen jól működik, mivel az útvonal-gyorsítótárban lévő bejegyzések hosszabb ideig is használhatók maradnak. Hátrányuk, hogy a csomagok a közbülső résztvevők címeit is tartalmazzák, amely jelentős többletterhelést (overhead) eredményezhet.

Az ***Ad hoc On-Demand Distance Protokoll (AODV)*** szintén reaktív protokoll, mely a DSR-hez hasonlóan üzenetszórásos útvonalkereső algoritmust használ, de a küldő nem határozza meg a csomag teljes útját. A közbülső résztvevők irányítják a csomagot, minden csomópont csak a következő ugrás irányát dönti el. A hurokmentesség problémáit az üzenetek sorszámozásával küszöbölik ki.

Útvonal-felderítés indítja a folyamatot útvonalkérő (route request – RREQ) üzenettel, ami csak a célállomás és a küldő címét tartalmazza. Minden résztvevő, aki megkapja a RREQ üzenetet, egy visszafelé mutató utat (reverse route) állít be oda, ahonnan az üzenetet kapta, majd szétszórja az üzenetet (1. ábra). Amikor a célhoz megérkezik egy ilyen RREQ üzenet, arra útvonalválasz (route reply – RREP) üzenettel felel. A RREP üzenet a visszafelé vezető útvonalon fut végig, miközben beállítja az előre vezető utakat. Minden útvonal-bejegyzésnek meghatározott élettartama van, adatforgalom hiányában az útvonalak helyi „Hello" üzenetek segítségével tarthatók érvényben.



1. ábra AODV útvonalválasztás

Ha valahol hiba lép fel a kapcsolatban (link error), akkor az elérhetetlenné váló célpontokról útvonalhiba (route error – RERR) üzenet keletkezik. Az AODV képes helyileg korrigálni a kieső kapcsolatokat, amit helyi javításnak (local repair) neveznek. Ha a helyi javítás sikeres, a végpontok nem is észlelik a változást.

Az AODV kis számítás- és memóriaigényű, jól skálázható protokoll, ami nagyobb mobilitású hálózatok esetén is megfelelően működik.

3. Ad hoc hálózatok biztonsági kérdései

Biztonsági célok

A különböző típusú hálózatoknak különféle biztonsági szempontokat kell figyelembe venniük.

Az *elérhetőség vagy használhatóság* (availability) azt jelenti, hogy a hálózat és annak szolgáltatásai mindenféle támadás és hibás működés ellenére mindig hozzáférhetőek és használhatók legyenek.

A *hitelesítés* (authentication) garantálja a résztvevőknek egy adat vagy egy másik résztvevő hitelességét.

A *bizalmasság* (confidentiality) az adat titkosságát biztosítja, ami azt jelenti, hogy az információ nem kerülhet illetéktelen egységek birtokába.

Az *integritás* (integrity) biztosítja az üzenetben történt módosítás felismerését. A módosításokat természetes környezeti hatások vagy szándékos támadások is okozhatják.

A *letagadhatatlanság* (non-repudiation) a részvételt rögzíti, ami egy későbbi bizonyítási folyamathoz lehet szükséges.

Számos kriptográfiai eljárás igényel *kulcs menedzsmentszolgáltatást* a kulcsok adminisztrálásához.

Ezeken kívül más szolgáltatásokra is szükség lehet, mint például a *hozzáférés-védelem* (authorization), ami a rendszer erőforrásokhoz való hozzáférését korlátozza.

Mobil ad hoc hálózatok veszélyforrásai a vezetékes hálózatokhoz képest újabb kérdéseket vetnek fel. A fő gond, hogy egy elosztott hálózatban a kommunikáció résztvevőinek meg kell bízniuk egymásban, ezért a megfelelő biztonság csak nehezen érhető el.

Szolgáltatás elleni támadások (Denial of Service Attacks – DoS) a hálózat működésének meggátlását célozzák. Ez a fajta támadás csökkenti a szolgáltatás elérhetőségét, a hálózat hozzáférhetőségét, nagy késedelmet generálhat. Romboló magatartásról lévén szó, együttműködést igénylő környezetben kiemelt veszélyt hordoznak és csak nehezen küszöbölhetők ki.

A közös kommunikációs csatornát megcélozva, az ellenség zajjal terhelheti a fizikai közeget, mellyel interferenciát vagy nagy információvesztést okozhat. Nehéz feladat a támadó zavarását a csatorna természetes zajától megkülönböztetni. A közeghozzáférés szabályainak be nem tartása szintén gátolhatja a normális kommunikációt.

A mobil eszközök kis CPU-, memória- és telepkapacitása miatt a nagy CPU teljesítményt igénylő folyamatok leterhelhetik a processzort, így gátolva más folyamatok haladását. A telepterhelést is optimalizálni kell, a használaton kívüli időben a készülék egyes részei alvó üzemmódba válthatnak át. Viszont a bejövő kérésekre válaszolni kell, így az energiakimerítési támadás egy másik veszélyforrás lehet.

Az útvonal-választási folyamat is megzavarható- ha a támadó egységek rossz útvonalakat terjeszthetnek szét. Rossz vagy régi információk bekerülve az útkere-

sési folyamatba hosszabb utakat, hurkokat eredményezhetnek, ami teljesítménycsökkenéshez vezethet. Végül az önzőség is gátolhatja a működést, amikor a node kihasználja az együttműködés lehetőségeit, viszont ő maga nem segíti mások kommunikációját, tehát nem továbbít csomagokat, kíméli saját erőforrásait.

Megszemélyesítés (impersonation) az, ha egy támadó másnak adja ki magát, mint aki valójában. Hitelesítés nélküli rendszerekben ez egy triviális támadás lehet, ám tanúsítványok használatával kiküszöbölhető. E tanúsítványok összerendelik a résztvevőket azonosítóikkal. A központi egység hiánya miatt ez egy nehéz feladat, mivel nincs egyetlen elem, mely tárolhatná ezen összerendeléseket. A megszemélyesítés másik fajtája a „man-in-the-middle" támadás, amikor egy ellenséges node beépül az útvonalba, majd szimulálja a másik oldalt.

Bizalmasság megsértésének (Confidentiality Violation) első esete, ha egy passzív lehallgató csak figyeli a kommunikációs csatornát, de egy aktív lehallgató módosíthatja is az átvitt adatokat. Ha a kommunikáció titkosított, kódtöréses támadás jöhet szóba, ezért elkerülendő a gyenge algoritmusok használata.

Üzenet megváltoztatása is lehet a támadó célja. Próbálkozhat új üzenetek beszúrásával, az információk megváltoztatásával vagy fontos adatok eltávolításával is.

Üzenetletagadás veszélye, hogy a kommunikációban való részvétel utólagos letagadhatósága teljesen lehetetlenné teszi egy későbbi bizonyítás és szankcionálás folyamatát.

Anonimitás megsértése is lehetséges, mert a hálózat résztvevői információkat gyűjthetnek szomszédaikról, a médium ugyanis osztott, és az identitás egy részének (cím) mindenképpen elérhetőnek kell lennie. Ezáltal egy célpont azonosíthatóvá vagy lokalizálhatóvá válhat. Ha az egységek azonosító tanúsítványokkal rendelkeznek, képesek egymásról információkat gyűjteni. Az azonosítókat lehallgatók is megszerezhetik.

Fizikai támadás esetén a támadó megszerezhet (pl. lopás) egy készüléket, melyben akár hardverben, akár szoftverben módosításokat végezhet. Gyenge védelem esetén így megtudhat titkos információkat (pl. kulcsok), de akár vírus vagy trójai faló is telepíthető a készülékbe.

Központi egység hiánya miatt a rendszer nem sebezhető a központ ellen induló támadásokkal. Így viszont számos probléma válik nehezebben megoldhatóvá.

Útvonalválasztás manipulálásával, vagyis rossz információk terjesztésével a támadó jelentős zavarokat okozhat. Megnövekedhet a csomagküldési idő, fontos adatok veszhetnek el, de akár a támadó is fontos információkat szerezhet meg.

A kulcsok kompromittálódása is komoly veszélyforrás, mert a titkosítási eljárások gyakran titkos információkat igényelnek.

4. Biztonsági mechanizmusok

A hagyományos biztonsági megoldások (pl. autentikáció, digitális aláírás, kódolás) megfelelő biztonságot nyújthatnak, ám legtöbbjük valamiféle menedzsmentszolgáltatást igényel. Ezen szolgáltatások ellátásához nagyrészt központi egység szükséges, ám ad hoc környezetben ilyenekre nem támaszkodhatunk. A feladatokat is elosztva kell végezni, ráadásul a rendszernek redundanciát is hordoznia kell, mivel résztvevők eltűnhetnek, szakadozhat a kapcsolat és a kompromittálódás veszélye is fennáll.

Nyilvános kulcsú infrastruktúra (Public Key Infrastructure – PKI) esetén minden résztvevő rendelkezik egy nyilvános és egy titkos kulccsal. A nyilvános kulccsal kódolt üzenet csak a neki megfelelő titkos kulccsal dekódolható, és a két kulcs egymásból nem származtatható.

Létezhet ezen felül egy központi Certificate Authority (CA), amely tanúsítványokat (certificate) állít ki nyilvános kulcsok és a megfelelő csomópontok összetartozásáról. A CA-nak on-line elérhetőnek kell maradnia, hogy ezen összerendeléseket biztosítani tudja, követnie kell a kulcsváltozásokat, valamint képesnek kell lennie tanúsítványok visszavonására.

A *küszöb kriptográfia (Thresold cryptography)* a bizalom szétosztásának egy lehetséges módja. Egy (n, t+1) séma lehetővé teszi egy kriptográfiai feladat n résztvevő közötti szétosztásának lehetőségét úgy, hogy azt bármely t+1 tagja sikeresen el tudja végezni, de ennél kevesebb tag már nem birkózhat meg a feladattal.

Ez esetben a kulcs menedzsmentszolgáltatás n szervere osztja szét egy bizonyítvány aláírásának jogát, vagyis a szolgáltatás titkos kulcsát. Ezt a szétosztást a k kulcs (n, t+1) szétosztásának nevezzük. Így minden egyes szerver az aláírásnak csak egy részét képes előállítani, melyek még akkor is sikeresen összekombinálhatók, ha t szerver kompromittálódott.

A küszöb kriptográfia egy nagyon hasznos tulajdonsága, hogy képes a résztvevők információdarabjainak frissítésére. Ez a mozgó támadók kivédését teszi lehetővé, melyek egymás után törik fel a szervereket, melyek mennyisége egy bizonyos idő elteltével meghaladhatná a kritikus t mennyiséget. Szétosztásfrissítéskor (share refresh) egy új küszöb kriptográfiai séma jön létre, mely kiválóan alkalmas a hálózat változásaihoz való alkalmazkodásra. Mivel e frissítés nem túl bonyolult művelet, egy kialakított konfiguráció hosszabb ideig is képes a hálózat változásait követni. Viszont ha túl sok résztvevő kompromittálódott, új sémát kell kialakítani.

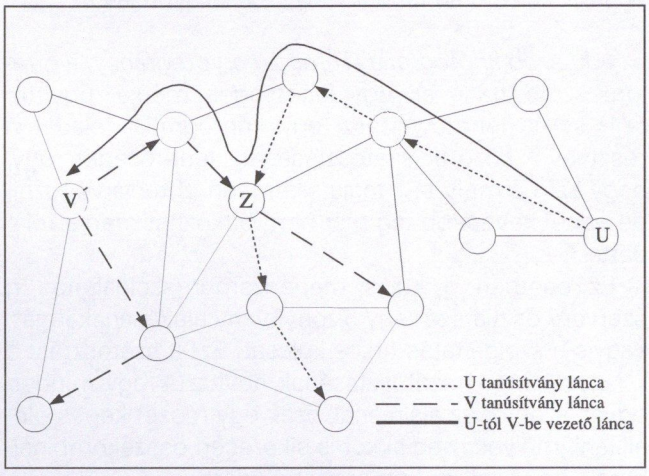
A megoldás súlyos problémája, hogy feltételezi a résztvevők szinkronitását, amely csak a legritkább esetben áll fenn. Egy csomópont megszüntetheti a kapcsolatot vagy lelassulhat egy DoS támadástól, és eközben előfordulhat, hogy a többiek véghezvisznek egy szétosztásfrissítést. Ezek után már nem lesz képes visszakapcsolódni a folyamatba, mert azóta egy újabb konfiguráció alakult ki.

A rendszer kialakítása is probléma, amikor még csak néhány résztvevővel rendelkezik a hálózat. Ki kezdemé-

nyezheti ekkor egy szétosztott CA létrehozását? Ráadásul később, ha egy kialakult rendszer diszjunkt részekre szakad, ezek külön folytathatják működésüket.

Önszervező PKI (Self Organizing Public Key Infrastructure) nem igényel megbízott harmadik felet, mely a tanúsítványalapú hitelesítés hátránya ad hoc környezetben. Léteznek bizonyítványalapú autentikációs megoldások (pl. PGP – Pretty Good Privacy), melyekben a bizonyítványokat a felhasználók saját maguk készítik el és terjesztésüket is megoldják.

Az önszervező PKI-ben minden felhasználó rendelkezik egy kisméretű helyi tanúsítványtárolóval, melyben korlátozott számú tanúsítványt képes tárolni. Amikor az *u* és *v* felhasználók kommunikálni szeretnének, egyesítik e tárolójukat, melyben *u* próbál keresni egy megfelelő tanúsítványláncot *v*-hez (2. ábra). Ilyen jó eséllyel biztosítható még akkor is, ha a tároló mérete kicsi a hálózat résztvevőinek számához képest.



2. ábra Önszervező PKI tanúsítványlánca

Ezen megközelítés jó megoldás lehet a teljes önszerveződés eléréséhez. Problémát a skálázhatóság jelent, mivel a hálózat növekedésével egyre növekszik a valószínűsége annak, hogy nem létezik megfelelő tanúsítványlánc.

Az *identitásalapú (ID-Based)* titkosítás alapötlete, hogy a nyilvános kulcs szerepét az egységek azonosítására szolgáló szintén nyilvános azonosítók töltik be, és ezzel a CA által biztosított összerendelő feladatkör szükségtelenné válik. Azonban a nyilvános kulcsú titkosítás biztonságossága éppen abból adódik, hogy a nyilvánosból a titkos kulcs (illetve fordítva) nem származtatható. Ez esetben egyetlen központi entitás képes arra, hogy egy publikus kulcsból (identitás) elkészítse annak titkos párját. Ezt a műveletet csak egyszer, a felhasználók regisztrációjakor kell elvégezni. A rendszerben részt vevő, egymást nem ismerő elemek ezután autentikusan képesek egymással kommunikálni, mégpedig harmadik féllel való interakció nélkül.

A valóságban viszont gondolni kell a kompromittálódott és később újra regisztrálni akaró résztvevőkre is, tehát a regisztráló központnak mégis elérhetőnek kell maradnia. További veszélyforrás, hogy e központ ismerni fogja a résztvevők titkos kulcsait is. A titkos kulcs elkészítése erőforrás-igényes feladat, hosszú időt vehet igénybe.

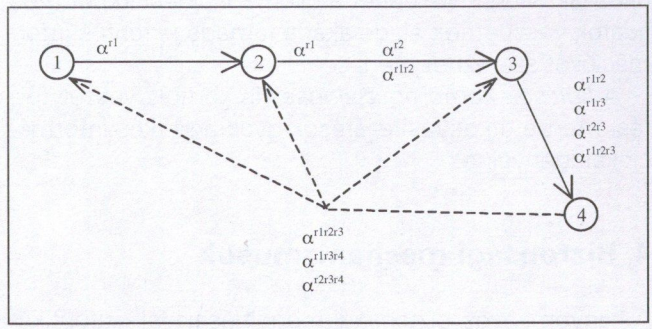
A *titkos (közös, szimmetrikus) kulcsú architektúra (Common key architecture)* az előzetesen megtörtént azonosítás után a titkosság megvalósításához egy kódolt kapcsolat felépítésére van szükség. Ilyen célra alkalmazzák a közös kulcsú titkosítási megoldásokat. Ehhez szükséges a résztvevőknek egy közös (osztott, szimmetrikus) titok birtoklása, melyet egyaránt használnak a kódolás és a dekódolás folyamatában. Titkos kulcsú architektúra alkalmazásával lehetővé válik a titkos multicast is. Korábban kialakultak már kulcs menedzsmentprotokollok, de az ad hoc hálózatok számos más igényeket is támasztanak.

Diffie-Hellman (DH) kulcs cserealgoritmus segítségével nyilvános csatornán lehetőség nyílik két résztvevő közötti közös titok kialakítására. Egy harmadik résztvevő, aki teljes egészében hallja a kommunikációt, nem lesz képes ugyanazon közös kulcs generálására.

A két résztvevő (*A* és *B*) megállapodnak egy q rendű G ciklikus véges csoporton és a csoport generátorelemén. Ezután mindkét résztvevő véletlenszerűen választ egy titkos kitevőt (A és B). Ezután A kiszámítja α^A-t, *B* α^B-t és átküldik egymásnak az így megkapott értékeket. Mivel a kitevők titokban maradtak, az átküldött értéket erre a hatványra emelve kialakul a α^{AB} közös titok, de egy lehallgató nem tudja meg magukat a kitevőket, így képtelen a kulcs létrehozására.

A kettőnél több résztvevős esetben is lehetséges a kulcs kialakítása. Ezek közül tekintünk most át néhányat.

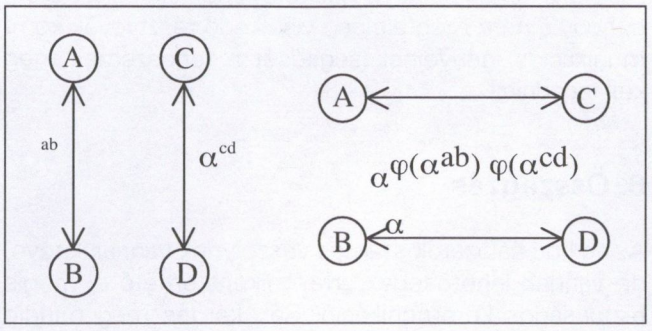
A *GDH.2 (Generalized Diffie Hellman)* a Diffie-Hellman kulcs csere algoritmus általánosítása több résztvevő esetére. A résztvevők egy láncot alkotnak, majd az első tag elküld egy Diffie-Hellman-szerű üzenetet a következő tagnak, majd az kiegészítés után továbbküldi. Ez a folyamat az utolsó elemig folytatódik, ami már képes lesz a közös titok létrehozására, és minden egyes résztvevőnek eljuttatja a számára hiányzó információtöredéket (3. ábra). A titok csak a megfelelő végpontok számára lesz ismert.



3. ábra A GDH.2 üzenetváltásai

Ezen eljárás fő problémája, hogy az utolsó entitás kitüntetett szereppel rendelkezik. Ezen túl viszonylag nagy mennyiségű adatot kell átvinni a csatornán.

A *Hypercube* algoritmus alapötlete párok kialakítása, melyek létrehozzák a közös titkaikat. Ezután a párokat is párokba rendezzük, melyek között ismét elvégezhető a kulcscsere kettesével, és így tovább (4. ábra). A probléma az, hogy a résztvevők száma csak 2^n lehet.



4. ábra A Hypercube üzenetváltásai

Az Octopus protokoll segít a fenti problémán, mely létrehoz egy Hypercube magot, amit karokkal egészít ki. Először e karok végeznek DH kulcscserét a nekik megfelelő központi elemekkel, majd e központ alakít ki Hypercube struktúrát, végül pedig közlik a karokkal a kialakult új kulcsot. A gond ezzel az, hogy a Hypercube mag központi szerepet játszik, valamint új résztvevők bevonása bonyolult.

