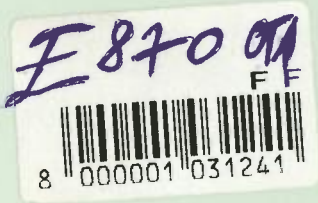


hírtechnika

VOLUME LV.

2000/11



® journal on
communications
computers
convergence
contents
companies

JOURNAL ON C⁵

A PUBLICATION OF THE SCIENTIFIC ASSOCIATION FOR INFOCOMMUNICATIONS, HUNGARY

SPONSORED BY

Főszerkesztő / Editor in chief

SIMONYI ERNŐ

Rovatvezetők / Senior editors

BARTOLITS ISTVÁN
KOSÁRSZKY ANDRÁS
TORMÁSI GYÖRGY
TÓTH LÁSZLÓ
ZSÓTÉR JENŐ

Munkatársak / Assistants

HOLLÓ KATALIN
LESNYIK KATALIN
SELMECZI VILMOS

Szerkesztőbizottság / Board

ZOMBORY LÁSZLÓ
elnök / president
ANTALNÉ ZÁKONYI
MAGDOLNA ügyvezető/
managing director
BOTTKA SÁNDOR
CSAPODI CSABA
DROZDY GYŐZŐ
GORDOS GÉZA
GÖDÖR ÉVA
KAZI KÁROLY
MEGYESI CSABA
PAP LÁSZLÓ
SALLAI GYULA

ALCATEL

ERICSSON

SAGEM

NOKIA

T N

SIEMENS



MOTOROLA



antenna © hungária

PANNON GSM
Az élvonal.



Szerkesztőség / Editorial office

Budapest, VI. Paulay E. u. 56. II.14/A.

Tel/Fax: (361) 341-6421. (361) 470-0713

Előfizetés / Orders to

Holló Katalin

1147 Budapest, Ilosvay Selymes u. 133.

Tel/Fax: (361) 470-0713

2000-ES ELŐFIZETÉSI DÍJAK

Hazai közületi előfizetők részére

1 évre 20000 Ft +12% ÁFA = Btto 22400 Ft

Hazai egyéni előfizetők részére

1 évre 4000 Ft +12% ÁFA = Btto 4480 Ft

Subscription rates for foreign subscribers

12 issues 100 USD, single copies 10 USD

HÍRADÁSTECHNIKA, JOURNAL ON C⁵ is published monthly, in English and in Hungarian

H-1164 Budapest, Csókakő u. 27. Phone: (361) 400-2166, 400-2167, Fax: (361) 400-2168. Printed by Regiszter Kft.

HU ISSN 0018-2020



Associate Editor: Tibor Trón

CONTENTS 1

CommunicationsPongrácz G.: *TCP/IP protokoll vizsgálata és optimalizálása ATM hálózat felett* 2**Computers**T. Nemetz, P. Papp: *Data-compression aiding data-security* 13Trinh Anh Tuan: *On the stability of TCP* 19**Convergence**Cser L.: *Gyorsulásmérők alkalmazási lehetőségeinek kutatása* 24Lencse L.: *Digitális képátviteli rendszerek* 29Urbancsek T., Lassó A.: *Mikromanipulátorok biotechnológiai kutatásokban* 36**Contents & Distribution of Multimedia**M. Hate: *Digital radio – an international broadcaster's perspective* 44Dósa Gy.: *Kettős rombuszantenna rendszer a hazai műsorszórásban* 52**Companies**K. Gyenes: *Intelligent vehicle follower system controlled by satellite* 61

Abstract: In this TDK (Scientific Research and Development of the University Students) special issue, edited by Associate Editor, Dr. Tibor Trón 9 papers are presented. Section **Communications** is represented by a single contribution on TCP/IP performance optimisation. Section **Computers** is comprised of 2 contributions on the Data-Compression Aiding, on the Data-Security Stability of TCP. Section **Convergence** is formed by 3 contributions written on the Applications Feasibility of the Acceleration Sensing, on the Digital Image Transmission Systems, on the Application of a Teleoperated Microrobot for Biotechnological Research. Section **Contents & Distribution of Multimedia** is now representing by 2 contributions on Digital Radio Perspective, on the Dual Rhombus Antenna System for the Hungarian Broadcasting. Section **Companies** is now dealing with a single contribution on the Intelligent Vehicle Follower System Controlled by Satellite from Technical University of Budapest. In this issue 4 senior reviewers scientifically evaluate all of the submitted papers. The reviewers as scientific contributions have accepted these papers. They are marked on their first pages by the sign of \mathcal{L} giving the evidence of the scientific applications result.

A TCP/IP PROTOKOLL TELJESÍTMÉNYÉNEK VIZSGÁLATA ÉS OPTIMALIZÁLÁSA ATM HÁLÓZAT FELETT

PONGRÁCZ GERGELY

BUDAPESTI MŰSZAKI ÉS GAZDASÁGTUDOMÁNYI EGYETEM
TÁVKÖZLÉSI ÉS TELEMATIKAI TANSZÉK
1117 BUDAPEST, PÁZMÁNY PÉTER SÉTÁNY 1/D
Gergely.Pongracz@eth.ericsson.se; pgeri@iq.sch.bme.hu

Napjaink adatkommunikációjának zöme a TCP/IP protokoll segítségével zajlik. Az Internet alapú adatkommunikáció iránti igény rohamos növekedésével azonban szükségessé válik egy mindenhol jelenlévő szélessávú hálózat kiépítése. A jelenlegi trendek szerint ennek a jövőbeni integrált hálózatnak az alapját a ma meglévő távközlési gerinchálózatok fogják adni. Ezen hálózatok egyre inkább elterjedt alatechnológiája az aszinkron átviteli mód (ATM). Munkánk, amely az Ericsson Traffic Lab kutatólaboratóriumában készült, ennek a két technikának az együttműködését vizsgálja. Az eredményekkel segítséget kívánunk nyújtani azoknak, akik adatkommunikációs hálózatokat ATM alapon tervezik kiépíteni, hiszen a két protokoll együttműködése nem zökkenőmentes.

1. BEVEZETŐ

Napjainkban a nyilvános adatátviteli hálózatok gerincét többnyire ATM alapú rendszerek alkotják, és a technológia várhatóan a telefóniában is hamarosan teret nyer. Ezek a rendszerek képesek már jelenlegi állapotukban is különböző karakterisztikájú forgalmi igények egyidejű kielégítésére. Az ATM így lehetőséget ad arra, hogy a napjainkban folyó hálózati konvergencia, azaz a jelenleg külön hálózaton zajló forgalmak egységes hálózaton való kezelése megoldható legyen technikailag.

Az IP protokollt használó adatátviteli eszközök és erre épülő szoftverek széles elterjedtsége megkívánja, hogy bármely jövőbeni alaphálózat képes legyen az IP alapú adatátviteli technikával nagyon jól együttműködni. A jelenlegi trendek azt mutatják, hogy az IP protokoll az adatátvitelben megőrzi kizárólagosságát, valamint esélye van arra, hogy más hálózati szegmensekben is (műsorszórás, hangátvitel) de facto szabvánnyá váljon.

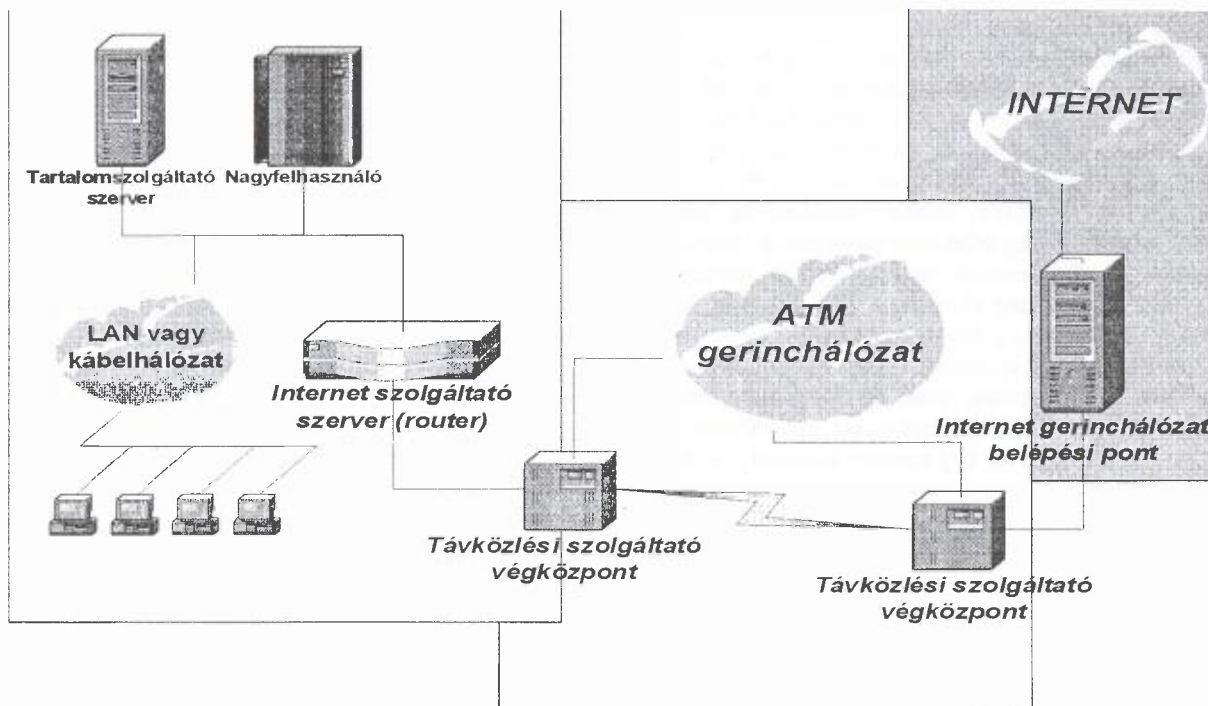
A fentiek értelmében az ATM és az IP együttműködésének kidolgozása kiemelt fontosságú. Az ezirányú kutatások célja egy olyan egységes hálózati kép kialakítása, amely gazdaságilag és technológiailag is optimálisnak mondható. A cél eléréséhez egy olyan integrált forgalommenedzsment architektúra kidolgozása szükséges, mely az ATM-ben már kidolgozott minőségi garanciákat elérhetővé teszi az IP technológia számára.

Ameddig azonban ez nem születik meg, az IP és ATM együttműködését a lehető leghatékonyabbá kell tenni, hiszen a jelen eszközei ezt a problémakört gyakorlati jelentőségűvé teszik. Nem megfelelő beállításokkal az IP proto-

koll nagyon csekély kihasználtságot tud csak elérni ATM hálózatok felett. A mérések célja a harmonikus együttműködéshez szükséges IP paraméterek megtalálása, illetve az IP és ATM paraméterek sávszélesség kihasználtságára gyakorolt hatásának bemutatása.

Vizsgálatunkkal elsősorban egy gyakorlati problémára szerettük volna megoldást találni. Ez a gyakorlati probléma az ATM gerinchálózat felhasználása IP adatátvitelre, azaz például egy cég földrajzilag különálló hálózatainak egyesítése a hálózatok közti ATM bérelt vonalakkal. Hasonló problémához vezet egy Internet szolgáltató problémája is, amely szerverei és a legközelebbi Internet belépési pont között szeretne igénybe venni ATM hálózatot (1. ábra). A forgalom típusától függően a kapott mérési eredmények segítségével javaslatok tehetők az ATM hálózatot felhasználni kívánó cégeknek az IP paraméterek helyes megválasztására, és a megfelelő ATM szolgálati osztály, illetve osztályhoz tartozó paraméterek kérésére a szolgáltatótól. Természetesen ez utóbbi már gazdasági jellegű döntést is kíván.

Vizsgálatunk célja adott paraméterekkel rendelkező ATM kapcsolaton a lehető leghatékonyabb IP átvitel elérése volt. Ezt a célt az IP paraméterek helyes megválasztásával próbáltuk elérni, hiszen az IP szintű paraméterek könnyedén konfigurálhatók, míg az ATM paraméterek elsősorban a szolgáltatótól függenek, tehát a felhasználók szempontjából adottan tekinthetők. Természetesen a méréseket különböző paraméterek esetén is elvégeztük, miáltal lehetőség nyílik az ATM paraméterek hatását is megfigyelni.



1. ábra. A vizsgált hálózat felépítése

A fő probléma az, hogy annak ellenére, hogy az ATM minőségi garanciákat (Quality of Service, QoS) biztosít a kapcsolatnak, ez a TCP szinten már nem jut érvényre. Ennek az az oka, hogy az ATM elvárja a kapcsolattól, hogy a forgalmi szerződésben rögzített feltételeket betartsa. A TCP nincs tisztában a forgalmi szerződéssel, csupán a küldési sebességét tudja változtatni, azt is csak csomagszinten, és nincs felkészítve állandó sebességű adatátvitelre. Mivel az ATM csak úgy tud garanciákat vállalni, ha a forrás betartja a forgalmi szerződést, így ennek megszegésekor a hibamentes, minőségi garanciákkal rendelkező átvitelt nem képes tovább biztosítani. Az elsődleges paraméter, amire optimalizálni kell a kapcsolatot az effektív TCP átvitel, illetve a rendelkezésre álló fizikai sávszélesség kihasználtsága, hiszen egy bérelt vonalakat használó cég elsősorban ezt igényli a hálózattól. Ezért a méréseinket ennek szellemében végeztük, és legjobb eredménynek azt könyveltük el, ahol az elért sávszélesség leginkább megközelíti a fizikailag adottat.

2. HÁTTÉRINFORMÁCIÓK

2.1. A TCP/IP protokoll

Az IP protokoll [1], [4] az Internet hálózati szintű protokollja. Feladata, a felső rétegektől érkező csomagokat a meghatározott címre továbbítani. Szükség esetén darabolhatja a csomagokat, majd a fogadó oldalán helyreállítja azokat. Nem tartalmaz semmiféle biztosítékot arra nézve, hogy a csomagok megérkeznek a célállomásra, illetve hogy sorrendhelyesen érkeznek meg.

A TCP protokoll [2], [4], megbízható, sorrendhelyes összeköttetést biztosít jó teljesítményt nyújtva a legkülönbözőbb kiterjedésű és sebességű hagyományos hálózatok felett. Röviden tekintsük át a TCP protokoll főbb elemeit, melyek segítségével a fenti tulajdonságokat megvalósítja:

- **Sorszámozás (sequence number):** A küldő a kapcsolat felépítésekor véletlenszerű kezdőértékről indít egy számlálót, melyet minden csomag elküldésekor a csomag által hordozott információ byte-ban mért méretével növel. Az elküldött csomagokban szerepel a számláló aktuális értéke, így állítható helyre a csomagok sorrendje, illetve ezen sorszámokat használja a nyugtázási mechanizmus is.
- **Nyugtázás (cumulative acknowledgement):** A fogadó nyugtázza az érkező csomagokat, mindig annak a byte-nak a sorszámát küldve vissza, amelyet a következő csomagban vár. Ezért egyetlen nyugta az összes addig elküldött csomagot nyugtázza, melyben a sequence number érték a nyugtázottnál kisebb.
- **Forgalomszabályozás:** A forgalomszabályozást a TCP két csúszó ablak segítségével végzi. Az egyik ablak (receive window – *rwnd*) jelzi a vevő, a másik (congestion window – *cwnd*) a hálózat szabad kapacitását. A hálózatban lévő csomagok mindenkori lehetséges számát a két ablak közül a kisebbik mérete határozza meg.
- **Slow start és congestion avoidance:** A TCP feladata, hogy garantált, összeköttetés-alapú szolgáltatást biztosítson, valamint hogy a kapcsolat sebességét szabályozva, minél jobban kihasználja az adott hálózat lehetőségeit. Az adatátvitel indulásakor a küldő egyetlen csomagot ad át a hálózatnak, majd minden sikeresen átjutott csomag (érkezett rá nyugta) után megnöveli egy csomag méretével az ablakának méretét, így növelve az adási sebességet. A küldő ablaka (*cwnd*) emiatt exponenciálisan növekszik egy körülfordulási idő (Round Trip Time, RTT) alatt. Ez könnyen belátható, hiszen egy körülfordulási idő alatt először 1, majd 2, 4, 8 stb. csomag lesz a küldő ablakmérete. Ezt a szakaszt hívjuk slow start szakasznak. Ez a növekedés addig folyhat, míg el nem ér egy küszöbértéket (Slow Start Threshold, *ssthresh*). Az *ssthresh* kiindu-

lási értéke 64kbyte, csomagvesztés esetén csökken. Amikor az ablak (cwnd) mérete eléri ezt az sstresh küszöbértéket, a növekedés exponenciálisból lineárisba vált. Ezt congestion avoidance szakasznak hívjuk. Célunk ezzel a kétszakaszos módszerrel az, hogy viszonylag gyorsan elérjünk egy komolyabb átviteli sebességet, de a növekedést lassítsuk le akkor, amikor számításaink szerint elértünk a fizikai közeg teljesítőképességének határára.

- **Torlódáskezelő algoritmusok:** timeout és fast retransmit [3]: Amikor egy csomag elveszik, a TCP feladata, hogy detektálja és korrigálja a hibát, lehetőleg minél rövidebb idő alatt. Egy csomag elveszhet egyszeri hiba miatt, amikor csak az adott csomag sérül, de történhet tömeges csomagvesztés is. Az első detektálása gyorsabban zajlik: a vevő észreveszi, ha egy csomag kimarad, és olyan nyugtát küld, amivel az előző (hiányzó) csomagot kéri. Ha ezt egymás után többször is megteszi, akkor az adó azt feltételezheti, hogy egy csomag elveszett, és ekkor ezt a hiányzó csomagot újraküldve megoldhatja a problémát. Ezt a működést implementálja a fast retransmit algoritmus. Amennyiben tömeges veszteség következik be, az adó csak a nyugták hiányát észlelheti, és egy bizonyos várakozási idő után úgy dönt, hogy az elküldött csomagok elvesztek, és újra kell adni őket. Ekkor az adás egy új slow start szakasszal kezdődik. Ezt az algoritmust timeout – slow start algoritmusnak hívjuk. Természetesen mindkét esetben csökkentjük az sstresh küszöbértéket is. A működés pontosabb leírása megtekinthető az RFC2001 szabványban.

2.2. AZ ATM jellemzői

Az ATM hálózati technikát [5] egységes megoldásnak tervezték, emiatt fel kellett készíteni a legkülönbözőbb forgalomkarakteristikával rendelkező kapcsolatok egyidejű fogadására. Az ATM hálózaton egyszerre bonyolódhat beszéd-, hang- és képátvitel, de fel van készítve a legkülönbözőbb jövőbeni igényekre is, legyenek azok késleltetésre érzékenyek, vagy érzéketlenek. Az ATM legfontosabb tulajdonsága, hogy képes szolgáltatásminőségi garanciákat adni, azaz a felhasználó számára garantálja hálózatminőségi paramétereit. Az ATM cellakapcsolt hálózat, amely átmenetet képez a hagyományos távközlési hálózatok, és az adatkommunikációban elterjedt csomag-kapcsolt hálózatok között.

A következő pontok röviden összefoglalják az ATM jellemzőit:

- **Statisztikus multiplexálás:** Nem fixen lefoglalt sáv szélességgel dolgozik (mint pl. az SDH), hanem virtuális erőforrás-foglalással dolgozik. Amikor egy kapcsolaton nincs forgalom, a sáv szélességet más használhatja.
- **Fix cellaméret (48/53 byte):** A fix cellaméret rendkívül gyors cellakapcsolást tesz lehetővé, hiszen a teljes folyamat hardverben végezhető.
- **Forgalmi szerződés, szolgáltatásminőség:** Minden kapcsolat számára forgalmi szerződést kell kötni a kiépítésekor, a hálózat ugyanis csak úgy képes minőségi garanciák adására, ha a forrás előre meghatározott módon forgalmaz. A forgalmi szerződésben a forrás megadja a hálózat számára a forgalmát leíró paramétereit, valamint azt, hogy milyen szolgáltatást, illetve szolgáltatási garanciákat

vár tőle. A forgalmi szerződésben tehát specifikáljuk a szolgáltatási kategóriát, majd ettől függően a megfelelő forgalomleírókat, és minőségi követelményeket.

- **Szolgáltatási kategória:** Az integráló szerep megkívánja, hogy a gyökeresen eltérő karakterisztikájú kapcsolatok az ATM rétegben különböző kezelésben részesüljenek. Ezek a kategóriák a következők:
 - **CBR (Constant Bit Rate):** konstans bitsebességű forgalmakhoz, mint pl. a mai PCM alapú beszédátvitel, a PDH és SDH rendszerekben. A legmagasabb prioritású osztály, ami fix sáv szélességet kap.
 - **VBR (Variable Bit Rate):** változó bitsebességű forgalmakhoz konstruált kategória. Két alfaja létezik, valósidejű VBR (real-time VBR, rt-VBR) és nem valósidejű VBR (non real-time VBR, nrt-VBR) kategória. Az rt-VBR összeköttetések számára cellakésleltetési garanciát is vállal a rendszer, valamint magasabb prioritású várakozási sort rendel hozzá, mint a csak cellavesztési garanciában részesülő nrt-VBR kapcsolatokhoz
 - **ABR (Available Bit Rate):** rendelkezésre álló bitsebesség. Részben garantált, értéknövelt "best-effort" szolgáltatási osztály. A kapcsolatok számára opcionálisan garantál egy minimális sáv szélességet. majd a kapcsolat élettartama alatt a hálózat folyamatosan értesíti a forrást az ezen felül rendelkezésére bocsátott sáv szélességről.
 - **UBR (Unspecified Bit Rate):** nem garantált minőségű szolgáltatás, a kapcsolat számára nem történik sáv szélesség-foglalás. A legalacsonyabb prioritású szolgáltatási kategória, cellákat először ezekből a folyamatokból dob el a forgalomszabályozó algoritmus túlterheltség esetén.
- **Forgalomleírók:**
 - **PCR (Peak Cell Rate):** maximális cellasebesség
 - **SCR (Sustainable Cell Rate):** átlagosan fenntartható cellasebesség
 - **MBS (Maximal Burst Size):** maximális löketméret
 - **CDVT (Cell Delay Variation Tolerance):** késleltetés ingadozás tolerancia
- **Szolgáltatásminőségi paraméterek:**
 - **CTD (Cell Transfer Delay):** átviteli késleltetés
 - **CDV (Cell Delay Variation):** késleltetés-ingadozás
 - **CLR (Cell Loss Ratio):** cellavesztés
- **Hívásengedélyezés:** A kapcsolatban résztvevő valamennyi kapcsoló ellenőrzi, hogy rendelkezésre állnak-e a szükséges erőforrásai, és bármelyük megtagadhatja a kapcsolat felépítését, ha úgy találja, hogy nem rendelkezik megfelelő erőforrással. Ekkor a kapcsolat nem épül ki, hiszen a hálózat nem tudná biztosítani a kért szolgáltatást. Ha egy kapcsolat sikeresen felépült, akkor a vállalt garanciák a felépüléstől kezdve élenek. Ez azt jelenti, hogy ettől kezdve amennyiben a kapcsolat nem szegi meg a forgalmi szerződést, a hálózat terheltségétől függetlenül mindig megkapja a számára garantált erőforrásokat.
- **Forgalomkorlátozás (Traffic Policing):** A hálózat figyeli, hogy az egyes kapcsolatokon az adatforgalom valóban megfelel-e a szerződésben foglaltaknak. Ha nem, akkor a szerződést megszegő cellákat eldobhatja, vagy alacsony prioritásúvá teheti őket, hogy a hálózat későbbi szaka-

szán egy másik kapcsoló dobja el. A forgalomkorlátozás a többi, szabályosan viselkedő kapcsolat védelmére hivatott, hiszen ha egy kapcsolat következmények nélkül megszeghetné a forgalmi szerződést, akkor ezzel túlterhelhetné az ATM kapcsolót, illetve a hálózatot, veszélyeztetve a vállalt szolgáltatásminőségi garanciákat.

- **ATM jelzésrendszer:** Az ATM alapvetően összeköttetés orientált, ami azt jelenti, hogy adatforgalmazás előtt a kapcsolatot fel kell építeni. Az ATM hálózaton kétféle virtuális kapcsolat építhető ki: állandóan fennálló (Permanent Virtual Circuit, PVC), illetve a kapcsolt virtuális összeköttetés (Switched Virtual Circuit, SVC). Az állandó kapcsolat az ATM hálózatot üzemeltető operátor által kerül beállításra, ettől kezdve rendelkezésre áll, míg a dinamikus kapcsolat felépítéséről és lebontásáról jelzésváltással kell gondoskodnia a felhasználónak. A statikus kapcsolat tulajdonképpen bérelt vonali összeköttetéshez hasonló, a dinamikus pedig a telefóniában megszokott kapcsolatorientált összeköttetést reprezentálja.

2.3. Klasszikus IP ATM felett

Az IP ATM hálózatok feletti kommunikációjának egyik lehetséges megoldását kínálja a protokoll, melyet az Internet Engineering Task Force dolgozott ki [6]. A kommunikációhoz a következőkre van szükség:

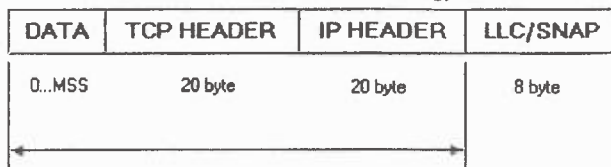
- **Címfeloldás:** Egy IP csomag küldésekor az ATM címfeloldó protokoll (ATM Address Resolution Protocol, ATMARP) a célállomás IP címe alapján megkeresi annak ATM címét. Amennyiben a küldő nem ismeri a célállomás ATM címét, akkor az ATMARP szervertől fordul a címfeloldási kérelmével (ennek ATM címét alapbeállításakor ismeri meg).
- **Encapsulation:** A csomag átküldése előtt egy fejléccet kap. A CIP az IEEE 802.2 LLC/SNAP encapsulation technikát használja. A teljes AAL5 csomag így a következőkből fog állni:
Adat + 20 byte (IP fejléc) + 20 byte (TCP fejléc) + 8 byte (LLC/SNAP fejléc)
Az ATM hálózatra kerülés előtt ezt az adatsomagot 48 byte-os részekre szabdalja az AAL réteg, ebből következően méréseink során figyelembe kell venni, hogy a hálózat akkor lesz a leginkább kihasználtság, ha az AAL5 csomag mérete egész számú többszöröse 48 byte-nak.

2.4. Teljesítményre jelentős hatással lévő TCP/IP, ATM és egyéb paraméterek

A TCP/IP és az ATM protokolljai is nagyszámú paraméterrel rendelkeznek. Ezek közül majdnem mindegyik hatásos van a kapcsolat minőségére, de van néhány paraméter, amelyek azt alapvetően határozzák meg. Ezek a következők:

- **Maximum Transmission Unit (MTU) mérete (byte):** Ez az interfészenként állítható paraméter az IP csomag maximális méretét adja meg. Amennyiben az IP réteg ennél nagyobb csomagot kap a felső szint felől, azt maximá-

lisan akkora darabokra darabolja, hogy az IP fejléccel együtt elférjen az MTU méret által meghatározott nagyságú helyen. A 2. ábra az így felépített TCP/IP csomagot mutatja a CIP LLC/SNAP fejléccel együtt.



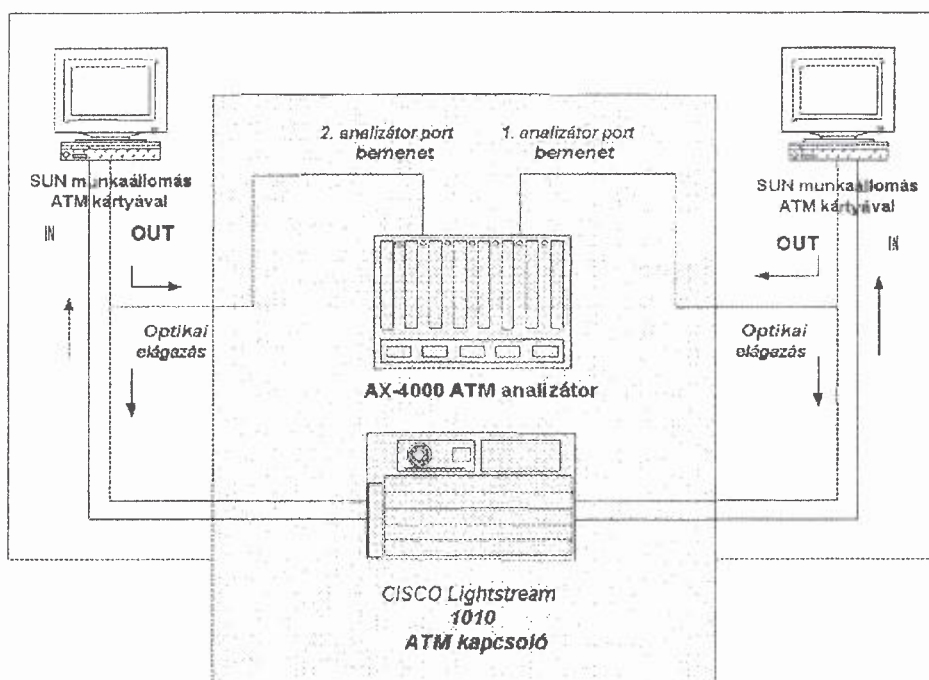
2. ábra. Egy AAL5 csomag felépítése

- **Küldő és vevőoldali puffer mérete:** Ez a paraméter adja a felső korlátját a TCP által hirdetett ablakméretnek, ami befolyásolja, hogy egy időben hány nyugtázatlan csomag lehet kint a hálózaton. Ezen paraméter növelésével növeljük a TCP forgalomszabályozó algoritmusainak mozgásterét.
- **ATM szolgálati kategória és forgalomleírók:** Az IP forgalmat hordozó ATM kapcsolat fogalmi szerződése erősen befolyásolja az alkalmazás-szintű kapcsolatminőséget. A szolgáltatási osztály és a lefoglalt sávszélesség általában az alkalmazási területtől és a rendelkezésre álló költségkerettől függ elsősorban, így méréseink arra irányultak, hogy adott szolgáltatási osztályon belül az ATM forgalomleírók milyen mozgásterünk van a minőség javítására.
- **Forgalomformázás, sávszélesség maximalizálás:** Általában felhasználói oldalon beállítható opció, amely egy megadható sebesség fölé nem engedi a TCP adási sebességet. Ezt egy beállítható kiszolgálási sebességű FIFO tároló segítségével éri el. A forgalomformázás lehetőséget ad arra, hogy az ATM forgalmi szerződést nem ismerő TCP/IP forgalmát úgy alakítsuk, hogy elkerülje az összeköttökést az ATM forgalomkorlátozó mechanizmusával.

3. A KÍSÉRLETI HÁLÓZAT FELÉPÍTÉSE

A méréseket az Ericsson Traffic Lab kutatólaboratóriumban végeztük a következő mérési összeállítás igénybevételével:

- 2 db Sun SPARCstation 5 [7] munkaállomás FORE Systems ATM kártyákkal felszerelve (ForeRunner SBA200E) [8],
 - 1 db Cisco LightStream 1010 ATM kapcsoló [9],
 - 1 db Adtech AX-4000 ATM hálózati analizátor berendezés [10],
 - a Hewlett-Packard által kifejlesztett Netperf hálózati teljesítménymérő szoftver [11], és
 - File Transfer Protocol az átviteli sebesség mérésére.
- A munkaállomások közti kapcsolatot PVC segítségével oldottuk meg, az egyszerűen kézbentartható konfigurálhatóság érdekében. Az átvitelre a Klasszikus IP ATM felett (Classical IP over ATM, CLIP) protokollt használtuk, szintén egyszerűsége miatt. A munkaállomásokon Solaris operációs rendszer futott.



3. ábra. A mérési elrendezés

4. TELJESÍTMÉNYVIZSGÁLATOK

4.1. A csomagméret hatása

A hálózatba kerülő csomagok méretére két módon is hatással lehetünk. Mérések, és tesztlések esetén lehetőség nyílik felhasználói szinten meghatározni, hogy mekkora adatot küldünk el egy blokkban, így az MS (Message Size, üzenetméret) segítségével kontrolláljuk az IP csomagok méretét. Egyes alkalmazások jellegükből adódóan meghatározzák az általuk használt üzenetek méretét. Tipikusan a távoli terminálélelést nyújtó alkalmazások rövid üzenetekkel dolgoznak, míg a blokk átvitelt nyújtó applikációk kihasználják a maximális üzenetméretet.

A maximális csomagméretet az MTU paraméter segítségével hálózati interfészenként tudjuk beállítani, így a felhasználó számára láthatatlan módon tudunk befolyással lenni az IP által használt csomagméretre. Az MTU méretet minden gépen a rendszergazda tudja állítani. Minden TCP/IP kapcsolat esetén a kommunikációban részt vevő felek menet közben dinamikusan felméri az útvonal során a legkisebb MTU méretet használó hálózat MTU méretét, és ezt teszik meg a közös MTU méretnek.

4.2. A forgalomformázás hatása

A FORE típusú ATM kártyán beállítható egyszerű forgalomformázási funkcióknak komoly jelentősége van abban, hogy a TCP/IP legalább a maximális sebességet ne legyen képes túllépni, hiszen az az ATM kapcsolóban cellák eldobásához, így a kapcsolat szempontjából csomagvesztéshez vezetne. A forgalomformázás hátránya, hogy némi késleltetést és késleltetés ingadozást visz be az átvitelbe, így az erre érzékeny alkalmazások (tipikusan beszéd vagy videó) esetén nem, vagy csak megkötésekkel alkalmazhatók. A forgalomformázás IP átvitelre gyakorolt hatásának nagysága erősen függ az alkalmazott MTU méretétől.

4.3. A maximális löketméret hatása

Az egyetlen ATM paraméter, amelynek hatását megvizsgáltuk a VBR kategória maximális löketméret paramétere. Ennek a paraméternek a növelésével az ATM kapcsolóban lévő puffereket terheljük, így növelésével óvatosan kell bánni, mert könnyen felesleges hálózati terheléshez vezethet. Emellett az MBS növelésének hatása lecsengő, azaz egy bizonyos határ felett már nem érdemes tovább növelni az értékét, mert az átvitel már nem javul.

Méréseinket a minőségi garanciákat adó ATM kategóriák esetére korlátoztuk, hiszen itt a forgalom jellegét nem befolyásolhatják a külső körülmények (kapcsoló és link terheltsége).

5. MÉRÉSEK

5.1. Állandó bitsebesség (CBR)

Méréseinket egy 34 Mbps maximális bitsebességet adó állandó virtuális áramkörön végeztük, amellyel egy E3-as bérelt vonalat szerettünk volna emulálni. Olyan forgalomkorlátozást használtunk, ami a szerződést megsértő cellákat azonnal eldobta. A csomagméretet az extrém kis esettől (88 byte) növeltük 8232 byte-os méretig. A kis csomagok esetén ügyeltünk arra, hogy a keletkezett AAL5 csomagok 48 byte egész számú többszörösei legyenek, így elkerülve a kitöltetlen csomagok által okozott torzítást. Nagyobb csomagok esetén erre már nem volt szükség, így itt inkább a gyakorlatban használt értékeket vettük alapul. Így a következő csomagméretek születtek: 88, 232, 472, 576, 1518 és 8232 byte. Ezekből az utolsó három jól ismert érték, az 576 byte az a minimális érték, amelyet minden hálózatnak támogatnia kell, az 1518 byte az elterjedt Ethernet hálózatok alapértéke, a 8232 byte pedig az alkalmazások által leggyakrabban alkalmazott üzenetméret.

Az előzetes várakozások szerint a nagyon kicsi csomagméret esetén gyenge lesz a kihasználtság, amit egyrészt a nagy fejléc/adatarány, másrészt a megnövekedő darabolási és összeállítási idő okoz. Közepes csomagok esetén várhatóan nő a kihasználtság, viszont egy csomagnak nagyobb esélye van arra, hogy megszegi a szerződést, és ezért a kapcsoló eldob belőle cellákat, így várhatóan az első csomagvesztés előbb következik be. Nagy csomagok esetén a kihasználtság jó, viszont ebben az esetben a veszteség is nő, mivel itt már az ATM forgalomkorlátozásával komolyan számolnunk kell. Esetleg már az első csomagok egyike is megsérülhet, amennyiben nem használunk forgalomformázást. Forgalomformázás esetén a csomagméret növelésével egyre jobb kihasználtságot várunk.

Az átvitel minőségét tehát meg kell nézni forgalomformázás esetén, és annak használata nélkül is. A forgalomformázás itt annyit tesz, hogy a forrás számítógépben egy 34 Mbps kiszolgálási sebességű várakozási soron vezetjük át a kapcsolatot, így kontrolláljuk annak maximális sebességét. Az 1. és 2. táblázat első oszlopa az *MTU méretet* mutatja byte-ban kifejezve, második és harmadik oszlopa a különböző módokon mért átviteli sebességet szemlélteti kbps mértékegységben, utolsó két oszlopa pedig kihasználtság értékeket mutat százalékos formában. A *Netperf* oszlop a Hewlett Packard által kifejlesztett Netperf teljesítmény-analizátor szoftver segítségével mért értéket, az *FTP* az FTP által mért átviteli sebességet mutatja meg. A kihasználtság értékek közül az első, a *Max. kihasználtság* azt mutatja meg, hogy elméletileg mekkora kihasználtságot lenne képes elérni a kapcsolat, amennyiben nem történne adatvesztés, tehát a kihasználtságot csak az átvitt fejlécek csökkentenék. Például 88 byte-os MTU esetén a hasznos adat 48 byte, az egy csomag átviteléhez szükséges adatmennyiség pedig $2 \cdot 53$, azaz 106 byte, így a hányados $48/106$, azaz 45 %. Ez azt jelenti, hogy ezzel az MTU mérettel 45 %-nál jobban semmiképp nem tudjuk kihasználni a hálózatot, hiszen a hasznos legjobb esetben is csupán az összes átvitt adat 45 %-a lesz. Kis csomagok esetén a veszteség jelentős, így a maximális kihasználtság csekély, míg nagy méretnél a veszteség gyakorlatilag az ATM fejléc által okozott veszteségre korlátozódik ($48/53$), itt a maximális kihasználtság lényegesen jobb. A *Kihasználtság* oszlop az FTP-vel ténylegesen mért érték és a garantált sávszélesség hányadosa. Ez a garantált sávszélesség állandó cellasebesség (CBR) esetén a maximális sávszélesség (PCR), változó cellasebességű kategória (VBR) esetén pedig az átlagos cellasebesség (SCR). A teljesítményértékek $3 \cdot 10$ független mérés átlagaként adódtak, mindkét módszerrel körülbelül 1 perces átviteleket teszteltünk.

Látható, hogy várakozásainknak megfelelően, ha a maximális adási sebességet lecsökkentettük PCR szintje alá a forgalomformázással, akkor a kapcsolat annál gyorsabb, minél nagyobb az MTU méret. Ebben az esetben a szűk keresztmetszetet a csomagdaraboló és egyesítő mechanizmusok működése jelenti a fejlécből adódó veszteségek mellett, hiszen ha csupán a nagyobb fejléc/adatarány miatt lenne lassabb a kis MTU mérettel rendelkező kapcsolat, abban az esetben a nyers átvitel azonos értéket adna.

Forgalomformázás nélkül ellenben az átvitel drámai visszaesését tapasztalhatjuk a nagy MTU méretekkel rendelkező kapcsolatok esetén. Ezt azzal lehet magyarázni,

hogy ezek a kapcsolatok már az első csomagokkal megsértik a forgalmi szerződést, így a TCP még a körforgulási időt sem tudja bemérni, és az újraküldés előtti várakozás egyre nő (az RTO 6 s-ról indul, de a sikeresen átvitt csomagok nyugtájának beérkezésekor csökken az értéke). Emiatt a veszteségek között nagy várakozások vannak (másodperces nagyságrendűek). A 8232 byte-os MTU méret esetén pedig látható, hogy már az első csomag is elveszik, így adatátvitel nem tud zajlani.

1. táblázat. TCP/IP teljesítmény CBR kapcsolaton, forgalomformázás használata esetén

Az MTA méret hatása				
MTU méret (byte)	Netperf (kbps)	FTP (kbps)	Max. kihasználtság, %	Kihasználtság, %
88	416	226	45,4	0,7
232	10394	13257	72,5	39,0
472	18034	18315	81,5	53,9
576	19873	21491	77,8	63,2
1518	21567	23768	87,2	69,9
8232	26877	27500	89,9	80,9

2. táblázat. A forgalomformázás nélküli TCP/IP teljesítmény

Az MTA méret hatása				
MTU méret (byte)	Netperf (kbps)	FTP (kbps)	Max. kihasználtság, %	Kihasználtság, %
88	412	225	45,4	0,7
232	12630	13821	72,5	40,7
472	15730	17321	81,5	50,9
576	15982	18532	77,8	54,5
1518	0	1500	87,2	4,4
8232	0	0	89,8	0,0

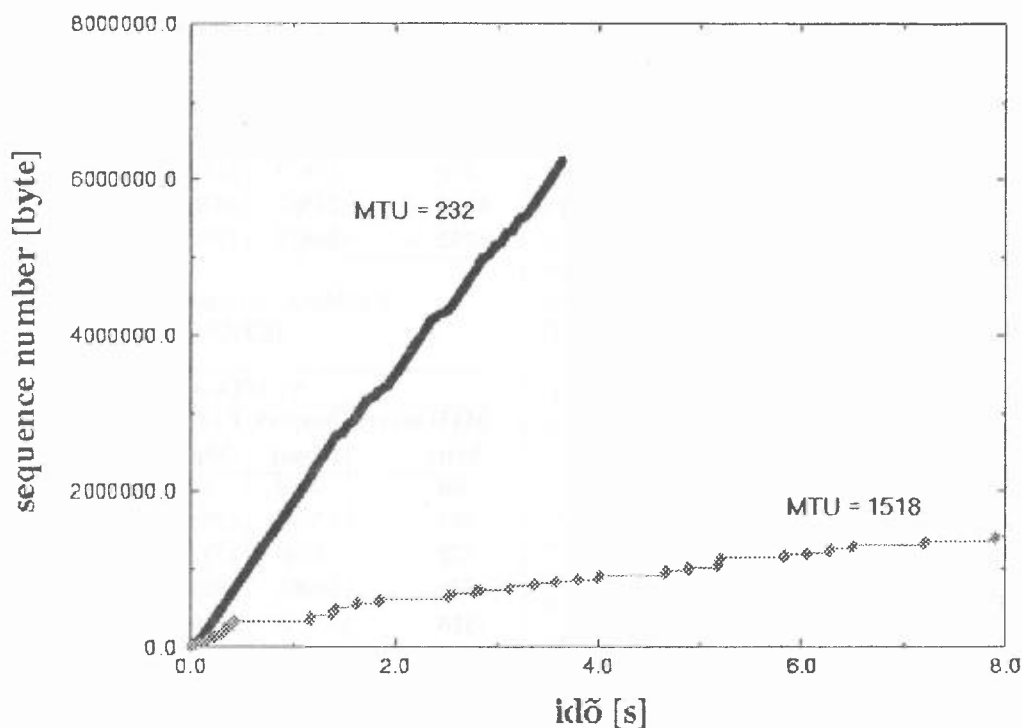
Kis csomagméret esetén a teljesítmény két dolog miatt nem esik vissza. Az egyik tényező, hogy az első csomagok sikeresen átjutnak, így a TCP beméri a körforgulási időt, lecsökkenti az RTO értéket, és veszteségek esetén nem vár olyan sokat. Emellett a csomagvesztést legtöbbször képes a gyorsabb algoritmussal korrigálni (Fast Retransmit és Fast Recovery). Nagy MTU esetén egy csomag elvesztése után már nincs lehetőség 3 további csomag sikeres átvitelére, így a küldő csak a nyugta hiányából értesül a hiányról. Ezzel szemben a kis csomagmérettel dolgozó kapcsolat azonos ablakméret mellett jóval több csomagot tarthat kint a hálózatban, és kisebb az esélye a többszörös csomagvesztésnek. Ebben az esetben egy csomag elvesztése után a vevő rövidebben észreveszi a veszteséget a hálózatban lévő többi csomag megérkezésekor, és ismételt nyugták küldésével kérheti az adótól az újraküldést (három egyforma nyugta érkezése esetén indul be a Fast Retransmit algoritmus az adónál). Ráadásul a Fast Retransmit algoritmus csak az egyszeres csomagvesztés kiküszöbölésében hatékony, így nagy csomagméret esetén, ahol gyakran alakul ki többszörös veszteség, kevésbé eredményes a működése.

A második tényező sokkal egyszerűbb. Mivel a csomagdarabolási és összeállítási idő lecsökkenti az adási teljesítményt azáltal, hogy időt vesz igénybe, így nagyon kis MTU méret esetén a kapcsolat önmagát szabályozza le egy bizo-

nyos sebesség alá azzal, hogy a forrás egyszerűen képtelen gyorsabban adni. Ezt úgy láthatjuk be, hogy összehasonlítjuk a két táblázatot. Látható, hogy 88 byte-os MTU méret esetén a két érték gyakorlatilag megegyezik és nagyon alacsony. Ebben az esetben valószínűleg egyéb operációs rendszertől függő problémák is befolyásolják az átvitelt, hiszen az értékek minden esetben irreálisan alacsonyak. 232 byte-os MTU esetben is gyakorlatilag megegyezik a két mérési eredmény. Látható, hogy először 576 byte-os MTU méret esetén már tapasztalható olyan mértékű visszaesés, amit nehéz mérési hibának kezelni, így itt már valószínűleg az elsőként említett okok lehetnek meghatározók.

A forgalomformázás nélküli esetet szemlélteti a 4. ábra,

ahol a vízszintes tengelyen az eltelt idő másodpercben, a függőleges tengelyen pedig az átküldött byte-ok száma látható. Az egyenes meredeksége az átviteli sebességgel (throughput) arányos. Az átküldött byte-ok számát a TCP fejléc sequence number sorszámból nyertük ki, így újraküldés esetén a függvény megtörik. A függvény ábrázolásához használt adathalmaz az FTP alkalmazás forgalmából vett minta. Látható, hogy az 1518 byte-os MTU mérettel rendelkező kapcsolat már viszonylag rövid idő alatt megszegi a forgalmi szerződést, minek következtében cellavesztés, és ebből fakadóan csomagvesztés lép fel. A hosszú, vízszintes szakasz a 4. ábrán a veszteség észrevételének idejét mutatja (RTO).



4. ábra. Forgalomformázás nélküli kapcsolatok CBR ATM linken

5.2. Változó bitsebesség (VBR)

Méréseinket egy 34 Mbps maximális (PCR) és 24 Mbps fenntartható (SCR) sávzélességet biztosító VBR állandó virtuális áramkörön végeztem. A maximális löketméretet három különböző szintre állítottuk, hogy ennek hatása is megfigyelhető legyen. Ebben a mérési környezetben is leteszteltük a forgalomformázás hatását, melynek beállításai az előző méréssel azonosak voltak, azaz 34 Mbps kiszolgálási sebességű volt a várakozási sor. Azért ezt a sebességet választottuk, hogy a forgalmi szerződés egyik paraméterét így biztosan betartsa a kapcsolat, de emellett megfigyelhesük, hogy a másik korlát, az átlagos sebesség mennyiben befolyásolja a TCP/IP működését. A mérések tehát ebben az esetben is két részre oszthatók a forgalomformázás állapotától függően. Az ATM kapcsolatban a forgalomkorlátozást az előző mérésnek megfelelően úgy állítottuk be, hogy a szerződészegő cellákat azonnal eldobja. A beállított MTU méretek is az előző mérésben használtakal egyeztek meg.

A VBR osztály használata már sokkal komolyabb elvárásokat támaszt a TCP által generált forgalommal szemben. Nem elég csupán a PCR értéket betartani, hanem még az átlagos sebességet sem szabad korlátlanul túllépni. Mivel a TCP csupán az adási sebességét tudja szabályozni, előzetesen azt vártuk, hogy a kapcsolatok általában rosszabbul fognak teljesíteni. Az MBS paramétert úgy választottuk meg, hogy megadott számú TCP/IP csomag férjen bele, ezzel biztosítva azt, hogy minden csomagméret esetén azonos mozgástere legyen a TCP forgalomszabályozási mechanizmusának.

Forgalomformázás nélkül az MBS paraméter jelentősége csökkent, a kapcsolatok MBS paramétertől függetlenül rossz átvitelt produkáltak, ám az átviteli teljesítmény romlása nem volt olyan rendkívül gyors, mint CBR esetben. A 4. táblázat ezt mutatja. A táblázat összevont, azaz a különböző MBS beállítások esetén mért értékek közül a legkedvezőbbet mutatja.

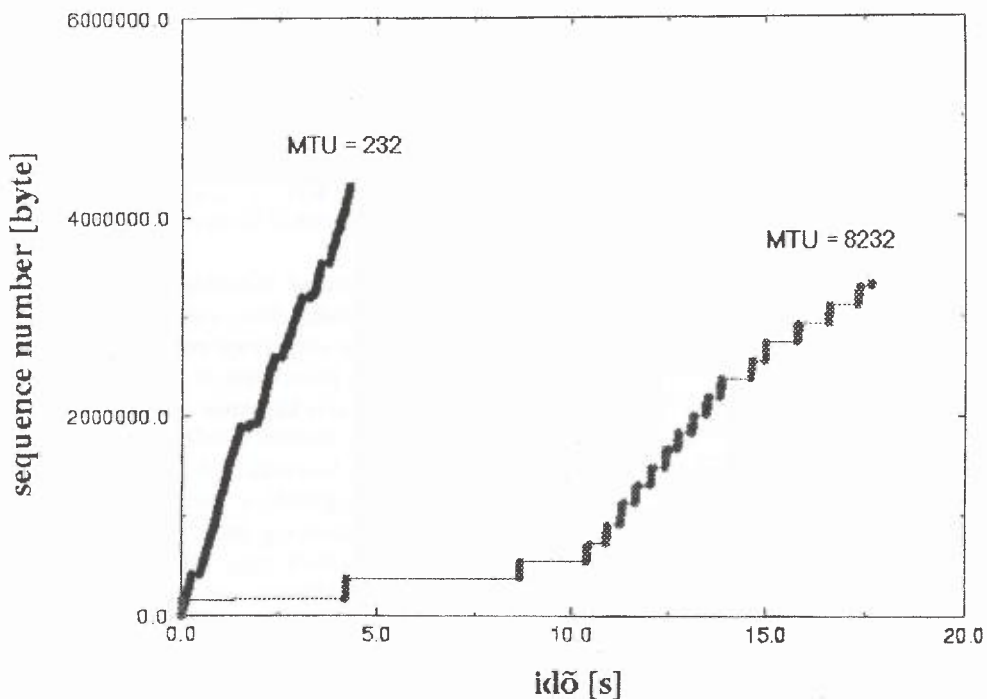
3. táblázat TCP/IP teljesítménye különböző VBR ATM linkeken, forgalomformázással

Az MTA méret hatása, különböző MBS beállítások esetén				
MTU méret (byte)	Netperf (kbps)	FTP (kbps)	Max. kihasználtság, %	Kihasználtság, %
MBS = 10*MTU				
88	227	236	45,5	1,0
232	11177	10920	72,5	45,5
472	4393	5651	81,5	23,5
576	2940	4693	77,8	19,6
1518	7133	5069	87,2	21,1
8232	6630	7264	89,8	26,1
MBS = 100*MTU				
88	407	240	45,4	1,0
232	12997	11035	72,5	46,0
472	9167	8088	81,5	33,7
576	8337	8427	77,8	35,1
1518	9277	9997	87,2	41,7
8232	7140	8424	89,8	35,1
MBS = 1000*MTU				
88	373	196	45,4	0,8
232	11163	9827	72,5	40,9
472	11050	8768	81,5	36,5
576	9143	9627	77,8	40,1
1518	8767	11629	87,2	48,5
8232	6704	7419	89,8	30,9

4. táblázat. Forgalomformázás nélküli TCP/IP teljesítménye VBR ATM linken (legjobb mért eredmény a 3 különböző MBS esetből)

Az MTA méret hatása, különböző MBS beállítások esetén				
MTU méret (byte)	Netperf (kbps)	FTP (kbps)	Max. kihasználtság, %	Kihasználtság, %
88	373	196	45,4	0,8
232	10062	10627	72,5	44,3
472	3781	4012	81,5	16,7
576	3580	3520	77,8	14,7
1518	2967	2736	87,2	11,4
8232	2211	1960	89,8	8,2

VBR esetben sokkal nehezebb következtetéseket levonni. Látható, hogy forgalomformázott esetben is problémás a kihasználtság, főleg nagy MTU méret, és kis MBS esetén. Ez a probléma az MBS növelésével csökkenthető, de meg nem oldható. A nagy MBS paraméter azt jelenti, hogy az ATM kapcsoló puffereit használjuk az átlagnál nagyobb mértékben. Tulajdonképpen ilyenkor a TCP/IP maximális sebességgel ad egészen addig, amíg túl nem lépi a forgalmi szerződésben foglalt maximális lötetméretet, majd ezután leáll, mivel csomagvesztést érzékel. Ameddig a TCP újraküldési időzítője lejár, a cellák is kiürülnek a kapcsoló puffereiből, így ismét kezdődhet az adás. Ezzel egy viszonylag determinisztikus viselkedésű ON-OFF jellegű kapcsolatot kapunk.



