



Hálózati Rendszerek és
Szolgáltatások Tanszék

Mesterpróba 2014

Tudományos konferencia végzős MSc és elsőéves PhD hallgatóknak

Távközlés és infokommunikáció témakörében

KONFERENCIA KIADVÁNY

Budapesti Műszaki és Gazdaságtudományi Egyetem

Villamosmérnöki és Informatikai Kar

BUDAPEST

2014. május 29.



Mesterpróba 2013

Tudományos konferencia végzős MSc és elsőéves PhD hallgatóknak

Távközlés és infokommunikáció témakörében

A Budapesti Műszaki és Gazdaságtudományi Egyetem Villamosmérnöki és Informatikai Kara által rendezett hallgatói konferencia, a Hírközlési és Informatikai Tudományos Egyesület tudományos bizottsága, az MTA Távközlési Tudományos Bizottsága, az IEEE Hungarian AP/ED/MTT/ComSoc/EMC chapter támogatásával.

A konferencia célja kettős. Egyrészt a legjobb végzős hallgatók számára fórumot biztosít, ahol a diplomamunkájuk készítése során elért kiemelkedő eredményeiket be tudják mutatni a szélesebb szakmai közönségnek. Másrészt támogatja a tudományos indíttatású hallgatók első konferencia részvételét. A konferencia tutorial előadásai és a bírálat menete hozzásegíti a hallgatókat a tudományos konferenciák sok esetben formai, technikai - de ezzel együtt tartalmi, minőségi kérdéseket is meghatározó – elvárásainak megfelelő szintű teljesítéséhez.

A benyújtott cikkek többfordulós bírálati folyamat minősíti, ennek eredménye dönt az elfogadásról (peer to peer review). A folyamat során a szerzők írásos visszajelzést kapnak a bírálóktól, a cikk elfogadását és a konferencián való részvételt hivatalos nyilatkozat igazolja.

Mesterpróba 2014

A KONFERENCIA TUDOMÁNYOS ÉS SZERVEZŐ BIZOTTSÁGA

Társelnökök

Gerhátné Dr. Udvary Eszter (IEEE MTT/AP/ED/ComSoc/EMC Chapter, BME-HVT)
Prof. Pap László (MTA TTB, HTE, BME-HIT)

Tagok

Győry Erzsébet (BME-TMIT)
Huszák Árpád (BME-HIT)
Prof. Pávó József (Periodica Polytechnica, BME-HVT)
Fekete Gábor (BME-HVT)

TÁMOGATÓK

Budapesti Műszaki és Gazdaságtudományi Egyetem Villamosmérnöki és Informatikai Kar
Hírközlési és Informatikai Tudományos Egyesület
MTA Távközlési Tudományos Bizottsága
IEEE Hungarian Joint AP/ED/MTT/ComSoc/EMC chapter
BME – Hálózati Rendszerek és Szolgáltatások Tanszék
BME – Szélessávú Hírközlés és Villamosságtan Tanszék

Program

TUTORIAL 1.

Berceli Tibor: Néhány tanács cikkíráshoz

IDŐPONT: 2014. március.27. 16:00

HELYSZÍN: BME V1 épület, 5. emelet, 519-es terem
Budapest, 1111, Egly József utca 18.

TUTORIAL 2.

Elek Nikoletta, Jász Bori: Előadás-technikai felkészítő: nem elég okosnak lenni, annak is kell látszani

IDŐPONT: 2014. május.22. 16:00

HELYSZÍN: BME V1 épület, 5. emelet, 502-es terem
Budapest, 1111, Egly József utca 18.

Program

A konferencia időpontja: 2014. május 29. csütörtök
A konferencia helye: BME V1. épület, 5. emelet, 502 terem
Budapest, 1111 Egrý József utca. 18.

9:20	Dr. Pap László	Megnyitó
------	----------------	----------

1. szekció V1. 502 elnök: Dr. Pap László 9:40-11:00

Bakró Nagy István	Elektrosztatikus mérés szimulációja multiprocesszoros környezetben
Botlik Barnabás	Comparison of the performance of tone reservation based PAPR reduction techniques
Bokori Sára	Implementation and Simulation Investigation of TCP Optimization Methods in LTE Networks
Meződi Máté	Thermal Characterization of System on Chip Packages
Kufcsák András	Comparison of positioning algorithms for PET detectors

2. szekció V1. 502 elnök: Dr. Veszely Gyula 11:20-13:00

Kalvach Arnold	Effektív anyagparaméter kinyerő módszerek összehasonlítása
Csajbók Viktória	Fémes nanorészecskék elektromágneses elrejtése
Vermes Anna	Nemkonvencionális fotolitográfia plazmonikus nanoszerkezetekkel
Budai Tamás	Végeselemes megoldó párhuzamos implementációja MPI segítségével
Unger Tamás	Nemlineáris anyagtulajdonságok modellezése a numerikus térszámításban
Friedl Gergely	Rádiófrekvenciás terek numerikus analízise

3. szekció V1. 501 elnök: Dr. Bíró József 11:20-13:00

Mihálka Bence	Virtuális forgalomtovábbítási táblák tömörítése
Majdán András	Benchmarking Succinct Data Structures
Solymos Szilárd	Digitális szökökút alapú transzport protokoll fejlesztése és elemzése
Szabó Gábor	Effects of including more words in text units on bilingual predictive dictionaries
Debreceni Csaba	Automated abstraction in model-driven engineering
Czékus József	TDM- és WDM-PON hálózatok analitikus költségelemzése

4. szekció V1. 502 elnök: Dr. Nagy Lajos 13:40-15:20

Gorócz Vilmos	Obszervatóriumi műszer fejlesztése és megvalósítása tellurikus áramok méréséhez
Szeile Aliz	Optimalizálási algoritmusok szenzorhálózat alapú pozicionálásra naprendszerbeli égitesten
Kiss Dóra	SSR alapú monopulse követő rendszer
Pető Tamás	DVB-T Alapú Passzív Radar
Farkas Pál	Tölcsérantennák irányítottságának növelése nulla törésmutatójú metaanyag segítségével
Simon Dániel	Radardetektor tervezése olcsó ISM rádió IC segítségével

16:00	Dr. Pap László	Eredményhirdetés és zárszó
-------	----------------	----------------------------

Mesterpróba 2014

DÍJAZOTTAK

Budapesti Műszaki és Gazdaságtudományi Egyetem, Villamosmérnöki és Informatikai Kar
díja

Kalvach Arnold
Effektív anyagparaméter kinyerő módszerek összehasonlítása
Konzulens: Dr. Szabó Zsolt

Pető Tamás
DVB-T Alapú Passzív Radar
Konzulens: Dudás Levente

Huawei Technologies Hungary Kft. különdíja

Solymos Szilárd
Digitális szökökút alapú transzport protokoll fejlesztése és elemzés
Konzulens: Dr. Molnár Sándor

Elektrosztatikus mérés szimulációja multiprocesszoros környezetben

Bakró-Nagy István

Szélessávú Hírközlés és Villamosságtan Tanszék
Budapesti Műszaki és Gazdaságtudományi Egyetem
Budapest
bakro.istvan@gmail.com

Reichardt András

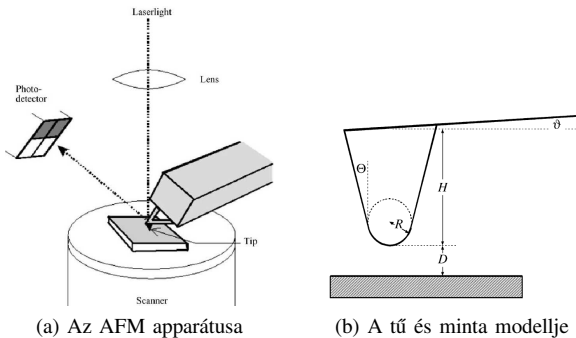
Szélessávú Hírközlés és Villamosságtan Tanszék
Budapesti Műszaki és Gazdaságtudományi Egyetem
Budapest
reich@evt.bme.hu

Kivonat—Atomerő mikroszkóppal való fémezett felületű minta felületi töltéssűrűségének mérése során kritikus a tű és a minta közötti kapacitás ismerete. A kapacitás értékét közelítések mellett lehetséges analitikusan kifejezni. Numerikus szimulációval pontosabban ismerhetjük az értéket, ezáltal nagyobb felbontás is érhető el. A szimuláció során a lehetséges párhuzamosításokat felhasználva a számítási időt elfogadhatóra csökkentjük.

Index Terms—AFM, felületi töltéssűrűség, numerikus szimuláció, párhuzamos számítás, OpenCL, MATLAB Parallel Toolbox

I. BEVEZETÉS

1986-ban Binning demonstrálta az atomerő mikroszkóp (AFM) ötletét [1], ami mára a nanotechnológia egyik legfontosabb eszköze lett. Felhasználható képalkotásra, nanolitográfiára és adott anyag alakítására [2]. Az AFM apparátusa az adott minta felülete és a felette pásztázó kantilever végére erősített tű kölcsönhatásának vizsgálatát végzi. (lásd 1b. ábra) A felület és a tű közötti domináns kölcsönhatás határozza meg, hogy az anyag melyik fizikai mennyiségét kaphatjuk meg.



1. ábra. Az AFM apparátusa látható az (a) ábrán. A lázernyaláb a kantilever felületéről tükröződve egy fotodetektorra irányul. A tű pozícióját ez alapján nagy pontossággal ismerjük. A (b) ábrán a minta és a felette lévő tű modellje látható a kapacitás analitikus számításához. A tű R sugarú H magasságú és D távolságra van a mintától. (Forrás: [3])

Az AFM felhasználása kontakt illetve kopogtató üzemmódú lehet. A kontakt mód során a felületen végighúzzuk a tűt és mérjük a z irányú elmozdulását. Így képesek vagyunk a minta felületén lévő atomok elrendezéséről magasságtérképet adni. Kopogtató mód [4] során a tűt elemeljük a mintától és f frekvenciával rezgetjük. A letapogatás során az átlagos minta-tű távolságot a kontakt módú magasságtérkép felhasználásával konstans értéken tartjuk. A kantilever dinamikáját ismerve a rezgetés frekvenciájának eltéréseiből számítható a tűre ható erő. Ezen erő nagyságát két tényező befolyásolja:

- a) a minta és a tű közötti feszültség,
- b) a minta felületi töltéssűrűség eloszlása.

A cikkben a felületi töltéssűrűség mérését tekintjük célunk. A [5], [3] szerint az erő a) komponensét a minta és a tű közötti kapacitásból az (1) szerint származtathatjuk.

$$F_s = -\frac{dE}{dD} = -\frac{d(CV^2/2)}{dD} = -\frac{1}{2} \frac{dC}{dD} V^2 \quad (1)$$

Ha a minta pásztázása során ezen a) erőkomponens konstansnak mondható, tehát a felületi érdesség kicsi, akkor a töltéssűrűg mérésében állandó hibát okoz. ami eliminálható. Az (1) számításában a kritikus elem a kapacitás értéke, amit numerikus számítás mellőzése esetén a [6] szerinti analitikus eredményt használhatjuk fel. A tű formáját a (1b ábra) szerintinek veszi és a mintát sík felületűnek feltételezi. Ezen utóbbi feltételezés legtöbb esetben helyénvaló, viszont a minta nagyfokú érdessége és a tű egyedi formája esetén érvényét veszti. Ilyen esetben a kapacitás értéke mintáról mintára változik és állandó hiba helyett, a mérést zajként terheli. A kapacitás numerikus szimulációjával ezen zajt is ki lehet küszöbölni. Persze ezen szimulációt minden egyes mérési pontban el kell végezni, aminek a kivitelezése csak multiprocesszoros környezetben lehetséges elfogadható idő alatt.

II. A FELADAT

A minta egy fémezett felület, aminek magasságtérképét mérések eredményeként adott pontossággal ismerjük. A mérések egy négyzetes háló felett történtek, amelynek a háló mindkét vízszintes irányában $\delta_x = \delta_y \simeq 180nm$ azonos felbontása, illetve a függőleges $\delta_z \simeq 20nm$ felbontása volt. A töltéssűrűség méréséhez szükséges második pásztázás során a tűt felemeljük és a mintához képest V_{tu} potenciálra kapcsoljuk. Ezután a fémezett felületről mindig azonos távolságra rezgetjük és ezen hosszú hegyes tűre ható erőt mérjük.

Végző cél olyan szimulátor építése, amelynek segítségével közel valós időben lehetséges a mérés alapján a felületi töltéseloszlásról korrigált/pontosabb információt kapni.

A cikkben felhasznált magasságtérkép-mérési eredmény (2. ábra) egy 512×512 méretű szürkeárnyalatos *.tiff állomány, amely értéke 0 – 255-ig terjed.

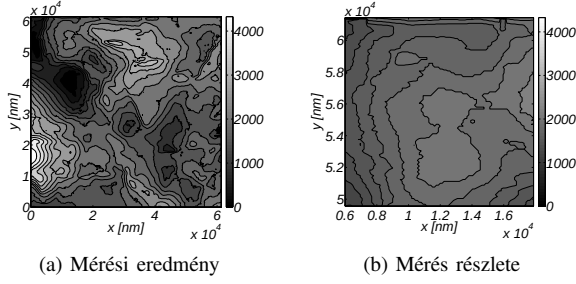
III. SZIMULÁCIÓ

III-A. Fizikai probléma matematikai formalizálása

A megoldandó feladat egy elektrosztatikus feladat. A minta és a tű közötti térben nincsenek töltések, így itt a Poisson egyenlet helyett a Laplace-egyenlet (2) érvényes.

$$\Delta V(x, y, z) = 0 \quad (2)$$

Az egyenletet későbbiekben részletezett (III-B) megfontolások végett egy redukált 3D-s téren oldjuk meg. Ezen 3D-s térre egy inhomogén



2. ábra. Méréssel kapott magasságtérkép. Felbontás $d_x = d_y = 120nm$ $d_z = 18.03nm$

ponthálót illesztünk, amelynek vízszintesen $d_x = d_y$, függőlegesen d_z a felbontása. A függőleges felbontás megegyezik a használt AFM apparátus felbontásával $d_z = \delta_z$. Az így kapott térbeli háló minden pontjához hozzárendeljük az $V_{i,j,k} \simeq V(id_x, jd_y, kd_z)$ potenciált. Dirichlet határfeltételek a felület fémezése, amely zérus potenciálú és az adott (V_{tu}) potenciálú tű fémes felülete. A térnek a minta felületétől különböző határfelületén homogén Neumann feltételt alkalmazunk a szimmetriák (végtelen tér) és a töltésmenetség miatt.

Az így adódó lineáris egyenletrendszer megoldására lehetséges direkt és iteratív megoldó algoritmusokat alkalmazni. A párhuzamosítási szándékok miatt az iteratív megoldást választottuk, mivel a multiprocesszoros környezetek tipikusan kevés fajlagos-memóriával¹ rendelkeznek. Ekkor nem teljesen pontos megoldást kapunk, azonban gyorsabban juthatunk el a kívánt eredményhez. A számítási pontosság növelhető az iterációt leállító konvergencia követelmény keményebb megszabásával, ami persze több iterációt jelent.

Az iteratív megoldás során a megoldás aktuális értékének kiszámításához az előző megoldásból indulunk ki. A (2) egyenletben szereplő deriválást az elsőrendű Taylor közelítés alkalmazásával a (3) szerinti 6-pontos sémát kapjuk.

$$V_{ijk}^{n+1} = \Delta_1 \cdot (V_{i-1,j,k}^n + V_{i+1,j,k}^n + V_{i,j-1,k}^n + V_{i,j+1,k}^n) + \Delta_2 \cdot (V_{i,j,k-1}^n + V_{i,j,k+1}^n) \quad (3)$$

ahol $V_{i,j,k}^n$ az az n -dik iterációs lépésben az i, j, k indexű pontban mérhető potenciált jelöli, Δ_1 a vízszintes felbontásból, Δ_2 a függőleges felbontásból adódó állandó.

III-B. Szimuláció felépítése

A felületmérés során a vízszintes felbontás jóval kisebb, mint a függőleges felbontás, azaz a pontok vízszintes távolsága sokkal nagyobb a függőlegesnél $d_x = d_y = 120nm \gg d_z = 10nm$ (lásd II. rész). A Coulomb-kölcsönhatás a távolság négyzetével fordítottan arányos, így az előbb említettek értelmében egy mérési pont szomszédjait, pontosabban egy redukált környezetét szükséges csupán szimulálni (Másképpen megfogalmazva a vízszintes mérési pontok távolsága jóval nagyobb mint a Coulomb-kölcsönhatás effektív távolsága). Ezzel az elhanyagolással a feladat már numerikusan kezelhető méretűre csökken.

Ezen módon egy mért pont 3×3 -as környezetét vesszük figyelembe, a középső pont felett lévő elektródát (tűt) feltételezve. A szimulációs tér alsó felületét a mért magassági 3×3 mintákból kell meghatározunk. Mivel ezen pontokra úgy is tekinthetünk, mint a

¹Fajlagos alatt az egy szimulációra jutó memóriát értem. (Természetesen ezen szimulációk egyszerre futnak, így a fajlagos-memória akkumulálódik.)

minta magasság-függvényének a mintavételezésével kapott mintáira, így a közbűlő pontokat interpolációval (mozgó átlagolással) kapjuk meg. Az interpoláció N_{ip} faktorával² lehetséges a szimulációs tér vízszintes felbontását $d_x = d_y = \delta_x / N_{ip}$ megkapni.

A szimulációk során a felület magasságának mérési adatait már ismertnek feltételezzük. A teljes magasságtérkép pontjait külön-külön vizsgáljuk. Egyetlen pontban a mérési eredmény kiszámításának lépései az alábbiak:

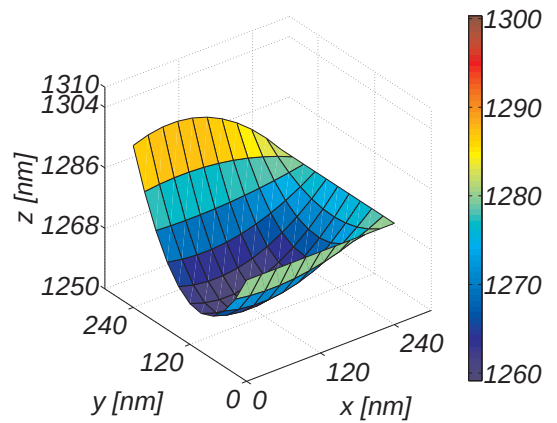
- 1) A pont körüli felület 3×3 -as mérési részének megállapítása,
- 2) Közbenső (virtuális) mérési pontokkal a belső felbontás növelése interpolációval,
- 3) Szimulálandó tér méretének számítása,
- 4) Direkt/iteratív megoldó algoritmussal a tér meghatározása, a tűre ható erő számítása illetve a tű alatti töltésmennyiség számítása,
- 5) Adatok mentése.

IV. A SZIMULÁTOR BEMUTATÁSA

A prototípus algoritmus fejlesztése MATLAB környezetben történt, ami később referenciaként szolgál. Alap MATLAB utasításokat használva több órát vesz igénybe a szimuláció futtatása. A MATLAB Parallel Toolbox-nak segítségével a szimulációt lehetséges párhuzamosan több processzormagon futtatni. Ezzel párszoros sebesség növekedés érhető el. A következőkben magát az algoritmust és az OpenCL keretrendszerben történő implementációját mutatjuk be. Majd az eredmények bemutatása során kerül összevetésre a MATLAB referencia, a MATLAB Parallel Toolbox segítségével, az OpenCL processzoron és az OpenCL GPU-n való futtatási ideje.

IV-A. A lépések részletezése

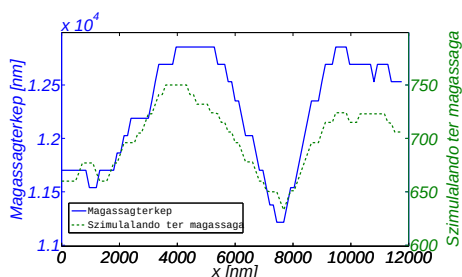
IV-A1. Interpoláció: A korábban elmondottak alapján a felületet további virtuális pontokkal egészítjük ki. Az virtuális mérési pontokat a legegyszerűbb síklapos közelítéssel alkothatjuk meg, a magasságmérési felbontás figyelembevételével. A szimulátorban egy általánosabb módszert alkalmazunk, ami egy 2D-s mozgó átlagoló szűrővel való simítás. A szűrővel aluláteresztést tudunk elérni, ami a minta magasságának mintavételezése utáni rekonstrukcióját jelenti. Egy ilyen interpoláció eredményét láthatjuk a 3. ábrán.



3. ábra. 3×3 mérési pont 11×11 pontba való interpolációja

IV-A2. A szimulálandó tér mérete: A szimulálandó tér (hasáb) alapja adott az előzőleg említett interpolált felületként, míg a magassága nem. Ezt a következő két mennyiség közül a nagyobbikkal határoztuk meg:

² N_{ip} -szeresére növeljük a pontok számát.



4. ábra. A mérési eredmény egy vonal menti részlete (folytonos vonal) és az ezen mérési pontokhoz számított szimulációs tér magassága (szaggatott-vonal)

- Középső pont fölött lévő tű közepének magassága,
- A (3×3) környezet legalacsonyabb és legmagasabb pontjának különbsége.

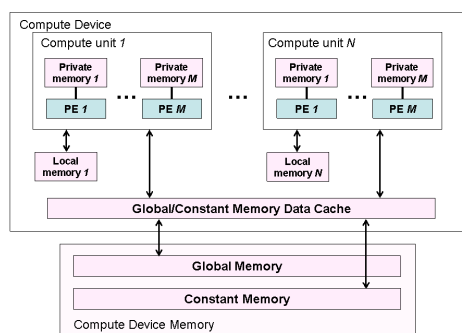
A magasságtérkép egy vonalának részlete látható a 4. ábrán, továbbá a szimulálódó tér (hasáb) magassága.

IV-A3. Iteratív megoldó algoritmus: Az iterációhoz a térháló pontjaihoz két mátrixot (tömböt) rendelünk, ami a pontok potenciáljának aktuális (U_{now}) és előző (U_{prev}) értékeit tartalmazza. Az aktuális értékeket (3) szerint számítjuk, majd az egész térre számítjuk az előzővel vett különbségének négyzetösszegét (normáját). E mérték képviseli a konvergencia szintjét, amit az iteráció során vizsgálva jutunk el a kívánt konvergencia szintre. Ha nem értük el a konvergencia szintet, akkor az előző két mátrixot felcserélve iterálunk tovább.

IV-A4. Adatok mentése: Tesztelhetőségi megfontolások végett nem csak a tűre ható erőt (villamos térerősséget) exportáljuk, hanem a konvergencia szintjének változását és az interpolált felületet is. Az exportálandó menyiségek „kis” mérete miatt egyszerű *.csv fájlként kerülnek mentésre. Ezen fájlok további poszt-processzálása MATLAB vagy munkalap kezelő szoftverrel is elvégezhető.

IV-B. OpenCL architektúrája

Az Open Computing Language (OpenCL) keretrendszer [7] közös nyelvet, magas szintű programozási interfészt és hardware absztrakciót nyújt a fejlesztőknek adat- vagy feladat párhuzamos számítások gyorsítására különböző számítógépsíkon (CPU, GPU, FPGA, DSP, ...). Az OpenCL modellje a különböző „device”-okra bontható, amik több „compute unit”-ot (processzor-magot) tartalmaznak és ezeket heterogén módon kezelik.



5. ábra. OpenCL „device” architektúra [7]

A „compute unit”-ok kiéheztetésének elkerülése végett több ezer „work-item” virtuális osztozik rajta. Továbbá ezen „work-item”-ek „work-group”-okba vannak rendezve (később részletezett megfontolások végett). A „compute unit” kiéheztetését a „device”-on található memória chippek lassúsága okozza. Ennek hárdiveres megoldása a több

szintű prediktív cache memória beiktatása a „compute unit” és a külső memória közé. Mivel a bank szervezett külső memóriák hozzáférési ideje relatíve nagy így a memória szervezésére nagy hangsúlyt kell fektetni.

Az OpenCL négy memória szintet különböztet meg, ami az II táblázatban és az 5. ábrán látható. Ahhoz, hogy a rendszerben rejlő teljesítményt kiaknázzuk, három fontos kérdést kell a szimulátor magjának implementálásakor megválaszolniuk:

- **Mennyit?:** Tisztában kell lennünk az aktuális memória fogyasztással és a szükséges memóriamérettel.
- **Honnan-hova?:** Fontos, hogy a lehető legközelebb legyen az adat a „work-item”-hez.
- **Mikor?:** Mivel a memória művelet alatt a „work-item” nem dolgozik, így átadja a helyét egy másiknak. Ennek a megfelelő szinkronizációjával nagyobb kihasználtság érhető el (load balance).

II. táblázat. OpenCL memória szintek

	Global	Constant	Local	Private
Host	Dinamikus R/W	Din. R/W	Din. R/W	
Kernel	R/W	Statikusan R	Satik. R/W	Statik. R/W
Sebesség	Lassú	Gyors	Gyors	Regiszter
Méret	1 Gbyte <	~ 64 Kbyte	~ 16 Kbyte	< 1 Kbyte

OpenCL keretrendszerben történő programozás során két programot kell írunk. Az egyik a „host”-on fut, ami elvégzi a probléma összeállítását, memória allokálását, argumentumok beállítását és a kernel (másik program) meghívását a „device”-on. A kernel futása végeztével a „host” program kiolvassa a „device”-ból a kívánt eredményt.

IV-C. Implementációhoz szükséges megfontolások

A következőkben egy kisebb teljesítményű notebook videokártyát veszek alapul a megfontolások demonstrálására. Ez az nVidia GeForce 330M, 575 MHz-en futó 48 CUDA core-al, 1024GB memóriával és OpenCL 1.0 kompatibilitással. A videokártya továbbiakban fontos paraméterei a III. táblázatban látható.

III. táblázat. nVidia GeForce 330M OpenCL tulajdonságai

MAX_COMPUTE_UNITS	6
MAX_WORK_GROUP_SIZES	512 512 64
GLOBAL_MEM_SIZE	1073020928
MAX_CONSTANT_BUFFER_SIZE	65536
LOCAL_MEM_SIZE	16384

Ha a tér, ahol a laplace egyenletet meg kell oldanunk nagyon nagy, akkor érdemes szétbontani kisebb alterekre és azokhoz rendelni egy-egy „work-item”-et. Mivel a diszkrét Laplace egyenlet egy pontja a szomszédos pontokkal szoros kapcsolatban van, így az összefüggő „work-item”-eket egy „work-group”-ba érdemes szervezni, mivel így az átlapolódó pontok értékét a szomszédos „work-item”-ek is tudják írni és olvasni. Az ilyen típusú problémának méretét a MAX_WORK_GROUP_SIZES tulajdonság korlátozza.

Jelen esetben a mérési eredmény egy pontjához tartozó tér átlagosan $11 \times 11 \times 30$ pontból áll. Tehát a korábbi nem áll fenn és egyszerű megfeleltetéssel szétoszthatjuk a feladatot. A teljes tér $512 \times 512 \times 11 \times 11 \times 30$ méretű, ami 951k pont. A tárolásához single-precision mellett ennek a számnak a 4-szerese szükséges byte-okban mérve. Mivel ez a videokártyán nem áll rendelkezésre, így szétbontjuk kisebb feladatrészekre.

I. táblázat. OpenCL futási idő eredmények 12×12 mérési pontra

	Globális memória	Lokális memória, ha befér	Lokális memória bufferelés
Globális tranzakciók száma átlagosan	$12 \times 12 \times 32.3$	$12 \times 12 \times 32.3$	$12 \times 12 \times 32.3$
Lokális tranzakciók száma átlagosan	0	$0.48 \times 12 \times 12 \times 30$	$2.08 \times 12 \times 12 \times 32.3$
Futási idő	5990 ms	2530 ms	510 ms
Fajlagos futási idő	410 ms	170 ms	3.5 ms

Ezen feladatrészek méretét egy paraméter állításával lehet változtatni és az implementált algoritmus ettől generikusan függ. Emellett az interpoláció mértéke N_{ip} is paraméterrel generikusan állítható. Az algoritmus generikusságát csupán a futási időben történő dinamikus memória allokációval lehetséges megvalósítani. A korábban említettek végett (II táblázat) az allokáció csak a „host” programban történhet.

IV-D. Memória szervezés

IV-D1. Csak globális memória használata: Az algoritmus pszeudó kódjának direkt leképezése esetén a „host”-on allokálunk memóriát a „device” globális memóriájában, majd a megfelelő adatokat ide másoljuk és a kernel is itt ír és olvas. A problémát a globális memória nagy hozzáférési ideje jelenti, ami miatt sok „work-item” tétlenül a memóriára fog várakozni. Ilyenkor az egy mérési pontra vonatkoztatott szimulációs idő a referenciánál is lassabb.

IV-D2. Globális memória és adott esetben lokális memória használata: Kis erőfeszítéssel nagy javulást lehet elérni, ha a mérési ponthoz tartozó szimulációs tér éppen belefér a lokális memóriába. Tehát, mielőtt az (3) szerinti iteratív megoldót futtatnánk először a globális memóriából a lokális memóriába töltjük át a kérdéses pontokat, majd számolunk rajta és a végén visszatöltjük a globális memóriába. E javítással a referenciával azonos sebességet tudunk elérni.

IV-D3. Globális memória és minden adódó alkalomkor a lokális memória használata: Nagyobb erőfeszítést igényel, hogy minden alkalommal a globális memóriával való kommunikációt a lokális memória közbeékelésével tegyük. Ezt úgy lehet felfogni, mintha a globális memóriát lokális memória méretű kvantumokban tudnánk csak elérni. Ekkor nagy odafigyelést kíván a memóriacímzés megfelelő prgramozása, de eredményképp gyorsulás érhető el.

Összegezve elmondható, hogy az aktuálisan használt adat tárolását a lehető legközelebb kell tartani a „compute-unit”-hoz.

V. EREDMÉNYEK

V-A. MATLAB implementációk

A referenciaként szolgáló MATLAB algoritmus lineáris program-szervezést alkalmazva az elérhető fajlagos futási idő $\sim 100ms$.

A kód minimális változtatásával elérhető a párhuzamos végrehajtás. Ezt a for ciklusok Parallel Toolbox beli `parfor` utasítására cserélve érhetjük el. 4 processzormaggal rendelkező PC-n legjobb esetben a negyedére csökkenhet a futási idő. Valójában ez sose történik meg. A MATLAB Parallel Toolbox-a az egyes szimulációkat adott processzormagra osztja el. Korábban láttuk, hogy ezen szimulációk lépésgénye nagyban eltér egymástól, így előáll az a sajnálatos eset, hogy 3 mag tétlenül a negyedikre vár, ami miatt ez nem tekinthető járható útnak.

V-B. OpenCL implementációk

OpenCL keretrendszer segítségével írt programot a GPU-n futtatva a I. táblázatban látható eredményeket kapjuk. Csupán a globális

memóriát használva a referenciához képest romlik a teljesítmény. Ezt a videokártya prediktív cache nélküli kialakításának és a globális memóriájának a kiéheztetésének tudhatjuk be. A lokális memória használata a futási időt drasztikusan le tudja csökkenteni, ami a korábban ismertetett memória szervezési megfontolások helyességét igazolja.

VI. ÖSSZEGZÉS

A cikkben összefoglaltam az AFM felületi töltéssűrűség mérés-technikáját, amiben azonosítottam a kapacitás értékének kritikus voltát. Ezidáig a kapacitás értékének a [6], [3] szerinti közelítéseket tartalmazó analitikus eredményt lehetett felhasználni.

Feladatként ezen kapacitás értékét számító szimulátor építését tűztam ki, aminek elfogadható időn belül kell eredményt szolgáltatnia. A párhuzamosítás lehetősége triviálisan adódott. A hordozhatóság és a gyorsabb végrehajtás végett a szimulátort OpenCL környezetben implementáltam, ami a szimulátor heterogén multiproszessoros környezetben való futtatását lehetővé teszi.

Az eredményeket ismertetve a szimulátor futási idejében látványos gyorsulást tapasztaltam, ami az érdesebb felületek esetén is elfogadhatóan pontos eredményt tud szolgáltatni.

A szimuláció felhasználásával történő töltéssűrűség származtatása még várat magára, ugyanígy ezen származtatás validálására való mérési összeállítás kidolgozása. A szimulátor magját képező lineáris egyenletrendszer iterációs megoldó konvergenciájának bizonyítása és az alternatív direkt megoldó vizsgálata is további feladat.

HIVATKOZÁSOK

- [1] G. Binnig, C. F. Quate, and C. Gerber, „Atomic force microscope,” *Phys. Rev. Lett.*, vol. 56, pp. 930–933, Mar 1986. [Online]. Available: <http://link.aps.org/doi/10.1103/PhysRevLett.56.930>
- [2] B. Vasić, M. Kratzer, A. Matković, A. Nevošad, U. Ralević, D. Jovanović, C. Ganser, C. Teichert, and R. Gajić, „Atomic force microscopy based manipulation of graphene using dynamic plowing lithography,” *Nanotechnology*, vol. 24, no. 1, p. 015303, 2013. [Online]. Available: <http://stacks.iop.org/0957-4484/24/i=1/a=015303>
- [3] H.-J. Butt, B. Cappella, and M. Kappl, „Force measurements with the atomic force microscope: Technique, interpretation and applications,” *Surface Science Reports*, vol. 59, no. 1–6, pp. 1 – 152, 2005. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0167572905000488>
- [4] Y. Martin, C. C. Williams, and H. K. Wickramasinghe, „Atomic force microscope–force mapping and profiling on a sub 100 a scale,” *Journal of Applied Physics*, vol. 61, no. 10, pp. 4723–4729, 1987. [Online]. Available: <http://scitation.aip.org/content/aip/journal/jap/61/10/10.1063/1.338807>
- [5] H.-J. Butt, „Measuring electrostatic, van der Waals, and hydration forces in electrolyte solutions with an atomic force microscope,” *Biophys J.*, vol. 60(6), p. 1438–1444., 1991. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1260203/>
- [6] S. Hudlet, M. Saint Jean, C. Guthmann, and J. Berger, „Evaluation of the capacitive force between an atomic force microscopy tip and a metallic surface,” *The European Physical Journal B - Condensed Matter and Complex Systems*, vol. 2, no. 1, pp. 5–10, 1998. [Online]. Available: <http://dx.doi.org/10.1007/s100510050219>
- [7] Khronos OpenCL Working Group, *The OpenCL Specification, version 1.0*, 6 August 2010. [Online]. Available: <http://www.khronos.org/registry/cl/specs/opencl-1.0.pdf>

Comparison of the performance of tone reservation based PAPR reduction techniques

Barnabás Botlik, Zsolt Kollár

Department of Broadband Infocommunications and Electromagnetic Theory
Budapest University of Technology and Economics
Budapest, Hungary

Email: barnabas.botlik@gmail.com, kollar@mht.bme.hu

Abstract— Orthogonal frequency division multiplexing (OFDM) is a commonly used modulation scheme for high speed digital communication. The high peak-to-average power ratio (PAPR) of the modulated signal is a major drawback, which may negatively affect the overall system performance if amplifiers with limited linear range are applied. One of the many PAPR reduction techniques is tone reservation (TR), which is also a part of the DVB-T2 standard. In this paper comparison and simulation results for different realizations of TR are presented. We propose implementation parameters for which the best PAPR reduction performance can be achieved. A novel TR technique based on enlipping is also investigated.

I. INTRODUCTION

Orthogonal Frequency Division Multiplexing (OFDM) is a widely used multicarrier modulation technique due to its spectral efficiency and robustness in multipath propagation scenarios. This modulation scheme is applied in many high rate wired and wireless communication systems such as Digital Video Broadcasting (DVB), Digital Audio Broadcasting (DAB), Wireless Local Area Network (WLAN) and in Asymmetric digital Subscriber Line (ADSL) as well. The usable frequency range is divided into multiple frequency bands which can be easily separated using Fast Fourier Transform (FFT). Each carrier frequency can be modulated using Quadrature Amplitude Modulation (QAM) signals. However, it suffers from high Peak-to-Average Power Ratio (PAPR), which may negatively affect the system performance if nonlinearities (e.g. nonlinear amplifiers, mixers) are present in the transceiver chain.

There has been numerous techniques presented in the literature for reducing the large dynamics of the modulated signal. The theoretical background of many techniques including amplitude clipping and filtering, coding, partial transmit sequence, selected mapping, interleaving, Active Constellation Extension (ACE) and Tone Reservation (TR) are summarized in [1]. Each technique apart from their PAPR reduction capabilities have disadvantages, some of them require additional power, others reduce the data rate and some may introduce nonlinear distortions. Currently, the ACE and TR are the only techniques which have a practical implementation in the terrestrial second generation DVB standard (DVB-T2) [2], [3]. In this paper we will investigate the TR technique.

The basic idea of the TR methods is that a previously defined number of tones – called reserved tones (RTs) – do not carrying any data, these are used to reduce the PAPR value of the OFDM symbols. The main advantage of this

technique is that the receiver can be unaltered without bit error rate performance degradation, only the indices of the RTs have to be known by the receiver. The performance of this method is strongly dependent on the number of RTs. The higher the number of these tones, the better improvement can be achieved, however this causes data rate loss. Finding the optimal modulation values on the reserved tones is not a straight forward task. There has been suboptimal solutions proposed in the literature [2]–[6]. In this paper these methods are investigated.

In [4] there are two implementation of this technique presented, namely the Clipping-based TR (TR-C) [4], [5] and the Kernel-based TR (TR-K) method [3]. In this paper a novel technique is also introduced: the Enlipping-based TR (TR-E) technique. TR-K is part of the DVB-T2 standard, and though it explains the major steps of the method, there are still a few parameters for the algorithm, which are not presented therein, and finding the proper values for them is still an open issue.

The paper is organized as follows. In Section II signal model of the OFDM transmission is given, and the statistical properties of the modulated signal is described. After that, in Section III a brief overview is given about the examined techniques. At the end, in Section IV the simulation results of the different methods are presented, where we also compare the presented TR schemes from a practical point of view. Finally, Section V concludes the paper.

II. OFDM SIGNAL MODEL

An OFDM symbol is generated using an N -point Inverse FFT (IFFT), which is basically a summation of modulated complex harmonics, and can be expressed as follows:

$$x_n = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} X_k e^{j2\pi \frac{kn}{N}}, \quad 0 \leq n < N. \quad (1)$$

In (1) the index of subcarriers is denoted by k , while the index of the discrete sample and the complex modulation value of the k^{th} subcarrier is n and X_k respectively. Calculating the value of the PAPR for OFDM symbols, we can come to the following expression:

$$\text{PAPR} = \frac{\max |x_n|^2}{\mathbb{E}\{|x_n|^2\}}, \quad 0 \leq n < N, \quad (2)$$

With this formula one can compute the probability, that the PAPR of an N sample long OFDM symbol exceeds a

previously determined PAPR_0 threshold as

$$P(\text{PAPR} > \text{PAPR}_0) = 1 - (1 - e^{\text{PAPR}_0})^N. \quad (3)$$

The Complementary Cumulative Distribution Function (CCDF) is often used to visually represent a random variable, namely the PAPR value, when the PAPR_0 value is known. This function can be expressed as follows:

$$P(\text{PAPR} < \text{PAPR}_0) = 1 - P(\text{PAPR} > \text{PAPR}_0). \quad (4)$$

III. ALGORITHMS FOR TONE RESERVATION

In this section three TR algorithms are presented. The TR concept is based on the idea to hold some positions of the subcarriers, called reserved tones (RT), and adding a specified signal to these positions in frequency domain helps to reduce the PAPR value in the time domain. To generate this signal, various implementations exist. First, a kernel based method is described which is used in the DVB-T2 standard, then a clipping based method is explained, finally a novel technique based on enlipping – which has not yet been used before – is introduced.

A. Kernel-based tone reservation

The TR-K technique is based on that principle, that high PAPR values can be reduced, when we reduce the peaks in the transmitted signal. The block diagram of the peak-cancellation algorithm is shown in Figure 2. In the first step of this low-complexity algorithm we have to generate a kernel, which has to be an impulse-like function. Ideally it should be one peak, but in reality we cannot achieve a signal like that, so we will approximate it as close as we can. The generation of the kernel function is proposed in [2]. First, we have to set every reserved tone ($\mathbf{P}_k, k \in k_{\text{RT}}$) to the value 1 and every other tone ($\mathbf{P}_k, k \notin k_{\text{RT}}$) to the value 0. Then applying an IFFT and normalizing this signal, we can get the reference kernel $\mathbf{p}_n$:

$$\mathbf{p}_n = \frac{\sqrt{N_{\text{FFT}}}}{N_{\text{RT}}} \text{IFFT}(\mathbf{P}_k), \quad (5)$$

where N_{FFT} is the IFFT size and N_{RT} is the number of the RT positions. A circularly shifted kernel signal is shown in Figure 1.

In the second step of the algorithm, we have to find the maximum value for every symbol of the input signal:

$$A^i = \max_n |x_n^i|, \quad (6)$$

$$m^i = \arg \max_n |x_n^i|, \quad (7)$$

where x_n^i is the n^{th} element of the vector $\mathbf{x}^i$, A^i is the maximum magnitude, while m^i is the position of the detected peak in the i^{th} iteration. As the third step, the peak of the reference kernel has to be circularly shifted to the m^i position, so it will be perfectly aligned with the peak of the input signal. After that it has to be scaled and phase rotated so, that the power of the peak should be reduced to a previously determined target clipping level. Then the kernel signal is subtracted from the original one, and the PAPR value is calculated. If it reaches the desired limit, then the iterations terminate.

$$\mathbf{x}^{i+1} = \mathbf{x}^i - \alpha^i \mathbf{p}_n(m^i), \quad (8)$$

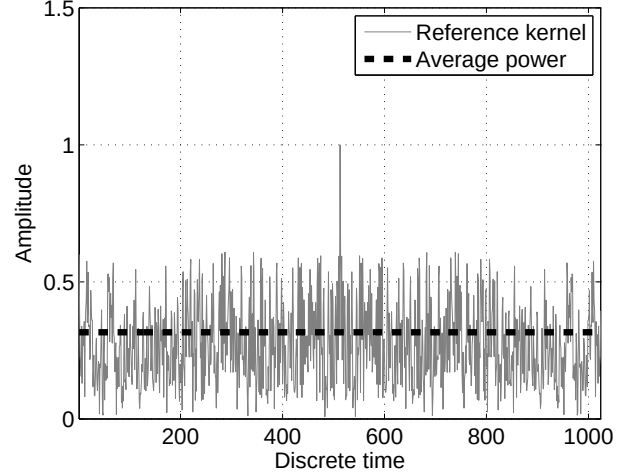


Figure 1: A circularly shifted kernel signal for $N = 1024$ subcarriers according to DVB-T2 [2]

where

$$\alpha^i = \frac{\mathbf{x}^i(m^i)}{A^i} (A^i - A_{\text{max}}), \quad (9)$$

where A_{max} is the clipping magnitude level. The clipping ratio (CR) is defined as:

$$\text{CR}_{\text{dB}} = 20 \log_{10} \frac{A_{\text{max}}}{\sigma_x}, \quad (10)$$

where σ_x is the square root of the average power of the signal. Subtracting the kernel from the input signal does not cause any distortion in the OFDM symbol, because circular shift and scaling in the time domain is only phase rotation at the RT positions in the frequency domain. Another condition for termination of the method could be the reaching of a predetermined iteration number.