Csoport kulcs menedzsmentprotokoll (Group Key Management Protocol-GKMP) a résztvevők egy csoportjának szimmetrikus kulcs menedzsmentfunkcióit látja el. A GKMP kulcs kialakítási mechanizmusa bármely két résztvevő közötti kooperatív eljárás, mint például a Diffie-Hellman kulcscsere. A kulcs létrehozása után a GKMP elterjeszti e csoportkulcsot az arra jogosult résztvevőknek. Lehetővé teszi ezen túl új résztvevő beléptetését, tag törlését és a csoport teljes újrakulcsolását. A GKMP teljes jogosultság-ellenőrző rendszert is tartalmaz, mivel a kulcsolás alkalmával jogosultsági tanúsítványok (Permission Certificate – PC) is létrejönnek. Bármely csomópont bárki jogosultságát ellenőrizheti, de módosítani nem képes azt. GKMP-vel lehetséges a kompromittálódott résztvevők kizárása, mivel a feltört résztvevők listáját (CRL – Compromise Recovery List) szétterjeszti a hálózatban.

Ez a protokoll megpróbálja a lehető legtöbb feladatot kiosztani a csoportnak, tehát igyekszik elkerülni bármiféle központi beavatkozást. Ennek ellenére néhány funkció, mint például a jogosultságok osztása, továbbra is központosított feladatkör maradt.

Az előzőkben láttuk, hogy a kommunikáció titkosságának biztosítása nem jelent komolyabb problémát, mivel közös titok, vagyis titkos csatorna kialakítására és kulcsainak létrehozására léteznek megoldások (DH, GDH.2, Hypercube, Octopus). A közös kulcs a végpontok között kialakítható, így a köztes csomópontok és más lehallgatók számára a kommunikáció rejtve marad.

Nagyobb kihívás az egyes résztvevők azonosítása, vagyis az autentikáció megvalósítása. Ad hoc hálózatokban nem támaszkodhatunk egy központi tanúsítványokat biztosító entitásra (Certificate Authority), bár láttuk, hogy megoldási lehetőségek elméletben már léteznek (Küszöb kriptográfia, Önszervező PKI, ID-Based PKI).

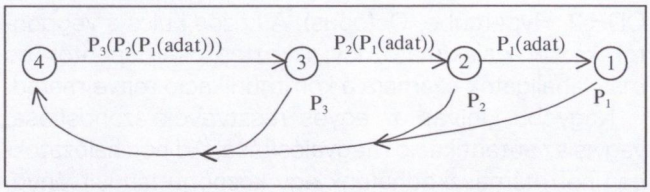
5. Biztonságos útvonalválasztás

Az ad hoc hálózatok biztonságosságának vizsgálata elvezet az útvonalválasztás kérdéskörébe, bár a jelenlegi útvonal-választási megoldások biztonsági szempontokat még nem vesznek figyelembe. A támadó megpróbálhat beépülni egy útvonalba, és őt ebben a jelenlegi megoldások egyáltalán nem korlátozzák. A támadónak elegendő lehet akár megközelíteni az útvonalat és egy onnan indított DoS támadással szabotálhatja a kommunikációt. Egy DoS támadás sokszor teljesen megkülönböztethetetlen a csatorna természetes minőségcsökkenésétől. Úgy is tekinthetjük a kérdést, hogy egy megfelelő minőségű csatornát szeretnénk felépíteni, ahol az útvonal megbízhatósága a rendszernek egy QoS paramétere.

A lényeg tehát az, hogy olyan útvonal-választási folyamatra van szükség, amelynél a kialakuló úton átvitt adatok megadott biztonsági szintje biztosítható. Emellett fontos még, hogy az útvonal igény szerint meghatározható legyen. Ezt elegendő információ birtokában előírhatja egy kijelölt node (pl. a küldő), de elosztott algoritmus is elképzelhető.

Az *Onion routing* egy olyan biztonságos útvonal-választási megoldás, mellyel egy üzenet csakis a megadott útvonalon utazhat. Ezt nyilvános kulcsú titkosítással oldják meg, melynek során a küldő összegyűjti a közbülső csomópontok nyilvános kulcsait és mindegyikkel egymás után kódolja az üzenetet. Az egyes publikus kulcsok hatását csakis a neki megfelelő titkos kulcs, vagyis a megfelelő közbülső résztvevő képes eltávolítani (5. ábra). A címzetthez eljutó adat csak abban az esetben feldolgozható, amennyiben végigutazott a meghatározott útvonalon. Ez akkor alkalmazható, ha már a küldő pontosan tudja, hogy csomagja milyen útvonalon fog végighaladni, amely a forrás útvonalválasztások sajátossága (pl. DSR).

A *Security-aware routing (SAR)* nemcsak az optimális útvonal megtalálását tűzi ki célul (legrövidebb vagy leggyorsabb), hanem meghatározott biztonsági igények teljesülését is szem előtt tartja. Egy útvonal biztonságosságának leírására többféle attribútum alkalmazható, mint például a megbízhatóság szintje (trust level) vagy biztonságosság szintje (security level). Az útvonalválasztás ezeket figyelembe veszi, és csak a megfelelő biztonsági jellemzőjű résztvevők szerepel-



5. ábra Onion routing

hetnek az útban. Természetesen e szinteknek manipulálhatatlannak kell lenniük ahhoz, hogy sem a csomópontok szintjei, sem az igényelt szint ne kompromittálódhasson.

SAR alkalmazásakor a küldő, aki útvonal-felderítést kér, az igényelt biztonsági szintet hozzáépíti a kéréshez. A közbülső résztvevők csak akkor továbbítják a kérést, amennyiben megfelelnek a biztonsági előírásoknak. Ha eljut egy kérés a címzetthez, akkor kialakul egy megfelelő biztonságú útvonal. A SAR kiegészítés szinte bármely igény szerinti (on-demand) útvonalválasztó protokollhoz kapcsolható.

A SAR megoldások nehézsége, hogy az egyes szinteket autentikálni kell, semmiképpen sem lehet a résztvevők felelőssége saját szintjükről nyilatkozni. Láttuk korábban, hogy ad hoc hálózatokban az autentikáció kérdésének megoldása nem triviális feladat, úgymint az sem, hogy megakadályozzuk a nem megfelelő biztonságú (és esetleg támadó) csomópontok beépülését az útvonalba. Mindezek tetejére már az is veszélyt jelenthet, ha egy node biztonsági paraméterei kiolvashatók, mivel ezek általában szoros összefüggésben vannak annak fontosságával.

A *Watchdog* a résztvevők jóságának folyamatos nyomon követését teszi lehetővé. A rádiós csatorna osztottsága révén minden résztvevő figyelemmel kísérheti közvetlen szomszédainak viselkedését, így könnyen észrevehet egy megbízhatatlan elemet. Ez a passzív figyelés a hálózatra nézve nem okoz további terhelést, de sajnos tévedhet (pl. aszimmetrikus link, ütközés), illetve megtéveszthető (pl. irányított antenna).

A *Pathrater* eljárás úgy valósítja meg a veszélyes elemek kirekesztését a felépülő kommunikációból, hogy az egyes utakhoz működésük során felállított statisztikák alapján jósági értékeket rendel. Az egyes utak jósági mutatóját megfelelő működés esetén folyamatosan növeli, míg hiba esetén csökkenti. Ilyen adatok ismeretében az útvonalválasztás során lehetőség van egy megfelelő út kiválasztásra.

Az eddig áttekintett biztonságos útvonal-választási mechanizmusok mind arra törekedtek, hogy a kialakuló utakban ne szerepelhessenek megbízhatatlan résztvevők. Egy támadó fejével gondolkozva rájöhetünk, hogy semmiféle hátrányba nem kerülünk, legfeljebb nem kell majd a későbbiekben nagy fontosságú csomagokat továbbítani.

Az ilyenfajta önző viselkedésminta csábító lehet például a telep kímélése céljából, de tömeges méretekben a hálózat működésének megszűnését is eredményezheti. Kiküszöbölése egyfajta virtuális fizetőeszköz, a *nuglet* alkalmazásával történhet. Ezzel a fizetőeszközzel vásárolják meg a résztvevők egymás szolgáltatását a hálózatban, így az önző egyedek nuglet-jei pedig idővel elfogynak.

A kommunikáció során tehát a nem megbízható résztvevők olyan szankcionálására van szükség, mellyel csakis a jó viselkedés kifizetődő, vagyis amely alkalmazásával a támadók idővel ellehetetlenülnek. Ez azt jelenti, hogy a nem megfelelően viselkedő résztvevők kommunikációs igényeinek segítését a rendszernek meg kell tagadnia.

6. Összegzés

Az ad hoc hálózatok számos veszélynek vannak kitéve, de vannak lehetőségek arra, miként érhető el mégis biztonságos kommunikáció. Sok kérdés még mindig nyitott, de e területen napjainkban is intenzív kutatás folyik.

Irodalom

Földesi András, Homolya György, Horváth Cz. János, dr. Imre Sándor: Bevezetés a mobil ad hoc útvonalválasztó protokollok világába (Híradástechnika, 2001. május)

Kata Molnár, László Zömbik: Security Issues in Mobile Ad hoc Networks. (Evolution in the military communications systems – trends and challenges in the XXI. century, 2001.)

Lidong Zhou, Zygmunt J. Haas: Securing Ad Hoc Networks (IEEE Network Magazine, vol. 13, no.6, November/December 1999)

Jean-Pierre Hubaux, Levente Buttyán, Srdan Capkun. The Quest for Security in Mobile Ad Hoc Networks (ACM Symposium on Mobile Ad Hoc Networking and Computing, MobiHOC 2001)

Ueli M. Mauer, Yacov Yacobi: A Non-interactive Public-Key Distribution System (Designs, Codes and Cryptography, 1996)

Maarit Hietalahti: Key Establishment in Ad hoc Networks (Proceedings of the Helsinki University of Technology, Seminar on Network Security fall 2000)

Seung Yi, Prasad Naldurg, Robin Kravets: Security-Aware Ad hoc Routing for Wireless Networks (Technical Report UIUCDCS-R-2001-2241(ps/pdf), August 2001)

Sergio Marti, T.J. Giuli, Kevin Lai, Mary Baker: Mitigating Routing Misbehavior in Mobile Ad Hoc Networks (In Proceedings of MOBICOM, 2000)

Sonja Buchegger, Jean-Yves Le Boudec: Nodes Bearing Grudges: Towards Routing Security, Fairness, and Robustness in Mobile Ad Hoc Networks (In Proceedings of the Tenth Euromicro Workshop on Parallel, Distributed and Network-based Processing, January 2002)

Az elektronikus aláírás alkalmazásának háttere

ERDŐSI PÉTER

CISA, NJSZT információrendszer-ellenőrzési szakértő
Insurance Technology Kft.
Peter.Erdosi@itech.hu

A 2001/XXXV. törvény hatályba lépése után az elektronikus aláírás Magyarországon is alkalmassá vált joghatás kiváltására. A szabályozási háttér kidolgozásától a gyakorlati megvalósításig vezető út azonban számos technikailag megoldandó feladatot állít a felhasználók elé. Az elektronikus aláírás fogalmának definiálása után a szükséges kriptográfia kiválasztása következik. A folyamat használhatósága függ a műveletek gépi végrehajtási idejétől, az algoritmusok működésétől, biztonságától. Fontos kérdés az is, hogy a használhatóság növelése érdekében alkalmazott szabványosítási törekvések mennyire befolyásolják a felhasznált kriptográfia által nyújtott biztonsági szintjét, és milyen ára van mindennek.

Az elektronikus aláírás definíciója

Az Európai Közösség 1993/93/EC direktívája alapozta meg az elektronikus aláírás jogi elfogadhatóságát az Európai Unió tagállamaiban. Ennek alapján elektronikus aláírásnak kell tekinteni azt az elektronikus formában létező adatot, amely hozzá van kapcsolva vagy logikailag társítva van egy másik elektronikus adathoz, és hitelesítési módszerként használható. Ebbe a körbe belefér az e-mail végére csatolt, névvel ellátott automatikus aláírásfájl, vagy a kézi aláírás digitalizált képe is. Fejlettebb az olyan elektronikus aláírás („advanced electronic signature"), mely egyértelműen köthető az aláíróhoz, képes az aláíró azonosítására, az aláíró a teljes ellenőrzése alatt tarthatja a készítési folyamatot, valamint oly módon kapcsolódik a vonatkozó adathoz, hogy az adaton minden későbbi módosítás észlelhető. A direktíva elméletileg platformfüggetlen módon fogalmaz, a terminológiájában azonban gyakorlatilag a Public Key Infrastructure, a PKI-rendszer elemei jelennek meg.

A magyar törvényalkotás az EU-Direktíva elveit követve alkotta meg az elektronikus aláírásról szóló törvényt. Értelmezésében elektronikus aláírás az elektronikus dokumentumhoz azonosítás céljából hozzárendelt és azzal elválaszthatatlanul összekapcsolt elektronikus adat, illetőleg dokumentum. Fokozott biztonságú lehet az olyan elektronikus aláírás, mely alkalmas az aláíró azonosítására és egyértelműen hozzá köthető. Olyan eszközzel hozták létre, mely kizárólag az aláíró befolyása alatt áll, és a dokumentum tartalmához oly módon kapcsolódik, hogy minden – az aláírás elhelyezését követően az iraton, illetve a dokumentumon tett – módosítás érzékelhető. Ezenkívül minősített aláírás a definíció szerint az olyan fokozott biztonságú aláírás, mely biztonságos aláírás-létrehozó eszközzel készült, és amelynek hitelesítése céljából minősített tanúsítványt bocsátottak ki.

Az ETSI ES 201 733 elektronikus aláírás-formátumokat tartalmazó szabvány (szabvány) az elektronikus kereskedelem sikerére és további fejlődésére koncentrál. Ebből a megközelítésből nézve az elektronikus aláírás „fontos biztonsági komponens, amely felhasználható az információ és az elektronikus üzletbe vetett bizalom védelmére". Ennek következtében a szabvány szerint készített elektronikus aláírások valamely kötelezettség explicit felvállalásának feldolgozható bizonyítékaként jelennek meg az elektronikus tranzakciókban részt vevő partnerek között. Más szóval a Szabvány az aláírás érvényességére koncentrál, azaz tartalmazza mindazon követelményeket, melyeket az aláírás készítése közben kell kielégíteni és azokat is, melyek ahhoz szükségesek, hogy a fogadó fél meggyőződhessen az aláírás érvényességéről. A szabvány kijelenti, hogy a szabványban használt „elektronikus aláírás" egyenértékű a direktíva „advanced electronic signature" fogalmával. Itt már nevesítve van a nyilvános kulcsú kriptográfia mint matematikai háttér, habár az aláírás-formátumok túlmutatnak a digitális aláíráson.

Kriptográfiai szempontból az elektronikus aláírás fogalmába definíció szerint beletartozik az üzenethitelesítő kód (Message Authentication Codes – MAC) és a digitális aláírás (digital signature – DS). A MAC olyan fix hosszúságú adat, melyet az üzenetből és egy titokból – jellemzően szimmetrikus kriptográfiai kulcs – képeznek. A fogadó fél az általa birtokolt titokból és a megkapott üzenetből újra generálja a MAC-kódot, aminek meg kell egyeznie a kapott kóddal, hacsak az üzenet nem változott a továbbítás során. Ez a módszer azonban nem alkalmas a letagadhatatlanság biztosítására, hiszen ugyanazon üzenet és titok birtokában többen is generálhatják ugyanazt a MAC-kódot. A digitális aláírás módszere alkalmas az eredet és a sértetlenség biztosítására is. A küldő az üzenet és az aláíró kulcs birtokában elkészíti az aláírást és elküldi az üzenettel

együtt a fogadónak. A fogadó az ellenőrző kulcs használatával tud meggyőződni az üzenet sértetlenségéről, és biztos lehet abban is, hogy az üzenetet az írta alá, akinek az aláíró kulcs a birtokában van.

A digitális aláírás elkészítéséhez több algoritmus is választható. A kriptológusok – egy kis túlzással – azt mondják, hogy a gyakorlatban használt algoritmusok 99%-a RSA, 1%-a DES, a maradék az összes többi. Ezért a vizsgálódás tárgya a továbbiakban csak az RSA algoritmus, és a segítségével megvalósított rendszerek lesznek.

Az RSA algoritmus röviden

Az RSA algoritmus a modulo m aritmetikát, a maradékos osztás elméletét használja fel működéséhez. Erősségét az egész számok faktorizálásának problémája adja.

Legyen p_1 és p_2 két tetszőlegesen kiválasztott prímszám. Legyen $m = p_1*p_2$ modulus, ekkor $\phi(m) = (p_1-1)*(p_2-1)$, az m szám Euler-függvénye, azaz 1,2,...,m között az m-hez relatív prímek száma. Válasszunk egy tetszőleges e természetes számot, melyre $1 <= e < \phi(m)$, valamint e és $\phi(m)$ relatív prímek. Ekkor meghatározható egyértelműen az a d szám, mely e-nek multiplikatív inverze, vagyis $d*e \equiv 1 \pmod{\phi(m)}$. Ekkor az egyik (pl. nyilvános) kulcs az (m, e) számpár, a másik (pl. titkos) kulcs pedig az (m, d) pár lesz.

Legyen a rejtjelezni kívánt üzenet M, ahol $1 \leq M < m$. A kódolás ekkor az $M_{enc} \equiv M^e \pmod{m}$, a dekódolás az $M \equiv M_{enc}^{\ d} \pmod{m}$ műveletek elvégzésével történik.

Az aláírás készítésének műveletigénye

A digitális aláírás elkészítése előtt rendelkezni kell legalább egy kulcspárral, és egy hash függvénnyel is. A kulcsok megfelelően nagynak választott prímszámok segítségével állnak elő, tehát az aláírási algoritmusnak valahol prímszám-generátort is tartalmaznia kell. A prímszámokat, illetve a segítségükkel előállított modulust úgy kell megválasztani, hogy az ismert faktorizáló algoritmusoknak ellenálljanak. Az ilyen prímek legalább 512 bináris jegyűek, a különbségük is legalább 511 bináris jegyű és emellett a peremfeltételek mellett véletlenszerűek. A prímfelbontás jelenlegi csúcsteljesítménye az RSA-155 (azaz 155 decimális számjegy szorzattá alakítása) 1999-ben, amikor egy 512-bites számot felbontottak két 78 decimális számjegyű prímszám szorzatára. A felbontás idejét 8400 MIPS évben jelölték meg, ami azt jelenti, hogy annyi műveletet kellett hozzá végrehajtani, amit egy másodpercenként 1.000.000 gépi műveletet végrehajtó számítógép 8400 év alatt hajtana végre. Ehhez 110 naptári napra volt szükség. Kisebb számokra valós eredmények alapján megállapítható, hogy egy RSA-80, azaz egy 80 jegyű tízes számrendszerbeli szám 21,5 óra, 70 jegyű 2 óra 6 perc és 40 jegyű 35 perc alatt állítható elő prímtényezők szorzataként. Megjegyezzük, hogy ez az idő nem feltétlenül a legrövidebb felbontási idő. Összehasonlításként azonban azt tudjuk, hogy az RSA-100 felbontásához szükséges „idő" 7 MIPS év.