5. ábra. TCP/IP viselkedés hosszú távon forgalomformázás nélküli VBR esetben

A forgalomformázás kiiktatása esetén a kapcsolatok mindkét korlátozást meg tudták szegni, amely elsősorban a nagy csomagokkal működő kapcsolatoknál jelentett komoly problémát. Ezt az esetet szemlélteti az 5. ábra. A probléma itt is különválasztható csomagméret szerint. Kis méret esetén a CBR esetben elmondottak itt is érvényesek. Közepes méret esetén a visszaesés lényegesen komolyabb, mint CBR esetben, nagy MTU méret esetén viszont kevésbé drámai. Az MBS méretének növelése némileg pótolja a forgalomformázás hiányát, nagyobb lefoglalt effektív sáv szélességgel (és bizonyára ennek megfelelő pénzüsszeggel) fizetünk. Az eredmények azt mutatják, hogy az MBS paramétert nem érdemes 100 csomagméretnél többre növelni, mert itt már csak az erőforrásokat pazaroljuk a határfok növekedése nélkül. Ez a megállapítás igaz a forgalomformázást használó esetekre is.

Ezen eset összefoglalásaként elmondható, hogy VBR esetben az MBS paraméter szerepe is lényeges, mégpedig elsősorban közepes és nagy MTU méret mellett. Kis csomagméret mellett az MBS paraméter hatása nem lényeges, bár ekkor is kb. 100 csomag méretének megfelelő MBS a legmegfelelőbb.

VBR esetben a forgalomformázás szerepe lényeges ugyan, főleg közepes és nagy csomagméret esetén, ám nincs akkora hatása, mint CBR esetben. Mivel a TCP algoritmust konstans bitsebességű hordozóhoz való illeszkedésre készítették fel, így csupán a maximális sebességét képes kontrollálni, de a borsztösségét nem. Ezért elmondható, hogy VBR kapcsolattípust csak akkor van értelme használni, amennyiben ezt a forgalom karakterisztikája indokolja, vagy olyan forgalomformázó berendezés áll rendelkezésre, amely a VBR kapcsolat forgalmi szerződésével konform kimenetet képes előállítani. Több TCP kapcsolat multiplexálása esetén az egyes kapcsolatok csúcsai valószínűleg nem egy időbe esnek, így ez a forgalom karakterisztikája alapján már alkalmas VBR típusú közeggel való jó együttműködésre.

6. ÉRTÉKELÉS

A következőkben megpróbálunk magyarázattal szolgálni az eddig még nem, vagy csak részben megmagyarázott eredményekre

A *88 byte-os MTU méretre* nagy valószínűséggel a munkaállomások nem voltak felkészítve, így az ezzel a mérettel történt mérések nem tekinthetők eredményesnek. Valószínűleg a daraboló és összerakó, vagy a forgalomformázó algoritmusok nem működtek megfelelően ezzel az értékű kis csomagmérettel.

Említést érdemel a *232 byte-os csomagmérettel* működő kapcsolatok ATM forgalmi szerződéstől és forgalomkorlátozástól való függetlensége. Ezt elsősorban azzal lehet magyarázni, hogy a nem túl gyors gépek miatt az adó egyébként sem volt képes teljesen kihasználni a rendelkezésre álló sáv szélességet, így nem volt képes megszegni a forgalmi szerződést. CBR esetben megfigyelhető, hogy a korlátozás nélküli esetben a mérés pontatlanságától eltekintve teljesen ugyanazokat az eredményeket kapjuk, mint ha bekapcsoljuk a korlátozást. VBR esetben pedig az is megfigyelhető, hogy sem a forgalomformázás, sem pedig az MBS méret nincs igazán hatással a kapcsolat sáv szélesség ki-

használására, amely minden esetben körülbelül 30–40 %. Ez a megfigyelés alátámasztja azt a feltételezést, hogy ebben az esetben nem csupán a TCP protokoll jó működése, hanem egyszerű teljesítményhiány is oka az egyenletes teljesítménynek.

A *közepes csomagmérettel* működő kapcsolatok (MTU = 472, ill. 576 byte) már sokkal inkább hasznosnak tekinthető eredményeket adtak. CBR esetben látható, hogy a forgalomformázás kikapcsolása visszaveti a kapcsolatot, de komoly veszteségek nem fordultak elő, ami két dologgal is magyarázható.

1. Egyrészt mire a TCP a megengedett fölé növeli az adási sebességét, már elég sok csomag sikeresen ájtut a hálózaton, hogy a forgalomformázó mechanizmus pontosan bemérje a körülfordulási időt és ezzel együtt elég kicsire csökkentse a kezdetben magas értékről indult várakozási időt (RTO). Így amennyiben tömeges veszteség keletkezik, a kapcsolat rövid várakozás után már újraindul.

2. Másik ok, hogy ebben az esetben mire a veszteség bekeövetkezik, a csomagok méretéhez képest már elég nagyra van nyitva a TCP torlódási ablaka, így viszonylag sok csomag van már kint a hálózaton. A szabvány szerint egy, csomag elvesztése esetén lehetőség van arra, hogy később kiadott csomagokra érkező nyugták ismétlődése (duplicate acknowledgements) alapján a küldő észrevegye, hogy csomagvesztés történt, és gyors algoritmussal korrigálja a problémát (Fast Retransmit). Az első forgalmi szerződést megszegő csomag után következő csomagok azért nem szegik meg feltétlenül a forgalmi szerződést, mert a csomagok kiadása nem tökéletesen egyenletes sebességgel zajlik, hanem kissé löketszerűen, így lehetőség van arra, hogy kisebb szünetek után az elvileg ugyanolyan sebességgel leadott csomag ne szegje meg a forgalmi szerződést, hanem átjusson a fogadóhoz.

VBR esetben a forgalomformázásnak valamint az MBS méretnek nagyobb szerepe van, kis MBS esetén gond lehet, hogy tömeges veszteség keletkezik viszonylag korán, és így várakozásra kényszerül a kapcsolat, ráadásul viszonylag korán, így az RTO a kezdeti magas értékénél is nagyobbra nő, ezzel rendkívül hosszú adási szünetekre kényszerítve a forrást.

Nagy csomagok használata esetén ajánlatos forgalomformázást használni, hiszen enélkül CBR esetben néhány csomag átvitele után (vagy rosszabb esetben az első csomagnál) veszteség keletkezik, így rendkívül hosszú várakozásra kényszerül a kapcsolat, mint az a 4. ábrán megfigyelhető. Ilyenkor tipikus ON/OFF jellegű viselkedés alakul ki, rövid, gyors löketekkel, és hosszú szünetekkel. VBR esetben sem megfelelő a korlátozás nélküli működés, hiszen ekkor is viszonylag hamar alakul ki veszteség. Korlátozást használó esetben nagy MBS esetén már használható hatékonyságot kapunk, ami főleg annak tudható be, hogy viszonylag hosszú idő telik el az első csomagvesztésig, hiszen idő kell, hogy a PCR sebességgel működő adó megszegje a forgalmi szerződést, így a TCP egyre rövidebb RTO értékekkel fog dolgozni, ami javítja a kihasználtságot. Sajnos ehhez az alkalmazkodáshoz viszonylag hosszú idő szükséges, ezért van az, hogy összességében elég rosszul működnek ezek a kapcsolatok VBR linken.

7. KÖVETKEZTETÉSEK

Ezek után a vizsgálatok után már viszonylag jó ajánlásokat tehetünk azok számára, akik ATM felett szeretnének IP alapú adatkommunikációt folytatni.

Amennyiben az adatkommunikáció multiplexált IP forgalom (például a bevezetőben említett Internet szolgáltató vagy a vállalati LAN-ok egymás közti forgalma), akkor a forgalmi jellemzőktől függően a következő megjegyzéseket tehetjük:

- Az Internet szolgáltató kliensei nagy valószínűséggel modem vagy ISDN feletti IP kapcsolattal rendelkeznek. Itt statisztikailag nagyobb biztonsággal lehet megjósolni az aggregált forgalmat, nem lesznek hirtelen kiugrások, minthogy egy kliens csekély sávszélesség felett rendelkezik. A szolgáltató a külvilággal a telephelyén telepített router-en keresztül tartja a kapcsolatot.
- Vállalati LAN-ok összekötése bonyolultabb probléma. Itt az összekapcsolandó belső hálózatok nagyon gyorsak is lehetnek, így sokkal komolyabb rövidtávú ingadozások lehetségesek.
- Mivel mindkét esetben a belső hálózat kifelé egyetlen eszközön tartja a kapcsolatot a külvilággal, illetve a többi telephely LAN-jaival, így a forgalomformázási funkciókat a külső hálózat felé kapcsolódó router-ben, vagy LAN emuláció esetén az ATM-kapcsoló kimenő interfésznél érdemes megvalósítani.
- Várhatóan az egyes állomások kifelé mutató forgalmi igényei nem egyidőben jelentkeznek, így a kimenő linken a terheltség középtávon (néhány perces időskálán) egyenletes lesz. Vállalatok számára azonban szükség lehet extra szolgáltatásokra is (pl. vállalati videokonferencia, nagymennyiségű adat gyors átvitele pl. adatbázis műveleteknél). Ekor a következő megoldás javasolható:
 - Amennyiben általában alacsony és középtávon egyenletes a forgalom, de rövidtávon (másodperces – tíz másodperces skálán) gyors és nagymértékű ingadozást mutat (például túlnyomóan HTTP használata, levelezés aggregált forgalma), akkor célszerű VBR kapcsolat alkalmazása viszonylag kis fenntartható sávszélesség és nagy maximális sávszélesség specifikálásával, maximális sebességre (PCR) beállított forgalomformázással, Ethernet kompatibilis MTU mérettel, és 100*1500/48 cellányi MBS mérettel. Ekkor viszonylag gyakran lesz cellavesztés, de mivel a forgalom karakterisztikája megfelel a VBR kategóriának, így olcsóbb azt alkalmazni, mint megpróbálni egy átlagot számolni, és annak megfelelő CBR kategóriát megfizetni.
 - Amennyiben a forgalom kiegyenlített (sok felhasználó esetén), vagy csak hosszú távon ingadozik (napszakonkénti ingadozások), akkor célszerű CBR kategóriát használni forgalomformázással, és Ethernet kompatibilis vagy 8232 byte-os MTU mérettel. Az utóbbi döntést az befolyásolhatja, hogy a belső hálózat Ethernet vagy ATM alapú-e.
 - Amennyiben láthatóan az átlagot jóval meghaladó, vagy más minőségi garanciákat követelő igény lép fel, célszerű erre az alkalomra egy külön SVC link

felépítése. Ilyen lehet az említett videokonferencia, ami ráadásul a késleltetést és annak ingadozását nem tolerálja, így forgalomformázás nem lehetséges. Ehhez kis MTU méret javasolt, és attól függően, hogy a jel állandó sebességű (pl. digitálissá alakított hagyományos videojel), vagy változó (pl. MPEG), CBR, ill. VBR kategóriát.

- Nagymennyiségű adat gyors átviteléhez is célszerű külön

SVC kapcsolatot felépíteni. Ehhez minél nagyobb MTU méret, forgalomformázás és CBR kategória használata javasolt. Erre jó példa a naponként automatikusan elvégzett adatmentés, akár banki adatbázisokat, akár a telefonközpontok számlázással kapcsolatos feljegyzéseit tekintjük. Ezeket a nagy adatátviteli igényeket azonban lehetséges úgy időzíteni, hogy elkerüljük a csúcsidőszakot, és a napszakonkénti ingadozásokat kihasználva akkor bonyolítjuk ezt a fajta forgalmat, amikor a sávszélesség nagy része egyébként is kihasználatlan. Ezzel egyenletesebbé tehetjük a hosszú távú forgalmi karakterisztikát, azaz kisímithatjuk a napszakonkénti ingadozásokat.

8. EGYÉB LEHETŐSÉGEK

Vizsgálatainkkal csupán a leglényegesebb TCP/IP és ATM paraméterek hatását vizsgáltam. Számos lehetőség van még, amivel lehetséges lenne növelni a kihasználtságot. Ezek a következők:

- Teljesértékű ATM forgalomformázás, amely nem csak a kapcsolat maximális sebességét tartja be, hanem az átlagos sebességet és a borsztösséget is kontrollálja. Ekkor a VBR szolgáltatói kategória által nyújtott flexibilitás és sávszélesség gazdaságosság teljes mértékben kihasználható a TCP/IP által is.
- RTO kezdőértékének kisebbre állítása: Legtöbb esetben a kezdeti veszteség utáni több másodperces várakozás teljesen szükségtelen, hiszen az ATM kapcsoló ennél sokkal gyorsabban feloldja a krízishelyzetet. Ezzel a csökkentéssel a nagyon rosszul viselkedő kapcsolatokon lehetne némi javítást eszközölni.
- A slow start algoritmusnál a kezdő ablakméret növelése: Kis MTU méret esetén lenne értelme, hogy a kapcsolat gyorsabban induljon.
- A fogadó ablakméretének állításával (receive window) meg lehet próbálni a kapcsolatot leszállítani egy maximális sebesség alá.
- Összefogott nyugtázás helyett szelektív nyugtázást használva kevésbé lenne hullámozó a TCP adási karakterisztikája, így talán kevesebb lenne a cellavesztés is.

A jövőben azonban az ATM technikát lenne célszerű úgy alakítani, hogy az IP-nek tökéletes közeget jelentsen. Ennek érdekében gerinchálózatot üzemeltetők nagyobb valószínűséggel képesek összehangolt lépéseket tenni, míg az IP ezirányú szervezett módosítása elterjedtsége és heterogén felhasználói tábora miatt nehezen képzelhető el.

IRODALOMJEGYZÉK

- [1] RFC791: Postel, J., Internet Protocol. September 1981.
- [2] RFC793: Postel, J., Transmission Control Protocol – DARPA Internet Program Protocol Specification, STD 7, DARPA, September 1981.
- [3] RFC2001: TCP Slow Start, Congestion Avoidance, Fast Retransmit, and Fast Recovery Algorithms, 1997.
- [4] RFC1180: A TCP/IP Tutorial, 1991.
- [5] Az ATM forum dokumentációs gyűjteménye, illetve FTP szervere:
<http://www.atmforum.com/atmforum/specs/approved.html> és
<ftp://ftp.atmforum.com/pub/approved-specs/>
- [6] RFC1577: Classical IP and ARP over ATM. M. Laubach. January 1994.
- [7] Sun SPARCstation 5 teljesítményjellemzői.
<http://www.sun.com/desktop/products/SS5/ss5.pdf>
- [8] ForeRunner™ SBA-200E ATM SBus Adapter User's Manual (MANU0069-01) April 1997. A FORE Systems, Inc. (www.fore.com) által az eszközhöz adott felhasználói dokumentáció.
- [9] LightStream 1010 ATM Switch Software Configuration Guide (Software Release 11.4(4)) A Cisco Systems, Inc. (www.cisco.com) által a kapcsolóhoz adott felhasználói dokumentáció.
- [10] AX4000 Network Analyzer User's Manual (Software Release 2.33.0) 1998. Az Adtech, Inc. (www.adtech-inc.com) által az eszközhöz adott felhasználói dokumentáció.
- [11] Netperf: A Network Performance Benchmark (Revision 2.1). Information Networks Division, Hewlett – Packard Company (www.hp.com), February, 1996.

TCP/IP PROTOCOL OPTIMIZATION OVER ATM

GERGELY PONGRÁCZ

The rapid advance of ATM in backbone networks and the massive position of TCP/IP in local networks predicts us a future when most of the data communication will use the IP protocol as a network layer protocol, and ATM technology as the transport network. This paper investigates the interworking of TCP/IP technology and ATM networks. We hope, that the results will help network engineers planning to build ATM based networks, because the operation of TCP/IP over ATM is not satisfactory at the present.

Why should we choose ATM? Why not stay with legacy networks?

The most serious problem with legacy networks (e.g. Ethernet) and TCP/IP is the lack of quality of service (QoS). It is cumbersome to differentiate between traffic sources of different priority, and solutions for service differentiation are proprietary, lacking a global standard. In the current, and widely used versions, there is no priority information in the frames and bandwidth reservation is impossible. Therefore a low priority batch job could destroy (or severely degrade) a very important videoconference session.

ATM was designed to integrate all types of connection services with completely different characteristics into a uniform, high-speed network. Because its uniformity, it has the ability to give certain QoS guarantees, and is able to do it in an easy and standardized way.

However TCP/IP in its current state is not capable of exploiting the QoS guarantees provided by ATM. As long, as this situation remains, ATM can not be a dominant network technology. Today TCP/IP has a lot of problems in working over ATM networks. If these problems can be solved, the last barrier that blocks the spreading of ATM technology will be moved away.

There are a lot of parameters in both TCP/IP and ATM, which affect the quality of the data transfer. In this paper the term quality refers to TCP/IP throughput and to TCP throughput divided by the theoretically available bandwidth (utilization). Some TCP/IP and ATM parameters can be configured easily, while others are nearly fixed. In this work we analysed the situation, where an ISP (Internet Service Provider) would like to connect to the nearest Internet peering point, or a medium sized company would like to interconnect its LANs. There is a telecommunication company (TC), which provides public services based on an ATM backbone. The TC can connect these users using ATM channels. When a customer subscribes a service (e.g. a leased line), the TC sets up a virtual ATM connection between the two endpoints.

The aim of the TC is to serve as many customers as he can, with the same resources, and do this in the least expensive way and without violating any of the QoS guarantees. The aim of the customer is to use the leased line with the best efficiency.

This paper inspects the achieved throughput of TCP/IP data transfer compared to the given upper bound coming from ATM layer. This is the TCP to ATM efficiency, and the user of the ATM line has to improve it as much as he can. Unfortunately in many cases TCP can only use a fraction of the available bandwidth. This efficiency can be increased by choosing the TCP and IP parameter values optimally and by using traffic shaping if possible.

However, not all parameters can be configured by the user. Some important ATM parameters can only be configured by the TC, thus it can be taken for granted by the user. The most important parameters, that can not be configured by the user, are the ATM traffic descriptors, like Peak Cell Rate (PCR) and Sustainable Cell Rate (SCR). In this case, we considered them fixed, and tried to reach the best efficiency in each case. These parameters mostly depend on the price of the connection, so changing them is expensive for the user. The only inspected parameter on the side of the TC is the Maximum Burst Size of the VBR class. It shows, how many cells can be delivered with PCR speed, and is used in traffic policing.

Our goal is to find easily configurable parameters, that can be adjusted at the user's side, and have a significant role in the efficiency of the connection. Finding such parameters would be a great help for ISPs and companies planning to connect LANs over ATM network, and for anyone using ATM to transfer TCP/IP traffic. Last, but not least, the perfect working of TCP/IP over ATM would solve TCP/IP's lack of QoS problem, and give us a completely universal transport protocol, that can deliver high-speed data with service differentiation.



Pongrácz Gergely 2000-ben diplomázott a Budapesti Műszaki Egyetem Villamosmérnöki Karán, Műszaki Informatika szakon. Jelenleg az Ericsson Traffic Analysis and Network Performance Laboratory-ban dolgozik, fő kutatási területe nagysebességű hálózatok teljesítménymérése és optimalizálása, valamint forgalomgenerálás és modellezés, különös tekintettel az ATM hálózatok teljesítményanalizálási problémáira.

DATA-COMPRESSION AIDING DATA-SECURITY

TIBOR NEMETZ*

RÉNYI MATHEMATICAL INSTITUTE, HUNGARIAN ACADEMY OF SCIENCE
NEMETZ@MATH-INSTHU

PÁL PAPP*

HUNGUARD LTD.
PAPP@HUNGUARD.HU

In this paper we deal with the effect of compression on the security of encryption from a theoretical and a practical point of view. Our first goal is to demonstrate that data-compression essentially improves the security-effect of encryption, helps to overcome technical difficulties, and cause lots of extra problems for the crypt-analysis. At present, applications of data-compression seem quite rare, and they even share some defects. We also propose a method to eliminate these negative effects.

An overview of the most frequently used cryptographic protocols and encryption software revealed that even the protocols are not supporting the combination of encryption with compression. We hope to be able to initiate a move towards the utilisation of this aid for data-security.

1. INTRODUCTION

Already during World War II. Shannon has observed the importance of text compression in secrecy systems, see [8]. His basic theory of information provided theoretical bounds on the achievable rates of compression, through the concepts of *entropy*, *redundancy of a source*. As an example, he has derived the entropy of printed English, [9]. The history has demonstrated that neglecting this fact may lead to fatal errors, like the one made by the Russians, see [1]. In this view it is rather surprising that the combination of data compression with encrypting is not every-day practice.

We suppose that the reader is familiar with the rudiments of information theory and cryptology, among them with the basic notions like source, alphabet, entropy, redundancy, codes, uniquely decodable codes, channel. No deeper knowledge is supposed. For the theory of cryptology, the reader is referred to [8], [5]. Familiarity with the notion and role of cryptographic keys would be useful.

Throughout this paper we refer to the original messages as *plain texts* (or clear text), the encrypted versions as *cryptotexts* (or cryptograms). Both the plain text and the cryptogram are strings of symbols from a finite set, called *alphabet*.

For the data compression literature we refer to [11], where appropriate further references are given. Our discussion may be followed, however, without knowing even the basic compression algorithms like the Huffman codes and the Lempel-Ziv type algorithms.

2. CRYPTOGRAPHIC SYSTEMS UNDER CONSIDERATION

In the "Information Society" the word data-security almost automatically relates to public key cryptography, in contrast to the traditional, secret key cryptography. This is justified by the open networks used for communication,

where eavesdropping is a common possibility. Nevertheless, a wide range exists between these two extremes, and the intermediate network types deserve special attention. In the present investigation we will therefore consider 3 types of network-systems:

- *Closed circuits* (networks, systems): Users (endpoints) of such systems are known, no new users are allowed to join the system on his/her own will. Typical examples are the diplomatic corps (embassies).
- *Open networks*, where anybody may join the system without asking special permission. The world of Internet serves as example.
- *Centrally controlled open networks*. In this case again anybody may join the system, but there is an authority that issues special means to the prospective users. E-commerce with obligatory certification belongs to this class.

The reason why we distinguish them is that they should be treated differently for solving security tasks. In the first case, single keys are used to encrypt a

$$\text{Unicity point} = \frac{\text{Entropy of the random key-variable}}{\text{Log (alphabet size) - laintext entropy}}$$

given message. Keys may be centrally generated and distributed, and kept secret. There exists a theoretically unsolvable secret code, the one time pad, see Shannon [8], or any basic book on cryptography. The one time pad is just one example. Essential is that the cryptographic algorithm and the underlying keys must be kept secret. An *absolute safe* channel should be used to interchange the keys.

For open networks, the usual way of ensuring secrecy is to publish a well-chosen "strong" common-to-all-users algorithm. Then all users generate two keys; one of them is used to encode, the other to decode messages. One of these keys is made public in a directory, which is accessible to anybody. Whoever wants to send him a secret message, encodes the open message with the public key and public algorithm. This idea is due to Diffie and Hellman, see [3]. An example is the RSA public key algorithm, see [7]. Here one can not speak about theoretically unbreakable codes; the *security is given by the lack of technical feasibility of breaking algorithms*.

In the case of controlled systems, *applicants must register* prior to joining the system. When registering, they may get special software and/or hardware that they are supposed to apply when initiating communication. This may contain some ID, special, individual algorithm for their own key generation, and special, individual secret key on *tamper-proof token* such that even the user can not learn. This gives a new dimension of security in comparison to public key systems, but allows the user to apply public key methods for exchanging keys on an *unsafe channel*. In

* Authors of the RITA System

this kind of systems usually RSA algorithm is combined by standard cryptographic procedures, like DES (3xDES), IDEA. It is essential for the present investigation that the controlled systems are established for special type of actions/transactions and therefore the communication apply *typical messages*, requests, answers.

3. ADVANTAGES OF USING DATA COMPRESSION

3.1. Closed networks

We analyse the three types of systems separately, starting with the one time pad. A short description of the method may be needed. We apply a known addition table (additive group) over the alphabet. For any plain text a key is chosen *randomly*, consisting of *independent, uniformly* distributed characters over the alphabet. Then, symbols-per-symbols one proceeds by adding the subsequent key symbols and plain text symbols. The arising sum gives the cryptogram. This encoding is theoretically unbreakable if the random key is used just once. Repeated use turns the system into an easily breakable one, as was demonstrated by the Venona-papers, [1], about USA cryptanalysis of Russian cryptograms.

The "one time" implies that the length of the random sequence is the same as that of the plain text. This random sequence should be delivered to the endpoints and there safely stored. Needless to explain that the shorter the random string the easier to overcome the difficulties. The only way to arrive at this goal goes by compressing the original text, by compression, which allows unique decoding. The larger the rate of compression the better. The question is what rate of compression can be attained. The answer highly depends on the type of the plain text. In the case of closed networks like in the diplomacy, the typical are the meaningful, written texts. In such cases the best rate of compression (in principle) is about 1/4 of the original length, see Shannon [9]. It should be noted, however, that such rates could be attained for very long texts, only. Practical cases call for texts of about 1000 characters in length, and then a compression rate of one half is a good achievement.

The advantage of preliminary text compression is obvious: *efforts* needed to generate random sequences, deliver them safely to the users, storing the key in a trusted media *are reduced to a half* as compared with the uncompressed case.

Due to technical difficulties, even in closed systems it is customary to use crypto-machines, finite cryptographic algorithms. Here the security is provided by an enormous key-space and the computational difficulty and the large complexity of possible cryptanalysis. The quality of such algorithms is measured — among others — by the *unicity point*, as detailed by Hellman, [5]. Here the *point* is a misleading word, since it gives a critical text length. If the plain text is shorter than that, then the cryptogram is even theoretically unbreakable, but there is no such assurance, if it is longer than that. The unicity point analysis assumes the usual information theoretical approach of specifying sources by probabilities. In this case one may speak about the entropy of a source, and the unicity point can

be defined formally as follows: Since the key is chosen randomly by uniform distribution, we have

Entropy of the key variable = log (Key-space size).

Suppose that the achievable compression rate is r , i.e. an uncompressed plain text string X_1, X_2, \dots, X_N of length N can be compressed into a sequence $Y_1, Y_2, \dots, \dots, Y_{rN}$ of length rN , $r \leq 1$. If the compressed sequence allows perfect reconstruction, then the uncertainty in the uncompressed string equals the uncertainty in the compressed sequence, yielding

$$H(X_1, X_2, \dots, X_N) = H(Y_1, Y_2, \dots, Y_{rN}).$$

where H (variable) denotes the entropy of the variable.

This gives for the rates

$$(1/N) * H(X_1, X_2, \dots, X_N) = r * (1/rN) * H(Y_1, Y_2, \dots, Y_{rN}).$$

The entropy rate for the uncompressed text equals r -times the entropy rate for the compressed text. This equality allows us to compare the unicity point for the uncompressed case to that of the compressed one:

$$\frac{U_{\text{comp}}}{U_{\text{cuncmp}}} = \frac{\log(|X|) - (1/r) * H_{\text{cuncmp}}}{\log(|X|) - H_{\text{cuncmp}}}$$

Here $|X|$, U_{comp} , U_{cuncmp} , H_{cuncmp} denote the alphabet size, the unicity point of compressed text, the unicity point of uncompressed text, entropy rate of uncompressed text, respectively. Depending on concrete values this fraction may be arbitrarily large. This can be interpreted as follows: approaching "perfect" compression the *length of plain text which can be encrypted safely by the same key may be very large*. The advantage is again obvious.

It is to be noted that Hellman's derivation is made for the attacks of "cipertext only". However, the same line can be followed for cases when some side information is available for the analyst. As an example, consider the attack, which utilises a list of probable text-fragments. The crypt-analyst shifts these fragments over the cryptotext, trying to match them, and, pretending that the match is perfect, solves corresponding equations for possible keys. Solutions are checked on the remaining part of the cryrogram. Now, *compression results in much shorter fragments*, which are typically shorter *than the unicity point*, even if the original is larger than that. Besides the same plain-text fragments will have different code fragments depending on their position in the text, *multiplying the list-size*.

In this case compression *increases the computational complexity* of testing if the received text is meaningful. While automatic test for the meaningful texts are straightforward, tests for compressed, statistically ruined texts are practically infeasible, see [4].

List-multiplication is crucial for the most effective methods of cryptanalysis, the differential and linear methods. Application of these methods requires the knowledge of a vast amount of matching plain-crypto pairs of blocks. Compression increases the difficulty of obtaining such pairs.

The same argument shows that the application of the chosen plaintext attack becomes also difficult. In this case the analyst artificially construct a plaintext, forces its encoding, thus obtaining a "sympathetic" plain-crypto text

pair. Trying to apply this method, he has to find a plaintext that would result in the desired text *after the compression*, which is a more difficult task.

We have to mention *some disadvantages when applying standard compression methods*. The reason lies in the application of obligatory compressed file-formats. For example, compressed files usually starts with a known string of

characters (and, possibly with other parameters, like the name of the file) which allows plain-crypto text matching, and, consequently known plaintext attacks. Fig. 1 and Fig. 2. show the structure of two popular compressing application. Bold text refer to parts, which may help the cryptoanalysis, by itself or together with some side information.

Structure of a gzip file (roughly, based on rfc 1952)
(Bold text refer to the parts which may help the cryptoanalysis)

A gzip file consists of a series of compressed data sets. Structure of one set:

ID1 (IDentification 1)=**31**
 ID2 (IDentification 2)=**139**
 CM (Compression Method)=**8**denotes the "deflate" (0-7 reserved)
 FLG (FLAGs) This flag byte is divided into individual bits. **Bits 5,6,7 are reserved, set to 0** :
 bit 0 FTEXT, bit 1 FHCRC, bit 2 FEXTRA, bit 3 FNAME, bit 4 FCOMMENT
 If FNAME is set, **an original file name is present terminated by 0**.
 If FCOMMENT is set, **a zero-terminated file comment is present**.
 MTIME (Modification TIME) in time_t UNIX format
 XFL (eXtra FLags) IF CM=8, then XFL = 2 max. compression. XFL = 4 - fastest algorithm
 OS (Operating System)
 0 - FAT ,1 - Amiga, 2 - VMS (or OpenVMS), 3 - Unix, 4 - VM/CMS, 5 - Atari, 6 - HPFS, 7 - Macintosh, 8 - Z-System, 9 - CP/M, 10 - TOPS-20, 11 - NTFS filesystem (NT),....(this value is easy to guess)
 XLEN (eXtra LENgth) If FLG.FEXTRA is set, this gives the length of the optional extra field.
 CRC32 (CRC-32)
ISIZE (Input SIZE)
 Extra field.....

Fig. 1. Structure of a gzip file (extract)

Structure of a zip file (extract)
(Bold text refer to the parts which may help the cryptoanalysis)

A zip file can contain several compressed files. Every file belongs to a local file header, and there is a central directory structure too. However we can't describe it in a detailed way due to lack of space. We highlight some elements which may help the cryptoanalysis.

<i>Local file header:</i>	Central directory
local file header signature (0x04034b50)	central file header signature (0x02014b50)
version needed to extract	version made by1
general purpose bit flag	0 - MS-DOS and OS/2,
compression method	1 - Amiga,
last mod file time	2 - VAX/VMS,
last mod file date	3 - Unix ,....
crc-32	11 thru 255 - unused
compressed size	..
uncompressed size	version needed to extract
filename length	compressed size
extra field length	uncompressed size
filename)	filename length
extra field	. 4 bytes
	filename (variable size)
	end of central dir signature (0x06054b50)

Fig. 2. Structure of a zip file (extract)

Very important special closed networks are the LANs. Even more specific is the channel between the user and his own data archives. It is well known, that the majority of computer crimes is committed by the personnel of companies. This naturally calls for encryption. Besides standard encryption schemes, here special, individual algorithms are also appropriate. Of course, whatever is true for the unicuity point in general, could be applied here. Especially, this means that the principle *first compress then encrypt* should be applied. We recommend, however, to use secret sharing methods about the keys. Let us recall, that secret sharing in cryptography means that a given secret is distributed among N parties in a way that any K of these N parties ($K < N$) together may reveal the secret, but no $K - 1$ or less parties could combine their knowledge to achieve any relevant part of the information.

3.2. Public networks

In the home-page-land the first goal of the owners is to distribute information about them, publicise their services, and attract attention. Therefore there is no need for hiding information there, and even data compression may make the access to the home page more difficult. Nevertheless, both cryptography and data compression is applied here, too. The owners have to defend themselves against hackers who may turn the content of the page into a garbage, and in such defends cryptographic methods are applicable.

On the other hand, files on home pages are usually rather long, and if someone wants to get them, then the time needed to it is important. Therefore together with html-files, it seems to be reasonable to store also their compressed versions as well. These compressed versions should utilise widely known algorithms, like Zip, Gzip, PKZip, ARC, LHarc, and ARJ. A complete reference about these algorithms is given in [11]. They are widely used on the Internet. Combination of cryptographic methods with compression is rare and we do not provide arguments in favour of such applications.

Another world is that of the e-mail correspondents. A large part of such connections is between people who know each other. We have made a small size opinion poll, and this shows that most of the e-mail users regularly apply compression techniques. Relatively small part of them combines compression with encryption algorithms like 3xDES, IDEA, or Blowfish. We have encountered cases, where the combination was in a bad order: first encryption then compression. No wonder, there were complaints about the lack of efficiency of compression.

In this area of users, the application of PGP and other commercially available plugins seem to get growing role. It is used to check the integrity of the message and the identity of the sender. Comments in this respect are outside of our main line.

3.3. Controlled open networks

The main feature of such networks that they bear high financial consequences (monetary, banking transaction, e-commerce, etc). Therefore the necessity of applying cryptographic tools is plain for everybody. Since these networks are essentially open, the building stones should be chosen according to certain standards, like the PKCS standards (Public Key Cryptography Standards) recommended by the RSA Laboratories. Connections are to be initiated by using PKI (Public Key Infrastructure) protocols. His calls for an overview of such protocols, checking if they offer compressions, either as optional or built-in. Such a short overview is given in the next section.

Of course, all advantages explained in connection with the closed systems, remain valid here, too. Here we would like to point out, that it is now possible to design and worth of designing task-oriented compression algorithms for the given purpose. The standard compression algorithms are of either statistical or universal kinds. In the present case, however, the messages are usually composed of typical sentences, requests, acknowledgements, etc. The designers of such PKI systems are strongly recommended to compile codebooks listing such typical expressions. This observation is rather old: shortly after Morse's invention of electronic telegraph, a "Secret Corresponding Vocabulary" was published for commercial use, see [10]. Our experience is that dictionary-look-up methods may result in compression rates 2-5 times over the general purpose ones, depending on the area of application.

Here we do not detail on the methods of constructing such vocabularies. Modifications of the Lempel-Ziv algorithm proved very useful for us. References [2] and [6] could be consulted.

4. SOME PROTOCOLS

In the previous section arguments were given showing the importance of data compression in increasing the degree of confidentiality. In this section we examine some widely used cryptographic protocols and techniques.

4.1. SSL

Netscape's Secure Socket Layer is one of the most widely used protocols ever. Latest version of this protocol is called Transport Layer Security protocol. However the differences between the SSL 3.0 and TLS is not important from our point of view.

The primary goal of the TLS Protocol is to provide confidentiality and data integrity between two communicating applications. Additionally, the protocol can provide authentication for the two parties.

The *typical application* is to provide confidential connection between a web server and a browser giving a "transport layer" for the http protocol. The protocol is above the transport level, the application can select between the secure transport layer provided by SSL, and the usual one. See Fig. 3.

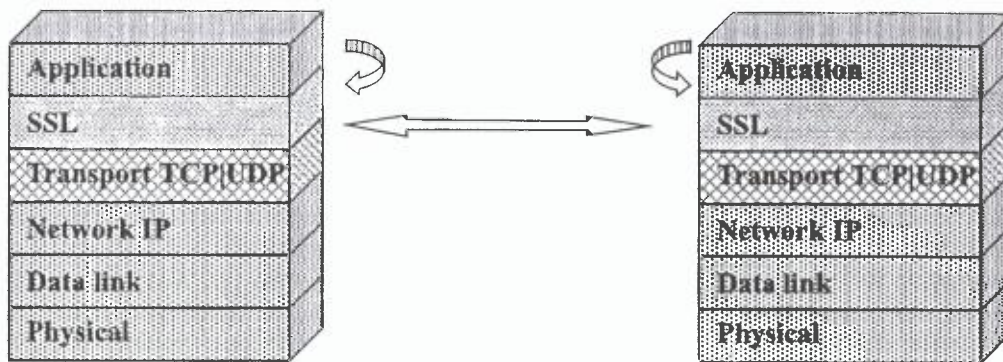


Fig. 3. SSL in the OSI model

Several encryption algorithms are supported with different key sizes due to the USA export regulations. These are: IDEA_CBC, RC2_CBC_40, RC4_40, RC4_128, DES40_CBC, DES_CBC, 3DES_EDE_CBC. It also supports the US Digital Signature Standard, the RSA, and the Diffie-Hellman system is used to perform digital signatures and key exchange.

The TSL protocol is *ready to use a compression* function. There is a compression phase between the plaintext and the ciphertext phase. It is a requirement that the compression must be lossless and may not increase the content length by more than 1024 bytes. The available compression methods are exchanged between the two parties during initialisation.

However *there is no standardised compression method*

in the TLS, so the usage of compression is an optional, theoretical possibility only.

4.2. SSH

SSH is a protocol for secure remote login and other secure network services over an insecure network. The SSH protocol consists of three major components: Transport layer protocol provides server authentication, confidentiality, and integrity with perfect forward secrecy. User authentication protocol, and connection protocol multiplexes the encrypted tunnel into several logical channels. This time we are interested in the SSH transport layer protocol which typically runs on top of TCP/IP. The protocol is above the transport level, the application can select between the secure transport layer provided by SSH and the usual one. See Fig. 4.

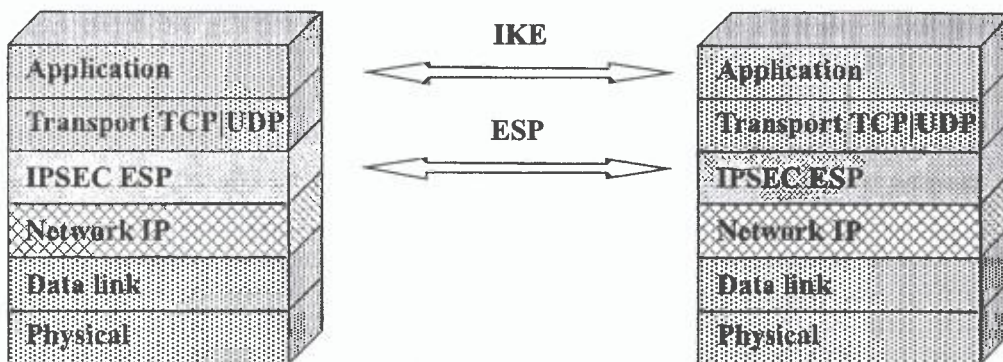


Fig. 4. IPsec protocols in the mirror of OSI layers

Several encryption algorithms are supported. However, only the Triple-Des algorithm is required to be implemented in cipher block chaining mode. The Blowfish and the Twofish is recommended to be implemented, while the Arcfour, IDEA, CAST128 are only optional. (Arcfour is believed to be RC4-compatible).

SSH uses Diffie-Hellman for key exchange providing perfect forward secrecy. The US Digital Signature Standard is used for signing. It is recommended to implement the widely used X509v3 certificate format, while the OpenPGP format is only optional.

Theoretically, it is possible to use compression in SSH. If compression has been negotiated, the payload field will be compressed using the negotiated algorithm. However, only one compression algorithm is named in SSH, the GNU zlib (lz77) described in RFC 1950 and in RFC 1951. The

main problem is that implementation of compression is not required, not recommended, only optional. That is, if an SSH implementation is negotiated with another SSH implementation, it is likely that one of them is not capable to compress packets, so compression won't be used.

4.3. IPSEC

IPsec is a set of protocols designed to provide interoperable, high quality, cryptographically based security for IPv4 and IPv6 packets. Nowadays the number of Virtual Private Network implementations is growing rapidly. There is a huge demand to use Internet to build secure connections between different companies or different departments of a company. IPSEC is expected to be the common denominator between different VPN solutions.

IPsec uses two protocols to provide traffic security.

Authentication Header (AH) and Encapsulating Security Payload (ESP). The IP Authentication Header (AH) provides connectionless integrity, data origin authentication, and an optional anti-replay service. The Encapsulating Security Payload (ESP) protocol may provide confidentiality (encryption), and limited traffic flow confidentiality. It also may provide connectionless integrity, data origin authentication, and an anti-replay service. The third protocol called IKE is responsible for key management. We focus on the ESP protocol with respect to compression.

IPSEC is above of the IP layer. This means that security features provided by IPSEC can be transparent for higher level protocols and applications. The IKE protocol is on the application layer.

ESP is *designed for use with symmetric encryption algorithms*. Because IP packets may arrive out of order, each packet must carry any data required to allow the receiver to establish cryptographic synchronisation for decryption. In practice the current key information comes from a table. Only the DES algorithm is required to implement for encryption. (Integrity is based on HMAC SHA and HMAC MD5)

There is *no compression in this protocol family*. On this network level each IP datagram should be compressed and decompressed by itself without any relation to other datagrams, as IP datagrams may arrive out of order or not arrive at all. (The SSH uses a stateful compression, i.e. there is no new compression table for a new packet, because SSH lies higher in the OSI layer, and TCP can guarantee the arrival of the packets.)

There is an IP payload compression protocol (RFC 2393) which provides compression for each packet. It is a good idea to use this in conjunction with IPSEC. However this protocol is new and rarely implemented.

4.4. OpenPGP

OpenPGP (RFC 2440) *provides confidentiality, digital signatures and authenticity* for messages between two communicating parties. OpenPGP software uses a combination of strong public-key and symmetric cryptography to provide these security services for electronic communications and data storage. OpenPGP is not a protocol (there is no online interactions between the two parties), but an application mainly used for file and message encryption and authentication.

REFERENCES

- [1] Benson, R. L., Warner, M. (eds): Venona. Aegean Park Press, 1996.
- [2] Cushing, J. A., Rikiso, Y.: Data Compression Algorithms for English Text Based on Word Frequency, Kobe University of Commerce, Institute of Economic Research, Kobe, 1990.
- [3] Diffie, W., Hellman, M.: New directions in cryptography. IEEE-IT 22 (1976), pp. 644-654.
- [4] Gerck, E., "Identification Theory", in two parts: <http://www.mcg.org.br/coherence.txt>, <http://www.mcg.org.br/coherence2.txt>
- [5] Hellman, M.: An Extension of the Shannon Theory Approach to Cryptography, IEEE-IT 23 (1977), pp. 289-294.
- [6] Horspool, N. R., Cormack, G. V.: Constructing Word-based Text Compression Algorithms, In: Proceeding of the

It MUST implement Triple DES, and SHOULD implement IDEA and CAST5. Implementations MAY implement other algorithms, like Blowfish (128 bit key 16 rounds), SAFER (13 rounds), ready for AES 128 bit, 192 bit, 256 bit key.

It supports the US Digital Signature Standard, the RSA, the ElGamal and the Diffie-Hellman system is used to perform digital signatures and key exchange. Ready to use Elliptic Curve systems (EC and DSA over EC).

OpenPGP *can use compression*. The message can be uncompressed, ZIP-compressed (RFC 1951) or ZLIB-compressed (RFC 1950). However it is not compulsory to implement these compression algorithms. According to the RFC, ZIP SHOULD be implemented and ZLIB MAY be implemented.

5. CONCLUDING REMARKS

We have provided arguments showing the advantages of data compression in increasing data security. An overview was given about usage of compression in today's popular cryptographic protocols. We found, that compression is very rarely used together with encryption, although some protocols are almost ready to use compression methods. The only drawback was the obligatory application of identifier-patterns. We may overcome this difficulty by applying a "smoothing transformation" after the compression, but before the encryption. This transformation may be implemented as a stream cipher with appropriate pseudo random key-stream. Technically it is not difficult to implement such an additional step. For example, one may choose a random seed as initial value, and use it to initiate the generation of the key-stream. This seed value should be included in the transmitted file.

Technically, a new function should be implemented in data compression applications, called "compress for encryption". This function would smooth the compressed text, making it ready for safer encryption. The generation may proceed according to the method described in the RSA standard PKCS#1, version 2. Here the generation is based on Hash functions. This may help uniformisation, since just a few Hash functions are generally accepted.

We should like to mention only, that such an application helps to prevent known plaintext attacks. The size of the paper, however, does not allow us to provide detailed discussion on these topics.

- 1992 Data Compression Conference, Storer, J. (ed), IEEE computer Society Press, 1992, pp. 62-71.
- [7] Rivest, R. L., Shamir, A., Adleman, L.: A method for obtaining digital signatures and public key cryptosystems, Comm. ACM, 21 (1978), pp. 120-126.
- [8] Shannon, C. E.: Communication Theory of Secrecy Systems, Bell System Techn. J., 28 (1949), pp. 656-715.
- [9] Shannon, C. E.: Prediction and Entropy of Printed English, Bell System Techn. J., 30 (1951), pp. 50-64.
- [10] Smith, O. J.: The Secret Corresponding Vocabulary: Adapted for Use to Morse's Electro-Magnetic Telegraph. Thurston, 1845.
- [11] Solomon, D.: Data Compression. Springer, 1998.

ON THE STABILITY OF TCP

TRINH ANH TUAN

BUDAPEST UNIVERSITY OF TECHNOLOGY AND ECONOMICS
DEPARTMENT OF TELECOMMUNICATIONS AND TELEMATICS
HIGH SPEED NETWORKS LABORATORY
H-1541 BUDAPEST, HUNGARY
E-MAIL: TUAN@TT-ATM.TTTBME.HU

TAMÁS ÉLTETŐ

ERICSSON TELECOMMUNICATIONS HUNGARY
TRAFFIC ANALYSIS AND NETWORK PERFORMANCE LABORATORY
H-1037 BUDAPEST, LABORC U. 1., HUNGARY
E-MAIL: TAMASELTETO@ETHERICSSON.SE

This paper discusses the stability of TCP. More precisely, we thoroughly examine the stationary behavior of the congestion window process. We measure real Internet traffic (bulk data transfer, FTP of large files) and process the data collected. An algorithm to approximate the actual congestion window will be presented. Statistical methods to analyze the data will be discussed in detail: change-point detection, distribution testing and parameter estimation. We show that the distribution of the congestion window process in stable period has bell-like shape and can be approximated by a normal distribution. We also present our interesting observations during the study and give possible explanation for them. Finally, we conclude the paper and discuss future works in building reliable and robust models for the TCP protocol.

1. INTRODUCTION

Stability is an emerging and important issue in science. It appears frequently in economics, physics, and generally speaking, in systems and processes that exhibit dynamical behavior.

The Internet with exponentially increasing number of users and trends to provide differentiated services [1], [2] is not an exception. Furthermore, a significant amount of today's Internet traffic, including WWW (HTTP), file transfer (FTP), email (SMTP), and remote access (telnet) traffic is carried out by the TCP transport protocol. TCP together with UDP form the very core of today's Internet transport layer. TCP is a window-based protocol. It uses its congestion window to control the sending rate at endpoints, and in consequence, together with queuing mechanisms at routers, to control the traffic over the Internet, [3], [7], [6]. Given the protocol, we want to predict the performance of the connections. This task, as it turned out, is not as simple as it would seem [4], [5].

In this paper we study the stationary behavior of TCP. Previous work that is closest to our study is the work by Teunis J. Ott, J. H. B. Kemperman and Matt Mathis in [9]. The authors also tried to find the stationary distribution of the congestion window process. The main difference between our study and their work is in the approaches. [9] used analytical approach and many assumptions to derive the stationary distribution. In other words, their approach is a theoretical approach. Our approach is an experimental approach. We measure real Internet traffic, process the data collected and use statistical methods to test it. Similar work related to throughput can be found in [10], [11], [12], [13].

The structure of this paper is the following. In Section 2 we discuss the architecture of our measurements over real Internet traffic. Here, we present an algorithm to approximate the actual congestion window of TCP. Section 3 introduces statistical methods to analyse stationary distri-

bution of the congestion window process. In Section 4, we present interesting observation during our study and give explanation for them. Section 5 concludes the paper and discusses future work in building reliable and robust models for TCP protocol.

2. MEASUREMENTS AND DATA PROCESSING

2.1. Architecture of the measurements

Our measurements was conducted over the real Internet. In order to do this, we need a clear measurement infrastructure. Such an infrastructure consisted of measurement platforms, which are dedicated work stations sited at different locations around the network. The platforms cooperate with one another to measure the properties of Internet paths as well as endpoints by exchanging test traffic among themselves.

Since our aim is to investigate the stationary behaviour of the TCP, the amount of data from measurements should be large to enable us to perform reliable data analysis. Internet measurements archive is also available publicly on the Internet, for example at Lawrence Berkeley Laboratory. This is the collection of the traces between nodes at the Laboratory and other nodes around the world. The data registered here only give us the information about the traffic passing the nodes at the Laboratory.

But for our specific purpose, we need to register the traffic at both endpoints, as explained later in data processing part. So we performed the measurements ourself. We did many measurement in various environments: one connection in LAN, many connections in LAN, one connection in WAN, many connections in WAN. The LAN environment in our case is the Local Area Network at Ericsson's Traffic Lab. The WAN environment is between Ericsson's Traffic Lab and the Technical University of Budapest. We also measured the traffic at different time of the day to capture the complex characteristics of the Internet traffic. The traffic we deal with is FTP of large files (bulk data transfer). The software to trace the traffic is the *tcpdump* program which is available in almost all Unix/Linux environments. Briefly, for each packet sent or received, *tcpdump* registers the time it is sent or received, it's sequence number, it's size, and some flags. We are interested mainly in the time stamp and the sequence number of the packets.

3. DATA PROCESSING

From the data collected, we want to compute the number of packets which have been sent at the sender but not yet arrived at the receiver as a function of time. This is approximately the congestion window size of the sender.

The algorithm:

1. *Synchronization.* First we need to synchronize the clocks at endpoints, since they show different time. We propose two methods. In the first method, we consider the case when two ends send each other SYN packets. Note that these packets don't contain data part, only header part. We assume that the time it takes to transfer these message is strictly smaller than the time it takes to send a data packet, and this time is approximately the same as of a PING packet. We used this time to synchronize the two clock. In the second method, at the sender we find the packet that have smallest round-trip time, we mark the time this packet arrive at the receiver and the time the sender send acknowledgment for this packet. We synchronize the two clocks by the middle of these two intervals.
2. *Marking packets.* After synchronization, the time a packet is sent at the sender is marked with +1 due to the fact that the number of packets outgoing is increased by one at that time. Similarly, the time a packet has arrived at the receiver is marked with -1.
3. *Summation.* We sum up all the +1 and -1 until a given time. The number of packets that has been sent but not yet received at a given time is equal to the sum.

First, we thought that this is already a correct algorithm. Later, we observed the "jump-ups" in the results. This is due to the losses. Once a packet is lost, sooner or later it will leave the line. For this packet, we only add +1 when sending it but never have the chance to mark -1 for it (since it does not arrive to the receiver). We don't know exactly when it's dropped from the line, but we know the upper bound for this. We assume that when the *next* packet sent arrives at the receiver, due to *pipe* nature of the connection, the lost packet was already dropped from the line. So we deliberately add -1 to the sum at that time. This completes our algorithm.

4. STATISTICAL ANALYSIS

4.1. Kolmogorov-Smirnov tests

4.1.1. One sample test

The one sample Kolmogorov-Smirnov [15] helps us to test the hypothesis that a X random variable's distribution is the same with a given continuous distribution.

$$H_0 : F(x) = P(X < x) = F_0(x)$$

The empirical distribution $\hat{F}_n(x)$ is constructed from (X_1, X_2, \dots, X_n) independent samples. The statics is :

$$D_n = \max_x |\hat{F}_n(x) - F_0(x)|$$

The limit distribution theorem states that:

$$\begin{aligned} \lim_{n \rightarrow \infty} P(\sqrt{n}D_n < y) &= \\ &= K(y), \quad 0 < y < \infty, \end{aligned}$$

where $K(y)$ is the Kolmogorov distribution.

4.1.2. Two sample test

The most important tool used by theoretical change-point detection is the (two sample) Kolmogorov-Smirnov test (see [15]). This section describes how this test works,

and also tells about the special case of equal sample series length.

Let

$$X_1, X_2, \dots, X_n$$

be the series of the first sample (of length n), and

$$Y_1, Y_2, \dots, Y_m$$

be the series of the second . Note that the two series are independent.

Furthermore, let $F_n(x)$ and $G_m(x)$ be the corresponding empirical distributions. Then the statistics is

$$D_{n,m} = \max_x |\hat{F}_n(x) - \hat{G}_m(x)|$$

The limit distribution theorem states that

$$\begin{aligned} \lim_{n,m \rightarrow \infty} P\left(\sqrt{\frac{nm}{n+m}} D_{n,m} < y\right) &= \\ &= K(y), \quad 0 < y < \infty, \end{aligned}$$

where $K(y)$ is the Kolmogorov distribution.

Assume that $n = m$. In this case $D_{n,n}$ is computed practically in the following way. By ordering the merged samples in an increasing order we have

$$Z_1^{(2n)} \leq Z_2^{(2n)} \leq \dots \leq Z_{2n}^{(2n)}.$$

Now, let

$$\delta_i = \begin{cases} 1, & \text{if } Z_i^{(2n)} \text{ comes from the series of } X_k \\ -1, & \text{otherwise} \end{cases}$$

and

$$S_j = \delta_1 + \delta_2 + \dots + \delta_j.$$

Here we get $D_{n,n}$ so that

$$D_{n,n} = \frac{1}{2n} \max_{1 \leq j \leq 2n} |S_j|.$$

This method computes a distance between two empirical distributions using the original data. The method was described assuming an equal number of observations but this is not the general case. The value of $D_{n,n}$ related to the difference between the originating distributions. The larger its value, the larger the difference.