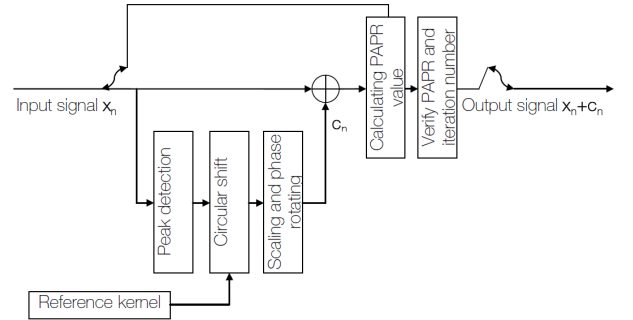


Figure 2: Block diagram of the TR-K algorithm [2]

B. Clipping-based tone reservation

The block diagram of TR-C is shown in Figure 3. In this method first a hard clipping is executed on the signal, which is then subtracted from the original signal. The result will be called as correction signal. We use here a classical amplitude

clipping function, which can be expressed as follows:

$$\mathbf{y}_n = \begin{cases} \mathbf{x}_n, & |\mathbf{x}_n| \leq A_{\max} \\ A_{\max} e^{j\varphi(\mathbf{x}_n)}, & |\mathbf{x}_n| > A_{\max} \end{cases}, \quad (11)$$

where $\mathbf{x}_n$ is the original signal, $\mathbf{y}_n$ is the clipped one, and $A_{\max}$ is the so-called clipping magnitude level. The correction signal $\mathbf{c}_n$, as it was mentioned above, is the difference of the input signal $\mathbf{x}_n$ and the clipped signal $\mathbf{y}_n$:

$$\mathbf{c}_n^i = \mathbf{x}_n^i - \mathbf{y}_n^i. \quad (12)$$

According to the TR concept, we have to use FFT (Fast Fourier Transform) on the correction signal, and we are keeping only the values at the RTs, while the other values will be set to zero:

$$\mathbf{C}_k = \text{FFT}(\mathbf{c}_n), \quad (13)$$

$$\hat{\mathbf{C}}_k = \begin{cases} \mathbf{C}_k, & \text{if } k \in k_{\text{RT}} \\ 0, & \text{if } k \notin k_{\text{RT}} \end{cases}. \quad (14)$$

Applying the IFFT on the signal $\hat{\mathbf{C}}_k$, the time domain filtered clipped signal $\hat{\mathbf{c}}_n$ can be calculated. At the end of each iteration, we have to add this resulting $\hat{\mathbf{c}}_n$ signal to the original $\mathbf{x}_n$ vector to form the new transmitted signal $\mathbf{x}_n^{i+1}$:

$$\mathbf{x}_n^{i+1} = \mathbf{x}_n^i + \mu \hat{\mathbf{c}}_n^i, \quad (15)$$

where μ is a parameter of the TR-C method.

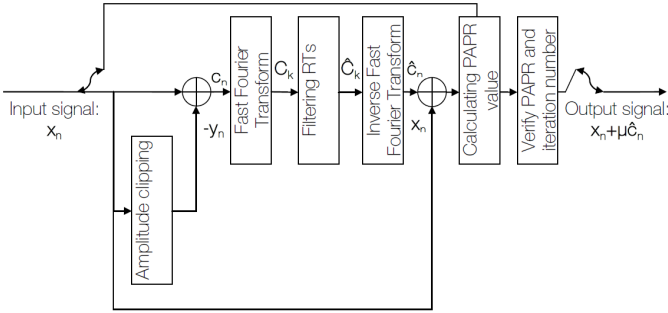


Figure 3: Block diagram of TR-C algorithm

C. Enlipping-based tone reservation

Enlipping is a novel clipping method originally introduced and used for ACE PAPR reduction [7]. The new idea was to apply this technique to the TR concept. The terminology *enlipping* is first used also in [7] and taken from there as well. The TR-E technique is similar to the TR-C, the only difference is that in the first step instead of clipping the high peaks to a desired clipping level, every amplitude value is set to $A_{\max}$. The enlipped signal $\mathbf{y}_n$ can be expressed with a minor modification of eq. (11) as

$$\mathbf{y}_n = A_{\max} e^{j\varphi(\mathbf{x}_n)}, \quad (16)$$

where $A_{\max}$ is the clipping magnitude level. The following steps are the same as presented in subsection III-B. With this technique we can reach lower PAPR values as for TR-C.

IV. SIMULATION RESULTS AND PRACTICAL COMPARISON OF THE PRESENTED TR SCHEMES

The performance of the three previously described techniques were compared through simulations. The simulation parameters are summarized in Table I. With higher number of RT indices better PAPR performance can be achieved, but in the first approach the choice of these variables are based on the DVB-T2 standard. In later examinations these simulations can be extended to further variations of the parameters.

FFT size	$N = 1024$
Number of RTs	10
RT indices	116, 130, 134, 157, 182, 256, 346, 478, 479, 532
Maximum amplitude of a RT	5
Modulation	4-QAM
Number of tones (N)	1024
Number of symbols	5000
μ for TR-C and TR-E	-1

Table I: Simulation parameters

Figure 4, 5 and 6 shows the relation between the CR and the achievable PAPR value at the probability of $10^{-1.8}$ in function of the number of iterations using TR-K, TR-C and TR-E techniques respectively. Choosing this limit to $10^{-1.8}$ seems to be a reasonable choice, because at larger probability levels the simulated CCDF curve is not reliable. For lower values longer simulations are required. In Fig. 4 it can be observed that the best result – approximately 8.4 dB – for TR-K can be achieved by choosing CR = 6 dB. In Fig. 5 and 6 it can be seen that TR-C and TR-E are less sensitive to the choice of the CR. For TR-C a clipping level of 4.8 dB leads to the best performance – 10.1 dB. In case of TR-E a wide range of CR=2-8 dB can be choose, all leading to the same result – 9.9 dB. It can be also concluded that the convergence of the TR-E is the fastest.

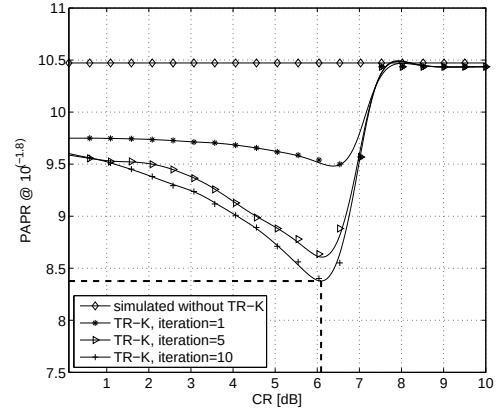


Figure 4: The achieved PAPR value at $P = 10^{-1.8}$ using TR-K method with various number of iterations

The CCDF curves of the three techniques are compared in Fig. 7 with different number of iterations. For each techniques a CR leading to the best performance was chosen based on the results presented in Figures 4-6.

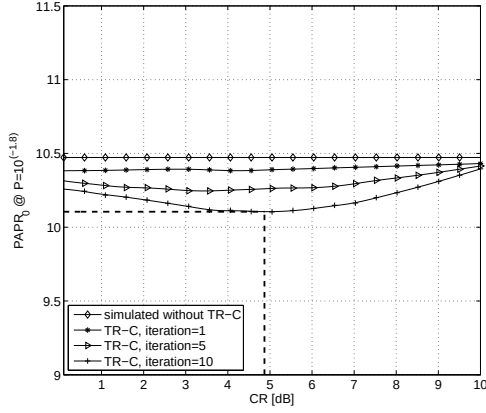


Figure 5: The achieved PAPR value at $P = 10^{-1.8}$ using TR-C method with various number of iterations

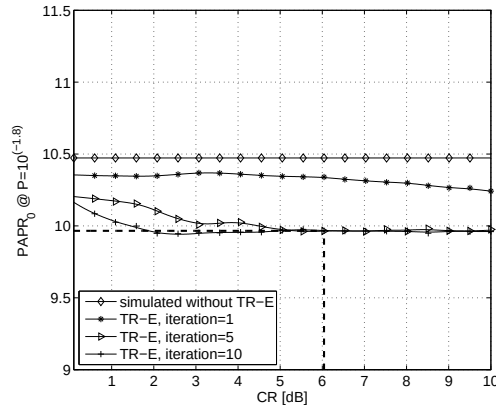


Figure 6: The achieved PAPR value at $P = 10^{-1.8}$ using TR-E method with various number of iterations

All of three presented TR schemes are capable of improving the signals' high PAPR. The TR-K algorithm has the best performance and the lowest computational complexity as it operates only in the time domain. In this paper we have investigated the TR for DVB-T2 standard which specifies the indices of the RTs. For an other set of parameters the position of the RTs is also an important question, and as it is described in [2], these indices are the results of a long optimization process. For a randomly chosen set of RTs an enlipping based TR is more suitable.

It can be also stated that TR-E outperforms TR-C and even TR-K in convergence steps and TR-C in performance as it requires less iterations and leads to lower PAPR values. On the other hand TR-E requires a higher computational complexity than TR-K as for each iteration an IFFT/FFT operations in needed. It can also be stated, that TR-C or TR-E should be used, when a kernel is not available or it cannot be generated for the given set of RT indices.

One additional parameter which has to be considered for the algorithms is the maximal allowed power of the RTs.

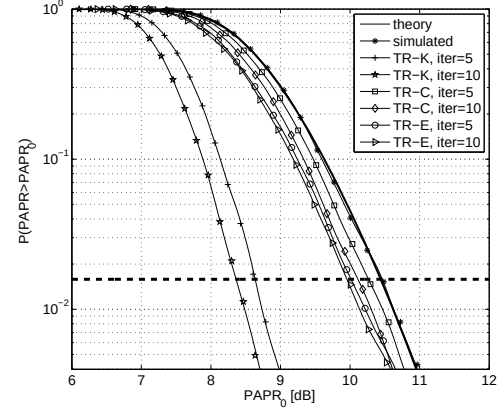


Figure 7: CCDF curves of investigated TR methods

Allowing higher values may lead to better performance but for practical implementation these values must be kept at a reasonable level.

V. CONCLUSION

In this paper three TR techniques for PAPR reduction of OFDM signals were described and compared. A novel enlipping based technique was introduced. The performance of these techniques were verified through simulations. It has been shown that each technique is strongly dependent on the chosen CR and their performance can be improved with increasing number of iterations. We have shown that the TR-K method has the best performance. The PAPR reduction performance of TR-E is slightly worse than TR-K but it can outperform the TR-C technique. The benefits of the TR-E method is that it has a fast convergence and it can operate with randomly selected TR indices as well.

REFERENCES

- [1] S. H. Han and J.-H. Lee, "An overview of peak-to-average power ratio reduction techniques for multicarrier transmission," *IEEE Wireless Communications*, vol. 12, no. 2, pp. 56–65, Apr. 2005.
- [2] ETSI, *Digital Video Broadcasting (DVB); Frame structure channel coding and modulation for a second generation digital terrestrial television broadcasting system (DVB-T2)*, ETSI EN 302 755 V.1.3.1, European Telecommunications Standards Institute Std., Apr. 2012.
- [3] R. Barsanti and J. Larue, "Peak to average power ratio reduction for digital video broadcast T2," in *Proceedings of IEEE Southeastcon 2011*, Mar. 2011, pp. 117–121.
- [4] T. Liu, X. Li, C. Chen, S. Cui, and Y. Liu, "PAPR reduction for 802.16e by clipping and tone reservation based on amplitude scale factor," in *ICPCA/SWS'12, Proceedings of the 2012 International Conference on Pervasive Computing and the Networked World*. Berlin, Heidelberg: Springer-Verlag, 2013, pp. 339–349.
- [5] S. Janaathanan, C. Kasparis, and B. Evans, "A gradient based algorithm for PAPR reduction of OFDM using tone reservation technique," in *Vehicular Technology Conference, 2008. VTC Spring 2008. IEEE*, May 2008, pp. 2977–2980.
- [6] M. Mroué, A. Nafkha, J. Palicot, B. Gavalda, and N. Dagorne, "Performance and implementation evaluation of TR PAPR reduction methods for DVB-T2," *International Journal of Digital Multimedia Broadcasting*, vol. 2010, pp. 1–10, 2010.
- [7] Y. Zhou and T. Jiang, "A novel clipping integrated into ACE for PAPR reduction in OFDM systems," in *International Conference on Wireless Communications Signal Processing. WCSP 2009*, Nov. 2009, pp. 1–4.

Implementation and Simulational Investigation of TCP Optimization Methods in LTE Networks

Sára Bokori

Department of Broadband Infocommunications and
Electromagnetic Theory
BME
Budapest, Hungary
sara.bokori@gmail.com

Árpád Drozdy

Department of Broadband Infocommunications and
Electromagnetic Theory
BME
Budapest, Hungary
drozdy@mht.bme.hu

Abstract — Due to the evolution of mobile technology and the emergence of terminals like smart phones, Internet based applications mostly accessed over the Transmission Control Protocol (TCP) have become dominant on mobile platforms. However, since the TCP flow control mechanisms were designed for wired environments, short-term degradations on the air interface of radio access systems, bad interaction with the unique characteristics of wireless networks, including fluctuating bandwidth, high and variable latency, and bursty bit-error rates or the TCP's timeout driven recovery mechanisms may lead to unnecessary performance drops and increased latency. TCP optimization of mobile networks is a timely and important research area, aimed at reducing access time to content and increase the efficiency of data transmission. TCP freeze is a possible solution to improve short-term degradations on the air interface of radio access systems by triggering the persist mode of the TCP sender, preventing TCP timeouts and rate reductions during poor radio conditions. This paper investigates a network element based implementation of the TCP freeze mechanisms. The solution is completely transparent to the TCP sender and receiver. The possible performance gain of the solution is investigated by simulations.

Keywords — TCP flow control mechanisms; TCP freeze; TCP performance improvement; Radio interface problems;

I. INTRODUCTION

Over the past few years, and especially with the mobile revolution, much economic and social activity has moved online (e.g., web browsing, video download, social networking, etc.), which can be reached over low latency, high capacity wireless access, provided by the evolved radio access systems beyond Wide band Code Division Multiple Access (WCDMA) such as High Speed Packet Access (HSPA) [1] and Evolved HSPA (HSPA+). The latter streamlines the evolution of 3G systems towards Long Term Evolution (LTE) [2]. The traffic mix served by these systems consists mostly of internet applications that use Transmission Control Protocol (TCP) for reliable data transfer; thus, the performance of the various TCP variants (Tahoe, Reno, New Reno [3], CUBIC [4], SACK, BIC etc. [5]) over these networks is important from system efficiency and user experience point of view. Thus user perceived Web performance is now the primary metric for modern network services.

In mobile environments, the requirement for efficient TCP operation raises several problems because originally TCP was designed for wired networks where packet losses are caused by congestion, and the probability of a packet loss due to bit errors is extremely low. However in wireless environments packet losses are frequently caused by bit errors occurring on the air interface. Reliable and efficient data transfer requires that TCP is able to dynamically adapt its rate to the available resources, detect data loss and that it is capable to trigger a TCP flow control action [6] in order to retransmit the missing data without unnecessary TCP rate reduction [7], [8].

TCP considers a segment to be lost whenever its Retransmission Timeout (RTO) timer expires or triple duplicate acknowledgments (ACKs) are received for a sequence number lower than that of the segment in question.

Congestion control in TCP is implemented through the use of two state variables; slow start threshold and congestion window. [6]

The RTO timer expiry (TCP timeout event) can be an indication of heavy congestion in the network, thus the TCP sender resets its congestion window (cwnd) to one segment size and starts increasing it according to slow start algorithm, while continues data transmission from the first unacknowledged segment until it reaches the one half of what its value was before the timeout occurred. After this point it enters the congestion avoidance phase.

Fast retransmit is triggered when three successive duplicate acknowledgements (DUPACKs) for the same sequence number have been received, i.e., without waiting for the retransmission timer to expire [6]. After a DUPACK-based retransmission, the TCP sender halves its load, and immediately enters the congestion avoidance phase.

If the cause of the packet loss is not the congestion in the wireless network than there are problems with the algorithms of TCP congestion control which will make TCP non-optimal for transferring data over mobile networks. Due to the exponential TCP back off algorithm [9], relatively long lasting coverage problems and repeated timeouts will result in long intervals between two TCP retransmission. Another corner case is the spurious fast retransmit and timeout which occur when packets get re-ordered beyond the DUPACK-Threshold in the network before reaching the receiver.

From user experience point of view, this causes increased latency and unresponsive data applications [11].

In response to increased web latency, some large Web service providers have made major structural changes to their service delivery infrastructure thus the latency due to TCP's design limits further improvement [11].

The performance improvement of the TCP protocol over wireless links has been a subject of intensive research in recent years. Proposals can be grouped into two categories: end-to-end and network supported solutions. End-to-end solutions require changes in the client and server TCP implementations, whereas network supported solutions require extra functionality at the underlying network. Important examples of end-to-end solutions are TCP Eifel [17] and TCP Westwood [18], TCP reactive, TCP proactive [11]. The network supported solutions are commonly referred as Performance Enhancing Proxies (PEP) [16].

In this paper the analyzed method for TCP performance improvement is called TCP freeze. TCP freeze is a possible solution to improve short-term degradations on the air interface of radio access systems by triggering the persist mode of the TCP sender, preventing TCP timeouts and rate reductions during poor radio conditions. This paper investigates the possible performance gain provided by a network side TCP freeze mechanism implemented in the radio access nodes such as LTE evolved NodeBs (eNB). The potential of the TCP freeze was investigated by simulations in an LTE network that had various coverage holes with poor radio signal conditions.

The rest of this paper is organized as follows. Section II provides an overview of existing TCP optimization methods. Section III presents the existing freeze related TCP improvements. In Section IV, the architecture and the details of the eNB based TCP freeze solution are explained. Section V shows the results of the performance evaluation of the TCP freeze solution. Finally, Section VI concludes the paper.

II. DECREASING WEB LATENCY BY FASTER LOSS RECOVERY METHODS

A. End-to-end supported solutions

TCP Eifel eliminates the retransmission ambiguity, thereby solving the problems caused by spurious timeouts and spurious fast retransmits. It can be incrementally deployed as it is backwards compatible and does not change TCP's congestion control semantics. [17]

TCP Westwood is a sender-side modification of the TCP congestion window algorithm. An important distinguishing feature of TCP Westwood with respect to previous wireless TCP "extensions" is that it does not require in section and/or interception of TCP packets at intermediate (proxy) nodes. Rather, it fully complies with the end-to-end TCP design principle. The key innovative idea is to continuously measure at the TCP source the rate of the connection by monitoring the rate of returning ACKs. The estimate is then used to compute congestion window and slow start threshold after a congestion episode, that is, after three duplicate acknowledgments or after a timeout. [18]

TCP reactive algorithm is a technique to mitigate RTOs that occur due to tail losses. It requires only sender-side changes and does not require any TCP options. It is designed for senders with Selective Acknowledgment (SACK) enabled because the SACK feedback of the last packet allows senders to infer whether any tail segments were lost.

The TCP reactive algorithm allows a sender to quickly detect tail losses without waiting for an RTO¹ [11].

TCP proactive takes the aggressive stance of proactively transmitting multiple copies of each TCP segment. The sender marks the copied segments by setting a flag in the TCP header using one of the reserved but unused bits. Sending duplicate segments can potentially increase congestion and consequently decrease goodput and increase response time. TCP proactive is designed only for latency-sensitive services on networks where these services occupy a small percentage of the total traffic. [11].

B. Network supported solutions

The network supported solutions are commonly referred as Performance Enhancing Proxies (PEP) [16].

The most prominent implementation of the PEP concept is the split connection TCP proxy, which divides traversing TCP connections into two independent connections, while being transparent to the end-points. The TCP proxy can improve end-to-end performance of a TCP connection in two ways. On one hand a proxy can separate the transmission path along network boundaries with significantly different characteristics and quarantine transmission problems in the affected loop. A proxy deployed at the boundary of the wired and wireless part of a network can isolate the impacts of the air interface problems (i.e., TCP congestion control problems) from the rest of the system and handle these locally.

III. TCP FREEZE RELATED IMPROVEMENTS

A possible way to prevent the TCP performance degradation in case of air interface problems is to "freeze" [10] the TCP senders whenever transmission problems are detected or predicted and resume the operation of the TCP when the system is able to operate efficiently again.

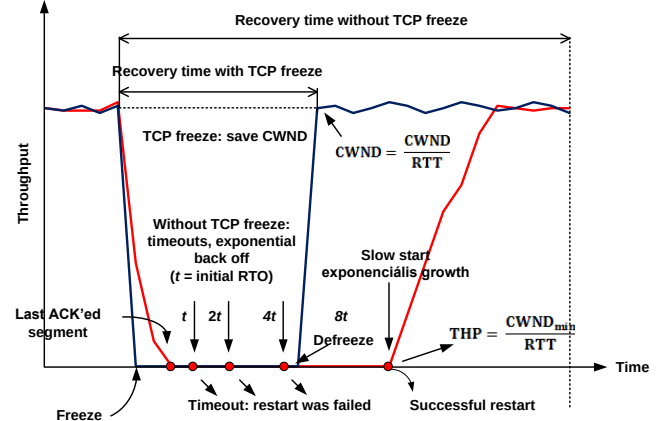


Figure 1. The possible gain which can be reached during poor radio conditions by using TCP freeze

A TCP sender is frozen if it receives an ACK segment with zero advertised window size (Zero Window Advertisement, ZWA). The ZWA forces the TCP sender to suspend all retransmission timers and enter persist mode, which lasts until the source receives a Non-Zero Window Advertisement (NZWA) segment, i.e., a TCP ACK with positive advertised window size. In between, the TCP sender does not send any new data and the RTO timer will not expire even if the persist mode lasts much longer than the RTO. When the NZWA is received, the data transmission continues with the same cwnd size (i.e., at the same sending rate) as before the freeze. Therefore, a timely freeze will minimize the probability of a TCP slow start caused by poor radio channel. The possible reachable gain of using TCP freeze is illustrated on Figure 1.

IV. THE ENB SIDE TCP FREEZE

The architecture of the eNB based freeze implementation is shown in Figure 2. The scope of the functionality, referred to as Freeze Agent (FA) here after, is to detect and/or predict the cases when a given TCP flow might experience performance degradation due to poor radio channel conditions and to freeze the TCP flow in order to prevent excessive performance degradation. The FA may handle both uplink and downlink TCP connections.

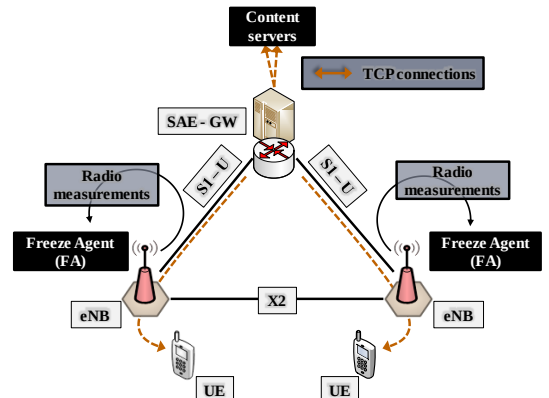


Figure 2. System architecture of the eNB based TCP freeze

The FA has two functional entities as depicted in Figure 3: the detection/prediction entity, which monitors events that may trigger the freeze/defreeze of a TCP connection, and the ACK management entity, which follows each TCP connection subject to the freeze mechanism and intercepts/processes its ACKs traversing through the FA by either forwarding them without modification or making the conversion to ZWA as well as sending an NZWA on behalf of the receiver if needed.

The detection and prediction require that the FA collects relevant radio network measurements, e.g., the Channel Quality Indicator (CQI) representing downlink radio signal quality in HSPA and LTE, measurements of the Signal to Interference and Noise Ratio (SINR) in uplink, retransmissions requested by the Hybrid Automatic Repeat

¹ Tail loss: either drops at the tail end of transactions or a loss of an entire window of data or acknowledgments

Request (HARQ) mechanism of the LTE radio stack, air interface load, etc. and handover management information.

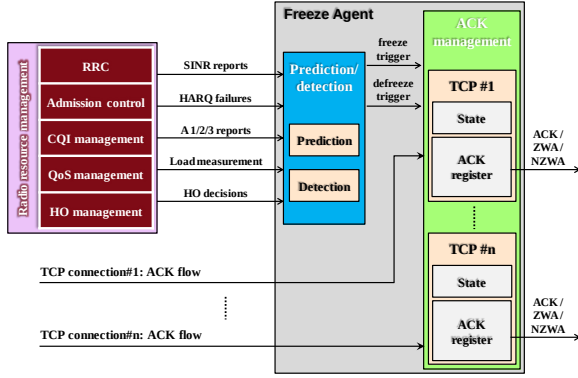


Figure 3. Internal architecture of the Freeze Agent

Based on these information, the relevant events that can lead to performance degradation can be detected or predictions can be made. These events are the freeze triggers.

The ACK management entity maintains two pieces of information for each TCP connection: its state, which is initialized to “normal” at connection setup, and an ACK register (that stores one ACK), into which it copies the latest ACK of the corresponding connection that traversed the node at which the FA is deployed. This operation is transparent to the connections as in normal state each received ACK is copied into the ACK register corresponding to the connection and forwarded unchanged by the system. [15]

Operation of the FA:

A. When a freeze trigger is received from the detection/prediction entity:

1. The state of the related TCP connection is set to “freeze”
2. The stored ACK of the connection is resent with its advertised window set to zero (ZWA)
3. The TCP sender enters persist mode: freezes the RTO timer and the cwnd size and it stops sending new data
4. Further ACKs may still be sent by the TCP receiver as a response to segments sent by the source before it has received the ZWA. These ACKs are also copied to the ACK register, so the content of the register is updated after a new ACK is detected by replacing the stored ACK with the new one and forwarded with advertised window set to zero
5. The FA keeps monitoring the relevant measurements in order to decide if the freeze state should be maintained or terminated

B. When the coverage problem is over:

1. A defreeze trigger is generated that causes the transition of the connection to “normal” state
2. The ACK management entity resends the stored ACK with nonzero advertised window size (NZWA)
3. The TCP sender leaves persist mode and starts transmitting new data [15]

V. PERFORMANCE EVALUATION

The performance of the TCP freeze mechanism in case of LTE with coverage holes was evaluated with a packet level simulator, which implements the S1-U interface, the LTE Uu interface and the X2 interface protocols in detail. The transport topology is shown on Figure 4. A Proportional Fair, Required Activity Detection [2] scheduler is assumed for air interface scheduling; the air interface model was based on [13] with a 10MHz bandwidth single input single output channel. The channel quality was simulated separately for each UE considering distance loss, shadow fading, fast fading and penetration loss.

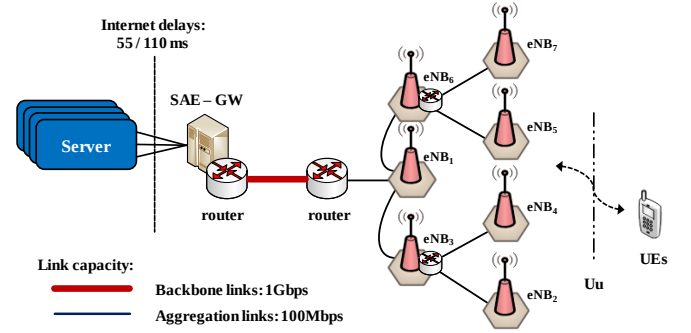


Figure 4. Simulation topology

Downlink inter-cell interference is calculated from the propagation loss, momentary transmit power and resource allocation of the interfering eNBs. Handover procedures are implemented; the handover decisions were based on the reported A3 radio measurements.

The simulated double ring radio topology consisted of seven active and twelve interference generator eNBs with 500m inter-site distance, each of them having a single cell setup. The transport topology is shown in Figure 4. In the beginning of the simulation, 30 users were randomly placed in each cell. During the 600 s simulation time, the users were moving according to the random way point mobility model with an average speed of 3km/h. Each user had an FTP application performing file downloads. The so-called Internet delay between the SAE-GW and the content servers was set to either 55ms or 110ms. The TCP variant was New Reno with the minimum RTO timer value set to 0.25 s or 1 s. In each active cell, there was a coverage hole at a random location where the shadowing was increased to the extent so that the received signal level of any of the cells was unsuitable for data transmission. The size of the coverage hole was either 5 X 5 m or 5 X 10 m. In case a user's trajectory with an active TCP connection lead through a coverage hole, its data download was abruptly interrupted since the data buffered in the serving eNB in downlink could not reach the UE anymore and also the ACKs could not be sent in uplink by the UE. In the reference case when there was no freeze mechanism in the eNBs, this caused multiple TCP timeouts and resulted in long recovery times after the user moved out of the coverage hole. When the freeze functionality was switched on, the FA in each eNB monitored the A2 and A1 events reported by the UEs. Both of these events were configured in the UEs by the eNB with a threshold of -5 dB and a hysteresis of 0.5 dB so that a UE sent an A2 report to the eNB whenever the RSRQ of the serving cell dropped below -5.5 dB (taken as a freeze trigger by the FA) and sent an A1 report if the signal level recovered above -4.5 dB (defreeze trigger). Both events were also configured with a 320 ms time-to-trigger value so the triggering condition should be continuously true for at least that long before the A2/A1 report would actually be sent by the UE.

Each simulation setup (Internet delay, RTO, coverage hole size) was run with 20 different random seeds both with and without the freeze mechanism.

In each simulation, the occasions when a user with an active file download crossed a coverage hole were monitored and the time required to download a given amount of data (100 kB, 250 kB, 500 kB and 1 MB) after moving out of coverage hole was measured.

According to the simulation results Figure 5. shows that the download time of 1 MB data was less (15 %) in case of TCP freeze than without using TCP freeze independently of the size of the coverage hole and the trajectory of the users.

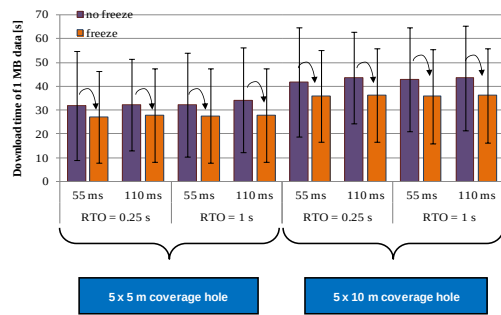


Figure 5. Download time of 1 MB data

Figure 6. and Figure 7. show that the download time gain was 2-6 sec when there were coverage holes with size 5 X 5 m and there was 4-7 sec download time gain in the other case.

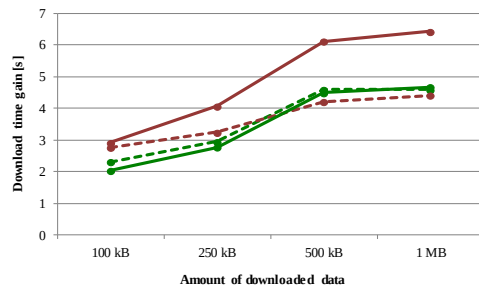


Figure 6. Download time gain of the TCP freeze by 5 X 5 m coverage hole

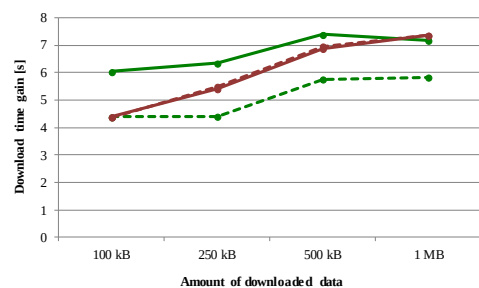


Figure 7. Download time gain of the TCP freeze by 5 X 10 m coverage hole

The time that the TCP connections spent in outage (i.e., without transmitting any data) due to the coverage holes was also measured as shown in Figure 8.

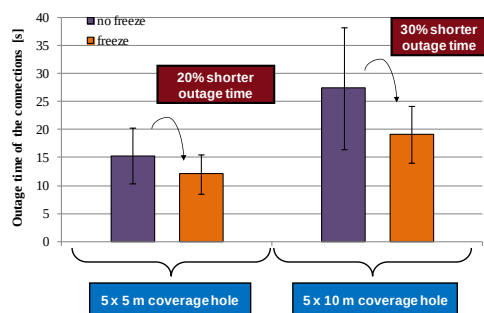


Figure 8. Time spent in outage by the TCP connections

When the TCP freeze was activated, the outage time of a connection was confined to the time during which it was in freeze state, which resulted in 20–30% shorter outage time on average (depending on the size of the coverage hole and the trajectory of the users) compared to the reference case with no TCP freeze. This is due to the multiple timeouts that occurred in the latter case, which resulted in exponentially increased RTO values during which the TCP sender made no attempt to restart the data transfer. On the other hand, the TCP freeze immediately restarted the TCP sender via the NZWA when the A1 trigger was reported by the UE.

VI. CONCLUSION

This paper provided an insight to the issues related to the TCP performance improvement over wireless links through

TCP reactive, TCP proactive and the TCP freeze mechanisms. A network side TCP freeze mechanism was described in detail. The performance of the mechanism in case of an example LTE network with coverage holes was investigated with simulations. Results show that when implemented in the radio access nodes, this functionality is suitable for monitoring radio conditions and minimizing the impact of temporary coverage problems caused by coverage holes. The freeze is able to suspend the data transmission of the TCP connections by triggering the persist mode of the TCP sender; consequently, when restarted, the transmission rate of the sender is not compromised by increased round-trip times or timeouts due to the temporary failure of connectivity between the receiver and the sender.

The outage time experienced by users with active data applications can be reduced significantly by the TCP freeze mechanism, that is, the freeze improves the responsiveness of the system. In addition to the improved interactivity, the freeze also provides higher overall throughput for the applications.

REFERENCES

- [1] H. Holma and A. Toskala, Eds., HSDPA/HSUPA for UMTS. John Wiley & Sons, 2006.
- [2] H. Holma and A. Toskala, Eds., LTE for UMTS. John Wiley & Sons, 2009.
- [3] S. Floyd, T. Henderson, and A. Gurtov, "The New Reno Modification to TCP's Fast Recovery Algorithm," RFC3782 (Proposed Standard), Internet Engineering Task Force, Apr. 2004.
- [4] S. Ha, J. Rhee, and L. Xu, "CUBIC: a new TCP-friendly high-speed TCP variant," SIGOPS Oper. Syst. Rev., vol. 42, no. 5, pp. 64–74, Jul. 2008.
- [5] Md. Shohidul Islam, M.A. Kashem, W. H. Sadid, M. A. Rahman, M.N. Islam, S. Anam "TCP Variants and Network Parameters: A Comprehensive Performance Analysis", Proceedings of the International MultiConference of Engineers and Computer Scientists 2009 Vol I IMECS 2009, March 18 - 20, 2009, Hong Kong
- [6] M. Allman, V. Paxson, and E. Blanton, "TCP Congestion Control," RFC 5681 (Draft Standard), Internet Engineering Task Force, Sep. 2009.
- [7] Y. Tian, K. Xu, and N. Ansari, "TCP in wireless environments: problems and solutions," Communications Magazine, IEEE, vol. 43, no. 3, pp. S27–S32, Mar. 2005.
- [8] H. Elaarag, "Improving TCP performance over mobile networks," ACM Comput. Surv., vol. 34, no. 3, pp. 357–374, Sep. 2002.
- [9] V. Paxson, M. Allman, J. Chu, and M. Sargent, "Computing TCP's Retransmission Timer," RFC 6298 (Proposed Standard), Internet Engineering Task Force, Jun. 2011.
- [10] T. Goff, J. Moronski, D. Phatak, and V. Gupta, "Freeze-TCP: a true end-to-end TCP enhancement mechanism for mobile environments," in INFOCOM 2000. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE, vol. 3, Mar. 2000, pp. 1537–1545.
- [11] T. Flach, N. Dukkipati, A. Terzis, B. Raghavan, et. al., "Reducing Web Latency: the Virtue of Gentle Aggression" ACM Conference of the Special Interest Group on Data Communication (SIGCOMM '13), ACM (2013)
- [12] E. Blanton, M. Allman, L. Wang, I. Jarvinen, M. Kojo, and Y. Nishida. A Conservative Loss Recovery Algorithm Based on Selective Acknowledgment (SACK) for TCP, 2012. RFC6675.
- [13] —, "Physical layer aspect for evolved Universal Terrestrial Radio Access (UTRA)," 3rd Generation Partnership Project (3GPP), TR25.814, Oct. 2006.
- [14]
- [15] Cs. Deák, Á. Drozdy, P. Szilágyi, Z. Vincze and Cs. Vulkán, "TCP Performance Improvement in Mobile Networks with Coverage Problems", Communications (ICC), 2013 IEEE International Conference, 9-13 June 2013.
- [16] V. Farkas, B. Héder, Sz. Nováczki, "A Split Connection TCP Proxy in LTE Networks", 18th EUNICE / IFIPWG 6.2, 6.6. International Conference, EUNICE 2012 Budapest, Hungary, August 29-31, 2012 Proceedings
- [17] Ludwig, R., Katz, R. H. "The Eifel algorithm: making TCP robust against spurious retransmissions", SIGCOMM Comput. Commun. Rev. 30(1), 30–36 (2000)
- [18] Mascolo, S., Casetti, C., Gerla, M., Sanadidi, M.Y., Wang, R. "TCP westwood: Bandwidth estimation for enhanced transport over wireless links", In: MOBICOM, pp. 287–297 (2001)

Thermal characterization of System on Chip packages

Based on Mixed Signal Boundary Scan Method

Máté Meződi

Department of Electronic Devices
Budapest University of Technology and Economics
Budapest, Hungary
mezodi.mate@gmail.com

Péter Onódy

Staff Design Engineer
Silicon Laboratories
Budapest, Hungary
peter.onody@silabs.com

Gábor Takács

Department of Electronic Devices
Budapest University of Technology and Economics
Budapest, Hungary
takacs@eet.bme.hu

Abstract— Thanks to the increasing rate of integration more and more circuit blocks can be implemented into a single silicon die in case of SoC (System on Chip) systems. The elements of the circuit are in a strong thermal coupling because of the common substrate they share, thus it is very important to map and investigate the thermal conducting paths in order to increase the reliability.

The work introduces a novel testing method which makes possible the thermal characterization of already manufactured and packaged mixed-signal circuits.

Keywords—component;band-gap, temperature sensor, SOC, PTAT, thermal characterization, mixed signal boundary scan

I. INTRODUCTION

Gordon Moore's statement made in the 1960-s alleges that the number of elements which can be integrated into a single silicon die doubles in every one and a half year. This statement is still accurate.

Thanks to the gradual decrease of size and the integration which leads beyond Moore's law more and more circuit blocks can be implemented into a single silicon die. A claim to integrate the subsystems of a whole electronic system into a single package occurred in the twentieth century.

This innovation led to the invention of SoC (System on Chip) systems. SoC systems represent one cost effective and efficient way to create complex, multifunction products [1]. Besides digital units the SoC systems consists of analog and RF (radio frequency) circuits, memories and MEMS (Micro Electro Mechanical System) structures.

Integrating these systems into a single package comes with a lot of difficulties and obstacles such as the technological incompatibility and the increase of design complexity.

The operation of the electronic devices depends on the temperature, especially in case of analog systems it is typical to optimize the operation to room temperature (27°C). This is the reason why the decrease of operating parameters can be experienced at different temperatures. This problem is less critical in case of digital systems.

The increase of the integration rate comes together with the increase of dissipation density. In some cases it reaches 100W/m² [2]. Thus it is not surprising that the dissipated thermal energy is responsible for the malfunctions of today's electronic devices in the greatest proportion.

In order to avoid such failures the heat generated by the dissipation need to be led out of the circuit on the shortest and smallest thermal resistant way. This is the reason why the packaging and the attach of the silicon die plays a key role for the sake of thermal conductance. With the increase of the complexity of the devices and packages new measurements and characterization methods are necessary to qualify thermal conductive paths inside and outside the package.

The purpose of this paper is to make this qualification possible for an arbitrary SoC system. One circuit of the chip shall serve as the trigger of the thermal dissipation. A design method will be introduced to realize a well scalable temperature sensor which can be integrated into a chip. This sensor is able to communicate with the external thermal transient tester module, which was invented to map the thermal conducting paths. The communication will be granted by the mixed signal boundary scan.

II. THERMAL TRANSIENT TESTING

A. The measurement

Thermal transient testing is a measurement method which is capable of mapping the thermal conducting paths inside and outside the package. The progress of the characterization: a stationary state system is thermally triggered by a step signal. By storing the transient response until the system reaches the new quasi stationer state the thermal resistance of the system can be determined from the temperature difference and dissipated power.

By measuring and storing the changes of the temperature until the thermal equilibrium the temperature-time response function of the system can be determined. Then the cumulative

structure function can be generated with the NID method [3]. This function gives the thermal resistance and conductance of the materials which can be found in the one dimensional thermal conducting path, so it describes the path on an explicit way. Degradations of the path can be determined from the step sizes of this function.

B. Thermal transient tester

The thermal transient tester is the product of the Mentor Graphics Company, which is capable of the thermal characterization of packaged integrated circuits.

The thermal transient tester is able to sense small voltage changes, the value of the least significant bit of its analog-digital converter is 25μV on the 100mV measurement range. Thus a typical pn junction diode with the thermal coefficient of -2mV/°C can reach 0.01°C resolution [4]. Although in order to properly qualify the thermal conducting path higher resolution may be necessary. Therefore instead of regular junction diode a temperature sensor with high sensitivity need to be used.

The long term purpose is to create a circuit which can be integrated into the device under the test. The output of this circuit shall be reachable from outside the package through a standard interface and the output level shall correspond to the input level of the thermal transient tester. This would mean a new thermal qualification method for SoC systems.

In order to create this corresponding, adequate quality input signal a sensor with high sensitivity and accuracy is necessary. The ideal characteristic would be a linear dependency on the temperature. Although reaching the straight line is not possible, but a slightly linear curve makes the characterization easier, so the output is intended to be as linear as possible hereinafter. The main objective is designing this sensor. In the paper mainly these design steps will be presented.

C. Mixed signal boundary scan

The IEEE 1149.4 standard. The aim is to grant a measurement environment in order to carry out analog measurements on chips attached to printed circuit boards. The printed circuit board has to be JTAG compatible, this grants two analog ports which are connected to the analog outputs of the chip [5].