Az aszimmetrikus titkosítás, legalábbis az ismert algoritmusok, sokkal költségesebbek, mint a szimmetrikus módszerek. Az RSA például n bites üzenetet $O(n^3)$ időben kódol, azaz 1024 bitet 10^9-el arányos időben. A többi kódolás sem gyorsabb, legfeljebb az az előnyük, hogy ugyanazon biztonság eléréséhez rövidebb kulcsot kell használni. Ez azonban nem mindig váltható át sebességnövekedésre, hiszen a kódolandó blokkok száma a modulus méretének változásával szintén változik. Ennek megfelelően a hardveres megoldások is sokkal lassúbbak. Stinson a könyvében azt írja, hogy egy DES csip körülbelül 1000-szer gyorsabb, mint egy RSA. Rivest, Shamir és Adlemann szerint az $M^e \pmod{n}$ kongruencia megoldása legfeljebb $2*\log_2(e)$ szorzás és $2*\log_2(e)$ osztás elvégzését követeli meg, ha az alábbi algoritmust használjuk (a kódolás és dekódolás műveletigénye hasonlóképpen számolható, ehelyett d-ből):

Step 1. Legyen $e_k e_{k-1} ... e_1 e_0$ a bináris reprezentációja e-nek.
Step 2. Inicializáljunk egy C változót, C = 1.
Step 3. Ismételjük a 3a és 3b lépéseket, for i = k; k-1;...; 0:
Step 3a. Legyen C a C^2 n-nel történő osztás utáni maradéka (C = maradék(n I C*C).
Step 3b. If e_i = 1, then C = maradék (n I C*M)
Step 4. Vége. Most a C felvette az M kódolt formáját.

Egyetemi körülmények között, nem túl erős, 450 MHz-es PC-n a közelmúltban elvégzett tájékoztató jellegű mérési eredmények tanúsága szerint az RSA algoritmuson alapuló műveletek elvégzésének időspektruma széles, és bizonyos függést mutat a paraméterek megválasztásától. A kulcsgenerálás ideje például egy 4096 bites RSA kulcs esetében 0,2 másodperctől 13 percig terjedő szórást mutatott, 5-6 perces átlag mellett. A 2048 bites kulcs generálási ideje ezzel szemben nem ment 1 másodperc fölé.

Az RSA algoritmus kódolásának futási ideje egy rövid adaton eltérő értékeket mutat, ha az exponens változott, és akkor is, ha a modulus hossza megnőtt. A kis exponensű (3, 65537) kulcsok esetében a futási idő jóval rövidebbnek bizonyult (előfordult 3 nagyságrendbeli különbség is), mint a véletlen előállítású nagy exponensű kulcsok esetében (6-7 másodperc). A kódolás ideje függést mutatott a kulcs hosszától is. Amíg 1024 bites kulcsokkal való kódolás szórása néhány tized másodperc, addig a 4096 bites kulcs esetén a szórás eléri a 6,5 másodpercet is, ami jelentős eltérésnek számít.

A dekódolás időigénye a mérési eredmények alapján nem a titkos kulcs exponensétől, hanem a modulus hosszától függ. Rövid kódolt üzenet dekódolása 2048 és rövidebb kulcshosszok esetében 1 másodperc alatti volt, míg a 4096 bites kulccsal történő dekódolás 4,5 másodpercnyi időt vett igénybe.

A digitális aláírás készítése azonban nemcsak az RSA kódolást jelenti, hanem az üzenetből képzett hash értéket is el kell előtte készíteni. A két művelet együt-

tes időigényére is jellemző volt, hogy a kulcs méretétől erősen függnek. Az aláírás elkészítése 512-bites kulccsal kb. 40 ezredmásodpercet, míg 4096 bites kulcs esetén kb. 4500 ezredmásodpercet, azaz 4,5 másodpercet vett igénybe. A lenyomatkészítő függvény választása a futási eredményt igen kis mértékben (<10%) befolyásolta.

Az aláírás készítése és ellenőrzése közötti összehasonlítás azt mutatta, hogy az aláírás ellenőrzésének ideje függ az exponens választásától, kis exponensnél gyorsabb volt az ellenőrzés, mint a készítés, nagy exponensnél pedig fordítva. Az aláírás és az ellenőrzés szélsőséges értéke is belül volt 7 másodpercen, jellemző értékük 1 másodperc alatti volt.

A felhasználás és szabványosítási kezdeményezések

A kriptográfia gyakorlati használhatósága tehát erősen függ a paraméterek megválasztásától. A gyakorlat azt igényli, hogy gyorsan, kényelmesen és hatékonyan lehessen megvalósítani nyilvános kulcsú rendszereket, azonban a biztonsági követelmények kielégítése már más jellegű feladat. A biztonság és a használhatóság egymásnak ellentmondó követelmények, így minden rendszer megbízhatósága attól függ, hogy milyen kompromisszumokat hoztak meg a tervezés során a használhatóság javára. El kell azonban ismernünk, hogy egy olyan rendszer, mely negyedórákig generál egy kulcsot, vagy egy aláírást, nem fog széles népszerűségnek örvendeni, másik oldalról azonban az a megvalósítás sem számíthat tömeges elterjedésre, melyről kiderül, hogy a kódolt üzenetek könnyen és gyorsan megfejthetők. Különösen fontosak az intelligens kártyákon kialakított rendszerek esetében a paraméterválasztások, mivel ezek az asztali gépek kapacitásánál kisebb erőforrásokkal rendelkeznek. Az intelligens – processzoros – kártyák olyan számítógépek, melyek nem rendelkeznek önálló áramforrással, azt kívülről kell biztosítani a számukra. Ezt szolgáltatja a kártyaolvasó. A tápfeszültség megjelenése után a kártya számítógépként viselkedik és kommunikál. A műveletek gyorsasága érdekében a paramétereket a lehető legkisebbnek, vagy binárisan könnyen kezelhetőnek szokták választani. Találkozhatunk például e=3 exponenssel megvalósított nyilvános kulcsú rendszerrel a piacon, habár Coppersmith 1997-es tétele alapján a kis exponenssel kódolt rövid üzenetek könnyen megfejthetők. Ez önmagában azonban még nem jelenti azt, hogy a kifejlesztett rendszer rossz, vagy nem használható. Azt azonban feltétlenül figyelembe kell vennie a rendszer használójának, hogy rá ez a veszély leselkedik.

Kripto-rendszereket általában a biztonság növelése érdekében szoktak üzembe állítani, a biztonság megteremtése azonban nem nélkülözheti a kockázatmenedzsmentet. Elképzelhető, hogy egy megvalósítás során a kis exponens felvállalható kockázatot jelent a megrendelő számára, egy másik helyen azonban ez már túl lesz az elviselhető fenyegetéseken. A rejtjelezéshez számos tévhit is kapcsolódik. Egyik legjobban elterjedt tévhit az, amikor a megfejthetetlenséggel a teljes kipróbálás módszerét azonosítják. Másik tévesen levont következtetés az, hogy ha a Föld atomjainak száma 10^{97}, a világűr élettartama 10^{18} másodperc, akkor az RSA-t 50 decimális számjegyű kulccsal csak annyi idő alatt lehetne megfejteni, amíg a Föld összes atomjait egyenként megszámolnánk.

A tévhitek számos esetben túlértékelik egy kriptorendszer biztonságát, ami nem nyilvánvaló módon megnöveli az alkalmazás kockázatát. A kockázatok megállapítása és kezelése azonban minden esetben helyzetfüggő és túlmutat a kriptográfia lehetőségein. A kockázatmenedzsment kialakítása a biztonsági rendszerszervezés területéhez tartozik.

A szerkesztő megjegyzése: a Híradástechnikában már többször foglalkoztunk a kódolással és ezen belül az RSA eljárással. A szerzők a prímszámválasztás fontosságát, nem elég körültekintően választott számok esetén a véges idejű feltörhetőséget ismertették. A digitális biztonság és digitális aláírás területén igyekszünk minden nézetnek helyt adni. Ugyanakkor ügyelünk arra, hogy ne foglaljunk állást olyan témában, melyben a legkiválóbb matematikusok [7, 8, 12, 13, 14] sem jutottak egyértelmű számszerű megállapításra. A biztonság és a kockázat számos mellékkörülménytől függ. Javasoljuk ezért, hogy gyakorlati döntéseiknél az olvasók saját feladataik ismeretében válasszanak, és ennek során a titkosítás, biztonsági eljárások különböző oldalú megvilágításait értékeljék.

Irodalomjegyzék

1. Directive 1999/93/EC of the European Parliament and of the Council of 13 December 1999 on a Community framework for electronic signatures (1999. 12. 13.)
2. 2001. évi XXXV. (IX. 1.) törvény az elektronikus aláírásról
3. ETSI ES 201 733 V1.1.3 Szabvány az elektronikus aláírás-formátumokról
4. Alfred J. Menezes, Paul C. van Oorschot and Scott A. Vanstone: Handbook of Applied Cryprography, October 1996, CRC Press
5. Jalal Feghhi, Jalil Feghhi and Peter Williams: Digital Certificates / Applied Internet Securtity, 4th Printing, April 2000, Addison-Wesley
6. D. R. Stinson: Cryptography, Theory and Practice, CRC, 1995.
7. R. L. Rivest, A. Shamir, L. Adlemann: A Method for Obtaining Digital Signatures and Public Key Cryptosystems
8. Pethő Attila: Paraméterválasztás nyilvános kulcsú kriptográfiai rendszereknél, SZTAKI Kriptográfia szemináriumi előadás, 2002. május 7.
9. Endrődi Csilla: Diplomaterv, Elliptikus görbén alapuló nyilvános kulcsú kriptográfiai algoritmusok elemzése, BME, 2001.

10. Stefania Cavallar, Bruce Dodson, Arjen K. Lenstra, Walter Lioen, Peter L. Montgomery, Brian Murphy, Herman te Riele, Karen Aardal, Jeff Gilchrist, Gerard Guiller, Paul Leyland, Joel Marchand, Francois Morain, Alec Muett, Chris and Craig Putnam, and Paul Zimmermann: Factorization of a 512-bit RSA Modulus; 1999. augusztus 22.
11. Horváth László, dr. Lukács György, dr. Tuzson Tibor, Vasvári György: Informatikai biztonsági rendszerek, Ernst&Young, 2001.

Felhívás

A **NOVOFER Alapítvány** kuratóriuma kéri a gazdasági tevékenységet folytató társaságok, a kutatással, fejlesztéssel, oktatással foglalkozó intézmények, a kamarák, a műszaki és természettudományi egyesületek, az érdekvédelmi szervezetek vezetőit, ill. tisztségviselőit, és a Gábor Dénes-díjjal korábban kitüntetett szakembereket, hogy az évente átadásra kerülő belföldi **GÁBOR DÉNES-DÍJ**-ra terjesszék fel azokat az általuk szakmailag ismert, kreatív, innovatív szellemű szakembereket, akik:

- kiemelkedő műszaki-szellemi tevékenységet folytatnak,
- jelentős szellemi alkotást hoztak létre,
- személyes közreműködésükkel hathatósan segítik az innovatív munkát,
- a környezet védelme területén kimagasló eredményt értek el,
- példamutató munkájukkal környezetükben élesztik a kreatív kedvet, alkotó szellemet,
- a vezetésük alatt álló szervezetnél meghatározó szerepet vállaltak az eredményesen végezhető alkotómunka infrastrukturális feltételeinek megteremtésében.

A felterjesztésnek tartalmaznia kell:

- adatlapot, az alábbi adatokkal:
  a jelölt neve (asszonyoknál leánykori nevet is), születési hely, év, hó, nap,
  pontos lakcím (irányítószámmal) és telefon,
  munkahely neve, címe, telefonszáma, munkahelyi beosztás,
  a felterjesztő (jelölő) személy és szervezet neve, a szervezet címe, a felterjesztő beosztása, telefonszáma, az ügyintéző neve, címe, telefonszáma,
- az ajánló szakemberek neve, munkahelye, beosztása, levelezési címe, telefonja,
- a jelölt szakmai képzettségének és munkásságának rövid ismertetését,
- a felterjesztés (jelölés) indokát a felterjesztő aláírásával (legfeljebb 3 db A4-es gépelt oldal terjedelemben), amelynek alapján a szakembert díjazásra javasolják. Ehhez mellékelhető az indoklásban hivatkozott alkotás(ok), ill. szakmai eredmények leírása (a jelentős újítások, találmányok, hazai és nemzetközi kutatási-fejlesztési projektek vagy nemzetközileg is elismert tanulmányok, a jelentősebb szakmai tudományos cikkek jegyzéke),
- két, a jelölt szakmájában országosan elismert, tekintélyes szakembernek a jelölt kitüntetését támogató ajánlólevelét.

A kitöltött adatlapot, a szakmai életrajzot, az indoklást, az ajánlóleveleket és az esetleges mellékleteket tartalmazó felterjesztéseket összefűzve, a NOVOFER Alapítvány címére (1112 Budapest, Hegyalja út 86.) kérjük megküldeni 1 eredeti és 2 másolati példányban. A felterjesztéshez csatolni kell a felterjesztő személy részére megcímzett és felbélyegzett 2 db kisméretű válaszborítékot is.

Az adatlap a www.novofer.hu internetcímről letölthető, vagy kérésre faxon továbbítjuk!

Beküldési vagy postára adási határidő: 2002. október 10.
Eredményhirdetés és díjátadás: 2002. december közepe.

A postán beérkezett felterjesztések átvételéről a felterjesztők, az elbírálás eredményről a felterjesztők, a kitüntetést elnyerők esetén a felterjesztők, az ajánlók és a díjazottak közvetlen értesítést is kapnak. A kitüntetettek személyét, a kitüntetés indokát a díjátadást követően, a szaksajtó segítségével is nyilvánosságra hozzuk.

További felvilágosítást ad: Kosztolányi Tamás titkár (Tel./fax: 319 8916, tel.: 319 8913).

Garay Tóth János
a kuratórium elnöke

Kriptográfiapolitika szeptember 11. előtt és után

DÉNES TAMÁS

E-mail: titoktan@freemail.hu

Budapest, 2002. május

A szeptember 11-i terrortámadás a XXI. század kezdetének, és remélhetőleg nem az egész századnak szimbólumává vált. Bár a sajtó, az elektronikus média foglalkozik e szörnyűséggel, azonban az ilyen típusú cselekmények előkészületeihez szükséges titkos kommunikációról tanulságokkal szolgálhat, eddig nem látott napvilágot. Ezért is érdekes és aktuális W. Diffie amerikai mérnök-kriptográfus [7] cikke, melyben e kérdéskörrel foglalkozik. Jelen cikkem W. Diffie gondolatmenetét és tanulságos következtetéseit foglalja össze a jövőbe vezető tágabb perspektívába, a biztonságos információs társadalom problematikájába helyezve e kérdéskört.

Míg a XX. század fordulóján a kriptográfia munkaigényes, nagy hibaszázalékkal dolgozó folyamat volt, amely nem volt képes többre, mint aránylag rövid írott szövegek kódolt formába való átalakítására, addig a XXI. század fordulóján a kriptográfia gyorsan, megbízhatóan, olcsón, számítógépesítve működik.

A telekommunikáció tökéletesedése és fontosságának növekedése a rendőri és hírszerző szervezeteket az elektronikus lehallgatás még kiterjedtebb használatára ösztönözte. Ugyanakkor éppen ezek a szervek tartanak attól, hogy a kriptográfia elterjedése az üzleti világban majd megfosztja őket olyan forrásoktól, melyekre eddig támaszkodhattak. Az eredmény (mondja W. Diffie, természetesen az USA-ra értve):

Küzdelem az üzleti világ (mely a kriptográfiát az elektronikus kereskedelem védelmére használja) és a kormányzati (titkosszolgálati) tényezők között (melyek felügyeleti lehetőségeik elvesztésétől rettegnek). Ezen harc egyik fő területe az exportkontroll.

2000. január 14-én az exportügyekkel foglalkozó hivatal (az USA-ban) a kriptográfiai hardverek és szoftverek exportjára vonatkozóan új szabályozást vezetett be. Ez a módosítás a rendszer erősségére való tekintet nélkül, korlátlanul engedélyezte a kriptográfiai termékek exportját, melyet a kriptográfiai „ipar" igazi győzelemként élt meg.

2001. szeptember 11-én az USA-t megtámadta az al-Kaida terrorista szervezet. S habár nincs megdönthetetlen bizonyíték arra, hogy a titkosítás szerepet játszott a hírszerzés ezen botlásában, a New Hampshire-i szenátor amellett érvelt, hogy ellenőrizni kell a titkosítási rendszereket. Ezt a felhívást azonban sem Bush elnök apparátusa, sem a kongresszus más képviselői nem támogatták.

Ebből a felvezetésből már világosan látszik W. Diffie gondolatmenete, amely az okokat a kriptográfia tömeges elterjedésében és az ellenőrzés liberalizálásában, azaz politikai okokban véli tetten érni. Legnagyobb meglepetésre cikkében a kriptográfiát homogén egységként tekinti, így annak lényegesen eltérő módszereit sem különbözteti meg.

Kriptográfiapolitika az USA-ban az elmúlt 30 évben

Az 1970-es években, sokéves hadiipari lappangás után, a kriptográfia kettős céllal jelent meg a nyilvánosság előtt. Elsőként létrejött a DES = Data Encryption Standard (Horst Feistel és az IBM más munkatársainak munkája), melyet azon kormányzati információk megóvására használtak, melyek védelmet élveztek, de nem tartoztak a titkos információk hivatalos kategóriájába. A fejlesztés második szakasza a nyilvános kulcsú titkosításhoz vezetett. Ahhoz a technológiához, amely ma az internetes kereskedelem biztonságának alapját szolgáltatja. A kormányzati válasz erre az volt, hogy megpróbálta „bekebelezni" a kriptográfiát. Az NSA (National Security Agency) remélte, hogy létrejön egy amerikai oktatási tanács, amelynek felügyelete alá tartozik a kriptográfiai kutatás és publikálás, így az NSA-n keresztülmenő kriptográfiai publikációk véleményezésével befolyásolható a lehetséges nemzeti biztonsági politika. Ez a kísérlet bukáshoz vezetett. Bár a kutatás és publikálás felügyelete kicsúszott a kormány hatásköréből, a katonai fejlesztések kontrollja azonban nem.

Az USA-nak tehát nemcsak megvolt a gazdasági ereje az export hatékony szabályozására, de hajlandósága is volt erre. Először a külügyminisztériumhoz tartozott az exportszabályozás, így a nemzetbiztonsági szempontok védelme is, mivel ezeket a termékeket a hadianyag kategóriába sorolták. Az exportfelügyelet természetesen erősen függött a termékek felhasználási céljától. A hidegháború alatt a legtöbb berendezés,

amely az USA és szövetségesei együttműködésében készült, COCOM listán volt. A különbségtétel a katonai és civil kriptorendszerek között nem történt meg. Néhány esetben, mint például az „ellenség-barát" (Friend or Foe) felismerő rendszereknél (amelyek a repülőgépeket azonosítják katonai radarokkal), a besorolás nem egyszerű, hiszen ezek a rendszerek polgári és katonai célokat is szolgálnak.

Az export célja egy olyan módszer kifejlesztése, amely a lehető legkevésbé avatkozik bele a nemzetközi kereskedelembe, miközben korlátozza más országok katonai lehetőségeit, arra nézve, hogy fenyegessék az USA érdekeit. Egy billió dolláros elektronikus átutalást védő kriptorendszer ugyanis nem különbözik egy olyantól, amely egy szigorúan titkos üzenet megóvására alkalmas.