4.2. Change-point detection

In order to investigate the stationary behaviour of TCP, we need to find an interval within which statistical behaviour is stable. We use the method proposed by T. Éltető and S. Molnár in [8] . We consider the data as a series of disjoint intervals within which our series can be regarded as stationary. When an interval ends and a new interval starts the characteristics of the data change. Here we concentrate on a method of finding these change-points. Our purpose is to find the edges of the intervals where the distribution remains the same [16], [14] . In order to determine these intervals the empirical distributions of two neighboring time windows are compared and these two windows are slid together along the series. In this way, we get another series, what we call test-series. (The comparison was done by two samples Kolmogorov-Smirnov test of the subsamples.).

This algorithm was applied to find possible stable intervals which are appropriate for examination of stability of TCP.

4.3. Distribution testing

The histogram of the congestion window visually suggests normal distribution.

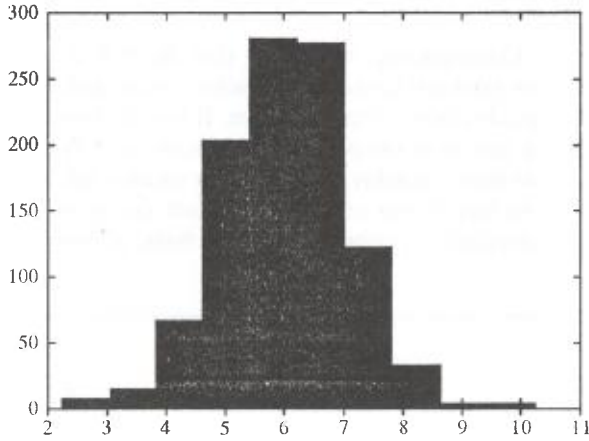


Fig. 1. A histogram of the congestion window

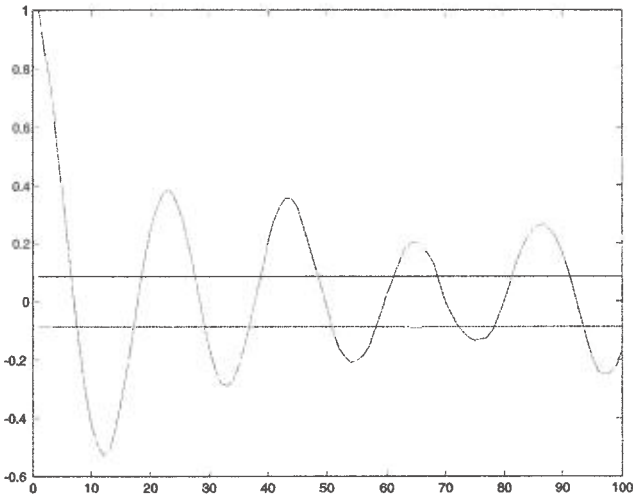


Fig. 2. Q-Q plot with a truncated normal distribution

We used Q-Q plot to test the distribution. The points roughly follow the line $u(x) = x$, that is a line with slope 1. This indicates that the distribution is similar to truncated normal distribution.

Table 1. One-Sample Kolmogorov–Smirnov Test

		X
N		151
Normal Param.'s	Mean	41.47
	Std. Dev.	6.66
Most Extreme Diff.'s	Abs.	.0071
	Pos.	.071
	Neg.	-.055
Kolmogorov–Smirnov Z		.875
Asymp. Significance (2-tailed)		.428

As we see in the table, the asymptotic significance of 0.428 suggests that it is hard to reject the null-hypothesis. This neither says that the distribution is

normal distribution, but strongly indicates normality.. Our conclusion for this part is that, the stationary distribution of congestion window process has the shape very similar to the shape of a truncated normal distribution.

4.4. Parameter estimation

The normal distribution has no limit below, however, the congestion window is never a negative number, so we have here a truncated normal distribution. In practice, when the congestion window process gets it stationary state, the lower limit is hardly 0. So, for generality, we discuss here a truncated normal distribution with lower limit c where $c \geq 0$. Assume a normally distributed random variable

$$Y \sim N(\mu, \sigma^2).$$

We generate a new random variable X from Y by setting

$$X = \begin{cases} c, & Y \leq c \\ Y, & Y > c \end{cases}$$

with a fixed c .

Now we estimate the parameters μ , σ and c .

Let the quantile be α such that

$$\alpha = \Phi\left(\frac{c - \mu}{\sigma}\right), \quad (1)$$

here the function $\Phi(\cdot)$ denotes the standard normal distribution function.

Again we assume that the originating distribution is Normal(μ, σ^2).

$$\begin{aligned} E(X|X > c) &= \\ &= \int_c^{+\infty} t \frac{1}{\sqrt{2\pi}\sigma(1-\alpha)} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt = \\ &= \frac{1}{1-\alpha} \int_{\frac{c-\mu}{\sigma}}^{+\infty} (\sigma u + \mu) \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du = \\ &= \frac{\sigma}{1-\alpha} \int_{\frac{c-\mu}{\sigma}}^{+\infty} u \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du + \\ &+ \frac{\mu}{1-\alpha} \int_{\frac{c-\mu}{\sigma}}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du = \\ &= \frac{\sigma}{1-\alpha} \frac{1}{\sqrt{2\pi}} e^{-\frac{(c-\mu)^2}{2\sigma^2}} + \\ &+ \frac{\mu}{1-\alpha} \left(1 - \Phi\left(\frac{c-\mu}{\sigma}\right)\right). \end{aligned} \quad (2)$$

If we define

$$\begin{aligned} a &= \frac{1}{\sqrt{2\pi}(1-\alpha)} e^{-\frac{(c-\mu)^2}{2\sigma^2}} = \\ &= \frac{1}{\sqrt{2\pi}(1-\alpha)} e^{-\frac{(\Phi^{-1}(\alpha))^2}{2}}, \end{aligned}$$

as a function of α .

By this definition and from (2) we get (3):

$$E(X|X > c) = \sigma a + \mu, \quad (3)$$

$$\mu = c - \sigma \Phi^{-1}(\alpha). \quad (4)$$

From (3) and (4) we obtain the following formula for σ

$$\sigma = \frac{E(X|X > c) - c}{a - \Phi^{-1}(\alpha)}, \quad (5)$$

and substituting σ in (4) we get μ .

Finally, we estimate α and $E(X|X > c)$ such that

$$\hat{\alpha} = \frac{1}{n} \sum_{i=1}^n \chi(X_i < c), \quad (6)$$

$$E(X|X > c) = \hat{E}$$

$$= \frac{1}{|\{1 \leq i \leq n : X_i > c\}|} \sum_{i=1}^n X_i \chi(X_i > c). \quad (7)$$

So we have estimated σ , and μ by (3), and (4), such that the estimates $\hat{\sigma}$ and $\hat{\mu}$ use the estimates $\hat{\alpha}$ and \hat{E} .

5. OBSERVATION

5.1. Synchronization of processes

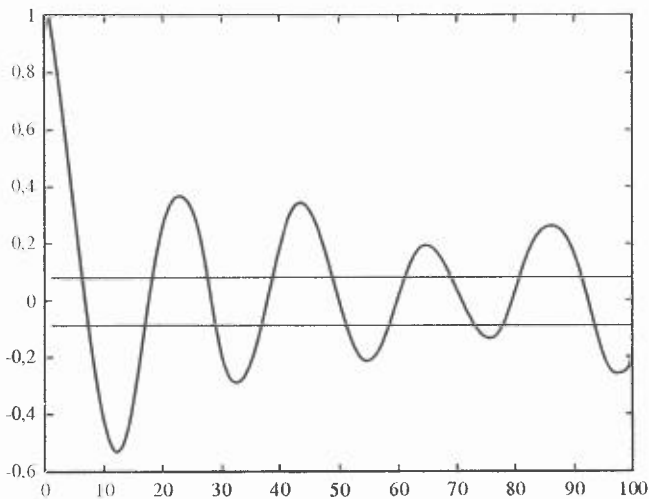


Fig. 3. The auto-correlation function of the sum

One of our experiment shows very interesting property: all of the 8 connections are synchronized, that is, the congestion windows go up and down together resulting in the sinus form of auto-correlation function of the sum of 8 connections. Later, we know that this phenomenon is not new. A possible explanation for this phenomena is the following. The packets from different connections form "clusters". One cluster contains packets from many connections. When the packets arrive at the router which employs the drop-tail queuing mechanism and is currently full, it drops all packets from a burst, and each packet dropped from connections cause the congestion window of that connections be reduced simultaneously.

5.2. On packet loss process

Our observation shows that individual packet loss probabilities are usually not sufficient for proper understanding of the system behaviour. In general, packets are more likely to be rejected because of buffer overflow as their rate of arrival to the buffer increases. Also, if a packet has been rejected, then it is more likely that consecutive

packets will be rejected. Practically, this phenomena can be explained by bursty nature of data traffic (i.e. packets arrive in burst) and drop-tail queuing mechanism that is employed by almost all routers. Theoretically, it can be shown that for the $M/M/1/K$ system for example, this is true that for all $K \geq 2$, i.e., the conditional probability of rejection given a prior rejection is strictly higher than the unconditional probability of rejection (even for arbitrary arrival). Consequently, it is clear that there is a strong correlation between consecutive packet losses, and therefore losses also exhibit bursty nature. It is well-known that the bursty nature of the packet loss process can effectively reduce the service quality, especially for sources which are sensitive to long bursts of losses. We can also recall here the synchronization phenomena due to burst of losses.

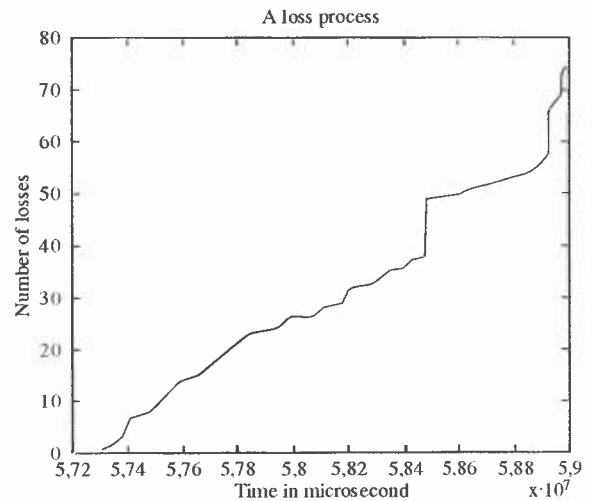


Fig. 4. Sum of losses

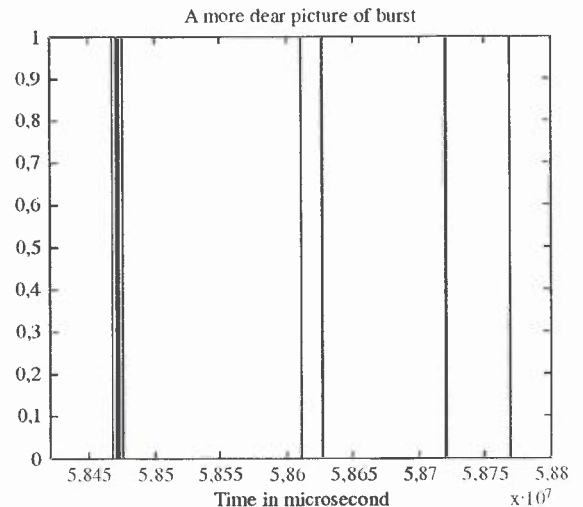


Fig. 5. A bursty period

6. CONCLUSION

In this paper, we have presented a measurement-base approach to analyse the stability of TCP. We showed that the distribution of TCP's congestion window process in stable period has bell-like shape and can be approximated by a normal distribution. Statistical methods to analyse

the data have been discussed in detailed. We showed interesting observations during our study and gave possible explanation for them. However, these steps are only initial steps in our approach. Many questions are still open. We do not know yet about the exact behaviour of the loss process, and therefore neither the congestion window process. Assumptions are always needed to ease the derivation. Our near future work is to investigate the nature of the loss process. TCP performance research over various environments is still in its initial state. This

REFERENCES

- [1] S. Schenker. Fundamental design Issues of the future Internet, *IEEE Journal on Selected Areas in Communications* 13, 1141-1149, 1995.
- [2] V. P. Kumar, T. V. Lakshman, D. Stiliadis: Beyond best effort: Router architectures for differentiated services of tomorrow Internet. *IEEE Communication Magazin*, Vol 36, No 5, May 1998.
- [3] R. Stevens, *TCP/IP Illustrated*, Volume 1, Addison-Wesley, MA Reading, 1997.
- [4] V. Paxson and S. Floyd. Wide-Area traffic: The Failure of Poisson Modelling, *IEEE/ACM Transactions on Networking*, 1995.
- [5] W. E. Leland, Murad S. Taqqu, Walter Willinger and D. Wilson: On the self-similar Nature of Ethernet traffic, *IEEE/ACM Transactions on Networking*, 1994.
- [6] D. Chiu and R. Jain: Analysis of the Increase/Decrease Algorithms in Computer Networks, *Journal of Computer Networks and ISDN*, 1989.
- [7] V. Jacobson: Congestion avoidance and control, *Proceedings of ACM SIGCOMM '88*, August 1988.
- [8] T. Éltető, S. Molnár: On the distribution of round-trip delays in TCP/IP networks, *International Conference on Local Computer Network*, 1999.

also motivates us on our future research. Finally, our ultimate goal is to build reliable and robust models for TCP protocols as well other congestion control schemes in telecommunication networks.

7. ACKNOWLEDGEMENT

We would like to express our thanks to Dr. László Györfi for his guidance and useful feedbacks. We also thank for the help and support from the High Speed Network Laboratory, Technical University of Budapest.

- [9] T. J. Ott, M. Mathis and J. H. B. Kemperman: The stationary behavior of ideal TCP congestion avoidance. <ftp://ftp.bellcore.com/pub/tjo/TCPwindow.ps>, 1996.
- [10] M. Matthijs, J. Semke, J. Mahdavi, T. J. Ott: The macroscopic behavior of the TCP Congestion Avoidance Algorithm, *Computer Communication Review*, July 1997.
- [11] J. Padhye, V. Firoiu, D. Towsley and J. Kurose: Modeling TCP throughput: a Simple Model and its Empirical Validation, *Proceedings of SIGCOMM '98*, September 1998.
- [12] A. Kumar: Comparative Performance Analysis of Versions of TCP in *IEEE/ACM Transactions on Networking*, August 1998.
- [13] T. V. Lakshman and U. Madhow: The performance of TCP/IP for networks with high bandwidth-delay products and random loss, *IEEE/ACM Transactions on Networking*, June 1997.
- [14] E. Parzen, "Nonparametric Statistical Data Modeling", *Journal of the American Statistical Association*, March 1979, 105-126.
- [15] M. Kendall, A. Stewart, *The Advanced Theory of Statistics*, Charles Griffin London, 1966.
- [16] M. Csörgő and L. Horváth, "Invariance Principles for Change-point Problems", *Journal of Multivariate Analysis*, 27, pp. 151-168.

A TCP PROTOKOLL STABILITÁSA

Cikkünk a TCP protokoll stabilitási vizsgálatával foglalkozik. Megvizsgáljuk a torlódási ablak (congestion window) folyamat stacionárius viselkedését. Valós Internet forgalmat (FTP) mérünk és feldolgozzuk a kapott adatokat. Javasolunk egy algoritmust, amelynek a segítségével a torlódási ablakot ki lehet számítani a mért adatokból. Részletesen tárgyaljuk azokat a statisztikai módszereket, amelyeket a vizsgálatok során alkalmaztunk. A vizsgálataink azt mutatják, hogy stabil szakaszban a torlódási ablak eloszlása egy haranggörbe és egy normális eloszlással approximálható. Megmutatjuk érdekes megfigyeléseinket, amelyeket a vizsgálatok során nyertünk, és lehetséges magyarázatokat is adunk hozzá. Végül összefoglaljuk a cikket, és tárgyaljuk jövőbeni munkánkat.



Trinh Anh Tuan is currently a 5-th year student at Faculty of Electrical Engineering and Informatics, Technical University of Budapest. This article is a revised version of his paper at Student Scientific Conference at Technical University of Budapest, 1999. The paper was awarded 2-nd prize. His main interests are performance evaluation and modeling of TCP, packet loss processes in IP networks and on-line estimation in time series.

GYORSULÁSMÉRŐK ALKALMAZÁSI LEHETŐSÉGEINEK KUTATÁSA

CSER LÁSZLÓ V. évf. villamosmérnök hallgató

BUDAPESTI MŰSZAKI EGYETEM
IRÁNYÍTÁSTECHNIKA ÉS INFORMATIKA TANSZÉK
1111 BUDAPEST, PÁZMÁNY PÉTER SÉTÁNY 1/D

A szenzorok egyre tökéletesebb gyártástechnológiájának köszönhetően számos, korábban nehézkesen mérhető mennyiség válik könnyebben mérhetővé. A gyorsulás az egyik olyan mennyiség, amelynek méréstechnikai fejlődését a gépjármű-ipar is ösztönzi, számos alkalmazási területet nyitva meg a tömegcikké váló gyorsulásmérők előtt. Az olcsó gyorsulásmérők alkalmazási lehetőségei a gépjármű-ipari alkalmazásokon túl felölelik a navigáció, térképészet, szeizmológia, kézírás alapú beviteli perifériák stb. területeit. Ebben a munkában a gyakorlatban használatos rezgés- és gyorsulásmérési módszerek áttekintése után egy tervezési feladaton keresztül mutatjuk be az ígéretes alkalmazási lehetőségeket kínáló Analog Devices gyorsulásmérő család ADXL05 modelljét. Az ADXL érzékelők különlegessége abban áll, hogy egyetlen szilíciumszeleten kialakítva tartalmazzák a gyorsulásszenzort és jelkondicionáló elektronikát. A tervezési feladat egy mozgó kerékpáron elhelyezett, ADXL05-ön alapuló műszer tervezése. A műszernek meg kell tudni mérnie, hogy a kerékpár a haladási irányában éppen emelkedőn halad felfelé vagy lejtőn halad lefelé, és egy szögértéket is szolgáltatnia kell, amely jellemzést ad az út meredekségéről. Először a mérési módszert tervezzük meg, amelyhez általános esetben is meg kell tudni válaszolnunk azt a kérdést, hogy pontosan milyen mennyiséget mér egy ideális gyorsulásmérő. Ha ezt megválasztottuk, akkor megmutatjuk, hogy az ideális gyorsulásmérőhöz képest egy valóságos eszköz miben tér el. A mérési módszer megtervezése után annak lehetséges hibáit is feltárjuk, majd a feladatban meghatározott követelmények, az eszközválasztásból adódó kényszerek és a mérési módszer kényszerei alapján egyre jobban megközelítjük a végső megoldást. A műszer tervét az egyes funkcionális egységek működése alapján tekintjük át. A terv egyes komponenseinek kidolgozása során helyenként indokoltnak véltük az apró részletekbe menő ismertetést, míg más helyeken inkább az elemző, rendszerező közelítésmódot alkalmaztuk.

1. BEVEZETÉS

A szenzorok egyre tökéletesebb gyártástechnológiájának köszönhetően számos, korábban nehézkesen mérhető mennyiség válik könnyebben mérhetővé. A gyorsulás az egyik olyan mennyiség, amelynek méréstechnikai fejlődését a gépjármű-ipar is ösztönzi, számos alkalmazási területet nyitva meg a tömegcikké váló gyorsulásmérők előtt.

Nemrég a gyorsulásmérőket emlegetve szinte kizárólag a piezoelektromos átalakítókra gondoltunk, amelyek a rajtuk fellépő gyorsulásnak csak a tranziens komponensét voltak képesek mérni. Az Analog Devices által gyártott ADXL sorozatú gyorsulásmérők azonban a gyorsulásjel stacionárius komponensét is mérni tudják. Igen nagy jelentőségű a $\Gamma \subset$ komponens mérése, hiszen ezek a gyorsulásmérők felhasználhatók szöghelyzet-érzékelésre, kiváltva a folyadékos, úszós pozíció- és szöghelyzet-érzékelőket. A megnyíló alkalmazási területek magukba foglalják az egyszerű vízmértéket, olajfűrófejek pozícionáló szenzorait, autóriasztókat, gépkocsik biztonsági-öv előfeszítő mechanizmusát, valamint a navigáció, földrengés-előrejelzés, számítógépes beviteli eszközök és perifériák, rezgésanalízis területeit, továbbá megkönnyítik gépek, mozgó, forgó alkatrészek tervszerű megelőző karbantartását stb.

Az alábbiakban röviden áttekintünk néhány alkalmazási lehetőséget.

2. NAVIGÁCIÓS ALKALMAZÁSOK

Navigáció alatt azon eljárások halmazát értjük, amelyek segítségével valamely jármű vagy objektum helyzetét és útvonalát meghatározzák, továbbá a járművet irányítják. Napjaink navigációs berendezései elsősorban GPS-vevőt

használnak elsődleges információforrásként. A GPS rádió-kommunikációs rendszer lévén ki van téve a csatornainagadozásnak, interferenciának. Számos tanulmány rámutat, hogy a GPS alapú helymeghatározás városi környezetben, erdőben megbízhatatlanná válhat, alagútban pedig egyáltalán nem lehetséges. Ugyanakkor az is ismert, hogy polgári célú alkalmazásokban a GPS által végzett helymeghatározás eredménye védelmi okokból szándékosan torzítva van (előreláthatólag 2005-ig). A többutas jelterjedés következtében a műholdról vett jelek a domborzatról és a környező magas épületekről visszaverődnek, pontatlan távolságokat eredményezve a helymeghatározásban.

A navigációs berendezések a GPS-vétel kiesésének idejére általában a dead-reckoning (kb. becsléssel történő helymeghatározás) módszert használják. Az eljárás segítségével a pozíciót a navigálandó objektum kezdőkoordinátáinak, mozgási irányának és sebességének ismeretében határozhatjuk meg. A kezdőkoordinátákat a vételkiesés előtti utolsó GPS mérés szolgáltatja. A mozgás irányát (azimut) általában elektronikus iránytűvel, vagy giroszkóppal határozzák meg, míg a sebességet a jármű mért gyorsulásának integrálásával kaphatjuk meg. A gyorsulásmérők alkalmazhatók még az irányinformáció meghatározásánál a giroszkópok inklinációjának mérésére.

Az előbb említett navigációs eljárásokat a térinformatikában; földi mobil térképező rendszerekben is használják. Itt a mozgó járműre erősített érzékelőkből álló műszergyűttes képes a jármű pályájának folyamatos meghatározására, s egyben a szenzorok fajtájától függő kiegészítő információ gyűjtésére. A begyűjtött adatok térinformatikai adatbázisba kerülnek.

3. INTELLIGENS BEVITELI PERIFÉRIÁK

Érdekes kutatás folyt az Arizona State University-n automatikus aláírásfelismerés témakörében. A kutatás célja a HMM-en (Hidden Markov Model) alapuló kézírásfelismerő rendszer implementációja volt, amely megkülönbözteti a hiteles aláírást a hamisítványtól. A beviteli eszköz egy egyszerű golyóstoll, amelynek végére gyorsulásmérő(ke)t helyeztek. A begyűjtött adatok alapján a kiértékelő szoftver megkeresi az eredeti aláírás jellemző vonásait, és később ezekkel hasonlítja össze az ellenőrizendő aláírást. Amennyiben megfelelően robusztus lesz az eljárás, sikerre számíthat a kereskedelmi és banki alkalmazásokban.

A mobil telefonok, PDA-k (Personal Digital Assistant) viszonylag kis méretű kijelzővel rendelkeznek. Amikor a kijelző méreténél nagyobb kiterjedésű információt kell megjeleníteni, általánosan használt eljárás a görgetés (scroll). Ezt általában gombok, görgők kezelésével tehetjük meg. A gyorsulásmérők itt is teret nyerhetnek, hiszen a képernyőtartalom mozgatásához elég lehet kézben tartva elmozdítani az eszközt a kívánt irányú görgetéssel ellentétes irányban, kikerülve a gombkombinációk megtanulásának nehézségeit. Ezzel sokkal egyszerűbb felhatalmazói felület valósulhat meg.

4. EGYÉB ALKALMAZÁSOK

Akár életeket is menthet a korszerű szenzorok és a mobiltelefonia integrációja pl. a közlekedésben. Ütközés esetén az extrém nagy gyorsulásértéket regisztráló szenzor utasítja a mobiltelefont egy előre beprogramozott segítség-szám felhívására, lerövidítve a balesetet szenvedők elsősegélyben részesüléséig eltelt időt. Az ütközés érzékelésére felhasznált szenzor beépülhet akár a mobil készülékbe (pl. az említett scroll miatt már egyébként is be van építve), vagy a gépkocsi buszrendszerére illetve összeköttetésben állhat az autó légzsákvezérlő elektronikájával. Amennyiben a szenzor beépül a mobil készülékbe, a véletlen leejtés hatására bekövetkező hamis riasztás ellen védekezni kell. Ez történhet úgy, hogy a leejtés után a készülék hangjelzésére a felhasználónak adott időn belül gombnyomással kell reagálnia. Amennyiben ez nem történik meg adott időn belül, megtörténik a riasztás.

A garanciális kérdések érvényesítésénél gyakran vitás kérdés a rendeltetészerű használat. A szenzorokkal ellátott készülék regisztrálhatja egyes paraméterek bizonyos küszöbérték feletti szintjét (pl. gyorsulás, hőmérséklet stb.). Amennyiben a garancia érvényesítésére kerülne sor, a hivatalos márkaszerviz ellenőrizhetné, hogy a készüléket valóban a rendeltetésének megfelelően, az ajánlott működtetési feltételekkel használták, és csak a jogos esetben végeznék el a garanciális javítást.

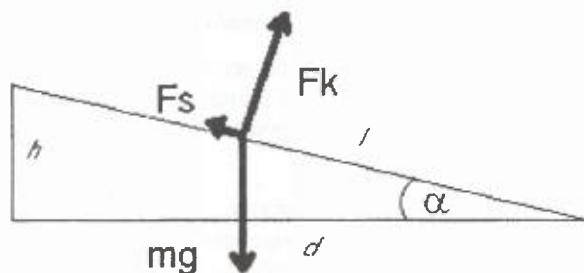
5. SPORT

Korunkban a hivatásos sportolókat edzések, felkészülések során műszerek egész arzenálja segíti: stopper, ódométer, magasságmérő, sebességmérő és számos egyéb menetparaméter regisztrálását végző eszköz. Jól tudjuk, hogy a sportban üzlet van, és a szponzor eredményt vár a befektetésért, így nem csoda, ha a kiélezett versenyben ilyen sokféle technikai segédeszközre van szükség.

A szélesebb körű alkalmazási kitekintés után most megvizsgáljuk a gyorsulásmérőkkel történő szöghelyzetérzékelés alapelveit. Valós sportolói igény alapján kezdtük kutatni azokat a módszereket, amelyek segítségével közúti járműből menet közben mérhető az út pillanatnyi meredeksége a menetirányban. Igyekezünk a gyakorlatban jól alkalmazható eljárásokat keresni.

6. A LEJTŐ FIZIKÁJA – POTENCIÁLIS MÉRÉSI MÓDSZEREK

Vizsgálódás elején a jármű által befutott utat úgy fogjuk fel, mint egy folyamatosan változó hajlásszögű lejtőt. Cél az adott pillanatban „rögzített” lejtő szögének előjelhelyes meghatározása méréssel. (Az előjel természetesen attól függ, hogy az út lejt-e vagy emelkedik.) Az 1. ábrán pontszerű testre ható erőket láthatjuk, rögzített lejtőn.



1. ábra. Pontszerű testre ható erők rögzített lejtőn. h : magasság, d : alap, l : hossz, $F_G = mg$: nehézségi erő, $F_K = mg \cos(\alpha)$: kényszererő, $F_s = -\mu mg \cos(\alpha)$: surlódási erő

Newton II. törvényét alkalmazva, az erők eredőjére felírható:

$$F_l^* = m\{g[\sin(\alpha) - \mu \cos(\alpha)] + a^*\} \quad (1)$$

A testre a lejtőn F_l hatásvonalában még egyéb (F^*) erő is hat (pl. légellenállás, testre ható gyorsító erő stb. összege) és az általa létrehozott gyorsulás a^* -gal vehető figyelembe.

7. GYORSULÁSMÉRŐK

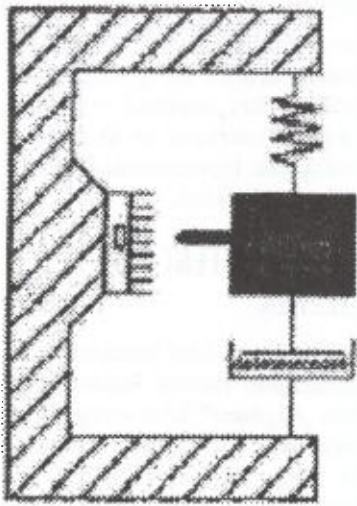
A gyorsulásmérő olyan lengő rendszer (2. ábra), amely valamilyen mozgás bizonyos gyorsuláskomponensét méri. A gyorsulásmérő a mozgó testhez van rögzítve.

A gyorsulásmérők rendelkeznek egy meghatározott geometriai tengellyel (ún. érzékeny tengely). A gyorsulásszenzor ideális esetben valamely gyorsulásvektornak azon komponensét fogja mérni, amely a gyorsulásmérő érzékeny tengelyére vetül.

Amikor a gyorsulásmérő gyorsul, akkor az érzékeny tengelye mentén a szeizmikus tömegre (m) tehetetlenségének következtében tehetetlenségi erő hat. Ez az erő a rugót kitéríti. A szenzor kimenete a rugó kitérésével, vagy a rugó tehetetlenségi erejével lehet arányos.

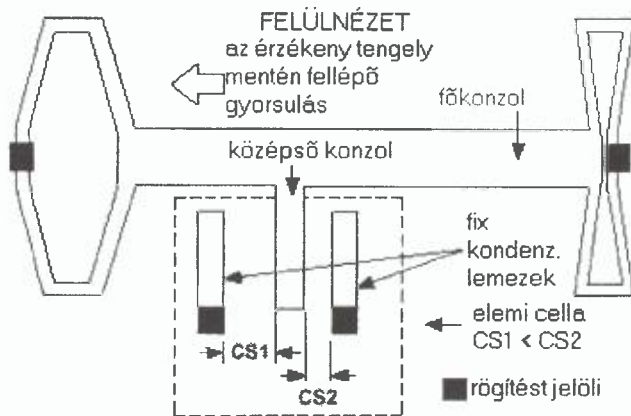
A gyorsulásmérők korai változatait lengő alkatrészekon fellépő gyorsulások mérésére használták. Ezek a szenzorok statikus gyorsulás mérésére alkalmatlanok voltak.

Az általunk vizsgált ADXL105 gyorsulásszenzor statikus gyorsulások mérésére is alkalmas.



2. ábra. Gyorsulásmérő

A 2. ábrán látható lengő rendszer az ADXL105 esetében a chipre van integrálva. Az inerciális tömeg egy differenciál kialakítású kapacitás fegyverzetei között mozdul el (3. ábra).



3. ábra. Az érzékeny tengely mentén gyorsulás lép fel

A kapacitás fegyverzeteit nagyfrekvenciás oszcillátor ellenfázisban hajtja meg. Az elmozduló konzolról (amely a diff. kapacitás része) levett jel szinkron demodulátorra kerül.

Amikor a szeizmikus tömeg (középső konzol) mozgásba jön, az elmozdulásnak megfelelő pillanatnyi kimeneti jel jön létre a demodulátor kimenetén. A szenzor belsejében a demodulátor kimenetét visszacsatoljuk, amely így elektrosztatikusan téríti vissza a szeizmikus tömeget a középső pozícióba. (Innen érthető az erőkompenzációs működési elnevezés, hiszen elektrosztatikus erővel kompenzálunk.) A demodulátor kimeneti feszültsége akkora értéket ér el, amekkora ahhoz szükséges, hogy a gyorsulás fellépte alatt nyugalomban tartsa a konzolt, így a fellépő gyorsulással arányos kimeneti jel jön létre.

8. SZÖGHELYZET-ÉRZÉKELÉS IDEÁLIS GYORSULÁSMÉRŐVEL

Először áttekintjük, hogy lejtőn nyugvó tárgyhoz erősített ideális gyorsulásmérővel hogyan tudunk lejtőszöget

mérni. (Az ideális gyorsulásmérő az érzékeny tengelyére merőlegesen fellépő gyorsulásra teljesen érzéketlen.)

Ha az 1. ábrán lévő lejtőre egy gyorsulásmérőt rögzítünk, akkor az ideális gyorsulásmérő az az F_k kényszere-ro¹ hatására abszolút értékét fogja mérni. (A képletben m az az inerciális tömeg, amire a gyorsulásmérő belsejében F_k hat.) Mivel F_{bk} függ a lejtő szögétől, egy gyorsulásmérő segítségével, valamint jelfeldolgozással² a dőlés szögét meg tudjuk határozni: $\alpha = \arccos\left(\frac{a_k}{g}\right)$. Tehát a mért gyorsulás számértékét el kell osztani a nehézségi gyorsulás értékével, majd a hányados arccos-át kell venni.

Egy másik lehetséges elrendezésben is lehetővé válik a lejtőszög mérése. Ha az érzékeny tengelyt a lejtő síkjában helyezük el, akkor az ideális gyorsulásmérő az bF_l lejtő irányú erő hatására létrejövő $a_l \frac{F_l}{m} = \frac{mg \sin(\alpha)}{m}$ gyorsulás abszolút értékét fogja mérni, mivel érzékeny tengelye az F_l erő hatásvonalával párhuzamos. A mért értékből a \sin függvény inverzének segítségével számítható ki a megfelelő szög (a g nehézségi gyorsulás számértékét ismerjük).

Az egyes elrendezések érzékenysége különböző. A mérendő szög tartomány alapján kell mérlegelnünk, hogy melyik elrendezést kell alkalmazni.

Most mozgó járműre rögzítjük a gyorsulásmérőt. Az imént tárgyalt elrendezés, amikor az 1. ábra jelölése szerint l -vel párhuzamosan helyezük el a szenzor érzékeny tengelyét, nem alkalmas lejtőszög mérésre, hiszen a súrlódás és az (1)-ben a^* -gal jelölt gyorsulás számunkra ismeretlen.

Azonban ha az érzékeny tengely a lejtő síkjára merőlegesen helyezkedik el, az ideális gyorsulásmérő érzéketlen lesz a lejtő síkjában fellépő gyorsulásokra. Feltételezve, hogy a jármű a lejtő síkjára merőlegesen nem gyorsul, α értéke meghatározható a \cos függvény inverzének ismeretében.

A valóságban a jármű esetleg gyorsulhat a lejtő síkjára merőlegesen is (egyenetlen útviszonyok), azonban ez a hatás jól megválasztott szűrési eljárással kiküszöbölhető.

Az említett elrendezésben fontos paraméterek még a szenzor érzékenysége és zaja, hiszen ezek határozzák meg a fenti eljárással mérhető legkisebb szöget.

Amennyiben egyetlen szenzor érzékenysége nem volna elegendő, próbálkozhatunk két gyorsulásmérő lejtőre helyezésével, amelyek érzékeny tengelyei egymásra merőlegesen, és a lejtő síkjával mindkettőn 45 %-ot zárnak be.

Feltételeztük továbbá, hogy a jármű oldalirányban nem dől meg. Ha a jármű a valóságban mégis megdől és így torzítja a mérés eredményét, további szenzorok felhelyezése és jelfeldolgozás lehet a megoldás.

9. MÉRÉSI HIBÁK

A valóságban a gyorsulásmérőknek van egy ún. keresztirányú érzékenysége, ami megmutatja, hogy az érzékeny tengelyre merőlegesen fellépő gyorsulás hány százaléka jelentkezik a kimeneti jelben. Az ideális gyorsulásmérővel elgondolt elrendezésünket tehát úgy kell, hogy módosítsuk,

¹ $F_k = mg \cos(\alpha)$, ahol m a gyorsulásmérőben lévő inerciális tömeg, amit F_{bk} gyorsít.

² A jelfeldolgozó egységnek \arccos funkciót kell megvalósítani.

³ A gyorsulásmérő belsejében lévő inerciális tömeg (amin fellépnek a gyorsulások) az összefüggésekből kiesik.

hogy a lejtő síkjára merőleges érzékeny tengellyel felállított gyorsulásmérő mellé elhelyezünk egy másik gyorsulásmérőt, amelynek érzékeny tengelye a lejtő síkjával párhuzamos.

Az elrendezést az üzembe helyezés során kalibráljuk: meghatározzuk a keresztirányú- és az érzékeny tengelyre eső érzékenységek viszonyát. Ezeket, mint A-priori ismereteket felhasználva most már kompenzálhatjuk a gyorsulásmérő kereszt-érzékenységi hibáját.

Egy további tényezőt is figyelembe kell vennünk: az eszköz hőmérsékletfüggését. Ez utóbbi hatás kompenzálását elvégezhetjük termosztálással – ez viszonylag nagy teljesítményt igényel.

Kompenzálhatunk azonban úgy is, hogy a gyorsulásmérő nyugalmi állapotában felvesszük az eszköz hőmérsékletfüggését. (A hőmérsékletet változtatjuk, és a szögmérési pontosság megkívánta sűrűséggel felvesszük a gyorsulásmérő kimeneti feszültségét a hőmérséklet függvényében.)

A jelfeldolgozási eljárásokban kihasználhatjuk azt a tényt is, hogy a hőmérsékleti folyamatok viszonylag lassan mennek végbe. Könnyíthet a helyzetünkön az ADXL105 szenzorba integrált hőmérő.

Ügyelnünk kell a gyorsulásmérők pontos pozícionálására mind egymáshoz képest, mind a lejtő síkjához képest, hiszen kis szögeknél a pontatlan elhelyezés hibát okoz. A mérési elrendezést sík talajon kalibrálhatjuk a földi nehézségi gyorsulás kihasználásával.

10. AZ ELJÁRÁS ALAPJÁN MÉRHETŐ TOVÁBBI PARAMÉTEREK

Elméletben lehetséges lenne további paraméterek mérése is, pl. a jármű által megtett relatív szintkülönbség mérése. Ezzel alternatívát mutatunk a barometrikus elven történő relatív magasságmérésre: a jármű által befutott út – mint változó hajlásszögű lejtő – tetszőlegesen sűrűn felosztható elemi szakaszokra és az egyes szakaszok viz-

gálhatók az 1. ábra jelölésrendszerével.

Az általunk javasolt megoldás a következő geometriai összefüggéseken alapszik:

$$h = \sin(\alpha) \cdot l \quad (2)$$

Az elemi lejtő magasságát h -val, hajlásszögét α -val jelöltük. Az elemi szakaszon a lejtőn megtett út: l .

Ha a lejtőn a megtett utat, valamint a lejtő hajlásszögét folyamatosan mérjük, akkor elvben tetszőlegesen kicsiny szakaszra ki tudjuk számítani az elemi lejtőszakasz magasságát. Ezeket az elemi magasságokat összegezve vagy kivonva (attól függően, hogy a jármű felfelé vagy lefelé halad, egy adott ponthoz viszonyítva), meghatározható a jármű magassága. (Az, hogy a jármű felfelé vagy lefelé halad-e, a lejtő szögének megfelelő eljárással történő mérése esetén megállapítható.)

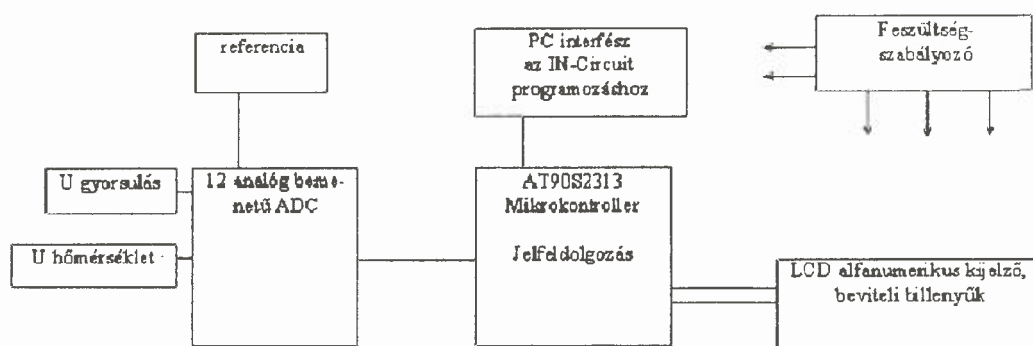
Ha $h = f[\sin(\alpha), l]$ és $b = \sin(\alpha)$ és l hibával terhelt, akkor $\Delta h = b\Delta b + l\Delta l$ lesz a mérési hiba egy elemi lejtőszakaszra.

A gyakorlatban azonban ez az eljárás csak korlátozott pontossággal alkalmazható; a szögmérés hibája egyre halmozódik és egy idő után elviselhetetlenül nagy lesz.

11. JELFELDOLGOZÓ EGYSÉG

Az önállóan tervezett jelfeldolgozó egység a 4. ábrán látható. A rendszer magja az Atmel cég AVR2313 típusú 8 bites mikrokontrollere. Azért választottuk ezt a korszerű RISC eszközt, mert a speciális DSP processzoroknál nagyságrendekkel olcsóbb, ugyanakkor igen hatékony (16 MIPS), így a szükséges jelfeldolgozási funkciók könnyen megvalósíthatók. A kártyán elhelyezkedik még egy 12 bites többszámú A/D átalakító, jusztirozható referenciák stb.

Az eszköz programja a mikrokontrollerbe integrált flash memóriában van, ami a PC-ből a soros letöltőkábelén keresztül frissíthető. A mért értékeket LCD kijelzőn jelenítjük meg.



4. ábra. Jelfeldolgozó egység

12. KIÉRTÉKELÉS

Az eszközzel a statikus méréseket egy fából készített kísérleti lejtőn folytattuk, amelyen a dőlésszög egy csavarral precízen beállítható volt.

Egy nagy felbontású mechanikai szögmérővel tudtuk a gyorsulásmérő által mért lejtőszög-értékeket összehasonlítani. A mechanikai szögmérőn a tolómércéhez hasonló nóniusz beosztás teszi lehetővé az igen kis szögek mérését.

A mechanikus szögmérő felbontása jobb, mint $0,2^\circ$.

Csak konstans gyorsulás mérésére is alkalmas gyorsulásmérővel lehet szöget mérni! Az ADXL szenzor természetesen méri a DC gyorsulást, így a mérőkörben is DC csatolást kell alkalmazni. A gyorsulásmérővel felépített szögmérő statikus érzékenysége eléri a $0,5^\circ$ felbontást.

A dinamikus körülmények közötti vizsgálatok diplomamunka keretein belül jelenleg is folynak.

IRODALOMJEGYZÉK

- [1] Michael Doogue, Michael Walsh: The Design of a Track Map Based Data Acquisition System, 1998.
- [2] Analog Devices: ADXL105 iMEMS@Single Axis Accelerometer with Analog Input, (Datasheet) 1999.
- [3] Holics László: Fizika – Klasszikus Fizika, Műszaki Könyvkiadó, Bp., 1986.
- [4] Dr. Fock Károly: Rezgésmérők dinamikus tulajdonságainak elemzése, Elektronet 1996/3 (13-15. o.)
- [5] Texas Instruments: TLC2543C 12-Bit Analog-to-Digital Converters with Serial Control and 11 Analog Inputs, 1997 (Datasheet)
- [6] Norm Bernstein: Interfacing the AD22100 Temperature Sensor to a Low Cost Single-Chip Microcontroller (Analog Devices Application Note AN-395)

FEASIBILITY RESEARCH OF APPLICATIONS USING ACCELERATION SENSING

L. CSER

DEPARTMENT OF CONTROL ENGINEERING AND INFORMATION TECHNOLOGY
TECHNICAL UNIVERSITY OF BUDAPEST
H-1111 BUDAPEST PÁZMÁNY PÉTER SÉTÁNY 1/D

Nowadays it is much easier to measure a number of physical quantities thanks to the progress made in sensor technology. The serious development in acceleration sensing is one of the benefits of automotive industry. Cheap accelerometers could be integrated into several applications such as navigation, mapping systems, seismology, signature analysis etc.

After reviewing the most common methods for acceleration and vibration sensing we would like to discuss accelerometer type ADXL05 from Analog Devices through a project. What makes the ADXL family a distinguished accelerometer is the integration of the sensor element and the signal processing circuitry on the same silicon surface.

The project is to design an ADXL05 based instrument which is going to be installed on a moving bicycle. The instrument should know if the bike is going uphill or downhill. The instrument should also show the tilt of the road.

The first thing to do is to design the measurement method to be used. We should also be able to answer the question "What does an ideal accelerometer actually measure?" After answering this question we are going to give a brief comparison of an ideal and a nonideal accelerometer.

Discussing the errors of the measuring method is an important issue. That is our next task. The specification itself, the constraint which comes from choosing the particular device (ie. ADXL05) and the constraints of the measuring method all contribute to the final solution.

The tilt instrument design is going to be discussed with the help of the functional blocks in the instrument. In some parts of the design we go into minute details and sometimes we take a high level approach if that is necessary.

DIGITÁLIS KÉPÁTVITELI RENDSZEREK

LENCSE LÁSZLÓ

BUDAPESTI MŰSZAKI ÉS GAZDASÁGTUDOMÁNYI EGYETEM
VILLAMOSMÉRŐNKI ÉS INFORMATIKAI KAR VILLAMOSMÉRŐNKI SZAK
1541 BUDAPEST

Digitális képek tárolása és továbbítása rendkívül költséges, hiszen általában egyetlen kép mérete is Megabájtos nagyságrendű. Ezért a legtöbb digitális képátviteli rendszerben általános cél, hogy a lehető legkevesebb bittel reprezentálják a képeket, és ezáltal növeljék a rendszer átviteli sebességét, illetve hogy minimalizálják a tárolási követelményeket.

A vektorkvantálás implementálásával az alkalmazás-orientált képkompresszió jelentős mértékben javítható napjaink általános célú képkompressziós eljárásaihoz képest. Megnőtt az igény az olyan alkalmazásokra is, melyben lehetőség van képjellemzők kiemelésén alapuló képviszakeresésre. A vektorkvantálás, mint hatékony lényegkiemelő módszer ezen a területen is jól használható.

Az Internetes technológiák gyors fejlődésének eredményeképpen egyre inkább elterjednek a különböző multimédia adatbázisok és alkalmazások. Az Oracle8 kiemelt helyen szerepel napjaink relációs adatbázisai között. Az Internetes adatbázisok történetében új fejezetet nyit mind az adatkezelés, mind az információ elérés területén. Képes kezelni a jelenleg érvényes általános képkompressziós technológiákat, valamint „hagyományos” eljárások alapján működő képkeresési rendszerrel rendelkezik.

A BME Irányítástechnika és Informatika Tanszékén jelenleg futó kutatás egyik célja, hogy megvizsgálja, hogyan lehetséges az Oracle8 relációs adatbázis képkeresési eljárásában a vektorkvantálás és egyéb alkalmazás-orientált eljárás szolgáltatása képjellemzők implementálása.

1. BEVEZETÉS

Az elmúlt évtizedben hihetetlen mértékben megnőtt a digitális képek használata. Ez főleg az informatika és a telekommunikáció gyors fejlődésének köszönhető. A képi információk kezelésében, hasznosításában legfőképp az Internet nyitott új lehetőségeket. Az Internetes technológiák, a hálózati képfeldolgozás, illetve multimédia adatbázisok szerepe különösen fontos szerepet játszik az orvosi telediagnosztikai alkalmazásokban.

A technológiai fejlődésnek eredményeképpen a képek nagyobb méretben bevihetők a számítógépekbe. Ez azonban új problémákat és kérdéseket vet fel. Digitális képek tárolása és továbbítása rendkívül költséges, hiszen általában egyetlen kép mérete is Megabájtos nagyságrendű. Ezért a legtöbb digitális képátviteli rendszerben általános cél, hogy a lehető legkevesebb bittel reprezentálják a képeket, és ezáltal növeljék a rendszer átviteli sebességét, illetve hogy minimalizálják a tárolási követelményeket. Ezek mellett fontos szerepe van a képviszakeresésnek, képfelismerésnek is. Gyakoriak az olyan alkalmazások, ahol néhány képi jellemző alapján kell megtalálni az adott képhez leginkább hasonló, az adatbázisban már meglévő másik képet, illetve képeket.

A BME Irányítástechnika és Informatika Tanszékén futó project célja egy automatikus rendszer fejlesztése a melanoma hatás korai felismerése céljából. A project olyan feladatok megvalósítását tűzte ki, amely korszerű technológiákat használ a hardware fejlesztés, adattárolás, titkosítás, kép kompresszió, automatikus diagnózis, hálózati lehetőségek kihasználása területén. Ezek közül a kompresszió és adattárolás témakörökkel foglalkoztam.

2. KOMPRESSZIÓ

A tárolási, hálózati továbbítási problémákat képkompressziós eljárások használatával lehet megoldani. Esetünkben állókép tömörítésről van szó, melynek alapvetően két típusa van: veszteséges és veszteségmentes. A veszteséges

mentesség azt jelenti, hogy a tömörített képből az eredeti kép bitről-bitre pontosan visszaállítható. Veszteséges esetben a kódolás után dekódolt kép nem egyezik meg az eredeti képpel, viszont nagyobb tömörítési arány érhető el. Önálló képek veszteséges tömörítése azért lehet hatékony, mert a szomszédos képpontok között az eltérés várhatóan kicsi, hiszen egy kép többnyire közel azonos színű területekből és ezek közötti zajszerű élekből áll. A kép ezen jellegzetessége a frekvenciatartományba transzformálva használható ki. A lassú változás kis, a gyorsabb pedig nagyobb frekvenciának felel meg. Mivel az emberi látórendszer a nagyobb frekvenciájú jelek esetén kevésbé érzékeny a torzításra, a frekvenciatartományban könnyen kezelhető a frekvenciától függő veszteség.

A veszteséges képtömörítőket sokféle szempont szerint értékelhetjük. A felhasználó számára a legfontosabb három paraméter a tömörítési arány, a végrehajtási idő és a képminőség. A tömörítés és dekódolás időigénye nagyjából egyforma lehet. További követelmény lehet, hogy a tömörített adatokhoz más típusú információ is illeszthető legyen (pl. szöveges adatok, geometriai, időbeli és numerikus információk, beszéd stb.), és a tömörítéshez csak közvetve kapcsolódó feladatokat is lehetővé tegyen (pl. zajszűrés, éldetektálás, mozgó objektum detektálása képsorozatokban stb.).

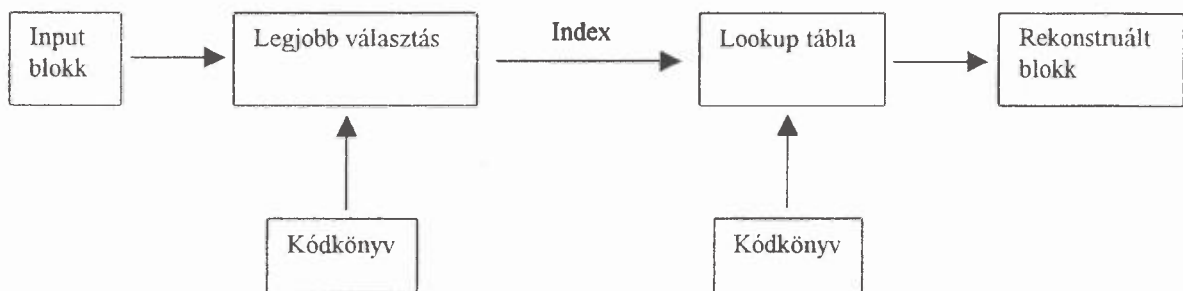
Napjaink aktuális veszteséges képkompressziós technikái közül a *vektorkvantálás* egy sokat ígérő eljárásnak mutatkozik. A kvantálás egy, az eredeti halmaznál lényegesen kisebb elemszámú halmazra való leképezés úgy, hogy az eredeti értékhez a halmazból a legközelebbi értéket rendeljük hozzá egy távolságkritérium (norma) alapján. A kvantálás végezhető önálló értékre (skalár kvantálás), ennél lényegesen hatékonyabb viszont, ha több értéket egy blokkba fogunk össze. Ezt nevezik vektorkvantálásnak. Képek esetén a blokkok tipikusan 2×2 , 3×3 vagy 4×4 -es méretűek. Minden blokkot egy $W = \{w_0, w_1, \dots, w_{(N-1)}\}$ kódkönyv segítségével lehet kódolni. A kódkönyv elemeit kódvektornak, vagy kódszónak nevezik. A kódolás után csak a legközelebbi kódvektor in-

dexe van továbbítva, vagy eltárolva. A dekódolás a kódoláshoz hasonlóan történik. A vektorkvantálás folyamatát az 1. ábra mutatja.

A legjobban illeszkedő kódszót egy bizonyos torzítás mérték alapján választják ki, amely általában a középérték hibája. A dekódoló a képet a kódkönyv kódszavain végigmenve egy lookup-table művelet alapján állítja vissza. A vektorkvantálás úgy biztosítja a tömörítést, hogy közvetlen megszünteti a redundanciát egy blokk pixelei között. Ennek ellenére a blokkok között szintén fennállhat korreláció.