This standard is widely used and effective method to examine analog and digital signals of integrated circuits. The standard interface is capable of communication with the thermal transient tester. The maximum voltage of the interface is the power supply, so this doesn't limit the output range of the sensor. The digital ports of the interface can be used to trigger the measurement by turning on the subsystems of the SoC system. The IEEE 1149.4 standard fulfills all the requirements which are necessary for this measurement.

D. The environment of the measurement

The elements of the environment are a packaged SoC system, basically the device under test with an integrated temperature sensor. This package is attached to a IEEE 1149.4 compatible printed circuit board. The analog output port of the board is connected to the thermal transient tester. The SoC system can be triggered from outside of the package through the digital ports of the board.

III. THE DESIGN OF THE TEMPERATURE SENSOR

A. The method of the temperature measurement

The aim is to create an analog integrated circuit which can provide an output signal from which with the help of an analog-digital converter the temperature of the circuit can be determined. The chosen temperature range is the military standard from -55°C to 125°C

At the Department of Electronic Devices of Budapest University of Technology and Economics the available technology is 0.35μm CMOS technology provided by the Austria Microsystems company (AMS). The circuit was designed with the Cadence 6.1 design software.

Accurate design with the whole, detailed inner structures of analog integrated industrial sensors are not available, because they are the properties of the producing companies. Thus their already packaged version with the output data can serve as the base of the design and the comparison [6]. There are digital output sensors available, but in this case analog output is adequate, because the thermal transient tester has an analog digital converter. The datasheets provide information about the temperature measurement method.

- The first method is based on a thermocouple, although the AMS library doesn't provide the means to design a circuit like this [7].
- The second method obtains temperature information from its pulse with modulated digital output's variation [8]. Although at this state of the measurement analog output is adequate as mentioned earlier.
- The third type's determines the temperature from the voltage difference between diodes biased by different current densities [9].

All of the three types have similar statistic, only slight differences can be noticed. Because of technological and simplicity reasons the third method was chosen.

B. Thermal voltage reference cell

The base of the thermal voltage reference cell is a regular pn junction diode. The voltage of the diode with respect of the bias current can be written as [10].

$$U = \frac{k \cdot T}{q} \cdot \ln\left(\frac{I}{I_0(T)}\right) \quad (1)$$

Where k is the Boltzmann constant, q is the charge of an electron and T is the absolute temperature. It is a simple model, the point is that thermal coefficient is not constant because of the saturation current's (I_0) temperature dependence. Although with taking the difference of two diode voltages the saturation current is eliminated from the equation (2). This can be realized by the following structure [10] (Figure 1.).

$$U_{out} = U_1 - U_2 = U_T \ln(n \cdot m) = \frac{k}{q} \ln(n \cdot m) T \quad (2)$$

The variable of the natural logarithm can be set with number of diodes (n) and the ratio of the bias currents (m).

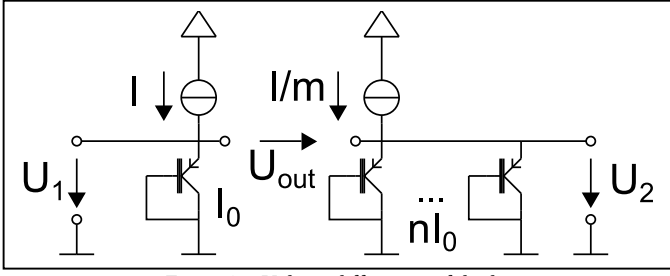


Figure 1. Voltage differences of diodes

The voltage difference of the diodes depends thermally only on the absolute temperature. It is also known as PTAT, e.g. proportional to absolute temperature. Although the output won't be perfectly linear, because secondary effect are neglected from this diode model. In order to ensure the correct operation of the cell the collector voltage of the bipolar transistor need to be equal. There are two main possibilities to ensure this, with an operational amplifier or a current mirror. Obviously the former choice is more stable.

The input of the thermal transient tester accepts only asymmetric signals, so the differential output need to be converted to current with the help of a resistor. The current can be mirrored out from the cell and can be converted back to voltage with another resistor. If the two resistors are made of the same material, then their thermal dependence will be eliminated from the voltage. Similarly to the voltage the current is PTAT as well.

The main problem of such a structure is that zero output voltage is at -273°C . I intend to maximize the sensitivity of the output, therefore it would be beneficial, if the zero point was at -55°C . This can be achieved by extracting a thermal independent current from the output current. Although to create such an independent current a reference source is necessary.

C. Temperature independent voltage source

Temperature independent voltage sources are created from thermal voltage reference cells. Both of the temperature sensor and the temperature independent voltage source require a thermal voltage reference cell [11]. It is possible to realize the sensor and reference source with using only one reference cell with the next structure (Figure 2.) [10].

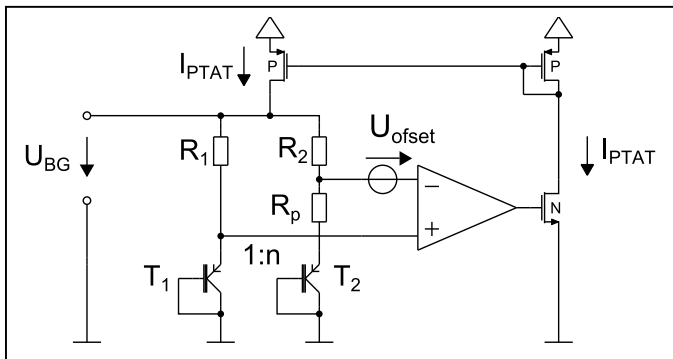


Figure 2. The band-gap circuit

Such a circuit is called band-gap reference circuit. The voltage U_{BG} is slightly dependent on the temperature. It can be expressed as:

$$U_{BG} = U_{EB} + (V_T \cdot \ln(n) + U_{offset}) \cdot (1 + \frac{R}{R_p}) \quad (3)$$

Where U_{EB} is the emitter-base voltage of T_2 bipolar transistor. The optimal choice for the ratio of R_1 and R_2 is equality, so R in the latter equation indicates the value of $R_1=R_2=R$. For the ratio of the diodes 8 was chosen, because of symmetry and layout design reasons. Another important question is the relation of power consumption and area. This can be determined from next equation (4).

$$I_{PTAT} = \frac{2U_T \ln(n)}{R_p} \quad (4)$$

The design software can estimate the area of the circuit. In order to find a fair compromise several simulations were made. The circuit hasn't been optimized at this state yet. The only objective was to determine how the consumption changes with the increase of the circuit area. These simulations indicated, that it was advantageous to increase the consumption until $50\mu\text{A}$, because it greatly decreases the area. However over the consumption of $100\mu\text{A}$ it is necessary to use unrealistic transistor sizes. The consumption of $50\mu\text{A}$ was chosen. The design software estimated $180 \times 180 (\mu\text{m})^2$ for the area of the circuit at 80% coverage.

D. The output circuit

The band-gap voltage (the output of the reference cell, U_{BG}) can be converted into current with a resistor. A follower amplifier is required to isolate the reference voltage, because a simple parallel resistor would degrade the operation of the thermal voltage reference cell. T_2 is necessary to isolate R_{out} because its voltage greatly varies with the temperature. The temperature dependence of R_{out} is also eliminated if all of the resistors are made of the same material (see equation (5), α denotes the thermal coefficient of the resistor).

$$U_{out} = \alpha R_{out} \left(\frac{2U_T \ln(n)}{\alpha R_p} - \frac{U_{BG}}{\alpha R_{BG}} \right) = R_{out} (I_{PTAT} - I_{BG}) \quad (5)$$

The whole output structure can be seen on next picture (Figure 3).

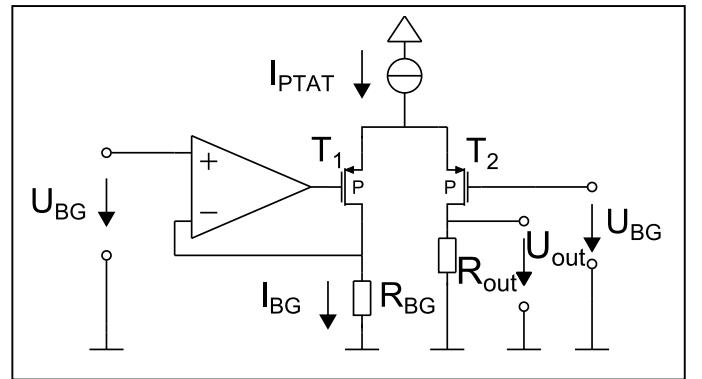


Figure 3. Output circuit

E. Accuracy

One of the most important parts of the design was the minimizing of distortion. The input offset of the thermal voltage reference cell's amplifier significantly influences the distortion of the I_{PTAT} current and the band-gap voltage. In order to

minimize such an effect, serious distortion considerations had been taken. As the result of design steps the distortion of the band-gap voltage was decreased to 1.3% in case of 3σ distortion. The curves are the results of Monte Carlo sweep simulation with 100 samples.

The thermal coefficient of the output voltage is about $10\text{mV}/^\circ\text{C}$. The maximum value of the distortion of the output voltage is 100mV on 125°C . This means 5.3% distortion at the upper end of the temperature range. Going downwards it keeps increasing. This is the reason why measuring the temperature with respect to the average of the result doesn't give utilizable information.

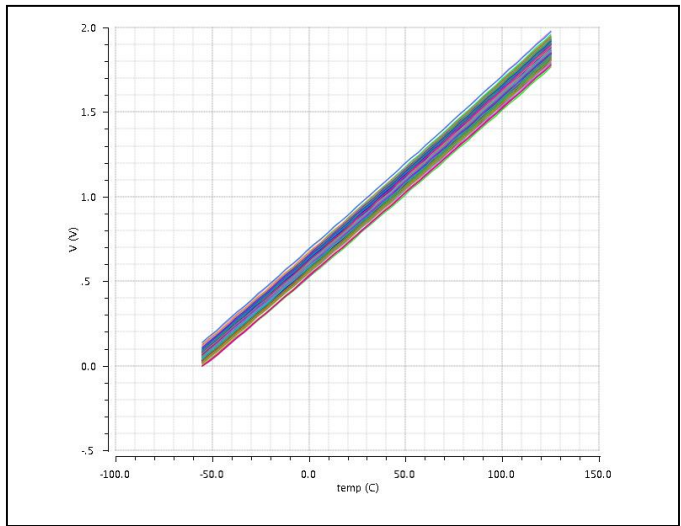


Figure 4. Monte Carlo simulation of the output voltage for 100 samples

To compare the temperature sensor of the Dallas Semiconductor company has a 12 bits digital output and value of the least significant bits is 0.0625°C . The maximum error is 0.8°C according to the datasheet [12]. The voltage change for the LSB can be estimated as $25\mu\text{V}$ assuming the 70mV PTAT voltage change. In contrast the current structure can show 0.0025°C temperature change in case of $25\mu\text{V}$ resolution. So the sensitivity increased to 25 times, the error to 11.25 times.

Although the nature of the error is offset failure. So the error can be greatly decreased with characterization. In case of taking ten samples of the output simulation and fitting a linear line to each of these values at the temperatures of 27°C and 42°C , the maximum error voltage is 17mV (Figure 6). 10mV error is equal to around 1°C . From -10°C to 90°C the error is less than 0.5°C , which corresponds to the error of industrial sensors [7].

A lot of technical papers also have been investigated in the literature in order to compare, although their main objective is quite different from this one's. Their output is digital and they focus on the accuracy of the temperature, not sensitivity. They do not give the accurate data about the analog output of the temperature measurement. Besides the published data are based on the information gathered from already manufactured circuits. So comparing to those sensors would be irrelevant.

I would like to emphasize, that these result are based on simulations. I could not guarantee their reliability in case of

realization. Although I have to mention, that the simulator works with complex mathematical models and gives good estimation of reality.

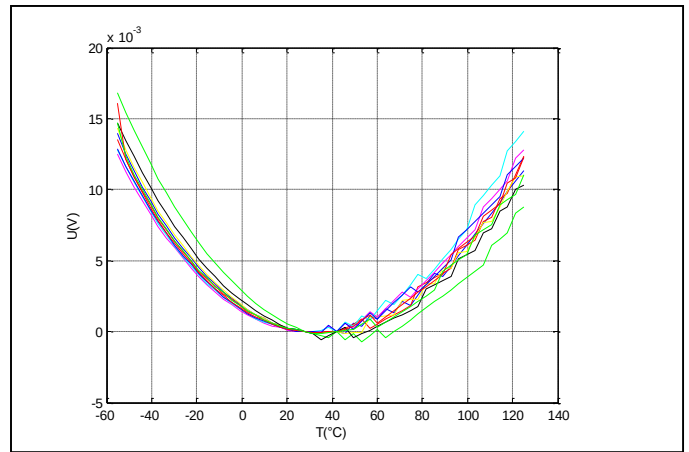


Figure 5. Error voltage after the characterization

IV. SUMMARY

As the result of the work a highly sensitive and linear thermal sensor has been designed which in case of realization in an actual circuit is capable of sensing the slight temperature changes of the silicon. The output of the sensor can be reached from outside with the mixed signal boundary scan interface. The thermal transient tester can map the thermal conducting path based on the information obtained from the sensor.

REFERENCES

- [1] Vijay K. Madiseti, "Electronic System, Platform and Package Codesign," IEEE Design & Test of Computers, Platform-centric codesign Georgia Institute of Technology, May-June 2006.
- [2] György Bognár, "A mikroelektronika egyes termikus problémáinak kezelése", Ph.D Dissertation, Budapest, 2009.
- [3] Vladimir Székely, "Identification of RC Networks by Deconvolution, Chances and Limits," IEEE Transactions on Circuits and Systems-I. Theory and applications, 1998
- [4] „Mentor Graphics - T3Ster - Transient Thermal Characterization,” [Online]. Available: http://s3.mentor.com/public_documents/datasheet/products/mechanical/products/t3ster-technical-info.pdf.
- [5] „Expanding boundary-scan to analog and mixed-signal testing: IEEE 1149.4,” JTAG Technologies, 2000. [Online]. Available: <http://www.jtag.com/en/content/ieee-11494>.
- [6] analog.com.\MEMS and Sensors\Temperature Sensors [Online]. Available: <http://www.analog.com/en/mems-sensors/digital-temperature-sensors/products/index.html>
- [7] Analog Devices, Precision Thermocouple Amplifiers with Cold Junction Compensation, AD8494/AD8495/AD8496/AD8497 Data sheet
- [8] Analog Devices, $\pm 0.5^\circ\text{C}$ Accurate PWM Temperature Sensor in 5-Lead SC-70, TMP05/TMP06, Data sheet
- [9] Analog Devices, 2-Terminal IC Temperature Transduce, AD590 Data sheet
- [10] B. Razavi, Design of Analog CMOS Integrated Circuits, McGRAW-HILL, 2001
- [11] P. Brokaw, A Transistor Voltage Reference, and What the Band-Gap Has To Do With It. [Performance]. 1989. Available: <https://archive.org/details/APaulBro1989>
- [12] Dallas Semiconductor, Maxim, High-Precision 3-Wire Digital Thermometer and Thermostat, DS1626/DS1726 Data sheet

Experiments with positioning algorithms for PET detectors

András Kufcsák

Budapest University of Technology and Economics

Budapest, Hungary

Email: kufcsak.andras@upcmail.hu

Abstract— In nuclear medicine, PET scanners equipped with solid state photodetectors form an active research area. A key task in this imaging is to determine the lateral position where gamma photons from the examined body impinge on the detector surface (point of interest, POI). Acquiring digital output from the solid state detector requires novel concepts and methods for this position estimation compared to conventional photomultipliers. On the other hand, exact and detailed system parameters and constructional properties draw limitations on the possible methods in multiple criteria such as accuracy, robustness, or computational complexity and time. In this study I revealed different positioning methods and compared them along different viewpoints. I also implemented the proposed solution on an FPGA connected to sensors established in the SPADnet project.

Keywords- PET, digital SiPM, positioning algorithm, subpixel resolution

I. INTRODUCTION

A. Basics of PET scanners

Positron emission tomography (PET) is an in vivo medical imaging method capable of producing three dimensional images of the examined body. The method discovers the functional condition of the body and enables the detection of several dysfunctions at an early stage. It has been successfully applied in different fields such as oncology, cardiology and neurology.

The imaging process is based on the detection of gamma photons originated from the reactions of a radioactive tracer and the tissue of the patient. The position information where gamma photons impinge on the surface of the detector ring around the patient makes it possible to determine the concentration distribution of the tracer inside the body. The detectors for gamma photons consist of scintillation crystals which yield light proportional to the energy of the absorbed gamma photon. Light sensors are used to detect the light of the crystals and the data of these sensors are used to determine the information of the absorbed gamma photon, namely the position, the time of the arrival, and the absorbed energy [1]. Traditionally the applied light sensors were photomultiplier tubes (PMT), however state-of-the-art solutions apply solid state photon counters [2].

B. Digital silicon photomultiplier (SiPM)

The digital SiPM is a novel type of solid state light sensor. Its utilization for scintillation light detection in PET application

has been proving promising results [3]. The benefit of SiPM-s compared to other sensors, like position sensitive photomultiplier-tubes (PS-PMT) is the improved resolution and the ability to operate in magnetic fields, opening the possibility of constructing hybrid PET-MRI scanners.

The SiPM utilizes single-photon avalanche diodes (SPAD). These are diodes which are reverse-biased over the breakdown voltage. SPADs are capable of detecting an incident light radiation through the avalanche current triggered by a photo generated carrier. Digital SiPMs are partitioned into an array of pixels consisting of a number of SPADs. The device takes care of quenching the avalanche current of the SPADs and so making them sensitive again. The digital concept makes it possible to establish further functions on the same substrate (e.g. handling dark current) and as a result the digital SiPM provides preprocessed digital numbers as time of scintillation and the count of collected light photons for each pixel [4][5].

C. Scintillator crystals

The output data of a certain light sensor is dependent not only on the sensor parameters but also on the used scintillator geometry. There are two mainstream detector construction concepts regarding the applied scintillator. One is to use a single crystal slab attached to the light sensor. In this case the scintillation light of the absorbed gamma photon spreads in the whole crystal slab above the sensor before reaching the sensor pixels.

The resolution in this case is mainly depending on the applied position estimation algorithm. The resultant resolution of the PET scanner is affected in several ways at different levels of the image reconstruction process [1]. These error sources bring an inevitable imprecision to the reconstruction process, hence there is no reason for exaggerated accuracy concerning the position estimation. A reasonably high resolution of the position estimation is submillimeter resolution.

In the other case crystal pins are used arranged into a matrix, where the crystal pins are optically insulated from each other. Generally there is too much effort in perfectly matching the crystal pins to pixel edges. Instead a thin glass is optically coupled to the crystal matrix guiding the light to pixels underneath and around the yielding crystal pin.

The range of firing pixels in this case covers a smaller area which can be used to assign the firing crystal pin to the observed light distribution using simple algorithms. Therefore

the obtainable resolution in gamma photon position estimation is limited by the size of crystal pins [6]. This however could be acceptable in certain applications because of other uncertainties arisen in the whole image reconstruction process.

II. DATASET

The dataset used in the examination of different positioning methods was made with simulations of the current sensor chip under development in the SPADnet project [7][8]. Simulations were carried out using the Zemax and Matlab software according to the specification and behavior of the first generation SPADnet1 light sensor [4][5]. The simulations not only took account detailed parameters of the detector construction but also circumstances of a real observation, e.g. the noise of the device.

When detecting a scintillation, the simulated device supplies the number of photon counts integrated for a given time. Hence the simulated data for a single measurement (scintillation at a given point in the crystal) consists of photon counts for each pixel.

The simulation model consisted of a tile built up from 5x5 chips, each of them consisting of 16x16 pixels, and a scintillator crystal slab over the sensor tile. The model was excited in 5x5 certain planar positions and in three depth levels along the z coordinate of the crystal slab for each of the twenty-five planar positions (Figure 1). There were 700 simulated measurements made for each of the resulting 75 simulated POIs.

III. COMPARISON VIEWPOINTS

To compare the efficiency of the potential algorithms, characterizing values have to be set as exact and measurable viewpoints. These are the average (1) and the corrected sample standard deviation (2) of the distance between the estimated position and the real POI. These values express the bias and the square root of corrected sample variance of the position estimator.

$$\bar{d} = \frac{1}{N} \sum_{j=1}^N d_j \quad (1)$$

$$\sigma_d = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (d_i - \bar{d})^2} \quad (2)$$

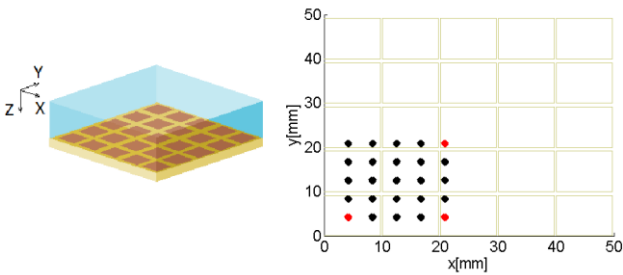


Figure 1. Simulated detector and POIs

$$d_i = \sqrt{(x_i - x_{POI})^2 + (y_i - y_{POI})^2} \quad (3)$$

N denotes here the number of simulated measurements for a given test point, x and y are metric coordinates according to Figure 1.

IV. POSITIONING ALGORITHMS

In the literature of PET detectors different positioning methods have been revealed. Many of these have outstanding accuracy, especially clustering and statistic based algorithms [9][10]. However these methods have a high computational complexity, preventing the possibility of an FPGA implementation and live estimation. Therefore in this paper I investigated simple and quick algorithms.

A. Maximum finding

The maximum of an ideal light distribution without noise would have the same planar position as the gamma photon. Hence the position of the pixel with the highest photon count could be used as position estimator with pixel size resolution. However, in a real measurement the pixel with the maximum photon count could be further away from the real POI because of statistical fluctuations and statistical uncertainty of low light yield of the crystal.

The above stated properties of the light distribution make this type of position estimation method inaccurate when used with slab scintillator based PET modules, however for the sake of further comparisons Table I. shows the bias (1) and the corrected sample standard deviation (2) values for this method. These characteristic values are calculated for three representative planar positions out of the twenty-five planar test points, which are marked with red colour in Figure 1. In each of the three planar positions, all three depth levels of gamma absorption were examined separately ($z_{POI} = 0.5000$ mm, $z_{POI} = 5.0000$ mm and $z_{POI} = 9.5000$ mm).

TABLE I. RESULTS OF THE MAXIMUM FINDING ALGORITHM

Bias of estimation (mm)								
corner			edge			middle		
$z = 0.5$	$z = 5.0$	$z = 9.5$	$z = 0.5$	$z = 5.0$	$z = 9.5$	$z = 0.5$	$z = 5.0$	$z = 9.5$
2.84	1.75	0.32	3.51	1.51	0.22	4.10	1.36	0.16
Corrected sample standard deviation of estimation (mm)								
corner			edge			middle		
$z = 0.5$	$z = 5.0$	$z = 9.5$	$z = 0.5$	$z = 5.0$	$z = 9.5$	$z = 0.5$	$z = 5.0$	$z = 9.5$
1.08	0.97	0.16	1.77	0.88	0.20	2.11	0.86	0.2

Values are rounded to two decimals

The results show that for scintillations in the scintillator crystal close to the sensor surface ($z_{POI} = 9.5$ mm), the estimation is giving better results than further away from the sensor surface. The reason for this is that light photons in this case are more concentrated, so the pixel with the maximum photon count has lower photon noise thus the estimation is more precise. This means that the method results lower bias

and variance for small area of firing pixels. Therefore this method might be acceptable for crystal pin matrices with small crystal pin pitch and light guiding glass of small thickness.

B. Maximum finding with smoothing

The effect of distortion in the light distribution could be reduced by smoothing the surface of pixel counts. This could be done by assigning the average of photon counts of the neighbouring pixels to each pixel, i.e. running a moving average window on the surface. Larger windows cover many pixels and hence the averages will be closer, i.e. the surface will be smoother. On too smooth surfaces multiple pixels will have maximal photon count and this will deteriorate the maximum finding method.

Figure 2 shows the mean distance between estimated and real positions depending on the applied window size. Distance values were averaged for 700 measurement simulations in the three representative planar test positions mentioned before ('corner', 'edge' and 'middle') at each depth level separately. It shows that the bias of the estimation decreases further from the sensor surface (blue and green lines) while having an acceptable increase close to it (red line), using a 15x15 pixel sized window. The bias and the corrected sample standard deviation values are encapsulated in Table II.

TABLE II. RESULTS FOR MAXIMUM FINDING WITH SMOOTHING

Bias of estimation (mm)								
corner			edge			middle		
z = 0.5	z = 5.0	z = 9.5	z = 0.5	z = 5.0	z = 9.5	z = 0.5	z = 5.0	z = 9.5
0.52	0.74	0.73	0.74	0.73	1.08	1.14	0.80	1.73
Corrected sample standard deviation of estimation (mm)								
corner			edge			middle		
z = 0.5	z = 5.0	z = 9.5	z = 0.5	z = 5.0	z = 9.5	z = 0.5	z = 5.0	z = 9.5
0.17	0.19	0.30	0.53	0.32	0.63	0.60	0.39	0.84

Values are rounded to two decimals

One solution to utilize more information of the light distribution while keeping the complexity low is to calculate a

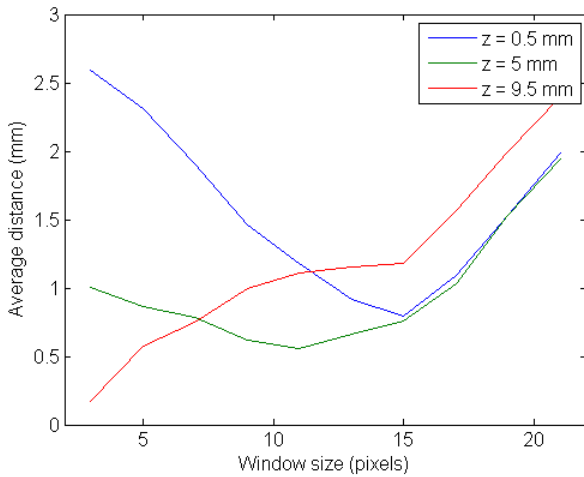


Figure 2. Mean distance in function of smoothing window size

weighted average of photon counts (center of gravity method).

C. Center of gravity

The center of gravity (COG) method calculates a weighted average of the pixel coordinates, where the weighting coefficients (p_i) are the photon counts on the pixels.

$$x_{COG} = \frac{\sum_{i=1}^N p_i x_i}{\sum_{i=1}^N p_i} ; y_{COG} = \frac{\sum_{i=1}^N p_i y_i}{\sum_{i=1}^N p_i} \quad (4)$$

The advantage of this solution is that the average coordinate could be fractional, thus the estimated position is continuous, not limited by the size of the pixels. An essential drawback is that noisy pixels highly affect the resulting coordinate. Even pixels with a small photon count could strongly modify the weighted sum if the noisy pixel is far away from the POI. In order to solve this problem some kind of noise reduction is needed.

If noise is equally distributed on the whole surface of pixels, subtracting the expected value of dark counts from each pixel would reduce the bias in COG estimation. The SPADnet sensors have an inner adjustable threshold level, and give no output signals if the total absorbed energy of photons is lower than this threshold level [5]. As a result, in certain cases when the energy over a sensor is low, this sensor will act as "blind sensor" of the tile. Subtraction of the expected value of noise from pixels of a blind sensor increases the bias of estimation. Therefore, a better way to reduce bias is to cut off the peak of photon counts, and use COG estimation only on this subset of pixels.

In this paper I investigated three different cut off schemes. In the first one a certain threshold value was applied on the pixels, photon counts under the threshold were omitted. In the second one a certain value was subtracted from each pixel's photon count, and values below zero were omitted (took into account as zero). This technique results the same shape of surface of photon counts over the sensor, but has a negative offset compared to the first technique and thus results a slightly different COG. The third scheme applied a threshold calculated as a threshold ratio times the highest photon count.

The solution proposed was chosen with respect to the lowest resulting bias in function of the threshold values. Figure 3 shows the mean distances calculated from the average for 700 simulated measurements at the three representative points

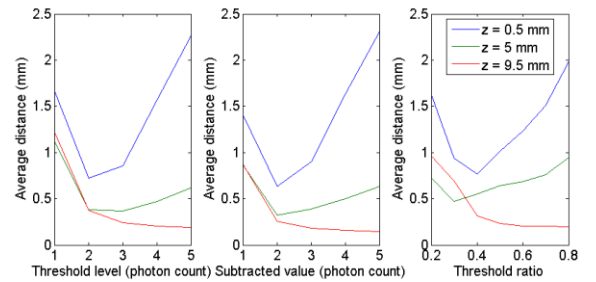


Figure 3. Results of COG estimation considering a smaller range of pixels around the peak

(‘corner’, ‘edge’, ‘middle’) at each depth level separately.

The results show that above a certain threshold level the bias of the estimation increases, especially in the case of scintillations further away from the sensor surface. The reason of this is that the material of scintillator crystal determines the approximate number of scintillation photons for a given energy level of the gamma photon. Hence it has the same order of magnitude independently from the depth level of scintillation in the crystal. Also, light photons of a scintillation further away from the detector surface have a higher spread before reaching the detector. This results in a higher number of firing pixels with a lower photon count compared to scintillations close to the sensor where there is a fewer number of firing pixels with a higher photon count. A high threshold value even filters out pixels of the photopeak, and the COG estimation result will be approaching the maximum finding estimation result.

Table III contains the bias and the corrected sample standard deviation values using a subtraction of 2 photons from each photon count as a cut off scheme (negative values are considered as zero).

TABLE III. COG RESULTS

Bias of estimation (mm)								
corner			edge			middle		
z = 0.5	z = 5.0	z = 9.5	z = 0.5	z = 5.0	z = 9.5	z = 0.5	z = 5.0	z = 9.5
0.54	0.27	0.15	0.58	0.29	0.30	0.77	0.41	0.31
Unbiased standard deviation of estimation (mm)								
corner			edge			middle		
z = 0.5	z = 5.0	z = 9.5	z = 0.5	z = 5.0	z = 9.5	z = 0.5	z = 5.0	z = 9.5
0.29	0.15	0.19	0.36	0.19	0.13	0.41	0.21	0.11

Values are rounded to two decimals

CONCLUSIONS

Comparing the above algorithms, it proved that the most useful solution of the three is COG estimation with the discussed cut off scheme.

For measurements with real device I implemented the proposed method on HDL level. By the time of this writing I am working on the integration to the existing firmware on an FPGA connected to the sensor. The current version of the positioning module supplies COG coordinates in fix point binary numbers with 10 bit fractional part. The module works in pipe-line processing mode, and can be rerun in every 20th clock cycle. It is being placed in a separate clock domain, which enables to tune the speed of the applied clock signal in order to fit processing capabilities to the frequency of gamma arrival. For a general tracer this requires a clock frequency in the order of magnitude of 100 MHz.

Further tasks of the position estimation are:

- to carry out measurements with the real device, and
- to examine the achievable precision with maximum likelihood (ML) estimation.

Although the ML method is too time consuming to be used in a live application, it gives the best result theoretically. An evaluation of the proposed COG method could be the comparison with the ML results.

ACKNOWLEDGMENT

I would like to thank Dr. Emőke Lőrincz, and Balázs Játékos for their assistance and guidance, as well as for submitting materials and simulation data for the investigation. I thank Zoltán Papp for the assistance concerning the operation of the firmware. I also thank Dr. Béla Fehér and Dr. István Kollár for constructive discussions during the process. This work is part of EU FP7 project SPADnet [11].

REFERENCES

- [1] Lewellen, T. K., “Recent developments in PET detector technology”, *Physics in Medicine and Biology*, **53**(17), R287-317, Sep. 2008, doi:10.1088/0031-9155/53/17/R01
- [2] Herman T van Dam, Giacomo Borghi, Stefan Seifert and Dennis R Schaart (2013), “Sub-200 ps CRT in monolithic scintillator PET detectors using digital SiPM arrays and maximum likelihood interaction time estimation”, *Physics in Medicine and Biology*, **58**, 3243–3257, Apr. 2013, doi: 10.1088/0031-9155/58/10/3243
- [3] Frach, T., Prescher, G., Degenhardt, C., de Gruyter, R., Schmitz, A., & Ballizany, R. (2009). “The Digital Silicon Photomultiplier – Principle of Operation and Intrinsic Detector Performance”. *IEEE Nuclear Science Symposium Conference Record*, 1959-1965, Orlando, FL, Oct. 24 - Nov. 1 2009, doi: 10.1109/NSSMIC.2009.5402143
- [4] Braga, L. H. C., Gasparini, L., Grant, L., Henderson, R. K., Massari, N., Perenzoni, M., Stoppa, D. and Walker, R., “A Fully Digital 8×16 SiPM Array for PET Applications With Per-Pixel TDCs and Real-Time Energy Output”, *IEEE Journal of Solid-State Circuits*, **49**(1), 301-314, Jan. 2014, doi: 10.1109/JSSC.2013.2284351
- [5] Braga, L. H. C., Pancheri, L., Gasparini, L., Perenzoni, M., Walker, R., Henderson, R. K., et al., “A CMOS mini-SiPM detector with in-pixel data compression for PET applications”, *IEEE Nuclear Science Symposium Conference Record*, 548-552, Valencia, Spain, 23-29 Oct. 2011, doi: 10.1109/NSSMIC.2011.6154110
- [6] Chaudhari, A. J., Joshi, A. A., Wu, Y., Leahy, R. M., Cherry, S. R., & Badawi, R. D., “Spatial distortion correction and crystal identification for position-sensitive avalanche photodiode-based PET scanners”, *IEEE Nuclear Science Symposium Conference Record*, 5045-5052, Dresden, Germany, 19-25 Oct. 2008, doi: 10.1109/NSSMIC.2008.4774373
- [7] Játékos, B., Erdei, G. and Lőrincz, E., “Simulation tool for optical design of PET detector modules including scintillator material and sensor array”, *2nd International Conference on Advancements in Nuclear Instrumentation Measurement Methods and their Applications (ANIMMA)*, 1-6, Ghent, 6-9 June 2011. doi: 10.1109/ANIMMA.2011.6172960
- [8] Bruschini, C., et al., “SPADnet: Embedded coincidence in a smart sensor network for PET applications,” *Nuclear Instruments and Methods in Physics Research A* **734**, 122–126, 2014.
- [9] Herman T. van Dam, Stefan Seifert, Ruud Vinke, Peter Dendooven, Herbert Löhner, Freek J. Beekman, and Dennis R. Schaart, “Improved Nearest Neighbor Methods for Gamma Photon Interaction Position Determination in Monolithic Scintillator PET Detectors”, *IEEE Transactions on Nuclear Science*, 2139 – 2147, Oct. 2011. doi: 10.1109/TNS.2011.2150762
- [10] T Ling, T K Lewellen and R S Miyaoka, “Depth of interaction decoding of a continuous crystal detector module”, *Physics in Medicine and Biology*, **52**(8) 2213–2228, 2007. doi: 10.1088/0031-9155/52/8/012
- [11] SPADnet consortium, “Fully Networked, Digital Components for Photon-starved Biomedical Imaging Systems”, Available: <http://www.spadnet.eu>

Metaanyag Paraméterek Homogenizálása Transzmissziós és Reflexiós Adatokból

Kalvach Arnold, Szabó Zsolt

Szélessávú Hírközlés és Villamosságtan Tanszék
Budapesti Műszaki és Gazdaságtudományi Egyetem
1111 Budapest, Egrý József utca 18.
kalvach@evt.bme.hu, szabo@evt.bme.hu

Absztrakt—A metaanyagok olyan hullámhossz alatti méretekkel rendelkező periodikus szerkezetek, amelyeknek különleges elektromágneses tulajdonságaik vannak, úgymint negatív vagy nulla törésmutató. Mivel a metaanyagok jellemző méretei nem sokkal kisebbek a vezetett hullámhossznál, ezért nincs általánosan alkalmazható módszer az effektív anyagparaméterek számítására, a homogenizálás körülményes és nagykörűt tekintést igényel. Ebben a cikkben két transzmissziós és reflexiós adatokat felhasználó homogenizálási eljárást vizsgálunk és hasonlítunk össze. A módszereket FishNet típusú metaanyagok effektív anyagparamétereinek kiszámításával szemlélítjük.

Kulcsszavak – metaanyagok, homogenizálás, effektív paraméterek

I. BEVEZETÉS

Az utóbbi két évtizedben nagy fejlődés ment végbe a metaanyag kutatás területén. Eleinte elméletben [1], majd gyakorlatban [2] mutattak be olyan hullámhossz alatti struktúrákat, amelyek különleges optikai paraméterekkel rendelkeznek, úgymint negatív [1],[3] vagy nulla törésmutató [4], illetve képesek az elektromágneses hullámok elhajlítására [5]. Ezen különleges tulajdonságok számos gyakorlati alkalmazásra adnak lehetőséget, mint a nagyobb felbontású fotolitográfia [6], antennák iránykarakterisztikájának növelése [7], vagy tárgyak elektromágneses elrejtése [5]. Éppen ezért a területen továbbra is széleskörű kutatás és nagy fejlődés várható.

A metaanyag kutatás egyik sarkalatos pontja az effektív paraméterek meghatározása. Ez azt jelenti, hogy a bonyolult hullámhossz alatti struktúrához olyan effektív anyagparamétereket rendelünk, amelyekkel aztán makroszkopikusan, a bonyolult struktúra elhagyásával tudjuk vizsgálni ezekben a szerkezetekben a hullámterjedést. Mivel ez igen jelentős problémakör, több megoldás is született a paraméterek kinyerésére. Ebben a cikkben transzmissziós és reflexiós adatokból való effektív paraméter meghatározási módszereket tárgyalunk, majd összehasonlítást adunk egy konkrét metaanyag struktúra, a FishNet vizsgálatával.

II. METAANYAGOK EFFEKTÍV ELEKTROMÁGNESES PARAMÉTEREINEK SZÁMÍTÁSA

Ha egy elektromágneses hullám olyan struktúrán halad keresztül, amelynek geometriai szerkezete jóval kisebb a hullámhossznál, akkor a struktúrában csak elsőrendű módusok

jönnek létre és a gerjesztett távöltér ugyanolyan frekvenciájú hullámokból áll, mint a gerjesztő tér. A lineáris összefüggés miatt a két tér összegét leírhatjuk az ún. makroszkopikus Maxwell-egyenletekkel úgy, hogy a struktúrához egy effektív permeabilitást illetve effektív permittivitást rendelünk. Természetes anyagok esetén ezt a permittivitást és permeabilitást az atomi struktúrák szabják meg. A metaanyagok olyan szerkezetek, ahol az inhomogenitás nem atomi szintű, ám még mindig kisebb, mint a hullámhossz, így ezekhez a struktúrákhoz hasonlóképpen rendelhető egy effektív permittivitás és permeabilitás, mint a természetes anyagokhoz. A homogenizálás célja, hogy ezen effektív paramétereket meghatározza.

Az effektív paraméterek meghatározásakor egy nem homogén struktúra (metaanyag) és egy effektív paraméterekkel rendelkező ideális homogén struktúra ugyanazon elektromágneses gerjesztésre adott válaszát hasonlítjuk össze. Amennyiben a két válasz megegyezik, az effektív paraméterek meghatározása sikeres volt. A legáltalánosabb módszer erre egy ismert vastagságú lemez szórás paramétereinek összehasonlítása a metaanyag szórás paramétereivel, azaz a reflexió (S_{11}) és transzmisszió (S_{21}) meghatározása, egy adott irányú (legtöbbször merőleges beesésű) síkhullám gerjesztés esetén. Egy ismert vastagságú ismert paraméterű homogén lemez szórás paramétere megározhatóak a Maxwell-egyenletekből és a folytonossági feltételekből [8]:

$$S_{11} = \frac{(1-\zeta^2)e^{jkd} - (1-\zeta^2)e^{-jkd}}{(1+\zeta^2)e^{jkd} - (1-\zeta^2)e^{-jkd}} \quad (1)$$

$$S_{21} = \frac{4\zeta}{(1+\zeta)^2 e^{jkd} - (1-\zeta)^2 e^{-jkd}} \quad (2)$$

ahol d a lemez vastagsága, merőleges beesés esetén $k = k_0 \sqrt{\epsilon_r \mu_r}$, k_0 a szabadtéri hullámszám, ϵ_r a lemez relatív elektromos permittivitása, μ_r a lemez relatív mágneses permeabilitása, valamint TM módus esetén és feltéve, hogy a lemezt vákuum veszi körül $\zeta = k/(\epsilon_r k_0)$.

Ugyanezen szórás paraméterek a metaanyagra méréssel vagy szimulációval határozhatóak meg. Homogenizáláskor azt szeretnénk elérni, hogy a mért/szimulált transzmisszió és reflexió minden frekvenciára megegyezzen a helyettesítő homogén rétegre kiszámolt értékekkel, azaz olyan ϵ_r -t és μ_r -t

keresünk, amire $S_{11,homogén}(f) = S_{11,mért}(f)$ és $S_{21,homogén}(f) = S_{21,mért}(f)$. Az effektív paraméterek kinyerése tehát egy inverz probléma, ahol az ismert transzmissziós és reflexiós paraméterekből határozzuk meg az azokat meghatározó elektromágneses anyagparamétereket.

Hogy egyszerűbb legyen invertálni a kifejezéseket, érdemes a szórési paramétereket az anyagparaméterek helyett származtatott mennyiségekkel kifejezni. Ilyen származtatott mennyiségek az effektív törésmutató ($n = \sqrt{\epsilon_r \mu_r}$), az effektív relatív hullámimpedancia ($\eta = \sqrt{\mu_r / \epsilon_r}$), a reflexiós tényező ($\Gamma = (\eta - 1) / (\eta + 1)$) és a transzmissziós tényező ($Z = \exp(jkd) = \exp(jnk_0 d)$).