Az alábbi három tényező következtében alapvetően megváltoztatta a kriptográfiai export-kontroll szerepét:

1. A tipikus amerikai cég áruinak több mint felét külföldön értékesíti, s ehhez versenyképes árat kell biztosítani.
2. A használhatóság és hatékonyság érdekében a biztonsági rendszert a kezdetektől fogva integrálni kell a termékhez. Egy már kész rendszerhez kriptográfiai rendszer hozzáadása nem kívánatos, még akkor sem, ha ez esetleg lehetséges lenne.
3. Egy termék két változatának előállítása költséges.

A bevezetett exportkontroll határt szabott az erős kriptográfiai eszközök alkalmazásának, nemcsak külföldön, de belföldön is. Az 1990-es évek elején változások történtek az exportpolitikában.

Első lépésként megegyezés jött létre (1992-ben) az NSA (nemzetbiztonsági szolgálat), a kereskedelmi minisztérium és az RSA (a kriptográfiai szoftverek vezető gyártója) között. Ez alapján jóváhagyták a 40 bitnél nem hosszabb kulcsú algoritmusok exportját.

Néhány hónappal Clinton elnök hatalomra kerülése után, a kormány a „Clipper" titkosító rendszert ajánlotta kompromisszumként. Ezt a rendszert erősen ellenezte az ipar és a Civil Szabadság Csoport, de végül elfogadták mint „Federal Information Processing Standard", ám soha nem terjedt el igazán a piacon. A Clipper által megnövekedett aggodalmakra válaszul az NRC (National Research Council) megjelentette a CRISIS (Cryptography's Role in Securing the Information Society) jelentést A kriptográfia szerepe az információs társadalom biztonságában címmel, 1996 nyarán.

A kormánypolitika ellenzői előrevetítették annak lehetőségét, hogy az NRC jelentés a Clinton-kormány kriptográfiapolitikáját nem támogatja. A jelentés az alábbi következtetésekre jutott:

- a kriptográfia széles körű használata több előnnyel jár, mint hátránnyal,
- a kormány jelenlegi politikája nem felel meg az információs társadalom biztonsági kívánalmainak,
- az exportlehetőségek megkönnyítése mellett foglaltak állást.

1996-ban az USA kongresszusa is foglalkozott a kriptográfiai exporttal, ám a törvényjavaslatok egyike sem kapott elegendő szavazatot. Ugyanebben az évben Daniel Bernstein a California Egyetem (Berkeley) matematikushallgatója elhatározta, hogy a törvény semmibevétele helyett – ahogy azt a legtöbb kutató tette – nyilvánosan követeli a jogot, hogy egy új kriptográfiai algoritmus kódját elektronikusan közzé tegye. Nem folyamodott exportengedélyért, jelezve ezzel azt a nézetét, hogy az exportkontroll alkotmányellenes. Ehelyett a szövetségi bíróságnál kereste igazát. Bernstein győzött a bíróságon. 1998-ban a Clinton-adminisztráció jóváhagyta, hogy nem kell ellenőrizni a „tömegpiac"-i kriptográfiát, amely 64 bitnél rövidebb kulcsot használ.

Ugyanebben az időszakban az USA által vezetett ECHELON rendszer (amely akkor már több mint 20 éve létezett) botrányos körülmények között napvilágra került. Az ECHELON rendszer egy UK–USA-megállapodás eredménye.

1999. szeptember 16-án Albert Gore, az USA alelnöke és elnökjelöltje kihirdeti, hogy a kormány kapitulál. A kulcs hossza többé nem volt döntő tényező a kriptográfiai termékek exportálhatóságában. Az új szabályok megosztották a piacot a vásárló típusa szerint. A kiskereskedelmi termékek is szabadon exportálhatók, például a Windows NT erős titkosítással sem képezi már az exportkontroll tárgyát. Az új szabályok okos kompromisszumnak mutatkoztak az üzlet kívánalmai és a biztonsági szervek között.

2000 júniusában a Miniszterek Európai Tanácsa kihirdette a kriptográfiai exportkontroll végét az EU-országokon belül és közeli partnerei között (pl.: Csehország, Lengyelország, Magyarország, Svájc, Japán, USA). 2000. július 17-én, válaszul az EU-liberalizációra, az USA hasonlókat produkált: többé nem kötelező az exportengedély a kriptográfiai termékek kivitelére a 15 EU-országba (és az EU-városokba). A váltás oka, hogy a szoftvergyártás ekkor vált igazán nagy üzletté. A legtöbb szoftvergyártó cég szétválasztotta a termékeket „tárgykód"-ra és a titkosan kezelt „forráskód"-ra. Éveken át a forráskód megosztása a felhasználókkal a „hobbistákra", néhány kutatóra és néhány „megszállottra" korlátozódott. Ez az 1990-es évek közepén változott meg, amikor néhány cég bevezette a nyitott forrású operációs rendszereket, amelyek egyre nagyobb részt hasítottak ki a szoftverpiacon.

Az eredmény

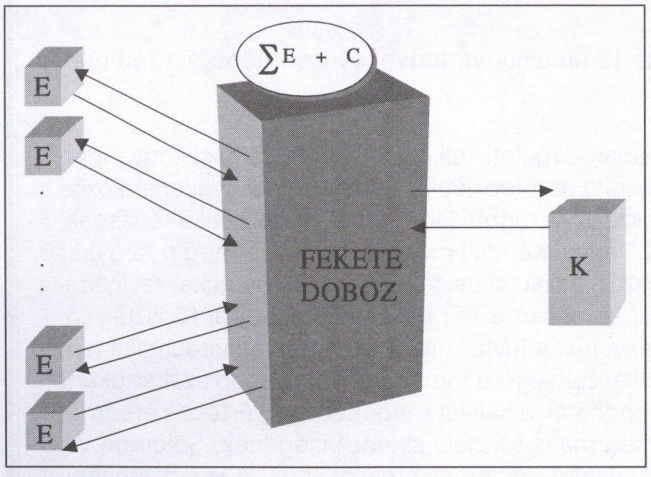
A szoftverek kontrollálhatóságának általános csökkenése különösen komoly fenyegetés a kormány azon erőfeszítéseit tekintve, melyek az erős kriptográfia exportját próbálták megtiltani. A nyílt forrású szoftverek széles körben terjedtek a weboldalakon. Ha egy program (főleg mint operációs rendszer) elhagyja az USA-t kriptográfia nélkül, a külföldi programozók jóval könnyebben készíthetnek hozzá saját kriptográfiai komponenst, melynek következtében „megszerzik" a

forrás operációs rendszer tulajdonjogát. Erre az USA exportkontrolljának minimális befolyása lehet csak.

Kriptográfia + politika ≠ kriptográfiapolitika

W. Diffie fenti gondolatmenete azt tükrözi, hogy az USA-ban a kriptográfia mint „iparág", élesen szemben áll a politikai érdekekkel, azaz erőteljesen érvényesül a fenti alcímbeli egyenlőtlenség. Megkockáztathatjuk azt az állítást, hogy ez minden olyan országra igaz, amelyben kialakulóban van a globális e-kommunikáció, azaz a napjainkban rohamosan terjedő információalapú társadalmi modell.

Mindez könnyen érthető, ha átlátjuk e modell lényegét, amely tulajdonképpen a digitális információk továbbításának és tárolásának dömpingjére épül. Ha a XX. század utolsó évtizedeire az információrobbanás volt jellemző, akkor a XXI. század első évtizedét nevezhetjük az „információ-láncreakció" évtizedének. Az egyének, a legkülönbözőbb társadalmi csoportok, szervezetek egyre több szálon kötődnek eme globális (idő- és térbeli korlátokat átívelő) rendszerekhez, így kialakul az információfüggőség. Az információ teljes személyiségünk, pszichikai, egzisztenciális létünk „digitális leképezésére képes". Tehát eme globális kommunikációs rendszerekkel olyan óriási fekete doboz modellt építünk (1. ábra), melynek birtoklása soha nem látott hatalomkoncentrációt eredményez.



1. ábra Globális kommunikáció modellje[1]

A fekete dobozban tárolt mérhetetlen mennyiségű információ az „egyszerű kommunikáló ember" (E) számára teljesen áttekinthetetlen, azaz az információk jelentős része nem áll össze ismeretté, hiszen ez csak bizonyos vonatkoztatási rendszerek (referenciainformációk) birtokában lehetséges. Tehát a fekete dobozt birtokló „tulajdonos" számára szinte tetszőleges manipulációs lehetőség kínálkozik. A digitális manipulációs eszközök tárháza viszont kimeríthetetlen, így ezek segítségével az „egyszerű felhasználók" számára észrevehetetlenül akár virtuális világok is generálhatók, melyek következtében ugyanezen felhasználók tömegeinek egzisztenciális biztonsága azonossá válik az információbiztonsággal. Ez az, ami miatt az információbiztonság és így a kriptográfia szerepe is jelentősen felértékelődik.

Szeretném azonban hangsúlyozottan felhívni a figyelmet a kriptográfiai eszközök két alapvetően különböző módszercsaládjára, nevezhetjük két különböző kriptográfiai filozófiának is. Az egyik a rejtjelezés eszközcsaládját fedi le, míg a másik a XXI. századi reneszánszát élő sztegonográfia (részletesebben lásd [2], [3], [4]). Az első módszercsalád a nyílt üzenet felismerhetetlenné tételét célozza meg, azaz olyan eljárásokat alkalmaz, amelyek eredményeként véletlen zajszerű jelsorozatot kapunk, melynek megértése csak bizonyos kulcsok segítségével válik hozzáférhetővé. A második, azaz a sztegonográfia eszközcsaládja ettől merőben eltérő filozófiát használ. Ekkor ugyanis a nyílt üzenetet nem rejtjelezzük, hanem egy másik szintén tökéletesen értelmes, úgynevezett „fedő szövegben", vagy „fedő üzenetben" rejtjük el. Így a kommunikációs csatornán küldött üzenetről egyáltalán nem eldönthető, hogy az tartalmaz-e rejtett üzenetet!

A gépi kapcsolatok egyre jobban kizárják a hagyományos értelemben vett személyes azonosítást, a tapasztalatokon nyugvó ellenőrzést, így különböző mesterséges azonosító eszközöket kell alkalmaznunk. A mesterséges azonosításhoz sok titkos kód, jelszó, kulcs megőrzésére, tárolására kényszerülünk, hiszen ezek mindegyike számunkra, vagy más közös érdekeltségű csoportok számára értékes információkat takar (hitelkártyák, telefonkártyák, igazolványkártyák, PIN kódok és jelszavas azonosítók stb.), akárcsak a fekete doboz „labirintusának titkos ajtaja". A kriptográfia valódi tömegcikké és egyben tömegszükségletté válik. Ezt állította középpontba W. Diffie elemzése szerint az USA politikája, a valódi kérdést azonban eltakarja a kriptográfiai exportkontroll középpontba állítása, mivel ez azt sugallja a politika számára, hogy az informatikai biztonság kézben tartható bizonyos rejtjelezési eljárások és azok kulcsainak szabályozásával. A digitális technika azonban a sztegonográfiának is kiváló lehetőségeket biztosít a titkolni kívánt üzenetek elrejtésére, méghozzá az alapvetően a rejtjelezett információk felfedésére és megfejtésére gyártott kriptológiai eszközök számára teljesen észrevehetetlen módon.

Itt érdemes felidézni W. Diffie alapgondolatát a XXI. századi kriptográfiáról:

[1] Az 1. ábra modellje tulajdonképpen egy megsokszorozott Turing-modell (részletes kifejtése megtalálható [2], [6]-ban), azaz pontosan úgy „működik", mintha sok-sok Turing-tesztet végeznénk párhuzamosan, ahol mindenki a géppel kommunikál elektronikusan, vagyis mindenki lehet kérdező (K) és kérdezett (E), a gép pedig összegyűjti és tárolja a információkat (ΣE). A Turing-teszt eredetileg a mesterséges és természetes intelligencia megkülönböztetését szolgálta, még a számítástechnika kezdetei előtt. Turing zsenialitása sem volt elegendő annak megsejtéséhez, hogy tesztje a XX. század végére hétköznapi gyakorlattá válik.

„... a XXI. század fordulóján a kriptográfia gyorsan, megbízhatóan, olcsón, számítógépesítve zajlik és az egész földön, másodpercenként billió bitet képes feldolgozni."

Mindezek fényében megkérdőjelezhető, hogy a szeptember 11-i terrortámadás előkészítésének informatikai okait valóban elegendő-e a kriptográfiai exportkontroll liberalizációjában keresni?!

Vagyis a modern, digitális sztegonográfia olyan titkosítási filozófia és eszközrendszer, amely rossz kezekbe kerülve valódi „csodafegyver", míg jó kezekben „csoda", azaz új lehetőség egy emberközpontú, biztonságos információs társadalom létrehozásához. Ez is tanulsága szeptember 11-nek, amely egyúttal felkiáltójel a jövő terrorizmus elleni küzdelméhez is!

Irodalom

1. Dénes Tamás: ECHELON az e-társadalom információpajzsa? Híradástechnika, 2001/6. 14–19.
2. Dénes Tamás: Biztonságos Információ(s) Társadalom INFO TÁRSADALOMTUDOMÁNY, 2001/53.
3. Dénes Tamás: SZTEGONOGRÁFIA – rejtett információk rejtjelezés nélkül. Híradástechnika, 2001/8. 15–21.
4. Dénes Tamás: Turing-teszt az információs társadalomban. A. M. Turing születésének 90. évfordulójára ajánlva E számunkban.
5. W. Diffie, S. Landau: September 11th did not change cryptography policy. Notices of the American Mathematical Society, Vol.49, april 2002.
6. I. S. Moskowitz: Information Hiding 4th International Workshop, Pittsburgh, USA (Proceedings).
7. Andreas Pfitzmann (Ed.): Information Hiding Proceedings of Third International Workshop, IH '99 Dresden, Germany, 1999.
8. M. Wu, B. Liu: Multimedia Data Hiding Springer-Verlag, New York, 2002.

Dr. Dénes József
1932–2002

Sokunk barátja és pályatársa, Dénes József augusztus 19-én elhunyt, hetvenévesen, több évtized után feltámadt súlyos betegségében.

Élete során az algebrai kódolásban és a kriptográfiában elméleti felkészültségére támaszkodva igen sok értékes eredményt ért el. Munkáját szorosan kapcsolta az informatikai és távközlési gyakorlat közös részeihez. Matematikai aktivitását fémjelzi, hogy társszerzője tudott lenni a két matematikai óriásnak, Erdős Pálnak és Turán Pálnak. A latin négyzetekről írt Dénes–Kendall-monográfia ma is még a tárgykör világszerte citált könyve. Éveken át sikeresen járult hozzá a hazai fejlesztésű számítástechnikai szolgáltatások kibontakoztatásához, utóbb az egykori SZKI-ban, korábban a KFKI számítástechnikai főosztályán. Azután is, hogy az SZKI-ból visszavonult, hozzájárulásai voltak a BME mikrohullámú híradástechnika tanszékének fontos kódolási projektjeihez, kódelméleti felkészültségére támaszkodva; később szakértőként különféle hazai és külföldi kriptológiai projektekhez. Mondhatni, a halálát megelőző évig értékes eredményeket ért el az IEEE Információelméleti Társaságot még ma is formáló személyiségének, Solomon W. Golombnak erősen Dénes József felkészültségét igénylő ipari akadémiai projektjeiben. A Híradástechnikában halála hónapjáig jelentek meg a távközlést és az informatikát művelők széles körének írt gondolatébresztő cikkei.

Akik igazán ismerték azoknak feltárult professzióján messze túlmenő érdeklődési köre, gazdag szellemi kultúrája és embersége. Fiával és unokáival együtt sokan gyászoljuk!

Csibi Sándor

Digitális szakadék

HORVÁTH GYULA

A Nemzetközi Távközlési Unió (ITU) legújabb kulcsszava a „digitális szakadék" (digital divide), ami azt fejezi ki, hogy a digitális távközlés, különösen az internet és a web használata a világ megosztója: a jól ellátott fejlett országok és a szükségesnél jóval kevésbé ellátott, főleg a legkevésbé fejlett országok[1] között szakadék tátong, aminek áthidalása szakmánk legfontosabb politikai föladata. A téma a „Hiányzó láncszem" címen híressé vált Maitland-jelentésig nyúlik vissza (1984), amiben a szakadékra még csak a vezetékes telefonon alig túlmenően hívták fel a figyelmet. A kérdés a világméretű szegénységre vezethető vissza.

A „Hiányzó láncszem" c. jelentés

A jelentést az ITU Meghatalmazottak Értekezletének 1982-ben hozott határozata alapján létrehozott Független Bizottság készítette Sir Donald Maitland vezetésével. Az elsők között tett javaslatokat a fejlődés fölgyorsítására, kitűzte a politikai célt: a 21. század első felére csaknem mindenki jusson hozzá telefonhoz és ésszerű határidőn belül más szolgálatokhoz is. A jelentés átfogóan foglalkozott a távközlés társadalmi jelentőségével, a fejlesztés műszaki és gazdasági problémáival és megfogalmazta a javaslatokat. Az egész világon nagy érdeklődéssel tárgyalták az addig szokatlan politikai üzenetet, fellobbantotta a képzeletet, pedig akkor még nem voltak ismeretesek a tíz évvel később napvilágra hozott, ebben a cikkben is megtalálható elgondolkoztató adatok.

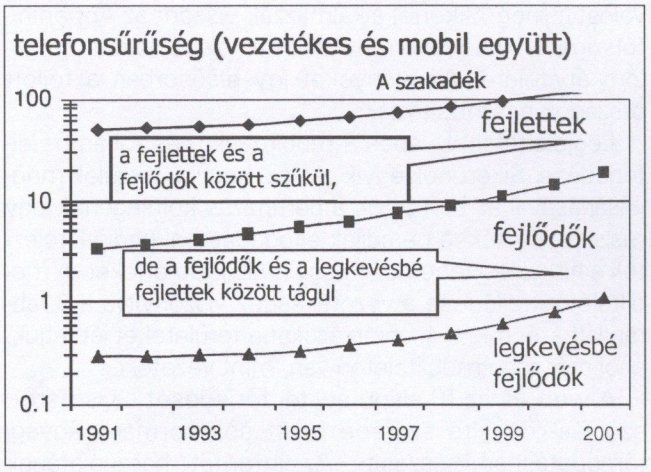
Miért nem hozott széles körű eredményt?

A nyers válasz egyetlen mondat: a hiányzó láncszem valójában a pénz, amit senki sem fordított megközelítőleg sem a szükséges mennyiségben a jelentésben fölsorolt teendők finanszírozására. A téma gazdája, az ITU nem pénzosztogató szervezet. Koncepcionálisan nem is lehet, mert erre a célra olyan nemzetközi intézmények vannak, mint a Világbank, az EBRD és mások. Még ma sincsenek egyértelmű számítások, amelyek a távközlés hasznát, illetve a hiánya miatt elmaradó gazdasági előnyöket mutatnák. Emiatt számos kormány nem tulajdonít megfelelő fontosságot a távközlésnek; saját forrásainak és az általa mozgósítható külföldi erőforrásoknak szükséges részét nem fordítja a távközlésre, inkább a sürgősebb és látványosabb feladatait részesíti előnyben.

A Maitland-jelentés ugyan adott segítséget a fejlődő országoknak azzal, hogy tekintélyes nemzetközi szervezetre lehetett hivatkozni, de csak javasolni tudta azt, hogyan jussanak a távközlés fejlesztését alátámasztó érvek birtokába. Bár a jelentés elfogadásakor az ITU-ra bízott feladatok teljesítése haladt legjobban, az érintettek közül sokak szemében mindez a döntéshozók pótcselekvésének látszott.