Véges Halmazú Vektorkvantálás (Finite State VQ: FSVQ) esetén memóriát viszünk a VQ sémába. Az FSVQ-ban minden blokk úgy van kódolva, hogy csak a kódszavak megfelelő részhalmazában keresi a legközelebbi kódvektort, ezt állapot kódkönyvnek, vagy rész kódkönyvnek

nevezik. Az FSVQ lényege, hogy az előzőleg kódolt blokkok alapján megjósolja a következő blokk pozícionak legvalószínűbb kódvektorait. Erre egy egyszerű, de hatásos megoldás a [2] alapján Feltételes Histogram (CH). A CH alap ötlete, hogy a kódvektor indexeinek sorrendisége elsőrendű Markov lánc szerint modellezhető. Feltételezve, hogy a kódszavak előfordulásának feltételes VALÓSZÍNŰSÉGE ISMERT, A KÖVETKEZŐ ÁLLAPOT FUNKCIÓ AZ M legnagyobb valószínűségű kódvektor kiválasztásából áll, hogy abból lehessen kialakítani az M elemű rész kódkönyvet. Ha a következő állapot funkció jóslása helyes, akkor az FSVQ a legjobban illeszkedő vektor választja, ami mindig benne van az állapot kódkönyvben. Rossz döntés esetén nem állítja elő ugyanazt a soron következő vektor sorozatot, mint azt az eredeti *memóriamentes* VQ.



1. ábra

Az eljárás ezen hibája azonban abban az esetben elkerülhető, ha a kódoló ellenőrzi, hogy a legjobb illesztés kívül esik-e a jelenlegi állapot kódkönyvön. A „Reversed Codevektor” (RC) egy lehetséges ellenőrzési módszer. A módszer részletes leírása a [3] cikkben olvasható. Ebben az esetben a kódvektorok száma a szuper kódkönyvben legyen N ($N = 2^n$), az állapot kódkönyvben pedig M ($M = 2^m - 1$). Ezután n és m bit ($m < n$) használható a szuper illetve állapot kódkönyv kódvektorának kódolásához. Ha a nyertes vektor az állapot kódkönyvben található, akkor m bit lesz átküldve, ha az állapot kódkönyvön kívül esik, akkor $n+m$ bit lesz átküldve. A „Reversed Bit” módszer esetén nem egy kódvektort használnak, hanem egy bit van fenntartva arra, hogy jelölje vajon hibás-e az állapot. A rész kódkönyv mérete ekkor $M = 2^m$. Ha a legjobb illesztés a rész kódkönyvben van, akkor egy „0” bit lesz a dekódoló felé továbbítva, ellenkező esetben „1”. Az RB megvalósításnál tehát legjobb esetben $1 + m$, legrosszabb esetben $1 + n$ bit kerül továbbításra. A két megvalósítás természetesen két különböző bit arányt biztosít. Ha gyakoriak a helytelen állapotok, akkor érdemes az RB ellenőrzést használni. Az RC-vel viszont kisebb bit arányt lehet elérni, ha a következő állapot funkció ritkán hibás.

A képtömörítés hatásossága függ a rész kódkönyv méretének megválasztásától. Túl nagy állapot kódkönyv (m nagy) esetén nem csökken nagyon a bit arány, vagyis kevésbé hatásos a tömörítés. Ha az állapot kódkönyv mérete túl kicsi, akkor a kép veszít az információtartalmából, hibás információkat fog tartalmazni. Az optimális megoldás a blokk indexek feltételes valószínűség eloszlásától függ.

Ha a feltételes valószínűség kevés kódvektorhoz koncentrálódik, akkor egy kis állapot kódkönyv képes tartalmazni a kódvektorokat, ekkor alacsony bit arány (magas tömörítési ráta) érhető el. Szélesebb eloszlású valószínűségeknél a nagyobb méretű kódkönyv az ajánlott. Létezik egy adaptív algoritmus, amely változtatja az állapot kódkönyv méretét az aktuális valószínűség eloszlásának megfelelően.

2.1. RC esetén

Legyen p_j az i -edik kódvektor feltételes valószínűségének eloszlása. Továbbá legyen a kódszavak valószínűsége $p_0 < p_1 < \dots < p_{N-1}$, valamint $p(m)$ legyen a valószínűsége a legjobban illeszkedő rész kódkönyv kódvektora. Ekkor $p(m)$ arányos m -mel és

$$\tilde{p}(m) = \sum_{r=0}^{2^m-2} p_r$$

Mivel az átvitt bitek száma hol m , hol pedig $m + n$, a várható értéke az átvitt biteknek az alábbi összefüggés szerint számítható:

$$E(m) = \tilde{p}(m) \cdot m + (1 - \tilde{p}(m)) \cdot (m + n)$$

Az optimalizáló eljárás ennek az értéknek a minimumát keresi meg. Ennek az értéknek kisebbnek kell lenni, mint n , hogy garantálja a kisebb bit arányt.

2.2. RB esetén

Az eljárás nagyon hasonlít az előzőre, de most 2^m kódvektor van.

A feltételes valószínűség és a várható érték az alábbiak szerint számítható:

$$\tilde{p}(m) = \sum_{\gamma=0}^{2^m-1} p_{\gamma}$$

$$E(m) = \tilde{p}(m) \cdot (1 + m) + (1 - \tilde{p}(m)) \cdot (1 + n)$$

Itt is a minimumot keressük és a kapott értékek kisebbnek kell lenni, mint n .

3. ADATTÁROLÁS ÚJ LEHETŐSÉGEI ORACLE8 PLATFORMON

Az Oracle8 kiemelt helyen szerepel napjaink relációs adatbázisai között. Új fejezetet nyit az Internetes adatbázisok történetében mind az adatkezelés, mind az információ elérés területén. Az Oracle8 nem csak fejlett eszközkészletet ad a mai web-oldalakra található szinte összes adattípus kezelésére, de rendelkezik a szükséges teljesítménnyel és skálázhatósággal, amely a nagyobb oldalak kezelését is támogatja. Az Oracle8 több mint egy relációs adattároló, a saját fejlesztésű adatbázis mellett olyan újításokkal rendelkezik, mint például az Internet File System és az Oracle Intermedia. A szerver/kliens architektúra kiépítésére több megoldást kínál, így az applikáció management egyszerűbbé vált.

A teljesítmény növelés érdekében az adatok olyan szervereken replikálhatók, amelyek közel vannak a felhasználóhoz és csak a számukra fontos adatokat tartalmazzák. A többretegű felhasználói környezet számára kifinomult hozzáférés szabályozással és továbbfejlesztett biztonsági modellel rendelkezik. A kifinomultság alapja egy paraméteres, bővíthető applikációs kontextus, ami lehetővé teszi az alkalmazások számára, hogy a felhasználók hozzáférést user-attribútumok alapján szabályozzák. A Web Application Server az adatbiztonság megteremtése érdekében a további lehetőségként támogatja az SSL-t, autentikációs sémákat, valamint IP- és domain alapú megkötéseket biztosít.

Az Oracle8-ban a képek tárolását, valamint a képi adatok feldolgozását az Oracle8 Image Cartridge [6], vagy az Oracle8 Visual Information Retrieval Cartridge (VIRC) [5] használatával lehet megvalósítani. A különbség a kettő között, hogy a Visual Information Retrieval Cartridge-ban lehetőség van a képek tartalom alapú visszakeresésére néhány lényeges vizuális tulajdonság alapján. A legfőbb előnye a tartalom alapú visszakeresésnek, hogy csökkenti a képi információk kinyeréséhez szükséges időt.

A VIRC fő szolgáltatásai a tárolás, visszakeresés, új-

raformázás, tömörítés, az adatmanipulációs szolgáltatások minimális része (pl. skálázás, képvágás). Minden képet az adatbázisba helyezéskor analizálnak, és ugyanekkor elkészül a kép tartalmának teljes reprezentációját jellemző vektor, vagy *signature*. A signature a következő vizuális tulajdonságokat tartalmazza: globális szín, lokális szín, textúra, struktúra, illetve ezek helyett szerepelhet az emberi arcot jellemző „facial”.

Az összehasonlító eljárás során a kiválasztott kép signature-je van összehasonlítva az adatbázisban szereplő többi kép signature-jével. A képek összehasonlításakor lehetőség van a vizuális tulajdonságokat súlyozni aszerint, hogy azok mennyire fontosak az aktuális képillesztési feladat szempontjából. Az összehasonlító eljárás eredménye a képek közötti hasonlóság mértéke a *score*. Ugyanakkor definiálva van egy küszöbérték is. Ha a score értéke kisebb a *threshold* értéknél, akkor az Oracle a képeket hasonlónak értékeli.

Mivel a felhasználó adja meg a vizuális jellemzők súlyának, illetve a *threshold*-nak az értékét, ezért az Oracle-ben a képi visszakeresés hatékonyan használata a felhasználó részéről nagyfokú szakértelmet követel. A kép jellegétől és a feladat fontosságától függően nem mindegy, milyen súllyal szerepelnek a vizuális jellemzők, illetve a *threshold* értéke hogy van megválasztva.

Az előzőekben leírtak alapján az Oracle8 használta tűnik a project számára legalkalmasabbnak az adattárolás területén, valamint a titkosítási és hálózati feladatok területén is jelentős előnyöket jelent a használata.

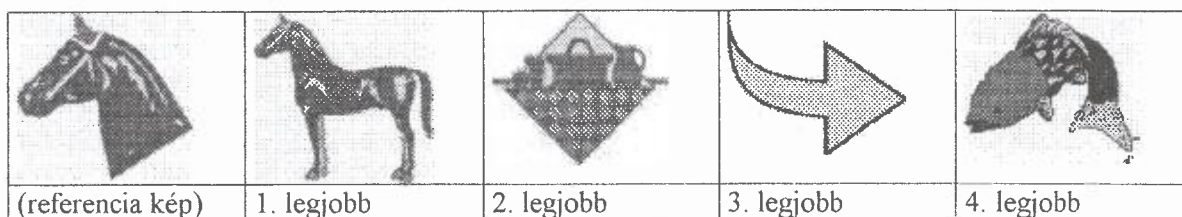
4. SZIMULÁCIÓS EREDMÉNYEK

Az Oracle8 képkeresési funkciójának tesztelésére létrehoztam egy mintaalkalmazást. Az eredmények helyességének ellenőrzéséhez a napjaink néhány jellegzetes kép visszakereső rendszerét [7], valamint az egyedi ellenőrzés céljából egy képmanipulációs szoftvert [8] használtam. A mintaadatbázis képei megegyeznek ezen rendszerek képeivel, bár az adatbázis mérete sokkal kisebb.

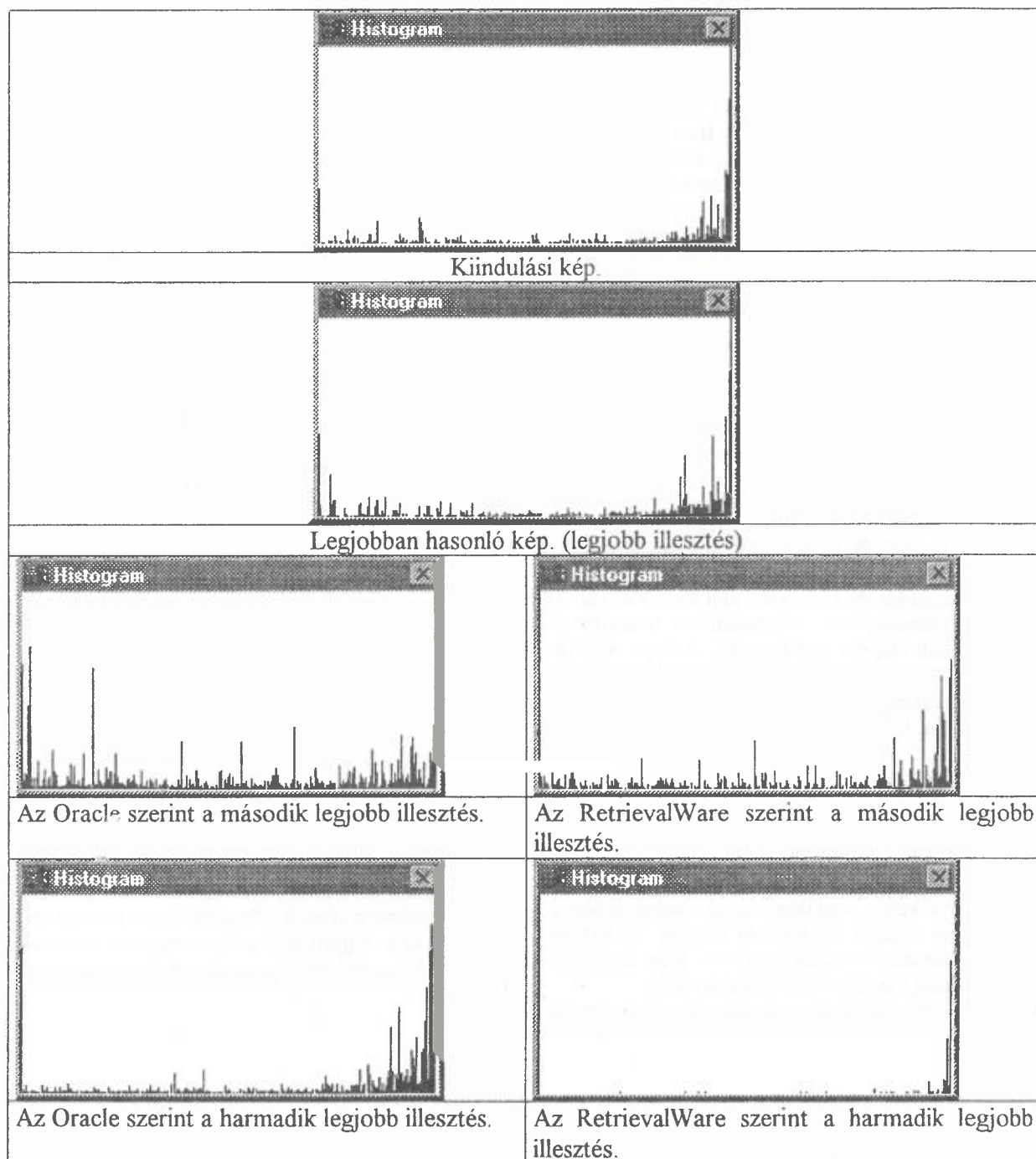
Az alkalmazás első tesztelési lépésében kiválasztva egy olyan képet mely szerepel a jellegzetes kép visszakereső szoftverek egyikében a program a 2. ábrán látható képeket találta a legjobban hasonlónak.

A 2. ábra alsó sorában a *score* értéke van megadva. A keresés során a vizuális jellemzőket azonos fontossággal voltak figyelembe véve, a jellemzők súlya minden esetben egységnyi volt. Ugyanezen képre elvégezve a keresést, a RetrievalWare rendszerben a 3. ábra szerinti eredményeket kaptam.

				
(referencia kép)	18.3983	20.7136	22.7958	24.2056



3. ábra



4. ábra

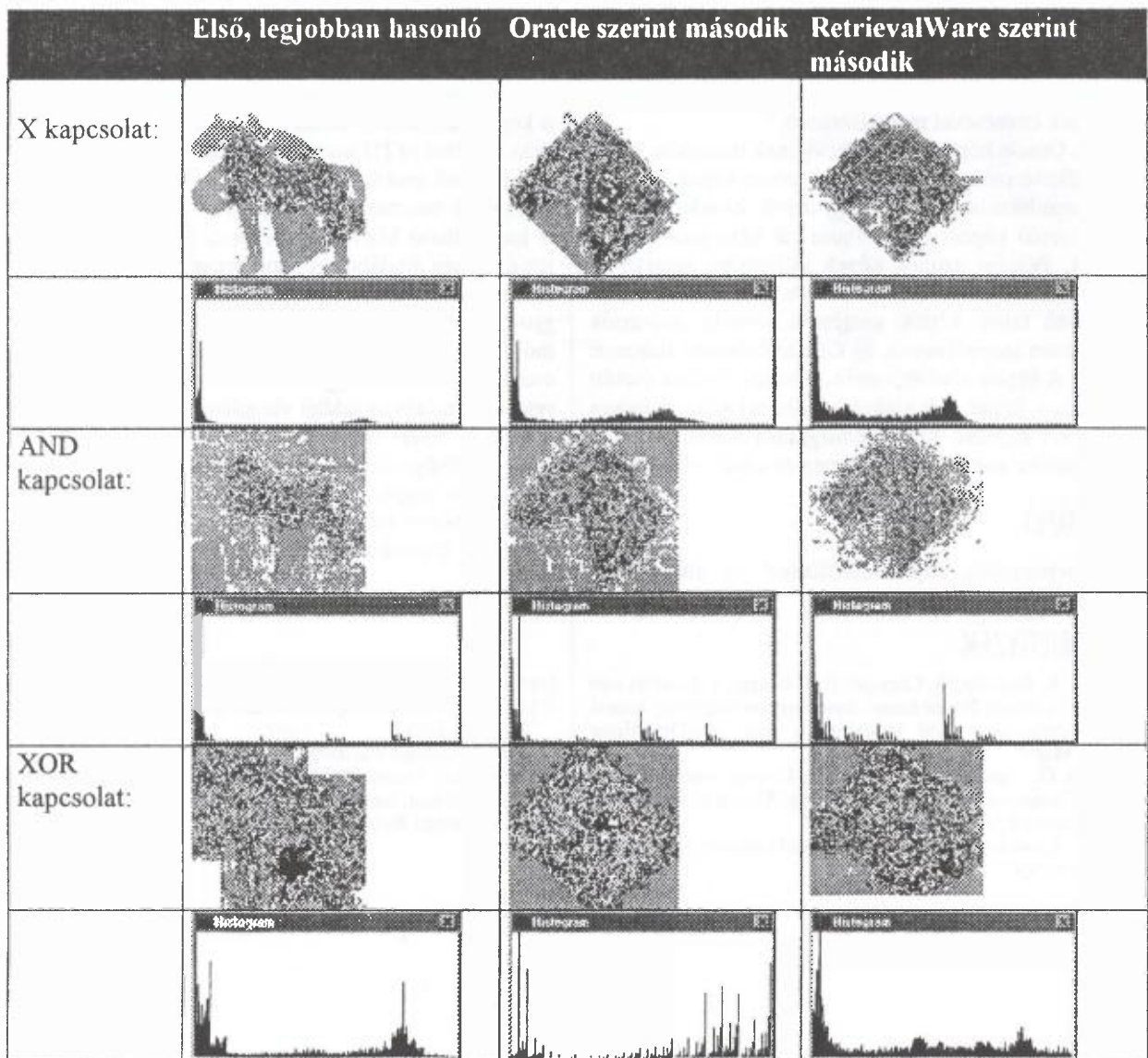
A RetrievalWare programban is van lehetőség a vizuális jellemzők súlyozására. Ebben a rendszerben vizuális jellemzőként a szín-, alak-, textúra-, fényesség-tartalom, színstruktúra és képarány szerepel. A keresés során mindegyik vizuális jellemző súlya megegyezett. Az egyes képek pontos elemzését elvégezve a megfelelő szoftverrel [8], a 4. ábrán látható histogramokat kaptam.

Tovább folytatva a vizsgálatot és elvégezve az egyes képek X, AND és XOR kapcsolatainak vizsgálatát, az 5. ábra szerintiakat kaptam.



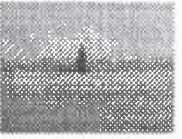
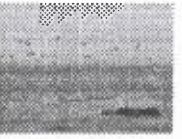






Jelen esetben az Oracle8 jobb eredményt adott a másik rendszernél, a két rendszert azonban nem lehet egyetlen

példa alapján minősíteni. A rendszerek a vizuális jellemzőkben is különböztek egymástól. A kutatás jelenlegi fázisáig nem állt rendelkezésre olyan dokumentáció, amely tartalmazná az egyes rendszerek képi jellemzőinek matematikai meghatározását. Ennek hiányában csak egy nagy-szabású tesztelés végrehajtása után lehet kijelenteni, melyik rendszer szolgáltatja a képvisszakeresés szempontjából a pontosabb eredményt.

Egy másik esetben, amikor clip art helyett egy fénykép volt a keresés alapja, a 6. ábrán látható eredményeket kaptam.



5. ábra

	(referencia kép)	1. legjobb	2. legjobb	3. legjobb	4. legjobb
Oracle8					
Ret.W.					

6. ábra

Látható, hogy a két rendszer közel azonos eredményeket ad a keresésre, a különbség a rendszerek képkeresési módszerének eltéréseivel magyarázható.

Hogy az Oracle képkereső funkciójának tesztelése minél átfogóbb, illetve széles skálájú legyen, olyan képek is szerepelnek az adatbázisban, melyek egy másik, az adatbázisban szintén szereplő képtől csak valamelyik képi jellemzőben különbözik. Például azonos képek különböző színekben, különböző torzításokban elforgatva stb.

Különböző színű képek esetén a vizuális jellemzők súlyát helyesen megválasztva, az Oracle helyesen felismeri a képeket. A képek továbbfestése, továbbrajzolása esetén (pl. valamelyik képre egy alakot rajzolunk) szintén helyes eredményeket kapunk. Képek elforgatása esetén azonban már nem minden esetben kapjuk meg az elvárt eredményt.

5. KONKLÚZIÓ

A vektorkvantálás implementálásával az alkalmazás-

orientált képkompresszió jelentős mértékben javítható napjaink általános célú képkompressziós eljárásaihoz képest és a kép visszakeresés területén is jelentős sikereket értek el vele. Például az [1] irodalom olyan, a tanszéken implementált eljárást mutat be, melyben a vektorkvantálást eredményesen használták arcképkeresésre. Az Oracle8 maga is használható képviszakeresésre. Képes kezelni a jelenleg érvényes általános képkompressziós technológiákat (pl. JPEG, GIF). Az Oracle8 kép visszakeresési módszere „hagyományos” eljárások alapján működik. Láttuk, hogy ez a módszer eredményesen használható, azonban nem minden esetben alkalmas a legjobb illeszkedés meghatározására. A vektorkvantálás az eddigi vizsgálatoknál igen hatékonynak bizonyult. Ezért további vizsgálatok tárgyát eredményezi, hogy lehetséges-e a két előnyös tulajdonság együttes kihasználása, vagyis az Oracle8 által elkészített signature fájlban valamilyen módon a vektorkvantálás szolgáltatja képjellemzők implementálása.

IRODALOMJEGYZÉK

- [1] A. Czihó, R. Ruiloba, G. Cazugel, B. Solaiman, I. Loványi and C. Roux: "Content Based Image Indexing: Investigating Joined Image Compression and Indexing by Use Self-Organizing Feature Maps"
- [2] A. Czihó, G. Cazugel, B. Solaiman, I. Loványi and C. Roux: "Signal Compression Using Finite State Vector Quantization With Optimized State Codebook Size"
- [3] A. Czihó, I. Loványi: "Image transfer and retrieval in multimedia applications"

- [4] Oracle8 Documentation
- [5] Oracle8 Visual Information Retrieval Cartridge User's Guide. Release 1.0.1
- [6] Oracle8 Image Cartridge User's Guide. Release 8.0.4
- [7] Yong Rui, Thomas S. Huang, Shi-Fu Chang: "Image Retrieval: Past, Present, and Future"
- [8] Scion Image Release 3b

DIGITAL IMAGE TRANSMISSION SYSTEMS

L. LENCSE

FACULTY OF ELECTRICAL ENGINEERING AND INFORMATICS
TECHNICAL UNIVERSITY OF BUDAPEST
H-1541 BUDAPEST

The storage and transmission of digital images is exceptionally expensive, since in general the size of a single image can be of the order of megabyte. Therefore there is an overall aim in most digital image transmission systems to represent the images with the minimum number of bits, in order to increase the transmission speed and to minimize the storage requirements respectively.

The implementation of vector quantization can improve the efficiency of the application oriented image compression, compared to nowadays' overall compression methods. The claim was grown to application, in which there is a possibility to content-based image retrieval. The vector quantization, like an efficient substance emphasis method, can be applied in this area.

The fast developments of Internet technologies yield spreading of different multimedia databases and applications. The Oracle8 stands in a respected place among up-to-date relation databases. It opens a new chapter in the story of Internet databases, both data management, and the area of information retrieval. It can incorporate the nowadays' general image compression methods, and has an image retrieval system based on traditional methods.

The purpose of the research controlled by Department of Control Engineering and Information Technology is, to examine how to implement vector quantization and other application-oriented image compression methods on Oracle8 Visual Information Retrieval Cartridge.



Lencse László 1995-ben kezdte meg tanulmányait a Budapesti Műszaki Egyetemen. Tudományos tevékenységét 1998-ban kezdte meg. A Híradástechnikai Tanszék Mobil Távközlési Laboratóriumában a mobil IP területen végzett munkájáért Andersen Consulting Ösztöndíjat nyert. 1999-ben részt vett az egyetemen megrendezett Tudományos Diákköri Konferencián. A Felharmonikus forrás azonosítása című munkája alapján jutalomban részesült. A 2000-ben megrendezett Végzős Konferencián a Telediagnosztikai célú információs rendszerek megvalósítása nyílt hálózatokon című munkájával lépett fel, melyben egy olyan kutatási irányzatot mutat be, ami napjaink egyik legdinamikusabban fejlődő területének számít.

MIKROMANIPULÁTOROK TELEOPERÁCIÓS ALKALMAZÁSA BIOTECHNOLÓGIAI KUTATÁSOKBAN

URBANCSEK TAMÁS IV. évf. villamosmérnök hallgató
LASSÓ ANDRÁS V. évf. villamosmérnök hallgató

BUDAPESTI MŰSZAKI ÉS GAZDASÁGTUDOMÁNYI EGYETEM
IRÁNYÍTÁSTECHNIKA ÉS INFORMATIKA TANSZÉK
1111 BUDAPEST PÁZMÁNY PÉTER SÉTÁNY 1/D

Három, napjainkban dinamikusan fejlődő tudományterület, a teleoperáció, a mikrorobotika és a biotechnológia legújabb eredményeinek integrációjából született az a rendszer, melynek tervezését, megvalósítását és egy valós kísérletben való alkalmazásának tapasztalatait foglaljuk össze cikkünkben.

Az emberi munkának automatikus berendezésekkel helyettesítését kétféle ok indokolhatja: megfelelő gépek alkalmazásával a munka hatékonyabban végezhető el, vagy a tevékenységet gépekkel vagyunk kénytelenek elvégeztetni (például a veszélyes környezet, szűk hely, nagy távolság miatt). Ha azonban a feladat bonyolultsága miatt, vagy váratlan események kezelésére nem alkalmazhatunk teljesen autonóm rendszert, távoli emberi beavatkozás lehetőségét kell biztosítanunk. Ezt a kérdéskört vizsgálja a teleoperáció tudományterülete.

Az általunk fejlesztett tele-mikrorobotikai rendszer elsősorban biotechnológiai feladatok megoldására készült, a témában folyó kutatómunka biztonságos, olcsó és rugalmas elvégzésére. A biztonságos működést a távolról irányítás teszi lehetővé, illetve ezt szolgálja az alkalmazott eszközök intelligenciája (ezek rendkívüli helyzetekben képesek önállóan is működni). Rendszerünk biztosítja a mikrorobot munkaterületének távolról megfigyelését, és a robot irányítását Internet hálózaton át. A speciális hardver eszközök működtetését, összekapcsolását és a hálózati kommunikációt saját fejlesztésű szoftvereink végzik.

1. BEVEZETÉS

A BME Irányítástechnika és Informatika tanszékén folyó MINIMAN mikrorobot fejlesztési projekt célja egy olyan univerzális, intelligens, távolról is irányítható mikromanipulációs munkacella létrehozása, amely alkalmas kutatási feladatok elvégzésére, különösen a biotechnológia területén. Ebben a témában folytatott kutatásaink során munkánk céljaként egy Interneten keresztül irányítható rendszer globális tervének kidolgozását, és egy működő prototípus létrehozását tűztük ki. Cikkünkben rövid áttekintést adunk az érintett tudományterületekről, majd bemutatjuk a megvalósított rendszer hardver- és szoftverelemeit, a megvalósításkor valamint az alkalmazás közben felmerült problémákat és megoldásait.

1.1. Telerobotika

A telerobotika egyike napjaink legkutatottabb, legdinamikusabban fejlődő területeinek. Ez talán nem is csoda, hiszen egy távolról irányítható manipulációs munkaállomás alkalmazási területek százain nyújthat nagy segítséget. A NASA telerobotikai programjának célkitűzése szerint 2004-re a világűrben az összes űrhajón kívüli tevékenység 50 %-át telerobotikai úton fogják elvégezni [15]. A telerobotikával lehetővé válik, hogy egy nagy tapasztalatú orvos a világ bármely pontjáról egy akár ezer kilométerrel odébb fekvő páciensen is segíthessen, vagy egy veszélyes rendőrségi akció emberélet kockáztatása nélkül lebonyolítható legyen, esetleg a Titanic több ezer méteres mélységben heverő roncsai között kincsek után kutathassunk...

De mi is az a telerobot? Egy olyan intelligens, ember által távolról irányítható robot, amely kiterjeszti az ember érzékelő és beavatkozó képességeit egy tőle távol fekvő területre. A lényeges különbség az ipari robotok és a telerobotok között az elvégzendő munka jellegében rejlik: az

ipari robotok sokszor ismétlődően hajtják végre ugyanazt a feladatot, jól meghatározott környezetben, sebességgel, pontossággal, megjósolható eredménnyel. Az emberi teleoperációban ezzel szemben a robotot távolról vezérlő személyé a legfőbb irányítói szerep. A robot szenzorai által küldött információk segítségével a kezelő személy a robot által, emberi képességeinek megfelelően végzi el a beavatkozásokat.

Miért kell egy robotot távirányítani? Azért mert az utóbbi évtizedekben a robotokkal foglalkozó kutatók felismerték, hogy a legtöbb feladatra nagyon nehéz teljesen önállóan működő robotokat készíteni. A telerobotika módszerei azonban lehetővé teszik, hogy a robot elvégezzen egyszerűbb részfeladatokat és kérje a felügyelő személy segítségét, ha valami nehézsége támad. A folyamatos kutatások eredményeként egyre több és egyre bonyolultabb feladatot lesznek képesek megoldani a robotok teljesen önállóan, ám a telerobotikának addig is, és azután is igen jelentős alkalmazási területei maradnak [5], [8]. Ezek a következők:

A. *Emberi tevékenység szükséges, vagy nem akarjuk kiváltani azt mesterséges eszközökkel.* Ha például egy összetett munkafolyamat során kevés számú ismétlődés lép fel, a környezet pedig változó és ismeretlen lehet, célszerűbb emberi felügyeleti irányítást alkalmazni, hiszen szükség van az emberi intelligenciára, felismerőkészségre, rugalmasságra. Bizonyos helyzetekben pedig éppen az emberi részvétel a lényege a tevékenységnek – például „személyes” találkozók megvalósításakor vagy szórakoztatás esetén.

B. *Ember valamilyen ok miatt közvetlenül nem végezheti el a feladatot.* Ilyen okok lehetnek:

- életveszélyes vagy egészségre ártalmas környezet;
- nem emberi léptékű feladat (nagy méretek, erőki-fejtés);



SAGEM

Mit kínálunk Önnek SDH és SDH access megoldásainkkal?

- Minőségi szolgáltatásnyújtás
- Gyors szolgáltatás bevezetés, illetve hibaelhárítás
- Kapcsolt és cella alapú átvitel 64 Kbit/s-tól 2,5 Gbit/s-ig
- Sok interfész (ATM, adatátvitel, ISDN, ADSL)
- Egységes intelligens központi felügyeleti rendszerrel biztosítható az SDH és az SDH access hálózat összes elemének felügyelete végponttól végpontig
- Jövőálló befektetés (64 Kbit/s-os cross-connect, SDH technológia előnyeinek meghosszabbítása az előfizetőig, ATM over SDH)

Bővebb információ:

SAGEM Magyarország Elektronikai Kft.

1045 Budapest, Istvántelki út 8. • Tel.: 399 1020 Fax: 399 1021

Internet: www.sagem.hu

Mozgó objektum követő rendszer Elektronikus menetlevél

Software:

- Kommunikációs program (a mobil egység adatainak kiolvasására)
- Grafikus feldolgozó program az útvonal megjelenítésére, és menetteljesítmény számítására

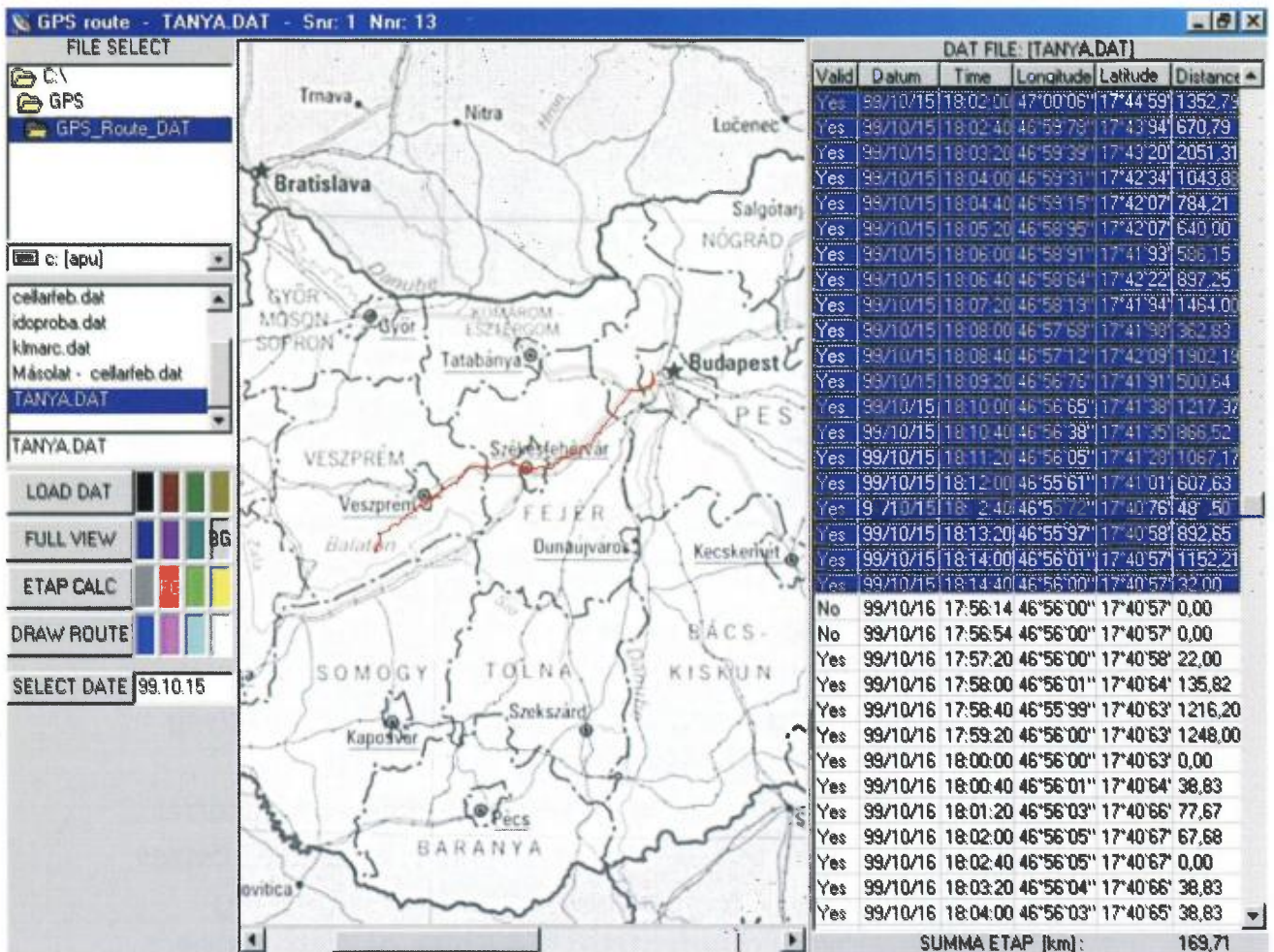


Fig. 23.
Data file selection for the COMMW program

HW és SW tervező: Dr. Gyenes Károly

CORROCONT GROUP

CATHODIC PROTECTION & NDT

CORROCONT HUNGARY H-1124 Budapest, Szendi u. 1

Tel: (0361) 214-2980, Fax (0361) 214-2982

e-mail: huoffice@corrocont.com



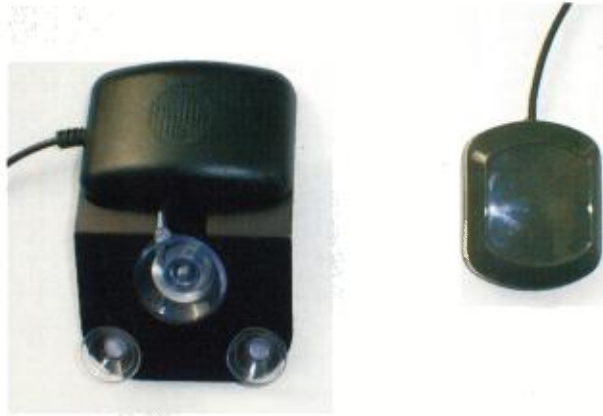


Fig. 2.
The photo of two antenna units

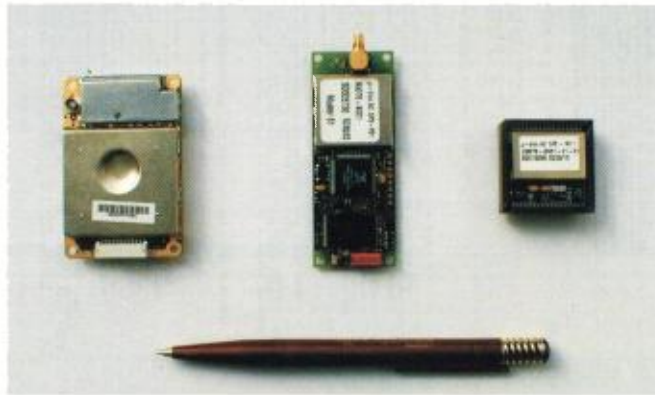


Fig. 3.



Fig. 4.

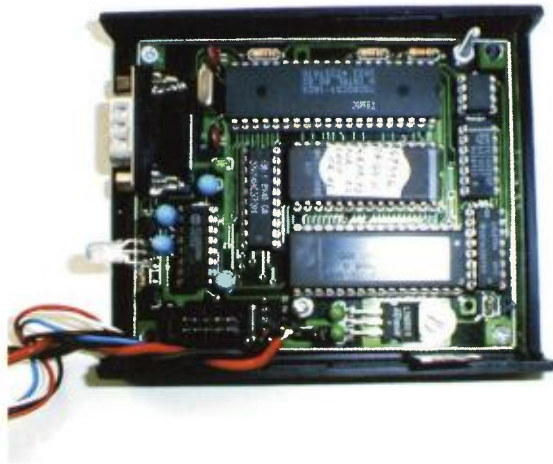


Fig. 5.

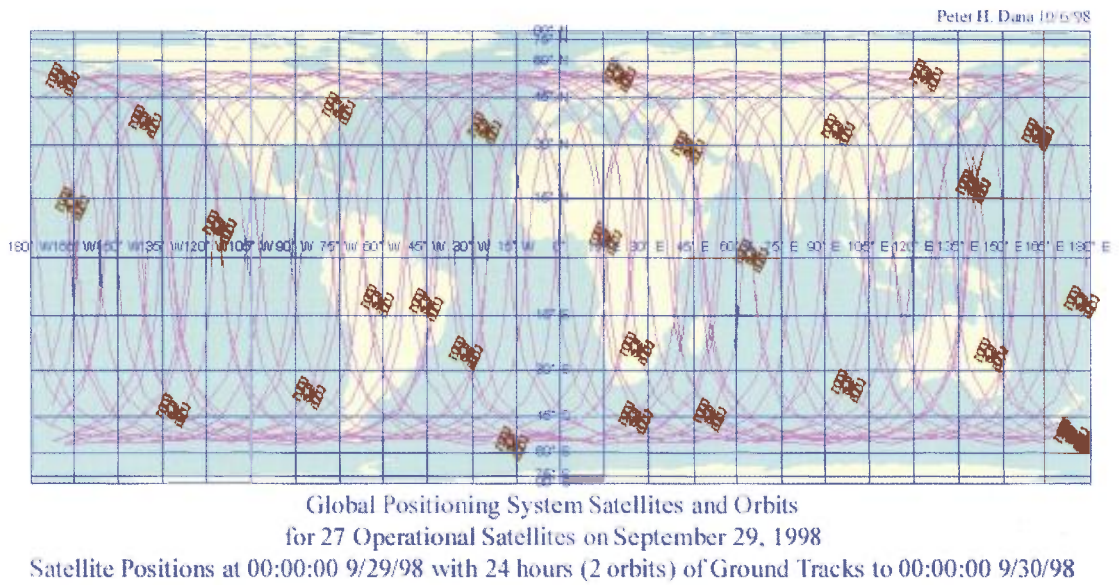
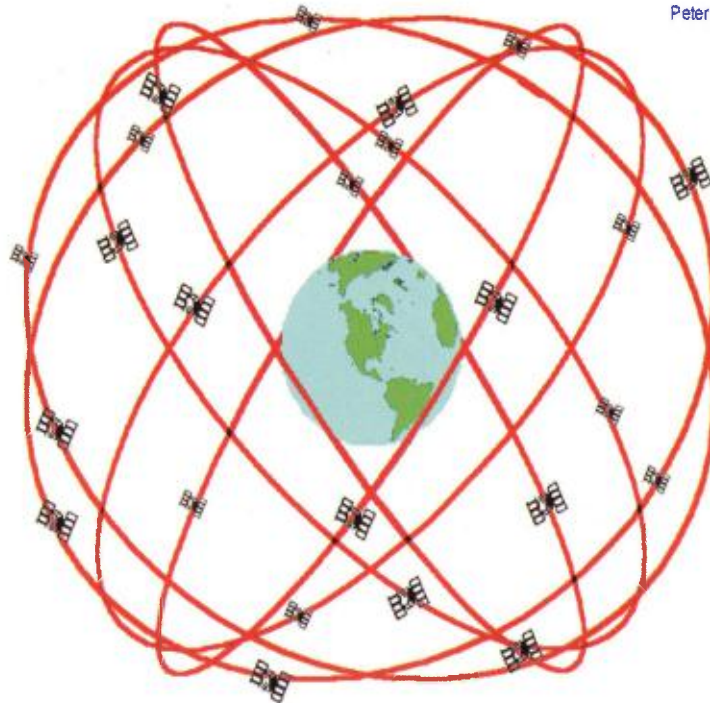
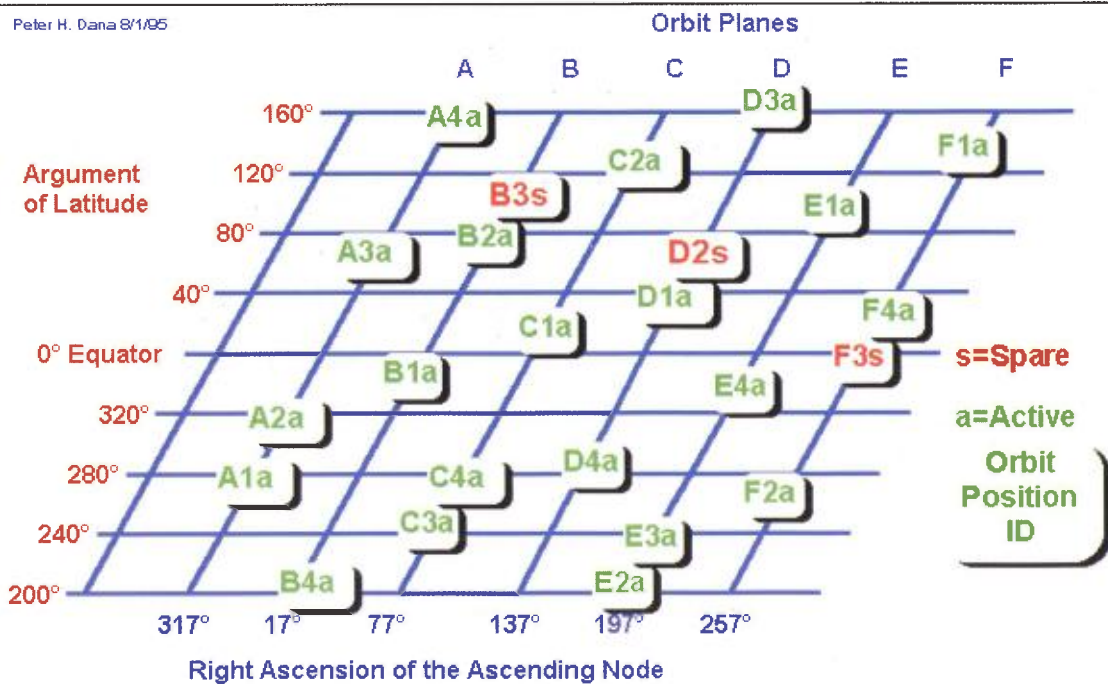


Fig. 6.



GPS Nominal Constellation
24 Satellites in 6 Orbital Planes
4 Satellites in each Plane
20,200 km Altitudes, 55 Degree Inclination

Fig. 7.
 The constellation of satellites and orbits



Simplified Representation of Nominal GPS Constellation

Fig. 8.
 The codification of satellites and orbits

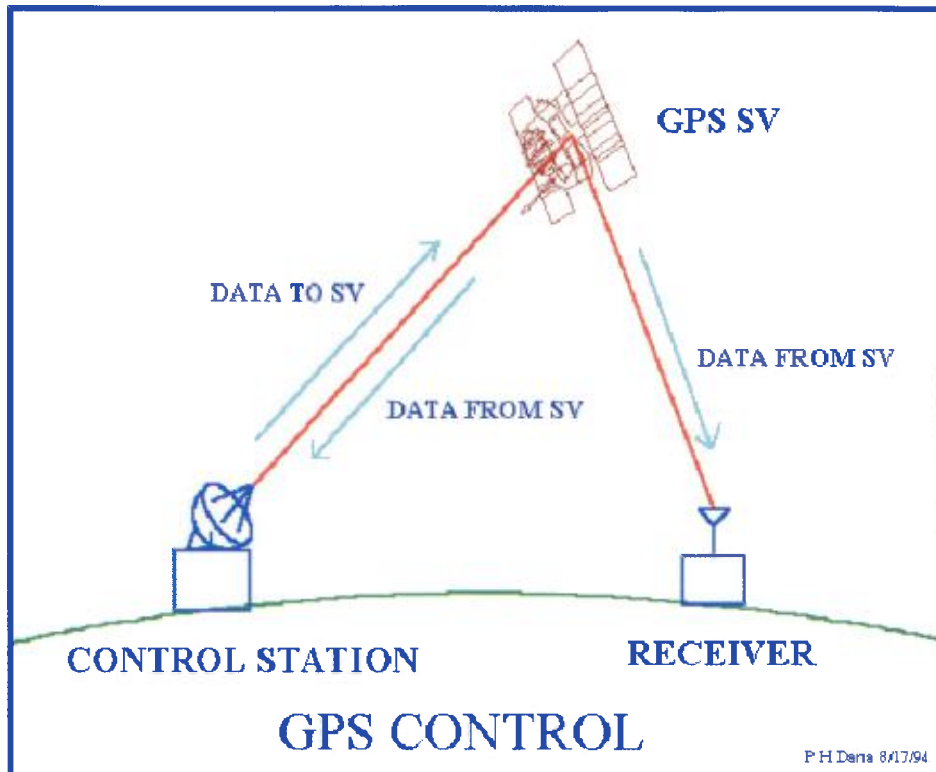


Fig. 9.
The control function



Fig. 10.
The five control stations of the NAVSTAR

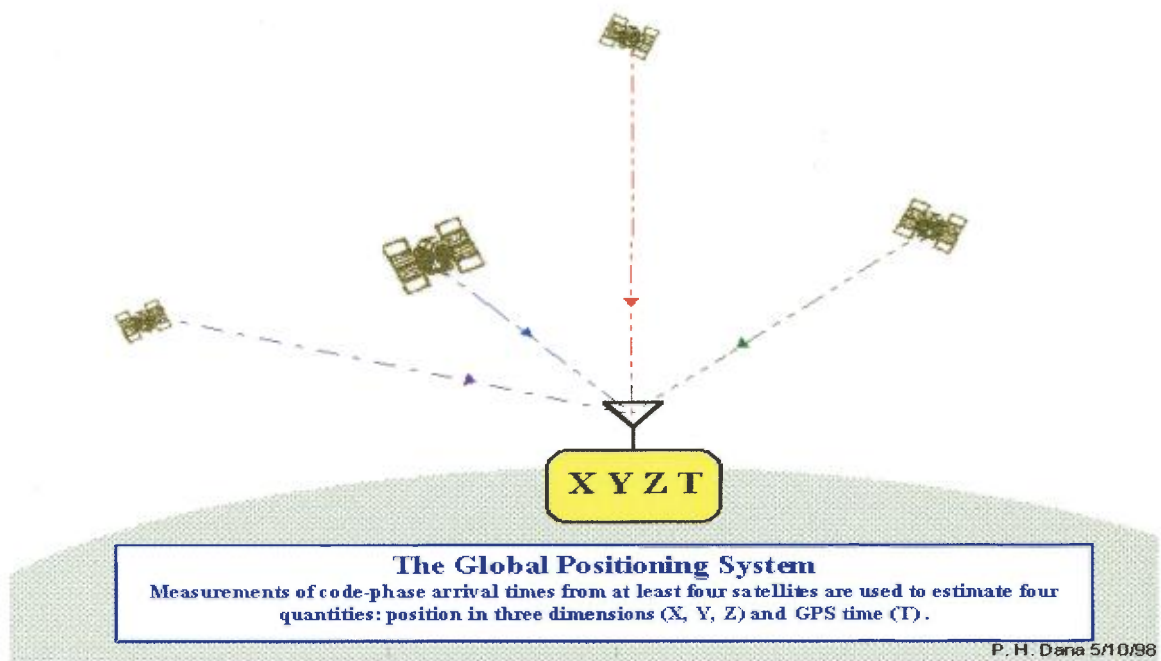


Fig. 11.
The optimal satellite and receiver constellation



Fig. 12.
Pocket GPS navigation instrument

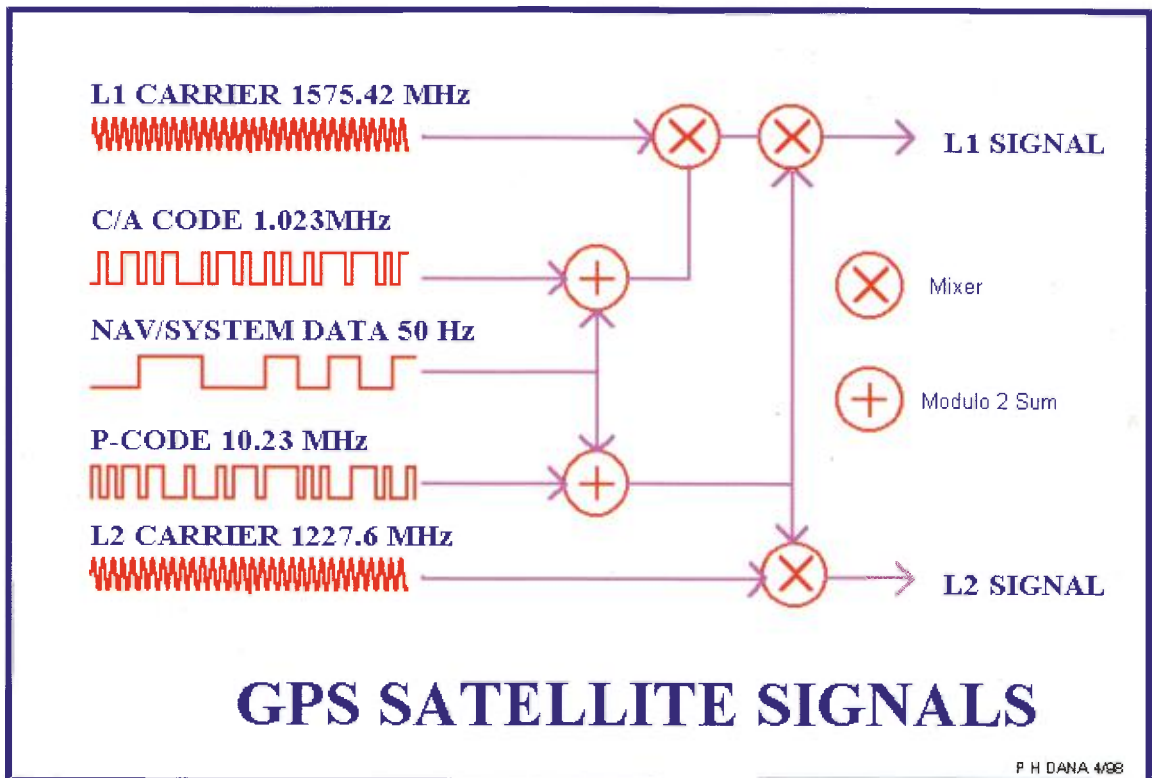


Fig. 13.
The GPS signals

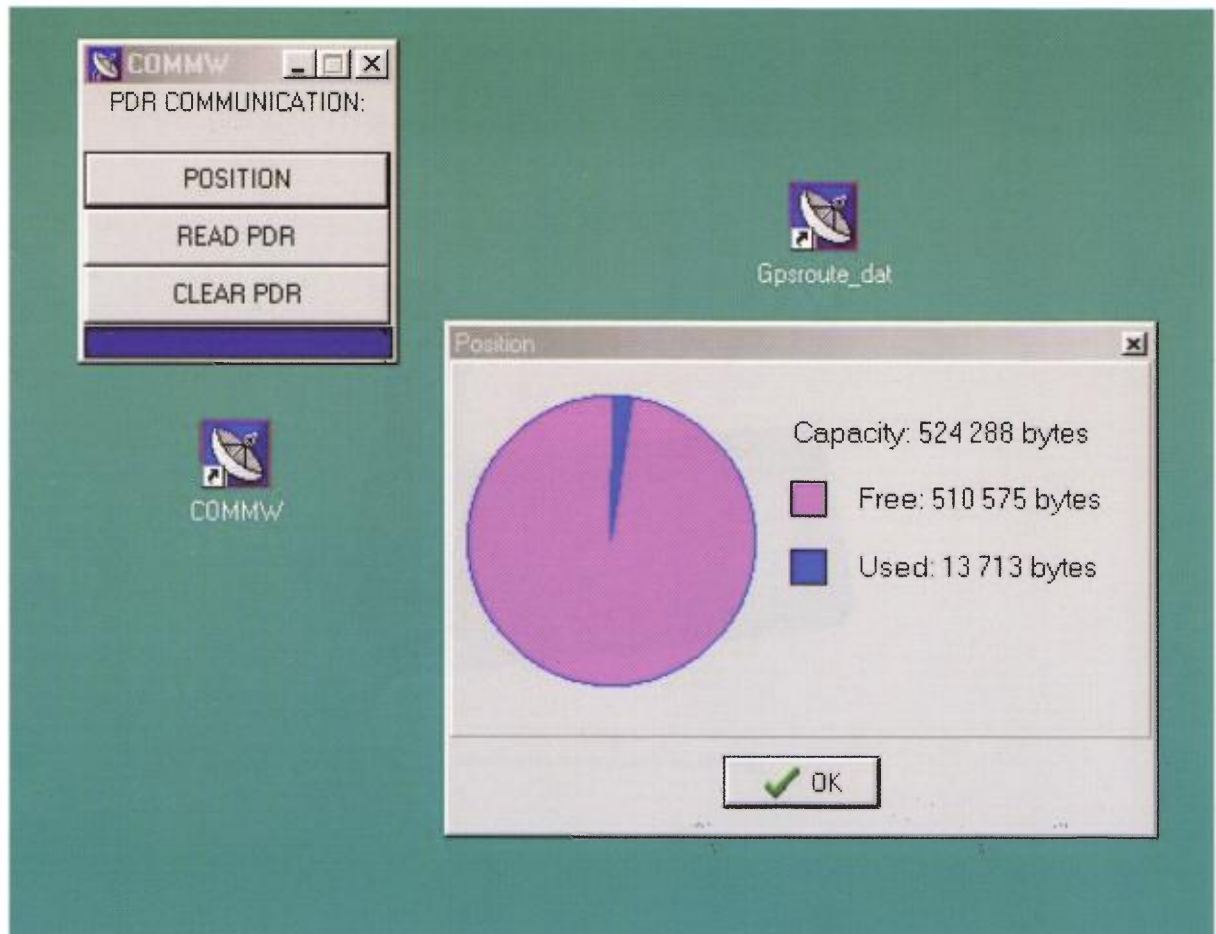


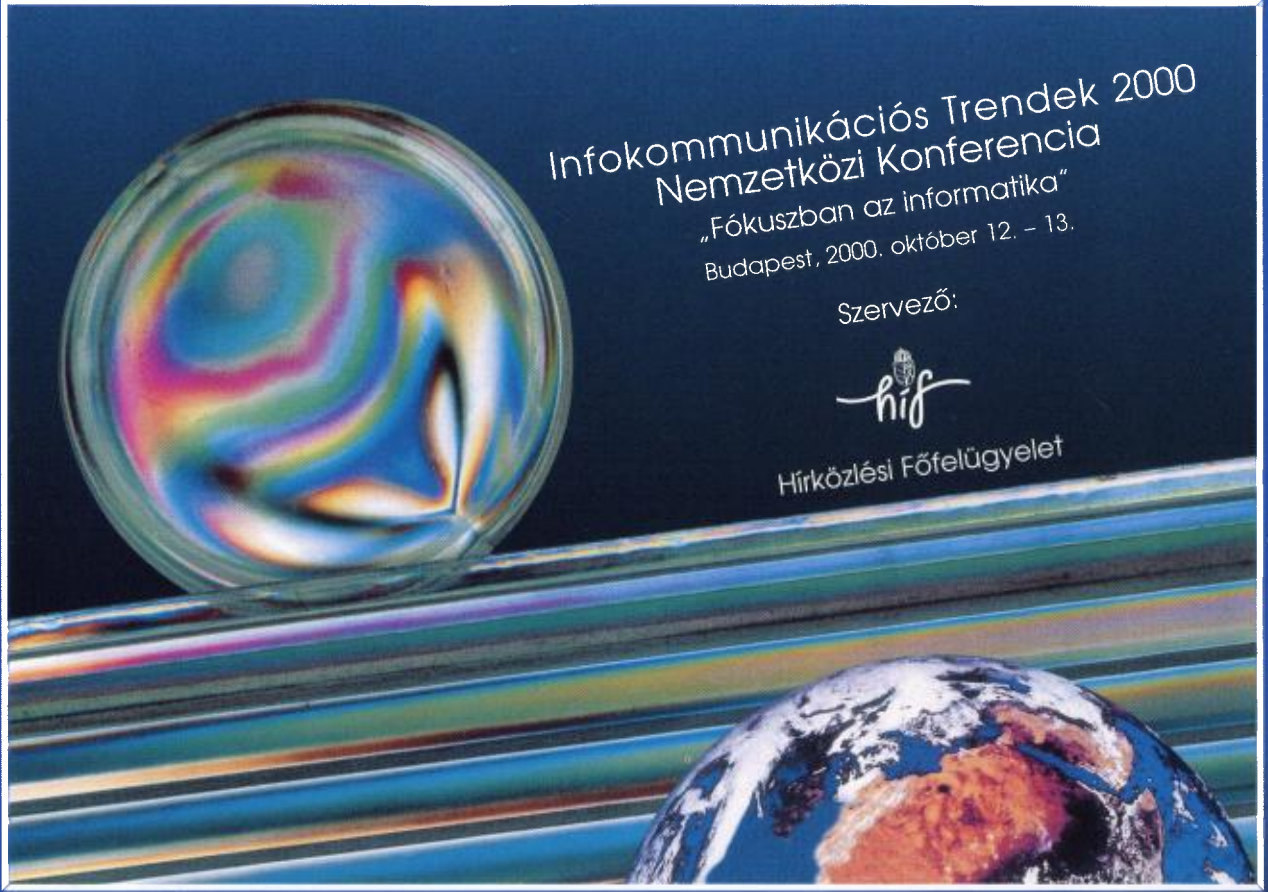
Fig. 22.
The screen of COMMW program



Fig. 24.