A szórési paraméterek a reflexiós tényezővel és a transzmissziós tényezővel kifejezve:

$$S_{11} = \frac{\Gamma(1-Z^2)}{1-\Gamma^2 Z^2} \quad (3)$$

$$S_{21} = \frac{Z(1-\Gamma^2)}{1-\Gamma^2 Z^2} \quad (4)$$

A. Vékony Metaanyag Struktúrák Homogenizálása

A szórési paraméterekből a reflexiós és a transzmissziós tényező többféleképp is kifejezhető. Egy ilyen formula a következő [9]:

$$\Gamma = \frac{Z-V_2}{1-ZV_2}, \quad (5)$$

$$Z = \frac{V_1-\Gamma}{1-\Gamma V_1}. \quad (6)$$

ahol

$$V_1 = S_{21} + S_{11}, \quad (7)$$

$$V_2 = S_{21} - S_{11}. \quad (8)$$

Ezekből kifejezhetőek a következő mennyiségek [10]:

$$1 - Z = \frac{(1-V_1)(1+\Gamma)}{1-\Gamma V_1} \quad (9)$$

$$\eta = \frac{1+\Gamma}{1-\Gamma} = \frac{(1+Z)(1-V_2)}{(1-Z)(1+V_2)} \quad (10)$$

Vékony rétegek esetén, ahol $\text{Re}\{k\}d < 1$, a transzmissziós tényező a vastagság lineáris függvényeként közelíthető:

$$Z \cong 1 + jkd. \quad (11)$$

Ezt behelyettesítve (9)-be és (10)-be a következő adódik a hullámszámmra és a hullámimpedanciára [10]:

$$k \cong \frac{(1-V_1)(1+\Gamma)}{-jd(1-\Gamma V_1)}, \quad (12)$$

$$\eta \cong \frac{2(1-V_2)}{-jkd(1+V_2)}, \quad (13)$$

ahonnan a törésmutató, és az effektív anyagparaméterek könnyen meghatározhatóak:

$$n = \frac{k}{k_0} \cong \frac{(1-V_1)(1+\Gamma)}{-jk_0 d(1-\Gamma V_1)}, \quad (14)$$

$$\mu_r = n\eta = \frac{k}{k_0} \eta \cong \frac{2(1-V_2)}{-jk_0 d(1+V_2)}, \quad (15)$$

$$\epsilon_r = \frac{n^2}{\mu_r} = \left(\frac{k}{k_0}\right)^2 \frac{1}{\mu_r}. \quad (16)$$

A relatív permittivitás azonban kifejezhető μ_r -hez hasonló formában is:

$$\epsilon_r \cong \frac{2(1-V_1)}{-jk_0 d(1+V_1)}, \quad (17)$$

felhasználva, hogy

$$\eta^2 = \frac{(1+V_1)(1-V_2)}{(1-V_1)(1+V_2)} = \frac{\mu_r}{\epsilon_r}. \quad (18)$$

vagy a következőképp:

$$\epsilon_r \cong \mu_r - j \frac{2S_{11}}{k_0 d}, \quad (19)$$

felhasználva, hogy S_{11} -et a következőképp közelíthetjük (3) és (11) alapján:

$$S_{11} = \frac{-2jkd(\eta^2-1)}{(\eta+1)^2-(\eta-1)^2}. \quad (20)$$

Fontos, megjegyezni, hogy míg (17) pontatlan a rezonanciák környékén a V_1 paraméter rezonanciára való érzéketlensége miatt, addig (19) kihangsúlyozza a rezonanciák hatását.

B. Többretegű metaanyag struktúrák homogenizálása

Bár az eddig ismertetett kifejezések vékony rétegek esetén nagyjából pontos értéket adnak, vastagabb rétegek esetén a (11) közelítés miatt nem alkalmazhatók. Ez többnyire akkor fordul elő, ha egy metaanyag struktúrából nem egy, hanem több réteget vizsgálunk. Ez esetben a pontos exponenciális kifejezéssel kell számolnunk.

A (3) és (4) egyenleteket invertálva (és a reflexiós tényező kifejtését felhasználva) a következő kifejezéseket kapjuk a hullámimpedanciára és a transzmissziós tényezőre [11]:

$$\eta = \pm \sqrt{\frac{(1+S_{11})^2 - S_{21}^2}{(1-S_{11})^2 - S_{21}^2}}, \quad (21)$$

$$Z = e^{jnk_0 d} = \frac{S_{21}}{1-S_{11}\Gamma}. \quad (22)$$

ahol a hullámimpedancia előjelét úgy kell megválasztani, hogy az fizikailag értelmes eredményt adjon. Mivel egy metaanyag mindig passzív, ezért a hullámimpedancia valós részének mindig pozitívnak kell lennie. Az anyagparaméterek egyértelmű meghatározásához a hullámimpedancia mellett

szükséges a törésmutató meghatározása is. A törésmutatót (22)-ből a következőképp fejezhetjük ki:

$$n = \frac{\Im\{\ln(Z)\} + 2m\pi - i\Re\{\ln(Z)\}}{k_0 d}, \quad m \in \mathbf{R}. \quad (23)$$

Ebből látható, hogy bár n képzetes része egyértelműen meghatározható

$$\Im\{n\} = \kappa = \frac{\Re\{\ln(Z)\}}{k_0 d}, \quad (24)$$

addig a valós résznek több megoldása van

$$\Re\{n\} = n' = \frac{\Im\{\ln(Z)\} + 2m\pi}{k_0 d}. \quad (25)$$

Ezek közül viszont csak egy az, ami fizikai tartalommal bír. A valós megoldás meghatározására felhasználhatjuk azt a feltételt, hogy a törésmutatónak folytonosnak kell lennie a teljes frekvenciatartományon. Ez viszont azt igényli, hogy legalább egy frekvencián ismerjük, melyik a valós megoldás, valamint iteratív módon tudjuk csak meghatározni minden frekvenciára a helyes megoldást. Egy könnyebben alkalmazható eljárás a Kramers-Krönig reláció felhasználása [12]. Ez kapcsolatot teremt egy analitikus függvény valós és képzetes része közt. Mivel a komplex törésmutató analitikus függvény, ezért felhasználható, hogy

$$\Re\{n(\omega')\} = 1 + \frac{2}{\pi} P \int_0^\infty \frac{\omega \Im\{n(\omega)\}}{\omega^2 - \omega'^2} d\omega \quad (26)$$

ahol P az improprius integrál főértékét jelenti. Mivel a törésmutató képzetes része egyértelműen ismert, ezért a Kramers-Krönig összefüggés segítségével a valós része is meghatározható. Gyakorlati akadályt jelent azonban, hogy a képzetes részt nem ismerjük a teljes frekvenciatartományon, és nem is ismerhetjük, mivel egy bizonyos hullámhossz alatt a struktúra nem homogenizálható. Így tehát az összefüggés nem használható a pontos érték meghatározására, azonban alkalmas a megfelelő megoldás kiválasztásához, ha kellően nagy frekvenciatartományon integrálunk, azáltal, hogy mindig a becsült értékhez legközelebbi ágat választjuk ki az összes megoldás közül.

Többrétegű struktúrák esetén további fontos tényező az effektív vastagság meghatározása. Míg homogén anyagok esetén a vastagság pontosan definiálható, addig inhomogén struktúrák esetén nem. Egyrészt maga a struktúra nem feltétlenül fix vastagságú, másrészt nehéz meghatározni azt a távolságot, ahol a struktúra közeltére már nem játszik túl nagy szerepet. Így hát a metaanyagokhoz egy úgynevezett effektív vastagságot rendelünk. Ez az a vastagság, amilyen vastagságú homogén lemezzel helyettesíteni fogjuk a struktúrát.

Egyrétegű metaanyagok esetén az effektív vastagság tetszőlegesen választható, az effektív paraméterek úgy fognak alkalmazkodni a vastagsághoz, hogy a szórási paraméterek mindig a mért eredményeket adják. Célszerű természetesen a struktúra méretéhez közeli effektív vastagságot használni. A mért szórási paramétereket mindig két olyan síkra kell

vonatkoztatni, amelyek egymástól effektív vastagságnyira vannak.

Többrétegű struktúrák esetén elvárjuk, hogy az anyagparaméterek ne változzanak a rétegek számának növelésével. Mivel azonban ez a feltétel általában nem teljesül, így a kiszámított effektív anyagparaméterek csak egy adott vastagságú metaanyagra vonatkoznak. Egy adott rétegszámból álló metaanyag effektív vastagságát legtöbbször optimalizálással lehet meghatározni [11], azt a legkisebb vastagságot választva, amelyiknél a hullámimpedancia abszolút értéke nem változik.

III. A FISHNET STRUKTÚRA VIZSGÁLATA

A homogenizálási módszerek gyakorlati bemutatására és azok eredményeinek összehasonlítására az ún. FishNet metaanyag struktúrát vizsgáltuk meg.

A FishNet egy olyan fém-szigetelő-fém struktúrát rétegekből álló planáris periodikus szerkezet, amelyben struktúrájából eredően adott polarizációjú gerjesztő tér esetén elektromos és mágneses rezonancia is létrejöhethet. Ezen rezonanciák helyét a struktúra méretei szabják meg, tehát azok tervezhetőek, és akár olyan nagy hullámhosszra is tehetőek, ahol a struktúra már leírható effektív anyagparaméterekkel. Az elektromos és mágneses rezonanciák eredménye, hogy a struktúra elektromos és mágneses választ is fog adni a gerjesztésre, tehát az effektív relatív permittivitása és permeabilitása is 1-től eltérő lesz. A rezonanciák tervezésével akár olyan választ is kaphatunk, amelyet nulla vagy negatív törésmutatójú homogén anyaggal lehet csak helyettesíteni.

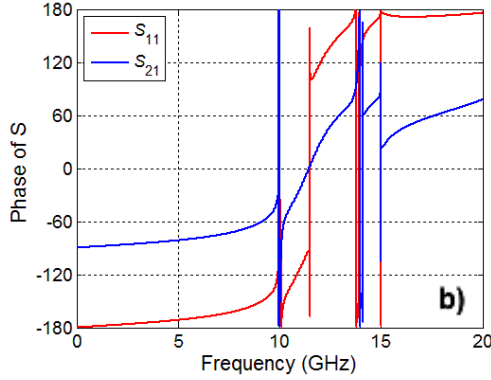
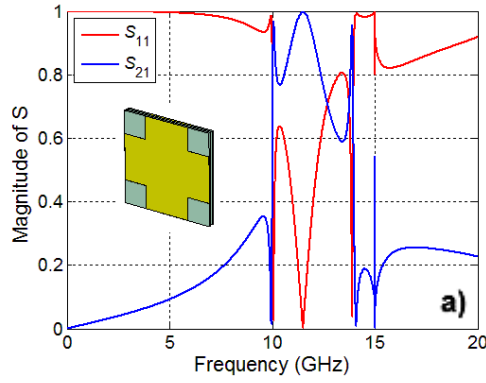
Az effektív anyagparaméterek kinyeréséhez kommersziális elektromágneses térszámító programsomag segítségével meghatároztuk egy 10GHz-es frekvenciatartományban működő FishNet struktúra transzmissziós reflexiós paramétereit, ahogy az 1. ábrán látható.

A FishNet struktúra elemi cellája az 1. ábra beillesztésében látható. Az elemi cella rácsállandója 20mm, a rézből készült kereszt szélessége 10mm, vastagsága 0.018mm, a hordozó FR-4, elektromos permittivitása 4.3, vastagsága 0.5mm. A struktúra két rétegből áll, amelyeket 0.5mm levegőréteg választ el [13]. A hátsó réteg hátulján szintén egy réz kereszt található.

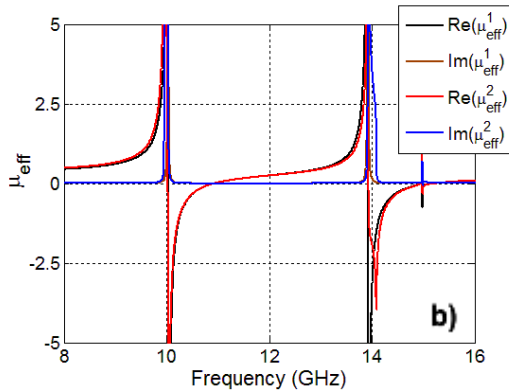
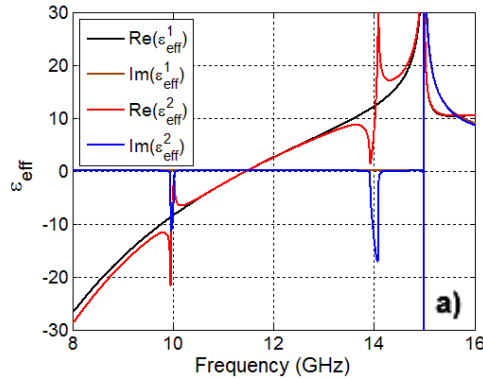
A két homogenizálási eljárással kapott elektromágneses paramétereket a 2. ábra mutatja, ahol az 1-es index a vékony rétegek esetén alkalmazható homogenizálási modellt [10], a 2-es index a tetszőleges vastagság esetén érvényes modellt [12] jelöli. A második modellel számolt törésmutatót a 3. ábra szemlélteti, megfigyelhető, hogy 10GHz környezetében létezik egy olyan frekvenciatartomány, ahol a törésmutató negatív.

IV. KONKLÚZIÓ

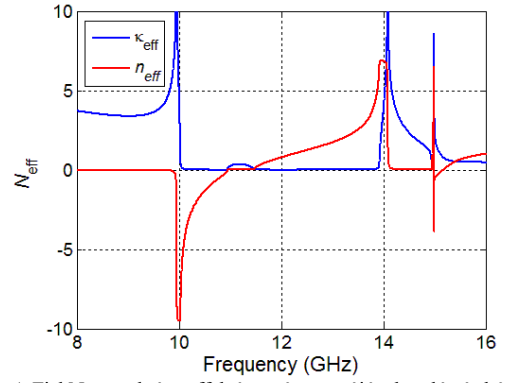
Bemutattunk két olyan eljárást, amelyek a transzmissziós-reflexiós adatokból homogenizálással meghatározzák metaanyagok effektív elektromágneses anyagparamétereit. A szimulációs eredményekből megfigyelhető, hogy kis frekvenciáknál a két modell hasonló eredményt ad, azonban nagyobb frekvenciáknál az eltérés számottevő. Látható továbbá, hogy a vékonyréteg modell nem veszi figyelembe az elektromos permittivitásban létrejövő antirezonanciákat.



1. ábra - A FishNet struktúra transzmissziós és reflexiós paraméterei



2. ábra - A homogenizálási eljárásokkal számolt effektív anyagparaméterek.



3. ábra - A FishNet struktúra effektív törésmutatójának valós és képzetes része

KÖSZÖNETNYILVÁNÍTÁS

Köszönjük a Magyar Tudományos Akadémia Bolyai János Kutatói Ösztöndíj és az EUREKA MetaFer projekt támogatását.

HIVATKOZÁSOK

- [1] J. B. Pendry, "Negative refraction makes a perfect lens." *Physical review letters* 85.18 (2000), pp. 3966.
- [2] R. A. Shelby, D. R. Smith, and S. Schultz. "Experimental verification of a negative index of refraction." *Science* 292.5514 (2001), pp. 77-79.
- [3] V. M. Shalaev, "Optical negative-index metamaterials." *Nature photonics* 1.1 (2007), pp. 41-48.
- [4] R. W. Ziolkowski, "Propagation in and scattering from a matched metamaterial having a zero index of refraction." *Physical Review E* 70.4 (2004): 046608.
- [5] D. Schurig et al., "Metamaterial electromagnetic cloak at microwave frequencies." *Science* 314.5801 (2006), pp. 977-980.
- [6] N. Fang, H. Lee, Ch. Sun, X. Zhang "Sub-diffraction-limited optical imaging with a silver superlens." *Science* 308.5721 (2005) , pp. 534-537.
- [7] D. Li, Zs. Szabó, X. Qing, E.-P. Li and Z. N. Chen "A high gain antenna with an optimized metamaterial inspired superstrate." *Antennas and Propagation, IEEE Transactions on* 60.12 (2012), pp. 6018-6023.
- [8] L. Solymar and E. Shamonina. *Waves in metamaterials*. Oxford University Press, 2009.
- [9] A. M. Nicolson and G. F. Ross. "Measurement of the intrinsic properties of materials by time-domain techniques." *Instrumentation and Measurement, IEEE Transactions on* 19.4 (1970), pp. 377-382.
- [10] R. W. Ziolkowski, "Design, fabrication, and testing of double negative metamaterials." *Antennas and Propagation, IEEE Transactions on* 51.7 (2003), pp. 1516-1529.
- [11] X. Chen, T. M. Grzegorzczak, B.-I. Wu, J. Pacheco and J. A. Kong, "Robust method to retrieve the constitutive effective parameters of metamaterials." *Physical Review E* 70.1 (2004): 016608.
- [12] Zs. Szabó, G.-H. Park, R. Hedge and E.-P. Li, "A unique extraction of metamaterial parameters based on Kramers–Kronig relationship." *Microwave Theory and Techniques, IEEE Transactions on* 58.10 (2010), pp. 2646-2653.
- [13] A. Ourir, R. Abdeddaim and J. de Rosny. "Planar metamaterial based on hybridization for directive emission." *Optics express* 20.16 (2012), pp. 17545-17551.

Plazmonikus nanorészecskék elrejtése látható tartományban

Csajbók Viktória és Szabó Zsolt

Szélessávú Hírközlés és Villamosságtan Tanszék
Budapesti Műszaki és Gazdaságtudományi Egyetem
Egry József utca 18, 1111 Budapest, Magyarország
csajbokviki@gmail.com, szabo@hvt.bme.hu

Absztrakt—Eljárást dolgozunk ki homogén fémből készült gömb alakú nanorészecskék koncentrikus rétegekkel történő elektromágneses elrejtésének vizsgálatára a látható fény tartományában. A Mie-elmélet felhasználásával kiszámítjuk síkhullámok szóródását a rétegzett objektumon. A minél tökéletesebb elrejtés megvalósítása érdekében optimalizáljuk a köpeny elektromágneses paramétereit és a geometria méreteit. A kidolgozott eljárást egy ezüst gömböt körülvevő elektromágneses köpeny tervezésével illusztráljuk.

Kulcsszavak: plazmonikus részecskék, elektromágneses elrejtés

I. BEVEZETÉS

Az elmúlt évek során egyre jelentősebb az érdeklődés a mesterségesen előállított anyagok elektromágneses tulajdonságai iránt, mind elméleti szempontból, mind pedig kísérleti szempontból [1, 2]. Az anyagok egzotikus elektromágneses tulajdonságainak tervezésével elérhetővé válik a tárgyra eső elektromágneses hullámok manipulálása. A mesterségesen előállított metaanyagok egyik legvonzóbb alkalmazása az objektumok láthatatlanná tévésének a lehetősége [1, 3]. Az elrejtetni kívánt objektum köré olyan köpenyt helyeznek, amely úgy módosítja a beeső hullám szórt terét, mintha ott nem lett volna semmi [4]. Ezáltal detektálhatatlanná válik az elrejtő köpenyen belüli tárgy [5].

Ideális esetben egy tárgy elrejtése maga után vonja a teljes szórás megszüntetését minden irányból, a környezettől és a megfigyelő pozíciójától függetlenül és mindezt széles frekvenciatartományon teljesíti, ezt azonban a jelenlegi technológiákkal még nem sikerült megvalósítani. A végső cél a teljes látószögű elrejtés a teljes látható frekvenciatartományban, ezidáig plazmonikus nanorészecskére sikerült elrejtő köpenyt készíteni [6-12].

Jelen cikkben kisebb gömb alakú fém részecskék elrejtésére keresünk megoldást. Gömb alakú objektumokat jól lehet tárgyalni a Mie-elmélet kiterjesztésével. [13] A gömböt körülvevő elrejtő köpenyt pedig a Drude-modell segítségével lehet könnyen modellezni, ha plazmafrekvenciája távol esik az elrejtési frekvenciától. [14, 15] A szórási hatáskeresztmetszet minimalizálásával, hullámhossznál kisebb fémes nanorészecskék elrejtését vizsgáljuk analitikus és numerikus szimulációkkal, valamint meghatározzuk egy adott vastagságú elrejtő köpeny elektromos permittivitását.

II. ELMÉLETI ÁTTEKINTÉS

A. Fémes objektumok szórás keresztmetszetének csökkentése

A tökéletes elrejtő köpeny megvalósításához a transzformációs elektrodinamika az ideális megközelítés, habár a transzformációs köpeny csak közelítőleg valósítható meg, így a szórt elektromágneses tér nem nulla. A transzformációs köpeny bármilyen praktikus realizációja valamilyen módon limitált lesz, és az apró tökéletlenségekre való nagyfokú érzékenysége miatt mindenképpen eredményezni fog valamennyi szórás. Ez a maradék szórás nem szabályozható, akár még nagyra is válhat [4].

Eltérő megközelítés szükséges plazmonikus részecskék és makroszkopikus testek álcázására, vagyis a szórási keresztmetszet csökkentésére [4]. A plazmonikus részecskék esetén általában elégséges minimalizálni a szórt tér multipólus sorfejtésbeli domináns tagjait. Hullámhossz alatti méretekkel rendelkező részecskék esetén ez egy nem rezonáns elrejtési eljárás és a sorfejtés elhanyagolt tagjai által létrehozott maradék sugárzás kellőképpen redukálható. Ebben az esetben az elrejtő köpenyt homogén, izotróp anyagok felhasználásával is lehet tervezni, szemben a transzformációs optika által eredményezett anizotróp anyagparaméterekkel rendelkező köpenyekkel.

Az elektromágneses szórás csökkentő elrejtő köpenyek tipikusan egy adott geometriával rendelkező tárgy elrejtésére működnek, más geometriával rendelkező tárgyak esetén az elrejtő köpeny paramétereit újra kell tervezni. Ha tetszőleges anyagparaméterekkel rendelkező tárgyat szeretnénk elrejtetni, célszerű vezető anyaggal bevonni a tárgyat. A hullámhosszal összemérhető nagyságú tárgyak esetén a domináns szórási harmonikusok száma gyorsan növekszik a tárgy méretével, ez pedig az elrejtő köpeny rétegeinek számának növekedését és a tervezés komplexitását eredményezi [7].

Tekintsünk egy monokromatikus $\exp(j\omega t)$ időfüggésű elektromágneses hullámot adott háttéranyagban, amely legyen ε_0 elektromos permittivitású és μ_0 mágneses permeabilitású, majd vizsgáljuk a szórását egy tetszőleges geometriájú és tetszőleges anyagparaméterekkel rendelkező tárgyon, amely egy gömbi koordináta-rendszer origójában van elhelyezve. A tárgyon szóródó síkhullám elektromos és mágneses komponense gömbi harmonikusokkal fejezhető ki [16]

$$\begin{aligned} \mathbf{E}_i &= \sum_{n=1}^{\infty} \sum_{m=-n}^n a_{nm} \nabla \times \nabla (\mathbf{r} \psi_n^m) - j\omega\mu_0 b_{nm} \nabla \times (\mathbf{r} \psi_n^m) \\ \mathbf{H}_i &= \sum_{n=1}^{\infty} \sum_{m=-n}^n b_{nm} \nabla \times \nabla (\mathbf{r} \psi_n^m) - j\omega\mu_0 a_{nm} \nabla \times (\mathbf{r} \psi_n^m) \end{aligned} \quad (1)$$

Az egyenletekben $\mathbf{r}$ a radiális vektor és ψ_n^m a skalár gömbi harmonikusok, amelyek a Helmholtz-egyenlet megoldásai. Az összeg minden kifejezése megfelel egy TE vagy TM modulusú gömbhullámnak. Tekintettel arra, hogy a Maxwell-egyenletek lineárisak és a gömbi harmonikusok ortogonális függvények, ezért a szórási probléma megoldható külön, minden egyes gömbi harmonikusra, a teljes szórási teret az összes multipól szuperpozíciójaként felírva [4]:

$$\mathbf{E}_s = \sum_{n=1}^{\infty} \sum_{m=-n}^n c_n^{\text{TM}} a_{nm} \nabla \times \nabla \times (\mathbf{r} \psi_n^m) - j\omega\mu_0 c_n^{\text{TE}} b_{nm} \nabla \times (\mathbf{r} \psi_n^m) \quad (2)$$

$$\mathbf{H}_s = \sum_{n=1}^{\infty} \sum_{m=-n}^n c_n^{\text{TE}} b_{nm} \nabla \times \nabla \times (\mathbf{r} \psi_n^m) - j\omega\mu_0 c_n^{\text{TM}} a_{nm} \nabla \times (\mathbf{r} \psi_n^m)$$

A c_n együttható határozza meg, hogy a szórás hányad része TE vagy TM módus. Ezek a szórási együtthatók csak az adott tárgy geometriájától és anyagparamétereitől, valamint a frekvenciától függenek. Egy adott a karakterisztikus méretű tárgy esetén a hozzájárulás mértéke dominál az összegben. A tárgy teljes láthatósága a teljes szórási hatáskeresztmetszettel adható meg:

$$\sigma_s = \frac{2\pi}{k_0^2} \sum_{n=1}^{\infty} \sum_{m=-n}^n (2n+1) (|c_n^{\text{TE}}|^2 + |c_n^{\text{TM}}|^2) \quad (3)$$

Tehát a c_n együtthatóktól függ, hogy látjuk-e a tárgyat, vagy sem.

B. Plazmonikus köpeny

A plazmonikus köpeny homogén, izotróp anyagból készült héjak felhasználásával valósítja meg a szórt tér csökkentését. A plazmonikus jelző arra utal, hogy úgy rejt el tárgyakat, hogy a köpeny relatív permittivitása kisebb, mint egy. [9, 16, 17]

Vizsgáljuk egy síkhullám szóródását egy a sugarú, ϵ permittivitású gömbön, amely egy dielektromos anyagból készült héjba van beágyazva, amelynek sugara a_c és a permittivitása ϵ_c . A gömbszimmetria miatt az (1)-(3) egyenletekben elhagyhatjuk az m indexet, illetve a szórási hatáskeresztmetszet együtthatói, c_n^{TE} és c_n^{TM} számolhatók a határfeltételekből. Mivel a héj csak egy szabadsági fokkal rendelkezik, ezért egy héj csak a multipólus sorfejtés egyik tagját képes csökkenteni. Hullámhossz alatti méretekkel rendelkező részecskék esetén a sorfejtésnek csak a dipólus módusa a domináns, ezért ebben az esetben egyetlen héj is elégséges a részecske elrejtéséhez. Egyetlen héj esetén az elrejtés akkor működik megfelelően, ha $k_0 a_c \ll 1$. Ebben az esetben

$$c_1^{\text{TE}} \approx 0 \text{ és } c_1^{\text{TM}} \approx -j(k_0 a_c)^3 f(\epsilon_c, a/a_c) \quad (4)$$

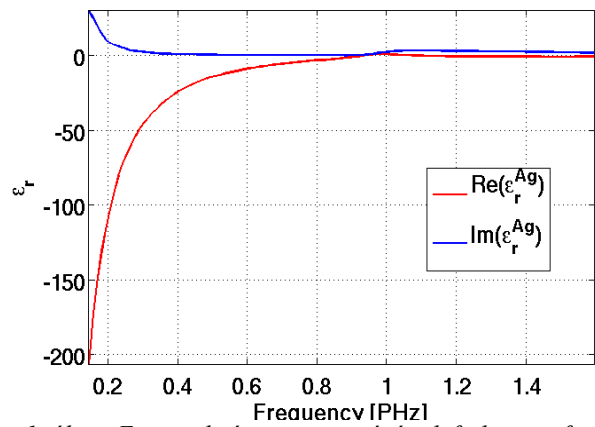
A c_1^{TM} kifejezhető egy valós értékű függvénnyel, ami pedig akkor válik nullává, amikor a következő feltétel teljesül

$$\left(\frac{a}{a_c}\right)^3 = \frac{(\epsilon_c - \epsilon_0)(2\epsilon_c + \epsilon_0)}{(\epsilon_c - \epsilon_0)(2\epsilon_c + \epsilon_0)} \quad (5)$$

III. FÉM NANORÉSZECSCKE ELREJTÉSÉNEK VIZSGÁLATA

A cél, hogy a látható fény tartományában homogén izotróp dielektrikumból készült elektromágneses köpennyel elrejtünk egy nanométeres nagyságrendű fémgömböt. Az (5) egyenlet alapján becslés adható az elrejtő héj optimális méretére és anyagparaméterére. A gömb méretét úgy választjuk meg, hogy kisebb legyen a megvilágító fény hullámhosszánál, de azért még kivitelezhető méretű legyen. Az ezüst gömb sugara legyen 50 nm, az ezt elrejtő köpeny sugara pedig 80 nm. 450 nm hullámhosszú síkhullámra kerestük a köpeny anyagparamétereit.

Az ezüst permittivitását széles frekvencia tartományon ismerjük, amint az 1. ábrán látható.



1. ábra: Ezüst relatív permittivitásának frekvenciafüggése

A. Elrejtő köpeny anyagparamétereinek számítása

Az elrejtő köpeny ϵ_c permittivitását az (5) egyenlet megoldásával kapjuk. A másodfokú egyenlet két megoldást eredményez. A két megoldás közül az egyik esetén teljesül a plazmonikus feltétel, vagyis a relatív permittivitás valós része kisebb mint egy. A másik megoldásnak általában nincs fizikai tartalma, vagy olyan permittivitás értékeket eredményez ami nem megvalósítható. A számolást elvégeztük a teljes látható tartománybeli hullámhosszokon több fém/dielektrikum sugárárány mellett. A köpeny permittivitásának értékei a 2. és 3. ábrán láthatók.

Az általunk beállított 0.625-es méretarányra és a 450 nm megvilágító hullámhosszra, a köpeny permittivitására $\epsilon_r^{\text{cloak}} = 0.3880 - 0.0194i$ adódott.

B. Köpeny diszperziós relációjának meghatározása Drude-moddellel

A dielektrikum modellezéséhez diszperzív Drude-moddellel használtunk. A vizsgált hullámhossz és az ott kapott permittivitás értékekkel kiszámolható a plazmafrekvencia és az ütközési frekvencia:

$$\omega_{\text{plasma}} = \omega \sqrt{1 - \epsilon_r' + \frac{\epsilon_r''}{1 - \epsilon_r'}} \quad (6)$$

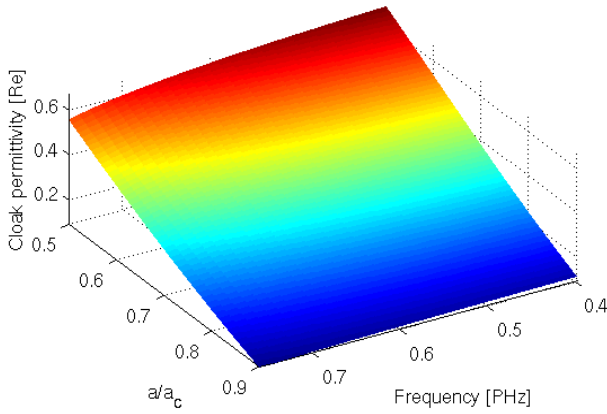
$$\gamma_e = \omega \frac{|\epsilon_r''|}{1 - \epsilon_r'} \quad (7)$$

amelyekkel a 4. ábrán látható frekvenciafüggő elektromos permittivitást kapjuk. Ezek az értékek megadhatók az általunk használt kereskedelmi forgalomban kapható elektromágneses térszámító programcsomagoknak.

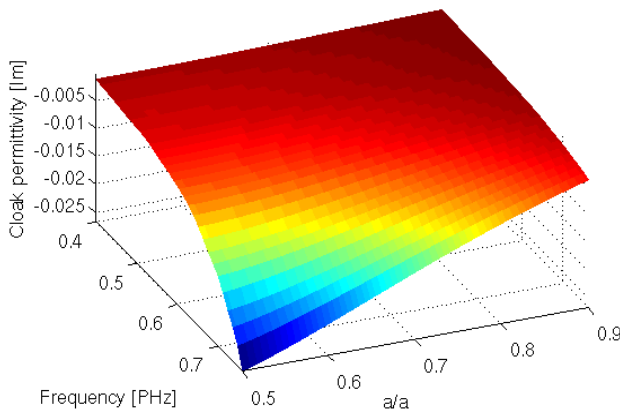
A 6. és 7. egyenlet alapján, az elrejtő köpenyre kapott plazmaplazmafrekvencia és ütközési frekvencia értékek:

- $\omega_{\text{plasma}} = 3.27 \text{ PHz}$
- $\gamma = 0.13 \text{ PHz}$

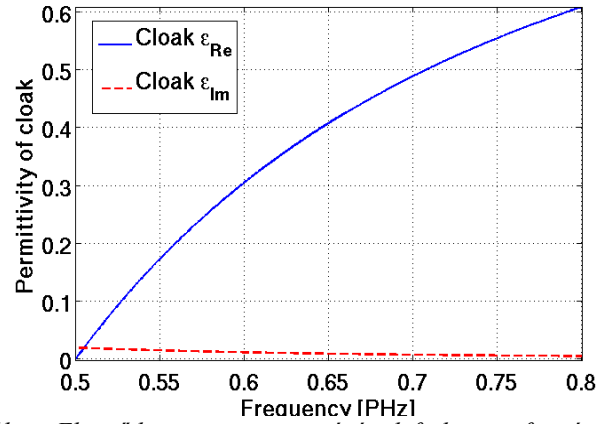
Látható tehát, hogy az elrejtő köpenyre kapott plazmafrekvencia nagysága távol van a vizsgált frekvenciától, vagyis alátámasztja, hogy a Drude-modell használható.



2. ábra: Elrejtő köpeny permittitásának valós része a méretarányok és megvilágító fény frekvenciájának függvényében



3. ábra: Elrejtő köpeny permittitásának képzetes része a méretarányok és megvilágító fény frekvenciájának függvényében

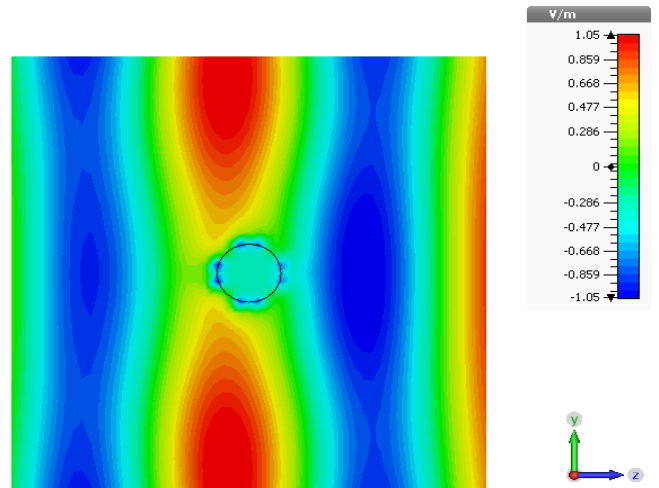


4. ábra: Elrejtő köpeny permittitásának frekvenciafüggése a Drude-modell alkalmazásával

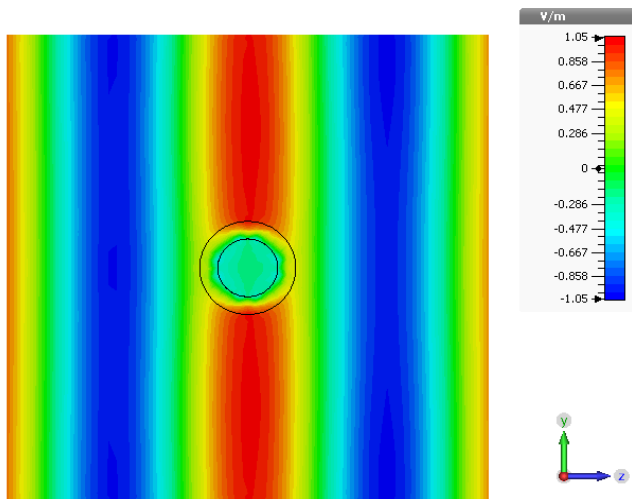
C. Szimulációs eredmények

A szimulációban egy z-irányba haladó, x- irányba mutató polarizációval rendelkező síkhullám szórását vizsgáltuk. Mivel el szeretnénk rejteni a gömböt, ezért azt várjuk, hogy a beeső síkhullám ne szóródjon a gömb körüli térrészben, vagyis a gömb előtt és mögött az elektromágneses téreloszlás legyen síkhullám jellegű.

Az 5. ábrán a síkhullám szóródása látható az ezüstgömbön elrejtő köpeny nélkül. A szórás számottevő, a gömb maga körül eltorzítja az elektromágneses teret. Az ezüst gömbön látható nagy térintenzitások a numerikus rács felbontásának köszönhetők, ideális esetben nem kellene ott lenniük. Ez abból ered, hogy a teret téglalap alakú ráccsal osztottuk fel, viszont annak érdekében, hogy véges idő alatt lefusson a szimuláció, kompromisszumot kellett kötni a felbontás mértékével.



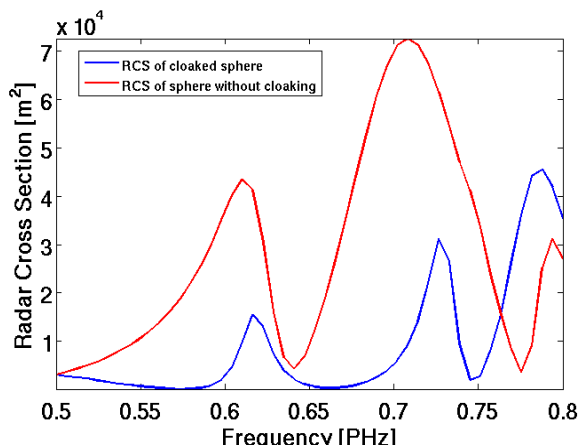
5. ábra: Ezüst gömbön szóródó elektromágneses síkhullám elektromos komponense



6. ábra: Elektromos térerősség a köpennyel körbevett ezüst gömbön való szóródás esetén

Ha az ezüst gömböt körülvesszük az elektromágneses köpennyel, akkor jelentősen megváltozik a gömb körüli elektromágneses téreloszlás. A 6. ábra az elektromágneses köpeny működését szemlélteti, ebben az esetben az elektromágneses téreloszlás a gömb körül alig tér el egy vákuumban terjedő síkhullámtól.

A 7. ábrán radarkeresztmetszet frekvenciafüggése látható önmagában álló és elrejtő köpennyel körülvett gömb esetén. A piros görbe az ezüst gömb körül radar keresztmetszetet, a kék görbe pedig a köpennyel körbevett gömbre vonatkozik. Összehasonlítva a két görbét megállapítható, hogy vizsgált tárgyakra érkező különböző hullámhosszú hullámok mennyire szóródnak, illetve reflektálódnak. Jól látható, hogy a keresett frekvenciaértékre, 0.6649 PHz esetén nagyságrendekkel csökkent a szórás, illetve a teljes szimulációs tartományon tapasztalható a köpeny elrejtő hatása.



7. ábra: Radarkeresztmetszet csak az ezüst gömb esetén, illetve az ezüst gömböt körülvető köpennyel együtt

IV. KONKLÚZIÓ

Egy analitikus formulát használva meghatároztuk plazmonikus nanorészecskék elektromágneses elrejtő köpenyének anyagparamétereit. Ezeket a paramétereket felhasználva egy kommersziális elektromágneses térszámító programcsomaggal szimuláltuk az elrejtő köpeny viselkedését, meghatároztuk a radarkeresztmetszetet, és azt összehasonlítva az önmagában álló plazmonikus nanorészecske radarkeresztmetszetével, nagyságrendekkel kisebb szórást kapunk.

A tervezési eljárás során meghatározott, elrejtő réteggel körülvett nanorészecskékké olyan nanokompozitok valósíthatók meg, amelynek elektromágneses tulajdonságai a mátrix anyag tulajdonságaihoz hasonlóak, viszont az anyagnak különleges mechanikai tulajdonságai lehetnek.

KÖSZÖNETNYILVÁNÍTÁS

Köszönjük a Magyar Tudományos Akadémia Bolyai János Kutatói Ösztöndíjának és az EUREKA MetaFer project támogatását.

HIVATKOZÁSOK

- [1] L. Solymár and E. Shamonina, *Waves in Metamaterials*, Oxford, University Press, 2009.
- [2] Marqués R., Martín F., Sorolla M., *Metamaterials with NegativeParameters*. John Wiley and Sons, 2008.
- [3] N. Engheta, R. W. Ziolkowski, *Metamaterials Physics and Engineering Applications*, John Wiley and Sons, 2006.
- [4] H. Chen, B.-I. Wu, B. Zhang, and J. A. Kong, "Electromagnetic Wave Interactions with a Metamaterial Cloak," *Phys. Rev. Lett.*, vol. 99, no. 6, p. 063903, Aug. 2007.
- [5] Martin Beech, *The Physics of Invisibility A story of Light and Deception*, Springer 2012
- [6] F. Monticone, C. Argyropoulos, A. Alu, „Layered plasmonic cloaks to tailor the optical scattering at the nanoscale”, *Scientific Reports*, 2012
- [7] A. Alù and N. Engheta, "Effects of size and frequency dispersion in plasmonic cloaking," *Phys. Rev. E*, vol. 78, no. 4, p. 045602, Oct. 2008.
- [8] Wenshan Cai, Vladimir Shalaev, *Optical Metamaterials: Fundamentals and Applications*, Springer, 2009
- [9] A. Alù and N. Engheta, "Plasmonic materials in transparency and cloaking problems: mechanism, robustness, and physical insights," *Opt. Express*, vol. 15, no. 6, pp. 3318–3332, Mar. 2007.
- [10] A. Alù and N. Engheta, "Cloaking and transparency for collections of particles with metamaterial and plasmonic covers," *Opt. Express*, vol. 15, no. 12, pp. 7578–7590, Jun. 2007.
- [11] A. Alù and N. Engheta, "Multifrequency Optical Invisibility Cloak with Layered Plasmonic Shells," *Phys. Rev. Lett.*, vol. 100, no. 11, p. 113901, Mar. 2008.
- [12] A. Alù and N. Engheta, "Theory and potentials of multi-layered plasmonic covers for multi-frequency cloaking," *New J. Phys.*, vol. 10, no. 11, p. 115036, Nov. 2008.
- [13] C. F. Bohren, D. R. Huffman, *Absorption and Scattering of Light by Small Particles*, Wiley-VCH, 1998
- [14] Zsolt Szabó, János Fűzi, „Equivalence of Magnetic Metamaterials and Composites in the View of Effective Medium Theories”, *IEEE Transactions on Magnetics*, Vol. 50, No. 4, April 2014
- [15] Zsolt Szabó, Yasaman Kiasat, Er Ping Li, „Subwavelength imaging with composite metamaterials”, Vol. 31, No. 5 / May 2014 / *J. Opt. Soc. Am. B*
- [16] C. H. Papas, *Theory of Electromagnetic Wave Propagation*. Courier Dover Publications, 2013.
- [17] J. D. Jackson, *Classical Electrodynamics*. Wiley, 1998.
- [18] Kormoczi K., Szabo Z., "Near-Infrared Invisibility Cloak Engineered With Two-Phase Metal-Dielectric Composites", *Magnetics, IEEE Transactions (Volume:50 ,Issue: 2)*, 14128047, Feb. 2014.

Nemkonvencionális litográfia plazmonikus nanoszerkezetekkel

Vermes Anna, Szabó Zsolt

Szélessávú Hírközlés és Villamosságtan Tanszék
Budapesti Műszaki és Gazdaságtudományi Egyetem
Egry József u. 18, H-1111 Budapest, Hungary

A jelen munkában egy újszerű plazmonikus nanoszerkezetet mutatunk be, mellyel diffrakciós határ alatti fotolitográfia valósítható meg. A javasolt szerkezet kialakítása a fotoreziszt felszínén önszerveződő nanogömb réteg segítségével történik, melyet a maszkoló gömbrétegen át fémréteg leválasztása követ. A gömbök eltávolítása után a létrejött bow-tie fémrészecske rendszer a fényt lokalizált felületi plazmon rezonancia útján fókuszálja, és a fotorezisztet exponálja. A rezisztben kialakuló intenzitáseloszlást a Maxwell-egyenletek numerikus megoldásával számoltuk, majd kioldási modellek segítségével megbecsültük a rezisztben előhívás után kialakuló profilt. A szimulációk segítségével meghatározhatók a litográfiai elrendezés optimális geometriai és anyag paraméterei, a megvilágító hullámhossz és a kioldási paraméterek a különböző periodikus mintázatok kialakításához.