A Maitland-jelentés közzététele óta elért eredmények

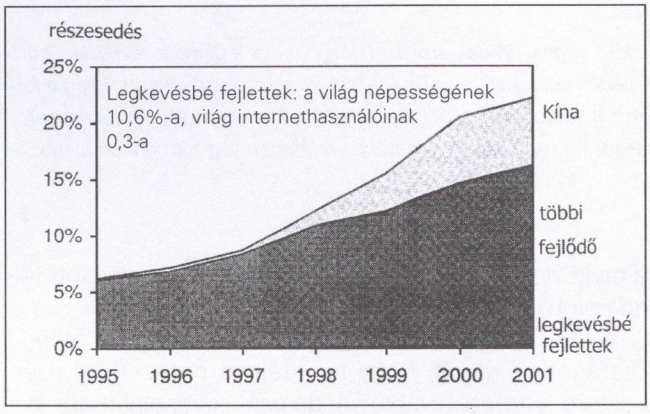
Mindenesetre a világ távközlése nagyot fejlődött, de nagyon nehéz kihámozni, hogy ebből mi köszönhető



1. ábra A szakadék a telefóniában zsugorodik

[1] Az ITU-dokumentumok szerint fejlett országoknak az Európai Unió, Izland, Norvégia, Svájc, Kanada, az Egyesült Államok, Japán, Ausztrália, Új-Zéland, Hongkong, Koreai Köztársaság, Szingapúr és Tajvan tekinthetők. A legkevésbé fejlett országok száma 49, a többi fejlődő.

a Maitland-jelentésnek. A leggyakrabban használt mutató, a telefonsűrűség növekedését az ITU adattára alapján szerkesztett diagram (1. ábra) mutatja. Eszerint a fejlett és a fejlődő országokbeli telefonsűrűséget ábrázoló görbék egymáshoz közelednek, míg a fejlődő és a legkevésbé fejlett országok görbéi még mindig távolodnak. A fejlődés legjelentősebb motorja a mobiltelefon színrelépése és gyors elterjedése. Az ITU következésképpen egyik háttéranyagában a „hiányzó láncszemet" a mobiltelefóniában véli megtalálni. Ugyanakkor a napjainkra már nagyon fontossá vált másik mutató, az internet elterjedése a digitális szakadék további tágulását jelzi (2. ábra).

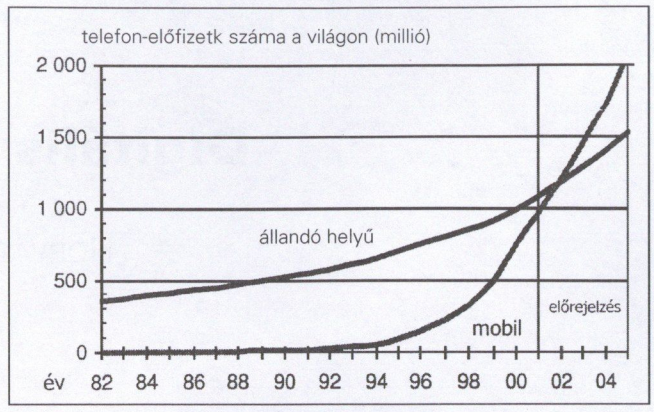


2. ábra Az internet és a rés bővül

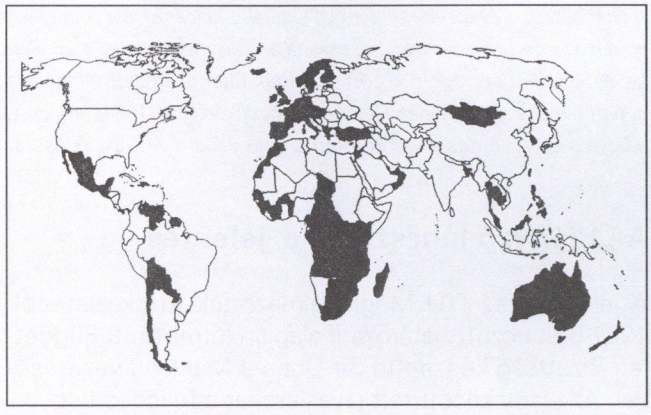
A fényvezető kábelek egyrészt jelentősen olcsóbbá tették a nagy távolságú, köztük a földrészek közötti telefonálást és az ebben az időben elterjedt távmásolást (faxot), másrészt átviteli kapacitásuk növelése elősegítette a gyorsan növekvő adatátvitel fejlődését, és átviteli kapacitást biztosítottak a terjedő internet számára. Például fölmerült az a gondolat, hogy egyes földrészeket, mint Afrikát és Dél-Amerikát a minden parti országba leágazó, tengerbe fektetett fénykábellel vegyenek körül. Ez anyagiak hiányában csak kis részben valósult meg. Sikerrel alkalmazzák viszont az Appenini-félszigetet körülvevő fénykábelt az olasz belföldi forgalom átvitelére. Az előnyöket így elsősorban a fejlett országok hasznosították.

Legjelentősebb lépés a mobil távközlés rohamos elterjedése. Sikerének egyik oka az, hogy a készülék megvásárlásával az előfizetők a beruházás költségének egy részét fedezték, a percdíjak elfogadásával pedig kifejezték e hírközlési lehetőség nagy használati értékét. A mobiltelefon fejlődése a vezetékeshez viszonyítva a 3. ábrán látható, míg a 4. ábrán azokat a területeket láthatjuk, ahol már több mobiltelefon van, mint vezetékes.

A web és az IP alapú átvitel terjedését az ismeretszerzés-könnyítő szerepe miatt sok kormány anyagi támogatásban részesíti. Az internet használatának akadálya a szoftverek és a számítógépek ára. Ez az újabb típusok, illetve változatok megjelenése során sem csökken, úgyhogy a támogatás jelentős részét számítógépek beszerzésére és a szolgáltatási díjakra kell fölhasználni.



3. ábra Mobil a világhálózatban



4. ábra Ahol már több mobil van, mint vezetékes (2001-ben, feketével jelölve)

A távközlés ma főként telefon, ezért használóinak esetleges írástudatlansága nem volt akadály. Az internetet azonban csak írni-olvasni tudók képesek használni, az írástudatlanok ezek áldásaiból írni-olvasni tudók segítsége nélkül ki lesznek rekesztve. A digitális szakadék betöltésének programja tehát az írástudatlanság felszámolásával együtt lehet csak teljes. Az igények kielégítettségével kapcsolatban nem feledkezhetünk meg arról, hogy egyes távközlési szolgáltatásokra sok embernek nincs is szüksége. Ezek számával nem érdemes növelni azok számát, akik még nem vették igénybe egyik vagy másik távközlési szolgálatot sem.

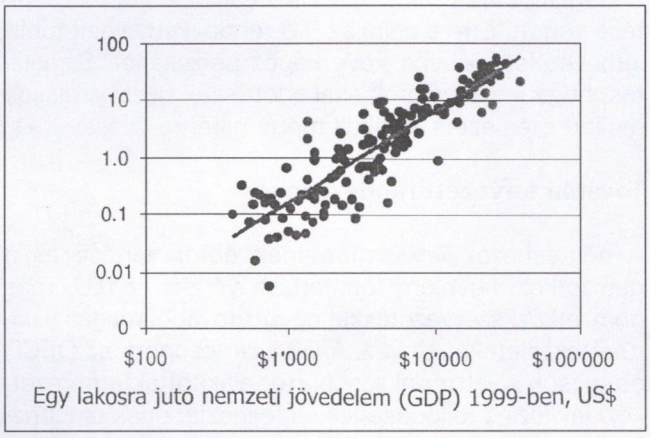
Csak új jelszó, vagy új álprobléma?

Amikor egy metaforikus elnevezés alatt futó tevékenység kifullad, gyakran új név alatt folytatják tovább. A szkeptikusok ezt a kudarc elkendőzésének, a gyér eredmények leplezésének tartják. Sokan inkább a Maitland-jelentésben foglalt javaslatok megvalósításának folytatására irányuló, kétségkívül jó szándékú törekvés új jelszavát látják a digitális szakadékban, az új metaforában.

Mások, utalva az azóta eltelt idő hozta, előbbiekben részletezett újdonságokra, a Maitland-jelentésben föltárt problémákat a digitális formájú információhoz való hozzáférésre terjesztik ki. Vannak, akik megjegyezték már, hogy ha azokra gondolunk, akik egész napon át az-

napi betevő falatjuk megszerzésével vannak elfoglalva (akiknek napi jövedelme legföljebb 2 dollár, mely utóbbi az ENSZ által meghatározott szegénységi küszöb), akkor számukra az információtechnikának semmi haszna nincs, a digitális szakadék csak egy álprobléma. Ugyanez az írástudatlanokra is érvényes. Ebben a megközelítésben az ITU, a mérsékelt eredményekre vezető okok elemezése után, a problémát a nemzetközi politikában azóta fölerősödött „szegény-gazdag országok" ellentéteként megfogalmazott problémakörbe helyezte át. Átalakította a jelen és a közeljövő problémájává, amivel lehetővé tette valós problémaként való kezelését.

Az egyes országok gazdasági ereje és bennük az internethasználat mértéke közötti összefüggést az 5. ábra szemlélteti.



5. ábra Internet és gazdaság

Az ITU szerepe

Az ITU sajtóközleményeinek végén saját magát így jellemzi:

„Az ITU világméretű szervezet, amely összefogja a kormányokat és az ipart a világ távközlési vállalatai és szolgálatai létrehozásának és működtetésének összehangolására; felelős a rádiós hírközlést is magában foglaló nemzetközi távközlés szabványosításáért, egyeztetéséért és fejlesztéséért, valamint a nemzeti távközlési politikák harmonizálásáért.

Küldetésének betöltése végett az ITU nemzetközi szabályzatokat és egyezményeket fogad el, melyek a frekvenciaspektrum földi és űrbeli használatát, valamint a geostacionárius pálya használatát irányítják, mint a nemzeti törvényhozás kereteit; szabványokat dolgoz ki tekintet nélkül az alkalmazott műszaki megoldásokra a távközlő rendszerek világméretű összekapcsolásának elősegítésére. Világméretű és regionális kiállításokat és fórumokat szervez, hogy összehozza a kormányok és a távközlési ipar képviselőit eszmék, ismeretek és technológiák cseréje végett az egész emberi közösség, különösen a fejlődő világ javára."

A Maitland-jelentésben foglalt javaslatok elfogadását és végrehajtását az ITU működésének alapjául szolgáló nemzetközi egyezményben lehetne kötelezővé tenni. Ennek megújítására a meghatalmazottak négyévenkénti értekezletén van alkalom, amire éppen 2002. szept. 23–okt. 18 között kerül sor. A javaslatok leglényegesebb pontjáról, a pénzügyi alapok megteremtéséről ennek az egyezménynek keretében gyakorlatilag nincs lehetőség, mert a nemzetközi pénzügyeket nem szokás a szakosított szervezetek útján intézni. Nincs olyan nemzetközi szervezet, ami a jelentésben foglalt javaslatok megvalósítását kötelezővé tehetné (az ENSZ-határozatok egy részét sem hajtják végre), másrészt sokkal súlyosabb világproblémák is vannak, mint például az éhezés, az ivóvízellátás, a levegő tisztaságának biztosítása stb.

WTDC02

A témának aktualitást ad az, hogy az ITU szervezésében 2002 márciusában zajlott le Isztambulban a harmadik Világ-távközlésfejlesztési Konferencia (World Telecommunication Development Conference, WTDC02), amelynek fő témája program kidolgozása volt a digitális szakadék áthidalására. Ezzel párhuzamosan a mexikói Monterreyben mintegy ötven államfő is a gazdagok és a szegények közötti rés csökkentéséről tárgyalt. A konferencia kézzelfogható sikerében természetesen elsősorban a fejlődő, még inkább a legkevésbé fejlett országok érdekeltek, bár a gazdasági válság hatására ma már a gazdag országok is úgy érzik, hogy piacaik bővítése és világméretű konfliktusok megelőzése végett érdekük a szegénység csökkentése.

A WTDC megnyitásakor az ITU főtitkára fölvetette a digitális szakadékkal kapcsolatban a távközlési közösség önvizsgálatának szükségességét arra vonatkozóan, hogy tudtán kívül nem tágította-e rendszeresen a digitális szakadékot. Szerinte az ITU erőfeszítései ellenére a kérdés bürokratikus útvesztőbe jutott. Most minden erő bevetésével támadást kell indítanunk, hogy 2005-re a világ mindegyik falujába eljusson a távközlés. Mind a közügyek intézői, mind a magángazdaságok mozgósítsák erőforrásaikat a közös cél, a lakosság javára. Hatékony cselekvési tervet kell végrehajtani, hogy 2005-re minden településen legalább egy, minden lakos által használható telefon legyen.

A két legnagyobb fejlődő ország, Kína és India hangoztatta annak szükségességét, hogy olcsó, szolgáltatásaiban az internet és az elektronikus levelezés lehetőségére összpontosító, lehetőleg helyben gyártott végkészülékekre van szükség. Szóba került a helyben hasznosítható tartalom előállításának problémája, de még nem merült föl az ezt szervezetten nyújtó információs közmű iránti igény.

A korábbi célkitűzések a telefonnal ellátott személyek minimális arányának elérésére vonatkoztak (min. 1%-os országos telefonsűrűség), most pedig a települések ellátásának mértékére. Ez több körülménynek tudható be. Az egyik az, hogy az eddigi norma szerint a telefonnal ellátott és a nem ellátott lakosok között keletkezik szakadék, amit a telefon-előfizetők korlátozott lehetőségeire tekintettel nagy közösségi beruházásokkal lehet csak mérsékelni. A szakadék országon belül is

megvan a társadalom anyagi javakkal jobban és kevésbé ellátott rétegei között. Utóbbiakat sok országban a magas távközlési, számos helyen az internet-díjszabások zárják ki eme eszközök használatából.

A másik körülmény a teleházak létesítése, ami nemcsak mindenki számára nyújt hozzáférési lehetőséget, hanem jelentősen megkönnyíti a jövőben a többi fontos távközlési szolgáltatás, elsősorban az internet bevezetését. Kiemelték, hogy a távközlő eszközök jelenléte a távgyógyítás, a távoktatás, távmunka stb. területén jelentősen hozzájárulhat az életminőség javításához.

Az isztambuli cselekvési terv[2]

Programok

A szabályozás problémái a szabad verseny előmozdítása során fölerősödtek. Nyilvánvaló a befészkelt (incumbent) szolgáltatók erőfölénye, ami az újonnan piacra lépők védelmét szolgáló aszimmetrikus szabályozást helyezi előtérbe. A szabályozás mértéke, a védelem formái jelentős vitatémává váltak. Nemzetközi szervezethez illően az ITU tagjainak abban kíván segíteni, hogy sajátos nemzeti körülményeik között leghatékonyabb szabályozást alakítsanak ki. Brazil tapasztalatok szerint a szabályozó hatóság legfontosabb jellemzői a következők legyenek: a) függetlenségét kormányváltások idején is folytonos működését biztosítva kell védeni, továbbá, hogy – csak bíróság előtt vitatható – határozatainak végrehajtása kikényszeríthető legyen; b) átlátható, a nyilvánosság elé tárt döntési folyamat; c) kiszámíthatóság; d) működés saját költségvetése szerint, szükségleteihez igazított ellátmányból; e) képzett és állandó személyzet, jogszabályban rögzített időtartamra kinevezett és politikai okokból nem elmozdítható vezetők; f) a fogyasztók védelme.

Technológiák és a távközlőhálózat fejlesztése, ami eddig is az ITU sikeresen megoldott feladata volt (ajánlások, kézikönyvek, együttműködési feltételek biztosítása).

Gazdaság és finanszírozás, a költségeket és díjszabásokat is beleértve. Cél a biztos hozzáférés a költséghatékony és értéknövelő, társadalmilag, gazdaságilag előnyös szolgálatokhoz és alkalmazásokhoz, ami csökkenti a társadalmi szakadékot. A finanszírozás vonalán kapjon prioritást a magántőke ott, ahol az állam szerepvállalására nincs kilátás, de törekedni kell az állami és a magántőke együttműködésére. A tagok támogatandók a sikerténvezők feltárásában és megvalósításukban avégett, hogy a szolgálatokat igazságos, megfizethető és költségorientált árakon lehessen nyújtani.

Emberi erőforrások kiépítése. Kulcsfontosságú, hogy a fejlődő országok a távközlés és az infokommunikációs technológia terén kompetenciára tegyenek szert, elsősorban a kormányzat és a szabályalkotók köreiben. Ez föltétele annak, hogy a nemzeti távközlési szabályozó hatóságok sikeresen működjenek[3]. Ugyanerre van szükség a szolgáltató szektor vezetőinek körében is.

Különleges programok a legkevésbé fejlett országok számára. A program segítséget irányoz elő, hogy a legkevésbé fejlett országok szükségletei városokban és vidéken ki legyenek elégítve. A korábbi „6 ország két évig" séma helyett „12 ország egy évig" sémára térnek át. Mivel eddig 20 ország már rész vett a programban, az új séma szerinti támogatás lefutása után már mindegyik legkevésbé fejlett ország részesül annak hasznából. Fontos az erőforrások lehetőleg közösen használható egységekbe, csapatokba szervezése.

A programok külön foglalkoznak a belvillongásokkal, háborúval, vagy természeti katasztrófákkal sújtott hat országgal, infrastruktúrájuk helyreállítása vagy újjáépítése végett. Erre a célra az ITU rendezvényeinek többletbevételét igénybe véve képez pénzalapot. Sajnálatos, hogy a szponzorok által adott szűk hozzájárulások mellett még ezt is indokolt megemlíteni.

További tervezett rendezvények

A döntéshozók lelkiismeretének ébren tartása és a nemzetközi figyelem fönntartása végett az ITU, más nemzetközi szervezetekkel együtt további magas szintű értekezleteket tervez. A G8 államcsoport, az OECD és mások saját munkacsoportot alakítottak a nemzetközi távközlés fejlődésének és fejlesztésének tanulmányozására, a fölmerült kérdések elemzésére és állásfoglalás kidolgozására. Néhány rendezvényük a táblázatban található. Ezen túlmenően számos nemzetközi szervezet, részben az ENSZ keretében, tart fönn a digitális szakadékkal vagy annak valamelyik részkérdésével foglalkozó munkacsoportot.

Szervezet ENSZ-közgyűlés	rendezvény tárgya ICT	időpont és helyszín New York, USA 2002. 06. 17–18.
Egyesült Nemzetek ICT munkacsoportja (UCT)	Világkonferencia	Tampere, Finnország 2002. 07. 28–31.
ITU	Meghatalmazottak Értekezlete	Marrakesh, Marokkó 2002. 09. 23.– 2002. 10. 18.
G8 (G7+Oroszország)	The Digital Opportunity Task Force	
ENSZ	Világ-csúcstalálkozó az információs társadalomról	Genf, 2003.12.10–12. Tunis, 2005.