Fig. 25.
The PDR instrument



Infokommunikációs Trendek 2000
Nemzetközi Konferencia
„Fókuszban az informatika”
Budapest, 2000. október 12. – 13.

Szervező:



Hírközlési Főfelügyelet

Infokommunikációs Trendek 2000 Nemzetközi Konferencia „Fókuszban az informatika”

A

Hírközlési Főfelügyelet

az idei évben is megrendezi az infokommunikáció
különböző területeivel foglalkozó nemzetközi konferenciáját.

A 2000. október 12-13. között megrendezésre kerülő kétnapos konferencia szekciói:

Szabályozási trendek

Mobil Internet

Információtechnológiai körkép

Az infokommunikáció biztonsága

A konferencia előadói között szerepelnek az Európai Bizottság Információs Társadalom Főigazgatóságának Igazgatója, több európai szabályozó hatóság vezetője, a WAP Fórum alapítójának képviselői és más nemzetközi szervezetek vezetői, valamint az infokommunikáció nemzetközileg elismert további magyarországi és külföldi személyiségei.

A rendezvényt a Magyar Tudományos Akadémia támogatja.

További felvilágosítás: Tel.: 457 72 75; Fax: 457 71 71; E-mail: szathmary.gabor@hif.hu

**A Konferencia szervezője:
Hírközlési Főfelügyelet (HIF)**

Fővédnök:

FRISCHMANN Gábor, Elnök - HIF

Védnökök:

Dr. BECK György, Elnök - Informatikai Érdekegyeztető Fórum (INFORUM)

Dr. KROÓ Norbert, Főtitkár - Magyar Tudományos Akadémia (MTA)

Dr. PAP László, Elnök - Hírközlési és Informatikai Tudományos Egyesület (HTE)

SÍK Zoltán, Informatikai Kormánybiztos

Dr. ZOMBORY László, Elnök - Nemzeti Hírközlési és Informatikai Tanács (NHIT)

Szervező és program bizottság:

Elnök:

Dr. SALLAI Gyula, Elnökhelyettes - HIF

Alelnök:

Dr. SCHMIDEG Iván, Igazgató - HIF

Tagok:

Dr. BARTOLITS István, Tanácsadó - HIF

Dr. ERÉNYI István, Főosztályvezető - Miniszterelnöki Hivatal (MEH)

Dr. RÉNYI István, mb. Igazgató - HIF

SZATHMÁRY Gábor, Osztályvezető - HIF

Hírközlési Főfelügyelet
Nemzetközi Igazgatóság

Cím:

1525 Budapest, P.f.: 75

Tel.: 457 72 75

Fax: 457 71 71

E-mail: contact@hif.hu

KONFERENCIA PROGRAM

2000. OKTÓBER 12.

Szekció 1: Szabályozási trendek

Szekcióelnök: Dr. SALLAI Gyula

Elnökhelyettes - HIF

Az EU-hoz csatlakozó országok kihívása: az Információs Társadalom felé vezető út
SÍK Zoltán, Informatikai Kormánybiztos

Az elektronikus kommunikáció új EU szabályozási rendszere
ARGYRIS, Nicholas, Igazgató - EC Információs Társadalom Főigazgatóság

Az infokommunikációs piac és szabályozása Magyarországon
FRISCHMANN Gábor, Elnök - HIF

Infokommunikációs szabályozási trendek Európában
FURRER, Marc, Vezérigazgató OFCOM, Svájc

A kommunikáció szabályozása Olaszországban
di ROBILANT, Filippo, Nemzetközi Igazgató -
Autorità per le Garanzie nelle Comunicazioni, Olaszország

*Szélessávú Internet hozzáférés:
a helyi hurok szétválasztása; harmadik generációs mobil távközlés*
Dr. ARNBAK, Jens C., Elnök OPTA, Hollandia

EBÉD

Szekció 2: Mobil Internet

Szekcióelnök: Dr. PAP László

Elnök - Hírközlési és Informatikai Tudományos Egyesület

WAP és a Mobil Internet: új szolgáltatások és trendek
DAHM, Cindy Nelson, Marketing Igazgató - Phone.com

WAP technológia a mobil Információs Társadalom szolgálatában
Dr. ÉLŐ Gábor, Kutatási és Fejlesztési Igazgató - Nokia

A Mobil Internet sikerének kulcsa
NASEEM, Asif, Főigazgató - Motorola

Elektronikus fizetés és számlázás mobil Internet és e-commerce környezetben
Dr. KOCSIS, Frank, Mobil Internet Igazgató - AMS

Mobil Internet: a kézzel fogható konvergencia
Dr. BARTOLITS István, Tanácsadó - HIF

Mobil Internet a médiaszociológus szemével
Dr. SZEKFŰ András, Adjunktus - Pécsi Tudományegyetem

ÁLLÓFOGADÁS

KONFERENCIA PROGRAM

2000. OKTÓBER 13.

Szekció 3: Információtechnológiai körkép

Szekcióelnök: Dr. ERÉNYI István
Főosztályvezető - Miniszterelnöki Hivatal

Új eszközök és alkalmazások az Internet világában

Dr. DÖMÖLKI Bálint, Elnök - IQSoft

IPv6 Az új generációs Internet

LADID, Latif, Alelnök - Ericsson Telebit A/S; Elnök - IPv6 Forum

Az Inforum szerepe a magyar informatikai közösségben

Dr. BECK György, Elnök - Informatikai Érdekegyeztető Fórum (INFORUM)

Technológiai szabályalkotás az infokommunikációban

Dr. KISS Ferenc, Tanácsadó - HIF

Az infokommunikációs jog aktuális kérdései az EU-ban és Magyarországon

Dr. ORMÓS Zoltán, Ormós Ügyvédi Iroda

Az Internet szerzői jogi kérdései, lehetséges hazai megoldások

Dr. GYERTYÁNFY Péter, Főigazgató - Artisjus Magyar Szerzői Jogvédő Iroda Egyesület

EBÉD

Szekció 4: Az infokommunikáció biztonsága

Szekcióelnök: Dr. SCHMIDEG Iván
Igazgató - HIF

Kompatibilis elektronikus aláírás infrastruktúrák és alkalmazások

Dr. ENDERSZ György, Elnök - ETSI SEC ESI; Tudományos Tanácsadó - Telia Research

Kulcsmenedzsment: a kriptográfiai megoldások kulcskérdései

Dr. VAJDA István, Egyetemi tanár - Budapesti Műszaki és Gazdaságtudományi Egyetem (BME),
Híradástechnikai Tanszék

Elektronikus aláírás: jogszabályi és szabványosítási kérdések

Dr. RÉNYI István, Informatikai szabályozási osztályvezető - HIF

A pénzüintézetek elektronikus kereskedelme

Dr. BAKONYI Péter, Elnök - Magyar Internet Társaság (MIT)

Információs társadalom az EU-ban

Dr. KROÓ Norbert, Főtitkár - Magyar Tudományos Akadémia (MTA)

ZÁRSZÓ



T-Network Kft.
1142 Budapest, Ungvár u. 64-66.
Tel: (+361) 460 9000 Fax: (+361) 460 900
Email: tnetwork@tnetwork.hu



T-Network Kft.
1142 Budapest, Ungvár u. 64-66.
Tel: (+361) 460 9000 Fax: (+361) 460 900
Email: tnetwork@tnetwork.hu



T  **N**



T-Network Kft.
1142 Budapest, Ungvár u. 64-66.
Tel: (+361) 460 9000 Fax: (+361) 460 900
Email: tnetwork@tnetwork.hu



T-Network Kft.
1142 Budapest, Ungvár u. 64-66.
Tel: (+361) 460 9000 Fax: (+361) 460 900
Email: tnetwork@tnetwork.hu





Foglaljon helyet az éterben!



Gondolt már arra, hogy létezik magasabb szintű kommunikáció?
Gondolt már arra, hogy kábeli zűrzavar nélkül is információt cserélhet?
Gondolt már arra, hogy egyszerűbben és gyorsabban teremthet kapcsolatot partnereivel?

Gondolt már az Antenna Hungáriára?
Mi műsorterjesztési berendezéseinkkel és digitális távközlési hálózatunkkal számtalan helyet kínálunk az éterben az Ön gondolatai számára is. Nálunk az adat adott, és minden információ házhoz jön. Légvonalban.

Tisztelt Előfizetők!

A 2000. évi Híradástechnika lapok 12 száma (több mint 1000 oldalon, aminek kb. a harmada 4-színes oldal) 2000. októberében már kézben lesz. A lap éves előfizetői ára bruttó 100 US\$ volt 2000-ben és annyi marad 2001-ben is. Időközben azonban az árfolyam-változások már jelentős (33%-os) különbséget eredményeznek a dollár alapú és az idei hazai, 22400Ft-os bruttó közületi előfizetői ár között. Ezért 2001. január 1.-től 30000.-Ft-ra kényszerülünk emelni az éves, számlával fizetett bruttó közületi, és 6000.-Ft-ra a bruttó egyéni előfizetői díjat.

Az előfizetői árváltozás áthidalására - akcióban - 2000. december 31.-ig a HTE-nél még a régi, 22400Ft-os bruttó áron lehet megrendelni a közületi, és 2240 Ft-os bruttó áron az egyéni előfizetéseket a 2001-es évre. (Ez az egyéni előfizetői ár a HTE 760Ft-os bruttó hozzájárulásainak levonásával alakult ki a 3000 Ft-os bruttó árból.) Az írásos megrendeléseket a 3416-421 és/vagy a 470-0713 fax számokra kérjük küldeni, vagy a sese@mail.mata.v.hu e-mail címre.

A Híradástechnika 2001-ben magas színvonalú tutorial sorozattal kívánja kiegészíteni témaköreit, amellyel a gyakran visszatérő továbbképzési igények rendszerezett és optimális kielégítésére törekszik.



www.antenna.hu

antenna  hungári
Teret adunk gondolatainak



Amint kiértél a partra,

telefonálj be, hogy beteg vagy,

aztán nézd meg a hálón,

milyen idő várható,

és azt is tudni fogod,

még hány napig leszel beteg.

A stresszes, hajszás üzleti világban mindent ki kell használni, amit csak lehet. A Motorola Timeport P7389 éppen ebben segít. Hangodra ugrik, ha a kezed foglalt, micro web keresőjével minden szükséges információ megtalálhatsz az interneten. Háromsávós, ezért egyaránt működik Európában, Ázsiában és Amerikában*. (Az csupán ráadás, hogy elég kicsike a kicsike, hogy a strandtáskába is behuppanjon.) Nos, a legközelebbi hullámlovaglást tekintsd úgy, mint egy kihelyezett igazgatósági ülést. A Motorola Timeport P7389. Ha nem volna, ki kellene találni.

WWW
MMM
MOBILE MEDIA MODE



MOTOROLA

TIMEPORT

GSM 900/1800/1900 MHz + WAP



* Ahol mobilszolgáltatójának barátságosi megállapodása van érvényben.

A Hotel Inter-Continental sikerei mögött ott van a Kapsch intelligens kommunikációs rendszere is.



Keresztes György,
a Hotel Inter-Continental
műszaki igazgatója

A Kapsch innovatív telefonközpontja többnyelvű hangpostarendszerével és speciális telefonkészülékeivel az Inter-Continental Budapest szállodában megnövelte a hatékonyságot, és csökkentette a működési költségeket. Valamennyi vendég állandóan elérhető és gyors Internet-hozzáféréssel rendelkezik. Kapsch. The communications company: (1) 209 2110, www.kapsch.net/kt_hu

 **KAPSCH**
the communications company



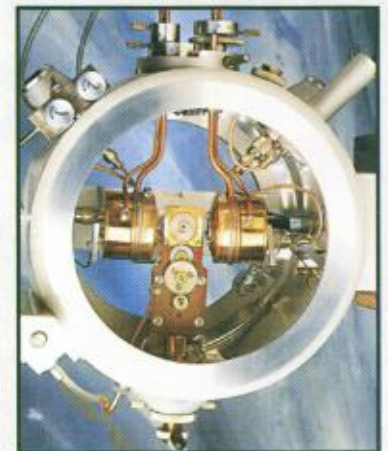
Transmission Electron Microscopy Excellence in specimen preparation

Facilities:

- TEMs working at 100 and 200 kV with different analytical tools
- 6 ion beam milling units for TEM sample preparation
- tools for the mechanical grinding and polishing of TEM samples
 - dimplers, ultrasonic cutter
- 3 different ion guns can be used for different purposes
- evaporators in the high vacuum and ultrahigh vacuum range
- magnetron sputtering systems for layer deposition



Top view of an ion mill showing the arrangement of the two ion guns and of the specimen stage. The unit was manufactured by Technoorg-Linda. Typical milling time is a few hours after mechanical treatment (polishing down to 50 μm).



Cross sectional specimen of a silicon integrated circuit prepared by a new Ar ion gun.



Typical fields of application:

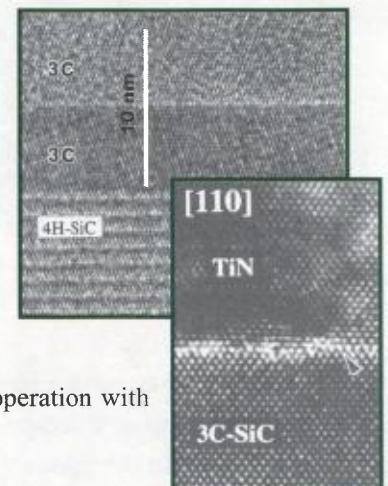
15 scientific researchers working in three major fields

Wide bandgap semiconductors, SiC and GaN

GaN and SiC are the most important wide bandgap semiconductors promising for high temperature, high power microelectronics. GaN with its direct bandgap is already one of the most important material of optoelectronic being the blue laser material. Growth of the above two materials achieved very different levels. Bulk growth of SiC was solved, while GaN is grown in the form of heteroepitaxial layers, on sapphire or SiC.

CVD grown, thick SiC layers have been grown onto hexagonal SiC substrates. To avoid the nucleation of 3C polytype in the hexagonal layers the optimum growth parameters have been determined.

High temperature stable ohmic contacts have been developed to SiC in cooperation with Linköping University, Sweden.



GaN activity is connected to THOMSON LCR, Orsay, France. The MOCVD growth of GaN onto miscut sapphire was optimised. Thick GaN layers with specular surfaces were grown. Inversion domains in GaN were found.

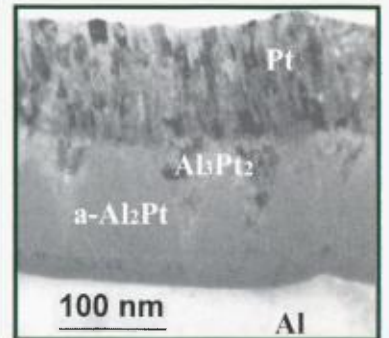


Structure, formation and characterisation of thin films

A new microcrystalline composite thin film coating (Al-Sn20) has been developed in the frame of EUREKA 338 Socomat project with TU Wien, MIBA Graitlager AG Austria and MTU Motoren und Turbinen Union, Germany for sliding bearings applied in combustion engines. This environmental friendly surface coating has also remarkably improved the wear/lifetime property of the bearings.



The solid phase reaction of Al and Pt at the high temperature sequential deposition results in the successive formation of amorphous, intermetallic and metallic phases. The dependence of the critical thickness of the amorphous phase on the parameters made possible the determination of the diffusion coefficient of Al in the Al₂Pt phase at first.



Growth and characterisation of CN_x hard coatings



*Cross sectional TEM image of a CN_x multilayer.
Cooperating partner:
Forschungsinstitut
Rossendorf*



*HRTEM image of
CN_x film deposited
onto NaCl substrate
at 0.1 mbar partial
pressure of N₂
at ambient temperature*

Cooperations

cc. 20 partnership with universities and research centers, like

- CNRS Grenoble, France
- THOMSON LCR Orsay, France
- Linköping University, Sweden
- Aristotle University of Thessaloniki, Greece
- Forschungsinstitut Rossendorf

TEM sample preparation of “difficult” samples
Samples can be prepared from wafers, wires, tapes and powders

Detailed TEM analysis of the samples in cross section or plan view.

Characterization of interfaces, layered structures, grain boundaries.

Phase determination of reaction products by selected area electron diffraction

Morphological analysis and separation of phases by bright field/dark field technique

Determination of the crystallographic orientation and polarity by CBED (Convergent Beam Electron Diffraction) or EDS (Energy Dispersive System)

Analysis of different regions to determine the components, elements

Failure analysis

Answers to your questions in materials science

Partnership in R&D projects, proposals

Feedback to your industrial processes under development

Optimization of new technologies

Involvement of Ph.D. students into collaborative research work

Experience in EU proposals/projects

Background of 30 years spent in microscopy, 15 years in ion milling

A világ tele van meglepetésekkel



Általában.



Nálunk azonban nem történik
semmi különös. A MATÁV vonalain
most is több millióan beszélnek,
akárcsak máskor.

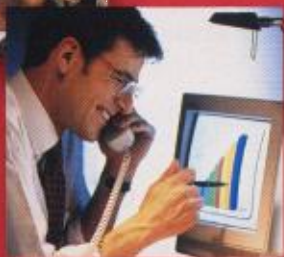
Közülük sokan veszik igénybe
az üzleti kommunikációs
szolgáltatásokat is, – Vsat, Internet,
Frame Flex, Zöld szám, Kék szám,
stb. – akárcsak máskor.

ISDN-vonalainkon gigabájtok milliói
száguldanak. Cégünk változatlanul
Közép- és Kelet-Európa vezető
telekommunikációs cége,
és továbbra is aktív szereplője
a nemzetközi pénzvilágnak.
Szóval nagyjából ennyi.

Ja! Azért egyvalamiről mégis
érdemes szót ejteni.

Ma kicsit lehűlt a levegő.

CRE ACTION



- nagy pontosság igénye;
- Túl drága vagy időigényes lenne az ember helyszínre szállítása;
- Környezet számára ártalmas az emberi jelenlét.

1.2. Mikrorobotika

A mikrorobot megnevezés a szakirodalomban kétféle értelemben is használatos. Tágabb értelemben arra utal, hogy ezeknek a robotoknak mind a mérete, mind az akciótere, mind pedig a pozícionálási felbontása kisebb, mint az általunk eddig „megszokott”, például ipari robotoknak. Szűkebb értelemben véve pedig a mikrorobotot a néhány mikrométer nagyságrendű manipulációt lehetővé tévő robotok osztályát jelöli. A mikrorobot kifejezést cikkünkben ezen túl az utóbbi értelmében használjuk.

Ezeknél a robotoknál már a hagyományos elveken alapuló érzékelő- és mozgó-berendezések nem alkalmazhatóak. Tervezésük és működtetésük során számolnunk kell azokkal a mikrojelenségekkel, amelyek a makrovilágban egyébként elhanyagolhatóak (felületi feszültség, adhéziós és kohéziós jelenségek). Pozíciójuk pontos érzékeléséhez külső szenzorokat, pl. mikroszkópra szerelt kamerát, vagy lézeres távolságmérőt használnak.

A technológia fejlődésével a mikrorobotika előtt egyre szélesebb alkalmazási területek nyílnak meg. Az iparban főleg szerelési munkálatok elvégzésére fejlesztenek

mikro-gyártósorokat; a kémiában, a gyógyszergyártásban és a biotechnológiában pedig mikrorobotok végezhetik kis mennyiségű anyagok megfelelő helyre juttatását.

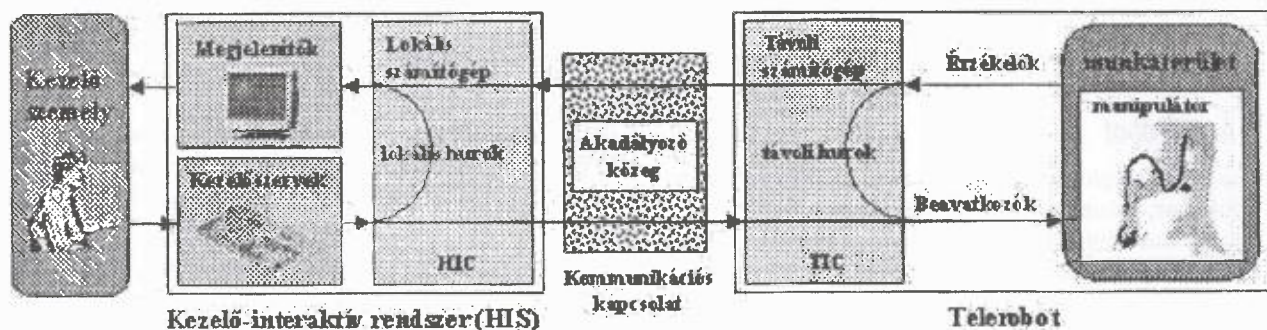
1.3. Robotok a biotechnológiában

A biotechnológia a biológiai és műszaki tudományok összefonódásából jött létre. Összefoglaló neve mindazon területeknek, amelyekben élő szervezeteket, azok részeit, vagy az azokban lejároló folyamatokat használnak fel hasznos anyagok, információk, esetleg más élő szervezetek előállítására. Ilyen terület például virológia és a genetika is.

A biotechnológiai alkalmazások gyakran igényelnek pontos mechanikai beavatkozást. Ennek nagyságrendje a néhány mikrométertől a milliméteres nagyságrendig terjed. Egy ebben a tartományban pozícionálható, intelligens, tanítható, megbízható mikrorobot nagymértékben meggyorsíthatja a munkát, legyen szó akár mesterséges megtermékenyítésről, egy vírustenyésztet gondozásáról, vagy akár egy DNS lánc azonosításáról [8].

2. TELEROBOTIKAI RENDSZER MEGVALÓSÍTÁSA

A telerobotikai rendszerek alapját az emberi felügyeleti irányítás általános modellje adja. Az 1. ábra egy telerobotikai rendszer felépítését szemlélteti.



1. ábra. Telerobotikai rendszerek általános felépítése [18]

A kezelő személy a rendszerrel felhasználói felületen keresztül érintkezik, amely a kezelő-interaktív rendszer (Human-Interactive System – HIS) része. Az általa küldött parancsok a kommunikációs kapcsolaton keresztül jutnak el a telerobotig, melynek érzékelői által szolgáltatott és feldolgozott jelek adnak hírt a parancs hatásáról a felhasználónak, a kommunikációs csatornán és a megjelenítőkön keresztül.

A kommunikációs csatorna a legtöbb telerobotikai alkalmazásban nem elhanyagolható késleltetéssel és csak korlátozott sáv szélességgel képes az adatokat továbbítani. Mivel konkrét rendszerünket Internetre tervezzük - amelynek kellemetlen tulajdonsága a viszonylag nagy és ráadásul dinamikusan változó holtidő –, így a közvetlen távirányítás lehetősége gyakorlatilag kizárt: Az embert és a holtidős kommunikációs csatornát is magába foglaló zárt szabályozási kör elfogadhatatlanul lassú lesz. (Egyszerű feladatok elvégzése is hosszú időbe telhet, ugyanis az instabilitás elkerülésére a kezelő személy általában a „move and wait” stratégiához folyamodik: beavatkozik, majd megvárja, amíg

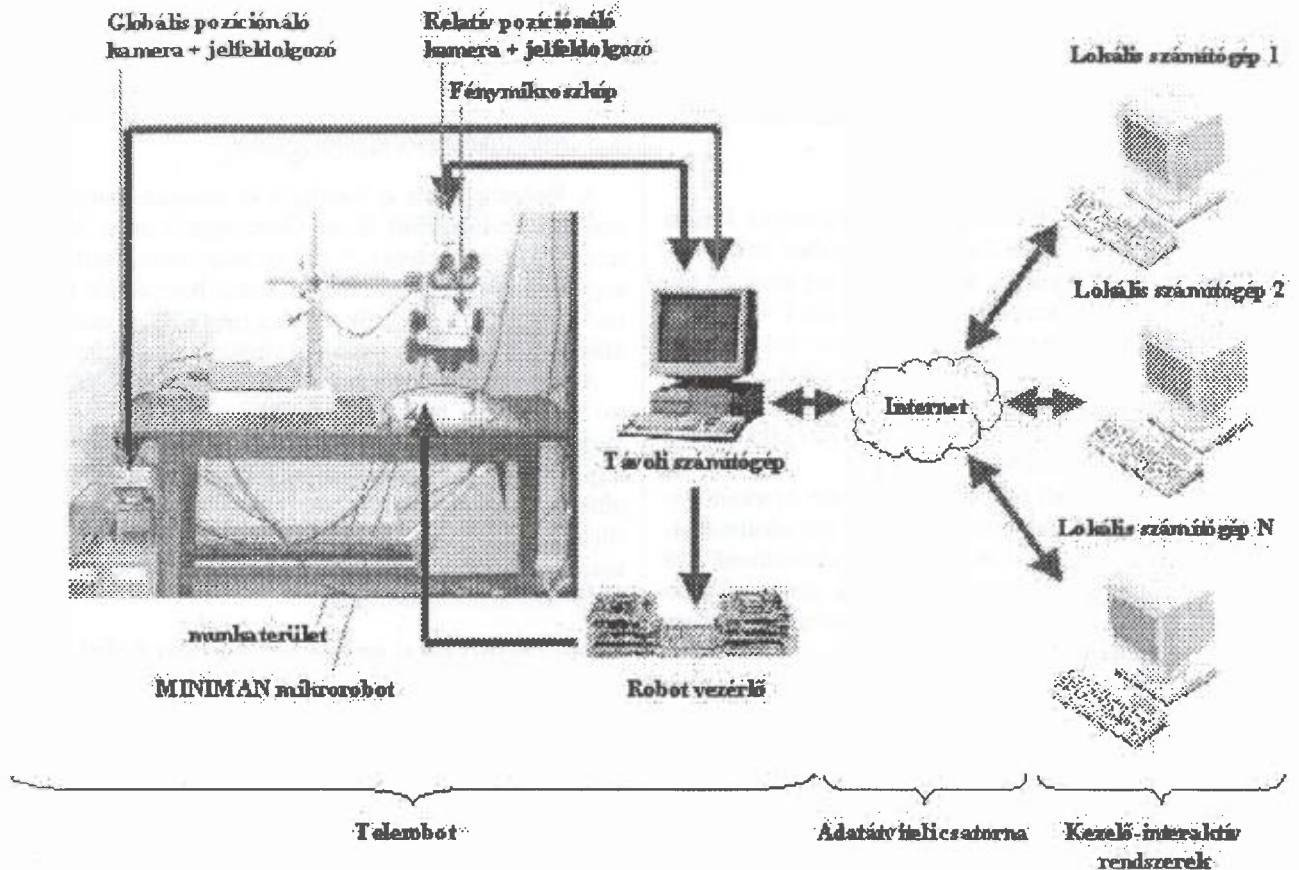
a művelet végrehajtásának eredményéről szóló információk megérkeznek.)

Ez a komoly szabályozástechnikai probléma két további visszacsatoló hurok beiktatásával oldható meg.

Az ún. távoli hurok egy olyan intelligens szabályzó kör, amely váratlan események bekövetkeztével a megfelelő gyors reagálásért, valamint a kommunikációs kapcsolat ideiglenes kiesése esetén a robot autonóm mozgásáért felelős. Az így már bizonyos autonómiával rendelkező telerobotot kezelő személynek a feladata széles skálán megválasztható (lehetőségek szerinti közvetlen irányítástól a stratégiai szintig), magasabb szintű irányítás.

A lokális hurok ergonómiai visszacsatolás a kezelő személy felé, amely gondoskodik a parancsok kiadásának, illetve a várható hatásának azonnali visszajelzéséről, lehetővé téve ezzel a kommunikációs kapcsolattól független folyamatos munkát.

Az általunk tervezett rendszer a 2. ábrán látható felépítést követi.

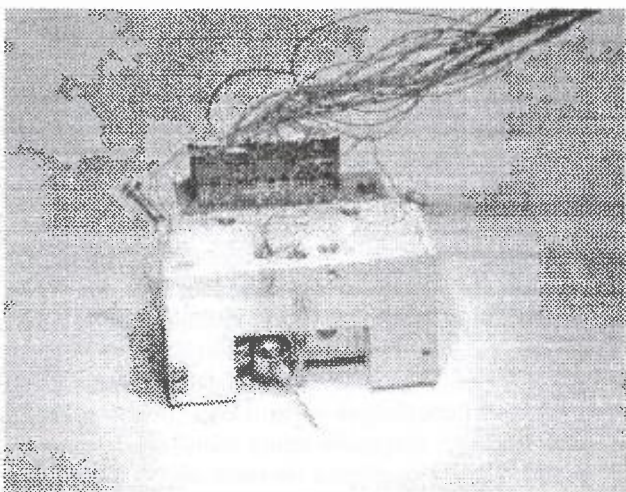


2. ábra. A tele-mikrorobotikai rendszerünk felépítése

2.1. A telerobot

A telerobot legfontosabb részei a manipulátor, az érzékelőrendszer, valamint a távoli számítógép.

Manipulátorként a tanszéken folyó nemzetközi projekt keretében kifejlesztésre került MINIMAN mikrorobotot alkalmaztuk (3. ábra).



3. ábra. A MINIMAN mikrorobot

A távoli számítógép megválasztásakor a megbízhatóság, teljesítmény, költségek, beszerezhetőség, hardver és szoftver támogatottság szempontjainak figyelembe vétele alapján egy egyszerű IBM PC-kompatibilis számítógép mellett

döntöttünk (Pentium osztályú processzorral, és megfelelő perifériaillesztőkkel ellátva). A szenzorrendszer szerepét két jelfeldolgozó processzorral ellátott intelligens kamera tölti be.

2.1.1. A telerobot manipulátora, a MINIMAN mikrorobot

Mikrorobotunk első prototípusát nemzetközi kutatási projekt keretében fejlesztették ki 1996-ban a Karlsruhei Egyetem Folyamatszabályozási és Robotika Intézetével együttműködve [11].

A robottest egy 30x30x60 mm élhosszúságú tömör alumíniumhasáb, amely három piezokristályból készült lábön áll. Az elülső láb mellett helyezkedik el a beavatkozószerv (kar), amely egy acélgolyóhoz, mint gömbcsuklóhoz rögzített orvosi injekciós tű. A robot lábainak mozgása az inverz-piezohatáson alapszik. Ha a lábakon kialakított elektródákra megfelelő feszültséget kapcsolunk, a lábak kissé elhajlanak. Az elhajlás mértéke és iránya a lábankénti négy elektródapárra kapcsolt feszültség nagyságától függ.

A robot úgynevezett „tapadó-csúszó” („Stick-and-Slip”) elven végzi a mozgását. A robotlábakra fűrészelj alakú, 150 V amplitúdójú jelet adunk. Ennek hatására a lábak a kívánt irányban lassan meghajlanak, a robottest pedig a tapadási súrlódásnak köszönhetően az elhajlás mértékében a megfelelő irányba mozdul. A lassú elhajlás után a láb hirtelen visszaugrik kiindulási helyzetébe, olyan sebességgel, hogy a robottest gyakorlatilag helyben marad, a láb pedig csúszik a talajon. A robot egy ilyen lépéssel néhány mikrométer távolságot tesz meg, a folyamatos haladás során ezt másodpercenként sok ezerszer ismétli.

A beavatkozószerv mozgatása is hasonló elven történik. Amint a 3. ábrán is látható, a robot karja gömbüzülettel kapcsolódik a robottesthez, s az ízület ugyanolyan piezolábak hajtják, mint a robottestet. A mozgó acélgömböt egy erős mágnes húzza a lábak felé.

A mikrorobot közvetlen vezérléséért, valamint a központi számítógép felé magas szintű felhasználói felület nyújtásáért egy párhuzamos mikrokontroller rendszer felelős. A vezérlőegység feladatai: a központi számítógép parancsainak fogadása párhuzamos porton keresztül, MINIMAN lábaira kerülő feszültség szintek előállítása, és státuszinformáció szolgáltatása soros porton keresztül az értelmezett és végrehajtott parancsokról.

A számítógép által a központi mikrokontroller kártyának küldhető parancsok segítségével beállítható a mikrorobot lábainak rezgési frekvenciája, amplitúdója és iránya, valamint lehetőségünk van a robot által megteendő lépések számának megadására is.

2.1.2. A telerobot szenzorrendszere [1], [2]

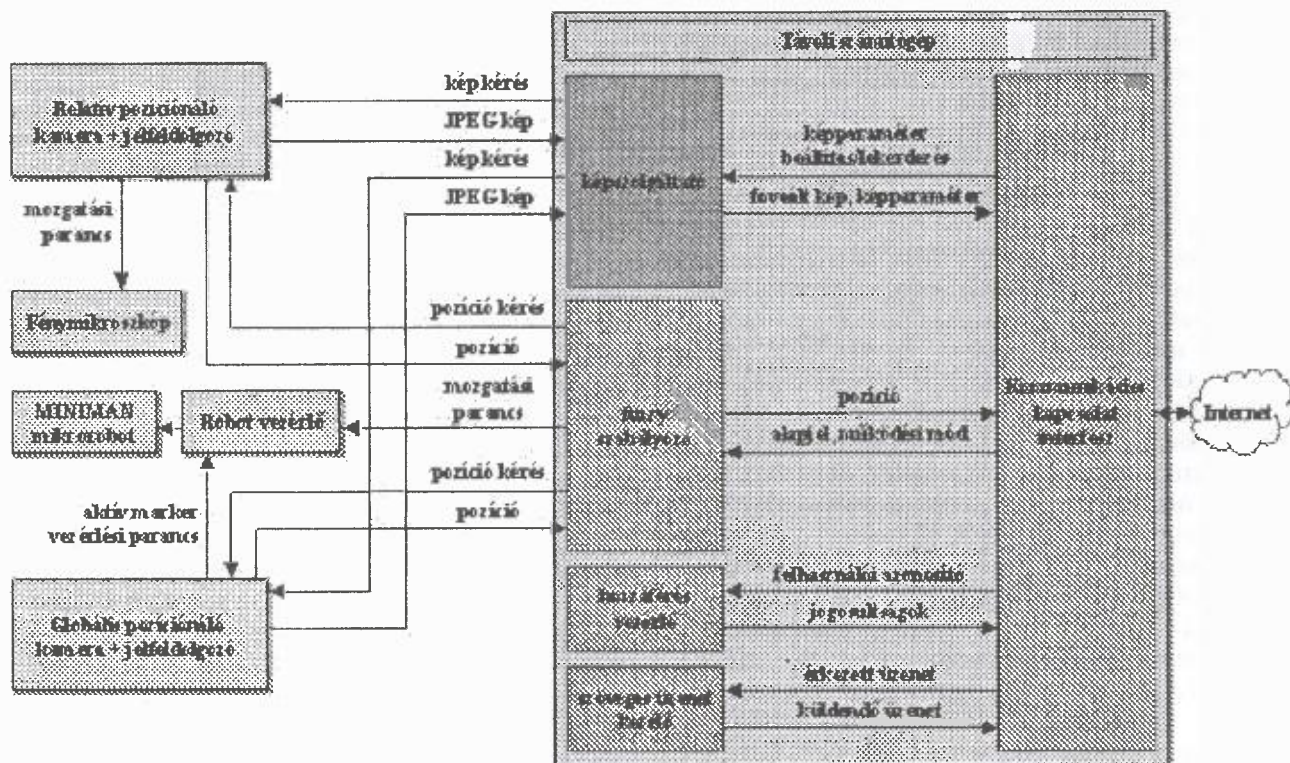
A MINIMAN munkaterülete egy 200x300mm-es üveglemez. A szenzorrendszer feladata, hogy ezen – a robot pontosságához képesti – nagy méretű területen a robottest és -kar helyzetének meghatározása mikrométeres pontossággal. Ehhez érintkezésmentes külső érzékelőket választottunk. Két videokamerát alkalmaztunk; az egyik a teljes munkaterületet átlátja, globális kép és robotpozíció

szolgáltatásáért felelős, míg a másik egy mikroszkópon keresztül egy kisebb területet figyel, és itt biztosítja a mikrométeres pontosságot. (Ez a kialakítás jól illeszkedik a tipikus „nagy méretű tárolóból anyag felvétele, és nagy pontosságú célba juttatása” feladathoz, illetve a kamera mozgatásával a nagy pontosságú pozícionálhatóság a munkaterület tetszőleges részén biztosítható.) Mindkét kamera egy PC-hez csatlakozik, amely a telerobotikai rendszerünkben a „távoli számítógép” (Task Interactive Computer – TIC) szerepét tölti be.

A rendszer hatékonyságának növelése érdekében a Vision Components cég VC11 típusú intelligens kameráit alkalmaztuk, melyet kifejezetten valós-idejű képfeldolgozási feladatokhoz fejlesztettek ki. Ezekbe a kamerákba a gyártók egy Analog Devices ADSP 2181 típusú jelfeldolgozó processzort integráltak. A pozíciók meghatározása így a kamerákon belül megvalósítható, ezzel tehermentesíthető a távoli számítógép, és képátvitelt biztosító szélessávú kapcsolat soros vonallal helyettesíthető, amelyen keresztül a kamera csak a képfeldolgozási algoritmus eredményét közli (de szükség szerint ezen a vonalon képek vagy képrészletek is továbbíthatóak).

2.1.3. A telerobot működése

A telerobot működését alapvetően a manipulátorral, az érzékelőkkel és a kezelő-interaktív rendszerekkel kapcsolatot tartó távoli számítógép határozza meg (4. ábra).



4. ábra. A telerobot és a távoli számítógépen futó szoftver működésének blokkvázlata

A telerobotot működtető szoftvert Windows NT operációs rendszer alá fejlesztettük ki, Visual C++ 6.0 programnyelv felhasználásával. A program öt alapvető funkció el látása szerint a következő modulokra osztható: képszolgáltató, fuzzy szabályozó, hozzáférés-vezérlő, szöveges üzenet kezelő és kommunikációs kapcsolat interfész.

A képszolgáltató modul rutinjai a beállított képfrissítési gyakoriságnak megfelelő időközönként aktiválódnak. Ekkor kerül sor a munkaterületről készült képek felvételére, és soros vonalon keresztüli továbbítására. A soros vonalak kezelése a program működésével átlapoltan zajlik, a 32 bites Windows rendszerek (Win32) által nyújtott kommuni-

kációs alapszolgáltatások igénybe vételével [6]. A képeket a a modul a kommunikációs kapcsolat interfésznek továbbítja. A fuzzy szabályozó implementálása jelenleg folyik. A modul feladata a kamerától pozíció adatok kérése és ennek alapján a manipulátor mozgatása a kezelő útmutatásai szerint, a robotvezérlőnek küldött parancsok által. A robotvezérlőnek a megfelelő parancsokat közvetlenül, a nyomtatókezelő megkerülésével küldjük el a párhuzamos vonalon [14]. A modul továbbá megállapítja a pontosnak tekintendő pozíció értékeket a kamerák erre vonatkozó mérési eredményeinek összevetéséből; és ezeket továbbítja a kommunikációs kapcsolat interfésznek.

A hozzáférés vezérlő gondoskodik arról, hogy csak érvényes felhasználói azonosítóval rendelkező személy használhassa a telerobotot; a szöveges üzenet kezelő pedig egyszerűen elküldi az egyes kezelő személyek által beírt üzeneteket az összes telerobothoz csatlakozott kezelőnek a kommunikációs kapcsolat interfész szolgáltatásain keresztül.

A kommunikációs kapcsolat interfész feladata a kezelő-interaktív rendszerektől érkező parancsok megfelelő moduloknak való továbbítása, illetve a modulok által szolgáltatott információk kezelő-interaktív rendszerekhez való eljuttatása. A hálózati kommunikációt a sok platformon támogatott, megbízható TCP socket architektúrával valósítottuk meg.

2.1.4. A mikrorobot pozíciójának meghatározása DSP kamerával

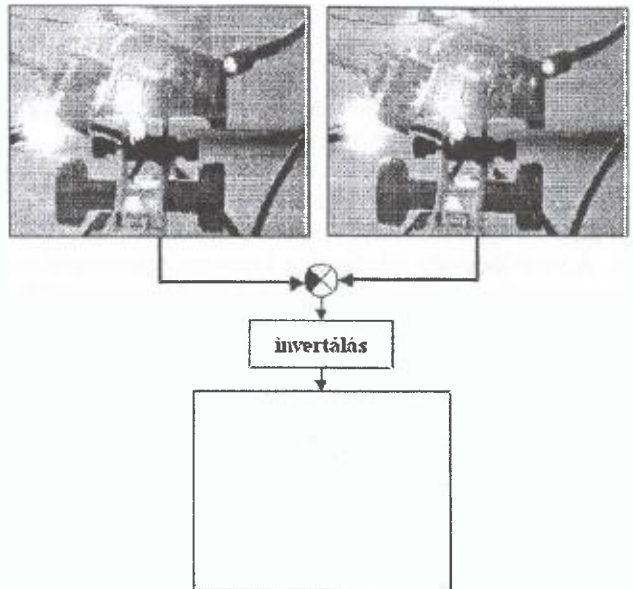
Ahhoz, hogy a mikrorobot bármilyen szintű önálló tevékenységet megbízhatóan tudjon végezni, irányítórendszere számára a munkaterület pillanatnyi állapotát leíró visszacsatoló információt kell biztosítani.

Az irányításhoz szükséges legfontosabb információk: a mikrorobot és a munkaterület többi objektumának

- abszolút pozíciója (mérsékelt pontossági követelményekkel),
- egymáshoz képesti relatív pozíciója, nagy pontossággal (esetünkben ez a néhány mikrométeres nagyságrend).

A robottest pozíciójának meghatározását a munkaterületen aktív markerek segítségével végezzük el. A marker olyan jellegzetes tárgy, amelynek helye egyszerűen meghatározható, mert speciálisan úgy készült, hogy könnyen megkülönböztethető legyen a környezetben lévő összes többi objektumtól. Az aktív szó arra utal, hogy ezt a tulajdonságát a marker valamilyen aktív működés végzésével éri el. A legelterjedtebben ilyen célra koncentrált fényforrásokat alkalmaznak. Ezt a módszert választottuk mi is: mikrorobotunk testére markerekként három világító LED diódát rögzítettünk. A mivel a három marker a robothoz képest sosem mozdulhat el, ezért ha ismerjük a jelölők pozícióját, ebből egyértelműen meg tudjuk határozni a robottest pozícióját.

A pozíciót a következőképpen nyerjük: egy kép helyett kettőt készítünk gyors egymásutánban, úgy, hogy az első képen világítanak a LED-ek, de a második kép felvétele előtt kikapcsoljuk őket, akkor a két kép elvileg csak a markerek helyén tér el egymástól. A két képet egymásból kivonva a markerek helye könnyen megállapítható. Ezt az eljárást szemlélteti az 5. ábra.



5. ábra. Világító és kikapcsolt markerek mellett készített két kép között különbség csak a markereket helyén jelenik meg

A mikroszkópon keresztüli kamerakép alapján igen pontos adatokat nyerhetünk ki a robot és különféle egyéb objektumok relatív helyzetéről. Ehhez a feladathoz is használhatunk markereket vagy különféle képértelmezési eljárásokat az adott alkalmazás függvényében.

2.2. A kommunikációs kapcsolat

2.2.1. Az adatátvitel módjának megválasztása

Az Interneten lehetőség van különféle szállítási protokollok segítségével az adatokat különböző megbízhatósággal és sebességgel továbbítani. A kettő rendszerint csak egymás rovására növelhető. Rendszerünk megvalósítása során a lehető legmegbízhatóbb átvitel mellett döntöttünk, a megfelelő sebesség elérése érdekében pedig nagy figyelmet fordítunk az átküldendő adatok megválasztására és jól kézbe tartható tömörítési eljárásokat alkalmazunk.

Az adatátviteli kapcsolat forgalmának túlnyomó részét a telerobot munkaterületéről érkező vizuális visszacsatolás (képek küldése) teszi ki. Az átvitel optimalizálásakor ezért kiemelt figyelmet kell szentelni a hatékony képátvitel megvalósításának.

Adott átviteli kapcsolat tulajdonságok mellett, a képátvitel sebessége a feldolgozási idő és az átvendő adatok mennyiségének csökkentésével javítható. Olyan algoritmusra van tehát szükségünk, amely a kódoláshoz/dekódoláshoz nem igényel túl sok számítást, valamint amelynél a tömörítés mértéke a minőséggel szemben paraméterezhető. Az ingyenes forráskódú JPEG algoritmus mellett döntöttünk, amellyel így akár 10-20:1 tömörítési arányt is el tudunk érni képkockánként.

Az átvendő adatok további csökkentését változó felbontású technikával értük el. A változó felbontású technika azon az elven alapul, hogy egy képnek különböző fontoságú területei vannak, és a kevésbé fontos "periférikus" területeket elég alacsonyabb felbontásban látnunk. A fontosabb részeket az algoritmus részletgazdagabban, nagyobb felbontással, míg a kevésbé fontosakat kisebb felbontással kódolja [9] [10]. Az algoritmus nagy előnye, hogy mivel a

kódolás eredménye több független kép, ez továbbítás előtt JPEG algoritmussal tömöríthető. Az így kapott algoritmus egyesíti a két módszer előnyeit és kivételesen jól konfigurálható és jó hatásfokú tömörítési módszerhez jutunk.