Kulcsszavak—fotolitográfia, plazmonikus nanoszerkezet, metaanyag

I. BEVEZETÉS

A hagyományos optikai litográfia felbontását az exponáló elektromágneses sugárzás hullámhossza szabja meg [1]. A diffrakciós határ alatti felbontás elérésére számos nemkonvencionális litográfiai technikát megvizsgáltak [1], [2]. Az elektron- és ionsugaras litográfia [3], az imprint litográfia [4] és a dip-pen litográfia [5] segítségével a nanométeres felbontás elérhető, de a nagy felületű, gyors és olcsó megmunkálásra ezek a technikák nem alkalmasak. Az önszerveződő nanogömb litográfiát sikerrel demonstráltak néhány száz nanométeres rácsávolságú periodikus struktúrák létrehozására [6]. Adott hullámhosszú expozíciónál a rácsállandó a nanogömbök méretével csökken, de egyre kisebb és kisebb gömbök alkalmazásakor a fókuszáló hatás fokozatosan elvész, és a hullámhossznál jóval kisebb nanogömbökből álló réteg homogén réteggé viselkedik.

Az elmúlt tíz évben a felületi plazmon effektus vizsgálata a nanolitográfiában is új lehetőségekre mutatott rá, ahol az egyik első jelentős eredmény a perforált fém szubsztrát réseiben kialakuló plazmonikus üregrezonátorokkal történő expozíció volt [7], [8]. A nanorészecske litográfiákban a fém nanorészecskék plazmonikus rezonanciái az előállítható minták méretének további csökkentését teszik lehetővé [9]. A plazmonikus rezonanciák igen erősen függenek a nanorészecskék méretétől és alakjától, és segítségükkel 100 nm-nél is kisebb rácsállandójú mintázatok alakíthatók ki [10]. A litográfiai elrendezés tervezése elektromágneses szimuláció

útján lehetővé teszi a kialakítandó mintázat megtervezését a részecskeméret függvényében.

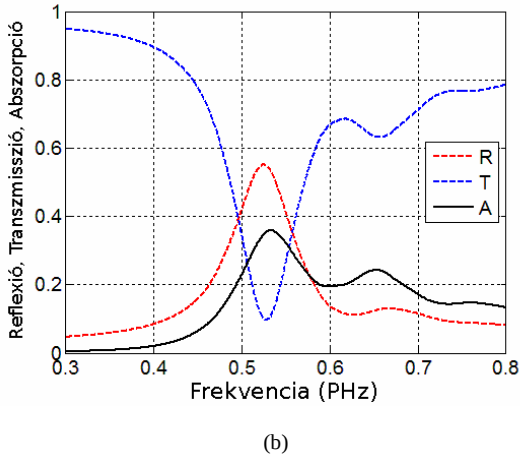
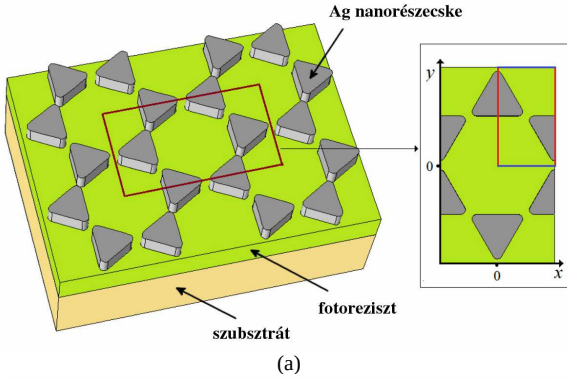
Az elektromágneses térszámítás fejlődésével számos különböző megoldó vált elérhetővé a Maxwell-egyenletekhez. A fém nanoszerkezetek szimulációja ezzel együtt speciális megfontolásokat tesz szükségessé. A mikrométeres és nagyobb hullámhosszak tartományán a fémek ideális vezetőként modellezhetők. Ahogy a hullámhosszat csökkentjük, és az összemérhetővé válik a behatolási mélységgel, a fémek átlátszóvá válhatnak, és optikai tulajdonságaik leírásához diszperzív anyagmodellre van szükség, mely a szabad és a kötött elektronok mozgását egyaránt figyelembe veszi. A fém szerkezetek nagy jósaági tényezőjű rezonanciái az időtartománybeli megoldók (pl. időbeli véges differencia megoldók) esetében hosszú futási időt eredményezhetnek. További kihívást jelent az, hogy a szerkezetek kis térfogatba képesek jelentős elektromágneses energiát koncentrálni, ami a frekvenciatartománybeli megoldóknál (pl. véges elemes megoldók) igen finom háló használatát teszi szükségessé. Az elektromágneses szimulációk eredménye a fotorezisztben kialakuló térerősség eloszlás, melyet különböző kioldást modellező algoritmusok bemenetként használva megbecsülhető a rezisztben előhívás során kialakuló mintázat.

Ebben a munkában egy újszerű bow-tie fémes nanorészecske rendszer elektromágneses tervezését mutatjuk be nemkonvencionális litográfia céljára. A fotorezisztben kialakuló elektromágneses tér eloszlását számoljuk, majd kétféle kioldási modell segítségével megbecsüljük a rezisztben előhívás után kialakult profilt.

II. FÉM NANORÉSZECSCKE LITOGRAFIA ELEKTROMÁGNES SZIMULÁCIÓJA

Kolloidkémiai eljárásokkal sokféle nanogömb készíthető, melyekből Langmuir-Blodgett technikával önszerveződő monoréteg képezhető a fotorezisztbe bevont üveg hordozó felszínén [1]. A maszkoló nanogömb monorétegen át fémet leválasztva, a gömbök közeiben bow-tie alakú nanorészecskék képződnek hatszöges rácsban. A gömbök eltávolítása után a fotorezisztet nagy felületen a bow-tie nanorészecskék rendszere fedi. A nanorészecskéket lekerekített csúcsú háromszög alapú hasábokkal modelleztük az 1.a ábra szerint. A háromszögek oldala 50 nm, vastagsága 20 nm, és a lekerekítés görbületi sugara 4 nm. A fotoreziszt vastagsága 20 nm.

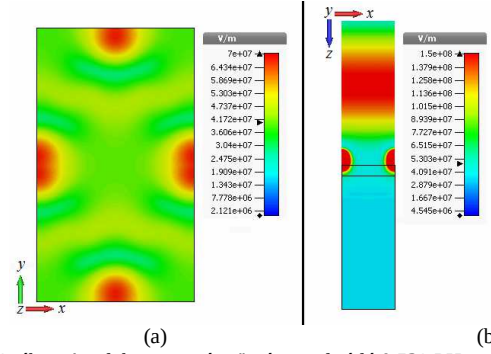
A modellezett nanostruktúra transzmisszióját és reflexióját a CST Microwave Studio időtartománybeli megoldójával számoltuk, és a frekvenciatartománybeli megoldóval ellenőriztük. A szimulációkat az 1.a ábrán látható szupercellán végeztük periodikus határfeltételek mellett. Az exponáló fény egy z-irányban haladó, lineárisan polarizált síkhullám, ahol x és y irányú polarizáció esetét vizsgáltuk. A struktúra periodicitását PEC és PMC határfeltételekkel modelleztük, és a számítási tartományt szimmetria feltételek előírásával negyedére csökkentettük. Az általunk vizsgált frekvenciákon a diffrakciótól eltekinthetünk, és ilyenkor ezek a határfeltételek megvalósítják a periodikus határfeltételt. A z irányban a számítási tartományt hullámvezető kapuk zárják le, melyek az exponáló elektromos sugárzást a rendszerbe bocsátják, a kimenő hullámot reflexió nélkül elnyelik, és automatikusan számolják a transzmisszió és reflexió adatokat. A bow-tie nanostruktúra anyaga ezüst, a fotoreziszt anyaga PMMA, a hordozó üvegből van. Az anyagok diszperzív elektromos permittivitása a SOPRA adatbázisából származik [11]. Az abszorpció spektrum, $A=1-T-R$ legmagasabb csúcsa 0.532 PHz-es frekvencián található mindkét polarizáció esetén, mely plazmonikus rezonancia jelenlétére utal ezen a frekvencián (1.b ábra).



1. ábra. A nemkonvencionális plazmonikus litográfiai elrendezés: ezüst bow-tie részecskerendszer az üveghordozóra felvitt fotoreziszt réteg felszínén és a szimulációkban használt szupercella felülnézeti képe (a). A számolt transzmisszió, reflexió és abszorpció spektrum y polarizációra (b)

Az elektromos térerősség amplitúdó eloszlás a rezonanciafrekvencián, y polarizáció esetén a 2. ábrán látható a számítási tartomány két síkmetszetére. A térerősség

maximumhelyei a nanorészecskék közeiben jönnek létre egy csatolt dipól plazmon módus eredményeképpen. A közökbén kialakuló erősítési helyeken a fotoreziszt lokálisan exponálódik, aminek hatását az előhívásra a következőkben előhívási modellek segítségével vizsgáljuk meg.



2. ábra. Az elektromos térerősség amplitúdó 0.532 PHz-en a nemkonvencionális litográfiai elrendezés két síkmetszetén: (a) az xy síkban a fotorezistréteg alján, (b) az xz síkban az y = 0 pozícióban

III. AZ ELŐHÍVÁS SZIMULÁCIÓJA

Az expozíció során a fotoreziszt kémiai tulajdonságai megváltoznak. Pozitív reziszt esetén, mint a PMMA, a reziszt oldhatósága megfelelő előhívó folyadékban az exponált területeken megnő. A reziszt előhívásának modellezésében az első lépés az elnyelt energiasűrűségből a lokális oldódási sebesség kiszámítása. A polimerreziszt oldhatóságának megváltozásában a legjelentősebb mechanizmus a polimerláncok megszakadása, mely lokálisan csökkenti a molekulatömeget az exponált területeken. Az oldódási sebesség a molekulatömeg függvényében az alábbi módon számolható [12]:

$$R(\mathbf{r}) = \left(R_0 + \frac{\beta}{M_f^\alpha(\mathbf{r})} \right) e^{-\frac{E_a}{kT}} \quad (1)$$

ahol R_0 , α és β egy adott előhívóra vonatkozó állandók. A lokális molekulatömeg M_f az abszorbeált energiasűrűség függvényében az alábbi kifejezéssel számolható [12]:

$$M_f(\mathbf{r}) = \frac{M_n}{1 + gW(\mathbf{r})M_n / \rho} \quad (2)$$

ahol M_n a fotoreziszt exponálás előtti átlagos molekulatömege, ρ a reziszt sűrűsége, A_0 az Avogadro-szám és g a reakció kémiai hatásfoka. A fentiek szerint számolt előhívási sebesség eloszlásból két kioldási modell segítségével szimuláltuk az előhívási profil időfejlődését.

A. Greneich-modell

Az első modell az elektronsugaras litográfiára kifejlesztett Greneich-modell módosított változata [12], [13]. A modell szerint az előhíváshoz szükséges időt a hely függvényében a következő képlet adja meg:

$$\tau(r) = \int_{r_0}^r \frac{1}{R(r)} dr \quad (3)$$

ahol r_0 a reziszt felszínén a maximális előhívási sebességű pont helye. A Greeneich-modell eredeti formájában felteszi, hogy a rezisztprofil fejlődés kezdetben a rezisztfelületre merőleges irányban halad. A modellt a fotolitográfiában jellemző profilok jobb közelítése érdekében úgy módosítottuk, hogy a (3) integrált egyenes vonalra végezzük el. Ebben a modellben elhanyagoltuk azt, hogy az előhívás több különböző úton különböző sebességgel egyszerre megy végbe, ehelyett csak a leggyorsabb terjedési úttal számolunk.

B. Sejtautomata-modell

A második előhívási modellünkben, a sejtautomata-modellben a reziszt térfogatot egyforma köbös cellákra osztjuk, melyek oldala a hosszúságú [10]. Egy adott (i, j, k) indexű cella állapotát a t időpontban definiálja a cellában található kioldott reziszt térfogat és a teljes térfogat aránya:

$$C_{i,j,k}^t = h_{diss} / a \quad (4)$$

ahol h_{diss} az előhívófolyadék ekvivalens magassága a köbös cellában. Minden cella kezdeti állapota nulla, és amikor a reziszt teljes egészében kioldódott egy adott cellából, akkor annak a cellának az állapota a további számolásban egy. A cellák állapota a szomszédos cellákból való előhívó átfolyás miatt iterációnként változik, melyet a differenciális cella állapot tagokkal veszünk figyelembe. A dC_{adj}^{dt} , dC_{edg}^{dt} , dC_{vtx}^{dt} differenciális tagok a lap-, él- és csúcshozsomszédok járulékát jelölik. Az (i, j, k) cella állapotát a t_0+dt időpontban a következők szerint számoljuk:

$$C_{i,j,k}^{t_0+dt} = C_{i,j,k}^{t_0} + dC_{adj}^{dt} + dC_{edg}^{dt} + dC_{vtx}^{dt} \quad (5)$$

Jelöljük az (i, j, k) cellában (1) szerint számolt előhívási sebességet $R_{i,j,k}$ -val! Ekkor a differenciális cella állapot tagok a következők:

$$\begin{aligned} dC_{adj}^{dt} &= \frac{\gamma_{adj} R_{i,j,k} dt}{a} (C_{i+1,j,k}^{t_0} + C_{i-1,j,k}^{t_0} + C_{i,j+1,k}^{t_0} + C_{i,j-1,k}^{t_0} + \\ &+ B_{i,j,k+1}^{t_0} + B_{i,j,k-1}^{t_0}), \\ dC_{edg}^{dt} &= \frac{2\gamma_{edg} R_{i,j,k}^2 t_0 dt}{a^2} (C_{i+1,j+1,k}^{t_0} + C_{i-1,j+1,k}^{t_0} + C_{i+1,j-1,k}^{t_0} + \\ &+ C_{i-1,j-1,k}^{t_0} + B_{i+1,j,k+1}^{t_0} + B_{i-1,j,k+1}^{t_0} + B_{i,j+1,k+1}^{t_0} + B_{i,j-1,k+1}^{t_0} + \\ &+ B_{i+1,j,k-1}^{t_0} + B_{i-1,j,k-1}^{t_0} + B_{i,j+1,k-1}^{t_0} + B_{i,j-1,k-1}^{t_0}), \\ dC_{vtx}^{dt} &= \frac{3\sqrt{3}\gamma_{vtx} R_{i,j,k}^3 t_0^2 dt}{8a^3} (B_{i+1,j+1,k+1}^{t_0} + B_{i-1,j+1,k+1}^{t_0} + \\ &+ B_{i+1,j-1,k+1}^{t_0} + B_{i-1,j-1,k+1}^{t_0} + B_{i+1,j+1,k-1}^{t_0} + B_{i-1,j+1,k-1}^{t_0} + \\ &+ B_{i+1,j-1,k-1}^{t_0} + B_{i-1,j-1,k-1}^{t_0}), \end{aligned} \quad (6)$$

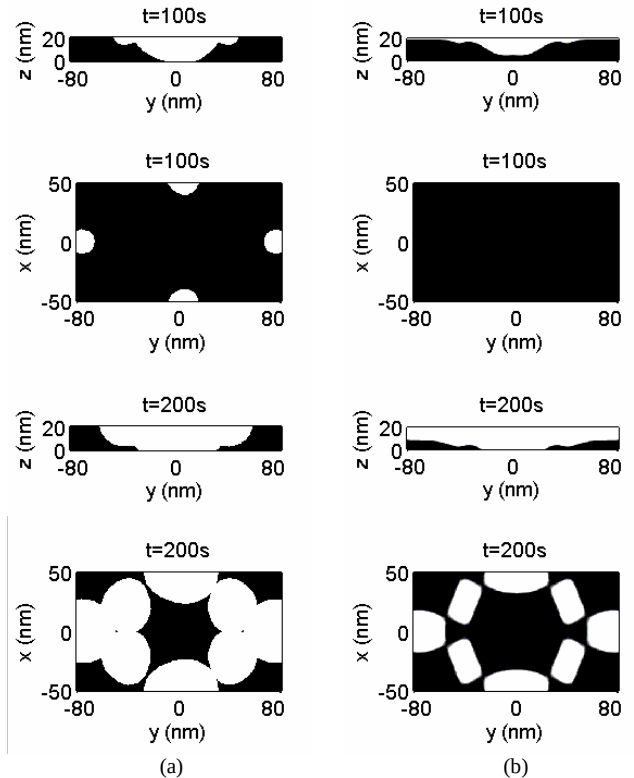
ahol a $B_{I,J,K}^{t_0}$ kétértékű függvény definíciója:

$$B_{I,J,K}^{t_0} = \begin{cases} 1, & \text{when } C_{I,J,K}^{t_0} = 1, \\ 0, & \text{otherwise.} \end{cases} \quad (7)$$

A kétértékű függvényt az algoritmus a $K = k-1$ és $K = k+1$ cellapozíciókban hívja meg, ami azt a modellfeltevést fejezi ki, hogy a cellákba alulról és felülről nincs előhívó átfolyás addig, amíg ezek a cellák nem teljes egészében kioldottak

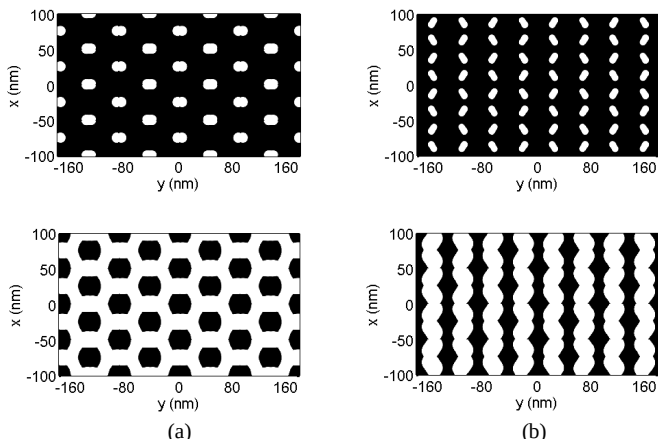
A kioldási modellek segítségével megbecsültük a 2. ábra szerint exponált reziszt előhívási profilját. Az előhívási sebesség számolásához használt paraméterek mindkét modellre: $R_0 = 4.326 \times 10^{19}$ Å/min, $\beta = 1.617 \times 10^{26}$, $\alpha = 1.5$, $E_\alpha = 1.04$ és $M_n = 10^5$ atomi tömegegység. A sejtautomata model paramétereit: $\gamma_{adj} = 10^{-4}$, $\gamma_{edg} = 0$ és $\gamma_{vtx} = 4 \times 10^{-6}$.

A számolt előhívási profilok két időpillanatban a 3.a és 3.b ábrán láthatóak y polarizációjú expozíció esetén. Az yz síkú metszetek az $x=0$ síkbeli állapotot, az xy síkú metszetek pedig a reziszt alját mutatják. Az előhívási kontrasztját a rezisztprofil meredeksége szabja meg. E fontos paraméter tekintetében a Greeneich-modellen alapuló szimulációnk nem ad megbízható eredményt az említett elhanyagolások miatt, de a mintázat jellegére validációként szolgál. A sejtautomata modellel számolt profil meredeksége sokkal kisebb, de ebben az esetben is lehet találni olyan előhívási időt (pl. 200 s), mellyel a mintázat kinyerhető.



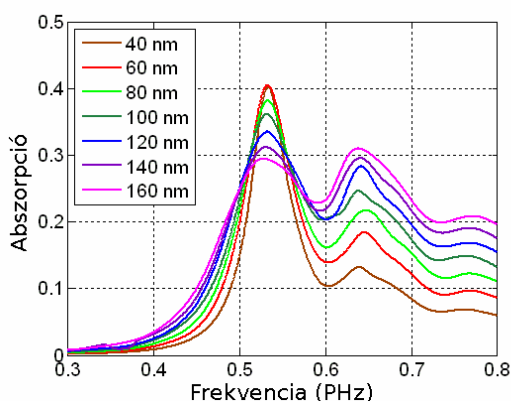
3. ábra. A bow-tie nanorészecske rendszerrel exponált fotoreziszt előhívási profilja két síkmetszetben. A 100 és 200 s-ig tartó előhívás utáni profil a Greeneich-modellel (a) és a sejtautomata modellel (b).

Az exponáló elektromágneses sugárzás polarizációjától függően kétféle rácsgeometriájú mintázatot állíthatunk elő. Y polarizáció esetén hatszögrácsú, x polarizációnál négyszöges rácsú a kialakuló mintázat. Az előhívási idő változtatásával a rácspontokban található bázis változtatható (4. ábra).



4. ábra. A bow-tie nanorészecske rendszerrel exponált fotoreziszt előhívási profilja az xy síkban a reziszt alján két előhívási időnél y polarizáció (a) és x polarizáció esetén (b).

A mintázat ráczállandójának változtatásához a nanorészecskék méretét változtathatjuk. Az 5. ábrán az elrendezés arányait állandónak tartva különböző méretű nanorészecskék mellett abszorpció spektrum látható. Ekkor a rezonanciafrekvencia azonos marad, de az abszorpció csúcs kisebb lesz, ahogy a plazmonikus rezonanciák mérettartományát elhagyjuk.



5. ábra. Az abszorpció spektrum a háromszögalapú nanorészecskék különböző oldalhosszáigainál. A nanorészecskék méretének növelésével az abszorpció csúcs magassága csökken, a szóródás válik a domináns kölcsönhatássá.

IV. KÖVETKEZTETÉS

Bemutattuk egy nemkonvencionális fotolitográfia elektromágneses és előhívás szimulációját [14]. A javasolt fémes bow-tie struktúra plazmonikus rezonanciát mutat, ami a nanorészecskék közeiben nagy elektromágneses tér erősítést eredményez. Így a reziszt réteg exponálható, és az előhívás után négy- illetve hatszögrácsban elhelyezkedő lyukakkal mintázott rezisztet kapunk, ahol a néhány száz tíz nanométeres ráczállandó is elérhető. A szimulációk segítségével meghatározható a nanorészecskék mérete, a reziszt vastagsága, az optimális megvilágító hullámhossz és az előhívás körülményei a kívánt mintázat elérése érdekében.

KÖSZÖNETNYILVÁNÍTÁS

Ezt a munkát a Magyar Tudományos Akadémia Bolyai János Kutatási Ösztöndíja és az EUREKA MetaFer projekt támogatásával jött létre.

HIVATKOZÁSOK

- [1] D. L. Schodek, P. Ferreira and M. F. Ashby, *Nanomaterials, Nanotechnologies and Design: An Introduction for Engineers and Architects*, Butterworth-Heinemann, 2009.
- [2] Wei, A. Katsnelson, O. G. Memis, H. Mohseni, "A deep sub-wavelength process for the formation of highly uniform arrays of nanoholes and nanopillars", *Nanotechnology*, Vol. 18, No 48, pp. 1-4, 2006.
- [3] H. D. Wanzenboeck and S. Waid, "Focused Ion Beam Lithography" in *Recent Advances in Nanofabrication and Applications*, InTech, 2011.
- [4] S. Y. Chou, P. R. Krauss, P. J. Renstrom, "Nanoimprint Lithography", *Journal of Vacuum Science and Technology B*, 14(6), 4129-4133, 1996.
- [5] H. Zhang, S. W. Chung, C. A. Mirkin, "Fabrication of sub-50-nm solid-state nanostructures on the basis of dip-pen nanolithography", *Nano Letters* 3 (1), 43-45, 2003.
- [6] W. Y. Fu and H. W. Choi, "Nanosphere Lithography for Nitride Semiconductors", in *Lithography*, Michael Wang, Croatia: Intech, 2010, pp. 615-628.
- [7] T. W. Ebbesen, H. J. Lezec, H. F. Ghaemi, T. Thio, P. A. Wolff, "Extraordinary optical transmission through sub-wavelength hole arrays", *Nature* 391, 667-669, 1998.
- [8] W. Srituravanich, N. Fang, C. Sun, Q. Luo, and X. Zhang, "Plasmonic nanolithography", *Nano Letters* vol 4 1085, 2004.
- [9] S. A. Maier, *Plasmonics: Fundamentals and Applications*, Springer, 2007.
- [10] V. M. Murukeshan, K. V. Sreekanth and J. K. Chua, "Metal Particle-Surface System for Plasmonic Lithography", in *Lithography*, Michael Wang, Croatia: Intech, 2010, pp. 598-614.
- [11] <http://www.spectra.com/sopra.html>, last visited: 27 December 2013.
- [12] Greeneich J. S., "Time evolution of developed contours in polymethyl methacrylate electron resist", *Journal of Applied Physics*, 45, No. 12, pp. 5264-5268, 1974.
- [13] M. Stepanova, T. Fito, Zs. Szabó, K. Alti, A.P. Adeyenuw, K. Koshelev, M. Aktary and S.K. Dew, "Simulation of electron beam lithography of nanostructures", *Journal of Vacuum Science and Technology B*, Vol. 28, pp. C6C48- C6C57, 2010.
- [14] A. Vermes, Zs. Szabó, "Unconventional photolithography with self-assembled plasmonic nanostructures", in *Proceedings of the COMPUMAG Conference on the Computation of Electromagnetic Fields*, 2013.

Végeselemes megoldó párhuzamos implementációja MPI segítségével

Budai Tamás

Elektromágneses Terek Laboratórium
Automatizálási Tanszék, Széchenyi István Egyetem
Győr, Magyarország
budait@maxwell.sze.hu

Prof. Dr. habil. Kuczmán Miklós

Tanszékvezető, Egyetemi tanár
Automatizálási Tanszék, Széchenyi István Egyetem
Győr, Magyarország
kuczmán@maxwell.sze.hu

Kivonat—Cikkünkben egy korábban *MATLAB* környezetben implementált végeselemes megoldó párhuzamosításának lépéseit mutatjuk be az *MPI*[5] szabvány és a *PETSc*[3] függvénykönyvtár felhasználásával C programnyelven. Az új megoldó eredményeinek helyességét az analitikus megoldással validált *MATLAB* megoldók segítségével ellenőrizzük. A működő, párhuzamos futtatásra alkalmas programot egy általános kiszolgálói hardverekből erre a célra kialakított tesztkörnyezetben egy ún. *Beowulf-cluster*-en futtatjuk, majd mérésrel megállapítjuk azt az optimális konfigurációt, amelyen maximális sebességnövekedés érhető el.

Kulcsszavak—*párhuzamos programozás, szimuláció, végeselem-módszer, MPI, cluster*

I. BEVEZETÉS

Napjainkban mind az ipar mind a tudományos kutatások területén széles körben alkalmaznak szimulációs eljárásokat egy-egy fizikai jelenség vizsgálatához. Ezen eljárások numerikus módszerek használatával képesek a fizikai jelenséget leíró bonyolult egyenletrendszerek közelítő megoldására. Az egyik legelterjedtebb ilyen eljárás a villamosmérnöki gyakorlatban az ún. végeselem-módszer[2]. A számítástechnika rohamos fejlődésének köszönhetően egyszerűbb modellek esetén a szimuláció akár egy általános használatra szánt „irodai” számítógépen is futtatható, ám a modellek komplexitásának növekedésével a megoldásra szánt idő is jelentősen növekszik, ami az alkalmazhatóság szempontjából kedvezőtlen. Mivel manapság az összes számítógép, illetve már az okos telefonok is többmagos processzorokkal kerülnek piacra, kézenfekvő irány a numerikus számítások gyorsítására a párhuzamos programozási technikák használata. Bár jelenleg a legnépszerűbb ilyen kutatások a grafikuskártyák használatát célozzák, fontosnak érezzük a CPU alapú párhuzamosítás alkalmazását is, mivel így a meglévő infrastruktúra számítási teljesítménye is kiaknázzható. Kutatásunk során ezért az *MPI*[5] szabvány segítségével olyan implementációt készítettünk, mely nem igényel különleges hardver-architektúrát és egyszerre több, hálózatba kapcsolt számítógépen párhuzamosan futtatható.

II. PÁRHUZAMOS PROGRAMOZÁSI TECHNIKÁK

II-A. Memória modellek

A párhuzamosan futó programok két nagy csoportba sorolhatók memóriamodelljük alapján. Az ún. *shared-memory* (megosztott memória) csoportba tartozó programok esetén a független végrehajtási *szálak* közös memóriaterülethez férnek hozzá. Ennek előnye, hogy a közös memóriaterületen

tárolt adatokhoz bármely *szál* hozzáférhet, hátránya viszont, hogy a program futtathatóságát egyetlen fizikai számítógépre korlátozza. Bár folynak kutatások olyan transzparens réteg létrehozására amellyel több különálló számítógép erőforrásai egyként használhatók, ezek nem kifarítottak. A másik megoldás az ún. *distributed-memory* (elosztott memória) modell használata. Ebben az esetben minden egyes végrehajtási *szál* saját memóriaterületet használ. A megoldás nagy előnye, hogy az egyes *szálak* akár több, hálózatba kapcsolt számítógépen azaz *CLUSTER*-en is futtathatók, azonban itt szükség van valamilyen kommunikációs csatornára az egyes *szálak* közti adatcsere lebonyolításához. Jelenleg ez a modell a legelterjedtebb, a világ nagy szuperszámítógépein is ezt alkalmazzák.

II-B. Az *MPI* szabvány

A végrehajtási *szálak* közti adatcsere egyik lehetséges módja az *MPI* szabvány[5] használata. Programjaink fejlesztése során a szabvány egyik nyílt-forráskódú implementációját az *OpenMPI*-t[6] választottuk, azonban a szabványosságnak köszönhetően a programok más *MPI* implementációk használata esetén is minimális módosítás után lefordíthatóak.

II-C. A *PETSc* függvénykönyvtár

A későbbiekben ismertetett számításokhoz az *MPI* szabványt is felhasználó, nyílt forráskódú *PETSc*[3] függvénykönyvtárt alkalmaztuk. Ez a C és FORTRAN programnyelvekhez készült függvénykönyvtár lineáris algebrai egyenletrendszerek szekvenciális és párhuzamos megoldásához is tartalmaz komponenseket és a mérnöki gyakorlatban számos helyen alkalmazzák meglévő szimulációs szoftverek párhuzamosítására[7][8].

III. A VÉGESELEM-MÓDSZER

A meglévő megoldók működésének megértéséhez, vizsgálatához és a párhuzamosításhoz szükséges a végeselem-módszer ismerete, ezért röviden bemutatjuk a módszer egyes lépéseit. Terjedelmi okból csak a párhuzamosítás szempontjából lényeges lépéseket tárgyaljuk részletesen. Részletesebb leírást ad [1] és [2].

III-A. Előfeldolgozás

Ebben a lépésben történik meg a modell specifikációja. A modell bonyolultságától függően a vizsgálandó geometriát általában valamilyen CAD szoftver segítségével definiáljuk,

majd ehhez hozzárendeljük a modellt alkotó egyes anyagok fizikai jellemzőit és definiáljuk a modell határait az ún. peremfeltételeket. Ebben a lépésben kihasználhatjuk a modellben található esetleges szimmetriákat és ezzel csökkenthetjük annak méretét.

III-B. Diszkrétizálás

A modellt egy ún. rácsgenerátor segítségével véges számú elemre bontjuk fel. A használt végelem típusok 1 dimenzióban a vonal, míg 2 dimenzióban leggyakrabban háromszög, három dimenzióban pedig tetraéder elemeket alkalmaznak. A rács elkészítésekor fontos ügyelnünk arra, hogy az nem tartalmazhat átfedéseket sem üres részeket. A cikkben ismertetett 2 és 3 dimenziós feladatok rácsoszáshoz a nyílt forráskódú *GMSH*[4] szoftvert használtuk fel. Ez a szoftver megfelelő beállítás esetén olyan formátumban szolgáltatja a rácsot, amelyből az általunk későbbiekben használt struktúrát viszonylag könnyen elő tudjuk állítani. Az egyes elemeket alkotó csúcspontokat *csomópontoknak* nevezzük.

III-C. Asszembhlálás

Ebben a lépésben a fizikai viselkedést definiáló egyenletekből, egy-egy elemre felírható *elemegyenleteket*[1] állítunk elő, majd az ezekből kialakuló egyenletrendszert egy mátrixba (ún. *együtthatómátrixba*) és két vektorba rendezzük. Az együtthatómátrix minden sora az egy-egy elemre felírt egyenlet bal oldalának együtthatóit, az $\underline{u}$ oszlopvektor az ismeretleneket, míg a $\underline{b}$ oszlopvektor a peremfeltételeket tartalmazza. Ezek alapján felírható a megoldandó egyenletrendszer, ahol az együtthatók és az ismeretlenek szorzata kielégíti a peremfeltételeket, azaz

$$\underline{K} \underline{u} = \underline{b}. \quad (1)$$

III-D. Megoldás

Az (1) egyenlet kézenfekvő megoldása előállítható

$$\underline{K}^{-1} \underline{b} = \underline{u}, \quad (2)$$

azonban a feladat és ezzel a mátrix méretének növekedésével az inverz-mátrix képzése egyre nehezebbé és időigényesebbé válik. Ezért az ilyen egyenletrendszerek megoldására a gyakorlatban különböző *iteratív* megoldó algoritmusokat használnak, amelyek közelítő megoldást adnak.

III-E. Utófeldolgozás

Ebben a lépésben a megoldó által szolgáltatott eredményt megjeleníthetjük a rácsra illetve, illetve további számításokkal egyéb paramétereket is meghatározhatunk. Az eredmény ismeretében képet kapunk a vizsgált modell fizikai viselkedéséről.

IV. A TESZTFELADAT

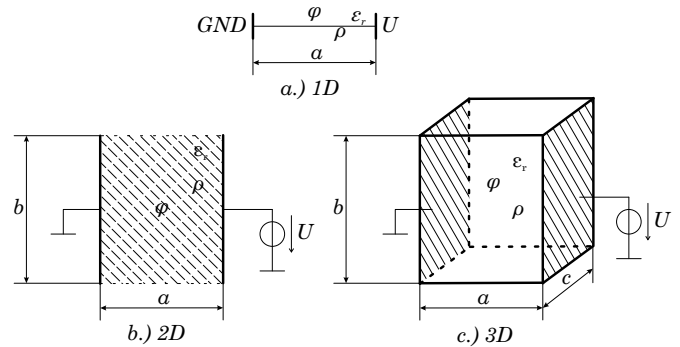
A megoldók teszteléséhez egy ismert példát választottunk az elektromágneses térszimuláció témaköréből, melyhez rendelkezésre állnak az általunk korábban MATLAB környezetben megvalósított, analitikus megoldás segítségével ellenőrzött megoldók. A feladat egy síkkondenzátor lemezei közötti dielektrikum kialakuló potenciáeloszlás meghatározása, amennyiben a fegyverzetek távolsága a , a kondenzátor egyik fegyverzete földpotenciál, a másik $U = 10V$ potenciál

van, a kitöltő dielektrikum relatív permittivitása $\epsilon_r = 10$, valamint a dielektrikumban $\rho = 10^{-2} \frac{C}{m^3}$ töltéssűrűség van jelen.

V. MEGOLDÁS VÉGESELEM-MÓDSZERREL

V-A. Modellalkotás

Bár ez az egyszerű feladat megoldható 1 dimenziós modellel is, a párhuzamosítás hatékonyságának vizsgálata céljából a feladatot 1, 2 és 3 dimenziós modellel is megoldottuk, különböző geometriai méretek és rácssűrűségek mellett. Ezen modelleket mutatja a 1. ábra.



1. ábra. A feladat modellje 1, 2 és 3 dimenzióban

Amennyiben a fegyverzetek a távolsága kellően kicsi a fegyverzetek kiterjedéséhez képest, a potenciáeloszlást homogénnek tekinthetjük a dielektrikum teljes hosszában, így egydimenziós modell használata elegendő.

Amennyiben a kitöltő dielektrikum teljesen homogén, a megoldás egyszerűen kiterjeszthető, hiszen a potenciál értéke továbbra is csak az a méret mentén felvett távolságtól függ. A kondenzátor fizikai méretei 2 dimenziós modell esetén $a = 10mm$ és $b = 15mm$, míg 3 dimenzió esetén $a = 10mm$, $b = c = 15mm$.

V-B. Adatstruktúra

Ahhoz, hogy az együtthatómátrixot fel tudjuk tölteni, szükségünk van az egyes csomópontok koordinátáira, valamint arra az információra, hogy egy-egy elem mely csomópontokból áll és melyik előre definiált tartományban található. Ehhez egy három részegységből álló adatstruktúrát hoztunk létre:

1) *Coords*: A csomópontok koordinátáit egy $n \times m$ méretű mátrixban tároljuk, ahol n a csomópontok, m a dimenziók száma.

2) *Connect*: Az egyes elemek és az azokat alkotó csomópontok relációját egy $i \times j$ méretű mátrixban tároljuk, ahol i az elemek száma, j pedig a felhasznált elemek típusától függ. Háromszögelemek esetén $j = 4$, ahol az első oszlop egy azonosító ami alapján eldönthető az elemről, hogy az mely tartományhoz tartozik, a másik három oszlop pedig az elemet alkotó három csomópont számát tartalmazza. Amennyiben ezzel a számmal indexeljük a *Coords* struktúrát, megkaphatjuk az egy-egy háromszöghöz tartozó csomópontok koordinátáit.

3) *Bnd*: A peremfeltételeket tartalmazó struktúra minden esetben $k \times 2$ méretű ahol k azon csomópontok száma, amelyekre peremfeltételt kívánunk előírni. Az első oszlop tartalmazza a perem azonosítóját, míg a második az adott csomópont sorszámát.

V-C. Megoldás

Az 1 dimenziós modellt a programban vonal elemekre bontottuk fel, majd az ezekre felírt elemegyenletekből állítottuk össze az együtthatómátrixot, majd a peremfeltételek definiálása után megoldottuk a feladatot. A 2 dimenziós modell esetén háromszög elemeket, 3 dimenzióban pedig tetraéder elemeket alkalmaztunk.

V-D. Műveletigény

A formafüggvények alkalmazása miatt, egy-egy elemegyenlet együtthatóinak meghatározása a dimenziók növelésével egyre nagyobb műveletszámot igényel. Míg 1 dimenziós vonalelem esetén 2 gradiens kiszámítására és két középpont meghatározására van szükség, addig 2 dimenziós esetben 6 gradiens és az adott háromszög területének kiszámítása is szükséges. A felépített egyenletrendszer megoldásának műveletigénye a dimenziókkal szintén növekszik, hiszen egy-egy csomópont több elemhez is kapcsolódik, így az együtthatómátrix egyre több nem-nulla elemet tartalmaz.

VI. PÁRHUZAMOSÍTÁS

Egy szekvenciális algoritmus vizsgálata során két fajta lépéstípust különböztethetünk meg adatfüggőség szerint. Amennyiben egy lépés végrehajtásához szükség van ugyan-ezen lépés korábbi végrehajtása során keletkezett adatra, a probléma *inherensen szekvenciális* és nem létezik mód a párhuzamosítására. Az elektromágneses térszimuláció területén az időtartománybeli analízis során alkalmazott időlépés tipikusan ilyen probléma. A másik esetben az adott lépés csak olyan adatokkal végez műveletet amelyek már rendelkezésre állnak. Ebben az esetben ezt a lépést azonos időben többször, azaz párhuzamosan is elvégezhetjük.

VI-A. Mester, szolga szálak

A programban megkülönböztetünk egy *mester* és nulla vagy több *szolga* szálát. Szekvenciálisan futtatva egy MPI-t használó program a 0 azonosítójú szálát hozza létre, így célszerű ezt kijelölni mesternek. Ilyen módon a párhuzamos algoritmusunk alkalmassá válik szekvenciális futtatásra is, megkönnyítve a hibakeresést.

A korábban MATLAB környezetben implementált 1, 2 és 3 dimenziós megoldókat vizsgálva és az ismertett kritériumokat figyelembe véve a következő egyszerűen párhuzamosítható részeket azonosítottuk:

- az elemegyenletek kiszámítása
- asszemblálás
- peremfeltételek megadása

Ennek megfelelően a megoldót C programnyelven újraírtuk és a felsorolt három lépés végrehajtását párhuzamosítottuk. Az asszemblálás közös együtthatómátrixba történik, az egyes

szálak pedig az elemek egy-egy diszjunkt részhalmazáért felelősek. Nagyon fontos, hogy egy-egy elemet csak egy szál dolgozzon fel, mert abból adódóan, hogy egy-egy csomópont több elemnek is része, K egy-egy cellájába több elem esetén is additívan helyezünk el értékeket. A peremfeltételek megadásának párhuzamosítását hasonlóan végeztük el. A b vektor közös, minden szál végiglép az összes peremfeltételen, de csak akkor helyez elemet a vektorba, ha az a hozzá kirendelt részben található.

A megoldáshoz minden esetben a PETSc Kyrlov-altér[3] módszert alkalmazó iteratív megoldóját használtuk, amely képes a megoldást is párhuzamosan elvégezni. Így elmondhatjuk, hogy a megoldó teljesítménykritikus részeit párhuzamosítani tudtuk.

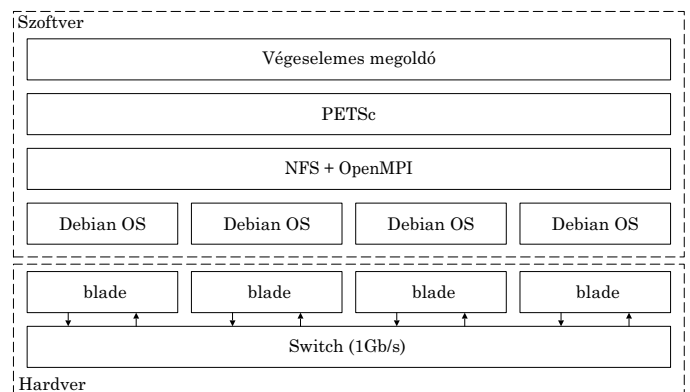
VII. SEBESSÉGMÉRÉS

VII-A. Futtatókörnyezet

Az elkészült programok teljesítőképességének vizsgálatához, egy ún. *Beowulf-cluster*[8]-t alakítottunk ki Egyetemünk Távközlés-Informatika Laboratóriumában található 4 darab IBM HS21-8853 típusú számítógép segítségével. Ezen gépek specifikációja egyenként a következő:

CPU: Intel(R) Xeon(R) CPU L5240 @ 3.00GHz
Memory: 8GB FBDIMM 667MHz
HDD: WD 80GB SAS
Network: BCM5708S Gigabit Ethernet

Operációs rendszerként Debian Linux 7-t használtunk, erre fordítottuk le forrásból az OpenMPI és a PETSc csomagokat. Az MPI szabvány megköveteli, hogy a futtatandó bináris minden gépen elérhető legyen. Ezt úgy teljesítettük, hogy a mesternek kijelölt gép /home mappáját *NFS* segítségével megosztottuk a gépek között. Az elkészült rendszer hardveres és szoftveres felépítését mutatja a 2. ábra.