1. táblázat

Multik és monopolhelyzetű cégek kapzsisága

A tőke működésekor saját magát igyekszik növelni, ezért olyan gazdasági tevékenységekbe igyekeznek

2 World Telecommunication Development Conference adopts comprehensive Action Plan to bridge the Digitla Division, 5 oldal, ITU preess release (2002-03-27); http://www.itu.int/newsroom/wtdc2002/

3 Lighting the Way: ITu and Telecommunication Regulators, ITU Feature story (2002. 03. 12.)

befektetni, amelyek viszonylag nagy haszonnal járnak. A műszaki fejlődés eredményeit is e cél elérése szempontjából értékelik. A tranzisztor föltalálása után a távközlési szakmában természetes cél lett a mikroelektronika elméleti és gyakorlati eszközeivel előállítható minden készülék, eszköz, berendezés megalkotása, de ebben sem mindig voltak tekintettel a társadalom tényleges szükségleteire. Kirívó példa erre a személyi számítógépek hardverének és szoftvereinek fejlesztése úgy, hogy azok szükségtelenül gyorsan elévüljenek, és használóik gyakran kénytelenek újabb változataikat megvásárolni. A WTDC02-n elhangzott politikai nyilatkozatokban ez a stratégia erős kritikát kapott.

A nyereség túlzásba vitt hajszolása gazdasági válságot okozott. Az alkalmazottak tömeges elbocsátása az államnak is kárt okoz, mert a munkanélküliek a közterhekhez való hozzájárulás helyett segélyek formájában az állami kiadásokat növelik.

Kitörési lehetőségek

Ázsiai és afrikai példák

Figyelemre méltó, hogy történelmi példák mellett a jelenkorban is vannak esetek, amelyekben szegény országok jelentős mértékben saját erejükből ki tudtak emelkedni szegénységükből. Szemünk előtt zajlott le a japán gazdaság és társadalom modernizálódása a 20. században, és annak utolsó évtizedeiben a délkelet-ázsiai „kis tigrisek" gyors fölemelkedése. Napjainkban újabb országok okoznak meglepetéseket. Kínában a távközlés fejlesztését, beleértve a hazai kutatást és gyártást, a kormány a nyitás óta központi kérdésként kezeli (2. ábra). A hazai kutatást és gyártást Indiában is megvalósították. A fejlődés hirtelen fölgyorsulására Brazília szolgáltatott legutóbb példát: 1999–2001 között annyi telefonvonal létesült, mint azt megelőzően 115 év alatt. E példák igazolni látszanak azt az ITU által levont következtetést, miszerint a kormányok hozzáállása igen nagy jelentőségű a digitális szakadék föltöltésében.

A példák azért is nagyon figyelemre méltók, mert olyan országokról szólnak, amelyek két-három évtizeddel ezelőtt elmaradott hálózatuk továbbfejlődését nagyrészt csak külső finanszírozással, importált berendezésekből fölépítve tudták elképzelni. Ez reményt ad arra, hogy az isztambuli cselekvési terv sikere számottevően nagyobb lesz, mint a Maitland-jelentésé volt.

Milyen mély a digitális szakadék?

Az infokommunikáció elterjedtségét kifejező statisztikai fajlagos mutatók nagy egységekre, országokra, országrészekre, nagyobb településekre vonatkoznak, elfedik az emberek szerint képezett csoportjain belüli különbségeket. A véleményekben, nyilatkozatokban gyakran előfordul a szolgáltatásokkal kapcsolatban a „megfizethető áron" jelző, de annak megjelölése nélkül, hogy kiknek a pénztárcájából fizethető meg. Ritkán utalnak arra, hogy az árnak nemcsak a létesítés költségeit, hanem az üzemeltetést is fedeznie kell. Érdemi válasz megalapozásához a szegénység olyan jövedelmi küszöbértékének ismeretére van szükség, ami lakó- és munkahelytől, foglalkozástól, műveltségi szinttől, nem utolsósorban egészségi állapottól és az elsőrendű létszükségletek (lakás, ruházat, élelem, egészségügyi ellátás) árától is függ.

Miért érdemes megfizetni?

Ma az internet használatára buzdító hirdetések hosszan sorolják az internet előnyeit. Akik az internetet elsősorban információforrásként használják, azok azonban azt is látják, hogy az internet nem szervezett, nem biztos, hogy hosszú idejű böngészés után meg fogják-e találni a keresett információt. Gutenberg óta kialakultak a lexikonok, enciklopédiák, kézi- és tankönyvek, hogy csak a legjobban szervezett tömeges információforrásokat említsük. Ezeknek az IT lehetőségeit kihasználó, korunk viszonyai között jobban használható (gyorsabban hozzáférhető, naprakész stb.) utódjai igazolhatják tartósan az infrastruktúra és a hozzá csatlakozó végberendezések költségeit. Már most vannak próbálkozások valamiféle információs közmű kialakítására, ahonnan alulról fölfelé meghatározott műveltségi és képzettségi szintig szükséges gyakorlati és elméleti ismeretek lekérdezhetők, nem csupán „megfizethető" áron, hanem lehetőleg ingyenes közszolgáltatásként. Nemcsak ez, hanem a tartalomnak nevezett, szervezett információfajták nagy mennyisége is meghatározza az internet iránti érdeklődést, mai gazdaságossági gondolkodás szerint azt, hogy mennyit érdemes az infrastuktúra használatáért fizetni.

Itt jelentkezik akadályként az írástudatlanság, melynek világátlaga a 15 évnél idősebb lakosok között 20,6%.

De még gondjuk van azoknak is, akik írásban nem tudnak érthetően fogalmazni, vagy más által fogalmazott szöveget megérteni. Ezek mentsvára a telefon, mert úgy, mint közvetlen személyes találkozáskor, beszélgetés közben a nem értett részletekre azonnal vissza tudnak kérdezni. Ezeknek az embereknek előbb képzettségben, műveltségben föl kell zárkózniuk azokhoz, akik között már beszélhetünk digitális szakadékról. A digitális szakadék nemcsak az infokommunikációval, hanem a szegénységgel és az élet számos más problémájával is kapcsolatos.

Globalizáció

A globalizációt a személyek, áruk és az információ szállításának minőségi ugrást jelentő fölgyorsulása tette lehetővé. Kérdése, hogy csak a legfejlettebb országok törekvése, vagy a többi ország lakosainak javára szolgál-e. Igennel csak akkor lehet e kérdésre válaszolni, ha az eladó országok hozzásegítik a vevő országokat ahhoz, hogy fölemelkedésükkel több, és az értéklánc maga-

sabb szintjén álló szolgáltatással és áruval ellentételezhessék a nemzetközi munkamegosztásban a gazdagabb országokból vásárolt árukat és szolgáltatásokat.

Érdemes idézni két józan javaslatot a globalizációs folyamat ügyében kidolgozott magatartásra. Anthony Giddens, a London School of Economics igazgatója, Tony Blair tanácsadója írta: „A piacnak nem lehet ellentmondani, az egyetlen lehetőség az, hogy a globalizáció folyamatát nemcsak gazdaságilag, hanem szociálisan és kulturálisan is az átlagember szolgálatára kényszerítsük." Ugyanez a gondolat Jacques Calvet (Peugeot) tollából: „Tudomásul kell venni, hogy menetel a globalizációs folyamat. A Franciaország előtti kihívás immár nem az, hogy harcoljon ellene, hanem az, hogy ellenőrizze, menedzselje és talán még civilizálja is."[4] E sorok írója szerint is a globalizáció során az egyes emberek a keletkezett lehetőségek ügyes kihasználása útján tudnak maguk számára a globalizációból hasznot húzni.

Az infokommunikációval foglalkozó embereknek tehát alapvető érdeke egyrészt, hogy a tartalom bővülése és minőségének javulása vonzza az információs infrastruktúrát használókat, másrészt, hogy a szegénység világméretekben annyira csökkenjen, hogy mindenkinek annyi távközlés használatára legyen lehetősége, amennyi saját felemelkedéséhez hozzájárulhat. Ez óriási feladat, mert sikeres megoldásához a szegénység káros hatását az élet többi területén is meg kell szüntetni. Enélkül viszont a globális információs társadalom megteremtése sem lehetséges. Végül tehát a szegénység általános, az élet minden fontos területére kiterjedő, világméretű fölszámolása szükséges.

[4] Népszabadság, 2002. június 15.

Hírek

A Siemens hozzájárul ahhoz a programhoz, amelynek célja több mint 30 ezer németországi iskola gyors széles sávú internetelérése. Ennek érdekében 15 000 DSL modemet bocsát térítésmentesen a Deutsche Telekom rendelkezésére. A tanárok és tanulók ily módon adatok, képek, filmek és hanganyag multimédia-feldolgozását végezhetik, és gyakorlati oktatási célokra használhatják a világhálót. Az év végéig 75%-ra növekszik a széles sávú internetelléréssel rendelkező német iskolák száma.

❖ ❖ ❖

A Siemens Corporate Research kutatói olyan megoldáson dolgoznak, amely lehetővé teszi a mobiltelefon és a kézi számítógép funkcionalitását kombináló készülékeken weblapok és háromdimenziós hatású hangzás megjelenítését. Ehhez a szerveroldalon a szövegeket és a futó képeket pontosan szinkronizálják a kísérő hangfrekvenciákkal. Az erre szolgáló utasításokat egy SMIL-dokumentum tartalmazza, melyet a felhasználó felhív. Az SMIL (Synchronized Multimedia Integration Language) forrásutasításai vezérelik a multimédia-objektumok (szöveg, kép, hang) visszaadását. Az audio- és videofájlokat freaming technológiával viszik át a végkészülékre. Ez a bevált eljárás az adatfolyamokat „löketenként" viszik át, de egy pufferelési technikával a lejátszás már átvitel közben a következő részletek letöltése alatt folyamatosan lehetséges. Mindehhez 128 kb/s átviteli sebességre van szükség.

Vállalati tervezési módszerek
Kezelhető-e a véletlen?

PÁTI BRIGITTA

Matáv Üzleti Intelligencia és Dokumentációs Igazgatóság
pati.brigitta@ln.matav.hu

Mindannyian szoktunk tervezni. Megtervezzük, hogy mit fogunk ma csinálni, a jövő héten, hová fogunk utazni a nyáron. Ha minden, amit teszünk és történik velünk, az előre eltervezett úton halad, elégedettek vagyunk. Ha azonban valami közbeszól, és kudarcba fullad a megtervezett program, csalódottak vagyunk. Ilyenkor egyik megoldás a rosszkedv, a másik, hogy gyors döntéssel megváltoztatjuk eddigi szándékainkat, és új célt tűzünk ki. Így működik a vállalatoknál is. A vállalat kijelöli céljait, programot dolgoz ki azok elérésére és megpróbálja azokat végrehajtani. Ha közben a körülmények változnak, a vállalat előtt két út áll: bénultan figyelni az eseményeket, vagy gyors döntéssel alkalmazkodni azokhoz.

Bevezetés

Stabil, kiszámítható kapcsolatrendszerekkel rendelkező piacon a szerződések megkötése után a vállalat egyetlen feladata a teljesítés. Ha a vállalat piacát nem fenyegeti támadás, a vevőért sincs versengés, technológiai változás sem borítja fel az elképzeléseket, akkor kevés döntéssel, gyorsan köttetnek az üzletek. Ha a piac változása folytán verseny kezdődik, sokasodnak a környezeti hatások, több szakemberre és egyre több alacsonyabb szintű döntés meghozatalára van szükség.

A növekedés és komplexitás miatt szükséges a tervezés a vállalaton belül. A nagyvállalatok a tervezéssel kívánják a környezeti változásokból adódó bizonytalanságot és ezzel járó kockázatokat csökkenteni, megszüntetni.

A vállalati tervezés fejlődési szakaszai

A XX. század elejéig a vállalatoknál az informális tervezés volt a jellemző. Ezt követően a formális tervezés (a tudományos tételek eredményeinek felhasználása, folyamatszabályozás és az eredmények írásban történő rögzítése) tapasztalati alapokról indult. Első megnyilvánulása a pénzügyi tervezés a század közepén jelent meg, majd beépült az ezt követő tervezési módszerekbe.

A tervezés, a tervezési szervezet a pénzügyi területről fejlődött ki. Kezdetben a terveket a pénzügyi-költségvetési osztály készítette, munkájáért a felső vezetésnek tartozott elszámolással. Később funkciója kibővült, koordinálta a tervezésbe bevont szervezetek munkáját és elvégezte azokat a tervezés-módszertani, fejlesztési és prognosztizálási tevékenységeket, amelyek szüksége-

Tervezés típusa	**Pénzügyi tervezés**	**Hosszú távú tervezés**	**Stratégiai tervezés**	**Stratégiai menedzsment**
Korszaka USA-ban Ny-Eu-ban	Kb. 1955–1960 Kb. 1955–1965	Kb. 1960–1970 Kb. 1965–1975	Kb. 1970–1980 Kb. 1972–1982	Kb. 1982-
A környezeti alkalmazkodás módja	Múltbéli teljesítmény javítása	Gyors növekedés passzív kihasználása	Aktív részvétel a versenyben	A vállalat minden funkciójára kiterjedő aktív alkalmazkodás, proaktivitás
Terv időhorizontja	1-2 év	5-10 év	2-5 év	Rugalmasan változó
Alkalmazott módszerek	Pénzügyi mutatók, kalkulációk (bázisszemlélet)	Trend, prognózis, pénzügyi terv, lineáris programozás, megtérülési mutatók	Prognózisok alapján célmeghatározás, ehhez eszközök rendelése	A menedzsment eszköztára

1. táblázat A tervezés fejlődési szakaszai [7]

sek voltak. A tervezők voltak felelősek a végrehajtásért és ellenőrizték a megvalósulás menetét.

Az egyes szakaszok természetesen átfedik egymást, a tervezési módszerek, a szemlélet változtatása időbe telik, tehát a fejlődés iteratív folyamatként fogható fel.

A stratégiai tervezés

Az USA-ban a 60-as évek végétől, Nyugat-Európában a 70-es évek közepétől lassult a gazdasági növekedés, és a túlzott termelőkapacitás sok termék piacán megindította a vevők kegyeiért a versenyt. A vállalat eddigi passzivitását a különféle tervezett és nem tervezett befolyások és a megváltozott környezet elkezdett élni, és ezáltal a vállalat és környezete közötti kapcsolat egyirányúból (vállalat → környezet) kétirányúvá, kölcsönössé változott.

A változások felismerése a vállalatok első, azok értékelése a második és az arra való reagálás a harmadik számú feladata. Számos nagyvállalat lassan, sokszor többéves késéssel ismerte fel a környezet befolyásoló erejét. A megkésett vagy téves reagálás könnyen a vállalat életébe kerülhet.

Az alkalmazkodás egyik tudományos megoldása a stratégiai tervezés, mint alkalmazkodási módszer. Segítségével meg lehet határozni, hogy a vállalat milyen akciókkal alkalmazkodjon a környezetéhez.

A stratégiai tervezés főbb jellemzői:

- Stratégiailag fontos területek kiemelt kezelése
- A célok megvalósításához szükséges eszközök meghatározása
- A tervezés bázisa a környezet jövőbeli alakulásának prognózisa
- A tervek időhorizontja a vállalati tevékenységekhez illeszkedő meghatározása
- A tervezésben a vállalat minden érintett területe részt vesz, a tervezési szervezet koordinál
- A terveket a felső vezetés hagyja jóvá
- Rendszerszemlélet: műszaki, pénzügyi, emberi, piaci tényezők együttes kezelése.
- Fókusza középtávra (2-5 év) irányul

A stratégiai tervezés általános modellje:

Az átfogó vállalati stratégián túl *részstratégiák* is készülnek egy vállalatnál, úgymint:

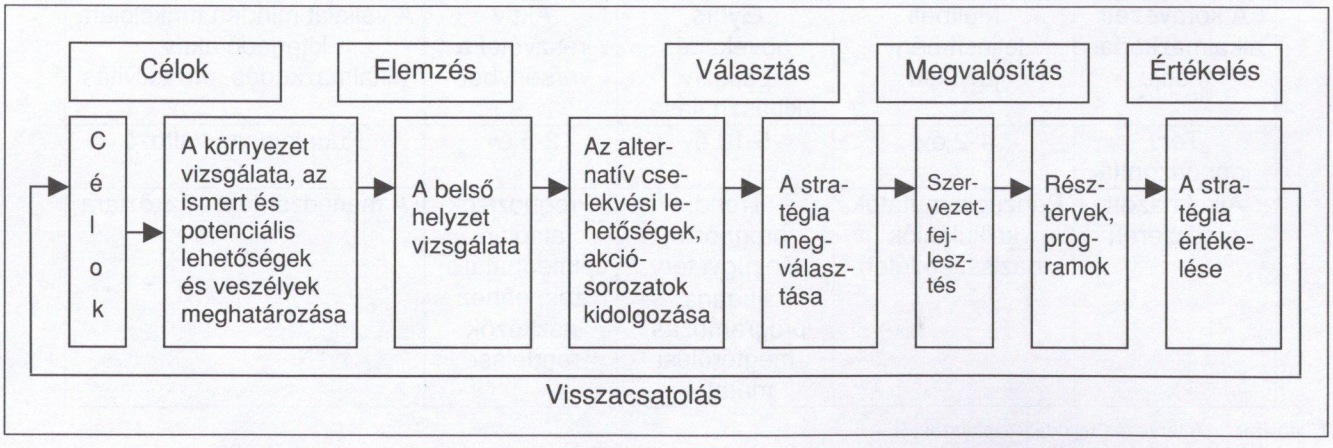
1. *Marketingstratégia*
2. *Termékstratégia*
   A termékstratégia kialakításánál meg kell találni az egyensúlyt a fogyasztói igények és a lehetőségek között. A stratégiaalkotás lépései: termékfejlesztés, termék-életciklusmenedzsment és termékkivonás.
3. *Értékesítési stratégia*
   Az értékesítési tervezés része az igény felkeltése, fogadása, az ösztönzőrendszer előkészítése és az értékesítést támogató IT koordináció.
4. *Hálózatfejlesztési stratégia*
   Speciális stratégia, például a távközlési szolgáltatások esetén kap kiemelkedő szerepet.
5. *Emberi erőforrás stratégia*
   A vállalat sikere nagyrészt a munkavállalókon múlik, a másik oldalról viszont igen jelentős ennek költségvonzata, ami a vállalat nyereségességét rontja.
6. *Finanszírozási stratégia*
   A beruházásokat, piaci terjeszkedést nem tudja kizárólag saját tőkéből finanszírozni a vállalat.
7. *Kockázatelemzés, a konkurencia játékelméleti vizsgálata.*

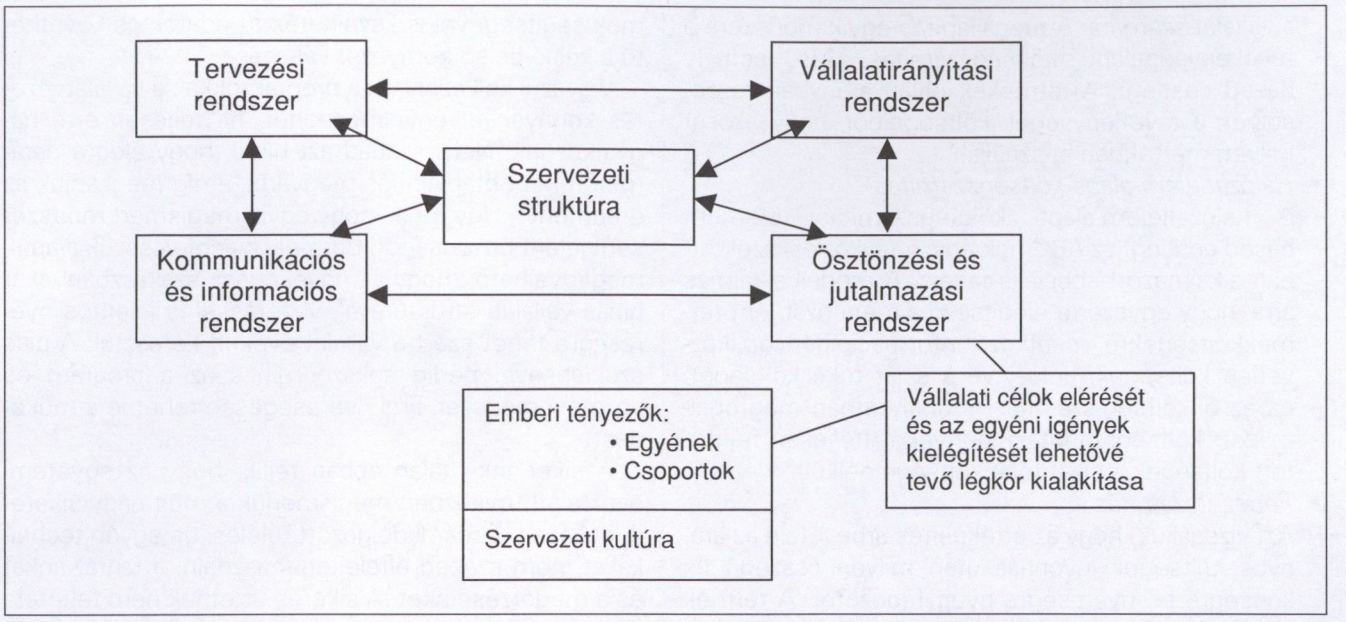
A stratégiai tervezés során a vállalat felvázolhat több irányvonalat a jövőre vonatkozó működésében. A stratégiai tervezés – optimális esetben – egyfajta szimulációs program. Probléma az elmélet megvalósításában, hogy az egyes változatok kidolgozásához a piac ritkán hagy időt.