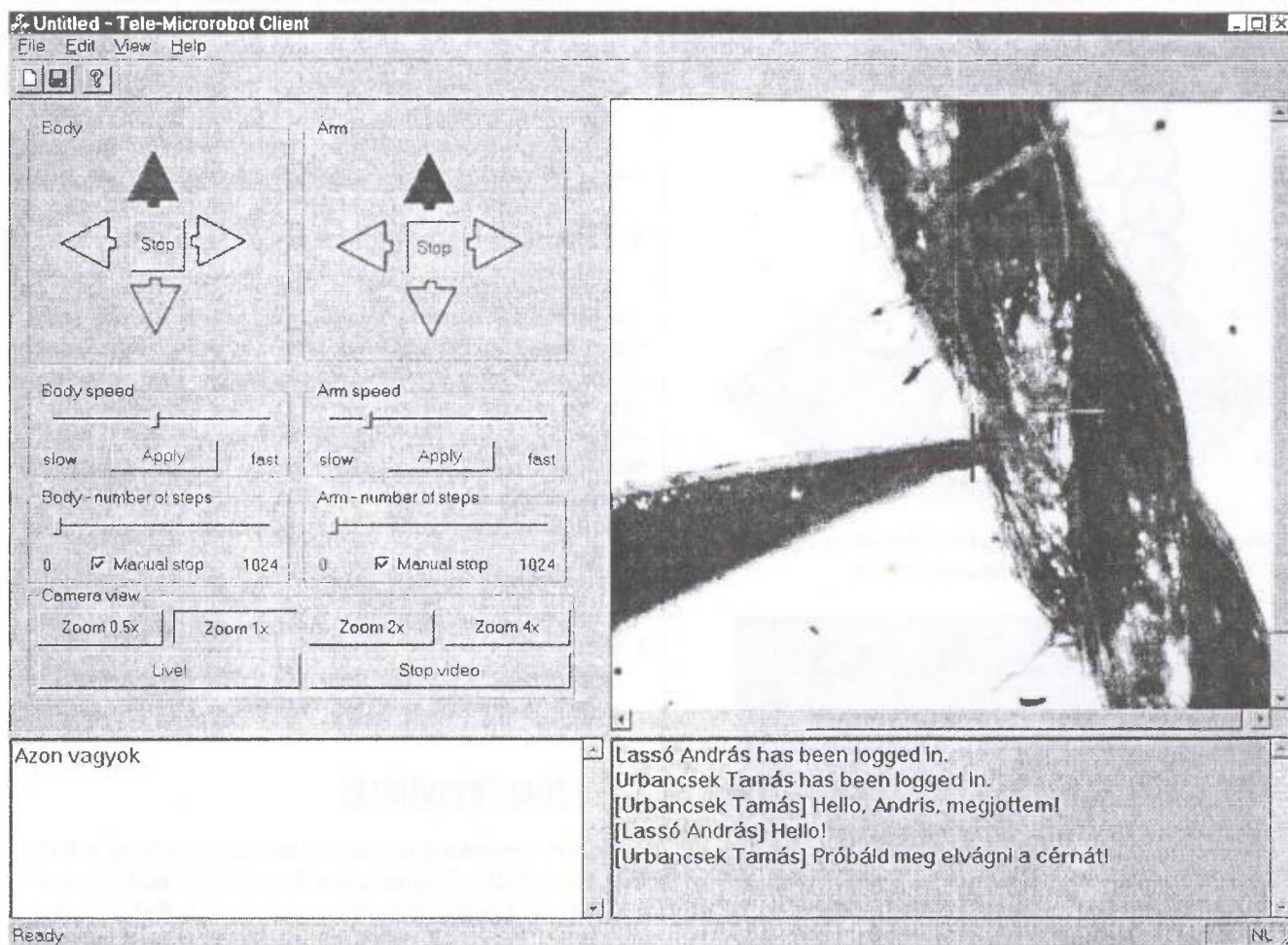
Döntéseinket kísérleti eredményeink is igazolták, ugyanis az MPEG tömörítésű mozgóképek Internet hálózaton elérhető néhány kép/másodperces képfriessítési frekvenciáját el tudtuk érni változó felbontású képek JPEG kódolásának alkalmazásával, jóval kisebb számítási kapacitás és jobb paramétereizhetőség mellett.

2.3. A kezelő-interaktív rendszer

A kezelő-interaktív rendszer hardver követelményeit úgy választottuk meg, hogy a telerobot által nyújtott lehetőségeket maximálisan ki tudjuk használni, de ne szűkítsük

le a felhasználók körét túlzottan magas igények támasztásával. Mivel minden felhasználónak saját kezelőfelületet kell létrehozni, így szoftver megoldásokat kerestünk a speciális hardver eszközök helyett. Egy átlagos IBM PC-kompatibilis számítógép Pentium processzorral Windows NT operációs rendszerrel már képes a szükséges teljesítmény nyújtására.

Visual C++ 6.0-ban megírtuk a telerobotot működtető szoftvert, amellyel a fent ismertetett kísérleti prototípus minden képessége kihasználható. A 6. ábrán látható az általunk készített szoftver felhasználói felülete. A programmal lehetőségünk van a robottestet és a kart a kívánt irányba, kívánt sebességgel és lépésszámmal mozgatni, a kamerakép zoomolására, valamint a többi bejelentkezett felhasználóval szöveges kapcsolattartásra.



6. ábra. A kezelő-interaktív rendszer felhasználói felülete

Külön dialógusdobozban lehetőség van a munkaterületről érkező képek minőségi paramétereinek beállítására. (a változó felbontás tömörítési rátái, az érdeklődés helye, a frissítési frekvencia)

3. GYAKORLATI ALKALMAZÁS

A MINIMAN projekt célja egy elsősorban biotechnológiai kutatásokra alkalmas tele-mikrorobotikai rendszer létrehozása volt. A több lehetséges alkalmazás közül kie-

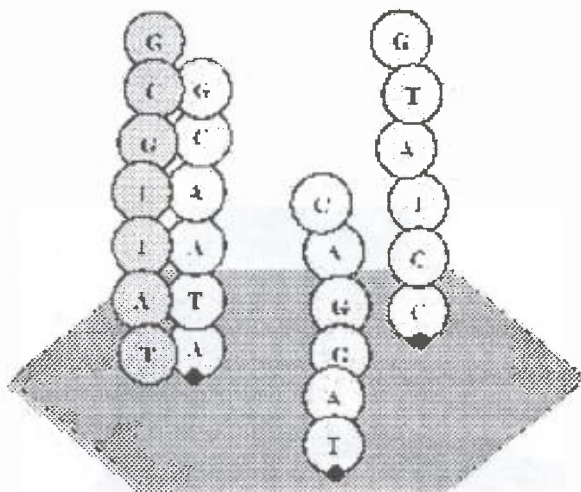
melkedik a SOTE Genetikai, Sejt- és Immunbiológiai Intézetrel tervezett közös kísérletsorozat, melynek kapcsán lehetőség nyílik az előző fejezetben tárgyalt megvalósított rendszerünk kiterjedt tesztelésének.

3.1. A DNS-chip [7]

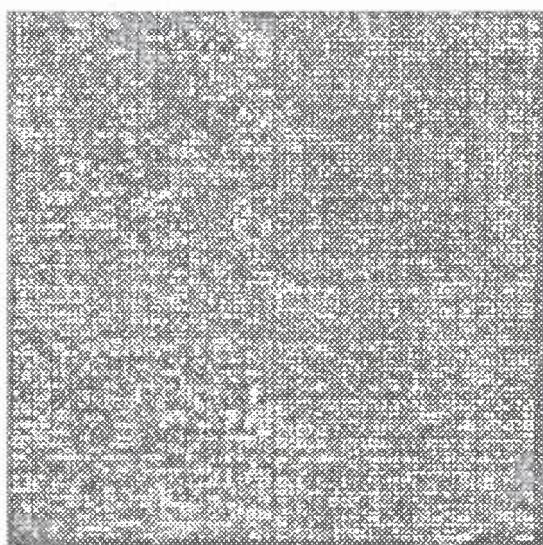
Az elmúlt években elterjedőben lévő új DNS technológia, a DNS (gén)-chip eljárás hatalmas technikai előrelépést jelent a genetikai azonosítás gyors és automatizálható kivitelezésében. Az azonosítási eljárás azon alapul, hogy

egy stabil DNS kettős láncát alkotó szemközti nukleotidok szükségképpen egymás komplementerei.

Az azonosítási eljárás úgy zajlik, hogy rendelkezésünkre áll egy eszköz, melynek parányi felületéhez (membránjához) ismert összetételű, azonos hosszúságú, egyszálú DNS-láncok vannak kötve általunk ismert elrendezésben. Ez a DNS-chip. A vizsgálni kívánt ismeretlen nukleotid sorrendű DNS-darabka két szála melegítés hatására kettéválik. Az egyiket megjelöljük (például fluoreszkáló festékkel), majd egy pontos mikromanipulátorral a membránra helyezünk. Egy mikroszkópra szerelt kamerával figyeljük a folyamatot, és leolvassuk, hogy a DNS-szál milyen koordinátájú másik DNS-hez kötődött hozzá. A reakciók koordinátáinak ismeretében egyértelműen megállapíthatjuk, hogy szerepel-e az analizálandó próba-DNS felépítésében egy bizonyos szekvencia (7. ábra).



7a. ábra. A DNS-chipen rögzített nukleotidláncok egy komplementer DNS-szekvenciát azonosítottak



7b. ábra. DNS chip felületéről fénymikroszkóppal készített felvétel

Ez az eljárás alkalmas többek között fertőző mikroorganizmusok, genetikai rokonságok, ismert genetikai hibák kimutatására és azonosítására. A módszer hatalmas előnye, hogy egyidejűleg óriási számú vizsgálat végezhető el, mindez automatizálható és az értékelés során a számítógép minden előnye (pl. hálózati adatbázisok) kihasználható.

Rendszerünk igen hatékony alkalmazhatóságát a DNS-chip technológiában a következők tények támasztják alá:

- A mikrorobot munkaterének és a robotot kezelő személynek térbeli elválasztása biztosítja mind az ember védelmét, mind a vizsgált DNS minták érintetlenségét.
- Rendszerünkkel lehetőséget tudunk teremteni olyan különleges felszereltséggel ellátott laboratóriumok létrehozására, amelyben a világ bármely pontján tartózkodó kutatók együtt tudnak dolgozni, és melyekben különleges biotechnológiai eszközökhöz széleskörű hozzáférés biztosítható.
- Valós körülmények közötti alkalmazhatóság: a manipulátor pozicionálási pontossága jól megfelel a feladat által támasztott igényeknek.

Összefoglalva elmondhatjuk, hogy a kitűzött céloknak megfelelő rendszert ezzel sikeresen megvalósítottuk.

4. TERVEZETT TOVÁBBFEJLESZTÉSEK

A jelenleg létrehozott telerobotikai rendszerben bár igen sok problémát megoldottunk, egyelőre csak erősen korlátozott keretek közt alkalmas arra a kutatási célra, amelyre készült. Ehhez a jövőben szeretnénk a rendszer összes eltervezett szolgáltatását megvalósítani. Mindenekelőtt a gyakorlati alkalmazhatóságához legszükségesebb teendőkhöz látunk hozzá és a rendszer eddigi felfedezett hiányosságait igyekszünk pótolni.

- Be kívánjuk befejezni a fuzzy szabályozó implementálását.
- Hatékonyabb, virtuális valóság alapú felhasználói felületet szeretnénk biztosítani a kezelő személyek számára.
- Szeretnénk „éles” biotechnológiai kísérletben kipróbálni rendszerünket; ehhez a robot jelenlegi beavatkozásszerve helyett megfelelő speciális tüt kell mikrorobotunkhoz illeszteni.

5. KÖSZÖNETNYILVÁNÍTÁS

Ezúton szeretnénk köszönetet mondani elsőként a *BME Irányítástechnika és Informatika Tanszékének*, német részről a *Karlsruhei Egyetem Folyamatszabályozási és Robotika Intézetének* a MINIMAN mikrorobot kifejlesztéséért, valamint a *Volkswagen Alapítványnak* a kutatás anyagi feltételeinek megteremtéséért.

IRODALOMJEGYZÉK

- [1] P. Arató, L. Vajta, G. Felső: Visual Sensoric for Microrobot Systems, Proc. of World Multiconference on Systemics, Cybernetics and Informatics, 7-11.06.1997, Caracas, Venezuela, Volume 3., pp. 1-5. ISBN 980-07-4175-5
- [2] P. Arató, L. Vajta, G. Felső: Micromanipulation Sensor Structure: a New Approach, Proc. of IFAC-IFIP MCPL 97 Conference on Management and Control of Production and Logistics, 31.08.-03.09.97. Campinas, Brazil, Volume 2., pp. 641-645
- [3] Bao, Y. (1997) Transmission of Real-Time Video Across the Internet using MPEG Encoding, <http://www.eecis.udel.edu/bao>
- [4] Heinz Wörn, Axel Bürkle: Kamera-basierter Sensorsystem eines mobilen Mikroroboters. 15. Fachgespräch Autonome Mobile Systems (AMS'99), München, 1999
- [5] Carter, M. (1996) Why use telerobotics for entertainment, <http://telerobot.mech.uwa.edu.au>.
- [6] Allen Denver (1995) Serial Communications in Win32. MSDN Library, Technical Articles.
- [7] Falus, A., Váradi, A., Raskó, I. (1998) Az orvosi genetikai diagnosztika új eszköze, a DNS-chip. Orvosi hetilap 139 (16), 957-960.
- [8] Friz, H. (1998) Design of an Augmented Reality User Interface for an Internet based Telerobot using multiple Monoscopic Views. Diplomamunka, Institute for Process and Production Control Techniques, Technical University of Clausthal.
- [9] Geisler, W. S. and Perry, J. S. (1998) A real-time foveated multi-resolution system for low-bandwidth video communication. Human Vision and Electronic Imaging, SPIE Proceedings 3299, 294-305.
- [10] Geisler, W. S. and Perry, J. S. (1999) Variable-Resolution Displays for Visual Communication and Simulation. The Society for Information Display, May 1999.
- [11] Institute for Process Control & Robotics – Universität Karlsruhe (IPR-UKA), home page: <http://www.ipr.ira.uka.de/miniman>
- [12] Kleinheinz Gábor: 3 dimenziós mérőrendszer fejlesztése mikromanipulátorhoz. Diplomamunka, Budapest, 1998
- [13] Kortum, P. T. and Geisler, W. S. (1996) Implementation of a foveated image-coding system for bandwidth reduction of video images. Human Vision and Electronic Imaging. SPIE Proceedings, 2657, 350-360.
- [14] Microsoft Knowledge Base (1998) HOWTO: Send Raw Data to a Printer by Using the Win32 API. MSDN Library, Article ID: Q138594
- [15] NASA homepage. <http://www.nasa.gov>
- [16] Saffiotti A.: The uses of fuzzy-logic in autonomous robot navigation. Soft Computing 1 (Springer-Verlag) 180-197, 1997
- [17] Jörg Seyfried: Die kamerabasierte Steuerung eines piezoelektrischen Mikroroboters in einer universellen Mikromanipulationszelle. Diplomarbeit, Karlsruhe, 1996
- [18] Sheridan, T. B. (1992) Telerobotics, automation and human supervisory control. Cambridge, MA: MIT Press.
- [19] T. Tanaka, J. Ohwi, L. V. Livintseva, K. Yamafuji, S. V. Ulyanov: Soft computing algorithms for intelligent control of a mobile robot for service use. Soft computing 1, 99-106, (Springer-Verlag), 1997
- [20] Vajda Ferenc, Kleinheinz Gábor: Flexibilis mikrorobot fejlesztése miniaturizált ipari technológiákhoz. TDK-dolgozat, Budapest, 1996
- [21] Vajda Ferenc: Pozícionáló szenzorrendszer mikrorobotos munkáállomáshoz. Diplomamunka, Budapest, 1998

APPLICATION OF A TELEOPERATED MICROROBOT FOR BIOTECHNOLOGICAL RESEARCH

T. URBANCSEK, STUDENT and A. LASSÓ, STUDENT

TECHNICAL UNIVERSITY OF BUDAPEST
H-1111 BUDAPEST, PÁZMÁNY PÉTER SÉTÁNY 1/D

We provide an overview of the design, implementation and a real-world application of our tele-microbotic system that incorporates the newest results of three emerging technologies: teleoperation, microrobotics and biotechnology.

Reasons for the application of machines instead of human work for a particular task is that: automated equipments can work more effectively, or humans simply can not accomplish the task (e.g. because of the dangerous environment, small sizes or great distances). However if we can not construct an autonomous system due to the complexity of the task or the system can not react properly to the unforeseen events we have to provide the possibility of human interference. These problems are studied by the researchers of teleoperation.

Our teleoperation system is built for accomplishing specific tasks in the field of biotechnology in a safe, economical and flexible way, by using a microrobot. Safety is guaranteed by the remote controlling and the application of intelligent devices (which behave properly in extraordinary situations). The system makes possible to observe the workspace and control the robot from a remote place via an Internet connection. We implemented all software elements for the operation and connection of special hardware elements and performing network communication.



Urbancsek Tamás a Budapesti Műszaki Egyetem negyedéves villamosmérnök hallgatója az irányítástechnikai és robotinformatikai szakirányon. Két éve tagja az Irányítástechnika és Informatika Tanszék Mobil- és Mikrorobotika Csoportjának, fő kutatási területe a MINIMAN mikrorobot irányítása.



kalkuláció, telerobotika és lehetséges alkalmazásainak területét.

Lassó András a Budapesti Műszaki Egyetem ötödéves villamosmérnök hallgatója az irányítástechnikai és robotinformatikai szakirányon. Képfeldolgozással, szabályozástechnikával és hálózati alkalmazások készítésével több éve foglalkozik az Irányítástechnika és Informatika valamint az Automatizálási és Alkalmazott Informatikai Tanszékeken. Másfél éve tanulmányozza a teleoperáció, telerobotika és lehetséges alkalmazásainak területét.

DIGITAL RADIO – AN INTERNATIONAL BROADCASTER'S PERSPECTIVE

MIKE HATE

BBC WORLD SERVICE

The views expressed here are solely those of the BBC World Service. They must not be regarded as the views or policy of any other organisation or consortium of which the BBC is a member. The author would like to express his thanks to the BBC and to all those who will recognise in the text, the results of conversations and discussions held with them.

Wherever the term BBC is used without any qualification it should be taken to mean the BBC World Service. The World Service is a division of the wider BBC, a large organisation whose interests include home service broadcasting throughout the United Kingdom – on television and radio – and international television. The primary concern of the World Service is international radio broadcasting. Where I have referred to other parts of the organisation I hope I have made this clear.

This paper was originally presented at the Regional Symposium on New Broadcasting Technologies (Sound and Television) organised by the Freidrtich Ebert Stiftung and the ITU in Santiago de Chile last August. I am sure that some readers will recognise it as being plagiarised from one that I presented earlier at the Broadcast Engineering Society of India's conference in Delhi. I have made a few minor changes to reflect subsequent developments.

1. SYNOPSIS

The BBC World Service currently serves a global audience of around 140 million. Traditionally, this audience has been served using skywave, HF and MF, transmissions originating from the BBC's own relay stations. It is apparent that the audience on short wave is declining; it seems that listeners are turning to alternative services provided on locally broadcast FM, Internet, TV, etc. The BBC itself attempts to provide a service of its most important programmes, to key audiences, on local FM transmitters wherever it can. At the moment about a quarter of its measured audience listens through this medium. At the same time developments in technology are opening up other possibilities for serving existing and new audiences. Among the more interesting to the World Service are potential developments in digital broadcasting. Ideally the BBC would benefit from a digital format that combined the coverage advantages of sky wave propagation with the quality and convenience of FM. Is this a realistic aspiration? At the same time the BBC recognises that the success of a new format will depend on the ability to populate the market with receivers capable of receiving the programmes. This in turn relies on the format being attractive to all of the broadcasters in a region, importantly the local and national ones. Speculating on what such a new digital format should offer an informal specification is suggested.

2. FOREWORD

When I was a newly qualified graduate my supervisor said to me, "remember, invention is the mother of necessity – you will get further faster by looking for problems to fit with known solutions than you will doing it the other way round". The broadcasting industry is a good example of this principle at work. Broadly, the invention of radio and television was not in response to a demand; demand followed from the availability of a service which was provided as much because it was possible to do it as for any other reason.

I am not aware of much listener demand for the BBC World Service to start digital radio services. It already uses digital techniques extensively in distributing and making its programmes but this is mainly because these methods offer economic benefits. If the BBC is to persuade its audiences to invest their hard-earned money in the equipment necessary to receive digital broadcasts some real benefit must be offered to those audiences. Equally, if the BBC is going to invest its (hard-won) money in digital transmission equipment it must be confident that it will play a part in securing those audiences.

3. INTRODUCTION

Broadcasting is a technology based industry and as such will continue to undergo rapid change as the relevant technologies develop. Among current technological trends there are two which seem to be of particular relevance to the broadcasting industry. Commercial pressures in the Information Technology business are bringing about very rapid advances in the available hardware; it is now easier, faster and cheaper to carry out by calculation, processes that were not so long ago the exclusive domain of analogue electronics. At the same time, TV and radio sets, recording media, PCs and the means of feeding them are converging into a seamless continuum; will the user know, or care, if he is playing a computer game or watching a TV programme. This is forcing the BBC World Service, along with most other broadcasters, to examine its markets, its audiences and the means for reaching them, in a way that it has not had to do before. Also, there is a movement around the world toward de-regulation of the airwaves so that existing broadcasters are finding new competitors offering new services in their traditional market places.

The following discussion will look at the way a broadcaster like the BBC World Service might decide its position in the market place and how best to use the available technologies to serve its audiences. As a starting point it

is assumed that the main 'product' of the BBC World Service is sound-radio programmes delivered straight to the listener. Having concluded that most of these listeners are either moving around while listening, or doing something else at the same time, or both, it goes on to see how these people are served now, why the methods used are successful and what the trends are. On this basis it goes on to speculate on the contribution digital broadcasting formats might make. Finally, and in a fairly qualitative way, a broad 'specification' for a digital broadcasting system is proposed, based on criteria evolved from the earlier discussion.

4. THE OBJECTIVES OF THE BBC WORLD SERVICE

4.1. *What do we mean by Radio?*

Today's analogue radio services try to reach audiences ranging from those listening on small portable radios, through car travellers to those listening on fixed High Fidelity receivers in their homes. Since the vast majority of the 'fixed location' audience currently chooses to watch television rather than listen to the radio, it is time for radio to accept that this audience might not be a prime target for its unique product. At least it might consider that there are better ways to serve it. For example, if one wants to listen to a classical music concert or an opera in high quality, the size of the currently available apparatus needed to reproduce the sound prevents it from being anything other than fixed. Such a service could just as well be delivered as part of a digital television multiplex through a fixed antenna, possibly a satellite dish, and who knows, there might even be pictures as well.

Radio must position itself as a medium which targets an audience listening on a portable receiver, possibly a 'walkman', or in a car. I use the radio to provide a continuous source of news (or any other programme of interest) while moving around the house; typically while getting ready to go to work in the morning. While choosing not to carry a receiver around — I prefer to have a set switched on in every room — the application is still more mobile than fixed.

Sound radio broadcasting should also recognise that as well as being able to move around with the listener its other unique feature is its availability while the listener is doing something else. In this respect, if it is to be anything more than a source of musical wallpaper, it has few competitors other than other radio stations. Even newspapers and magazines require their readers undivided attention and television and the Internet (certainly using today's technology) target audiences in locations which are more or less fixed. This does not mean that the radio broadcaster should not address the 'fixed' audience if it is seen as important, more that the way to do it might not be the same as that for a car driver.

4.2. *Who Are Our Audiences and How Do We Serve Them Now?*

One way to divide the community of broadcasters is on the basis of the geographical areas they are trying to cover. Three of these will typically be of, and originate material from, within their service area:

- Local Broadcasters try to reach an audience confined to a particular town, city or perhaps a rural area with a specific identity. Their main focus is matters of interest to the local community.
- Regional Broadcasters are looking to reach a larger audience in a wider geographical location. A number of countries naturally divide into regions for various reasons and there is a market for broadcasting which targets particular interests within them.
- National Broadcasters target whole nations.

The fourth:

- International Broadcasters, usually generate their programme material outside their target area but might want to address any or all of the groups listed above as well as, for reasons of economy, whole continents or areas which share a common language, hopefully, with a single transmission.

BBC radio falls into all of these categories. The BBC Home Services seek to be all of the first three while the BBC World Service is definitely the fourth.

Traditionally, the World Service has tried to serve its audiences using terrestrial, sky wave (MF or usually HF) transmissions. This is not because of any emotional attachment to HF broadcasting or anything to do with the modulation formats involved, it is simply that sky wave propagation allows very large areas to be covered from transmitter sites outside the target area¹. Clearly having the transmitter sites remote from the target means that the BBC can isolate itself from the political climate and, significantly, can cover the whole of the world from a relatively small number of locations. There are a number of reasons why sky wave broadcasting has not been the propagation medium of choice for the BBC's own domestic services: the geographical size of the United Kingdom and the need to maintain regional identities figure alongside the technical limitations and tuning difficulties. However, in very large countries with uneven population densities, sky wave could be, and sometimes is, an attractive option for national broadcasters as well.

It is becoming clear that while sky wave broadcasting might look good in terms of coverage maps and population counts within the target areas, there is no guarantee that anybody will be listening. In some instances the BBC's own research suggests that very few people are listening and HF broadcasting is an expensive business in terms of investment and running costs. Much as with a commercial broadcaster, the size of the audience influences the BBC's access to funding. This leaves the question of what new strategy should be adopted to make a service widely available in a format that people will listen to. In practice this usually means looking at the listening habits in the area concerned and at the opportunities that are available.

¹ The BBC World Service does not see itself as an 'HF broadcaster'. It tries to reach its audiences by whatever means it has at its disposal at any particular time and using the techniques which it judges to be most appropriate. Currently, HF, despite its obvious shortcomings is still the most tractable way for the World Service to address the majority of its listeners.

As things stand today, the range of options available to international broadcasters to counteract the decline in HF listenership is small, typically one, the exclusive or shared use of VHF band II (FM) channels in certain population centres. While this usually offers a readily available, good quality signal in the local area it is not without a degree of risk. The BBC must rely heavily on the goodwill of its partners, the local regulatory authorities and the political regime. Also local broadcasting brings a whole host of problems with maintenance contracts, deployment of maintenance personnel, spare parts, delivery of a suitable stream of programmes and a means of getting accurate data on continuing operation and performance. Available transmission sites do not always offer ideal coverage and agreements with partners might preclude the BBC from being able to transmit its most important programmes because air time during peak listening periods is at a premium. Having said all this, the BBC now has around a thousand such arrangements. The vast majority have proved to be very successful and the actual number of listeners tends to increase dramatically when services are available on 'local air'. Around twenty five per cent of the BBC's global audience now listens on FM.

FM services are only able to cover relatively small areas. They are best used in areas of high population density but this still presents the BBC with the problem of covering the less densely populated areas as well. Unfortunately, the only way to do this, and mitigate all of the other risks, is to retain the HF services. In the densely populated areas, where the BBC is effectively competing with itself, the HF audience drops even further because people obviously choose the more accessible medium. Expensive HF transmitters have to be kept in service even though the measured audiences for them can be quite small.

4.3. Other Possibilities

The discussion so far has concentrated more on propagation than the actual modulation format. Digital broadcasting has barely been mentioned at all. It is clear that neither of the two predominant methods of reaching the audience is entirely satisfactory. Sky wave services are not managing to retain their audiences and FM services suffer from a whole host of difficulties associated with operating from remote locations. Before going on to see how digital formats could be exploited by the international broadcaster to address these problems, it is necessary to mention at least two other possibilities.

An obvious candidate to replace sky wave transmission is direct satellite broadcasting. Like terrestrial sky wave this can be used to cover vast geographical areas but it can also make more selective use of the available transmission power. Unfortunately, the signal strengths available from satellites in geo-stationary and highly elliptical orbits (GEOs and HEOs) are so small that reception appears only to be a realistic proposition where there is a line of sight to the satellite. Even then the receiving antennas have to be a little more elaborate than the standard whip on the typical portable or car receiver. Car reception is likely to be difficult in urban areas without the use of terrestrial repeaters, which might themselves be FM transmitters, and anyway will suffer from all of the same logistic

difficulties. Satellites in low earth orbit (LEOs) could offer greater signal strengths but having several in the same orbit, a problem shared with HEOs, would tend to make them rather expensive. Further to this, broadcasters rarely own their own satellites, the BBC certainly doesn't, and so they have to rely on third party operators to carry their programmes. While the broadcasters should ensure that they are protected contractually in such arrangements, it is very difficult to think of every eventuality. How can services be protected if the ownership of the facility transfers to another organisation who is openly hostile to the broadcaster concerned, is a competitor or merely sees different commercial opportunities?

Having painted a somewhat bleak picture of the prospects for Satellite Radio Broadcasting it must be stressed that this does not mean the BBC is dismissive of the technology. At the present time the BBC is watching developments with considerable interest and an open mind, conscious of the fact that in spite of the apparent (and largely speculative) difficulties this could become a very important way of reaching particular target audiences. The World Service is currently in discussion with WorldSpace and had a pilot service available on the AfriStar satellite.

The BBC, certainly in the UK, has already established itself as an Internet / Webcaster news service. As well as comprehensive text and images, 24 hour live audio services are available on the BBC World Service website in English, Arabic and Chinese, with Spanish and Russian to follow late in 2000. A number of other languages are also available in audio at various times of day.

Thousands of radio stations are connected to the Internet in some way or another and this number is growing all the time. At the last estimate by the EBU, nearly two fifths of these stations used the world wide web for delivering audio services. This is an exciting development and improvements in technology are allowing faster access by the user, including video as well as files². While the restricted capacity of the currently available circuits into peoples homes can still limit the audio quality and server speed the time of access, broadband and higher modem speeds continue to improve the user experience. Furthermore, while the medium potentially suffers from the 'gatekeeper' problems associated with FM and satellites, this in itself is a double edged sword as it is through the Internet hosts, such as Broadcast.com, that the radio services can raise their profile in new markets. The technologies needed to access the programmes are quite expensive, by radio standards, and current access arrangements mean that the listener has to pay for the connexion even if the raw programme is, as most radio programmes are, 'free to air'.

A lot of this might well change. The cost of connexions is falling and the capacity will increase to the point where it will be, for all practical purposes, infinite; certainly greater than the demands that are ever likely to be placed upon it by the average household.

Unfortunately, I can not claim to be an expert on the subject of Internet broadcasting or Webcasting. It is also

² A recent survey reported that 35 % of Internet users in the US had accessed Internet audio in the previous month (Jupiter/NFO Consumer Survey, Spring 1999).

not part of the main thrust of this paper and really justifies a separate treatment of its own. For this reason I do not intend to elaborate further on the subject here. However, the BBC does regard this as an important development, which is having a major impact on the whole of the broadcasting industry.

5. THE DIGITAL OPPORTUNITY

5.1. *What It Might Offer*

Shannon's theorem tells us that the information carrying capacity of any communication channel is ultimately limited by the bandwidth of the channel and the ratio of the potential received signal power to the noise power in the channel. With analogue formats, the scope for exploiting the implicit trade offs here is limited. Digital formats allow much greater control and this is one of the fundamental reasons for using them. Provided that the channel has sufficient capacity to carry it, the received quality of a digital signal can be made to be almost entirely defined by the quality of the conversion processes and not by the vagaries of the channel.

Another great advantage (possibly the greatest) of using digital formats is that they allow the use of sophisticated compression techniques to reduce the rate of information flow in the signal with a controlled, if not always small, impairment to the perceived output quality. This allows digital services to be made much more efficient in their use of RF spectrum (or cable capacity) than their analogue equivalents. They ought to be able to put apparently higher quality services into the same (spectral) space or programmes of the same quality into a smaller space³. As well as this, techniques developed to counteract the damaging effects of multipath propagation can be adapted to allow the use of synchronous networks; the same RF signal can be transmitted from two geographically separate sites, each covering the same or overlapping areas. The same frequency can be used to cover a bigger area without having to worry about the effects of co-channel interference; spatial diversity of the transmitting sites might be used as a defence against the uncertainties of propagation. It is the use of single frequency, synchronous networks that allows wide area DAB (Eureka 147) networks to achieve spectral efficiency.

While not uniquely their domain, digital formats usually offer the facility to send auxiliary data streams alongside the audio (or video). The broadcaster can choose to enhance the service with additional background information, like biographical details of a musician or a still picture, or even provide separate parallel services such as stock market reports. While able to carry all sorts of information, which may or may not be related to the programme itself, the most useful application for this data, or at least some of it, is assisted or automatic tuning.

³ The success of cable and satellite television has been founded in the availability of a wider choice of services offered by the new formats.

5.2. *What We Might Want From It*

Clearly, for a whole host of coverage, administrative, political, and economic reasons, sky wave broadcasting in the HF, MF and possibly LF bands is an attractive proposition for a broadcaster like the BBC World Service. It is also clear that the current audience for these services is falling as they turn to FM, which is itself frequently offered by the BBC in competition with itself. Can digital modulation techniques provide a service in the sky wave bands that will offer the listener those features of FM broadcasting which make it the medium of choice? What is it about FM that makes it so popular?

Ignoring the editorial quality of the programmes, it is likely that the popularity of FM stems from a combination of three factors:

- Potentially very high sound quality. Stereo is often heralded as a selling point for FM stations but it is not clear how much of the audience genuinely benefits.
- Easy accessibility. Most people can tune an FM radio to find the station they want and once they have found it, it remains relatively stable.
- Cheap and readily available receivers. While the best FM tuners (plus amplifiers, loudspeakers and other paraphernalia needed to go with them) are relatively expensive it is possible to buy modest receivers practically anywhere in the world.

It is difficult to say which of these factors, if any, is the most important, I can only guess, and it is likely that different markets would rank them differently anyway. The digital system is probably going to have to address all three. Not only must it address them, it must address them in a way that will convince the listener that the benefits are sufficient to warrant buying a new receiver.

The success of the Digital Radio Broadcasting system is eventually going to be measured by the number of listeners. The BBC's international radio audience is currently around 140 million from an estimated world wide HF receiver population of around ten times that figure. The approximate life of a current generation HF transmitter is around 25 years and it would be nice to think that the new format would be universally accepted in a comparable time frame. This translates to selling 56 million new receivers to new customers every year for the next 25 years.

5.3. *Specification For a Sky Wave Digital Broadcasting System*

For reasons already explained, the BBC is looking for a digital broadcasting system to operate with sky wave propagation; this probably means in the HF, MF and possibly the LF bands. The following informal specification is offered with that in mind. In spite of this there is nothing to say that such a system, perhaps with minor modifications, could not be used in other bands; indeed, its introduction as a commercial reality probably depends on it. Elements of the specification are discussed as they are presented.

5.3.1. Basic Requirements

Taking a lead from the speculative list of benefits associated with FM in section 3.2, the take up of the digital radio system will probably depend on three critical factors:

- Assisted tuning: A real criticism that can be levelled certainly at present day HF broadcasting is the need to be something of an expert to be able to find the required programme in the first place and have the resolve to stay with it as transmission frequencies are changed. A parallel situation exists with MF and FM in a car when moving from the service area of one transmitter to another.
- Improved audio quality: Current trends suggest that however good the programme content, if there is no special incentive to seek information from an obscure source, the digital service must compete in an environment where FM broadcasting has become the norm. The digital system should seek to approach the perceived quality of FM.
- Entry level receiver: While there are, potentially, a number of advanced facilities that a digital radio system might be able to offer, their incorporation must not preclude the development of a simple and inexpensive receiver. It might not be able to make full use of any advanced facilities but should offer assisted tuning and improved audio quality.

Having said earlier that the quality of a digital transmission system ought to be dependent on the quality of the conversion processes rather than the quality of the channel, sky wave channels are notoriously hostile. The digital modems should take account of the possible degradations that the channel might impose and ultimately, the automatic tuning arrangements ought to be able to seek an alternative source of the programme if a better one is available.

Further to the factors mentioned already, the broadcasting bands are crowded and developments in the global media will inevitably result in pressure for more and more services to be accommodated. This means that:

- The RF band occupancy of the broadcast channel must be kept as low as possible commensurate with being able to meet the critical factors listed above.

These four factors, taken together, define the basic requirements of the digital transmission system. They are covered in greater detail along with the supporting arguments in the next four sections.

Assisted Tuning

From the point of view of finding and staying with the required station, the digital receiver will be more like a television set (or some FM tuners), with push buttons allocated to the programme streams. As with the current generation of television receivers, the fact that a particular station is being received from UHF bands 4 or 5, or even from satellite on Ku band or C band, should be quite transparent to the user. Ideally, having once told the digital receiver which programme stream is associated with a particular button (or some other selection device) the receiver will stay with that selection, finding the best

available signal, whichever band it happens to be in, and stay with it as frequencies, and even bands, are changed.

To make proper use of such a facility the broadcasters may well have to rethink the way they package their programme streams, at least from the point of view of the way they are publicised. It will be necessary to define an easy to understand package of unique streams which the listener, and the receiver itself, can readily identify. Since assisted tuning is seen as fundamental to the success of the venture it is important that it is designed into the concept at an early stage.

At least some of the information necessary to achieve assisted or automatic tuning will have to be carried on the digital multiplex in such a way that it is extremely robust and available quickly. It is not acceptable for the receiver either to get this wrong, even in difficult reception conditions, or to leave the user for more than a second or so without some acknowledgement of a request to change stream. One needs to imagine, for instance, the situation where a listener is scanning the band to see what is available; how long should the receiver have to sit on a particular frequency to see if there is anything of interest there or not? It is suggested that the potential time taken for the receiver to lock to a compressed and possibly time interleaved audio signal will be far too great.

Audio Quality

Given that there are probably better ways of delivering very high quality audio to fixed installations it is assumed that the digital transmission system will not be High Fidelity. Since the target audience is likely to be listening on a portable receiver or in a car, the highest quality that the system needs to be able to deliver is the highest quality that is achievable on a portable receiver or a car radio. Further to this, even the quietest car is not a high quality listening environment so it is likely that the most stringent requirement will be the quality expected from a good quality portable receiver, probably a stereo machine of the 'big box' variety or a 'Walkman'.

As much as possible of the available capacity should be used for carrying audio when high quality is needed, leaving additional data to be carried only at times when the audio requirement is less stringent. While possible to conceive of a system allocating bits between the audio and the data stream dynamically on a near instantaneous basis, it is probably easier to define quality grades for sections of the audio programme and allocate bits according to the defined grade. The following are suggested as suitable grades. I have used analogue criteria because I understand them. I know that this is far from ideal in the digital world and the criteria will almost certainly have to be adapted to make sense:

- Grade 1: High quality programme: music or drama. Audio bandwidth 12 kHz — stereo⁴.
- Grade 2: Medium quality programme: studio generated speech. Audio bandwidth 7 kHz — mono.
- Grade 3: Low quality programme: reports from jour-

⁴ Fairly obviously, when grade 1 programmes are being transmitted only the very highest priority data will be able to find a place in the multiplex. Other data will have to take its turn when lower grades are being transmitted.

nalists in the field delivered by telephone, sat phone etc. Audio bandwidth 3.5 kHz — mono.

- Grade 4: No audio: an opportunity to download data either as a separate exercise in its own right or as an adjunct to a radio programme⁵.

The question of dynamic range is more difficult but the guiding principle should still be 'as good as than that which can be reasonably achieved by a good quality portable receiver in a realistic listening environment'. Since the car radio audience as an important target, it must also be borne in mind that even grade 1 programmes will almost certainly be compressed.

Entry Level Receiver

A significant proportion of the BBC World Service audience is in the developing world. It is usually assumed that many of these people have a relatively low disposable income and that their propensity for spending any of it on a (new) radio receiver is likely to be limited. In the medium to long term the requirement must therefore be for a low cost receiver. While such a receiver must offer assisted tuning and audio quality approaching that of 'in car' FM, the entry level receiver would almost certainly not have to accommodate pictures and some of the other more esoteric possibilities the system might be able to offer.

Nearly all of the processing necessary in a digital receiver will be carried out in one or a few digital devices. To reduce its cost as much as possible most of the receiver should be accommodated in this device or devices leaving only those essential elements of RF filtering and audio output to analogue electronics. The economics of the silicon chip industry suggest that circuit complexity is only one of many factors affecting the eventual price of the device. Yield and production volumes (and sales) probably have a greater impact. Inexpensive receivers are only likely to emerge if it is possible to make and sell them in large numbers. This in turn will depend as much on the system being able to meet a set of commercial objectives as it is on technical criteria alone.

It is also important that there should not be too much diversification of standards around the world. Increasing the number of standards is likely to increase the number of different receivers which will have to be developed for different markets. This can only serve to push the cost up. While it must be expected that there will be some variation in standards (see 3.3.2 below), it is important that the same fundamental building blocks are used; the same audio compression coding system for instance. This will allow receivers to be constructed from a selection of standard chips or a single chip with all of the standard blocks built into it.

From a technical standpoint, it must also be considered that undue complexity might result in excessive power consumption, an important factor where portable receivers are likely to predominate.

⁵ An example of this might be an educational programme where it is appropriate to download text (lecture notes), numeric information (mathematical formulae) or even software for assimilation by a PC.

RF Bandwidth

Recent developments in the field of perception coding of audio signals suggest that very high compression ratios can be achieved while retaining a reasonable degree of fidelity in the resulting audio. Current indications are that to achieve audio quality comparable with stereo FM in a car, the bitrate of the compressed signal will be of the order of 40 to 50 kbits per second. This translates to a modem capability of 4 to 5 bits/Hz for a 10 kHz wide RF channel and 2 to 2.5 bits/Hz for a 20 kHz channel. Current indications also suggest that these target figures for modem performance are possible for ground wave and sky wave respectively.

It must be concluded that while it might be possible to accommodate a ground wave service within the existing 10 kHz (approximately) allocation for a single channel, a sky wave service will ultimately need a larger (at least 20 kHz) allocation. Given that there will be a bitrate overhead for tuning data etc. even these bandwidths are not generous. They do, however, compare very well with 200 kHz (nominally) for an FM signal and 250–300 kHz (per channel — actually 1.5 MHz for a five or six channel multiplex) for Eureka 147 DAB⁶.

From the above figures it seems plausible to start ground wave services, possibly MF, within the existing channel allocations. While difficult under the current regulatory arrangements, sky wave services which improved on current HF standards could also be introduced and better quality achieved by concatenating existing channels. To sell receivers in sufficient numbers the new service must not be perceived as being 'not as good as we have now' (on FM?).

5.3.2. Operational Modes

The HF bands are the only area where the whole world now uses exactly the same modulation format for broadcasting. The other broadcasting bands (radio and television) all accommodate formats which can vary from one part of the world to another and, sometimes, even within the same country. Given that a digital modulation regime can be optimised in different ways to work with the different conditions that are likely to be found on different propagation paths, it seems sensible to capitalise on this. I have already suggested that the channel characteristics of an MF ground wave path will permit a higher bit rate within a given RF bandwidth than an HF sky wave path.

Clearly it would be a considerable benefit for any receiver to work with any of the different formats. I do not know how the market would react to having to buy a different receiver for each different modulation format but I know how I would react as a customer. It is likely that any general purpose receiver will be designed to work with all of the digital formats and probably the analogue ones as well. Even though my colleagues at the BBC Research and Development department tell me that current chip technology will allow several different decoding arrangements to be incorporated on a single

⁶ Through repeated use of the same channel, synchronous operation, Eureka 147 DAB is itself an attempt to address the somewhat wasteful spectrum allocation currently needed for FM in Band II.

chip, possibly all working at the same time, common sense, and these same colleagues, suggest that the optimum strategy would be a unified system with the ability to accommodate different modes of operation. The different modes would cover different RF bandwidths, different bit rates and possibly (different) ways of allowing simultaneous transmission of analogue services to maintain compatibility with existing receivers.

The mode structure might also be sub-divided on a performance basis. I have tried already to define the audio signal in terms of a set of performance specifications; specifically, a number of grades of performance. Clearly the less stringent performance levels are likely to require a smaller bandwidth allocation or be more robust in difficult propagation conditions. Without compromising the ability of the system specification to meet defined performance targets, it may be that some broadcasters might only need the lesser grades of performance.

5.3.3. Regulatory Considerations

The primary objective of the regulatory environment in which any broadcasting system exists should be to support the interests of the audience and therefore, indirectly, the broadcasters. The converse is only true in so far as the 'anarchic' consequences of not adhering to the regulatory environment would be intolerable for all. The existing regulatory environment in the LF, MF, HF and VHF (FM band II) bands has evolved to support analogue radio broadcasting formats. These formats are now more than 50 (FM) and 80 (AM) years old and it is clear that new, digital, systems can offer improved audio quality along with more effective use of frequency spectrum and transmitter power. Unfortunately, in view of the significant differences between the analogue and digital radio technologies it is unlikely that the regulatory environments, which have evolved to support analogue, will also support digital in a way that best serves the needs of the listener. While possible to engineer a digital service into a 10 kHz HF channel it is unlikely that this would offer the opportunity properly to exploit the possibilities of a new digital modulation scheme.

Further to this the regulatory environment in different parts of the world is different now and will probably continue to be so. This is to support broadcasting formats which have been conceived with different commercial objectives in mind. If an arbitrarily imposed, or existing, regulatory regime prevents the commercial imperatives from being met there is a strong possibility that the whole venture will fail. A carefully conceived mode structure may well allow the digital service to work within different regulatory arrangements.

The existing regulatory environment serves analogue radio and its listeners reasonably well. Since the adoption of digital radio is likely to be a fairly lengthy process and analogue formats will continue to be the most important ones, in audience terms, for a considerable time, it is imperative that the needs of the existing analogue services are catered for throughout any transition period. Digital formats will necessarily have to be introduced into the existing analogue regulatory regime. It is to be hoped that

this restriction will not compromise the performance of the digital systems to such an extent that it seriously inhibits their take up in the market place and, in the fullness of time, that the regulatory environment can evolve to exploit the full potential of the new systems.

5.3.4. Transition Strategies

If a new digital broadcasting standard is to be introduced, perhaps the most difficult challenge to be faced is that of migration from the old standard to the new one. A fundamental problem is the need for each existing 'station' to be broadcast twice during the transition, once using the new digital system and again using its existing analogue allocation, at least until the bulk of its audience is equipped with a digital receiver. It would be commercial suicide to do otherwise.

One possibility is to attempt to accommodate both formats within the existing channel allocations. This might be called simulcasting but is better described by the term 'in band on channel' or 'IBOC'. While attractive this has drawbacks. The quality of the digital system is likely to be compromised through being unable to make full use of the available channel and in most parts of the world the currently available bandwidths would render both analogue and digital services degraded. This situation would clearly correct itself over time as the analogue services were discontinued, but it is important to offer the best available quality during the early years of the standard when the market is being developed. Further, it does not provide any obvious way for making more efficient use of the spectrum or introducing a wider choice of stations.

An alternative would be to allocate channels progressively for digital only use. Again this is likely to be very difficult because the spectrum is already crowded and finding new channels could be problematic. A large receiver base would have to exist before existing stations could give up their analogue allocations.

The digital system should be designed with this in mind. Automatic tuning would make the frequency and even the band being used completely transparent (irrelevant) to the listener as receivers should be able to find the required station regardless of the frequency. More use of spectrum efficient technologies in VHF bands II and III and 'L' band might create space to allow stations which have traditionally been 'AM' to relocate a digital service and the use of ground wave digital services in the HF band might be explored⁷.

The change is likely to be long and problematic. Any system will need to be able to work with both the current regulatory regime, any future one which might be designed to accommodate digital services properly and any intermediate situation. As well as this there will almost certainly be development in the available technology during the transition phase. This means that any standard will have to allow for a deal of flexibility, being able to optimise performance in the future while

⁷ The 26 MHz band is not widely used by international broadcasters; certainly not during periods of low sun spot activity. The possibility of using frequencies in this band for quasi VHF local services, possibly with concatenated channels has been suggested.

not being compromised in a way that would impede its commercial success in the short term.

6. CONCLUSION

Despite the fact that the BBC's HF audience is decreasing, the propagation characteristics of sky wave transmission still make it an attractive proposition. Local FM broadcasting, which is currently very successful in attracting new audiences, and hijacking existing ones, is accompanied by all manner of technical and administrative problems. While it is too early to make an informed judgement, direct satellite broadcasting and cabled services (the Internet) look as if they will serve new audiences and not take over the role of sound radio.

Perhaps the way forward for the international broadcaster is through digital formats. Digital formats, while retaining the propagation and coverage benefits of wide area HF, MF and LF propagation, could also offer those advantages of FM broadcasting (assuming you know what they are) which make it the medium of choice.

In order to attract listeners in the numbers that will make the BBC feel confident about the viability of the digital system, it will be necessary to generate a substantial market for receivers. This commercial imperative is a more important than any technical factors; the specification I have offered here is formulated as much with this in mind as any of the performance parameters the BBC would like to have associated with its programmes. Before anyone is going to buy a new receiver there must be a range of programmes available that can be accessed and 'show case' the capabilities of the new format.

It is therefore essential that the eventual specification for the system takes into account the needs of all broadcasters, local national and international, as many as possible, as well as the receiver manufacturers who will have the problem of actually doing the selling.

Ezt a cikket – amely a BBC World Service érvrendszerét mutatja be – eredetileg az ITU Új Műsorszóró Technológiák című, Santiago de Chilében, tavaly augusztusban megtartott konferenciáján adták elő.

7. WHAT THE BBC WORLD SERVICE IS DOING

The BBC continues to operate a large worldwide network of HF and a few MF transmitters providing analogue AM services around the world. As well as this it is continually seeking and exploiting opportunities to offer its services on local FM, and occasionally MF, using stations in population centres either with its own dedicated transmitters or sharing time with rebroadcasting partners. A limited number of services are available from sound sub-carriers on direct to home TV broadcasting satellites. The World Service is increasing the range and diversity of its programmes which are available as audio sources on the Internet and in parallel, the BBC as a whole, through its BBC Online services, is positioning itself as a leading Internet content provider through its websites.

As well as this, the World Service sees the development of digital radio formats, particularly ones that can exploit the benefits of sky wave propagation as a key to the continuing viability of its radio services. For this reason it is an active member of the DRM (Digital Radio Mondiale) consortium⁸ and gives positive support to the ITU Task Group TG 10-6 (and formerly Special Rapporteur Group SRG 10A-5)⁹, both of which are committed to establishing a unified world wide standard for digital radio broadcasting.

Not only does the BBC regard digital formats as essential to the continuing commercial success of the medium, it also sees the establishment of a unified global standard as important not only to its own aspirations as a global broadcaster but also to the development of a manufacturing base for affordable receivers.

The BBC is committed to plurality in the media and so actively supports all developments which have the potential to allow a wider range and quality of services to be offered through more effective use of the radio spectrum. The indications are that digital radio formats will further this aim.

⁸ Through chairmanship of the Commercial Committee and active membership of most sub-committees and working groups. The BBC was a founder member of DRM.

⁹ Through chairmanship and active participation in TG 10-6.

Mike Hate studied Electronics and Communications at London University between 1967 and 1975 gaining BSc and PhD qualifications. During the mid 1970s he worked briefly at the Research and Development Department of British Rail and later with the London Underground in the field of vehicle dynamics. He joined the BBC in 1977 in the Transmission Capital Projects department and has subsequently worked in various capacities and in various departments associated with the capital investment programme. Following the privatisation of the BBC's transmission interests in 1997, he has been a part of the small team in the World Service which manages contracts with outside suppliers and commissions work on those elements of the transmission network which the BBC still owns. With the average life expectancy of much of the installed plant being twenty years and more, a strategic interest in technical developments, and the way in which these will impact on the BBC's current and future purchasing decisions, is clearly an important aspect of this work.

KETTŐS ROMBUSZANTENNA RENDSZER ALKALMAZÁSA A HAZAI RÁDIÓKOMMUNIKÁCIÓS ÉS MŰSORSZÓRÓ SZOLGÁLTATBAN

DÓSA GYÖRGY

1061 BUDAPEST, ANDRÁSSY ÚT 37. TEL.: 352-8030
OKL. VILLAMOSMÉRNÖK: OKL. HÍRKÖZLÉSI ÉS MŰSORSZÓRÓ SZAKMÉRNÖK

Jelen cikk a székesfehérvári rádióállomáson már korábban üzembe állított rombuszantenna rendszer üzemi működési frekvenciatartományának kibővítési lehetőségét mutatja be és vizsgálja. A meglévő – eredeti – rombuszantenna rendszerbe egy kisebb geometriai méretű, de nagyobb frekvenciatartományra méretezett és üzemelő rombuszsugárzó került beépítésre a meglévő tartószerkezetek felhasználásával és távvezérelt antenna átkapcsolási lehetőséggel. Az így kialakított kettős rombuszsugárzó rendszer, antennaváltással (külső vagy belső) a 4,5–26 MHz sávban biztosítja a sugárzást átfedéssel.

Az eredeti antenna a 4,5–15 MHz, míg az új sugárzó a 8,5–26 MHz frekvenciasávban való sugárzásra vált alkalmassá.

A cikk részletesen foglalkozik a kettős rombuszantenna tervezési feltételeivel, tervezésével, kialakításával, valamint a kiépítés utáni vizsgálatokkal, mérésekkel és az üzemi tapasztalatokkal.

1. ELŐZMÉNYEK

Az ötvenes évek második felében a Nemzetközi Rádió-távíró szolgáltatásunk jelentősen felfutott, az összeköttetések (viszonylatok) száma elérte a 16-ot, és a forgalom mennyisége, ill. időtartama is növekedett.

Jelentős konkrét új összeköttetési – szolgáltatási igények merültek fel a Posta Távíró Hivatal és az MTI részéről európai, de különösen Kína és Észak-nyugat Afrika felé – 60°/240°-os főirányokra – rádiókommunikáció forgalmazására.

Ugyancsak felmerült további műsorsugárzási igény is Spanyolország, ill. Dél-Amerikai célterületekre.

A rádiókommunikációs (rádió-távíró) szolgáltatásokat a székesfehérvári rádióállomás 15 kW-os Philips adóberendezései bonyolították le főleg rövidhullámon, csak kis részben hosszuhullámon (Európa).

A hullámterjedési számítások [5] alátámasztották és igazolták, hogy a szóban forgó viszonylatokra – célterületekre – (Kína – Észak-nyugat Afrika – Dél-Amerika – Spanyolország) kommunikációs összeköttetésekre – optimális forgalmazásra – a 17, ill. 21 MHz-es frekvenciatartomány használata a legkedvezőbb a napszak jelentős részében, különösen a napfolt maximumos időszakokban.

Dél-Amerika – Spanyolország célterület műsorsugárzására 15, 17, 21 és a 26 MHz-es műsorszórási sávok használata a legkedvezőbb évszaknak és napszaknak megfelelően.