2. ábra. Az elkészült Beowulf-cluster

VIII. A MÉRÉS MENETE

A futási idő méréséhez a linux *time* parancsát használtuk a valóságban eltelt idő (*Wall-time*) mérésére, mivel a megoldó végfelhasználása szempontjából ez a leglényegesebb paraméter. Mivel az operációs rendszerben található processz-ütemező és gyorsítótárak befolyásolhatják egy program futási idejét, minden esetben 4 mérést végeztünk és ezek átlagát vettük

alapul. A mérések alatt a gépeken semmilyen más felhasználói folyamat nem futott.

A sebességnövekedést arányszámban fejeztük ki

$$S = \frac{T_s}{T_p} \quad (3)$$

szerint, ahol T_s a szekvenciális, míg T_p a párhuzamos program futásideje.

IX. MÉRÉSI EREDMÉNYEK

IX-A. Megoldások validálása

A párhuzamos megoldók által szolgáltatott eredményeket a MATLAB környezetben elkészített megoldók eredményeihez hasonlítottuk, azonos rácsok és peremfeltételek mellett. A maximális eltérések nagyságrendje minden esetben 10^{-11} vagy kisebb nagyságrendű volt, így a gyakorlat szempontjából ezen eltérések elhanyagolhatóak. Kijelenthetjük tehát, hogy a párhuzamos implementáció a gyakorlatban használható eredményt szolgáltat.

IX-B. Futási idő

Annak érdekében, hogy átfogó képet kapjunk az elkészített programok teljesítményéről, sűrű rácsok használata mellett mértük a futási időket. Míg egy gép használata esetén a szálak kommunikációs sebességét csak a memória sávszélesség korlátozza, addig több gép használata esetén az azokat összekötő hálózat szűkebb keresztmetszetet jelent, így az egyes méréseket 1, 2, 3 és 4 gép használatával is elvégeztük. A mért futásidőket az alábbi táblázatokban foglaltuk össze.

I. táblázat. 1D szimuláció, 100000 vonal elemmel

szálak	1	2	4	8	16
gépek	Végrehajtási idő [mp]				
1	2,478	1,299	1,247	1,766	3,796
2		0,452	1,406	1,779	3,481
3			1,472	2,317	5,844
4			0,457	1,776	3,33

II. táblázat. 2D szimuláció, 162626 háromszög elemmel

szálak	1	2	4	8	16
gépek	Végrehajtási idő [mp]				
1	52,633	39,658	27,906	38,404	47,083
2		41,744	30,785	25,034	27,4155
3			43,351	133,851	64,231
4			34,404	134,026	181,809

III. táblázat. 3D szimuláció, 94616 tetraéder elemmel

szálak	1	2	4	8	16
gépek	Végrehajtási idő [mp]				
1	46,437	8,558	6,363	8,093	11,032
2		8,235	6,079	5,047	6,762
3			6,504	6,181	7,815
4			5,352	5,416	7,936

Látszik, hogy 1D szimuláció esetén, mivel egy-egy elem-egyenlet együtthatóinak számítási igénye csekély, a szálak közti adatcsere időigénye 4 szál felett már nagyobb, mint a párhuzamosítás által elérhető csökkenés, így ezen szimulációkat nem éri meg ennél több szálon futtatni. A 2D szimulációnál

ezzel szemben már 8 párhuzamos szál és 2 gép használata esetén érhető el a legjobb eredmény. A 3D szimuláció esetében is ugyan ez a konfiguráció bizonyult optimálisnak. A szálak, illetve a számításba bevont gépek számának növelésével növekszik az adatcserehez szükséges idő is, ez pedig mindhárom esetben kedvezőtlenül befolyásolja a futásidőt.

Mivel a felhasznált gépek egyenként 4 magos processzorokat tartalmaznak, az eredményekből az is jól látszik, hogy a legjobb eredmény akkor érhető el, ha egy-egy gép minden processzormagjához egy végrehajtási szál tartozik. Az optimális konfigurációk mellett elért maximális sebességnövekedéseket az IV. táblázatban foglaltuk össze.

IV. táblázat. Sebességnövekedés

Feladat	konfig. [szál/gép]	T_s [s]	legjobb T_p [s]	S
1D	2/2	2,478	0,452	5,582
2D	8/2	52,633	25,034	2,102
3D	8/2	46,437	5,047	9,200

A legnagyobb nyereséget a 3 dimenziós szimuláció esetében sikerült elérni, ahol ideális hardverkonfiguráció esetén a párhuzamos megoldó 9-szer gyorsabban szolgáltatott eredményt.

X. ÖSSZEGZÉS

Bemutattuk egy végeselemes megoldó párhuzamosításának lépéseit és az elérhető sebességnövekedést egy egyszerű feladaton keresztül. A párhuzamos megoldók továbbra is helyes eredményeket produkáltak miközben a futtatási idő jelentősen lecsökkent. Az implementált megoldók a későbbiekben alapul szolgálhatnak egy teljesen dinamikus, gyakorlati feladatokat rövid idő alatt megoldani képes párhuzamos szimulációs szoftverhez, amelyet olcsó hardverelemekből épített *cluster*-en futtathatunk.

KÖSZÖNETNYILVÁNÍTÁS

A bemutatott munka a *TÁMOP-4.2.2.A-11/1/KONV-2012-0012: Hibrid és elektromos járművek fejlesztését megalapozó kutatások* projekt keretében, a Magyar Állam és az Európai Unió támogatásával, az Európai Szociális Alap társfinanszírozásával valósult meg.

HIVATKOZÁSOK

- [1] M. Kuczmán, A. Iványi. „The Finite Element Method in Magnetism” Akadémiai Kiadó, Budapest, 2008.
- [2] Iványi A. „Folytonos és diszkrét szimulációk az elektrodinamikában” Akadémiai Kiadó, Budapest, 2003.
- [3] S. Balay, M. F. Adams et. al. PETSc Users Manual. ANL-95/11 - Revision 3.4 <http://www.mcs.anl.gov/petsc>. 2013.
- [4] C. Geuzaine and J.-F. Remacle. „Gmsh: a three-dimensional finite element mesh generator with built-in pre- and post-processing facilities” International Journal for Numerical Methods in Engineering 79(11), pp. 1309-1331, 2009.
- [5] Message Passing Interface. <http://www.mpi-forum.org/>
- [6] OpenMPI. <http://www.open-mpi.org/>
- [7] S. Abhyankar, B. Smith et. al. Using PETSc to Develop Scalable Applications for Next-Generation Power Grid. Proceedings of the first international workshop on High performance computing, networking and analytics for the power grid pp. 67-74, 2011.
- [8] V.E. Sonzogni, A.M. Yommi et. al. A parallel finite element program on a Beowulf cluster. Advances in Engineering Software vol. 33 pp. 427-443. 2002

Nemlineáris anyagtulajdonságok modellezése a numerikus térszámításban

Unger Tamás István
Automatizálási Tanszék
Széchenyi István Egyetem
Elektromágneses Terek Laboratórium
Győr, Magyarország
ungert@maxwell.sze.hu

Prof. Dr. habil. Kuczmann Miklós
Automatizálási Tanszék
Széchenyi István Egyetem
tanszékvezető, egyetemi tanár
Győr, Magyarország
kuczmann@maxwell.sze.hu

Kivonat—A cikk az alacsonyfrekvenciás elektromágneses térszimuláció témakörében íródott, legfontosabb célkitűzése a nemlineáris anyagtulajdonságok modellezési lehetőségének felvázolása. Bemutatásra kerül a mágneses térerősség és a mágneses indukció között definiált skaláris hiszterézis mérésének elve, valamint a mérési eredményekből felépíthető nemlineáris, analitikus modell paramétereinek identifikációja, továbbá a modellből származtatható permeabilitásfüggvény és jelentősége is. Prezentálásra kerül a direkt modell illesztésének lehetősége egydimenziós és kétdimenziós végelelemes szimulációhoz, mely kapcsán felvázolásra kerül az illesztéshez szükséges iterációs séma, továbbá a térszimulációhoz használt parciális differenciálegyenletek, azok gyenge alakjai, implementációjuk végelelem-módszer segítségével, végül kiértékelésre kerülnek a szimulációs eredmények.

Kulcsszavak—ferromágneses hiszterézis, végelelem-módszer, elektromágneses terek, nemlinearitás, szimulációs módszerek

I. BEVEZETÉS

A számítástechnika rohamos fejlődésének és az iparban elterjedt szimulációs szoftverek térnyerésének köszönhetően napjainkban már alapvető feladat és elvárás a kutatási és fejlesztési területen a különböző anyagparaméterek jellemzőinek szimulációja, modellezése és az adott feladathoz optimalizálása. A villamos gépek, motorok és transzformátorok tervezése során különösen fontos az adott feladathoz illeszkedő (például kis veszteségű) vasanyag kiválasztása.

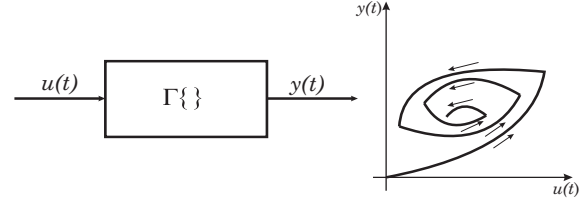
Az elektromágneses szempontból értelmezett anyagparaméterek optimalizálását numerikus térszámítással, így végelelem-módszerrel is hatékonyan lehet segíteni. A cikkben felvázolásra kerülő modell-illesztési módszer rendkívül rugalmas, mert a kereskedelmi forgalomban kapható szimulációs szoftverekkel ellentétben bonyolultabb összefüggések megvalósítására is lehetőséget nyújt.

II. A SKALÁRIS HISZTERÉZIS ÉS MÉRÉSE

Skaláris hiszterézissel bír az 1. ábrán látható általános, egy bemenetű, egy kimenetű rendszer, melynek $\Gamma\{\}$ operátora az

$$y(t) = \Gamma\{u(t)\} \quad (1)$$

alakban felírható, nemlineáris, többértékű, memóriával rendelkező gerjesztés-válasz kapcsolatot definiálja [1].



1. ábra. Egy általános hiszterézises rendszer

Ilyen típusú jelátalakítók a ferromágneses anyagok, amennyiben gerjesztésnek $\vec{H}(t)$ mágneses térerősség, válaszjelnek $\vec{M}(t)$ mágnesezettség, vagy $\vec{B}(t)$ mágneses indukció kerül megválasztásra [2]:

$$\vec{B}(t) = \mu_0 \left(\vec{H}(t) + \mathcal{H}\{\vec{H}(t)\} \right) = \mathcal{B}\{\vec{H}(t)\}. \quad (2)$$

Skaláris kapcsolat definiálható a térintenzitás és a gerjesztettség mennyiség között, amennyiben a vizsgált anyagon belül vektoraik párhuzamosak. Skalárkarakterisztika felvételére kiválóan alkalmas a toroid transzformátor [3], mely elrendezés esetén a mágneses térerősség

$$H(t) = \frac{N_p I(t)}{2R_k \pi} \quad (3)$$

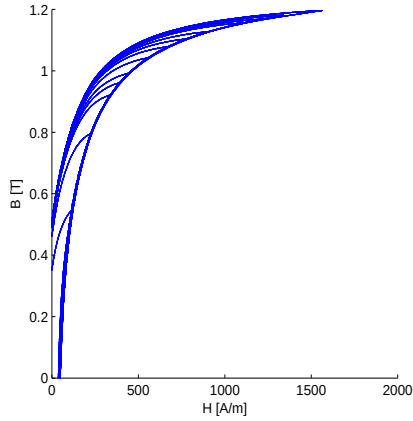
formula alapján a primer oldali menetszám, a gerjesztőáram és a közepes sugár ismeretében számít-

ható, továbbá a mágneses indukció

$$B(t) = B_0 + \frac{1}{N_{sz}A} \int_0^t u_i(\tau) d\tau \quad (4)$$

összefüggéssel, az indukált feszültség idő szerinti integrálásával számítható [3].

A mérés LabVIEW számítógépes mérőrendszer segítségével történt, mely rendszerrel egy áramgenerátor a primer oldali árama került vezérlésre, továbbá a szekunder oldali indukált feszültség került mérésre és feldolgozásra. A mérés során a modell identifikációjához szükséges koncentrikus skalárgörbék kerültek felvételre, mely görbéket a gerjesztőáram amplitúdójának fokozatos növelésével lehet meghatározni. A mérési eredmények a 2. ábrán láthatók.



2. ábra. A mért koncentrikus görbék

III. A MODELL IDENTIFIKÁCIÓJA

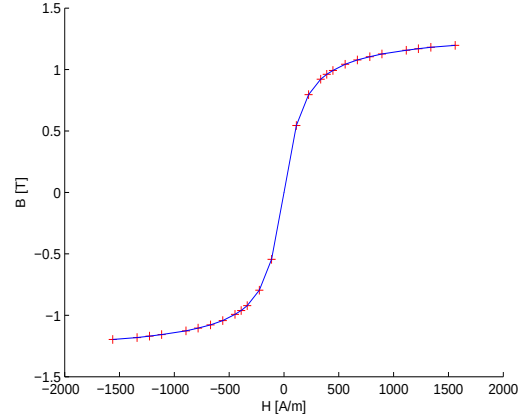
A szimulációhoz illesztett nemlineáris, egyértékű karakterisztikát az i darab mért koncentrikus $B(H)$ -függvények ($\arg \max_H B_n(H)$, $\max_H B_n(H)$), valamint ($\arg \min_H B_n(H)$, $\min_H B_n(H)$) pontjai határozzák meg ($n = 1, 2, \dots, i$). Ezen pontokra analitikus formában leírható, paraméterezett inverz tangens-görbe illeszthető. A

$$B(H, t) = \frac{2B_s}{\pi} \arctg \left(\frac{H(t)}{H_0} \right) \quad (5)$$

alakban felírható direkt modell formulájában B_s paraméter a telítési indukcióérték, valamint H_0 a karakterisztika meredekségét befolyásolja.

A modell paramétereinek identifikációja a Matlab *Curve Fitting Toolbox* moduljának segítségével történt meg. A vizsgált anyag paraméterei: $B_s = 1,25805 \text{ T}$, $H_0 = 146,47722 \frac{\text{A}}{\text{m}}$. A mért

karakterisztikák pontjai, valamint az illesztett görbe a 3. ábrán látható.



3. ábra. A mért pontok és az illesztett modell

A permeabilitásfüggvény a karakterisztikát leíró összefüggés H -szerinti parciális deriválásával fejezhető ki:

$$\mu(H) = \frac{\partial B}{\partial H} = \frac{2B_s}{\pi H_0} \frac{1}{1 + \left(\frac{H}{H_0} \right)^2}. \quad (6)$$

A függvény maximuma $H = 0 \frac{\text{A}}{\text{m}}$ -nél $\mu_{max} = 1,3009 \frac{\text{Vs}}{\text{Am}}$, minimuma $H \in [-2000 \frac{\text{A}}{\text{m}}, 2000 \frac{\text{A}}{\text{m}}]$ esetben $\mu_{min} = 0.0054 \frac{\text{Vs}}{\text{Am}}$. Az intervallum kikötése azért szükséges, mert a formula határértéke $H \rightarrow \pm\infty$ esetén 0.

IV. A MODELL ILLESZTÉSE A SZIMULÁCIÓHOZ

A szimuláció során alkalmazott potenciálformalizmustól függően a megoldás adódhat H , illetve B formájában is, így direkt karakterisztika alkalmazását feltételezve kétféle eset kezelése szükséges. A mágneses térerősség és a mágneses indukció is felírható egy lineáris, illetve egy nemlineáris, úgynevezett reziduális komponens összegeként [1]:

$$\begin{aligned} \vec{B} &= \mu_{opt} \vec{H} + \vec{R}, \\ \vec{H} &= \nu_{opt} \vec{B} + \vec{I}. \end{aligned} \quad (7)$$

A reziduál értéke mindkét esetben iteratív úton, konvergencia sorozatos közelítés segítségével kerül meghatározásra. Amennyiben a formalizmus a mágneses térerősségre kerül felírásra, úgy az n -edik lépésben a következőket kell végrehajtani:

- 0) while $\|\vec{R}^{(n-1)} - \vec{R}^{(n-2)}\| > \varepsilon$:
- 1) $\vec{H}^{(n)}$ meghatározása a Maxwell-egyenletek

alapján:

$$\vec{H}^{(n)} = \mathcal{M}\{\vec{R}^{(n-1)}\}; \quad (8)$$

2) Direkt modell alkalmazása:

$$\vec{B}^{(n)} = \mathcal{B}\{\vec{H}^{(n)}\}; \quad (9)$$

3)

$$\vec{R}^{(n)} = \vec{B}^{(n)} - \mu_{opt} \vec{H}^{(n)}. \quad (10)$$

Vektorpotenciálra ($\vec{A}$) épülő formalizmus esetén ($\vec{B} = \nabla \times \vec{A}$) [5] az iteráció a következőképp alakul:

0) while $\|\vec{I}^{(n-1)} - \vec{I}^{(n-2)}\| > \varepsilon$:

1) $\vec{B}^{(n)}$ meghatározása a Maxwell-egyenletek alapján:

$$\vec{A}^{(n)} = \mathcal{M}\{\vec{I}^{(n-1)}\} \rightarrow \vec{B}^{(n)} = \nabla \times \vec{A}^{(n)}; \quad (11)$$

2) $\vec{H}^{(n)}$ approximálása:

$$\vec{H}^{(n)} = \nu_{opt} \vec{B}^{(n)} + \vec{I}^{(n-1)}; \quad (12)$$

3) Direkt modell alkalmazása:

$$\vec{I}^{(n)} = \vec{H}^{(n)} - \nu_{opt} \mathcal{B}\{\vec{H}^{(n)}\}. \quad (13)$$

A permeabilitásfüggvény jelentősége, hogy az iteráció konvergenciáját μ_{opt} és ν_{opt} paraméterek helyes megválasztásán keresztül biztosítja [1]:

$$\mu_{opt} = \frac{\mu_{max} + \mu_{min}}{2}, \quad \nu_{opt} = \frac{2 \frac{1}{\mu_{min}} \frac{1}{\mu_{max}}}{\frac{1}{\mu_{min}} + \frac{1}{\mu_{max}}}. \quad (14)$$

V. VÉGESELEM-MÓDSZER

A végeselem-módszer numerikus technika parciális differenciálegyenletek közelítő megoldására. A módszer alkalmazása során a differenciálegyenlet $\vec{W}$ súlyozófüggvénnyel megszorozott, a vizsgált tartományra integrált gyenge alakja [5] kerül megoldásra az egydimenziós esetben szakaszokra, kétdimenziós esetben például háromszögekre diszkrétizált problémamartományon. A módszer az eredeti differenciálegyenletet általánosan a

$$\underline{\underline{K}} \underline{u} = \underline{b} \quad (15)$$

alakban felírható lineáris algebrai egyenletrendszerre képezhető le, ahol $\underline{\underline{K}}$ az együtthatómátrix, $\underline{u}$ az ismeretleneket, $\underline{b}$ pedig a gerjesztést és a peremfeltételeket tartalmazó oszlopvektor.

A *Galjorkin-módszer* értelmében a súlyozófüggvény és az elemeken értelmezett formafüggvények megegyeznek [4]. A cikkben bemutatott szimulációk Matlab-script formájában, csomóponti végeselemek használatával kerültek implementálásra.

VI. NEMLINEARITÁS EGYDIMENZIÓS SZIMULÁCIÓBAN

Klasszikus elektrodinamikai probléma az x -irányú vastagságához képest z -irányban végtelenül hosszúnak tekinthető, z -irányú váltakozó mágneses térbe helyezett vezető lemez. Az elrendezésből adódóan a mágneses térerősség $\vec{H} = H_z(x) \vec{e}_z$ alakban írható fel, ami lehetővé teszi az elrendezés egydimenziós szimulációját a vastagság szakaszokra történő diszkrétizálásával. A problémát a

$$\nabla \times \nabla \times \vec{H} + \mu_{opt} \sigma \frac{\partial \vec{H}}{\partial t} = -\sigma \frac{\partial \vec{R}}{\partial t} \quad (16)$$

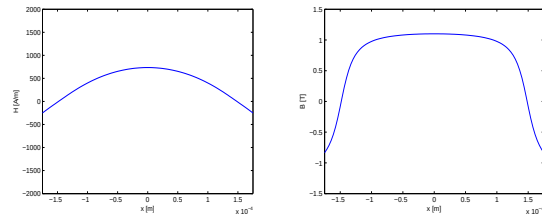
differenciálegyenlet írja le. Az egyenlet gyenge alakja a

$$\begin{aligned} \int_{\Omega} \nabla \times \vec{H} \cdot \nabla \times \vec{W} \, d\Omega + \mu_{opt} \sigma \int_{\Omega} \vec{W} \cdot \frac{\partial \vec{H}}{\partial t} \, d\Omega = \\ = -\sigma \int_{\Omega} \vec{W} \cdot \frac{\partial \vec{R}}{\partial t} \, d\Omega \end{aligned} \quad (17)$$

alakot ölti. A formulát egyetlen egydimenziós elemre felírva

$$\begin{aligned} \int_{\Omega} \left\{ \begin{bmatrix} \nabla W_1 \\ \nabla W_2 \end{bmatrix} \cdot \begin{bmatrix} \nabla W_1 & \nabla W_2 \end{bmatrix} \right\} d\Omega \begin{bmatrix} H_1 \\ H_2 \end{bmatrix} + \\ + \mu_{opt} \sigma \int_{\Omega} \left\{ \begin{bmatrix} W_1 \\ W_2 \end{bmatrix} \cdot \begin{bmatrix} W_1 & W_2 \end{bmatrix} \right\} d\Omega \begin{bmatrix} \frac{\partial H_1}{\partial t} \\ \frac{\partial H_2}{\partial t} \end{bmatrix} = \\ = -\sigma \int_{\Omega} \left\{ \begin{bmatrix} W_1 \\ W_2 \end{bmatrix} \cdot \begin{bmatrix} W_1 & W_2 \end{bmatrix} \right\} d\Omega \begin{bmatrix} \frac{\partial R_1}{\partial t} \\ \frac{\partial R_2}{\partial t} \end{bmatrix} \end{aligned} \quad (18)$$

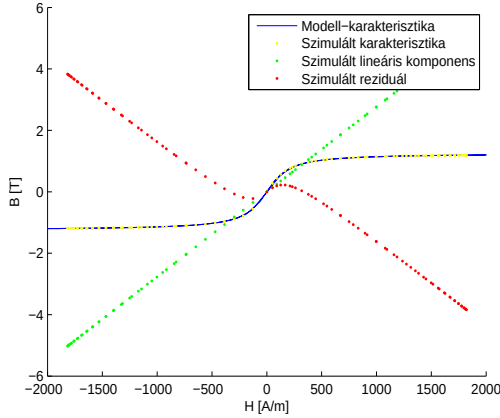
adódik. A gradiensket tartalmazó tag elsőfokú formafüggvények használata esetén konstans, kiértékelése visszavezethető egy téglalap területének meghatározására, a másik két tag viszont szakasz feletti numerikus integrálási technikát igényel. A mágneses térerősség és a mágneses indukció értékének alakulása egyetlen időpillanatban a 4. ábrán tanulmányozható. A 5. ábrán a szimulált tartomány egyetlen



4. ábra. $\vec{H}$ és $\vec{B}$ alakulása

csomópontjának modell alapján számított karakterisztikáját mutatja. Az ábrán megfigyelhető mind a μH lineáris, mind az R reziduális komponens alakulása, melyek szuperpozíciója a szimulált idő-

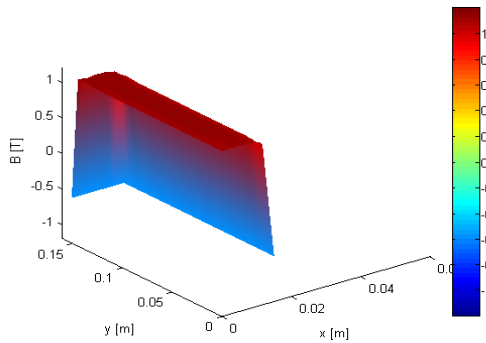
pillanatok mindegyikében kiadja a modell karakterisztikájához tartozó indukcióértéket.



5. ábra. Egyetlen csomópont karakterisztikája

VII. NEMLINEARITÁS KÉTDIMENZIÓS SZIMULÁCIÓBAN

Az előzőekben ismertetett probléma kétdimenziós szimulációval is megvalósítható, ezt mutatja be a 6. ábra. A vizsgált geometria egy téglalap alakú vezető lemez negyede, amely háromszög elemekre került felbontásra.



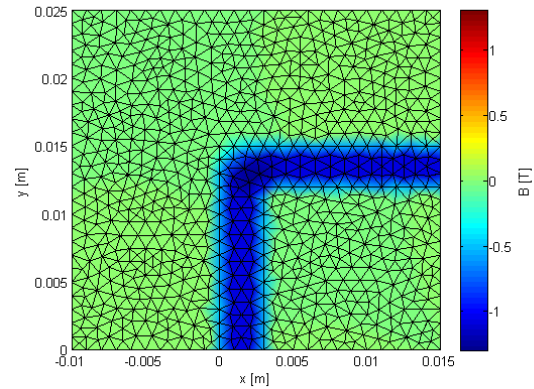
6. ábra. $\vec{B}$ alakulása egyetlen időpillanatban

A nemlineáris karakterisztika a IV. pontban bemutatott módon illeszthető az A -formalizmushoz is. Erre példa a négyzet alakban elrendezett vezető anyagból kialakított Epstein-keret [7], mely az anyagra felvezetett tekercselés segítségével árammal

kerül gerjesztésre. A probléma gyenge alakja a

$$\begin{aligned} \int_{\Omega} \nu_{opt} \nabla \times \vec{A} \cdot \nabla \times \vec{W} d\Omega + \sigma \int_{\Omega} \vec{W} \cdot \frac{\partial \vec{A}}{\partial t} d\Omega = \\ = - \int_{\Omega} \vec{I} \cdot \nabla \times \vec{W} d\Omega + \int_{\Gamma} \vec{W} \cdot \vec{K} d\Gamma \end{aligned} \quad (19)$$

formában írható fel, ahol $\vec{\Gamma}$ a tekercselést reprezentáló irányított felületet, $\vec{K}$ a felületi áramot jelöli.



7. ábra. $\vec{B}$ alakulása egyetlen időpillanatban

VIII. KONKLÚZIÓ

A cikkben bemutatásra került a nemlineáris anyagtulajdonságok modellezésének lehetősége véges elem-módszer és fixpontos iteráció használatával. A felvázolt eljárás a piacon kapható szoftverek megkerülésével kínál rugalmas, sokoldalúan használható szimulációs lehetőséget, többek között lehetőséget ad többértékű histerézis modell illesztésére is.

HIVATKOZÁSOK

- [1] Kuczmann M. *A histerézis karakterisztikával jellemzett rendszerek és a numerikus térszámítás kapcsolata*. Habilitációs disszertáció, Győr, 2010.
- [2] Della Torre E. *Magnetic Hysteresis*. IEEE Press, New York, 1999.
- [3] Pólik Z., Kuczmann M. *Measuring and control the hysteresis loop by using analog and digital integrators*. Journal of Optoelectronics and Advanced Materials, 2008.
- [4] Amália Iványi. *Hysteresis Models In Electromagnetic Computation*. Akadémiai Kiadó, Budapest, 1997.
- [5] Kuczmann M. Iványi A. *The Finite Element Method In Magnetism*. Akadémiai Kiadó, Budapest, 2008.
- [6] Bastos J.P.A., Sadowski N. *Electromagnetic Modeling By Finite Element Methods*. Marcel Dekker, New York, 2003.
- [7] Bertotti G., Mayergoyz I. D. *The Science of Hysteresis*. Academic Press, 2006.

Rádiófrekvenciás terek numerikus analízise

Friedl Gergely
Automatizálási Tanszék
Széchenyi István Egyetem
Győr, Magyarország
Email: friedl@maxwell.sze.hu

Prof. Dr. habil. Kuczmann Miklós
Automatizálási Tanszék
Széchenyi István Egyetem
Győr, Magyarország
Email: kuczmann@sze.hu

Kivonat—A cikkem a rádiófrekvenciás elektromágneses terek numerikus analízisének témakörében íródott. Bemutatom az idő- illetve frekvenciatartománybeli problémák végeselem-módszerrel történő szimulációjának alapját képező egyenletrendszereket lineáris és izotróp esetben, valamint a szimulációk menetét csomóponti és élmenti formafüggvények alkalmazásával. Az időtartománybeli szimuláció tárgyalása során említést teszek az inhomogén másodrendű differenciálegyenletek közelítő megoldására alkalmazható úgynevezett *Newmark- β* módszerről. Ezután egyszerűbb példákon keresztül idő- illetve frekvenciatartománybeli problémák szimulációját mutatom be. Konklúzióként összefoglalom a két módszer előnyeit és hátrányait, azok alkalmazhatóságát a különböző jellegű problémákra. Munkám során minden végeselemes-hálót GMSH [1] környezetben készítettem el, a végeselem-módszer egyéb lépéseit pedig Matlab [2] függvényekkel valósítottam meg.

Kulcsszavak—végeselem-módszer, elektromágneses terek, rádiófrekvencia, antenaszimuláció

I. ELMÉLETI ÁTTEKINTÉS

I-A. Végeselem-módszer

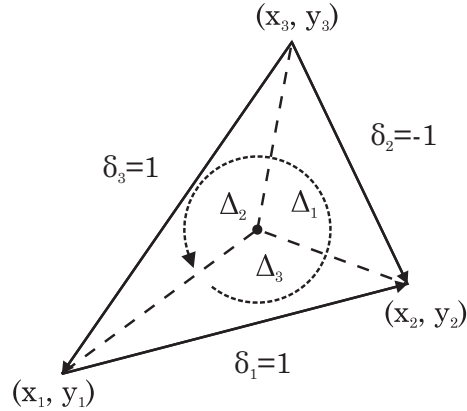
A súlyozott maradék elvének [3]–[6] családjába több numerikus technika is sorolható. Az elv működése abban rejlik, hogy a differenciálegyenletek és peremfeltételek egy tetszőleges súlyfüggvénnyel vett belső szorzatát teszi nullával egyenlővé. Amennyiben ez a súlyfüggvény a *Galjorkin*-eljárásnak megfelelően megegyezik az elemekre felírható forma-függvényekkel, végeselem-módszerről beszélhetünk. Léteznek úgynevezett csomóponti és élmenti forma-függvények is. A munkám során háromszögelemekkel foglalkoztam. A kétdimenziós elsőrendű csomóponti forma-függvények felírhatók a baricentrikus koordinátáknak megfelelően

$$N_i = \frac{\Delta_i}{\Delta} \quad (1)$$

alakban, melyek alapján a nulladrendű élmenti forma-függvények is bevezethetők

$$\vec{W}_i = l_i (N_i \nabla N_j - N_j \nabla N_i) \delta_i \quad (2)$$

formula segítségével. Itt Δ az elem területét jelöli, l_i az i -edik él hossza, δ_i pedig ± 1 attól függően, hogy a globális élírányok megegyeznek-e a háromszög körüljárási irányával (lásd: 1. ábra). Minden feladat kétdimenziós *Descartes*-koordináta-rendszerben lett megoldva.



1. ábra. A baricentrikus koordináták, és a globális élírányok háromszögelemen

I-B. Rádiófrekvenciás elektromágneses terek

A rádiófrekvenciás problémák általánosított képét mutatja a 2. ábra. A *Maxwell*-egyenletek [7] általános differenciális alakjából egyszerűen levezethetők az idő- és frekvenciatartománybeli hullámegyenletek. A hullámegyenlet időtartományban [8], [9]

$$\epsilon \frac{\partial^2 \vec{E}}{\partial t^2} + \sigma \frac{\partial \vec{E}}{\partial t} + \nabla \times \left(\frac{1}{\mu} \nabla \times \vec{E} \right) = -\frac{\partial \vec{J}}{\partial t}, \quad (3)$$

frekvenciatartományban pedig

$$\nabla \times \left(\frac{1}{\mu_r} \nabla \times \vec{E} \right) - k_0^2 \epsilon_r \vec{E} = -jk_0 Z_0 \vec{J} \quad (4)$$

formulával írható le. A gerjesztéstől végtelen távolságban az elektromos térnek ki kell elégítenie az

úgynevezett *Sommerfeld* sugárzási feltételt:

$$\lim_{r \rightarrow \infty} r \left[\nabla \times \vec{E} + jk_0 \vec{n} \times \vec{E} \right] = 0. \quad (5)$$

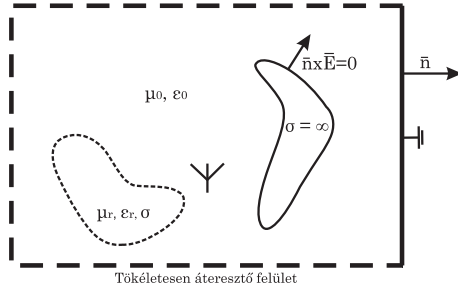
Végeselem-módszerrel végtelen távolság szimulációjához az (5) összefüggésnek egy véges távolságban vett approximációját kell alkalmazni. Ez időtartományban

$$\vec{n} \times \left(\frac{1}{\mu_0} \nabla \times \vec{E} \right) + Y_0 \vec{n} \times \left(\vec{n} \times \frac{\partial \vec{E}}{\partial t} \right) \approx 0, \quad (6)$$

frekvenciatartományban pedig

$$\vec{n} \times \nabla \times \vec{E} + jk_0 \vec{n} \times \vec{n} \times E \approx 0 \quad (7)$$

alakban állítható elő. A fenti két közelítés tulajdonképpen a véges távolságban található perem reflexiómentességét követeli meg.



2. ábra. A rádiófrekvenciás problémák általánosítása

I-C. Gyenge alak

A frekvenciatartománybeli hullámegyenletet megszorozva egy tetszőleges $\vec{W}$ tesztfüggvénnyel, és integrálva a teljes térfogatra előállítható a direkt alak [9]:

$$-\int_V \vec{W} \cdot \left(\nabla \times \frac{1}{\mu_r} \times \vec{E} \right) dV + \int_V \vec{W} \cdot k_0^2 \epsilon_r \vec{E} dV = \int_V \vec{W} \cdot j\omega \mu_0 \vec{J} dV. \quad (8)$$

Vektoralgebrai azonosságokat felhasználva, illetve behelyettesítve (7) összefüggést, a frekvenciatartománybeli analízis gyenge alakja:

$$\begin{aligned} \int_V (\nabla \times \vec{W}) \cdot \frac{1}{\mu_r} (\nabla \times \vec{E}) dV - \int_V \vec{W} \cdot k_0^2 \epsilon_r \vec{E} dV + \\ + jk_0 \oint_A (\vec{n} \times \vec{W}) \cdot (\nabla \times \vec{E}) dA = \\ = \int_V \vec{W} \cdot jk_0 Z_0 \vec{J} dV. \end{aligned} \quad (9)$$

Az időtartománybeli hullámegyenletet megszorozva egy tetszőleges $\vec{W}$ tesztfüggvénnyel, és integ-

rálva a teljes térfogatra előállítható a direkt alak [9]:

$$\begin{aligned} \int_V \vec{W} \cdot \nabla \times \left(\frac{1}{\mu} \nabla \times \vec{E} \right) dV + \int_V \vec{W} \cdot \epsilon \frac{\partial^2 \vec{E}}{\partial t^2} dV \\ + \int_V \vec{W} \cdot \sigma \frac{\partial \vec{E}}{\partial t} dV = - \int_V \vec{W} \cdot \frac{\partial \vec{J}}{\partial t} dV. \end{aligned} \quad (10)$$

Szintúgy vektoralgebrai azonosságokat felhasználva, illetve alkalmazva (6) összefüggést, az időtartománybeli analízis gyenge alakja felírható:

$$\begin{aligned} \int_V (\nabla \times \vec{W}) \cdot \frac{1}{\mu} (\nabla \times \vec{E}) dV + \int_V \vec{W} \cdot \epsilon \frac{\partial^2 \vec{E}}{\partial t^2} dV + \\ \int_V \vec{W} \cdot \sigma \frac{\partial \vec{E}}{\partial t} dV + Y_0 \oint_A (\vec{n} \times \vec{W}) \cdot \left(\vec{n} \times \frac{\partial \vec{E}}{\partial t} \right) dA = \\ - \int_V \vec{W} \cdot \frac{\partial \vec{J}}{\partial t} dV. \end{aligned} \quad (11)$$

alakban.

Időtartománybeli analízis gyenge alakja egy másodrendű inhomogén differenciálegyenletre vezet. Ennek megoldása az úgynevezett *Newmark-β* módszerrel [9], [10] lehetséges. A módszer szerint az első és a második deriváltakat a centrális differenciák módszerével kell közelíteni az alábbiak szerint:

$$\frac{\partial u(t)}{\partial t} \approx \frac{u(t)^{n+1} - u(t)^{n-1}}{2\Delta t}, \quad (12)$$

$$\frac{\partial^2 u(t)}{\partial t^2} \approx \frac{u(t)^{n+1} - 2u(t)^n + u(t)^{n-1}}{\Delta t^2}, \quad (13)$$

valamint veszi az időlépések eredményének súlyozott átlagát

$$u \approx \beta u^{n+1} + (1 - 2\beta)u^n + \beta u^{n-1}, \quad (14)$$

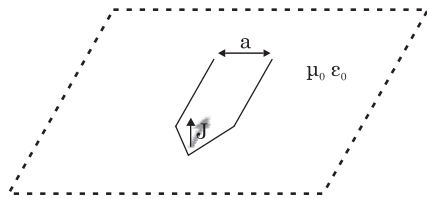
így a probléma visszavezethető egyszerű egyenletrendszer alakjára. Itt β egy 0 és 1 közötti paraméter. Amennyiben $\beta \geq 0,25$, a módszer feltétel nélkül stabil.

II. SZIMULÁCIÓK

II-A. Frekvenciatartománybeli analízis

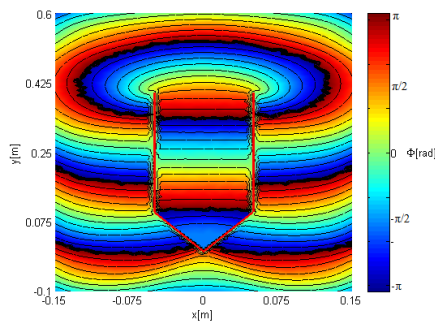
Frekvenciatartománybeli analízis segítségével szűrők illetve tápvonalak frekvenciatartománybeli viselkedését, valamint antennák iránykarakteristikáját érdemes vizsgálni.

II-A1. Nyitott végű csőtápvonal. Egy egyszerű négyszögletes csőtápvonal vizsgálata könnyen megvalósítható csomóponti formafüggvények alkalmazásával is, mivel az elektromos tér szempontjából elegendő a betáplálásra merőleges metszetben vizsgálni a rendszert (lásd: 3. ábra). A szimulált tápvonal szélessége 0,1 méter, így a tápvonal által áteresztett jel minimális frekvenciája 1,5 GHz. A tápvonal zárt



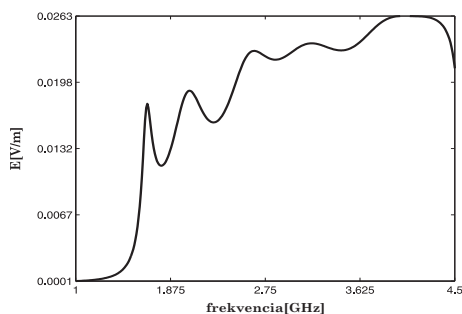
3. ábra. A nyitott végű négyszögletes csőtápvonal modellje

végének folyamatos kiszélesedése az állóhullámok hatásának csökkentésére szolgál. Az első egyfrekvenciás vizsgálat 2 GHz-es teszttel mellett történt. Ekkor egy valóságos tápvonalon belül mérhető hullámhossz 0,227 m. Frekvenciatartománybeli analízis eredménye az elektromos térerősség komplex csúcserőértéke. Az elektromos térerősség fázisának alakulását mutatja a 4. ábra radiánban kifejezve. A fázisábra



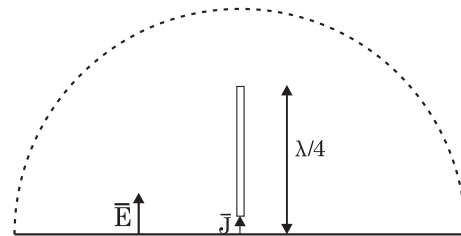
4. ábra. A szimuláció eredményeként kapott komplex érték fázisa

alapján mérhető a szimulált csőtápvonalban kialakult jel hullámhossza, amely 0,2268 méternek bizonyult, ami jól fedti a valóságot. Ugyanezen csőtápvonal végén kialakuló elektromos térerősség frekvenciatartománybeli alakulása látható a 5. ábrán.



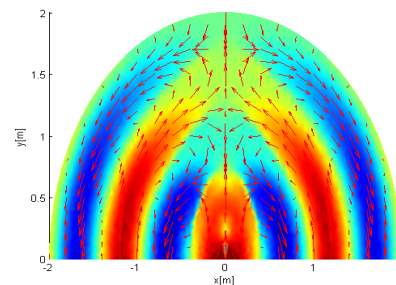
5. ábra. Az elektromos térerősség értéke a frekvenciatartományban a tápvonal végén

II-A2. Monopól antenna vizsgálata. Az antennák iránykarakteristikáját frekvenciatartománybeli analízis segítségével érdemes meghatározni. Egy negyedhullámú monopól antenna modelljében az antenna egy végtelen nagy, tökéletesen vezető felületre merőlegesen található (lásd: 6. ábra). A megoldás csak élmenti formafüggvények segítségével számítható ki, mivel a gerjesztés árama ebben az esetben a szimuláció síkjába esik. A szimulációt 300 MHz



6. ábra. Monopól antenna modellje

frekvenciára méreteztem, ennek következtében a monopól antenna magassága 0,25 méternek adódott. Fontos leszögezni, hogy a cikkben bemutatott módszer alkalmazása során az antennában áram nem folyik, hanem az anyagparamétereinek megfelelően eltorzítja a gerjesztésként alkalmazott áramelem által létrehozott elektromágneses teret. Az áram értéke utólag számítható. A peremen található élekre kiszá-

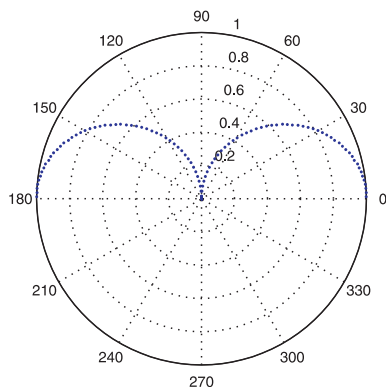


7. ábra. A monopól antenna által sugárzott tér

mított elektromos térerősség abszolút értéke alapján meghatározható a monopól antenna közelítési iránykarakteristikája, ez látható a 8. ábrán.