A stratégiai tervezés módszerének egyik hiánya, hogy a méret növekedésével a részstratégiák nem mutatnak egy irányba, rosszabb esetben nincsenek összhangban a vállalati stratégiával. A szakemberek megpróbálták kiküszöbölni ezt a stratégiai menedzsment módszerének megalkotásával.

Következő lépcsőfok: a stratégiai menedzsement

A *stratégiai menedzsment* egységes folyamatba foglalja a stratégiai tervezést és vezetést, újfajta kapcsolat





1. ábra A stratégiai menedzsment tényezői [11]

jön létre a stratégiai tervezés és az operatív döntéshozatal között, mely kapcsolat a vezetés tudatos munkájának eredménye.

A *stratégiai menedzsment* a stratégiai tervezés jó tulajdonságainak továbbfejlesztésével egyidejűleg annak hátrányaitól való megszabadulást hirdeti.

Hax és *Majluf* szerint a stratégiai menedzsment a vállalati döntéshozatali rendszer, a struktúra és a kultúra integrációját jelenti.

Az eszközök a szervezeti kultúra kivételével már ismertek voltak és egymástól elkülönülten működtek a vállalatoknál. A stratégiai menedzsment egyesíti a hat eszközt és rávilágít arra, hogy az eddig mindenhatónak hitt *tervezés* csupán egy része a *környezethez való alkalmazkodásnak*. Fontos és a sikerhez nélkülözhetetlen a tényezők közötti összhang folyamatos biztosítása. Valamely tényező változását a többi ötnek is követnie kell, megfelelő működés esetén az eszköztár a folyamatos változás állapotában van, vagyis a rendszerszemlélet valósággá válik.

Tervezést segítő technikák és módszerek

A stratégiai tervezést segítő módszereket két csoportba sorolhatjuk. Az első csoportba a célkitűzést, elemzést és stratégia kiválasztását segítő módszerek tartoznak, a másodikba a megvalósítást és értékelést támogató módszerek.

Célkitűzést, elemzést, stratégia kiválasztását segítő módszerek

- *SWOT analízis*
  A SWOT elemzés a vállalati erős és gyenge pontok (belső környezet) elemzését és a környezeti lehetőségek és veszélyek felmérését (külső környezet) jelenti. Egy gondolkodáselemzési rendszer, amely sokféle részmódszert foglal magába, a vállalat és környezete összefüggéseit vizsgálja és a harmónia megteremtését segíti.
- *Portfólió módszerek*
  Gyakran használt módszer a BCG mátrix, mely alkalmas termékek, termékcsoportok, szolgáltatások és vállalati egységek helyzetének elemzésére.
  A módszer továbbfejlesztett változata a *GE-McKinsey mátrix,* ahol a két fő vizsgálati tényező a környezeti lehetőségek és a versenyhelyzet. Ezeket olyan elemekkel írhatja le a vállalat, amelyek helyzetére a legjellemzőbbek, legfontosabbak. Ezáltal az elemzés testreszabottabb, így hatékonyabb és pontosabb. A módszer érzékenységvizsgálatra is alkalmas, azonban hibája, hogy még mindig statikus, nem követi a változásokat.
- *Pénzügyi mutatók*
  Számos fajtája létezik, gyakran használt mutatók: pénzügyi arányszámok, duPont modell, százalékos jövedelemkimutatás, cash-flow kimutatás. Általában a különböző arányszámokat használják a leggyakrabban.
- *Termékéletgörbe-elemzés*
  Minden terméknek ábrázolható a piaci életútja az idő függvényében. Általánosságban négy szakaszra bontható egy termék élete: bevezetés, növekedés, érettség és hanyatlás. Az egyes életszakaszoknál eltérő tervezési megfontolásokat ajánlatos alkalmazni.

Megvalósítást és értékelést támogató módszerek

- *Tevékenységalapú költségszámítás*
  A termékstratégia kialakításakor kell dönteni a termékek bevezetéséről és kivonásáról. A megfelelő döntés meghozatalához alapvető fontosságú annak ismerete, hogy mennyibe kerül az adott termék

a vállalat számára. A megállapítás egyik módszere a tevékenységalapú költségszámítás (ABC=activity based costing). A termékek olyan arányban részesülnek a tevékenységek költségeiből, hogy azokat milyen mértékben használják.

- *Hálózatielem-alapú költségszámítás*
  A hálózatielem-alapú költségszámítás (Element based costing) az ABC-nek egy, a távközlési szektorban alkalmazott speciális fajtája. A modell alkalmas arra, hogy egyszerre elégítse ki a menedzsment termékköltségekre vonatkozó információigényét (közvetlen költségek, beleértve a saját tőke költségét) és az önköltség-számítási szabályzatban megfogalmazott költségszámítási igényeket (teljesen felosztott költségek a saját tőke költsége nélkül).
- *Fedezetszámítás*
  Azt vizsgáljuk, hogy az értékesítés árbevétele az arányos költségek levonása után milyen összegű fix költségre és nyereségre nyújt fedezetet. A termék egy egységére eső fajlagos fedezet rangsorolásával megállapítható a jövedelmezőségi rangsor, kimutatható, hogy mely termékekből származik a vállalat árbevételének jelentős része.
- *BSC elemzés*
  A balanced scorecard, kiegyensúlyozott stratégiai mutatószám-rendszer a stratégiai célok és a hozzájuk rendelhető gyakorlatban alkalmazható mérőszámok módszere.

A tervezési tevékenység informatikai támogatása

A tudományos alapokon születő vállalati alkalmazkodási módszerek egyre több tényezőt ölelnek fel. A kutatók a környezet bizonytalanságának csökkentését oly módon kívánják elérni, hogy minél több, meghatározónak vélt tényezőt vonnak be a modellbe. Az egyre összetettebb módszerek születése a számítástechnika fejlődésével vált lehetségessé. A valóban hatékony, rugalmas, sok tényezőt kezelni tudó vállalati tervezési tevékenység elképzelhetetlen a számítógép használata nélkül. A vállalatok számára készült tervezési programok segítségével is kezelhetők az adatok és követhető a külső-belső környezet változása.

Vigyázni kell azonban a programokkal, a vállalatvezetés könnyen elkényelmesedhet, ha teljesen erre hagyatkozunk. Nem szabad azt hinni, hogy elég a paramétereket betáplálni és megvárni, amíg megkapjuk az eredményt. Úgy tűnik, hogy egyik megismert módszer sem jelent biztonságot. Bármely piacot vesszük alapul, megfigyelhető, hogy a növekedési szakaszban akár hibás vállalati struktúrával, vezetéssel is jelentős nyereségre tehet szert a vállalat éveken keresztül. A piac szűkülésével pedig sokszor nincs az a program és vezetési módszer, ami nyereségessé tehetné a működést.

A siker titka talán abban rejlik, hogy az egyetemi évek alatt, miközben megismerjük a múlt nagy sikereit, bukásait, a már kidolgozott túlélési és egyéb technikákat, nem szabad elfelejteni használni a fantáziánkat és a megérzéseinket. A sikeres üzletnek nem feltétele a közgazdászi végzettség, a tapasztalatokat felhasználó számítógépes eljárás, az csupán könnyítheti az utat. Jó esetben.

Irodalomjegyzék

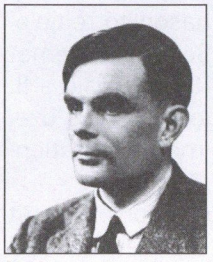
1. Ladó László: Szervezéselmélet és módszertan, KJK Budapest, 1986.
2. Horváth László: Stratégia–vállalati tervezés a gyakorlatban, KJK Budapest, 1987
3. Mohai György: A vállalati tervezéstől a stratégiai módszerekig, KJK Budapest, 1989
4. Horváth László–Csáth Magdolna Stratégiai tervezés Elmélet és gyakorlat KJK 1983
5. Hax, A.–Majluf, N.: Strategic Management: An Integrative Perspective, Prentice-Hall, Inc., New Jersey 1984. 75.oldal
6. Barakonyi Károly–Peter Lorange: Stratégiai management, KJK Budapest 1993.
7. Kubitsch Róbert: Tevékenységalapú költségszámítás (ABC), Magyar Távközlés 2001/1 szám, 25–27. oldal
8. Kubitsch Róbert: Hálózatielem-alapú költségszámítás (EBC), Magyar Távközlés 2002/1 szám, 37–40. oldal

A. M. Turing születésének 90. évfordulójára emlékezve

DÉNES TAMÁS

matematikus

E-mail: titoktan@freemail.hu



Alan Mathison Turing kilencven éve született a londoni Paddingtonban 1912. június 23-án. Apja, Julius Mathison Turing az Indian Civil Service (Indiai Polgári Szolgálat) tagjaként sokat tartózkodott külföldön. Anyja, Ethel Sara Stoney a Madras vasutak főmérnökének lánya volt, így érthető, hogy Alan szülei Indiában találkoztak és házasodtak össze, de Angliában éltek. Tehetsége hamar megmutatkozott, ezért a magas színvonalú Hazlehurst Preparatory Schoolba íratták, ahol több tantárgyból elnyerte a kiváló tanuló címet. Még az iskolában elkezdett érdeklődni a sakk iránt, ami aztán végigkísérte egész életét. 1926-ban beiratkozott egy patinás középiskolába. Ez az év az általános sztrájk éve volt Angliában, így Turing rendszeresen tett meg 60 mérföldet kerékpáron, otthona és az iskola között. Talán ennek is köszönheti, hogy később hosszútávfutásban ért el kitűnő eredményeket.

Turing tehetsége korán kirajzolódott, és annak ellenére, hogy tanárai megrótták kézírásáért, angoljáért és azért is, mert a matematikai problémák megoldásában saját egyéni megoldásokat választott, hamarosan saját útját kezdte járni. Elnyert minden lehetséges matematikai díjat, amit a Sherborne Schoolban lehetett.

Turing már iskolai évei alatt elmélyült a matematikában, ezért tanárai abban támogatták, hogy önmagát képezze, saját elképzelései szerint haladjon. 1928-ban szoros barátságba került egyik iskolatársával, Christopher Morcommal, akivel meg tudta beszélni tudományos gondolatait. Ez emberileg is sokat jelentett számára, ezért viselte meg Morcom betegsége, majd 1930-ban bekövetkező halála.

1931-ben beiratkozott a cambridge-i King's College matematika szakára. Egy olyan „renitens" gondolkodású embernek, mint Turing, sok szempontból Cambridge könnyebbnek bizonyult az addigi iskoláknál. Itt kifejthette gondolatait, már 1933-ban Russellt (Introduction to mathematical philosophy), valamint Neumann János írásait olvasta a kvantummechanika matematikájáról.

1933-ban került hatalomra Hitler Németországban, és ekkor indult meg a háborúellenes mozgalom Angliában, melynek Turing is tagja lett. 1934–35-ben hallgatta Max Newman (1897–1984) előadásait a matematika alapjairól. Ezen a kurzuson foglalkoztak Gödel nem teljességi elméletével és Hilbert eldönthetőségi problémájával *(Entscheidungsproblem)*. Hilbert felvetette, hogy lehetséges-e általános algoritmust adni a matematikai problémák megoldására, vagy egyáltalán létezhet-e ilyen algoritmus (eljárás)? Igazán nagy jelentőségű lépést e kérdés megválaszolásában egy osztrák matematikus, Kurt Gödel 1931-ben bebizonyított tétele jelentett, mely szerint minden axiómarendszerben megfogalmazható olyan állítás, amely az axiómarendszer segítségével se nem bizonyítható, se nem cáfolható.

Első látásra az „eldönthetőség" egy egyszerű kérdés, hiszen egy adott matematikai állításhoz kell találni egy algoritmust, amelynek segítségével eldönthető, hogy az állítás igaz, vagy hamis. Ekkor kezdett Turing foglalkozni az algoritmus (eljárás) pontos meghatározásával. 1937-ben megjelent cikke [1] nagy vihart kavart. Ebben a cikkében vezette be az *absztrakt gép fogalmát,* amelyet máig is *Turing-gép*nek neveznek. A Turing-gép tulajdonképpen egy „darab absztrakt matematika", és bár elnevezése erre utal, nem technikai eszköz. A Turing-gép, mint minden igazán zseniális elképzelés, könnyen leírható:

Képzeljünk el egy olyan automatát, amely véges sok szimbólumot (jelet) tud feldolgozni úgy, hogy egy adott időpillanatban egyetlen szimbólumot képes leolvasni, vagy felírni egy elvileg végtelen szalagra. A feldolgozást egy speciális jel, a STOP feldolgozásakor fejezi be.

Ebben az absztrakt definícióban valóban benne van a jelek hosszabb jelsorozatokká való összeláncolásának és így tetszőleges bonyolultságú utasítások létrehozásának és tárolásának, a végrehajtás közben keletkezett jelek (adatok) tárolásának lehetősége. A Turing-gép tehát valóban egy absztrakt automata, amelyre teljesül az a meghökkentő tétel, melyet Alonzo Church (1903–1995) amerikai matematikus 1936-ban állított fel, és amely szerint minden programhoz található egy azzal ekvivalens Turing-gép és fordítva, minden Turing-gép egy programot (algoritmust) valósít meg, azaz a Turing-gép tökéletes modellje a program fogalomnak.

Ugyanekkor Turing definiálta az úgynevezett *„kiszámítható számot"*, mint olyan valós számot, amelynek tizedes jegyei előállíthatók egy Turing-géppel. Megmutatta, hogy a π kiszámítható, de csak megszámlálhatóan sok ilyen valós szám van, ennél „jóval több" valós szám nem kiszámítható. Turing pontosan leírta, hogy milyen egy nem kiszámítható szám és megjegyezte, hogy az paradoxonhoz vezet, ha véges sok jellel le tudja írni azt a számot, ami „nem írható le" véges számú jellel.

Magát a paradoxont már 1905-ben felvetette Jules Antoine Richard (1862–1956) francia matematikus, de egzakt tárgyalása Turing munkásságához kötődik, aki a már említett cikkében összefoglalta a matematika és számítástudomány számára alapvetően fontos tételnek a bizonyítását. Ez adta az indítékot Alonzo Churchnek, hogy közzétegye az American Journal of Mathematicsban *An unsolvable problem in elementary number theory* című cikkét, amelyben bizonyítja, hogy az aritmetikában nincs olyan eljárás, amellyel ez a probléma eldönthető.

Turing többszörösen módosított 1936-ban írt cikke, amely 1937-ben jelent meg, már tartalmazta Church eredményeinek elemzését. Ez jó ajánlólevél volt ahhoz, hogy 1936-ban felvették a Princeton University munkatársai sorába, ahol a Church vezette kutatásokban vett részt, majd 1938-ban visszatért Angliába. Legjelentősebb publikációja, amely betekintést ad princetoni munkájába, 1939-ben jelent meg. Newman így ír erről a cikkről:

„Ez a cikk tele van figyelemre méltó gondolatokkal és ötletekkel... ezek még világosabbá teszik Turing elképzeléseit a matematikai bizonyítás területéről."

A programozható gépekkel kapcsolatban, szintén a XX. század 30-as éveiben vetődött fel a kérdés, hogy létezik-e (létezhet-e) olyan programozási feladat, amely nem oldható meg, azaz a Church-tézis szerint létezik-e olyan programozási feladat, amelyhez nem található Turing-gép? 1937-ben A. M. Turing bebizonyította, hogy a válasz „igen", mivel azok, és csak azok az algoritmusok programozhatók, melyekhez úgynevezett rekurzív függvények tartoznak. A matematikának azt a területét, amely eme kérdések tárgyalását tűzte ki céljául, kiszámíthatóságelméletnek, algoritmuselméletnek, illetve Turing előbbi tétele szerint a rekurzív függvények elméletének nevezzük. Ezek az elméleti területek leegyszerűsítve a következő kérdéssel foglalkoznak: *Melyek azok a számítások, amiket a számítógép el tud végezni, ha minden gyakorlati jellegű korláttól eltekintünk (mint például a rendelkezésre álló idő és tárkapacitás)?* Munkásságának legnagyobb jelentősége, hogy megelőzte korát, hiszen leírta a modern számítógép lényegét jóval azelőtt, hogy annak technikai feltételei ebben az időben adottak lettek volna.

Amíg Turing Princetonban volt, játszott a gondolattal, hogy tervez egy működő számítógépet. 1938-ban, mikor visszatért Cambridge-be, valóban elkezdett építeni egy analóg mechanikus berendezést a Riemann-hipotézis tanulmányozására (ez a mai napig a matematika egyik leghíresebb megoldatlan problémája). A Riemann-hipotézis a prímszámok számával, illetve azok eloszlásával kapcsolatos, amely problémakör nem csupán a matematikusok fantáziáját mozgatta meg, hanem a hipotézis empirikus ellenőrzése rendkívüli számítási kapacitásokat igényel. Nem véletlen tehát, hogy sokszor a prímszámokkal kapcsolatos problémák inspirálták a számítástechnika fejlődését. [5]

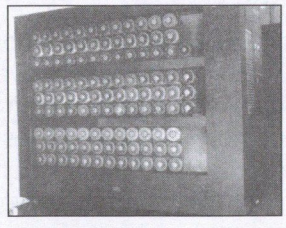
Nem sokkal ezután Turing tevékenysége egészen új fordulatot vett, amikor kapcsolatba került a Government Code and Cypher Schoollal (az angol titkosszolgálat rejtjelfejtő szolgálata), akik felkérték, hogy segítsen a német Enigma rejtjelező rendszer feltörésében.

Az Enigma egy külsőre írógéphez hasonlító rejtjelező gép volt [6, 7, 8], melyet Arthur Scherbius német elektromérnök szabadalmaztatott 1918-ban, majd a II. világháborúban a német vezérkar szigorúan titkos üzeneteinek rejtjelezésére használták, megfejthetetlennek tartották.