Ez időszakban komoly sugárzási problémát jelentett, hogy a székesfehérvári rádióállomáson az optimális összeköttetések biztosítására a kívánt közép- és nagytávolságú főirányokra (pl. 60°/240°) megfelelő nyereséges – irányított antennák nem álltak rendelkezésre. További probléma volt az is, hogy az állomás rövidhullámú adóparkja csak 5–15 MHz frekvenciatartományban tudtak üzemelni, de a napfolt maximumos időszakokban, főképp a közép- és nagytávolságú viszonylatok optimális forgalmazásánál, ill. a műsorsugárzásnál feltétlenül szükségessé vált tehát a 17–26 MHz frekvenciatartomány használata is.

A székesfehérvári rádióállomáson a 60°/240°-os főirányokra csak egy rombuszantenna állt rendelkezésre, de ez is csak 60°-os főirányba, és csak kb. 5–15 MHz frekvenciatartományban biztosította a sugárzást.

1959-re kiépült egy 17, illetve egy 21 MHz-es rezonáns collineáris félhullámú dipólokból kialakított HRR 4/2/0,5 típusú síkantenna egység közepes nyereséggel, azonban ezekkel nem lehetett a felmerülő igényeket kielégíteni, ill. biztosítani (időben, mennyiségben, frekvenciában és lefedésben sem tudták a felmerült új igényeket kielégíteni).

A felmerült célterületeknek az optimális frekvenciákon történő besugárzására, ill. az összeköttetések biztosítására részletes vizsgálat történt a megfelelő optimális és gazdaságos sugárzó rendszer (antenna) meghatározására, ill. kiválasztására [1], [5].

Tekintettel arra, hogy 15–26 MHz-es frekvenciatartományban is kell biztosítani a sugárzást két főirányba, közép- és nagytávolságokra (3000 km, 8000 km, 12000 km), ezért aránylag nagy nyereségű, irányított szélessávú, lapos kilövellési szögű sugárzó kerülhetett alkalmazásra.

Kétféle antennatípus került részletes vizsgálatra, elemzésre, mégpedig:

- félhullámú elemekből kialakított irányított (reflektoros) szélessávú síkantenna rendszerek,
- haladóhullámú, szélessávú rombuszantenna rendszerek.

A vizsgálatok, számítások, elemzések és megfontolások alapján a meglévő 60°-os rombuszantenna rendszer üzemi frekvenciatartományának kibővítése látszott a legkedvezőbbnek mind műszakilag, mind gazdasági szempontból, azáltal, hogy egy másik, új – geometriailag kisebb méretű –, kb. 12–26 MHz-es sávra méretezett rombuszantennával bővülne.

Ez a másik, kisebb méretű rombuszantenna a meglévő rombuszantenna belsejébe kerül elhelyezésre, a meglévő tartószerkezetekre felfüggesztve.

Az antennák váltását – külső, ill. belső – és a kétirányú, 180°-os irányváltást távvezérelt irányváltó kapcsolókkal lehet biztosítani.

Így létrejönne egy ún. kettős rombuszantenna rendszer, melynél a régi – külső – antenna kb. 5–16 MHz tartományban, a belső – új – antennarendszer pedig kb. 9–26 MHz tartományban üzemelhetne.

A meglévő rombuszantenna kibővített alkalmazásának döntésénél az értékelések összefoglalása a következő.

- A szóban forgó sugárzási igények kielégítésére legalább egy HRR 4/4/1 típusú kb. 18 dB nyereségű többsávú irányváltós síkantenna rendszer jöhetett volna szóba (2 db 60 méteres öntartó rácsos tartószerkezettel kb. 300 m légtápvonal kiépítésével, szélessávú illesztő rendszerrel, 30 kW teljesítményre az új EMV adók telepítése miatt).

Az antennarendszer jelentős területet igényelt volna, mely nem állt rendelkezésre, továbbá igen jelentős beruházási költséggel lehetett volna megvalósítani. Tehát alkalmazása nem volt gazdaságos.

- Egy rombuszantenna kiépítése szintén nagy területet igényelt volna, azonban megfelelő szabad terület az állomás belső részén nem állt rendelkezésre.

Jelentős beruházási költségkihatás az antenna kiépítésénél, miután 4 db 25 m-es öntartó vasszerkezeti tartótorony, kb. 400 m légtápvonal, távvezérelt irányváltó kapcsolórendszer vált volna szükségessé.

A meglévő rombuszantenna rendszer üzemi frekvenciatartományának kibővítése új kisebb geometriai méretű – magasabb frekvenciatartományban üzemelő – rombuszsugárzó rendszer az alábbi jelentős előnyöket jelentette:

- a meglévő rombuszantenna rendszer tartószerkezetei felhasználhatók voltak a statikai vizsgálatok alapján,

- újabb területet nem kellett felhasználni, ill. kisajátítani az antenna építéséhez,
- új tápvonalrendszert nem kellett kiépíteni és a fentiek alapján az új antennarendszer kiépítése tehát jelentősen kisebb költséggel valósulhatott meg és kedvezően biztosítva lett a kívánt irányokba a magasabb frekvencián történő sugárzás.

A meglévő (rég) rombuszantenna rendszer összefoglaló műszaki paramétereit az 1. táblázatban foglaltuk össze.

1. táblázat

Typus	RHO 2,9/0,55/63
Működési frekvenciatartomány	kb. 4,5–16 MHz
Maximális üzemi teljesítmény	30 kW
Nyereség (G)	11–18 dB között
Vízszintes nyílásszög (γ)	24°-os,
	3 dB-es pontok
Kilövellési szög változás az üzemi frekvenciatartományban (γ)	13–46°
Antenna oldalhossza (l)	101 m
Csúcsszög (θ)	54°
Oldalszög (ψ)	63°
Nagyátló hossza (D)	180 m
Kisátló hossza (d)	90 m
Az antenna felfüggesztési magassága (h)	19 m
Sugárzó vezetékek száma	2 db
Sugárzási főirányok	60°/240° csak 60°-os főirányba üzemelt

A 2. táblázat tartalmazza a sugárzási viszonyok alakulását a régi antenna esetében.

2. táblázat

l/λ	h/λ	$\lambda(m)$	ϑ°		A főnyalábhoz viszonyított melléknyaláb amplitúdó	G (dB)
			főnyaláb	melléknyaláb		
1,5	0,298	67	45°	86°	30 %	12,1
2	0,396	50,5	35°	75°	20 %	13,2
2,85	0,564	35,44	26°	57°	10 %	15,2
4	0,796	25,1	20°	68°	16 %	16,4
5,39	1,07	18,75	13°	45°	20 %	17,4

$$\psi = 63^\circ, \quad \lambda_0 = 35,44 \text{ m}, \quad h = 0,55 \lambda_0$$

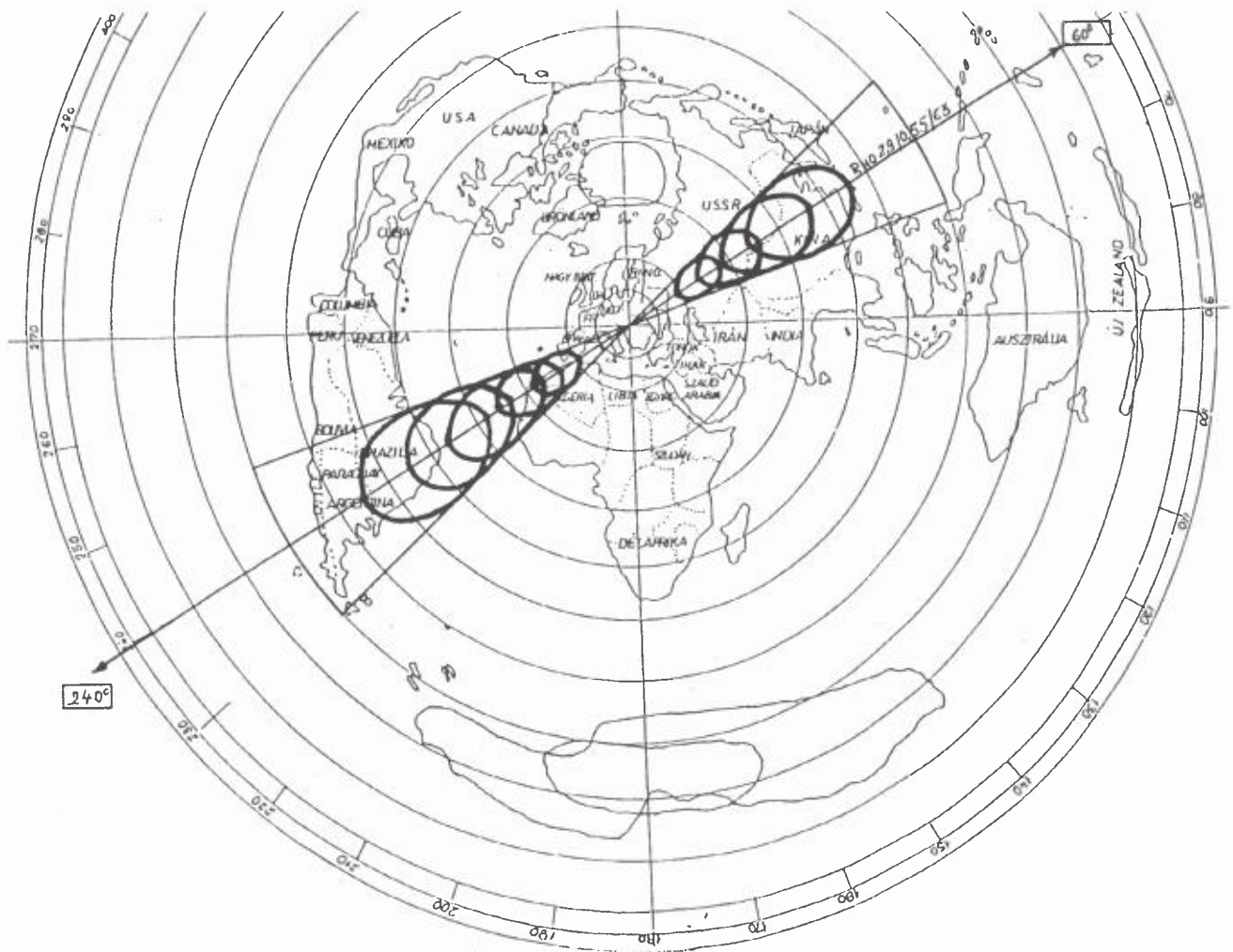
A meglévő (rég) antennarendszer sugárzási hatáskörzetét (9 MHz-en) az 1. ábra (térkép), a számított vertikális és horizontális karakterisztika alakulását a 2/a. és a 2/b. ábrák szemléltetik, az antennarendszer kialakítását pedig a 3. ábra mutatja be.

2. AZ ÚJ, KETTŐS ROMBUSZANTENNA RENDSZER KIÉPÍTÉSE

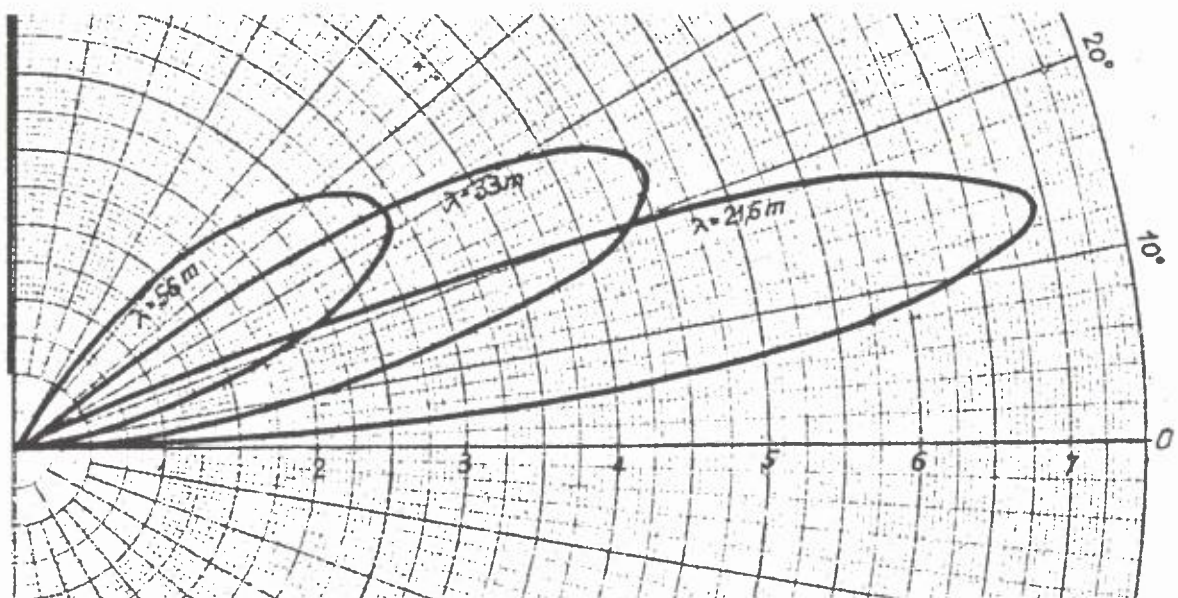
A döntés alapján tehát a meglévő 60°-os rombuszantenna frekvenciasávjának kibővítése, kisebb méretű rombuszantenna beépítésével a tartószerkezetek felhasználásával, vagyis egy síkú, kettős rombuszrendszer megvalósítása

az alábbi munkálatok, vizsgálatok és mérések lebonyolítását tette szükségessé:

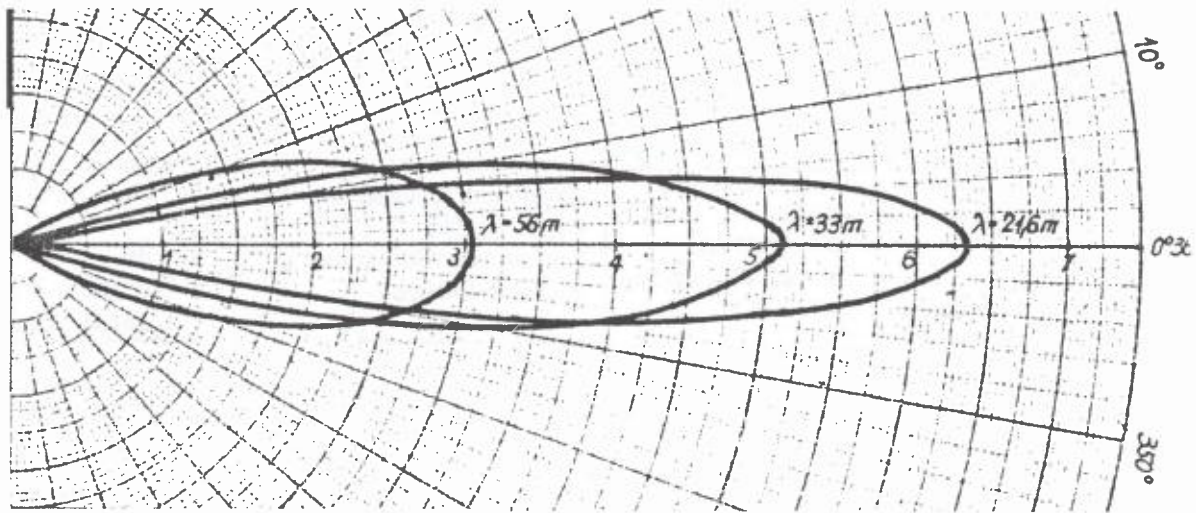
- megtervezése az új magasabb frekvenciatartományú rombuszantenna rendszernek, figyelembe véve a meglévő (alacsonyabb frekvenciatartományú) rombuszsugárzó rendszert,
- a tervek alapján a meglévő 60°-os rombuszantenna rendszerbe a tervezett új rombuszsugárzó kivitelezése, beépítése, ill. végleges kialakítása irányváltó kapcsolókkal,
- a meglévő (rég), ill. a beépített új rombuszantenna rendszeren az antenna jellemzők mérése, ill. a kialakított kettős rombuszsugárzó rendszeren próbaüzem és üzemi mérések lefolytatása.



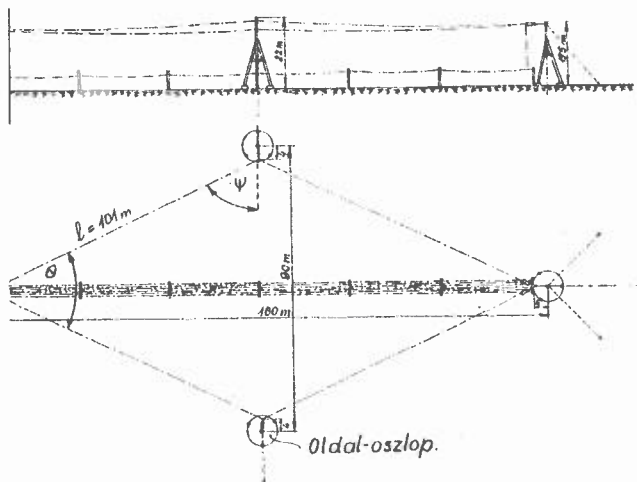
1. ábra. A meglévő (rég)i rombuszantenna sugárzási hatáskörzete



2a. ábra. A meglévő (rég)i rombuszantenna elméleti vertikális karakterisztikája



2b. ábra. A meglévő (régi) rombuszantenna elméleti horizontális karakterisztikája



3. ábra. A régi rombusz sugárzó kialakítása

2.1. Belső, új rombusz tervezési feltételei, kialakítása

Figyelembe véve, hogy a meglévő sugárzó rendszer tartószerkezete (nyelves faoszlopok) adott volt, az új belső antenna elektromos tervezése ezen meglévő adottságból és a célterületi igények optimális sugárzás biztosításából kellett, hogy kiinduljon.

Alapértelmezésnél a lefedés biztosításához a célterületi távolság az útvonal és az elevációs szögek meghatározásából kellett kiindulni.

A geometriai terjedési út meghatározásánál a rétegek állandó magasság változása miatt átlagértékkel kellett számolni. Csak F2 réteges terjedésre 320 km-es átlagos magasság lett figyelembe véve.

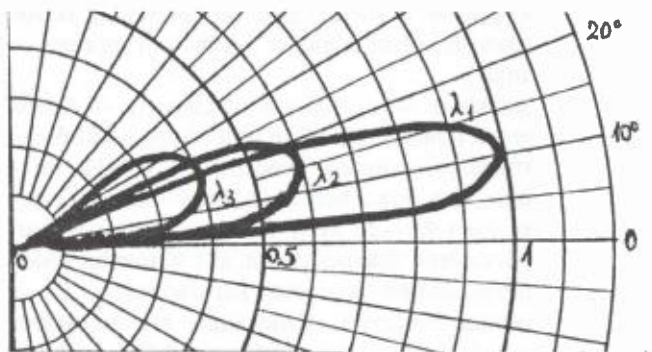
A tervezésnél tehát alkalmazkodni kellett a meglévő antennához, ez kétféleképpen lett figyelembe véve:

- a meghatározásra került a régi (külső) sugárzó rendszer paraméterei és egy hasonló jellemzőkkel bíró (ψ ; $1/\lambda$; h/λ), de frekvenciában eltolt antennát kellett tervezni,
- a régi (külső) rombusz és a felfüggesztési magasság megszabta keretek között biztosítani – nem azonos paraméterekkel – a maximális üzemi feltételeket. Kétvagy többutas terjedéssel – magasabb üzemi frekven-

ciák esetében –, alacsonyabb kilövellési szöggel tehát elérhető több ezer kilométeres célterületek [3], [5].

A tervezésnél figyelembe kellett venni, hogy a belső sugárzó rendszer optimális felfüggesztési magassága alacsonyabbra adódott, mint a meglévő külső rombusz magassága, azonban a belső rombusz is azonos magasságra felfüggeszthető, miután az állandó h magasság esetében a hullámhossz csökkenésével a kilövellési szög laposabb, ami megfelelt a kívánt célterületi igényeknek.

A belső, új rombuszantenna számított elméleti vertikális karakterisztikájának alakulását a 4. ábra szemlélteti.



4. ábra. Az új (belső) rombuszantenna elméleti vertikális karakterisztikája

A részletes és bonyolult tervezés alapján több változat került vizsgálatra [2], [3].

A végleges, optimális terv alapján a belső, kisebb méretű rombuszsugárzó rendszerre a 3. táblázat szerinti geometriai méretek, ill. paraméterek adódtak.

A sugárzási viszonyok alakulását az új rombuszantenna esetére pedig a 4. táblázat foglalja össze.

Az új antennarendszer tervezése 1965. szeptemberében megtörtént. Az új, belső sugárzó rendszer legyártása, kivitelezése és beépítése a meglévő rendszerbe 1966. áprilisára valósult meg.

Az új sugárzó rendszer 7x1,5 mm-es sodrott bronzhuzalból került kivitelezésre és a hullámmellenállás stabilizálása érdekében kétszálas kivitelenben. A két távolság (az oldalpontoknál) a végleges beállítás után 2,5 m-re adódott. Az

emésztő tápvonal pedig speciális horganyzott vashuzalból került kivitelezésre végelegesen 510 méter hosszban.

3. táblázat

Típus	RHO 3,64/0,974/63
Működési frekvenciatartomány	kb. 8,5–26 MHz
Maximális üzemi teljesítmény	30 kW
Nyereség (G)	13–18 dB között
Vízszintes nyílásszög (γ)	24°-os, 3 dB-es pontok
Kilövellési szög változás az üzemi frekvenciatartományban (γ)	10° ~ 27°

Antenna oldalhossza (l)	71 m
Csúcshög (θ)	54°
Oldalszög (ψ)	63°
Nagyító hossza (D)	127 m
Kisító hossza (d)	64 m
Az antenna felfüggesztési magassága (h)	19 m
Sugárzó vezetékek száma	2 db
Sugárzási főirányok	60°/240° főirányba üzemelt

4. táblázat

l/λ	h/λ	$\lambda(m)$	ψ°		A főnyalábhoz viszonyított melléknyaláb amplitúdó	G (dB)
			főnyaláb	melléknyaláb		
2,13	0,57	33,33	26°	66°/17°	4%/13%	13,2
2,5	0,67	28,4	22°	52°/35°	6%/12%	14,4
3	0,8	23,7	18°	44°/36°	5%/8%	15,1
3,5	0,93	20,3	15°	40°/60°	9%/10%	15,8
4	1,07	17,75	14°	37°/24°	7%/18%	16,2
4,5	1,2	15,8	13°	—	—	16,9
5	1,34	14,2	12°	42°/10°	7%/10%	17,4
6,5	1,61	11,8	10°	—	—	18,2

$$\psi = 63^\circ, \quad \lambda_0 = 19,5 \text{ m}, \quad h = 0,974 \lambda_0$$

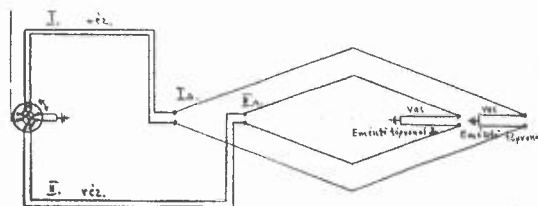
A tervezés alapján a kettős rombuszsugárzó rendszer-technikai kialakítását egy-, ill. kétirányú sugárzás lehetőségére az 5. és 6. ábra szemlélteti.

A sugárzási hatáskörzet alakulását pedig a 7. ábra (térkép) mutatja be a sávközépen (16 MHz), a kettős rombuszantenna rendszer végleges geometriai méreteit a 8. ábra mutatja.

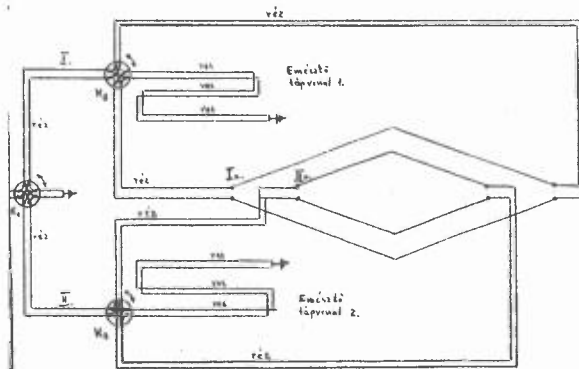
Az így kialakított – egy síkban elhelyezett – kettős rombuszantenna rendszer lehetővé tette, hogy a külső nagyobb méretű rombuszsugárzó rendszer 5–16 MHz frekvenciatartományban, míg a belső (új), kisebb méretű rombuszantenna pedig 8,5–26 MHz frekvenciatartományban üzemeljen, távvezérelt átkapcsolással, két főirányba történő irányváltással, szintén távvezérelt kapcsolással.

A két antenna sugárzási viszonyainak eredményeiből – főnyalábok eleváció szögeinek változásaiból – látható, hogy a külső antennarendszer főnyalábja 13–46° között változik az adott frekvenciatartományban, míg az új (belső) rombuszsugárzó főnyaláb elevációs szöge 10–27° között változik a magasabb frekvenciatartományban. Ez biztosítja, hogy az új, belső antennával, magasabb frekvenciákon, többutas terjedés esetén 8–12000 km-es célterületek érhetők el (7. ábra).

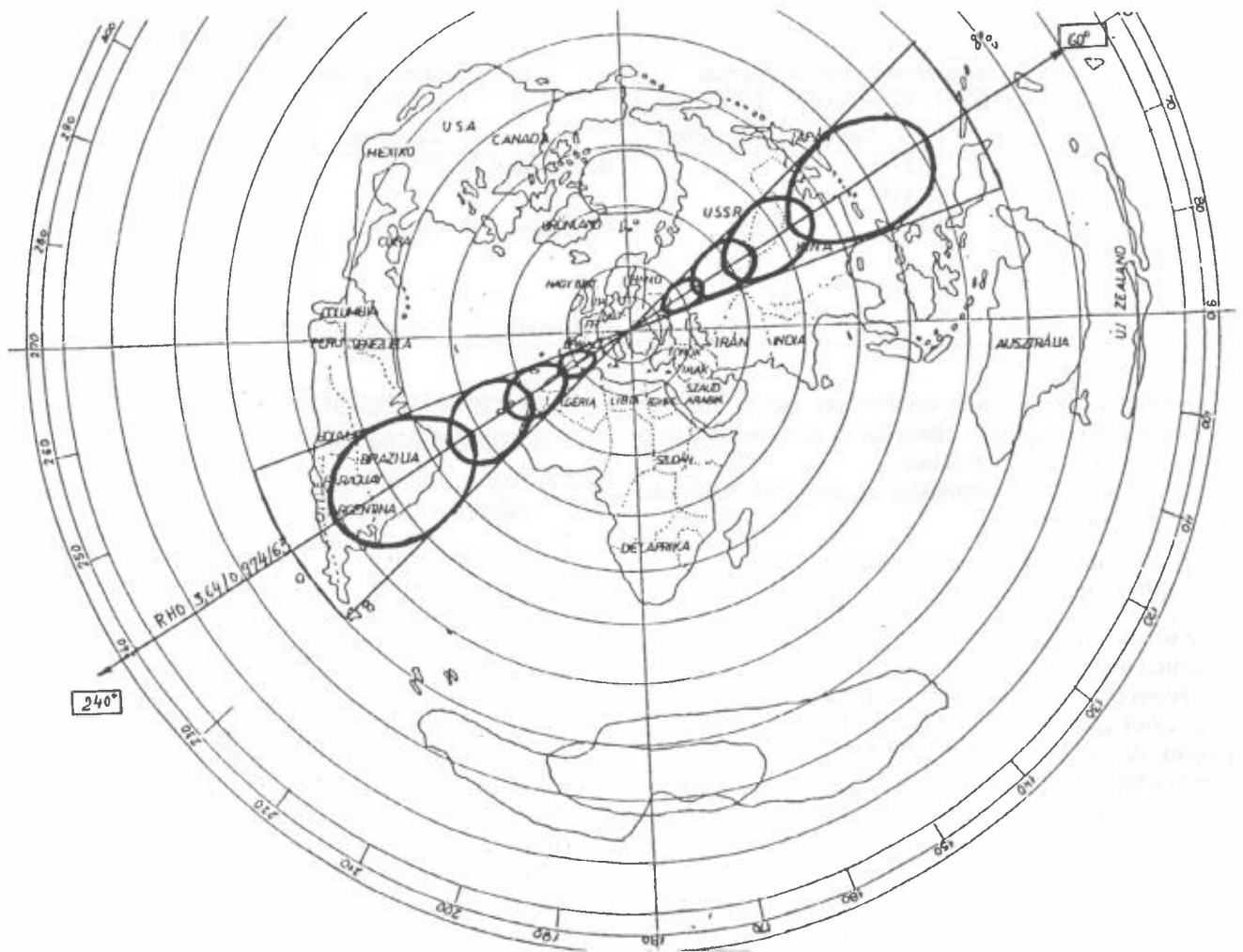
A kívánt célterületre számított rombuszsugárzó fő paramétereit: ($\psi = 63^\circ$; $l/\lambda = 3,64$; $h/\lambda = 0,974$), tehát nagytávolságú sugárzásra optimális.



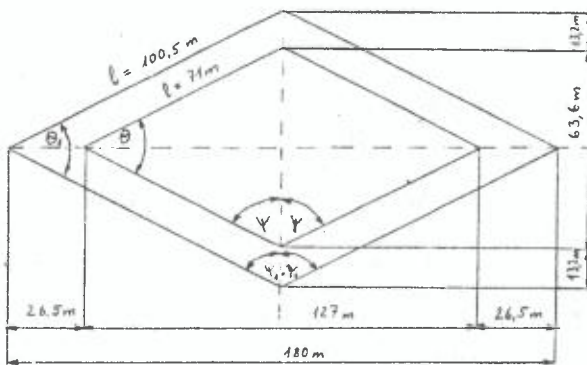
5. ábra. A tervezett új kettős rombusz sugárzó rendszer-technikai kialakítása egyirányú sugárzási lehetőségekre



6. ábra. A tervezett új kettős rombusz sugárzó rendszer-technikai kialakítása kétirányú sugárzási lehetőségekre



7. ábra. Az új rombuszantenna rendszer sugárzási hatáskörzete



8. ábra. A kettős rombuszantenna rendszer elvi kialakítása

2.1.1. A sugárzási és hullámellenállás értékek ellenőrzése az új és a régi antennarendszeren

Az antenna hatások megállapításához, ill. az ellenőrző értékelésekhez szükséges a sugárzási ellenállás konkrét meghatározása, ill. annak változása a frekvencia függvényében.

A hullámellenállás értéke a rombusz mentén nem állandó, a rombuszhuzalok közötti legnagyobb távolságig nő. A legnagyobb értékét a kisátlónál éri el, majd ismét csökken. Ezen ingadozás csökkentése érdekében kellett a rombusz szárait két vezetékből kialakítani, így a hullámellenállás ingadozás kisebb értékben változik.

A sugárzási ellenállás (R_s) az alábbiak szerint határozható meg:

$$R_s = 240 \left[\ln \left(4 \Pi \frac{1}{\lambda} \cos^2 \psi \right) + 0,577 \right]$$

ahol l az antenna oldal hossza; ψ a rombusz tompaszöge.

A rombuszantenna átlagos hullámellenállása pedig:

$$Z_0 = 138 \log \frac{a \sqrt{a^2 + b^2}}{b \cdot d}$$

alapján határozható meg, ahol a a kisátló hossza; b az oldalhuzalok közötti távolság; d a huzalátmérő.

Az 5. táblázat összefoglalóan tartalmazza a számítások alapján mindkét rombuszantenna sugárzási ellenállás értékeit a sáv elején, közepén és végén, valamint a hullámellenállás változását az antenna hosszában.

5. táblázat

f (MHz)	Sugárzási ellenállás változás		Hullámellenállás változás		
	Külső ant. R_s (ohm)	Belső ant. R_s (ohm)		Külső ant. Z_0 (ohm)	Belső ant. Z_0 (ohm)
4,5	278	—	Bemenetnél	593	533
8,5	345	—			
9,0	—	317	Félhossznál	716	676
15,5	—	373			
16,0	411	—			
25,5	—	425	Kisátlónál	758	720

Az ellenőrző számítások eredményei szerint mindkét sugárzó rendszer sugárzási ellenállás és hullámellenállás a szokásos értékeknek megfeleltek.

A fenti adatokkal, számítással az antennák hatásfoka 50–65 % között változik.

2.1.2. Az emésztővezeték paramétereinek ellenőrzése

Az új antennarendszer emésztővezeték hullámellenállása szintén 600 ohmra lett választva.

A beeső és a visszavert hullám viszonyára – csillapításra – az előírt érték 0,1 volt, ehhez 490 méter tápvonalhossz adódott. Az emésztővezeték a megfelelő csillapítás végett 4 mm-es horganyzott vashuzalból került kialakításra.

A csillapítás és a hossz ellenőrzése:

Az emésztővezeték hullámellenállása meghatározható az alábbiak szerint:

$$Z_0 = 120 \ln \frac{2a}{d} \quad (\text{ohm})$$

ahol a a vezetékek távolsága, d a huzalátmérő.

30 cm vezeték-távolság és 4 mm huzalvastagság esetén tehát az emésztővezeték hullámellenállása:

$$Z_0 = 603 \cdot \text{ohm}$$

A fajlagos ellenállása az emésztő vezetéknek az alábbiak alapján:

$$R_1 = \frac{5,5 \cdot 10^3}{r} \sqrt{\frac{\mu_r \cdot \rho}{\lambda}} \quad (\text{ohm/m})$$

A számítások alapján a fajlagos ellenállás értéke tehát:

$$R_1 = 1,3 \text{ ohm/m}$$

Az emésztővezeték hullámellenállásának és a fajlagos ellenállásának ismeretében a csillapítási tényező:

$$\alpha = \frac{R_1}{Z_0} = \frac{1,35}{603} = 2,25 \cdot 10^{-3}$$

Az emésztővezeték szükséges optimális hossza meghatározható az alábbi összefüggésből:

$$l = \frac{\ln \frac{1}{2\alpha}}{2\alpha} \quad (\text{m})$$

Az adatok ismeretében a belső rombuszantenna emésztővezetékének végleges optimális hossza tehát $l_e = 510$ m-re adódott és a végleges beállítás is erre a hosszra történt meg.

2.2. Mérések és vizsgálatok az elkészült kettős rombuszantenna rendszeren

2.2.1. Impedancia és vízszintes karakterisztika ellenőrző mérések

Az elkészült egysíkú kettős rombuszantenna rendszeren az előzetes elméleti megfontolások az összehasonlítás és használhatóság szempontjából a mérések, ill. a vizsgálatok az alábbiakra terjedt ki:

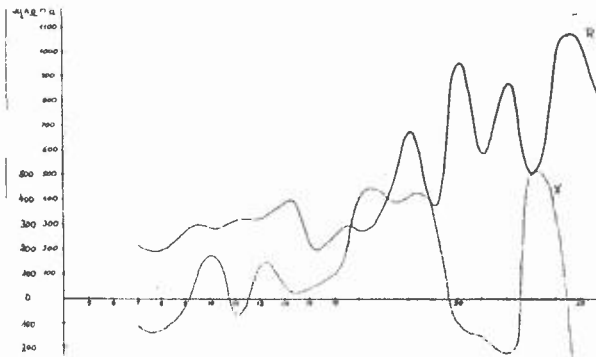
- a meglévő külső rombuszantennán impedancia mérés az üzemi frekvenciatartományban, az új, kisebb méretű rombuszantenna beépítése előtt és után,
- az új – kisebb – méretű rombuszantenna impedancia mérése a frekvencia függvényében a beépítés után,
- a meglévő – külső – rombuszantenna vízszintes iránykarakterisztikájának ellenőrzése 5, 10, 15 MHz frekvenciákon az új kisebb méretű – belső – rombuszsugárzó beépítése előtt és után,
- az új – belső – rombuszantenna vízszintes iránykarakterisztikájának ellenőrző felvétele 10, 15, 19 és 23 MHz-en,
- a kettős rombuszantenna rendszeren próbaüzem és üzemi mérések elvégzése, ill. próbaadások lefolytatása több üzemi frekvencián.

A kettős rombuszantenna rendszeren a fentiekben megadott részletes mérések és vizsgálatok 1966 április és szeptember hónapokban kerültek lebonyolításra.

A 9. ábra a belső, új rombuszsugárzó impedancia érték alakulását mutatja be a beállítás utáni üzemi frekvenciatartományban. Az impedanciaellenőrző mérésekből megállapítható volt, hogy a belső rombuszantenna beépítésével mindkét rombuszsugárzó tápponti impedancia értéke a gyakorlatnak megfelelő és a külső sugárzó impedancia változás jellege ugyanaz maradt.

A belső sugárzó impedancia változásának maximális és minimális értékeiből meghatározásra került a maximális állóhullámarány értéke, mely 2,2-re adódott, ami elfogadott érték.

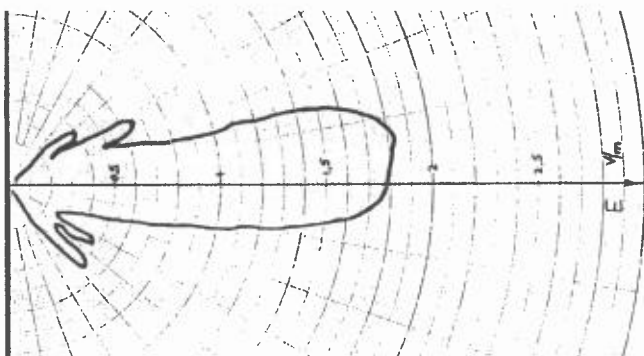
A vízszintes iránykarakterisztika ellenőrző mérések arra terjedtek ki, hogy a külső rombuszantenna vízszintes iránykarakterisztikájának alakulásában történt-e eltérés vagy erős változás azáltal, hogy az új – belső – rombuszsugárzó beépítésre került, azaz összehasonlítható legyen a külső rombuszsugárzó vízszintes iránykarakterisztikája a beépítés előtti, ill. utáni állapottal.



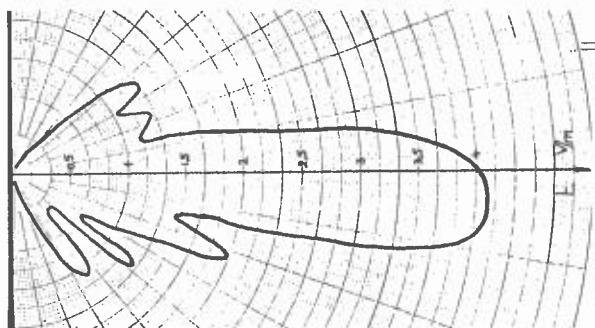
9. ábra. A belső (új) rombuszantenna impedancia értékének alakulása a frekvencia függvényében

A vízszintes iránykarakterisztika felvétele a külső rombuszantenna esetében 5, 10, 15 MHz frekvenciákon történt az új antenna beépítése előtt, ill. után.

A külső rombuszantenna vízszintes iránykarakterisztikájának alakulását szemlélteti a 10. és 11. ábra beépítés előtti és utáni állapotban 15 MHz frekvencián.

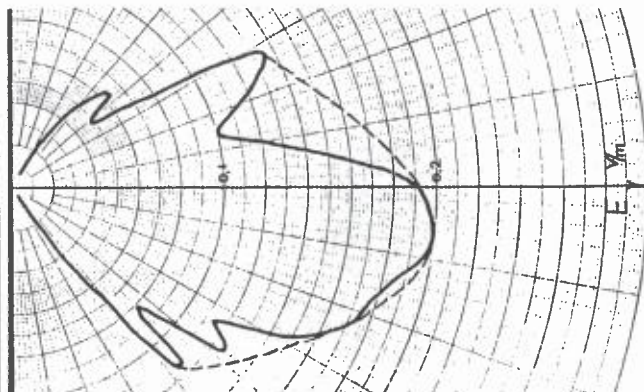


10. ábra. A meglévő (régi) rombusz sugárzó mért vízszintes iránykarakterisztikájának alakulása a belső (új) sugárzó beépítése előtt



11. ábra. A meglévő (régi) rombusz sugárzó mért vízszintes iránykarakterisztikájának alakulása a belső (új) sugárzó beépítése után

Az elkészült belső rombuszsugárzó vízszintes iránykarakterisztikájának alakulását pedig a 12. ábra mutatja be 23 MHz frekvencián.



12. ábra. A tervezett új (belső) rombuszantenna mért vízszintes iránykarakterisztikájának alakulása 23 MHz frekvencián

A vízszintes karakterisztikák ellenőrző felvétele a BHG 15 kW-os, ill. az 1963-ban üzembe helyezett 1 kW-os Vateg rövidhullámú adókkal történt, melyek 5–27 MHz tartományban üzemeltek.

A vízszintes karakterisztika mérési diagrammok eredményei bizonyították, hogy a környező kisebb zavaró tényezőktől származó ingadozásoktól eltekintve (környező tápvonalak, növényzet alakulása, terepadottságok stb.) a rombuszantennák vízszintes iránykarakterisztikái megfelelően alakultak és a beépített kisebb méretű rombuszsugárzó semmiféle változást, rendellenességet nem eredményezett a külső rombuszantenna vízszintes iránykarakterisztikájában.

2.2.2. A kettős rombuszantenna rendszer próbaüzemi vizsgálata különböző frekvenciákon

A beépített új, ill. a külső (régi) rombuszantennákon különböző frekvenciákon a próbaüzemi vizsgálatok a 15 kW-os BHG, ill. az 1 kW-os Vateg rövidhullámú adókkal történtek üzemi beállításban. Ezen vizsgálatok arra terjedtek ki, hogy az elkészült új rombuszsugárzó, ill. a két antenna – mint kettős rombuszsugárzó – külön-külön 15 kW-os, ill. 1 kW-os üzemi adókra történő kapcsolása esetén különböző frekvenciákon, üzemszerű állapotban való használat esetén az adóberendezések beállítási értékei mennyire egyeznek vagy változnak meg a szokásos üzemi értékekhez képest.

6a. táblázat (15 kW-os adó üzeme esetén)

Megnevezés:	Üzemi frekvenciák (kHz)		
	7464	10911	15670
I_a "A"	1,4 A	1,17 A	0,88 A
I_a "B"	1,67 A	1,64 A	0,6 A
U_a	9,5 kV	9 kV	9,6 kV
I_{tp} "A"	6 A	6,7 A	3,5 A
I_{tp} "B"	5 A	5,4 A	6,1 A
Antennaáram és feszültségértékek a rombuszsugárzó táppontjánál			
I_{tp} "B"	6,5 A	7 A	5,5 A
U_{tp} "A"	2,3 kV _{cs}	2,4 kV _{cs}	2,1 kV _{cs}
U_{tp} "B"	2,27 kV _{cs}	2,15 kV _{cs}	2 kV _{cs}
N_{ki}	20 kW	17,7 kW	11 kW

6b. táblázat (1 kW-os adó üzeme esetén)

Megnevezés:	Frekvenciák (kHz)	
	23,48 MHz	19,02 MHz
I_k "A"	210 mA	210 mA
I_k "B"	290 mA	195 mA
$U_a =$	3,2 kV	3,1 kV
U_{tp} "A"	0,27 kV _{cs}	0,28 kV _{cs}
U_{tp} "B"	0,26 kV _{cs}	0,26 kV _{cs}
N_{ki}	0,95 kW	0,96 kW

A 6. táblázat összefoglalóan tartalmazza az új antennarendszernek 15 kW-os, ill. 1 kW-os adóval történő próba-

adások eredményeit, üzemszerű sugárzás esetén különböző üzemi frekvenciákon.

Az új antennarendszeren a próbaadások 15, ill. 1 kW-os teljesítménnyel különböző frekvenciákon kedvező eredménnyel bonyolítottak le, az antennarendszeren rendelkezés és hiba a próbaadások alatt nem történt.

2.2.3. Próbaüzemi sugárzások az új, kettős rombuszantenna rendszeren

Az üzemi mérések és vizsgálatok befejeztével az új kettős rombuszantenna rendszer megkezdte próbaüzemét, azaz a rendszeres Nemzetközi Postai távíró adásait Peking felé $f = 15,550$ kHz frekvencián. Az adások 1966 júliusában kezdődtek meg 15 kW teljesítménnyel, ill. kis részben 1 kW teljesítménnyel.

A több hónapos próbaüzem alatt az antennarendszeren semmiféle rendellenesség, hiba nem történt és a szóban forgó rádiótávíró összeköttetéseket az Üzemközpont tájékoztatása szerint naponta, rendszeresen jó eredménnyel vették.

IRODALOMJEGYZÉK

- [1] Dósa György – Kiss Lajos: Javaslat a meglévő székesfehérvári, 65°-os rombuszantenna rendszer üzemi és frekvenciasávjának kibővítésére; az ezzel kapcsolatos tervezések és vizsgálatok. Tanulmányterv: Budapest, 1964. szept.
- [2] Dr. Istvánffy Edvin: Tápvonalak, antennák, hullámterjedés. Tankönyvkiadó. Budapest, 1967.

Dósa György a Budapesti Műszaki Egyetem Villamosmérnöki Karán 1955-ben szerzett oklevelet. 1970-ben pedig Rádióműsorszóró és Hírközlő Szakmérnöki oklevelet. 1955-től a Posta Rádióműszaki Hivatalban, ill. a Posta Rádió és Televízió Műszaki Igazgatóság Műszaki Osztályán kezdett dolgozni. Munkaterülete közép- és rövidhullámú műsorszóró, ill. hosszú és rövidhullámú kommunikációs adóberendezések, antennarendszerek üzemeltetése, korszerűsítése, valamint új adóberendezések és antennarendszerek telepítése. 1962-től a Posta Rádió és Televízió Műszaki Igazgatóság (1992. június 30-tól Antenna Hungária Rt.) fejlesztési osztály

1967-től a Nemzetközi forgalmazás kibővült 17630 kHz frekvenciára is, majd az MTI is megkezdte a kommunikációs sugárzást Peking felé 17468 kHz frekvencián.

A hatvanas évek végétől a kettős rombuszantenna rendszer belső sugárzója időközönként 21685 kHz frekvencián Kína felé, ill. Dél-Amerika és Spanyolország célterületekre 17795 kHz frekvencián magyar és angol nyelvű sugárzást is végzett időszakosan kedvező eredménnyel.

3. ÖSSZEFOGLALÁS

A székesfehérvári rádióállomáson a meglévő 60°-os rombuszantenna rendszerénél a megadottak szerint kialakított egysíkú kettős rombuszsugárzó rendszerrel tehát lehetővé vált, hogy kb. 4,5–26 MHz frekvenciatartományban lehessen sugározni optimálisan a két antennával, átkapcsolással 60°/240°-os főirányokba irányváltással, szintén távvezérelt átkapcsolási lehetőséggel.

A kettős rombuszantenna rendszer kialakításának rendkívüli előnye volt, hogy az új – belső – sugárzó a meglévő antennatartó szerkezetre került felszerelésre, tehát új tartószerkezetekre és tápvonalra nem volt szükség, új területet sem kellett felhasználni vagy kisajátítani. Kiépítése tehát gazdaságos is volt.

A kettős rombuszantenna rendszer üzembe helyezésétől igen kedvező eredménnyel végzett postai és MTI rádió-kommunikációs sugárzást 17 és 21 MHz-es sávban Peking, ill. Marokkó viszonylatokba, ill. műsorsugárzást részlegesen Dél-Amerika és Spanyolország felé. A kiépített kettős rombuszantenna rendszer és ezzel kapcsolatos tervezés, vizsgálatok és alkalmazása hazai vonatkozásban egyedülálló volt.

A kettős rombuszantenna rendszer 1966-tól a kilencvenes évek elejéig, majdnem 30 évig, kedvező eredménnyel üzemelt a rádiótávíró és a műsorszóró szolgálatában.

- [3] E. A. Laport: Radio Antenna Engineering. London–New York, 1952.
- [4] Zuhrt: Strahlungswiderstand und Gewinn von Rombusantennen mit angenaherter Berücksichtigung der Strahlungsdampfung. AEÜ 1955.
- [5] Kenneth Davies: Ionospheric radio propagation. NBS 1965.

vezető-helyettese 1996 júliusáig. Ezen időszakban a hazai közép és rövidhullámú adó és antennahálózat fejlesztési, beruházási és rekonstrukciós munkáival, valamint hullámterjedési és hálózatfejlesztési vizsgálatokkal foglalkozott. 1962-től tagja a Híradástechnikai Tudományos Egyesületnek és ezen belül esetenként részt vett a hazai rádió műsorsugárzási kérdések vizsgálatában. Jelenleg tagja a HTE Szenior Klubjának. 1996-ban HTE Ezüst Jelvény kitüntetést kapott. Több szakcikk szerzője, illetve társszerzője.

INTELLIGENT VEHICLE FOLLOWER SYSTEM CONTROLLED BY SATELLITE

K. GYENES

TECHNICAL UNIVERSITY OF BUDAPEST
BERTALAN L. STREET 2. 1111-H BUDAPEST, HUNGARY
TEL: (36-1) 463-1993, FAX: (36-1) 463-3087
E-MAIL: GYENES@KAUTKKA.BME.HU

This paper presents the most important aspects of up-to-date computerized solution of the vehicle follower system. Nowadays there is a great demand from the side of official and private transport firms to follow their vehicles. Very necessary the knowledge of the real distance and the duration time of transportation. The discussed system provides a real data stream about the moving of a target (car, lorry, train, tram, etc.).

The system presented in this article monitors the position, time and speed coordinates of moving object. The system uses messages send by the popular satellite system called GPS (Global Positional System). The messages decoded by a local microcomputer, and the compressed information stored in a non-volatile memory for the elaboration. There will be discussed the hardware and software components and the questions of data transmission, too. The paper also deals with the practical application of the discussed method.

Keywords: GPS systems, Remote Control, Microcomputer, Data Transmission

1. INTRODUCTION

It is fact, that the number of the vehicles in the last some years increased, and the importance of the road transportation became higher and higher. The requirements of the traffic are proportionally higher. Therefore the aim of the high level control of the traffic raised.

The base of the traffic control is the real time data base. In the years of 1990's the base of the information was the paper documentation made by the driver and at case of lorries and camions the registrants derived from tachograf.

Nowadays the high level microelectronics provides many new elements, which gives the possibility to develop a new generation of the vehicle follower systems having a lot of advantages. These are the high performance, low consumption, high safety, small dimension and last but not least the low price.

The base idea comes from the war industry, where the satellite control is used more then twenty years. The GPS satellite system with reduced accuracy can used for civil application. For our purpose this reduced position accuracy (cca. 50 meters) is enough.

Our system needs neither telecommunication network (telephone line) nor radio connection. Therefore not necessary any permission from the authority.

I would like to express my special thanks to my students, *K. Szőnyi*, *B. Kiss*, *G. Hidegkuti* and *C. C. Obi* for their diligent and useful participation in this project.

2. THE PRINCIPLE OF THE VEHICLE FOLLOWER SYSTEM

The portable instrument is able to follow any moving object, recording its four dimensions (x, y, z and time coordinates). In the later part we call this instrument PDR = Portable Data Recorder. The functional elements of the PDR are:

- antenna
- GPS receiver
- microprocessor unit
- storage memory
- communication unit

The functional connection diagram is shown in Fig. 1.

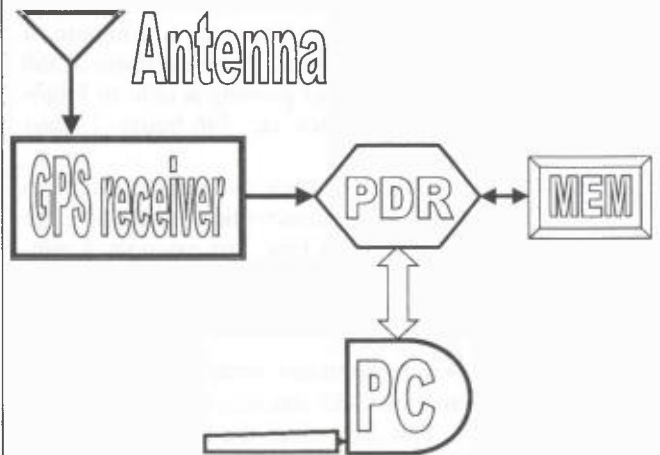


Fig. 1. The logical scheme of the PDR system

The antenna unit consists of a small parabolic collector and a high frequency amplifier. The necessary supply voltage is lead to this unit from the GPS receiver by a coaxial cable, and this cable leads the high frequency signal back.

This unit has to place on the top of the car by a magnetic fixer, or it is possible to fix on the car window by tree plastic legs. In Fig. 2 (see in colours) are shown two antenna units: the GA27 low profile GPS antenna placed on left side produced by firm Garmin, on the right side Model SM25 made by Hirschmann.

The sign derived from the collector enlarged by the preamplifier is lead to the GPS receiver. This unit demodulates the complex sign, selecting the useful information from the high leader frequency.

After decoding the microprocessor of the GPS receiver elaborates the source sign, and according to his own program builds up the different messages. Finally the message will send to the PDR by the selected protocol. The user can select the kind of message from a set, the speed of transmission and the frequency of the message (i.e. the repetition time).

The most of the GPS receivers have two serial communication channels, one for the ASCII character stream, and one for the binary coded communication.

In the Fig. 3 (see in colours) shows three GPS receivers produced by tree manufacturer. The receiver on the left is a Garmin GPS 25LP, in the middle the Marconi μ -blox GPS-PS1, on the right is the Marconi GPS-MS1 receiver, which is the smallest receiver has been made till 1998.