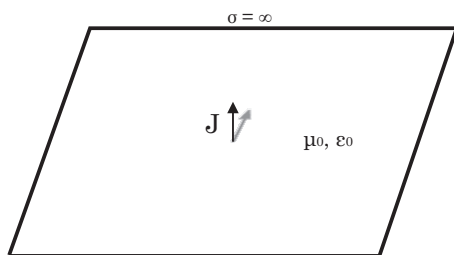
II-B. Időtartománybeli analízis

Az időtartománybeli viselkedés olyan esetekben lehet érdekes, amikor valamilyen tranzienst jelenségen alapszik egy eszköz működése, illetve akkor, ha az adott eszköz viselkedése nemlinearitást mutat. Az időtartománybeli analízis hátránya, hogy minden időlépésben külön el kell végezni a számításokat, mely hosszadalmas szimulációkhoz vezethet. A tranziensek hatása jól szemléltethető egy üregrezonátor



8. ábra. A monopól antenna közeltere

működésén keresztül. Egy egyszerű négyszögletes üregrezonátor kétdimenziós modelljét mutatja a 9. ábra. Amennyiben az üreg fala elektromosan tö-



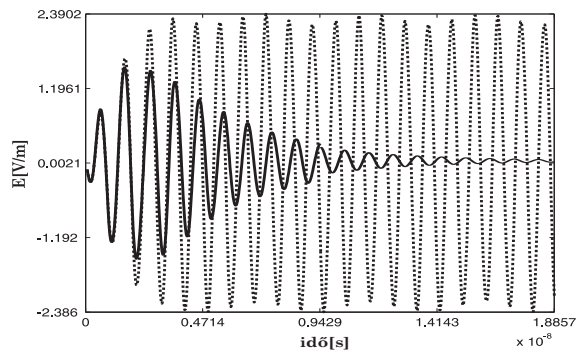
9. ábra. Üregrezonátor modellje

kéletesen vezető, az elektromágneses hullám teljes egészében reflektálódik róla, az elektromágneses teljesítmény nem tud eldisszipálódni, viszont amennyiben az üreg fala véges vezetőképességgel rendelkezik, az elektromágneses tér behatol az anyagba, ahol nagy veszteséget szenved.

Egy tökéletes és egy mikrohullámú eszközhöz képest kifejezetten veszteséges üregrezonátor kerül összehasonlításra (lásd: 10. ábra). Az üregek modellje mindkét esetben 0,1 méter oldalú négyzet. Veszteségmentes esetben az üreg falának vezetőképessége végtelen, míg veszteséges esetben ez az érték $\sigma = 20 \frac{S}{m}$. Az extrém alacsonynak beállított vezetőképességnek köszönhetően az elektromágneses teljesítmény jól láthatóan eldisszipálódik az üregből.

III. KONKLÚZIÓK

A rádiófrekvenciás elektromágneses terek végelelemes analízise kidolgozásra és megvalósításra került csomóponti és élmenti formafüggvények segítségével egyaránt. Megállapítható, hogy az eredmények tükrözik a valóságos viselkedést. A frekvenciatartománybeli analízis gyors és tranziensektől men-



10. ábra. Veszteséges és veszteségmentes rezonátorban az elektromos tér alakulása

tes eredménnyel szolgál, nemlinearitás kezelése viszont nem lehetséges alkalmazásával. Időtartománybeli analízis lassú, viszont tetszőleges bemeneti jel és anyagparaméter mellett is lehetőség van egy rendszer viselkedésének analízisére.

KÖSZÖNETNYILVÁNÍTÁS

A kutatás a TÁMOP 4.2.4.A/2-11-1-2012-0001 azonosító számú "Nemzeti Kiválóság Program - Hazai hallgatói, illetve kutatói személyi támogatást biztosító rendszer kidolgozása és működtetése konvergencia program" című kiemelt projekt keretében zajlott. A projekt az Európai Unió támogatásával, az Európai Szociális Alap társfinanszírozásával valósul meg.

HIVATKOZÁSOK

- [1] C. Geuzaine and J.-F. Remacle. Gmsh: a three-dimensional finite element mesh generator with built-in pre- and post-processing facilities. International Journal for Numerical Methods in Engineering 79(11), pp. 1309-1331, 2009.
- [2] Matlab, www.mathworks.com, (utolsó látogatás: 2013.10.31.).
- [3] Kuczmann M. Iványi A. „The Finite Element Method in Magnetics”. Akadémiai kiadó, Budapest, 2008.
- [4] Iványi A. „Folytonos és diszkrét szimulációk az elektrodinamikában”. Akadémiai Kiadó, Budapest, 2003.
- [5] J. P. A. Bastos, N. Sadowski. „Electromagnetic Modeling by Finite Element Methods”. Marcel Dekker, New York, 2003.
- [6] J. Jin. „The Finite Element Method in Electromagnetics”. John Wiley and Sons, New York, 2002.
- [7] J. C. Maxwell. „A Dynamical Theory of the Electromagnetic Field”. Philosophical Transactions of the Royal Society, London, pp. 459-512, 1865.
- [8] D. M. Pozar. „Microwave Engineering”. John Wiley and Sons, New York, 2012.
- [9] J. Jin. „Finite Element Analysis of Antennas and Arrays”. John Wiley and Sons, New York, 2009.
- [10] Newmark, N. M. „A method of computation for structural dynamics”. Journal of Engineering Mechanics, ASCE, 1959.

Virtuális forgalomtovábbítási táblák tömörítése

Mihálka Bence, Rétvári Gábor

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

Kivonat—Az utóbbi időben a hálózat-virtualizáció egyre nagyobb figyelmet kap, mivel többek közt a felhő alapú rendszerekben is fontos szerepet játszik. Ezen technológia kulcsa a virtuális útválasztók kezelésében rejlik, ami nem más, mint egyetlen fizikai eszközön több virtuális útválasztó egyidejű futtatása. Ugyanakkor a felhasznált nagysebességű memóriamodulok mérete limitálja a virtuális útválasztók maximális számát, mivel fenn kell tartaniuk egy-egy forgalomtovábbítási táblát (VFIB – Virtual Forwarding Information Base), így ezek tárolása skálázhatósági problémákhoz vezet. Jelen cikkben olyan algoritmusokat mutatunk be, melyek kiterjesztik a jól ismert prefix fa hajtogatás módszerét a VFIB-ekre, jelentősen csökkentve azok memóriagigéjét, és bemutatunk egy, a VFIB-ek átcímkezésén alapuló módszert is, amivel tovább csökkenthető a szükséges memória mennyisége. A tömörítések hatásfokát egy szimulátorral vizsgáltuk, amely segítségével valós és generált VFIB-eket is tömörítettünk. Az eredményeink alapján elmondhatjuk, hogy módszereink praktikus megoldásul szolgálhatnak a virtuális útválasztók memória-skálázhatósági problémájára.

Kulcsszavak: virtuális útválasztó; forgalomtovábbítási tábla; IP-cím keresés

I. BEVEZETÉS

A hálózat-virtualizáció, és ezen belül a virtuális útválasztók használata ígéretes technológia, mivel segítségükkel számos hálózati technika realizálható, többek közt felhasználó-specifikus, ill. szabály-alapú útválasztás. Emellett a VPN-ek (Virtual Private Network), illetve felhő alapú rendszerekben is fontos szerepet játszanak.

A megvalósítás lényege, hogy ugyanazon a fizikai eszközön több virtuális útválasztót futtatunk párhuzamosan. Az útválasztó a csomag fejlécében található információk - például MPLS címke (Multi Protocol Label Switching), Ethernet VLAN (Virtual Local Area Network) tag – alapján dönti el, hogy az melyik virtuális példányhoz tartozik. Ez alapján az útválasztó az eltárolt VFIB (Virtual Forwarding Information Base) segítségével megállapítja, hogy melyik szomszédnak kell továbbítania a csomagot. Ezzel egyetlen fizikai eszköz képes lesz résztvenni többféle hálózati topológia forgalomtovábbításában, amely a fent felsorolt hálózati technikák egyik alapkövetelménye.

Ennek megvalósításakor a szűk keresztmetszetet a fizikai eszköz memóriája adja. Amellett, hogy rendkívül gyorsnak kell lennie ahhoz, hogy a csomagtovábbítás késleltetése minimális legyen, az összes VFIB-et is fenn kell tartani benne. Ezen két tényező együttesen skálázhatósági problémákhoz vezet, limitálva az alkalmazható virtuális példányok számát.

Erre a problémára szolgáltat megoldást, ha a VFIB-eket betömörítjük, vagyis átalakítjuk az adatstruktúrájukat úgy, hogy továbbra is ugyanazt a továbbítási logikát valósítsák meg, mint a tömörítetlen esetben, de lényegesen kevesebb memória felhasználásával. Az alapötletet természetesen a virtualizációtól függetlenül is széleskörűen alkalmazzák az útválasztókban [1] [2][3]. A szakirodalomban számtalan megközelítéssel

találkozhatunk, jelen cikkünkben az egyik jól ismert módszert, a prefix fa hajtogatást terjesztjük ki a VFIB-ek tömörítésére. Azért választottuk ezt az algoritmust, mert információelméleti szempontból bizonyított, hogy entrópia-közi tömörítés érhető el használatával. [4]

Az útválasztó architektúrájának szempontjából fontos különbséget jelent, hogy az egyes virtuális példányok teljesen elkülönülnek egymástól, vagy képesek egymással adatokat megosztani. [5]-ben megmutatták, hogy amennyiben képesek közös memórián osztozni az egyes virtuális példányok, úgy hatékonyabb tömörítés érhető el. Ezt az ötletet is alkalmaztuk, amikor a prefix fa hajtogatást átültettük a VFIB-ek tömörítésére. Emellett bemutatunk egy olyan heurisztikus módszert is, amelynek használatával további jelentős memória-megtakarítás érhető el.

A cikk további részei a következők szerint épülnek fel: a II. fejezet a szükséges háttérter tartalmazza: első részében a prefix fát, második részében a prefix fa hajtogatás algoritmusát írjuk le, harmadik részében pedig bemutatjuk a virtuális útválasztók architektúráit. A III. fejezet tárgyalja a különböző VFIB tömörítési módszereket, míg a IV.-ben teljesítményanalízisnek vetjük alá őket, a mérési környezet és az adatok forrásának leírásával együtt. Végül az V. fejezetben a konklúziók levonásával, és a jövőbeli munka leírásával zárjuk cikkünket.

II. HÁTTÉR

A. FIB, prefix fa

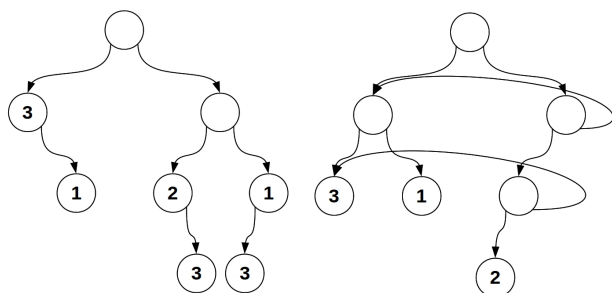
Akár virtuális, akár valós forgalomtovábbítási tábláról beszélünk (FIB, Forwarding Information Base), ez az adatszerkezet felelős azért, hogy egy beérkező IP csomag a megfelelő szomszéd felé legyen továbbítva. A bemenő adatok olyan bejegyzésekből állnak, amelyek egy IP címhez és prefix hosszhhoz egy nexthop értéket rendelnek, ami nem más, mint az útválasztó valamely szomszédjának IP címe. Példaként két darab FIB bejegyzéseit láthatjuk az 1. táblázatban, a későbbiekben a tömörítések során is ezen két példát fogjuk felhasználni. Ezekben a nexthop érték nem egy IP cím lesz, az egyszerűség kedvéért ehelyett egész számokat használtunk. Mint később, a tömörítési módszerek során látni fogjuk, ez nem csak a jobb átláthatóságot segíti, de további optimalizálási lehetőséget is rejt magában.

VFIB #1		VFIB #2	
IP cím/prefix hossz	Nexthop	IP cím/prefix hossz	Nexthop
0.0.0.0/1	3	0.0.0.0/3	1
64.0.0.0/2	1	0.0.0.0/2	2
128.0.0.0/2	2	0.0.0.0/1	3
160.0.0.0/3	3	64.0.0.0/3	2
192.0.0.0/3	3	128.0.0.0/1	1
192.0.0.0/2	1	128.0.0.0/2	2

1. táblázat: két darab FIB bejegyzései

Az adatstruktúrán értelmezett műveletek közül a legfontosabb a keresés, amely a beérkező csomag cél IP címe alapján visszaadja azt a nexthop értéket, amely a lehető legtöbb bitben egyezik a cél IP címmel (LPM, longest prefix match). Amennyiben több bejegyzés is illeszkedik az IP címre, úgy a legspecifikusabbat kell visszaadni. Emellett szükséges még a frissítés művelete is, amellyel bejegyzéseket adhatunk hozzá, módosíthatunk, vagy törölhetünk.

A FIB-ek tárolására leggyakrabban használt adatszerkezet a prefix fa. A fenti táblázat első VFIB-jének bejegyzéseit az 1. ábrán látható módon tároljuk a prefix fában. Az IP címet bináris számként értelmezve lépünk a fában balra vagy jobbra, a megadott prefix hossz mélységig, és ott eltároljuk a nexthop azonosítóját.



1. ábra: az első FIB tömörítetlenül 2. ábra: az első FIB tömörítés után

B. A prefix fa hajtogatás algoritmus

Ezen módszer lényege, hogy megkeressük az azonos felépítésű részfákat, azokat csak egyszer tároljuk el, és oda több csomópontból is mutatókat irányítunk. Az eredményül kapott adatszerkezet már nem prefix fa, hanem DAG (Directed Acyclic Graph, irányított körmentes gráf). A keresés és frissítés műveletek a tömörített adatszerkezeten is értelmezhetőek és megvalósíthatóak, ezek részletes leírása [6]-ban található. Az 1. ábrán látható FIB tömörítés utáni formáját a 2. ábra mutatja.

A példák szempontjából fontos megemlíteni, hogy a prefix fa hajtogatás módszer része egy másik algoritmus is, amely úgynevezett komplett prefix fává alakítja a bemeneti prefix fát. Ez annyit jelent, hogy minden belső (nem-levél) csomópontnak pontosan kettő gyermeke lehet, és nexthop címkét csak a levelek tartalmazhatnak. Minden prefix fához megalkotható egy ilyen komplett prefix fa: ha egy csomópontnak csak egyetlen gyermeke van, akkor létrehozunk mellé a másikat is, a nexthop címkéket pedig "lefelé toljuk".

A FIB-ek méretének csökkentésére számtalan különböző algoritmus létezik. Felmerül tehát a kérdés, hogy miért pont a prefix fa hajtogatást alkalmazzuk a VFIB-ek tömörítésére? A válasz részletes indoklása szintén [4]-ban található, melynek lényege, hogy a módszer bizonyíthatóan entrópia-közelítő tömörítés elérésére képes, információelméleti szempontból alátámasztva ezzel az algoritmus hatékonyságát.

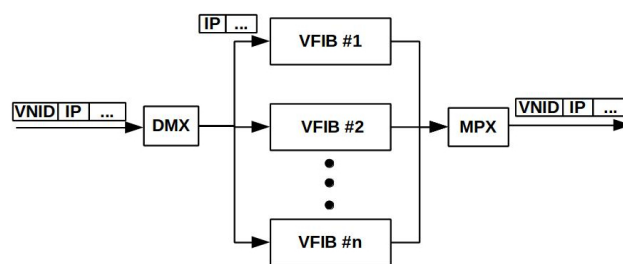
C. Virtuális útválasztó architektúrák

A virtuális útválasztók megvalósításánál a naív megoldás az lenne, hogy az egyes példányokat egymástól elkülönítve tároljuk, és ezek be- illetve kimeneteit egy demultiplexer/multiplexer páros kezeli, amik a virtuális hálózat azonosítója alapján adják át a csomagot a megfelelő virtuális példánynak. Ennek vázlatát mutatja a 3. ábra. Jelen cikkben a virtuális hálózat azonosítóját VNID-vel jelöltük (Virtual Network Identifier), de ezt a gyakorlatban gyakran az Ethernet keret VLAN mezeje, vagy az MPLS címke határozza meg. Ebben az elrendezésben a VFIB-ek pontosan ugyanolyan bejegyzésekből épülnek fel, mint a hagyományos FIB-ek, és ugyanazokat a műveleteket kell támogatniuk.

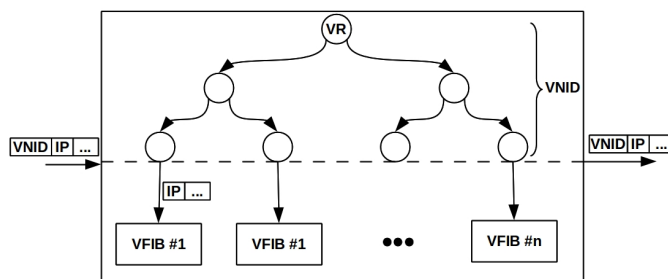
Mindezt az útválasztóban hardver szinten speciális, nagy sebességű memóriamodulokkal valósítják meg, amiben az összes virtuális hálózat bejegyzéseinek eltárolása skálázhatósági problémákat okoz. Az Internet növekedésével együttjáró FIB méret-növekedése már a virtualizációtól függetlenül is problémákat okozhat, pláne, ha egyszerre többet is fent szeretnénk tartani egyetlen eszközben. Ezen túlmenően, mivel minden VFIB egyszerre aktív, így annak is romlik az esélye, hogy különböző szintű gyorsítótárakban megtaláljunk egy bizonyos bejegyzést.

Ezen problémákon részben ugyan segít, ha a VFIB-eket külön-külön tömörítjük, de hatékonyabb eredményt érhetünk el, ha megengedjük, hogy az egyes virtuális útválasztók osztozzanak az adataikon [5]. Ehhez azonban az is szükséges, hogy módosítsuk az útválasztó architektúráját, melyet a 4. ábra szemléltet. Itt a virtuális hálózatok azonosítóját nem egy előfeldolgozó egység kezeli, ehelyett az összes VFIB-et bekötjük egy VR (Virtual Root) nevű közös szülő csomópontba. Így a VNID már az IP cím keresés (LPM) részévé válik. Ennek révén már az egész adatszerkezetre, mint egy nagyobb egységre tekinthetünk, amely lehetővé teszi az adatok cseréjét a virtuális útválasztók között. Mint látni fogjuk, ez jelentősen javítja a tömörítés határfokát.

Bármelyik útválasztó architektúrát is használjuk, fontos, hogy a VFIB-ek frissítése megfelelően megoldott legyen. Ehhez az első esetben a VNID-k kezelése az előfeldolgozó egység feladata, a másodikban pedig ezen bitek a teljes, közös VFIB része lesz, ezért így is kell kezelni. Megjegyezzük, hogy a hajtogatott prefix fán értelmezett frissítés algoritmus létezik, és a tömörítetlen esettel hasonló nagyságrendű időt vesz igénybe. A frissítés részleteit az olvasó [4]-ben megtekintheti.



3. ábra: elkülönített VFIB-ek, VNID kezelése: demultiplexer/multiplexer-rel



4. ábra: VFIB-ek egyben, VNID kezelése: a prefix fa részeként

III. A PREFIX FA HAJTOGATÁS ALKALMAZÁSA VFIB-EK TÖMÖRÍTÉSÉRE

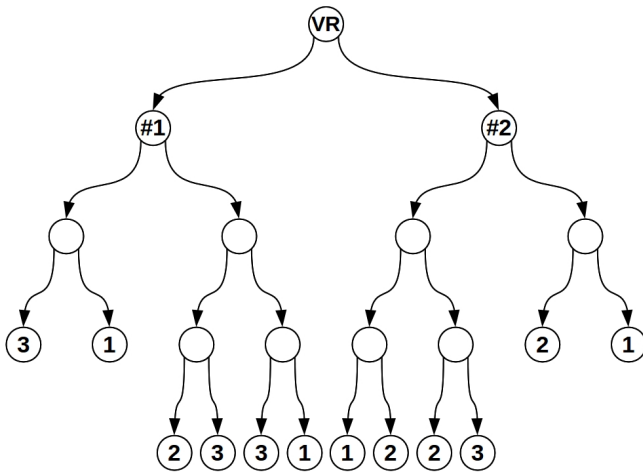
Ebben a fejezetben alkalmazzuk a megismert prefix fa hajtogatás módszerét a VFIB-ekre, tekintetbe véve a különböző virtuális útválasztó architektúrákat is.

A. A virtuális útválasztók nem osztanak meg adatokat

Amennyiben a fizikai eszköz felépítése nem teszi lehetővé, hogy a VFIB-ek adatokat osszanak meg egymás között, akkor az egyes VFIB-eken külön-külön tudjuk lefuttatni a prefix fa hajtogatást. A példák közül az 1. számú VFIB tömörített alakját a 2. ábrán már láttuk, a 2. VFIB pedig ezzel teljesen azonos módon képezhető. Mint azt a későbbiekben látni fogjuk, ez a módszer kevésbé hatékony, mint a másik két bemutatásra kerülő algoritmus, ugyanakkor kifejezett előnye, hogy az útválasztó architektúrájának szintjén nem szükséges változtatás a hagyományos kialakításhoz képest.

B. A virtuális útválasztók osztoznak az adatokon

Ha az útválasztó architektúrája lehetővé teszi, hogy az összes VFIB-et egy közös gyökérrel kössük össze, akkor az egész struktúrára, mint egy nagyobb egységre tekinthetünk, amelyen végrehajthatjuk a prefix fa hajtogatást. Ekkor a VFIB-ek ugyanolyan felépítésű részfáikat megosztják egymás között, ezáltal kevesebb csomóponttal leírhatóak a bejegyzéseik. Példaként tekintsük az 5. ábrán látható prefix fát: ez az 1. táblázat két VFIB-jének bejegyzéseiből áll, amelyeket a könnyebb áttekinthetőség érdekében átalakítottuk komplett prefix fává. Miután elvégezzük a prefix fa hajtogatást, a 6. ábrán látható eredményt kapjuk.



5. ábra: teljes VirtualRoot-os prefix fa 2 darab VFIB-ből

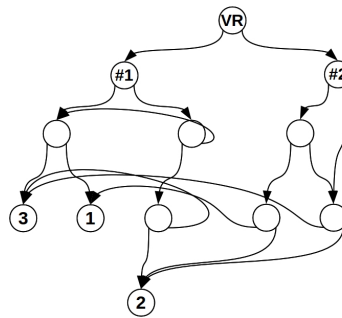
C. Gyakoriság alapján újraosztott nexthop címkek

A VFIB-ek bejegyzéseiben szereplő nexthop érték a gyakorlatban egy IP cím, ami az útválasztó valamely szomszédját azonosítja. Azonban a prefix fában nincs szükség az IP cím értékre, helyette elegendő egy azonosítót (egész számot) eltárolni címkéként, és egy kis méretű segéd-adatszerkezetben párosítani a konkrét IP címekhez.

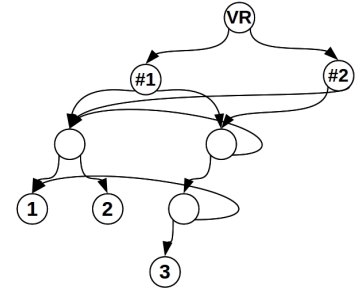
Az előző pontban bemutatott algoritmusnál ezen címkek kiosztása pusztán a beolvasott bejegyzések sorrendjétől függött, vagyis meglehetősen véletlenszerűen történt. A mi ötletünk ezzel szemben az volt, hogy osszuk újra ezen azonosítókat úgy, hogy az adott VFIB-ben leggyakrabban szereplő nexthop IP cím kapja az "1" azonosítót, a második a "2"-t, stb. Ettől a heurisztikus elrendezéstől azt várjuk, hogy mivel feltehetően többször fognak szerepelni a leggyakoribb címkek kombinációiból álló részfák, ezért azok megosztásával még kevesebb csomópontot igénylő szerkezetet kapunk.

Ha az 5. ábrán látható szerkezetben VFIB-enként újraosztjuk a nexthop címkeket gyakoriság szerint, akkor a 7. ábrán szereplő eredményt kapjuk.¹ A példában valóban sikerült

tovább csökkenteni a csomópontok számát, de felmerül a kérdés, hogy a címkek különböző eloszlásai alapján mennyit használ valójában az újraosztásuk. Az intuíció azt sugja, hogy egyenletes eloszlás esetén nem képes javításra, de – a valós FIB-ekben gyakran fellelhető – hatványtörvény eloszlás esetén már jelentős memóriaigény-csökkenés várható.



6. ábra: egyben tömörített VFIB-ek



7. ábra: újraosztott nexthop címkek

IV. TELJESÍTMÉNYANALÍZIS

Ebben a fejezetben bemutatjuk az elvégzett mérések hátterét és eredményeit, majd kiértékeljük ezeket.

A. Mérési környezet, bemeneti adatok

Az eddig bemutatott algoritmusok hatékonyságának mérése érdekében egy szimulációs szoftverben implementáltuk őket. A különböző tömörítések futtatása után statisztikát készítettünk az eredményül kapott adatszerkezetekről. A számunkra jelenleg legtöbb információt hordozó számérték a memóriaméret, melyet az alábbi képlet alapján számoltunk, ahol M a memóriaméret, P a szükséges mutatók száma, N a csomópontok száma, NH pedig a nexthop címkek száma:

$$M = P * \log(N) / \log(2) + NH * \log(NH) / \log(2) \quad (1)$$

A bemeneti adataink két helyről származnak. A felhasznált valós adatok a [4]-ben ismertetett nagy nemzetközi távközlési szolgáltatóktól kapott hozzáférési hálózat-beli FIB-ek. Emellett generált VFIB-eket is készítettünk, melynek alakját szintén egy valós FIB adta, de amelynek nexthop címkeit egyenletes, valamint hatványtörvény eloszlás szerint osztottuk ki annak érdekében, hogy mérésekkel tudjuk igazolni korábbi feltevésünket.

A méréseket egy Intel core i5-ös processzorral és 8 GB memóriával felszerelt számítógépen, Ubuntu 13.10 operációs rendszer alatt, user space-ben végeztük el.

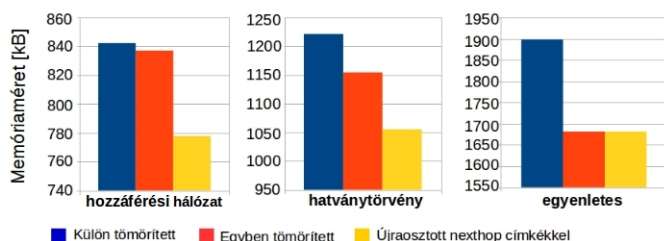
B. Eredmények

A mérések eredményeit a 2. táblázatban foglaltuk össze. Jól látható, hogy feltevéseink beigazolódtak mind a valós FIB-eket, mind a generált VFIB-eket tekintve. Például amikor az egyes VFIB-eken külön-külön végeztük el a prefix fa hajtogatást, a tömörítetlen esetnek mindössze 8.2%-ára tömörítette, ami 842 kB memóriát igényel, mégis elmarad az egyben tömörítettől 8 kB-tal, illetve 65 kB-tal az újraosztott nexthop címkek módszerét alkalmazó algoritmustól. A jobb átláthatóság érdekében a tömörített memóriaméretekről diagramot készítettünk, ez a 8. ábrán látható.

¹ első VFIB: 3 → 1, 1 → 2, 2 → 3; második VFIB: 2 → 1, 1 → 2, 3 → 3

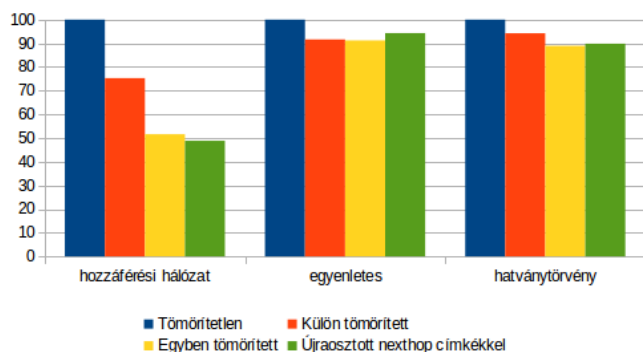
	Tömörítetlen		Külön tömörített		Egyben tömörített		Újraosztott nexthop	
	<i>M</i>	<i>n</i>	<i>M</i>	<i>n</i>	<i>M</i>	<i>n</i>	<i>M</i>	<i>n</i>
hozzáférési hálózat	10166	3809704	842	196358	836	195189	777	182147
hatványtörvény	4568	1801135	1221	276751	1154	262635	1055	241696
egyenletes	4568	1801135	1898	416515	1681	372157	1681	372143

2. táblázat: 6 db valós FIB, illetve 6 db különböző eloszlású generált FIB tömörítésének eredményei. (*M*: memóriaméret kB-ban; *n*: prefix fa pontjainak száma)



8. ábra: a tömörített VFIB-ek memóriaméretei [kB]

Bár jelen munkánk elsődleges célja a VFIB-ek memóriaigényének csökkentése volt, a gyakorlatban legfontosabb jellemző mégis az LPM sebessége. Ennek érdekében megvizsgáltuk, hogy az egyes tömörítési módszerek hatására hogyan változik a sebesség, ezt a 9. ábrán szemlélítettük.



9. ábra: egymillió IP kereséshez szükséges sebesség százalékos formában

A mérés elvégzéséhez mindegyik VFIB-en egymillió LPM keresést futtatunk. Mivel egy valódi útválasztóban sokkal nagyobb sebességű memóriát alkalmaznak, ezért az eredményeket százalékos formában közöljük (a szimulátorban az egymillió keresés körülbelül 1-1.5 másodpercet vett igénybe). Jól látható, hogy a tömörített VFIB-eken gyorsabb lett a keresés, különösen a valós FIB-ekből álló esetben. Ez annak köszönhető, hogy a tömörített VFIB-eknek nagyobb töredéke tud elférni a gyorsítótárak hierarchiájában.

V. KONKLÚZIÓK

Cikkünk során bemutatottuk, hogy a valós FIB-ek tömörítésének világából már jól ismert prefix fa hajtogatás módszere a virtuális útválasztók forgalomtovábbítási tábláinak tömörítésénél is megállja a helyét. Amennyiben a fizikai útválasztó architektúrája miatt nem megoldható, hogy a

VFIB-ek adatokat osszanak meg egymás között, úgy mindegyikre külön-külön futtatva az algoritmust jelentős tömörítést sikerült elérni. Ha azonban lehetőség adódik az adatok közös kezelésére, úgy mindegyik VFIB-et egyben tömörítve jobb hatásfokot kapunk. Különösen akkor igaz ez, ha az általunk javasolt nexthop címke újraosztást is végrehajtjuk: ebben az esetben a valós FIB-ek tömörítése során több, mint 65 kB memóriát sikerült pluszban megspórolni az első esethez képest, ez további 8%-nyi javulást jelent.

Ahhoz, hogy a bemutatott módszereket a gyakorlatban is alkalmazni lehessen, mindenképpen ki kell térni a tömörített adatszerkezeten értelmezett frissítés műveletére is, hiszen egy valós útválasztóban is folyamatosan frissíteni kell a bejegyzéseket. Ugyan ezt jelen munkánkban hely hiánya miatt nem tettük meg, de a szintén [4]-ben ismertetett algoritmus alapján a frissítés művelete könnyen elvégezhető, nagyságrendileg ugyanolyan gyorsan, mint a tömörítetlen esetben. Ez az algoritmus hagyományos FIB-ek frissítésére szolgál, tehát ahhoz, hogy a VFIB-ekre is alkalmazható legyen, kisebb módosítások szükségesek, különös tekintettel a VNID kezelésére.

Mindezek alapján kijelenthetjük, hogy a bemutatott algoritmusok alkalmasnak tűnnek arra, hogy segítségükkel virtuális útválasztókat valósítsunk meg egyetlen fizikai eszközön belül.

További munkánk során szeretnénk egyéb optimalizálási stratégiákat is hozzáadni a meglévő algoritmusokhoz (például szint-tömörített prefix fa [4]), illetve méréseinket elvégezni további valós VFIB-eken is, valamint további méréseket végezni az imént említett műveletekkel kapcsolatban is.

IRODALOMJEGYZÉK

- [1] Song, H., Kodialam, M., Hao, F., & Lakshman, T. V. "Building scalable virtual routers with trie braiding." *INFOCOM, 2010 Proceedings IEEE*. IEEE, 2010.
- [2] Ferragina, P., Luccio, F., Manzini, G., & Muthukrishnan, S. "Compressing and indexing labeled trees, with applications." *Journal of the ACM (JACM)* 57.1 (2009): 4.
- [3] Draves, Richard P., et al. "Constructing optimal IP routing tables." *INFOCOM'99. Eighteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*. Vol. 1. IEEE, 1999.
- [4] Rétvári, G., Tapolcai, J., Körösi, A., Majdán, A., & Heszberger, Z. "Compressing IP forwarding tables: towards entropy bounds and beyond." *Proceedings of the ACM SIGCOMM 2013 conference on SIGCOMM*. ACM, 2013.
- [5] Fu, Jing, and Jennifer Rexford. "Efficient IP-address lookup with a shared forwarding table for multiple virtual routers." *Proceedings of the 2008 ACM CoNEXT Conference*. ACM, 2008.
- [6] Rétvári, G., Csernátó, Z., Körösi, A., Tapolcai, J., Császár, A., Enyedi, G., & Pongrácz, G. "Compressing IP forwarding tables for fun and profit." *Proceedings of the 11th ACM Workshop on Hot Topics in Networks*. ACM, 2012.

Benchmarking Compressed and Succinct Data Structures

András Majdán

BME Dept. of Telecommunication and Media Informatics
Budapest University of Technology and Economics
Budapest, Hungary
majdan@tmit.bme.hu

This research is supported by the Google Faculty Research Award programme
"Searchable Compressed Data Structures for the Mainstream".

Abstract — In theory, queries on compressed bitmaps may easily outperform those on plain, uncompressed ones on today's heavily cached CPU architectures. The reason is that queries on memory-efficient compressed bitmaps result less time-consuming cache misses than on large, uncompressed forms. An additional benefit is that compressed bitmap data structures provide a richer query API (namely, rank and select operations) on top of the usual random access. In this study, we introduce a sophisticated performance benchmarking tool that we use to validate the above theoretical assumption. Our benchmarks indicate that current compressed bitmap implementations' computing overhead for random access on dense bitmaps is greatly outweighs the effects of improved cache affinity. On the other hand, on sparse bitmaps this computing overhead is getting lower and provides comparable performance to uncompressed cases. Besides this rank and select queries can be orders of magnitudes faster on compressed bitmaps than on plain ones.

Keywords – *benchmark; compressed; succinct; bitmap; rank; select; cache*

I. INTRODUCTION

Bitmaps are amongst the most important data structures used essentially everywhere in the theory and practice of computer science. Suppose, for instance, that the task is to design a data structure for storing telephone numbers, which would allow to effectively identify unallocated number that can assigned to a new subscriber.

In Hungary, a subscriber of a mobile network operator is identified by a 7 digit number. Storing such a number requires at least 24 bits per number, so storing all telephone numbers in a file would require a whopping 28.61 Megabytes of memory.

A significantly more space efficient storage scheme would be to use a bitmap: associate a bit to each number between 0 and 9999999 and if a telephone number is already assigned than set the corresponding bit to 1, and otherwise set the bit to zero. Using an uncompressed bitmap 1.92 Megabytes memory is enough, which gives a significant space reduction compared to native scheme. In the example of Fig.1., the zero-th bit is unset so number 0000000 is unassigned, while the second bit

is set to 1 so number 0000002 is already allocated. Correspondingly, the status of a particular telephone number is easy to determine by a simple *random access* to the corresponding position in the bitmap.

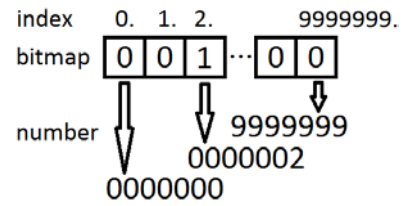


Figure 1. Each bit represents a telephone number.

Now, suppose that we want to make further queries to our data structure. For instance, we would like to know how many telephone numbers are still free, or how many numbers are assigned from, say, 9000000 to 9500000. The so called *rank* (or popcount) query is a simple operation that allows to answer such question. Namely, *rank* tells how many ones or zeros are present in the bitmap till a given index (see Fig.2.). Hence, determining the number of free telephone numbers boils down to the query $10000000 - \text{rank}(9999999)$, while $\text{rank}(9500000) - \text{rank}(9000000)$ gives the number of allocations in the corresponding range.

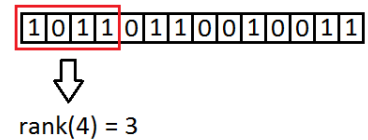


Figure 2. Three bits set till index four.

Select is essentially a query dual to *rank*. In particular, *select(n)* tells at which index occurs the *n*-th bit set. In our example, *select* could be used to answer question like "Which is the third assigned telephone number?". The corresponding query would be *select(3)* (see Fig.3.).

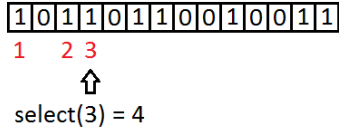


Figure 3. The third bit is set at index four.

Unfortunately, plain bitmaps, while support *random access* queries extremely fast, are extremely inefficient in *rank* and *select*. For answering a *rank* query, one would need to iterate over the entire bitmap until queried position and increase a counter by 1 whenever a set bit is encountered. In other words, in a bitmap string n bits *rank* and *select* terminate in $O(n)$ in the worst case, which is hardly affordable in any real-time application.

Compressed bitmaps are a relatively new development in computer science, which allow to store bitmaps in significantly less space than the naive scheme while also supporting all the usual queries, namely, *access*, *rank*, and *select*, in constant $O(1)$ steps [1][2], that is, independent of the length. The small size and rich query API make compressed bitmaps extremely appealing for the ever more space-hungry applications of our days. Unfortunately, a performance evaluation that would reveal the hidden costs and performance bottlenecks of compressed bitmap implementations is still missing in literature.

In this paper, we provide such a benchmark, using the most popular compressed bitmap data structure, the RRR [3] data structure. For this, we have developed a unique performance benchmarking tool that allows to evaluate running time of programs *at the granularity of CPU cycles*.

The rest of the paper is organized as follows. In Section 2, I introduce RRR and describe how different queries are implemented, in Section 3 we present our benchmarking tool, in Section 4 we present our performance evaluations, and finally, Section 5 concludes our work

II. RRR

Our focus in this paper is on RRR, a static succinct data structure which not only compresses to entropy but also provides fast *random access*, *rank* and *select* queries.

The compression of the bitmap is done by classifying blocks of bits using its weight (also called popcount) and coding just the class identifier and the offset (see Fig.4.).

On Figure 4. the block length is three. In real word scenarios of course wider blocks are used.

Block	Class	Offset
000	0	0 (none)
001	1	0
010	1	1
100	1	2
011	2	0
101	2	1
110	2	2
111	3	0 (none)

Figure 4. Coding of blocks by classifying.

Class identifiers require

$$classwidth = \lceil \log_2(blockwidth) \rceil$$

bits while the offset identifier's width depends on the class. The number of offsets of a given class (to calculate its width) can be read from the Pascal triangle or also can be calculated.

$$offsetwidth = \lceil \log_2 \left(\frac{blockwidth}{weight} \right) \rceil$$

The compression is greater when the block contains few or a lots of ones in it and worse when contains equal number of zeroes and ones (because we need more bits to represent the offset). It is the property which makes this data structure to compress to entropy.

To support fast *rank* and *select* queries RRR uses an array of superblocks (see Fig.5.) These superblocks divide the blocks into groups.

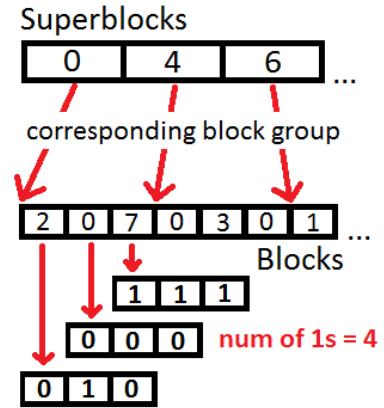


Figure 5. Superblock array of RRR

A superblock stores the *rank* value (number of ones) till its corresponding blocks' start. To answer a *rank* query we must calculate the superblock's index and then iterate through its corresponding blocks until we reach the query's target block. By adding the value in the superblock and all popcount values we have met through iterating we get the *rank* value until our target block. The final step is counting the ones till the desired bit in this block and add it the previously calculated value. That is the *rank* value.

Select can be done in constant $O(1)$ steps by using a three-level directory tree, the Clark's structure [4] which is far more complicated to include in this paper. A simple $O(\log n)$ version is to binary search in the superblocks until finding the greatest *rank* value which is lower than we are looking for. Then we must iterate through its corresponding blocks until we first count more ones than we are looking for. The answer is in the previous block, so we count in it till the desired bit and give back its global bit index as answer.

It is obvious that larger groups of blocks (fewer superblocks) results in slower *rank* and *select* queries because of long paths to iterate through. However the more superblocks present, the more memory need.

III. BENCHMARKING FRAMEWORK

We have made a benchmarking framework (cyclecount) to measure and compare the performance of compressed bitmap libraries. This framework can provide results at cycle accuracy and also make diagrams from them.

For the benchmarking part we have used and modified the sophisticated measurement library of Agner Fog. He has made his library to document instruction latencies and throughputs of different CPU-s [5]. It consists of a kernel driver and a user space part and uses the processor's Performance Monitor Counter (PMC) and the CPUID instruction to reliably measure the elapsed cycles. The PMC is a hardware counter which is presents in modern processors to be able to optimize the program for the CPU by measuring internal events (like failed branch predictions, cache misses, etc.). The CPUID is an instruction for determining the processor's type and features but in this case the library is taking advantage of a special byproduct of this instruction: it flushes the pipeline. In this way frequency scaling cannot alter the results.

However interrupts are not disabled so spikes can appear even though setting a higher priority to the process. To work around this problem our framework runs the test cases multiple times and calculates mean values.

Our benchmarking framework has multiple options for benchmarking a code:

- Compiling the code within the framework.
It can make use of inline expansions for further optimization. It is limited to open source codes by its nature.
- Use a code from a compiled object file
It is useful when no source code is available or it is not intended to share it with the testers.
- Use codes from libraries
Multiple libraries can be linked to the application.

It can also be given tasks to do before compilation and for cleaning. For example it can automatically make the library before using it and can clean temporary files after all.