Az Enigma megfejtésére irányuló angol projektet ULTRA-nak nevezték és központja a fenti elnevezésű „iskola" volt.

Az ULTRA-ban részt vett három lengyel rejtjelfejtő (Rejewski, Rozycki és Zygalski), akik 1938-ban a német megszállás elől Franciaországba menekültek, majd Angliába települve részt vettek az ULTRA-ban. Amikor kitört a II. világháború, Turing azonnal teljes idejét a Bletchley Parknak szentelte, ahol az angol titkosszolgálat rejtjelfejtő központja működött, briliáns ötletei a kódok megfejtésében és az elektronikus számítógép kifejlesztésében nagyban hozzájárultak az Enigma rendszer feltöréséhez. Paradox módon, eme időszak nagyon boldog periódusa volt életének, melyről Newman így ír: *„Talán életének legboldogabb időszaka volt, amelyben legjobban kamatoztathatta sziporkázó kreativitását."*

W. G. Welchmannal együtt, a „lengyel csoport" korábbi munkájára alapozva megtervezte azt a rejtjelező gépet, amely 1940 végétől dekódolta az összes üzenetet, amelyet a német légierő az Enigmával rejtjelezve küldött. Ezt a berendezést Bombe-nak nevezték el (1. ábra).



1. ábra A Bombe rejtjelfejtő gép

Ezt a Bletchley Park nagy sikereként tartja számon a történelem. Kevesebb szó esik azonban arról a Turing által vezetett másik tevékenységről, amely a világ első elektronikus számítógépének, a Colossusnak a létrehozása volt. A szűkszavúság annak köszönhető, hogy a Colossus megtervezése, létrehozása és üzemeltetése szigorúan titkos volt. Olyannyira, hogy csupán néhány eredeti fotó készült róla (ezek egyike látható a 2. ábrán, forrás [8]).

A háború befejeztével meghívták a londoni National Physical Laboratoryba számítógépet tervezni. Erre vonatkozó terv javaslatát *Automatic Computing Engine* (ACE) címmel 1946 márciusában adta le. A tervek gyors elkészítése a Colossus megépítésének tapasztalatai után már érthető. Nem csoda tehát, hogy Turing eredeti, részletekbe menő és számos rajzmelléklettel gazdagon illusztrált dokumentációja egy igazi modern számítógépet mutatott be. Ugyanakkor az ACE tervezett tároló (memória-) méreteit túlzottnak tartották a döntéshozók, így ez a projekt megvalósítását késleltette.

Turing 1947-ben visszatért Cambridge-be, ahol a számítógépektől és a matematikától látszólag távoli területeket kezdett tanulmányozni, mint a neurológia és pszichológia. De nem feledkezett el a számítógépekről ezalatt sem, mivel számítógépes programozási „kódokat" készített.

Turing a háború után komolyan foglalkozott a tudományon kívüli világgal is. Tagja volt a Walton Athletic Clubnak és rekordidővel nyerte a 3 és 10 mérföldes futóbajnokságot. 1947-ben maratoni futásban is negyedik helyezett lett.

1948-ban Newman volt a University of Manchester matematikaprofesszora, így felajánlott egy állást Turingnak, amit ő el is fogadott. Newman így ír erről:

„A számítógépek konstrukciója már megkezdődött F. C. Williams és T. Kilburn tevékenységével. Az elképzelés az volt, hogy Turing átveszi a munkák matematikai vezetését és ezt egy pár éven át folytatja. Először elkészítették egy ilyen géphez szükséges óriási program szubrutinjait, majd kialakították e munkák egységes rendszerét és kiterjesztették a számítási eljárások általánosabb problémáira is."

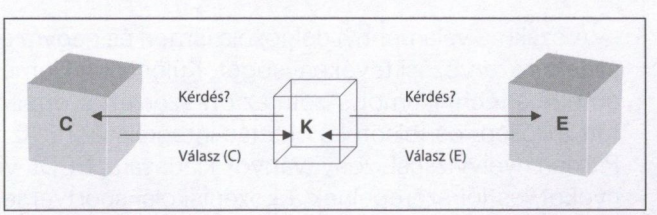
1950-ben publikálta Turing *Computing machinery and intelligence* című cikkét [9], amelyet életének kiemelkedő teljesítményeként tart számon a mai tudomány. Ebben a cikkében írta le az általa javasolt Turing-tesztet, amely a mesterséges és természetes intelligencia megkülönböztetését helyezte egzakt alapokra. Itt vetette fel elsőként azt a kérdést, hogy mit is jelent a „gépi intelligencia". Az első megválaszolásra váró kérdés persze az volt, hogy létezik-e ilyen, hiszen máig fennáll az a többségi felfogás, miszerint intelligenciával csupán az ember rendelkezik, ezért a „gépi intelligencia" szóösszetétel értelmetlen. Turing azt is látta, hogy az intelligencia és gondolkodás fogalmak egymástól elválaszthatatlanok, ezért fogalmazta meg 1950-ben megjelent, klasszikussá vált cikkében egyetlen mondatba sűrített kérdését: *„... tudnak-e a gépek gondolkodni?"* Turing szerint a „gondolkodni" szó inkább érzelmi kérdéssé teszi e kérdéskört, ezért el is veti, mint túlságosan bizonytalan (szubjektív) fogalmat. Ugyanakkor az 1950-es években sokan úgy gondolták, hogy Kurt Gödel nem teljességi tétele a mesterséges intelligencia lehetetlenségét is bizonyítja:

A mesterséges intelligencia mindig „egy program", azaz egy Turing-gép (Church-tézis). Az ebben a gépben tárolt axiómarendszer meghatároz egy „nyelvet", amely nyelven megfogalmazható olyan kérdés, amelyre ebben az axiómarendszerben nem vezethető le igennem jellegű válasz (Gödel-tétel). Tehát e mesterséges intelligencia számára érthető nyelven megfogalmazható olyan kérdés, amelyre nem tud sem igennel, sem nemmel válaszolni! Ezen érvelés legfőbb hibája, hogy ha a mesterséges intelligenciát mint az emberi intelligenciát utánzó konstrukciót fogjuk fel, akkor ennek megvalósíthatatlanságát nem bizonyítja az az érv, hogy bizonyos kérdésekre nem tud egyértelműen felelni, hiszen ez az emberi gondolkodásnak is jellemzője.

A rekurzív függvények elméletének, a matematikai nyelvészetnek jelentős alakja, a magyarországi kibernetikai iskola megalapítója, Kalmár László (1905–1976) az 1948-as amszterdami filozófiai kongresszuson tartott előadásában bebizonyította, hogy a Church-tétel a Gödel-tételből levezethető, így Church tétele nem bizonyíthatja abszolút eldönthetetlen probléma létezését. Kalmár László hangsúlyozta, hogy ezeket a tételeket (Gödel, Church) szabatosan úgy kellene megfogalmazni, hogy a kérdéses problémasereg általános rekurzív eljárással nem oldható meg, nem pedig abszolút megoldhatatlan [10, 11].

Turingot az ellenvetések és főleg a „gépi intelligencia" fogalmának bizonytalansága csak inspirálta egy új megközelítés felvetésére. Ennek lényege, hogy e szubjektív és ezáltal tudományosan megfoghatatlan fogalmak helyett egy olyan módszert kell konstruálni, amelyet jól definiált technikai fogalmakkal lehet leírni. Javaslata szerint ez az általa „utánzási játéknak" nevezett módszer, melyet manapság Turing-teszt, vagy Turing-próba néven ismerünk:

Képzeljük el, hogy egy C számítógép és egy E ember két külön helyiségben van elkülönítve és mindketten elektronikus kapcsolatban vannak egy harmadik helyiségben levő K személlyel, aki elektronikus úton kérdéseket tehet fel mindkettejüknek. K-nak az a célja, hogy a kérdéseire érkező válaszokból meg tudja különböztetni, hogy mely válasz származik C-től és melyik E-től (2. ábra).



2. ábra A Turing-teszt vázlata

A mesterséges intelligencia kutatások célkitűzése tehát a gépek alkalmassá tétele arra, hogy az embert minél pontosabban tudjuk utánozni. Turing eme korszakos cikkében kifejezte meggyőződését, hogy a XX. század végére a gépek már elég jól fogják játszani ezt a játékot ahhoz, hogy egy átlagos kérdezőnek nem lesz 70%-nál több esélye az azonosításra 5 percnyi kérdezés után.

Turingot a Royal Society of London 1951-ben tagjává választotta, alapvetően a Turing-gépekre vonatkozó munkásságának elismeréseként.

1951–52-ben a matematika biológiai alkalmazásaival, az élő organizmusok modellezésének kutatásával foglalkozott.

Turing 1954. június 7-én kálium-cianidmérgezésben halt meg, amikor éppen egy elektrolízis-kísérleten dolgozott. A mérget egy félig elfogyasztott almában találták meg mellette. A nyomozás önkezűséget állapított meg, de édesanyja meg volt győződve arról, hogy baleset történt.

Irodalom

1. A. M. Turing: On computable numbers, with an application to the Entscheidungsproblem Proc. Lond. Math. Soc., 1937. (ser. 2) 42, 230–265.
2. Dénes József: Adalékok a párhuzamos architektúrájú számítógépek történetéhez. Híradástechnika, 2002/5.
3. Dénes Tamás: TITOK TAN avagy Kódtörő ABC KRIptográfia MIndenkinek. Bagolyvár Könyvkiadó, Budapest, 2002.
4. A. M. Turing: Systems of logic based on ordinals Proc. Lond. Math. Soc., 1939. 45, 161–228.
5. Dénes Tamás: Turing-teszt az információs társadalomban (e-világi gondolatok). Megjelenés alatt.
6. A. Hodges: Alan Turing: The Enigma. Burner Books Ltd., London, 1983.
7. F. W. Winterbotham: The Ultra Secret. Futura Publications Limited London, 1975. Magyarul: Az Ultra titka. OMIKK, Budapest, 1996.
8. T. Enever: Britain's Best Kept Secret Ultra's base at Bletchley Park Alan Sutton Publishing Limited, Dover, 1994.
9. A. M. Turing: Computing Machinery and Intelligence Mind, 9 (1950), 433–460.
10. Kalmár László: Egyszerű példa eldönthetetlen aritmetikai problémára. Mat. és Fiz. Lapok, 50, 1943. 1–23.
11. Kalmár László: Integrállevél. Gondolat, Budapest, 1986.

Gratulálunk

Dr. Horváth László, a Puskás Tivadar Távközlési Technikum igazgatója augusztus 16-án Kovács Kálmán informatikai és hírközlési minisztertől átvette a

Magyar Köztársasági Arany Érdemkeresztet.

A szakma valamennyi dolgozója ismeri és nagyra értékeli széles körű szakmai, és az ehhez kapcsolódó társadalmi szervezési tevékenységét. Különösen kiemelkedik az utolsó tíz év, amikor igazgatóként régi rangjára emelte a technikumot. Azóta az ott szerzett érettségit az egyetemi felvételek során nagyra értékelik. Sokat tett a közép- és felsőfokú oktatás integrálásáért. Az iskola jogosult Euro-minősítésű technikusi oklevelek és Pitman nyelvvizsgabizonyítványok kiadására, ECDL vizsgáztatásra. A tanulók sorra nyerik a tanulmányi versenyeket, és jól szerepelnek a középiskolai sportversenyeken. A „Puskás" az Arany János Tehetséggondozó Programhoz az ország egyetlen szakközépiskolájaként csatlakozott.

Mind „Laci bácsinak", mind az iskolának további szép sikereket kívánunk.

Pályázati felhívás

A Huszty Dénes Alapítvány Kuratóriuma
az alapító okiratában foglaltak szerint pályázati felhívást tesz közzé

Huszty Dénes (1927–1979) gépészmérnök, akusztikus. A II. világháború utáni Magyarországon az elektroakusztika egyik legjelentősebb szakmai vezetője, kutatója, szerzője, és nemzetközileg is nagyra becsült szaktekintélye volt. Egyaránt maradandót alkotott a tudományos kutatásban, a műszaki fejlesztésben, a szakmához kapcsolódó gyártástechnológiában és gyártásban, a hírközlésben és stúdiótechnikában, a nemzetközi és hazai szabványosításban, az elektroakusztika menedzselésében és üzletpolitikájában. Az egykori „új gazdasági mechanizmus" adta lehetőségek között munkatársait már 1968–69-ben a mai igényeknek megfelelő piaci és minőségpolitikai szemlélettel igyekezett a várható fejlődésre felkészíteni. Az akkori gazdasági korlátok miatt tanítványai a nagyra törő és a hazai jólétet megalapozó iparszervezési elképzeléseit csak mintegy 20 év múlva kezdhették el megvalósítani.

Váratlan halála miatt 1979-ben megszakadt az a 30 éven át tartó gazdag szakmai tevékenység, amely az ORION Rádió és Villamossági Vállalattal, a VIDEOTON elődjével (Vadásztölténygyár), az Elektroakusztikai Gyárral (BEAG), a Magyar Rádióval, a MTA Akusztikai Kutatólaboratóriummal, és a Bolgár elektroakusztikai iparral kapcsolódott össze. Huszty Dénes minden fent említett ipari helyen a hangszórók és hangsugárzók tömeggyártását megalapozó technológiát igyekezett teremteni. Alapvetően az ő munkásságának eredményeként a 70-es évek elején hazánkban évente közel 1,3 millió darab hangszórót gyártottak. Ezekből akkoriban százezres darabszámban exportálhattunk az USA-ba is. A hangátvitel és az ahhoz kapcsolódó technológiai területeken közel 100 szabadalom bizonyítja mérnöki tehetségét. A nemzetközi szabványosításban ma is több élő, bevált kezdeményezését sikerült elfogadtatni. A stúdiótechnikai ún. OIRT szabványosításnak elfogadott és elismert szaktekintélye volt.

A tudományos kutatásban iskolaalapító intenzív munkásságát 68 közlemény bizonyítja. Az MTA Akusztikai Komplex Bizottságban, az OPAKFI-ban, a Híradástechnikai Tudományos Egyesületben, az Audio Engineering Society-ben aktív szakmai – társadalmi tevékenységet fejtett ki. Igyekezett a fiatalok számára is vonzóvá tenni a számára életpályát jelentő akusztikus szakmát. Ezt a munkásságát a Petzval József díj és a Békésy Díj fémjelzi.

A további részleteket illetően lásd. a Kép és Hangtechnika XXV. (1979) célszámát.

A pályázat részletezése

1. A 2001 novemberében bejegyzett Huszty Dénes közhasznú alapítvány célja, hogy az akusztika vagy az elektroakusztika tématerületén tevékenykedő fiatal szakemberek, felsőfokú tanulmányaikat éppen befejező diplomatervezők vagy már végzett fiatalok, és a pályázat beadásakor 35. életévüket még be nem töltött fiatal akusztikusok olyan kiemelkedő eredményeit jutalmazza, amelyek hozzájárulnak az akusztika egyetemes fejlődéséhez. Az Alapítványnak további célja hogy emléket állítson Huszty Dénes munkásságának, aki az 1950–1979 közötti időszak kiemelkedő akusztikai szaktekintélye volt.
2. A Huszty Dénes Emlékdíj 2003-ban emlékplakettből és 100 000 Ft pénzjutalomból áll.
3. A Kuratórium a 2003. évben legalább egy díjat ad ki.
4. Az Alapítvány 2003. évi díjaira pályaművel lehet pályázni.
5. Pályázhatnak végzett, elsősorban mérnökök, fizikusok saját önálló munkájuk összefoglaló dolgozatával, szakirányú lapban megjelent cikkeikkel vagy új dolgozattal. A pályázóknak lehet más szakirányú diplomája is, de a pályázó tevékenységét az akusztika területén kell, hogy kifejtse.
6. A jelen 2003. évi pályázat kiemelt témakörei:
   hanginformációk feldolgozása az elektroakusztikában
   a hangminőséggel kapcsolatos kutatások és fejlesztések legújabb eredményei.
7. A pályázatok beérkezésének határideje: 2002. december 10.
8. A pályázatok beadhatók személyesen az Alapítvány székhelyén: a Hírközlési és Informatikai Tudományos Egyesület irodájában, 1055 Budapest, Kossuth tér 6–8., vagy postai úton. A postán feladott küldemények esetén a feladás legkésőbbi időpontja 2002. december 7.
9. A pályázatokat a Kuratórium által felkért értékelő bizottság értékeli és javaslatot tesz a díjazásra a Kuratóriumnak.
10. A Huszty Dénes Emlékdíj ünnepélyes kiosztására 2003. január 21-én, kedden, az Alapítók képviselőinek jelenlétében kerül sor. A díjkiosztás helyszínéről a 11. pont szerint jelenik meg tájékoztató.
11. A jelen pályázati felhívás a Hírközlési és Informatikai Tudományos Egyesület valamint az Optikai, Akusztikai és Film- és Színháztechnikai Tudományos Egyesület tájékoztató Hírlevelében és a Zajinfóban, valamint az Egyesületek honlapjain jelenik meg.

Budapest, 2002. augusztus hó

A Kuratórium nevében:

Dr. Illényi András sk.
a Kuratórium Elnöke

Contents

Szerkesztőség
HTE Budapest V., Kossuth L. tér 6–8.
Tel.: 353 1027, Fax: 353 0451
e-mail: hte@mtesz.hu

Hirdetési árak:
1/1 (205 x 290 mm) 4C 120 000 Ft + áfa
Borító 3 (205 x 290 mm) 4C 180 000 Ft + áfa
Borító 4 (205 x 290 mm) 4C 240 000 Ft + áfa

Cikkek eljuttathatók az alábbi címre is
BME Szélessávú Hírközlő Rendszerek
Budapest XI., Goldmann Gy. tér 3.
Tel.: 463 1559, Fax: 463 3289
e-mail: zombory@mht.bme.hu

Előfizetés
HTE Budapest V., Kossuth L. tér 6–8.
Tel.: 353 1027, Fax: 353 0451
e-mail: hte@mtesz.hu

2001-ES ELŐFIZETÉSI DÍJAK
Hazai közületi előfizetők részére
1 évre bruttó 30 000 HUF

Hazai egyéni előfizetők részére
1 évre bruttó 6 000 HUF

Subscription rates for foreign subscribers
12 issues 150 USD, single copies 15 USD

www.hte.hu

Felelős kiadó: MÁTÉ MÁRIA
Lapmenedzser: Dankó András

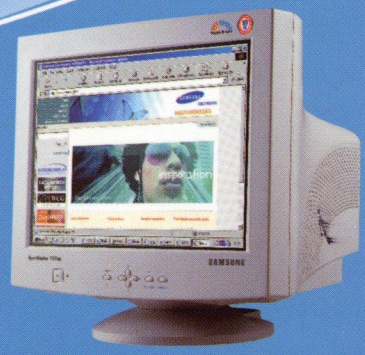
Design by: Kocsis és Szabó Kft.
HU ISSN 0018-2028

Printed by: Regiszter Kft.

SAMSUNG

MagicBright

DigitAll *varázslat*



Szövegszerkesztés →

← *Internet*

Film ↑

Varázslatos változatok.

A SyncMaster MB monitorok MagicBright technológiájukkal gombnyomásra változtatnak a munkastílusodon. Szövegszerkesztés? Internet? Film? A MágikusRagyogás varázsütésre biztosítja a felhasználás módjához leginkább alkalmas fényerőt, kímélve ezzel szemeidet és kihozva a maximumot monitorodból. Három típus, egy árért? Kár lenne kihagyni.

SAMSUNG DIGITall
everyone's invited™

www.samsung.hu