The message derived from the GPS receiver will read by the PDR microprocessor via an internal serial communication line. This processor stores the message in a temporarily buffer. Controlled by its program the processor checks the message, in aspect of its intact, selects the necessary information, and stores it in compressed format in the non volatile memory. The processor stays in idle mode until arriving the next message from the GPS receiver.

The compression is very important because of the storage capacity of the memory is not too high. The frequency of the messages influences the working time of the PDR. Let us see an example for the calculation. Supposing, that one compressed message needs 16 bytes memory room, and the message frequency is 1 minute, if the memory capacity is 512 Kbytes, there are room 32768 messages, also the PDR without clearing is able to follow a vehicle through 32768 minutes, i.e. 546 hours, 22 days continuously working.

If the user needs more accuracy can degrees the frequency, but if he needs longer observation period time, he can program higher repetition time, for example 5 minutes. In this last case the observation time can raise up to 110 days, i.e. more than 15 weeks. In practice this time enough for a long foreign trip.

The PDR has two independent serial communication channels, one toward the GPS receiver the other for the PC communication. This unit senses the line from the PC, and if any character arrives from it, immediately realizes it and changes his mode to PC communication.

In this mode the receiving of messages from the GPS receiver will suspend. The PDR wait command from the PC. This commands will discussed later, in Chapter 5. The PDR never initializes communication neither the GPS nor the PC. The serial communication line toward the PC is a standardized RS232 interface. The parameters of the PDR-PC communication programmable, too.

3. OVERVIEW OF THE GPS SYSTEM

Global Positioning Systems (GPS) are space-based radio positioning systems that provide 24 hour three-dimensional position, velocity and time information to suitably equipped users anywhere on or near the surface of the Earth (and sometimes off the earth). Global Navigation Satellite Systems (GNSS) are extended GPS systems, providing users with sufficient accuracy and integrity information to be

useable for critical navigation applications. The NAVSTAR system, operated by the U.S. Department of Defense (DoD), is the first GPS system widely available to civilian users. The Russian GPS system, GLONASS, is similar in operation and may prove complimentary to the NAVSTAR system.

The Global Navigation System comprises of parts usually referred to as segment. As such, we have the Space, Control and User segments. In the next part we discuss the three segments detail.

3.1. The Space Segment

This consists of the GPS satellites, usually known as Space Vehicles (SV), which send radio signals from space. The nominal GPS Operational Constellation consists of 24 satellites that orbit the earth in 12 hours. The Picture 6 (see in colours) shows these satellites and orbits.

These satellites orbit at an altitude of 11 000 miles (10 900 nautical miles), have an orbital plane inclined at 55 degrees to the equatorial plane, and a life span of 7.5 years. There are often more than 24 operational satellites as new ones are launched to replace older satellites. The satellite orbits repeat almost the same ground track (as the earth turns beneath them) once each day. The orbit altitude is such that the satellites repeat the same track and configuration over any point approximately every 24 hours (4 minutes earlier each day). There are six orbital planes (with nominally four SVs in each), equally spaced (60 degrees apart). This constellation provides the user with between five and eight SVs visible from any point on the earth.

The Picture 7 (see in colours) shows the constellation of the satellites and orbits around the earth.

The position of satellites is denoted by letter-number combination. This codification used at the sign of satellite and orbits. The codification is shown in the Fig. 8 (see in colours).

The name of a satellite reflects its status, too. For example *D4a* means that the satellite identified by 4A position is active (lowercase *a*).

3.2. The Control Segment

The Control Segment consists of a system of tracking stations (GPS Master Control and Monitor Network) located around the world. The Master Control facility is located at Schriever Air Force Base (formerly Falcon AFB) in Colorado. These monitor stations measure signals from the SVs, which are incorporated into orbital models for each satellite. The models compute precise orbital data (ephemeris) and SV clock corrections for each satellite. The Master Control station uploads ephemeris and clock data to the SVs. The SVs can then send subsets of the orbital ephemeris data to GPS receivers over radio signals.

The functional connection of control station, satellite and GPS receiver is shown in Fig. 9 (see in colours).

The American NAVSTAR system uses five monitor stations: Hawaii, Ascension Island, Diego Garcia, Kwajalein and Colorado Springs. In Fig. 10 (see in colours) is shown the placement of these control stations.

3.3. The User Segment

The GPS User Segment consists of the GPS receivers and the user community. GPS receivers convert SV signals into position, velocity, and time estimates. Four satellites are required to compute the four dimensions of X, Y, Z (position) and Time.

The optimal satellite and receiver constellation is shown in Fig. 11 (see in colours).

GPS receivers are used for navigation, positioning, time dissemination, and other research. Navigation receivers are made for aircraft, ships, ground vehicles, and for hand carrying by individuals. The usage of the GPS system becomes popular all over the world. The portable systems used for human navigation. In the Fig. 12 (see in colours) is shown a pocket Magellan navigator instrument.

3.4. The History of GPS

The Global Positioning System, or GPS, was born as a result of the problems experienced by the US military forces during the Vietnam conflict. One of the main difficulties for the troops on the ground was how to keep in contact with each other, especially due to the harsh jungle terrain. A localized LORAN system was in use, but this was subject to the errors common to all radio systems, such as ground wave deflection and poor radio reception at night and in bad weather.

The US then experimented with a system of 4 satellites, initially named TRANSIT. These were in high orbit above the earth and available to marine users as well as the military. However, the system was largely inaccurate, as position fixes could only be obtained every 2 hours at best.

The NavStar system was developed next and was operational in a limited way from 1986, but there was only 3-4 hours coverage per day due to the small number of satellites in orbit. (The main reason for the delay in putting additional satellites into orbit was due to the Challenger Space shuttle disaster of 1988, as the shuttle was the main launch vehicle for the satellites.)

The GPS system became "partially operational" when hostilities began in the Gulf in 1990. Here, experimental Block 1 satellites were used in addition to the established Block 2 satellites, thus giving a useable constellation of 21 satellites. The US Defense Department made the system operational for civilian users in 1990, which is the same GPS system we enjoy today [7].

3.5. The Working of GPS

The GPS satellites orbit the earth twice a day, 11,000 miles above the earth, transmitting their precise position and elevation. The GPS receiver acquires the signal, then measures the interval between transmission and receipt of the signal to determine the distance between the receiver and the satellite. Once the receiver has calculated this data for at least 3 satellites, its location on the earth's surface can be determined.

Every satellite transmits ALMANAC and EPHEMERIS data. Almanac data is general information on the location and the health of each satellite in the constellation, which can be received from any satellite. A receiver with a

current almanac in its memory knows where in the sky to look for satellites, given its last known position and the time of day. Ephemeris data is the precise satellite positioning information that is used by the GPS receiver to compute its position. Each satellite transmits its own ephemeris data.

3.6. GPS Signal Transmission

There are 2 distinct signal types emitted from the satellites; CA (Coarse Acquisition) and PPS (Precise Positioning System). CA coded signals can give 15 meter RMS (Root Mean Square) accuracy. However, the US Defense Department has introduced a random error into the system, known as Selective Availability. This means that the satellites will randomly give out an error signal, thus degrading the accuracy of the signals to 100 meters officially, although accuracy is usually 50 meters. PPS is only available to licensed, mainly military, users and can give sub-1 meter accuracy.

3.7. The GPS Signals

The satellites transmit on two L-band frequencies: $L1 = 1575.42$ MHz and $L2 = 1227.6$ MHz. Three pseudo-random noise (PRN) ranging codes are in use. The coarse/acquisition (C/A) code has a 1.023 MHz chip rate, a period of one millisecond (MS) and is used primarily to acquire the P-code. The precision (P) code has a 10.23 MHz rate, a period of seven days and is the principle navigation ranging code. The Y-code is used in place of the P-code whenever the anti-spoofing (A-S) mode of operation is activated. The C/A code is available on the L1 frequency and the P-code are available on both L1 and L2. The various satellites all transmit on the same frequencies, L1 and L2, but with individual code assignments. Due to the spread spectrum characteristic of the signals, the system provides a large margin of resistance to interference. Each satellite transmits a navigation message containing its orbital elements, clock behavior, system time and status messages. In addition, an almanac is also provided which gives the approximate data for each active satellite. This allows the user set to find all satellites once the first has been acquired.

The Fig. 13 (see in colours) shows the signal stream of the GPS system.

3.8. Selective Availability, Anti-Spoofing

Selective Availability (SA), the denial of full accuracy, is accomplished by manipulating navigation message orbit data (epsilon) and/or satellite clock frequency (dither). Anti-spoofing (A-S) guards against fake transmissions of satellite data by encrypting the P-code to form the Y-code. SA was activated July 4, 1991 at 0400 UT. Presently, due to the fact that SA has not been able to fully achieve its aim since the use of DGPS became widespread, it will most likely be disabled by the year 2000, in order to reduce costs.

3.9. Advanced Forms of GPS

The quest for greater and greater accuracy has spawned an assortment of variations on basic GPS technology.

One technique, called "Differential GPS, (DGPS)" involves the use of two ground-based receivers. One monitors variations in the GPS signal and communicates those variations to the other receiver. The second receiver can then correct its calculations for better accuracy.

The DGPS is a system, which aims to correct the random signal errors caused by Selective Availability. A series of land-based beacons transmit exact position information to an optional radio beacon receiver attached to the GPS receiver, thus enabling the receiver to give a position fix accurate to less than 15 meters.

Another technique called "Carrier-phase GPS" takes advantage of the GPS signal's carrier signal to improve accuracy. The carrier frequency is much higher than the GPS signal that means it can be used for more precise timing measurements.

The aviation industry is developing a type of GPS called "Augmented GPS" which involves the use of a geo-stationary satellite as a relay station for the transmission of differential corrections and GPS satellite status information. These corrections are necessary if GPS is to be used for instrument landings. The geo-stationary satellite would provide corrections across an entire continent.

4. OVERVIEW OF THE GPS RECEIVER SENSOR BOARDS

Having dealt with an introduction to GPS in the former chapter, this one will go further to introduce the GPS receivers produced by different companies, and provide information on how they work and how can be interfaced with PC-s.

The most type of GPS sensors designed for a broad spectrum of OEM (Original Equipment Manufacturer) system applications. The receivers simultaneously track up to twelve satellites providing fast time-to-first-fix, one second navigation updates and low power consumption. Their far-reaching capability meets the sensitivity requirements of land navigation as well as the dynamics requirements of high performance aircraft.

The sensor boards design utilizes the latest surface mount technology (SMD) as well as high level circuit integration to achieve superior performance while minimizing space and power requirements. This HW capability combined with SW intelligence makes the board set easy to integrate and use.

A minimum GPS system must provide the sensor board with conditioned input power and RF signal. The host system may communicate with the board via a choice of two serial bi-directional channels. A highly accurate one pulse per second (PPS) output can be utilized in applications requiring precise timing measurements.

In Fig. 14. is shown the logical connection diagram of the GPS receiver board.

An on board memory allows the sensor board to retain critical data such as satellite orbital parameters, last position, date and time. Non volatile memory is also used to retain board configuration settings even if backup battery power fails. The user interfaces such as display and keyboard are added by the application designer.

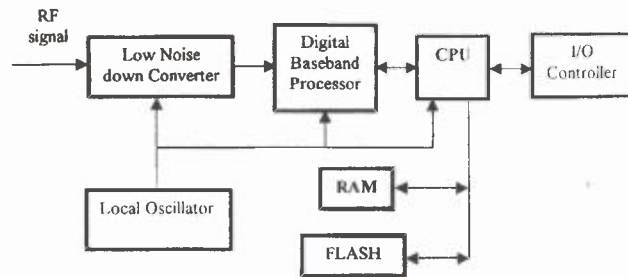


Fig. 14. The block diagram of the receiver board

The typical interface circuit diagram to the host is shown in Fig. 15.

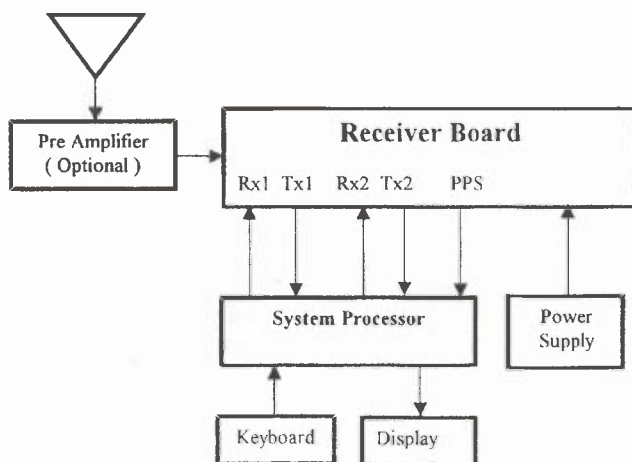


Fig. 15. Connection diagram to the host system

4.1. The Functional Modes of GPS Receivers

GPS receivers' function in four major different modes amongst others and these depend mainly on the actual state of the receiver. These four major modes are:

- Self-test,
- Initialization,
- Navigation and
- Data collection.

These modes of operation will now be briefly introduced using Garmin's GPS receiver (GPS 25LP sensor board).

4.1.1. Self Test

After input power has been applied to the GPS 25LP sensor boards and periodically thereafter, the units perform critical self-test functions and report the results over the output channel(s). The following tests will be performed:

- 1) RAM check
- 2) FLASH test
- 3) Receiver test
- 4) Real-time clock test
- 5) Oscillator check

In addition to the results of the above tests, the board sets reports software version information.

4.1.2. Initialization

After the initial self-test is complete, the GPS 25LP will begin the process of satellite acquisition and tracking. The acquisition process is fully automatic and, under

normal circumstances, will take approximately 45 seconds to achieve a position fix (15 seconds if ephemeris data is known). After a position fix has been calculated, valid position, velocity and time information will be transmitted over the output channel(s).

Like all GPS receivers, the GPS 25LP utilizes initial data such as last stored position, date and time as well as satellite orbital data to achieve maximum acquisition performance. If significant inaccuracy exists in the initial data, or if the orbital data is obsolete, it may take 1.5 minutes to achieve a navigation solution. The GPS 25LP AutoLocate™ feature is capable of automatically determining a navigation solution without intervention from the host system. However, acquisition performance can be improved if the host system initializes the board set following the occurrence of one or more of the following events:

- 1) Transportation over distances further than 1500 kilometers.
- 2) Failure of the on-board memory battery.
- 3) Stored date/time off by more than 30 minutes.

4.1.3. Navigation

After the acquisition process is complete, the GPS 25LP will begin sending valid navigation information over its output channels. These data include: 1) Latitude/longitude/altitude

- 2) Velocity
- 3) Date/time
- 4) Error estimates
- 5) Satellite and receiver status

The GPS 25LP sensor boards automatically select the optimal navigation mode (2D or 3D) based on available satellites and geometry considerations. When navigating in the 2D mode the board set utilizes the last computed altitude or the last altitude supplied by the host system, whichever is newer. The host system must ensure that the altitude used for 2D navigation is accurate since the resulting position error may be as large as the altitude error.

The GPS 25LP switches to automatic differential mode by default "looking" for real-time differential corrections in RTCM SC-104 standard format using message types 1,2,3 or 9 and attempting to apply them to the satellite data in order to produce a differential (DGPS) solution. The host system, at its option, may also command the board set to choose differential only mode. When navigating in the differential only mode, the board set will output a position only when a differential solution is available.

4.1.4. Satellite Data Collection

The GPS 25LP sensor boards will automatically update satellite orbital data as they operate. The intelligence of the board set combined with its hardware capability allows these data to be collected and stored without intervention from the host system. A few key points should be considered regarding this process:

- 1) If the sensor board is not operated for a period of six (6) months or more, the unit will "search the sky" in order to collect satellite orbital information. This process is fully automatic and, under normal

circumstances, will take 3-4 minutes to achieve a navigation solution. However, the host system should allow the board set to remain on for at least 12.5 minutes after the first satellite is acquired.

- 2) If the memory backup battery fails, the sensor board will search the sky as described above. The system designer should be aware of the availability of standby power input to the board set to prevent this situation.
- 3) If the initial data is significantly inaccurate, the board set will perform an operation known as AutoLocate™. This procedure is fully automatic and, under normal circumstances, will require 1.5 minutes to calculate a navigation solution. AutoLocate™, unlike search the sky, does not require that the sensor board continue operating after a fix has been obtained.

Other functional modes do exist, especially for the receivers produced by other companies. For example, the SiRF protocol receiver also works in a power saving mode called TricklePower mode. Another SiRF mode is the Push-to-Fix mode, wherein the system stays in sleep mode, until awakened by an external request.

5. THE STRUCTURE OF THE DIFFERENT TYPES OF PROTOCOLS

The protocol is a well defined format of the communication between two or more devices. If the programmer sends a message having syntactical mistake the GPS will refuse it. On the other side the user has to understand the message providing the GPS for the correct elaboration.

That is why necessary know the structure and the content of protocols byte by byte.

5.1. Garmin (GPS 25LP) Protocol

One of the most popular GPS receiver is the Garmin produced GPS 25 LP type. This very small, high performance unit, used by a single 5 V power supply. The consumption can not be eliminated, cca 100 mA in normal mode.

The Garmin interface protocol design on Tx1/Rx1 is based on the National Electronic Association's NMEA 0183 ASCII interface specification, which is fully defined in [4].

In the next part we demonstrate some of the most important message sentences which are realises the host-GPS communication. First we discuss the program sentences dedicated to change the configuration and working mode of the GPS receiver. The comma indicates the different fields, while hh denotes the check summa (this field is optional, used for the noise channels only).

5.1.1. Command Sentences

Almanac information (ALM) command

\$GPALM,<1>,<2>,<3>,<4>,<5>,<6>,<7>,<8>,<9>,<10>,<11>,<12>,<13>,<14>,<15>,<hh>,<CR>,<LF>

This command initializes the sensor board (if battery back-up has failed).

- <1> Total number of ALM sentences.
- <2> Number of current ALM sentence.

- <3> Satellite PNR number: 01 to 32.
- <4> GPS week number.
- <5> SV health, bits 17-24 of each almanac page.
- <6> Eccentricity.
- <7> Almanac reference time.
- <8> Inclination angle.
- <9> Rate of right ascension.
- <10> Root of semi major axis.
- <11> Omega, argument of perigee.
- <12> Longitude of ascension node.
- <13> Mean anomaly.
- <14> Af_0 clock parameter.
- <15> Af_1 clock parameter.

Sensor initialization information (PGRMI) command

\$PGRMI,<1>,<2>,<3>,<4>,<5>,<6>,<7>,
hh,<CR>,<LF>

This command initializes the sensor board position and time. If there are no errors in the sentence, it will echo upon receipt.

- <1> Latitude, ddmm.mmm format (leading zeros must be transmitted).
- <2> Latitude hemisphere, N or S.
- <3> Longitude, ddmm.mmm format (leading zeros must be transmitted).
- <4> Longitude hemisphere, E or W.
- <5> Current UTC date, ddmmyy format.
- <6> Current UTC time, hhmmss format.
- <7> Receiver command, A = Auto locate, R = Unit reset.

Sensor configuration information (PGRMC) command

\$PGRMC,<1>,<2>,<3>,<4>,<5>,<6>,<7>,<8>,
<9>,<10>,<11>,<12>,<13>,<14>,hh,<CR>,<LF>

This command configures the sensor board operation. These parameters are stored in the NVM (FLASH memory). If there are no errors in the sentence, it will echo upon receipt.

- <1> Fix mode, A=Automatic, 2 = 2D (two dimensional), 3 = 3D mode.
- <2> Altitude above/below mean sea level, -1800 to 18 000 meters.
- <3> Earth datum index.
- <4> User earth datum semi-major axis (0.001 meter resolution).
- <5> User earth datum inverse flattening factor (10^{-9} resolution).
- <6> User earth datum delta x earth centered coordinate (1 meter resolution).
- <7> User earth datum delta y earth centered coordinate (1 meter resolution).
- <8> User earth datum delta z earth centered coordinate (1 meter resolution).
- <9> Differential mode, A = Automatic, D = Differential.
- <10> NMEA Baud rate, 1=1200, 2=2400, 3=4800, 4=9600, 5=19600, 6=300, 7=600.
- <11> Velocity filter, 0 = No filter, 1 = Automatic filter, 2-255 = Filter time in seconds.
- <12> PPS mode, 1 = No PPS, 2 = 1 Hz.

- <13> PPS pulse length, $0 - 48 = (n + 1) * 20$ msec (e.g. 4 \rightarrow 100 msec).
- <14> Dead reckoning valid time 1-30 sec.

The Baud rate and PPS mode change take effect on the next power cycle or external reset event.

Additional configuration information (PGRMC1) command

\$GPRMC1,<1>,<2>,hh,<CR>,<LF>

This command configures the sensor board operation. These parameters are stored in the NVM. If there are no errors in the sentence, it will echo upon receipt.

- <1> NMEA output time 1-900 sec.
- <2> Binary Phase Output Data, 1 = Off, 2 = On.

Configuration changes take effect on the next power cycle or external reset event.

Output sentence enable/disable (PGRMO) command

\$GPRMO,<1>,<2>hh,<CR>,<LF>

This command provides the ability to enable or disable specific output sentences.

- <1> Target sentence description (e.g. GPGSA, GPGSV, etc).
- <2> Target sentence mode, where:
 - 0 = disable specific sentence
 - 1 = enable specific sentence
 - 2 = disable all output sentences
 - 3 = enable all output sentences (except GPALM).

5.1.2. Received sentences

Second we discuss the most important sentences transmitted by the GPS sensor board.

GPS Fix Data (GGA) sentence

\$GPGGA,<1>,<2>,<3>,<4>,<5>,<6>,<7>,<8>,
<9>,M,<10>,M,<11>,<12>,hh,<CR>,<LF>

- <1> UTC time of position fix, hhmmss format.
- <2> Latitude, ddmm.mmm format (leading zeros will be transmitted).
- <3> Latitude hemisphere N or S.
- <4> Longitude, ddmm.mmm format (leading zeros will be transmitted).
- <5> Longitude hemisphere E or W.
- <6> GPS quality information, 0 = Fix not available, 1 = Non DGPS fixed, 2 = DGPS fix available.
- <7> Number of satellites in use 00-12.
- <8> Horizontal dilution of precision, 0.5 to 99.9.
- <9> Antenna height above/below mean sea level (-9999 to 99999 meters).
- <10> Geoidal height -999.9 - 9999.9 meters.
- <11> DGPS data age, number of seconds since last valid transmission.
- <12> Differential Reference Station ID, 0000 to 1023.

Recommended Minimum Specific GPS Transmit Data (RMC) sentence

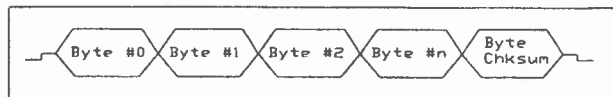
\$GPRMC,<1>,<2>,<3>,<4>,<5>,<6>,<7>,<8>,
<9>,<10>,<11>,hh,<CR>,<LF>

- <1> UTC time of position fix, hhmmss format.
- <2> Status, A = Valid, V = NAV receiver warning.

- <3> Latitude ddm.mmm format (leading zeros will be transmitted).
- <4> Latitude hemisphere N or S.
- <5> Longitude, ddm.mmm format (leading zeros will be transmitted).
- <6> Longitude hemisphere E or W.
- <7> Speed over ground, 0.0 to 999.9 knots.
- <8> Course over ground, 000.0 to 359.9 degrees (leading zeros transmitted).
- <9> UTC date of position fix, ddmmyy format.
- <10> Magnetic variation 000.0 to 180.0 degrees (leading zeros transmitted).
- <11> Magnetic variation direction, E or W (westerly variation adds to true course).



Header Format



Data Format

Fig. 16. The Header and Data format of Lowrance protocol

Table 1. The LSI 100 commands

Command	Description
0x0008	Read from a memory location
0x0009	Write to a memory location
0x000D	Wake up external serial port
0x010A	Change NMEA Baud Rate
0x0301	Request Screen Pointer do download the current Screen
0x0302	Unfreeze the current Screen (happens after downloading Screen)
0x0303	Get a Waypoint
0x0304	Send a Waypoint
0x0305	Get a Route
0x0306	Send a Route
0x0307	Get plot trail Pointer
0x0308	Get number of Icons
0x0309	Get Icon Symbol
0x030A	Set number of Icons
0x030B	Send Icon
0x030C	Get number of Graphical symbols
0x030D	Get Icon Graphical symbol
0x030E	Get product information
0x0312	Get Plot Trail Origin (Protocol version 2.0)
0x0313	Get Plot Trail Details (Protocol version 2.0)
0X0314	Set Plot Trail Origin (Protocol version 2.0)
0x0315	Set Plot Trail Details (Protocol version 2.0)

Track Made Good and Ground Speed (VTG) sentence

\$GPVTG,<1>,T,<2>,M,<3>,N,<4>,K,hh,<CR><LF>

- <1> True Course over ground, 000 to 359 (leading zeros transmitted).
- <2> Magnetic course over ground, 000 to 359 (leading zeros transmitted).
- <3> Speed over ground, 0.0 to 999.9 knots.
- <4> Speed over ground, 0.0 to 1851.8 km/hour.

5.2 Lowrance (LSI 100) Protocol

Some receiver uses the standardized LSI 100 protocol. The LSI 100 was designed to be a byte protocol, thus all transferred data will always be in bytes. It is made up of 2 sections, the "Header" and the "Data". The Header is made up of eight bytes that tell the unit/PC system exactly what the interfacing device wants. The header is made up of five parts.

The first part is called the preamble, is 2 bytes long and is used to signal the start of an LSI 100 message. The second part is the "Message Command" which is a word that indicates a specific message requested or sent. Bit #7 within the message command word is the acknowledge bit. The next part, the "Data Count" is a word length data byte count, which shows the number of bytes of information that will be transmitted in the "Data" structure. If the data count is set to zero, then only the header information will be transmitted, without data.

The next field is a reserved field that is always set to zero. The last part is the byte field called the "Header checksum". The checksum is calculated by summing up all the header's bytes (with the exception of the checksum byte) and taking the 2's compliment of that number. If the checksum is computed correctly, then the total summation of all header bytes should always be zero.

Fig. 16. shows the header and data format of the Lowrance protocol.

For short information in Table 1 we demonstrate the LSI 100 commands.

Finally let's take a look at an example for both header and data streams. The following section shows what the host system transmitted to the GPS unit, and what the unit sends back. The command opens a session. The GPS responds back with bit 7 in the message command. All data are given in hexa format.

Host TxD → Header: 8155 Cmd=000D
 Cnt=0000 Id=00 Chksum=1D
 GPS resp. ← Header: 8155 Cmd=008D
 Cnt=0001 Id=00 Chksum=9C
 C6 3A (sessin ID + chksum)

Looking at the data stream that followed the received header above, we can see that the unit transmitted a C6 and the chechsum 3A.

On the serial line we can monitor the next byte format:
 Host TxD → 55 81 0D 00 00 00 00 1D
 GPS resp. ← 55 81 8D 00 01 00 00 9D C6 3A.

The detail subscription of this protocol can found in reference [6].

5.3. SiRF (μ -blox GPS-MS1 and GPS-PS1) Protocol

This is the default protocol used by both members of the SiRF family (μ -blox GPS-PS1 (82.5mm x 32mm x 8.5mm) and μ -blox GPS-MS1 (30.2mm x 29.5mm x 7.55mm), with

the later being one of the smallest receivers ever produced. NMEA is also used, enabling the possibility of switching protocols.

5.3.1. Transport Message

The transport layer of the protocol encapsulates a GPS message in two start characters and two stop characters.

The values are chosen to be easily identifiable and such that they are unlikely to occur frequently in the data. In addition, the transport layer prefixes the message with a two-byte (15-bit) message length and a two-byte (15-bit) check sum. The values of the start and stop characters and the choice of a 15-bit values for length and check sum are designed such that both message length and check sum can not alias with either the stop or start code.

5.3.2. Message Validation

The validation layer is of part of the transport, but operates independently. The byte count refers to the payload byte length. Likewise, the check sum is a sum on the payload.

5.3.3. Payload Length

The payload length is transmitted high order byte first followed by the low byte. This is the so-called big-endian order:

High Byte	Low Byte
0x00...0x7F	0x00...0xFF

Even though the protocol has a maximum length of $(2^{15} - 1)$ bytes practical considerations require the SiRF GPS module implementation to limit this value to a smaller number. Likewise, the SiRF receiving programs (e.g., SiRFdemo) may limit the actual size to something less than this maximum.

5.3.4. Payload Data

The payload data follows the payload length. It contains the number of bytes specified by the payload length. The payload data may contain any 8-bit value. Where multi-byte values are in the payload data neither the alignment nor the byte order are defined as part of the transport although SiRF payloads will use the big-endian order. It should be noted that some processors do not allow arbitrary byte alignment of multi-byte data and therefore care should be used when reading data delivered as payload data.

5.3.5. Checksum

The check sum is transmitted high order byte first followed by the low byte (the former so-called big-endian order). The check sum is 15-bit checksum of the bytes in the payload data.

5.4. Star (OmniSTAR – 3000LCE) Protocol

The interface protocols of GPS receivers are usually designed to comply with the National Marine Electronic Association's NMEA 0183 ASCII interface specification. NMEA 183 is a simple, yet comprehensive ASCII protocol,

which defines both the communication interface and the data format.

NMEA 0183 allows a single source (talker) to transmit serial data over a single twisted wire pair to one or more receivers (listeners). The NMEA 0183 protocol covers a broad array of navigational data. This is separated into discrete messages, which convey a specific set of information. The NMEA 0183 message structure is outlined on the next page.

$\$IDMSG,D1,D2,D3,D4,...,Dn*CS[CR][LF]$

"\$" The "\$" signifies the start of message.

ID The Talker identification is a two letter mnemonic which describes the source of the navigation information. The GP identification signifies a GPS source.

MSG The message identification is a three letter mnemonic, which describes the message content and the number and order of the data fields.

"," Commas serve as delimiters for the data fields.

Dn Each message contains multiple data fields (Dn), which are delimited by commas.

"*" The asterisk serves as a checksum delimiter.

CS The checksum field contains two ASCII characters, which indicate the hexadecimal value of the checksum.

[CR] The carriage return [CR] and line feed [LF]

[LF] combination terminates the message.

NMEA messages vary in length, but each message is limited to 79 characters or less. This length limitation excludes the "\$" and the [CR][LF]. The data field block, including delimiters, is limited to 74 characters or less.

5.5. Marconi Rt-Star Serial Protocol

This protocol used by the GPS products of firm Marconi [8]. The description of this protocol can be found detail in the document of User's Manual P/N 220-600944-40X in the Internet WEB Site [9].

6. THE SOFTWARE OF THE PDR

In this chapter we overview briefly the SW running on the PDR data acquisition system.

6.1. The Logical Structure of the SW

We demonstrate the tasks and their relations to each other by a logical connection diagram shown in Fig. 17.

Since the PDR works in event controlled mode, in the most time spends in an so called *idle* state. In this state the receiver needs a reduced power consumption. If an event happens the processor wakes up, and starts his work. This event made by three causes. The higher priority event the command arrives from the host. In this case the user realizes a serial connection to the PDR. In this case the PDR is waiting command from the user SW (see in the chapter 7), running on a PC computer. In this situation the PDR SW receives one of the possible commands from it.

In Fig. 18. show the allowed commands and the response of the PDR unit.

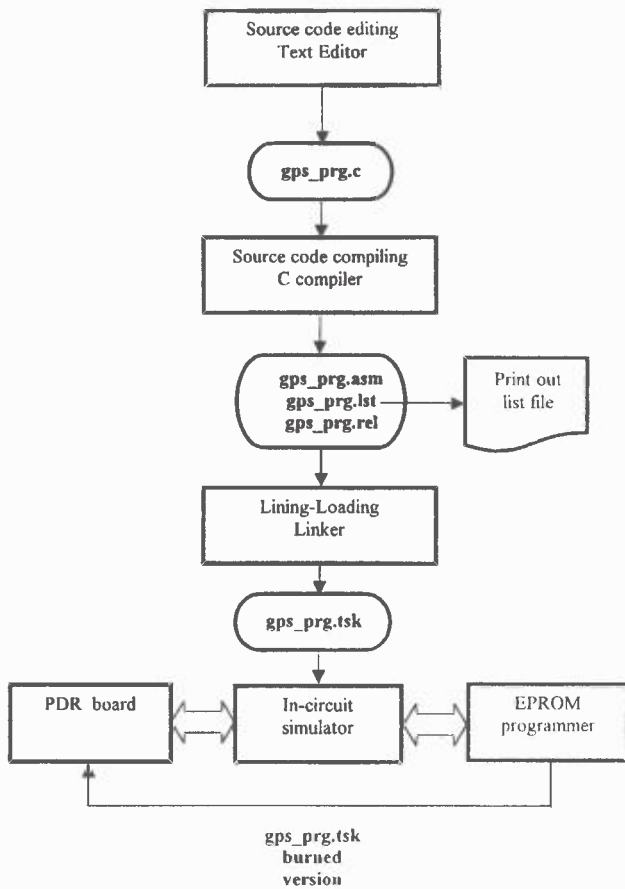


Fig. 17. The flow chart of the PDR SW

We can see, that there are altogether four possible command:

- Get PDR internal memory position, i.e. the host want to know how many empty memory room is available. The answer is a pointer value, pointing to the first empty place.
- Get PDR collected data. The response is a pre-elaborated, condensed data stream, having all position, time and speed coordinates recorded since the previous clear operation. This stream can be very long up to 500,000 messages.
- Clear command for clearing the whole PDR internal memory content. After executing this command all data will lose.
- Disconnect command cuts the communication with the host, and the PDR turns back to its internal mode (i.e. idle state).

This communication is solved by a special protocol, applying hand-shake signs, time out, ACK (Acknowledge), NAK (Not acknowledge), Fatal transmit Error and time-out sequences for the shake of error free data communication.

The GPS data collection includes the next tasks:

If the GPS sends message (according to the previously programmed time interval) the PDR wakes up from its idle state. First stores the whole arrived message in a temporarily buffer. Second analyses the message, checks it syntactically (checks the parity and checksum values). If the message formally correct, cuts it into his fields, selects the different information, i.e. date, time, longitude,

latitude, speed coordinates. After compress them using a special binary codification and stores it to the next free record in the internal non volatile memory place, and obviously increments the pointer value.

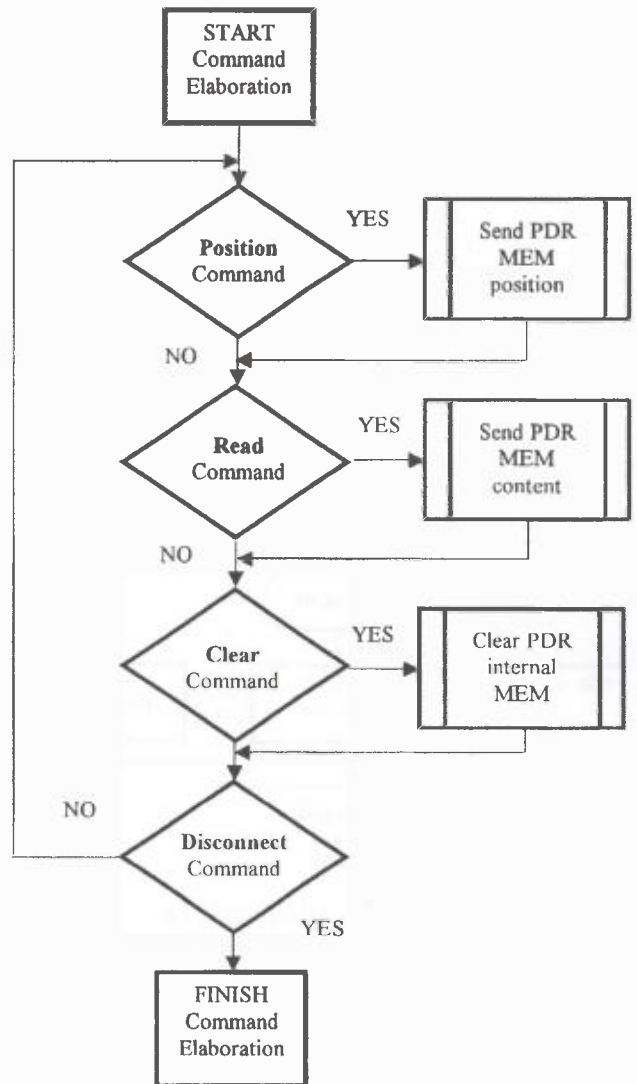


Fig. 18. The flow chart of the command elaboration

The third statement from host is the Clear memory command. After receiving this command the PDR SW clears its internal storage memory, i.e. fills it a pre defined byte pattern (0xFF).

The fourth and last possible statement is the Disconnect command. The effect of this command is clear, sign to the PDR, that the communication with the host has finished, and the PDR can revert his stand-by (idle) state.

6.2. The SW Developing Environment

Since the PDR apply an INTEL 80C31 single-chip microcomputer, the SW consists of its machine code. This code can be translated from the high level C language using the IAR-ICC8051 integrated SW developing environment [10]. The editor, compiler and loader run in a PC compatible desk-top computer. The executable binary form is loaded to an in-circuit on-line PROM simulator for testing. If the test procedure has already finished, the

executable code is burn into an EPROM chip, and placed in the PDR board.

The developing process is demonstrated in the Fig. 19.

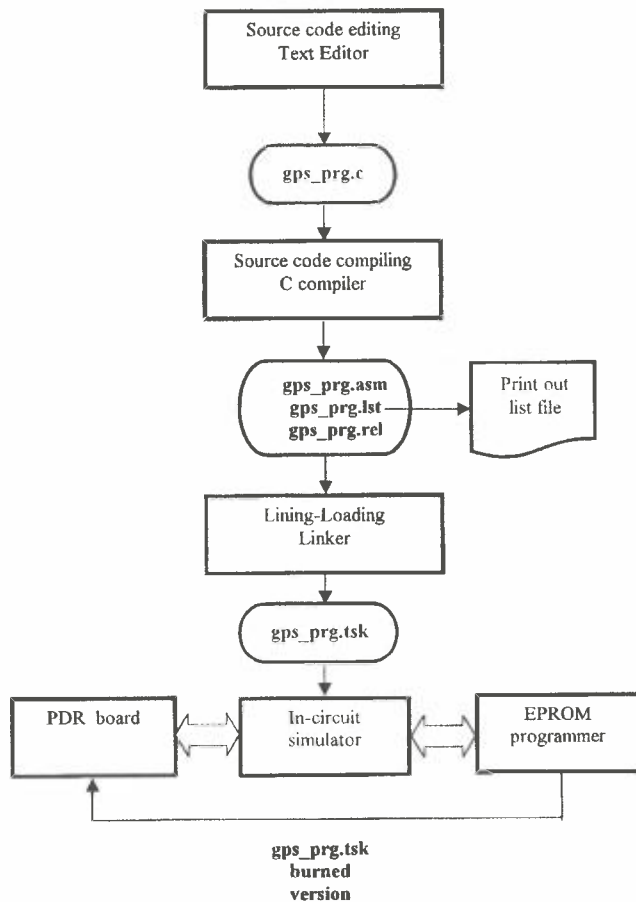


Fig. 19. The SW developing process

7. DATA TRANSMISSION SYSTEM

If a message is sent from one point to another, we must always count with the data corruption. This is due to the noise, interference, cable error etc. Generally we are not able prevent the corruption, but there is possible to detect them. After covering the transmission error, we can eliminate its influence by correction of the message, or repeat the transmission.

The basis of the safety is the high level error detection. It is not possible to prevent the appearing of errors, but the system must cover them. We have to choose a suitable coding-decoding system which able to detect any random, systematically, single or burst corruption which appear due to noise, line error, distortion or cross talk.

At the same time have to be regard to the requirements of transmission capability, transparent and bandwidth.

These requirements are opposite to each other, but the planner should find the optimal solution.

In Fig. 20. we can see the theoretical scheme of the data protection.

The wide-spreading tool for the error detection is application of data redundancy. More or less bits will append to the binary information to help in the error detection. The usage of the applied redundancy can by classified into three main groups:

- Systematical data repetition
- Convolutional codification
- Block codification

The systematical data repetition method applies in the level of codification well definite number (2 or 3) data repetition. In Fig. 21. are shown some possible methods supposing that the number of repetition equal three.

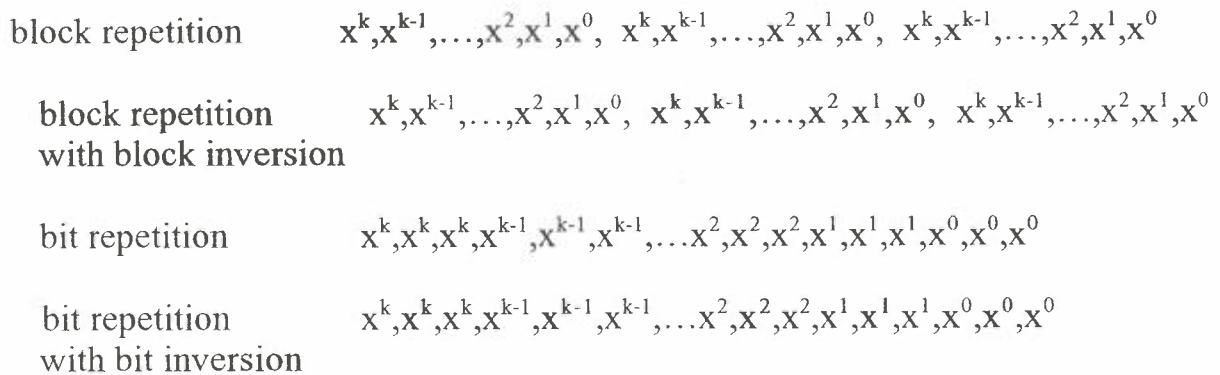


Fig. 20. The scheme of data protection

In the Fig. 21. the bold characters denote the inverted data bits. This technique is economical if the channel not too noisy, the corruption is not systematical, and the transfer rate is not an important parameter. In other words this method does not provide the necessary level of protection if the burst and systematical errors can be supposed.

The convolutional codification has the largest literature, started by the publication of Wozencraft thought Viterbi till Forney. These algorithm operate with the inserting of

few redundant bits, therefore the length of the information block is limited. The redundant bits are generated from some neighbor's information bits (generally 2...5) and the original bit pattern is replaced by this modified pattern. The error detection capability of this codification strongly depends on the applied algorithm. The random, the systematical, the single and burst errors need different algorithms. That is why we do not discuss this method further.

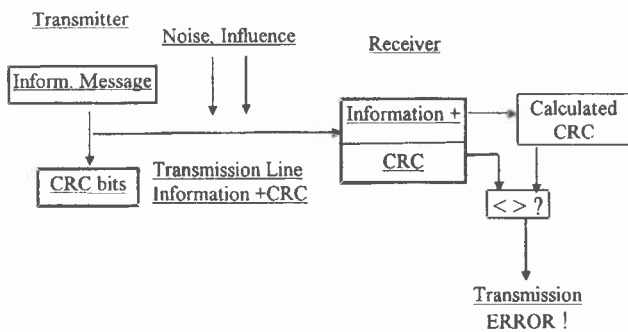


Fig. 21. Some repetition methods

The block codification method have been in use in communication systems for many years. In most cases the block coding is applied in conjunction with retransmission. The rule of the codification is the error detection, and the correction should be solved by the protocol applying the retransmission.

There are many block codification technique, but all of them apply CRC (Cyclic Redundancy Code) bits generated from the information message by a well definite algorithm, and send toward the receiver together with the information. In the receiver side using the definite algorithm the message will decoded, and the intact of information tested. If any violation is detected, error message is sent toward the higher level (the protocol). The protocol makes the administration of the error, and arranges for the correction of the failure (for example sends back special message sequence: NACK). The method is detail explained in the [11].

For the practice the necessary protection level is provided by using the check-summa control. The check-summa is an unsigned integer in two bytes of length, which derived from the sum of the bytes of the given information block. The 2-s compliment of this sum will append to the message. On the receiver side also is made this sum, and adding to the arrived check-sum have to get zero result. If this sum will not equal 0, the message will regard corrupted, and have to repeat the transmission.

Let's see the next example:

Supposing the message have to transmit is: 27 3A D5 F9,
 The calculated check-summa will: 02 2F
 The compliment of the summa: FF D1
 The redundant message: 27 3A D5 F9 FF D1
 On receipt checking: 02 2F + FF D1 = 00 00

At the check summa calculation the carry will not take account.

8. ELABORATION OF THE DATA

The collected data can use by the user reading it via a serial transmission line. For this elaboration enough to apply a PC desk-top or laptop computer. The PC serial communication port will connect to the PDR second serial port (see Fig. 1.).

The SW, controlling the PC-PDR communication has two parts. The first builds up the physical connection, tests the line correctness, and send pre-defined commands toward the PDR. These commands were detail discussed

in the chapter 6.1. This program runs under Windows operation system providing a comfortable man-machine interface for the user. The Fig. 22 (see in colours) shows the screen appearing after starting this COMMWE.XE program.

On the top left side appears the program handling window offering the choose one of the available three commands: Position, Read, Clear. The bottom right window shows the answer if previously the Position command was selected. The user is informed about the PDR whole capacity, the used and free territory. The pie chart demonstrates the percent of reservation. In this case the PDR memory is almost empty.

If the user executes the Read command, the program ask a filename using an input window (see Fig. 23 in colours), and automatically transfers the whole memory content and save into the definite file. A small scroll bar follows the transmission.

If the data transmission was successful, the user can decide whether he Clears the PDR memory or not. If yes, all data in the PDR will loss. If not, after disconnection the PDR continues its data acquisition.

Having the data file, the user can start running the elaboration program, called GPSROUTE.EXE. In the Fig. 24 (see in colours) is shown the screen, appearing after the start.

There are three fields on the screen. The group of windows on the left side serves for the map and input data file selection, there are command keys (Load Map, Read Data File, Day Selection, Calculate length of road section, Select Color and Draw Route).

In the middle part we can see the selected map, in this case the map shows Hungary. The selected route appears with a line with selected color.

The window on the right side shows the decoded data file, providing the validity, date, time, longitude, latitude information. The most right column gives a calculated value, which gives the distance in meters between the previous and current row. On the right bottom the user gets the summa distance relating to the selected day, or any group of selection.

These programs was written in Delphi language, and provides a comfortable (mouse and keyboard handled) and nice interface for the user.

9. APPLICATION EXAMPLES

The instrument discussed above is a portable, small unit needs neither handling nor maintenance procedure, is shown in Fig. 25 (see in colours).

The internal panel without the GPS receiver board is shown in Fig. 26 (see in colours).

The PDR currently is used for car tracking, mainly at small firms having no more than 30 number of cars. According to the applied technology, the car drivers make way-bill, and finishing their trip among the country, turn back to his firm. A dedicated worker reads the PDR data (without removing it from the car) using a portable lap-top computer. In the office his way-bill will compare with the PDR data, and by this process will their trip verified. The allowed difference is less than 10 %.

There are many field for application of this technology. In the Lorry, Bus, Taxi, Tram vehicles can be tracked by the system. Therefore it can be interesting first of all for transportation firms, but also can used in the field of

service (gas company, electricity company, quick mobile service company, rent a car company etc.), too.

Since the system very cheap, it will soon produce its investigation.

REFERENCES

- [1] Dr. Gyenes Károly: A CRC blokk kódolás hiba analízise számítógépes szimulációval (Error Analising of CRC Block Code with Computer Simulation), Vezetékek Világa 98/1 sz.
- [2] Transmission of Safety Information ORE Red Books 155/RP A 2,8,10,13, 1987.
- [3] The Elaboration of the Safety Informations (A biztonsági információk feldolgozása és átvitele), UIC 738 R Recommendation 1992.
- [4] NMEA 0183, Version 2.0 protocol definition, P.O. Box. 50040, Mobile, Al. 36605. USA.
- [5] RTCM Recommended Standards for Differential Navstar GPS Service, Version 2.1. RTCM Special Committee No. 104.
- [6] Lowrance Serial Interface Protocol LS100. <http://www.lowrance.com/software>
- [7] M. Szilvási, T. Szőke, R. Hargitai: Tracking and Control of Moving Objects, Híradástechnika 1998/1-2. pp. 55-56.
- [8] Omni Spectra, 100. Chelmsford St. P.O. Box 3295. Lowell, MA 01853-9910.
- [9] WEB Site: <http://www.marconi.ca>
- [10] IAR C Compiler & Microprocessor Development Tool. User's Manual IAR Systems AB. IPM-00042E 1989.
- [11] Dr. K. Gyenes: Remote Controlled Traffic for Small Railway Stations, Periodical Polytechnic Vol.26.PP. 89-100. 1998.
- [12] GPS Utility: <http://www.gpstoday.com/utilities.htm>
- [13] GPS Education Material: <http://www.host.cc.utexas.edu/ftp/pub/gps.html>
- [14] GPS Manufacturers, suppliers: <http://www.thegpsstore.com/site/>

INTELLIGENS MŰHOLDAS JÁRMŰKÖVETŐ RENDSZER

GYENES KÁROLY

BUDAPESTI MŰSZAKI ÉS GAZDASÁGTUDOMÁNYI EGYETEM

A cikk bemutat egy korszerű mikroszámítógépes adatrögzítő rendszert, amelynek célja járművek követése. Számos hivatalos és magán cég rendelkezik járművekkel, amelyek valóságos mozgását szeretné követni. Az itt bemutatott rendszer nem valós idejű, hanem a jármű mozgásának utólagos követését teszi lehetővé. Integrált vállalatirányító rendszerekbe a menetlevél feldolgozás és tényleges menet teljesítmény számítását segíti elő. Nagyon jól használható a járművezetők aktivitásának ellenőrzésére is. Ugyancsak felhasználható az esetleges sebességtűlépés kiszűrésére is.

A kisméretű, igénytelen és olcsó berendezés a jelenlegi benzinárak figyelembe vételével rövid időn belül megtermeli saját beszerzési költségét. A berendezés off-line működési módja következtében nem igényli semmilyen engedély beszerzését ellentétben a rádiós rendszerekkel, nem kell telefonos hálózat működéséhez, végül nem igényel diszpécser központot, ahol figyelő személyzet állandó szolgálatban követi egy jármű mozgását. A berendezés egy GPS műholdvevő által szolgáltatott idő, hely és sebesség koordinátákat tárolja saját memóriájában. Ez a memória kialakításánál fogva nem veszti el a tárolt adatokat. A rendszer programozható ismétlési idő alapján tárolja a koordinátákat. Ha a gyakorlatban jól bevált 2 perces ismétlődési idővel számolunk, akkor ez egy jármű több mint 2 heti mozgását teszi lehetővé.

A tárolt adatok a jármű telephelyre való megérkezésekor egy kommersziális PC segítségével olvashatók ki. A kiolvasott adatok file-ba kerülnek rögzítésre. Ezután egy Windows operációs rendszer alatt futó feldolgozó program olvassa fel az előbbi adat file-t. Ez a feldolgozó program a felhasználó igényei szerinti szolgáltatásokat nyújt. Alapfunkciói: adatfile olvasása, térkép kiválasztása és betöltése, adatfile rendezett listázása, adatbejegyzések kijelölése, a kijelölt adatok összes menet teljesítményének számítása, grafikon szín kijelölése, a megjelölt adatok alapján az útvonal térképre rajzolása. Ezekon kívül integrált vállalatirányító rendszer esetén a kijelölt adatok menetlevéllel való összevetése, egyes járművezetők menet teljesítményének kalkulálása, sebességellenőrzés, vezetés idejének ellenőrzése stb. is lehetséges.

A cikk részletesen ismerteti a GPS rendszer működésének alapjait, az egyes gyártók készülékeinek jellemzőit. Megismerteti a GPS vevő programozásának módját néhány közismert protokoll bemutatása által. Kitér az adatátvitel legfontosabb kérdéseire, ismerteti a PDR HW és SW komponenseit. Végül a javasolt felhasználási területéről számol be, és néhány jellemző irodalmi hivatkozást közöl.



Károly Gyenes engineer electric, has graduated at the Technical University of Budapest in the year of 1968. Five years later he has finished his postgraduate course and got his second diploma of Electronic Integrated Circuits. In 1976 absolved his thesis of doctor's degree. He works at the Technical University of Budapest, at the Faculty of Transportation Engineering at the Department of Transport Automation as associate professor. He is lecturer of Computing, and Computer Hardware. His investigation field is the computerised automation (mainly microcomputer controlled) and control systems for transportation. Till now he published 2 technical books, 5 lecture notes, 4 foreign, and 10 Hungarian language papers, and a lot of conference lectures.



A kezében van a helyzet megoldása

Az új Nokia TETRA professzionális mobil rádió megsokszorozza a gyorsaság, a higgadság, a döntésképeség, a csapatmunka erejét.

A Nokia TETRA rendelkezik az Európai Távközlési Szabványügyi Intézet (ETSI) által specifikált új digitális technológia minden lényeges elemével. Számos helyzetben jó szolgálatot tesz a rövid hívásfelépülési idő, a hang- és adatátviteli lehetőség, a megosztható felhasználás, a hatékony frekvencia-kihasználás, a prioritási szintek beállításának lehetősége.

A Nokia a teljes rendszerek szállítása és a rádiós berendezések területén szerzett rendkívüli tapasztalatával és nemzetközi ügyfélszolgálati hálózatával segíti az Ön munkáját.



Nokia TETRA. M megbízható technológia egy megbízható cégtől.



NOKIA
CONNECTING PEOPLE



journal on

communications

computers

convergence

contents and distribution

companies

5