We have introduced a special macro (`#interval`) which can be used to measure a code with different parameters. In this case the code compiled and ran with different define macros and results are grouped together on the output graph. It is useful for example comparing a query's performance on different bit lengths or block sizes.

For the output part multiple formats are created:

- Simple text file for human reading
- XML file for machine processing
- Input file for GNU Plot
- Diagram (in EPS format for documents and PNG format for web)

The diagrams are created by using GNU Plot so only the input files will be created without that. However someone can use these files to create the diagrams on a different computer where it is installed.

Note that we have used our framework for measuring performance of compressed bitmaps but it is a universal benchmarking framework and not restricted to this task. It is written in C++ and Python and should work on all operation systems where these compilers (or interpreters) present.

Large bitmap files should not be sent in emails or uploaded with the paper to be able to reproduce the experiment because it's wasting space and network bandwidth.

We have written a pseudo random bitmap generator (randbitmap) as part of this framework which solves this problem. With this program only the seed, size and density values must be given and it will generate the same bitmap on all platforms. It is reached by using an internal pseudo random number generator instead of the platform's one. It uses mmap (file memory mapping) from POSIX standard so it should run on all POSIX compliant operation systems (like Linux). On Windows however a POSIX emulation library must be installed (e.g. Cygwin).

To find out whether queries on compressed data structures can outperform uncompressed ones we have chosen two common compressed bitmap libraries: libcds and sdsl. The query routines on uncompressed bitmaps were extracted from the Linux kernel (bitmap.c).

The Linux kernel developers are not following the terminology *rank* and *select*, but they name these functions *bitmap_pos_to_ord* and *bitmap_ord_to_pos*. Also they use more optimization technics for ffs (find first set bit) and that is why multiple test cases exit.

The *rg* and *bvi* structures are just indexes for accelerating *rank* and *select* queries on top of uncompressed bitmaps. Therefore random access in these structures is same as accessing an uncompressed bitmap. The *d*, *sd*, *gap* and *rrr* structures are compressed ones with support for *rank* and *select* queries and fast *random access*.

IV. PERFORMANCE EVOLUTION

In our performance evolution we use random filled bitmaps with predetermined densities. The database is like the allocated numbers example in Section 1 but with a different range.

We have found that only the RRR succinct data structure implemented in sds library can provide comparable *random access* query to the uncompressed one (see Fig.6.)

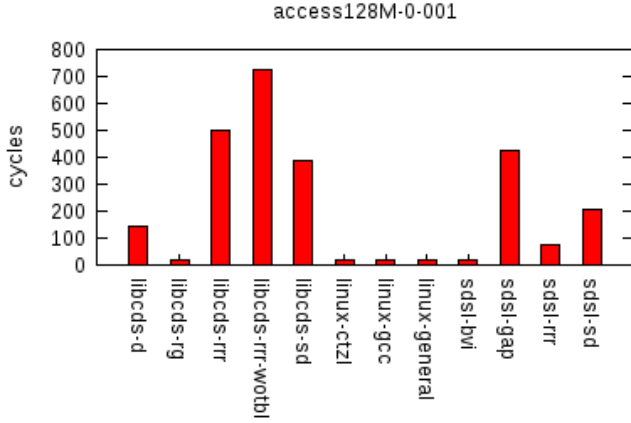


Figure 6. Random access in a 128 Megabit bitmap with density of 0.1%

By further analyzing the RRR structure *random access* performance by using different densities we have found that it is only effective at very sparse bitmaps. On dense bitmaps *random access* takes significantly more time (see Fig.7.)

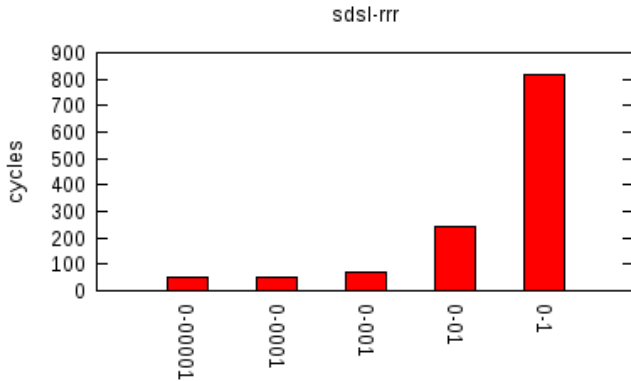


Figure 7. *Random access* of bitmaps with different densities.
0-00001: 0.001% dens., 0-0001: 0.01% dens.
0-001: 0.1% dens., 0-01: 1% dens., 0-1: 10% dens.

At 0.01% bitmap density RRR’s *random access* can reach as low as 53 cycles per query which is close to the 18-20 cycles of access to uncompressed bitmaps (Linux bitmap.c).

As we can see on Figure 8., *rank* is significantly faster on compressed structures and uncompressed structures with indexes than on plain bitmaps. The difference is even getting higher as size of the bitmap grows.

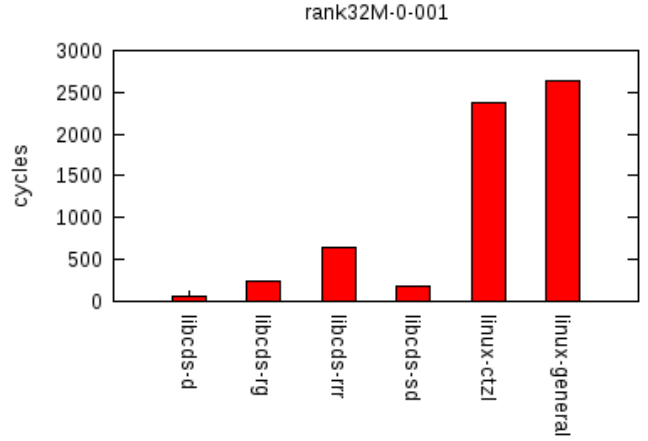


Figure 8. *Rank* query on a 32 Megabit bitmap with density of 0.1%

The performance graph of *select* was not fit in this paper however the scale would be logarithmic of such a diagram. The fastest *select* on an uncompressed bitmap was linux-ctzl with 69188 cycles while RRR was finished with 2556 cycles in the same scenario as we have tested *rank*.

These tests were run on an Intel Q6600 processor (an old Kentsfield core). Compressed data structures should perform better on the newer Haswell microarchitecture because it has a new bit manipulation instruction set (ABM, BMI1, BMI2). The population count, count leading or trailing zeros and bit field extraction are now parts of the processor as instructions.

V. CONCLUSION

The current compressed bitmap implementations’ computing overhead for *random access* on dense bitmaps is greatly outweighs the effects of improved cache affinity. On the other hand, on sparse bitmaps this computing overhead is getting lower and provides comparable performance to uncompressed cases. Besides this *rank* and *select* queries can be orders of magnitudes faster on compressed bitmaps than on plain ones.

REFERENCES

- [1] S. Vigna, “Broadword Implementation of Rank/Select Queries”, Proceedings of the 7th international conference on Experimental algorithms, Berlin, 2008, pp.154–168.
- [2] R. González, S. Grabowski, V. Mäkinen, G. Navarro, “Practical implementation of rank and select queries” in Poster Proceedings Volume of 4th Workshop on Efficient and Experimental Algorithms, Greece, 2005.
- [3] RRR – A succinct data structure proposed by R. Raman, V. Raman, and S. R. Satti in their “Succinct indexable dictionaries with applications to encoding k-ary trees, pre x sums and multisets.” paper, ACM Transactions on Algorithms, 2007
- [4] D. Clark, “Compact Pat Trees”, PhD thesis, University of Waterloo, 1996
- [5] A. Fog, “Lists of instruction latencies, throughputs and micro-operation breakdowns for Intel, AMD and VIA CPUs”, http://www.agner.org/optimize/instruction_tables.pdf, 1996–2014

Digitális szökökút alapú transzport protokoll fejlesztése és elemzése

Solymos Szilárd

Távközlési és Médiainformatikai Tanszék
Budapesti Műszaki és Gazdaságtudományi Egyetem
Budapest, Magyarország
solymos@tmit.bme.hu

Dr. Molnár Sándor

Távközlési és Médiainformatikai Tanszék
Budapesti Műszaki és Gazdaságtudományi Egyetem
Budapest, Magyarország
molnar@tmit.bme.hu

Kivonat—Az Internet folyamatos fejlődése során megjelent számos különböző TCP (Transmission Control Protocol) verzió nem képes optimális megoldást nyújtani a jelenlegi folyamatosan változó, heterogén hálózati környezetek esetén. Erre a kérdésre egy olyan új koncepció adhat választ, amely a torlódásszabályozás helyett hatékony hibajavító kódolás alkalmazását javasolja. Az utóbbi években egy ilyen elven működő, DFCP-nek (Digital Fountain based Communication Protocol) elnevezett, szállítási rétegbeli protokollt fejlesztünk a tanszéken, amely Raptor kódolás alkalmazásával képes hatékony működést elérni. Ebben a cikkben egy olyan elemzést mutatunk be, amely kitér a protokoll működésére, fejlesztésére, és annak teljesítményét elterjedten használt TCP verziókkal hasonlítja össze szimulációk és teszthálózati mérések segítségével.

Kulcsszavak: *jövő Internetje, hibajavító kódolás, szökökút kódolás, transzport protokoll*

I. BEVEZETÉS

Jelenleg a TCP (Transmission Control Protocol) az egyik legszéleskörűbben használt szállítási rétegbeli protokoll, amely kapcsolat-orientált, és megbízható adatátvitelt valósít meg. A TCP torlódásszabályozást alkalmaz annak érdekében, hogy minél hatékonyabb működést érjen el a jelenlegi nagysebességű, nagy késleltetésű, vezeték nélküli és egyéb média esetén. A hatékonyság növelése céljából számos, gyakran igen eltérő teljesítményű, együttműködési problémákat okozó, egyre nagyobb komplexitású TCP változat jelent meg [1], [2], amelyek egy adott környezet esetén ugyan képesek jobb működést elérni, de nem képesek univerzális, optimális megoldást nyújtani napjaink hálózatainak kihívásaira.

A TCP esetén fellépő hatékonysági és együttműködési kérdések megoldása céljából jelenleg számos kutatás folyik és új koncepciók jelentek meg. Az egyik ilyen, a GENI (Global Environment for Network Innovations) által javasolt, elképzelés szerint egyáltalán ne alkalmazzunk torlódásszabályozást [3]. Az ötlet ígéretesnek tűnik, de eddig néhány, ezzel az elvvel kapcsolatos munkán kívül semmilyen realizációt, vagy további finomítást nem publikáltak, így a koncepció részletes elemzése és hozzá kapcsolódó releváns következtetések nem álltak rendelkezésre.

Raghavan és Snoeren tanulmányozta a torlódásszabályozás nélküli hálózat nyújtotta előnyöket, és bemutatott egy *torlódásmentesítőt*, mint egy lehetséges megoldás alapját [4].

Bonald és mások szintén megvizsgálták a torlódásszabályozás nélküli hálózat viselkedését [5]. Az ő eredményük megmutatta, hogy nem igaz az a hiedelem, amely szerint a torlódásszabályozás nélkülözése a hálózat összeomlásához vezetne. López és mások a játékelméleten keresztül vizsgálták meg egy szökökút kódoláson alapuló protokoll teljesítményét [6]. Megmutatták, hogy elérhető egy olyan Nash-egyensúly, amelynél a hálózat teljesítménye hasonló ahhoz, mintha minden végpont TCP-t alkalmazna. Úgy tűnik, hogy a ráta nélküli kódok jól alkalmazhatóak hibajavító kódolásként. Például [7] bemutatja az élő videó közvetítés egy ráta nélküli kódoláson alapuló forgatókönyvét és tartalmazza a hozzá kapcsolódó kísérleti eredményeket is. Botos és mások egy olyan transzport protokollt mutattak be, amely nagy csomagvesztésű hálózaton alkalmaz ráta nélküli kódolást [8]. Kumar és mások [9] pedig egy vezeték nélküli hálózatokon ígéretes teljesítményt nyújtó transzport protokollt fejlesztettek ki, amely szökökút kódoláson alapul.

A továbbiakban azt tételezzük fel, hogy egyáltalán nem alkalmazzunk torlódásszabályozást. Ebben az esetben a hálózatban minden entitás maximális sebességgel küldheti az adatokat, az átvitel során fellépő, főleg borsztös csomagvesztések hatékony hibajavító kódolás segítségével állíthatók helyre. Az igazságos működéshez a hálózatban igazságos ütemezést tételezünk fel, mint például a WFQ (Weighted Fair Queuing). Mivel ennek az új koncepciónak nem volt elérhető további, részletesebb leírása, implementációja, egy teljesen új protokollt terveztünk meg és valósítottunk meg, amely a DFCP (Digital Fountain based Communication Protocol) nevet kapta.

A cikk a következőképpen épül fel. A következő, II. részben a DFCP-t megalapozó koncepciót, a III. fejezetben a protokoll fejlesztését és megvalósításának alapjait ismertetjük. A IV. fejezet az elvégzett szimulációs és teszthálózati mérések eredményeit mutatja be és összehasonlítja a protokoll teljesítményét néhány elterjedten alkalmazott TCP variánssal. Végül az V. fejezet összefoglalja a leírtakat.

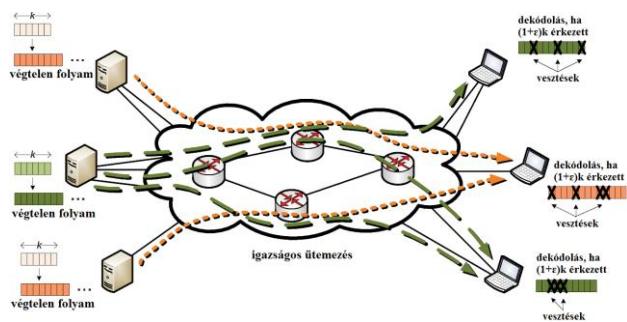
II. SZÖKÖKÚT KÓDOLÁSON ALAPULÓ ARCHITEKTÚRA

Napjainkban az Internet fő szabályozó mechanizmusa a torlódásszabályozás, amely minél hatékonyabban próbálja kihasználni a hálózati erőforrásokat. A különböző hálózati környezetek jelentette problémák megoldására számos

különböző TCP verzió jelent meg, amelyek az adott viszonyok között jobb teljesítményt tudnak ugyan nyújtani, de az is látható, hogy ezek a protokollok nem mindig képesek az igazságos együttműködésre és a különböző TCP változatok nagyon eltérő teljesítményt érnek el.

A jövő hálózataira vonatkozóan a GENI adott egy olyan javaslatot, amely szerint a torlódásszabályozás helyett alkalmazzunk hibajavító kódolást. Ekkor a hálózat minden végpontja maximális sebességgel küldhet adatokat, tehát amíg csak van rendelkezésre álló adat egy adott végpont esetén, addig olyan gyorsan történhet a küldés, amennyire csak lehetséges. Természetesen ha minden végpont maximális sebességgel küld adatokat, akkor nagymértékű, főleg csomós jellegű csomagvesztés keletkezhet az erőforrások túlterhelése miatt. A GENI javaslata alapján ezt a csomagvesztést hatékony hibajavító kódolás alkalmazásával ellensúlyozhatjuk. Az eddig ismertetett megoldásnak számos előnye van. Az egyik a hatékonysága, ugyanis ez a módszer minden hálózati erőforrást mindig teljesen kihasznál és azonnal felhasználja a rendelkezésre álló új kapacitásokat is a hálózatban. Újabb előnye az egyszerűsége, valamint az, hogy a csomagvesztések hatékony hibajavító kódolással történő helyreállítása következtében támogatja a kis méretű bufferek alkalmazását a routerek esetén. Végül fontos a módszer stabilitása is, mivel a maximális sebességű adatküldés alkalmazása sokkal előrejelezhetőbb forgalmat jelent a TCP-hez képest, amely esetén a küldési sebesség nagy mértékben ingadozhat, ami ezt megnehezíti. Az itt leírt előnyök különösen kedvezőek az optikai hálózatokra nézve, ahol csak kisebb méretű bufferek alkalmazására van jelenleg lehetőség [10].

Az 1. ábra egy szökökút kódoláson alapuló hálózati architektúrát ábrázol. A küldő folyamatok az adatot Raptor kódolással kódolják [11], majd maximális sebességgel küldik. Ezt a sebességet két dolog korlátozhatja, az egyik maga a küldő alkalmazás, a másik pedig a link kapacitása. A Raptor kódolás lehetővé teszi azt, hogy a beérkező kódolt folyam bármely $(1 + \varepsilon)k$ méretű részéből nagy valószínűséggel visszaállíthassuk az eredeti k szimbólumot, bármely kis $\varepsilon > 0$ valós szám esetén. Ez a tulajdonság rendkívül hibátűrő adatátvitelt tesz lehetővé még akkor is, ha a csomagvesztés mértéke dinamikusan változik, vagy nagymértékben csomósodik. Ha ekkor a dekódolás sikertelen lenne, akkor a vevő megpróbálhat újabb kódolt szimbólumokat gyűjteni, és újra próbálkozhat a dekódolással.



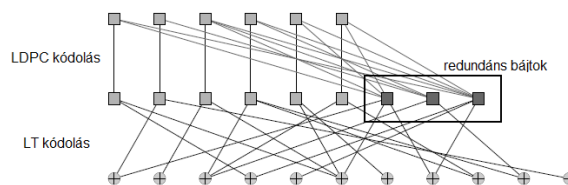
1. ábra Szökökút kódoláson alapuló architektúra

A maximális sebességű adatküldés felveti az alkalmazott mechanizmus igazságosságának kérdését. A versengő

folyamok hozzáférése különböző sebességű lehet. Ez egy szűk hálózati keresztmetszet esetén azt jelentheti, hogy a mohóbb folyamatok kiéhezetik a kevésbé agresszív folyamatokat. Ahhoz, hogy megoldjuk a maximális sebességgel küldő folyamat esetén ezt a problémát, igazságos ütemezést tételezünk fel a routerek esetén, mint például a WFQ, vagy a DRR (Deficit Round Robin) ütemező [12].

III. DIGITÁLIS SZÖKŐKÚT ALAPÚ TRANSPORT PROTOKOLL

A Digitális szökökút alapú transport protokoll (Digital Fountain based Communication Protocol, DFCP) egy kapcsolat-orientált, megbízható működést biztosító transport protokoll [13]. A TCP-vel szemben ez a protokoll egyáltalán nem alkalmaz torlódásszabályozást, hanem Raptor kódoláson alapuló hibajavító kódolás segítségével képes hatékony működést elérni. A DFCP implementációja a 2.6.26-2 verziójú Debian Linux kernelben történt C nyelven. A protokoll működése három fázisra bontható, amelyek a *kapcsolat-felépítés*, az *adatküldés* és a *kapcsolatbontás*. Ha a kapcsolat létrejött, akkor megkezdődhet az *adatátvitel*. Ebben a fázisban a küldő tetszőleges mennyiségű adatot küldhet, amelyet a kernel előre meghatározott, fix méretű blokkokra bont, majd a blokkokat egyenként Raptor kódolás alkalmazásával kódolja a 2. ábrán látható módon.



2. ábra A DFCP által alkalmazott Raptor kódolás

Az alkalmazott kódolás egy *előkódolásból* és egy *LT (Luby Transform) kódolásból* tevődik össze [11]. Az előkódolás gyakorlati realizációja LDPC (Low-Density Parity-Check) kódolással történik [14]. Az előkódolás során a blokkot alkotó bájtokhoz redundáns bájtokat rendelünk hozzá, amelyeket az eredeti blokk bájtjaiból származtatunk egy adott valószínűség-eloszlás szerint, XOR művelettel. Az előkódolás kimenetét LT kódolás segítségével elméletileg végtelen hosszú, a gyakorlatban korlátos, egy paraméter által szabályozható hosszúságú, kódolt bájtokból álló folyamra transzformáljuk. A kódolt bájtokat csomagokba, majd keretekbe szervezve a vevő felé továbbítjuk. A vevő a kapcsolat-felépítés során szinkronba kerül a küldővel, így nagy valószínűség szerint képes a kódolt bájtok dekódolására.

Minden csomag tartalmaz egy olyan *sorszámot*, amely meghatározza, hogy az adott csomag mely blokkhoz tartozik. Először is, ha a csomagok nem megfelelő sorrendben érkeznek meg a vevőhöz, a sorszám alapján a blokkot megtudhatja. Másodszor, a blokkokon belül nem szükséges a csomagok sorrendjének visszaállítása, mert az alkalmazott Raptor kódolás tulajdonságai biztosítják a protokoll számára, hogy csak a kódolt bájtok mennyisége számítson, így a sorrendjük nem lényeges.

A DFCP által alkalmazott *folyamvezérlés* szerepe, hogy megóvja az esetleg lassabb vevőt az elárasztástól. A DFCP

esetén egy csúszóablakos mechanizmus került megvalósításra, ahol az ablak mérete meghatározza azt, hogy hány blokk kerülhet elküldésre, mielőtt nyugtára kezd várakozni a küldő oldal. A vevő oldal minden blokk sikeres vételéről nyugtát küld, amelyek megérkezése esetén a küldő további blokkokat küldhet.

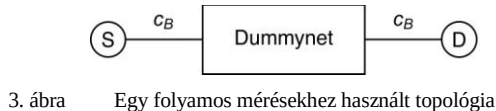
IV. MÉRÉSI EREDMÉNYEK

A DFCP teljesítményének elemzése céljából összehasonlító jellegű méréseket végeztünk az eddig elkészült megvalósítással szimulációk és teszhálózati mérések által [15]. Ennek során két elterjedt TCP verzióval hasonlítottuk össze a DFCP-t. Az egyik TCP verzió a TCP NewReno volt, amely a hagyományos TCP hatékonyságát növelő mechanizmusokat tartalmaz, mint például a SACK (Selective Acknowledgment) [16]. A másik a TCP CUBIC volt, amely jelenleg az alapértelmezett TCP variáns a Linux kernelben. A mérések során hálózati topológiákat hoztunk létre, majd különböző egyenletes eloszlás szerinti csomagvesztés és késleltetés értékeket emuláltunk a hálózatban. Megmértük az említett három protokoll *goodput* értékét, amely az egy másodperc alatt továbbított hasznos bitek mennyiségét jelenti, és ezáltal összehasonlítottuk a protokollok által elért teljesítményt.

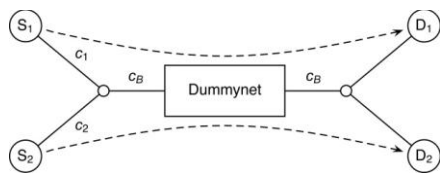
A. Mérési környezet

A teszhálózati mérések során a hálózati paraméterek emulációjához a *Dumynet*et [17], a szimulációk esetén pedig az *ns-2* programot használtuk fel [18]. Az elvégzett mérések között volt olyan, amikor egyszerre csak egyetlen folyam volt a hálózaton, illetve megvizsgáltuk azt is, mi történik versengő folyamatok esetén, amikor két, ugyanazt a protokollt alkalmazó folyam van egyszerre jelen a hálózaton.

Az egy folyam mérésekhez a 3. ábrán látható topológiát, a versengő folyamatokhoz a 4. ábrán látható topológiát használtuk. A bottleneck link kapacitása $C_B = 1$ Gbps, a mérések időtartama 60 másodperc volt.



3. ábra Egy folyam mérésekhez használt topológia

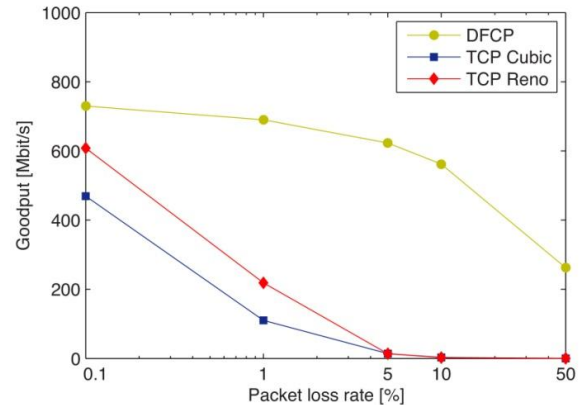


4. ábra Két folyam mérésekhez használt topológia

B. Az elvégzett mérések

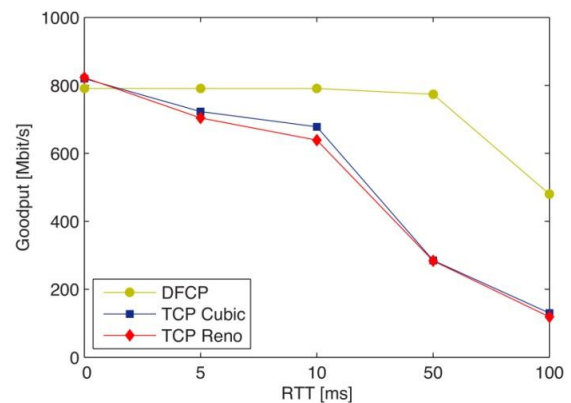
Először az egy folyammal elvégzett mérések eredményeit mutatjuk be. Az 5. ábrán azt láthatjuk, hogy változó, egyre növekvő csomagvesztés esetén hogyan változik az elért *goodput* érték az egyes vizsgált protokollok esetén. Itt minden csomagvesztés értéknél egy optimális redundancia értéket állítottunk be a DFCP esetén, amely azt jelenti, hogy minden blokkból azt a rögzített, minimális számú csomagot küldte ki a protokoll, amennyi éppen elegendő a sikeres dekódoláshoz. Ezt az optimális redundancia értéket előzetesen, külön-külön minden csomagvesztés értékhez meg kellett keresni, és ezen

értékek használatával történtek a mérések. Azt a következtetést lehet levonni, hogy a DFCP esetén nagyobb csomagvesztésnél sem csökken le szignifikáns mértékben a *goodput*, így sokkal kevésbé érzékeny a protokoll a csomagvesztésre, szemben a TCP-vel, amelynél jelentős csökkenés figyelhető meg már kisebb csomagvesztés értékeknél is. Azt is láthatjuk, hogy a DFCP még igen magas csomagvesztésnél is működőképes marad.



5. ábra Az elért *goodput* a csomagvesztés függvényében

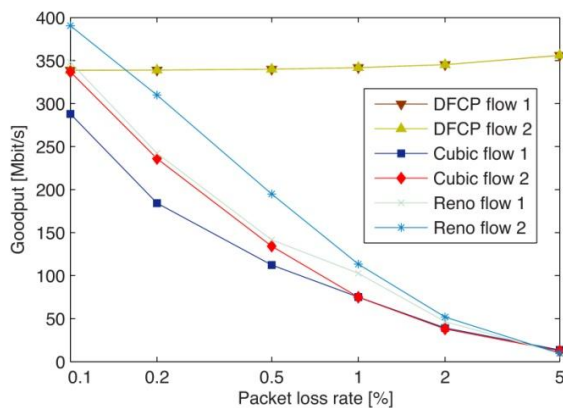
A 6. ábrán a késleltetés volt a vizsgált paraméter. Megfigyelhető, hogy a növekvő késleltetés ellenére a DFCP teljesítménye közel állandó, szemben a TCP-vel, amelynek teljesítménye megint fokozatosan lecsökken, ahogyan nő a késleltetés. Kezdetben a DFCP esetén a kódolási overhead miatt a TCP kicsit jobban teljesít, de utána látható, hogy csak az 50 ms értéknél kezd csökkenni a DFCP teljesítménye, amely annak tudható be, hogy a DFCP ablakmérete ebben az esetben korlátozta az adatküldés sebességét.



6. ábra Az elért *goodput* a késleltetés függvényében

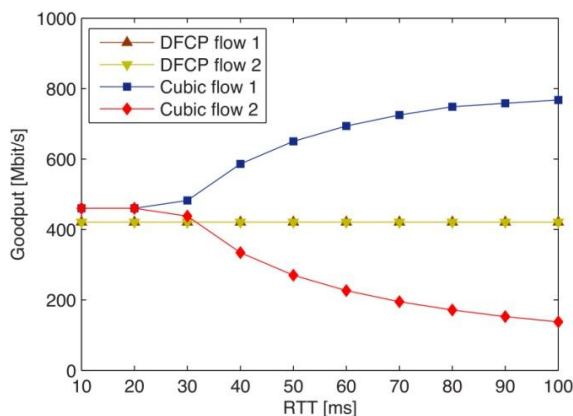
A következő két esetben több folyam volt jelen egyszerre a hálózaton, a mért jellemző itt szintén a *goodput* a csomagvesztés vagy a késleltetés függvényében. A két folyam elindítása egy időben történt, a hálózatban WFQ ütemezőt alkalmaztunk azonos súlyozással. A 7. ábrán egy olyan esetet láthatunk, amikor a korábban említett két TCP verzióval vetjük össze a DFCP teljesítményét. Minden vizsgált protokoll esetén két folyam volt egyszerre jelen a hálózaton. A vizsgált TCP verziók nem voltak képesek igazságosan megosztani a sávszélességet, amíg a két DFCP folyam teljesítménye végig közel azonos volt, ez amiatt történt, mert a DFCP esetén

állandó redundancia értéket állítottunk be, amely az 5%-os csomagvesztéshez tartozó optimális redundancia érték volt, így a DFCP esetén az overhead mindig azonos volt.



7. ábra A goodput a csomagvesztés függvényében, két folyam

A 8. ábrán a változó késleltetés hatásait elemeztük a TCP CUBIC és DFCP esetén. Az első folyam mindig 10 ms késleltetésű, míg a második folyam késleltetése folyamatosan változott 10 ms és 100 ms között. Látható, hogy a TCP folyamok 20 ms-ig képesek igazságosan megosztani a sávszélességet, de ennél nagyobb késleltetéseknél a fix, 10 ms késleltetésű folyam átveszi az irányítást a bottleneck link felett. Ez az úgynevezett *RTT unfairness* probléma a TCP esetén [19]. A DFCP esetén a két folyam végig képes volt igazságosan megosztani a rendelkezésre álló sávszélességen, mivel a DFCP sokkal kevésbé érzékeny az eltérő késleltetés értékekre.



8. ábra A goodput a késleltetés függvényében, két folyam

V. ÖSSZEFOGLALÁS

A cikkben egy olyan új koncepciót mutattunk be a jövő Internetje számára, amely hibajavító kódoláson alapul és képes hatékony, igazságosság szempontjából is kedvező működést elérni. Megvalósítottunk egy transzport protokollt, amely ezt az elvet alkalmazza és különböző hálózati paraméterek emulációjával megvizsgáltuk a működését, viselkedését, mind egy folyamatos, mind több folyamatos körülmények között. A kezdeti mérések eredményei alapján ígéretesnek tűnik az elképzelés, amely szerint a torlódásszabályozást elhagyhatjuk, mivel számos esetben a DFCP teljesítménye jobb a jelenleg elterjedten használt TCP verzióknál. A jövőben a protokollt

folyamatosan fejlesztjük és további méréseket végzünk többféle topológián (pl. parking lot), illetve többféle ütemezőt tartalmazó architektúrákon (pl. WFQ, DRR és Droptail), összevetve a teljesítményét a jelenlegi protokollokkal annak érdekében, hogy megtaláljuk az esetleges gyenge pontokat mind a TCP, mind a DFCP jelenlegi megvalósítása esetén, így minél jobb megoldást adva a jövő Internetje számára.

IRODALOMJEGYZÉK

- [1] Y.-T. Li, D. Leith, R. N. Shorten, "Experimental Evaluation of TCP Protocols for High-Speed Networks", *IEEE/ACM Transactions on Networking*, vol. 15, no. 5, pp. 1109–1122, 2007.
- [2] S. Molnár, B. Sonkoly, T. A. Trinh, "A Comprehensive TCP Fairness Analysis in High Speed Networks", *Computer Communications, Elsevier*, vol. 32, no. 13–14, pp. 1460–1484, 2009.
- [3] D. Clark, S. Shenker, A. Falk, "GENI Research Plan (Version 4.5)", April 23, 2007.
- [4] B. Raghavan, A. C. Snoeren, "Decongestion Control", *Proceedings of the 5th ACM Workshop on Hot Topics in Networks*, pp. 61–66, Irvine, CA, USA, 2006.
- [5] T. Bonald, M. Feuillet, A. Proutiere, "Is the 'Law of the Jungle' Sustainable for the Internet?", *Proceedings of the 28th IEEE Conference on Computer Communications*, pp. 28–36, Rio de Janeiro, Brazil, 2009.
- [6] L. Lopez, A. Fernandez, V. Cholvi, "A Game Theoretic Comparison of TCP and Digital Fountain Based Protocols", *Computer Networks, Elsevier*, vol. 51, no. 12, pp. 3413–3426, 2007.
- [7] Shakeel Ahmad, Raouf Hamzaoui, and Marwan Al-akaidi. Robust Live Unicast Video Streaming with Rateless Codes. In *Proceedings of 16th International Workshop on Packet Video, PV 2007*, pages 78–84, Lausanne, Switzerland, November 2007.
- [8] A. Botos, Z. A. Polgar, V. Bota, "Analysis of a Transport Protocol Based on Rateless Erasure Correcting Codes", *Proceedings of the 2010 IEEE International Conference on Intelligent Computer Communication and Processing*, vol. 1, pp. 465–471, Cluj-Napoca, Romania, 2010.
- [9] D. Kumar, T. Chahed, E. Altman, "Analysis of a Fountain Codes Based Transport in an 802.11 WLAN Cell", *Proceedings of the 21st International Teletraffic Congress*, pp. 1–8, Paris, France, 2009.
- [10] N. Beheshti, Y. Ganjali, R. Rajaduray, D. Blumenthal, N. McKeown, Buffer sizing in all-optical packet switches, in: *Optical Fiber Communication Conference, OFC, Anaheim, CA, OThF8*, 2006.
- [11] A. Shokrollahi, "Raptor Codes", *IEEE Transactions on Information Theory*, vol. 52, no. 6, pp. 2551–2567, 2006.
- [12] M. Shreedhar and G. Varghese, "Efficient fair queueing using deficit round robin." in *Proc. ACM SIGCOMM*, 1995, pp. 231–242.
- [13] S. Molnár, Z. Móczár, B. Sonkoly, Sz. Solymos, T. Csicsics, "Design and Performance Evaluation of the Digital Fountain based Communication Protocol", *Technical Report*, 2012. <http://hsnlab.tmit.bme.hu/~molnar/files/DFCPTechReport.pdf>, accessed April 5, 2014
- [14] A. Shokrollahi, "LDPC Codes: An Introduction", Technical Report, Digital Fountain Inc., 2003.
- [15] S. Molnár, Z. Móczár, A. Temesváry, B. Sonkoly, Sz. Solymos, T. Csicsics, "Data Transfer Paradigms for Future Networks: Fountain Coding or Congestion Control?", *Proceedings of the IFIP Networking 2013 Conference*, pp. 1–9, New York, NY, USA, 2013.
- [16] T. Henderson, S. Floyd and A. Gurtov, The Newreno Modification to TCP's Fast Recovery Algorithm, RFC 6582, April 2012, 1–16
- [17] Dummynet Network Emulator, <http://info.iet.unipi.it/~luigi/dummynet/>, accessed April 5, 2014
- [18] ns-2 Network Simulator, <http://www.isi.edu/nsnam/ns/>, accessed April 5, 2014
- [19] E. Gvalsetz and J. Kaur, "Decomposing RTT-Unfairness in Transport Protocols," *IEEE Workshop on Local and Metropolitan Area Networks*, pp 1–6, May 2010.

Effects of including more words in text units on bilingual predictive dictionaries

Gábor Szabó, Sándor Juhász Dr.

Department of Automation and Applied Informatics
Budapest University of Technology and Economics
Budapest, Hungary
harry.bme@gmail.com, Juhasz.Sandor@aut.bme.hu

Abstract—In this paper we summarize the research in the field of giving predictive translation typing capabilities to computer-assisted translation tools, including the TransType project. Then we present our dictionary based method that uses simpler statistical techniques to extract possible text unit translations of a large, segmented, bilingual corpus. We investigated how using different types of text units (only words, bigrams or trigrams) affects the size and the goodness of the resulting dictionary. For measuring size of the dictionary both entry number and physical size are considered, and for measuring goodness we chose the ratio of number of keystrokes saved by translators to the length of the translation. Measurements were carried out for English-German and English-Hungarian language pairs.

Keywords—predictive typing; dictionary based; multi-word text units; statistical translation extraction

I. INTRODUCTION

Computer systems have an important role in people's life: make it easier and more efficient. Predictive typing or automatic completion of text input is one tool to make human-computer interaction more efficient. It also helps to reduce the amount of human errors.

One outstanding example of this is using devices with limited text input capabilities. Predictive typing had an important role in speeding up text input on mobile phones with hard keys [7], by allowing to press a key on the phone only once for each character and disambiguating using a dictionary, i.e. trying to determine which word the user wanted to type. How and Kan suggested automatic word completion based on the keys typed and the previous word [4]. Masui investigated the text-input productivity gain for English and Japanese language by using automatic word completion on pen-based soft key input devices, based on a predictive dictionary that stores words with their context [5]. Approximate string matching based on spatial key layout was also presented in that work, to make the system tolerant for typos.

It is an interesting idea to make predictive typing available between different languages. It can be a powerful feature to make translator's work more productive inside a computer assisted translation tool, i.e. automatic completion is offered while the user is typing the translation of a text.

The groundbreaking research in predictive translation typing was the TransType project [2][3]. Researchers of that project used machine translation systems to produce possible translations that could complete the text being typed by the translator. Words or shorter sequence of words were offered as translation completion to the user in that project.

Research were continued in the TransType2 project, a successor of TransType. As a new approach, full sentence translation completions were offered to the translator. In each iteration she accepts some prefix of the completion, then types some other text, and the computer gives another full sentence completion. Alignment templates, phrase-based models and stochastic finite-state transducers were examined in TransType2 as translation models [1].

Machine translation engines are complex systems. This paper presents a simpler approach for predictive translation typing: it uses a dictionary to offer typing completions. In that sense, it is similar to the typing completion solutions used in mobile phones.

There are multiple aspects of creating such dictionary. One of them is how long text fragments are stored in the dictionary. The measurement presented in this paper compares three kinds of dictionaries: one that uses only one-word sequences, another that includes bigrams, and a thirds that includes bigrams and trigrams as well.

II. DICTIONARY BASED PREDICTIVE TRANSLATION TYPING

A. Concept

The concept of dictionary based prediction is to create a dictionary from some source once, and use it to give predictions many times later. A bilingual dictionary is created and used in the presented case. The source of the dictionary is a large, segment-aligned, bilingual corpus. The corpus is processed with statistical methods in order to find translations of text fragments.

Typing completions can be offered based on these text fragments: if a known text fragment is found in the source language sentence and the translator types something that is the beginning of a translation in the dictionary, then this known

translation can be shown as a possible completion. Text fragments are called text units in this paper. The exact definition of text unit is presented at the measurement's description, but they can be thought of like some words from a sentence that are linked together in a way, e.g. they form an expression or subsentence.

An entry in the dictionary is a source language text unit coupled with a list of possible translations. Each possible translation in an entry consists of a target language text unit and a score representing the probability of being a correct translation. The score is used when the typing completions are sorted and filtered. Filtering to the best n completions is necessary (n is around 5-6), otherwise too many completions would just hinder the translator's work. An example dictionary is presented in Table 1.

TABLE I. SAMPLE ENGLISH-GERMAN DICTIONARY FOR PREDICTIVE TYPING. EACH SOURCE LANGUAGE TEXT UNIT HAS A LIST OF POSSIBLE TRANSLATIONS ALONG THEIR SCORES

Source language text unit	Possible translations
I would	(möchte; 0.45), (möchte ich; 0.31)
new	(neuen; 0.55), (neue; 0.49)
Parliament	(Parlament; 0.75), (Parlaments, 0.40)
and I	(und ich; 0.54)
and	(und; 0.88), (der; 0.61), (die; 0.59), (in; 0.46), ...

The structure of the dictionary implies that the method is asymmetrical, i.e. the dictionary can only be used for translation from the source language to the target language.

B. Dictionary creation

The dictionary is created in roughly three steps. Firstly, text units are collected from each side of the bilingual corpus individually (i.e. the source and target language). The "rare" text units are filtered out. Here, being "rare" is defined by the number of occurrences in the corpus. If the occurrence falls under a given threshold o_{min} , then the text unit is considered rare.

After that, the remaining text units of the source language text are paired with text units of the target language text, and their concomitant occurrences are counted (i.e. how often they appear together in a segment and its translation).

Then, a score is assigned to every text unit pair based on the individual and concomitant occurrences. If this score falls under a given threshold s_{min} , then the pair is omitted, otherwise it is added to the dictionary as a possible translation entry (along with the score). The score is assigned to a (s ; t) source and target language text unit pair by applying the following formula:

$$score = 2 \cdot N_{s,t} / (N_s + N_t). \quad (1)$$

Here $N_{s,t}$ denotes the number of sentence pairs where s and t occurs together, and N_s and N_t denotes the number of sentences where s or t occur, respectively. Multiple occurrences in a sentence pair count as one.

This score is called the Dice coefficient comparator [6]. It falls in the $[0; 1]$ interval and expresses the ratio of concomitant occurrences to the sum of the individual occurrences.

III. MEASUREMENT

There are basically two aspects of creating a dictionary: the exact definition of a text unit and the thresholds. These parameters can be exactly set before the measurement and they can alter the size of the resulting dictionary. The measurement presented in this paper uses a fix set of thresholds and three kind of text unit: words, words and bigrams, and words, bigrams and trigrams. Bigrams and trigrams cannot overlap subsentence boundaries.

The goal of the measurement is to investigate how using these text units affect the size and goodness of the dictionary. Size can be measured both in physical size (e.g. in megabytes) and in entry number.

The metric for the goodness of the dictionary is keystroke saving ratio, i.e. how much part of the typing can be saved if the translator uses the dictionary. It is automatically measured with a bilingual corpus. The source language sentences of the bilingual corpus are the sentences the hypothetical translator wants to translate, and the target language sentences are the translations the translator would like to type. When the simulated translator starts translating a new sentence, the sentence is split into text units, and these text units are looked up in the dictionary for possible translations. If the translator starts typing a new word, and it is the prefix of some possible translations, those possible translations are offered as completions. The best six completions is shown to the translator, and she chooses the longest appropriate completion. The number of keystrokes saved by the translator is: [length of completion] – [length of typed prefix] – [index of completion in the offered list].

Expanding the definition of text unit by including bigrams and trigrams obviously results at least the same dictionary size as the one using only words. This is because the text units, which are kept in the first step of dictionary creation using only words, are also kept including bigrams and trigrams. It is expected that some actual bigrams and trigrams survive the filtering of the first step and the scoring of the last step, meaning the dictionary size will definitely increase. The questions is whether the keystroke saving ratio will also increase.

The measurement program was written in C# language targeting .NET Framework 4.5. Measurements were carried out on a 64-bit Microsoft Windows 8.1 operating system, using an Intel Core i7-2630QM 2.00 GHz CPU with 8 GB RAM.

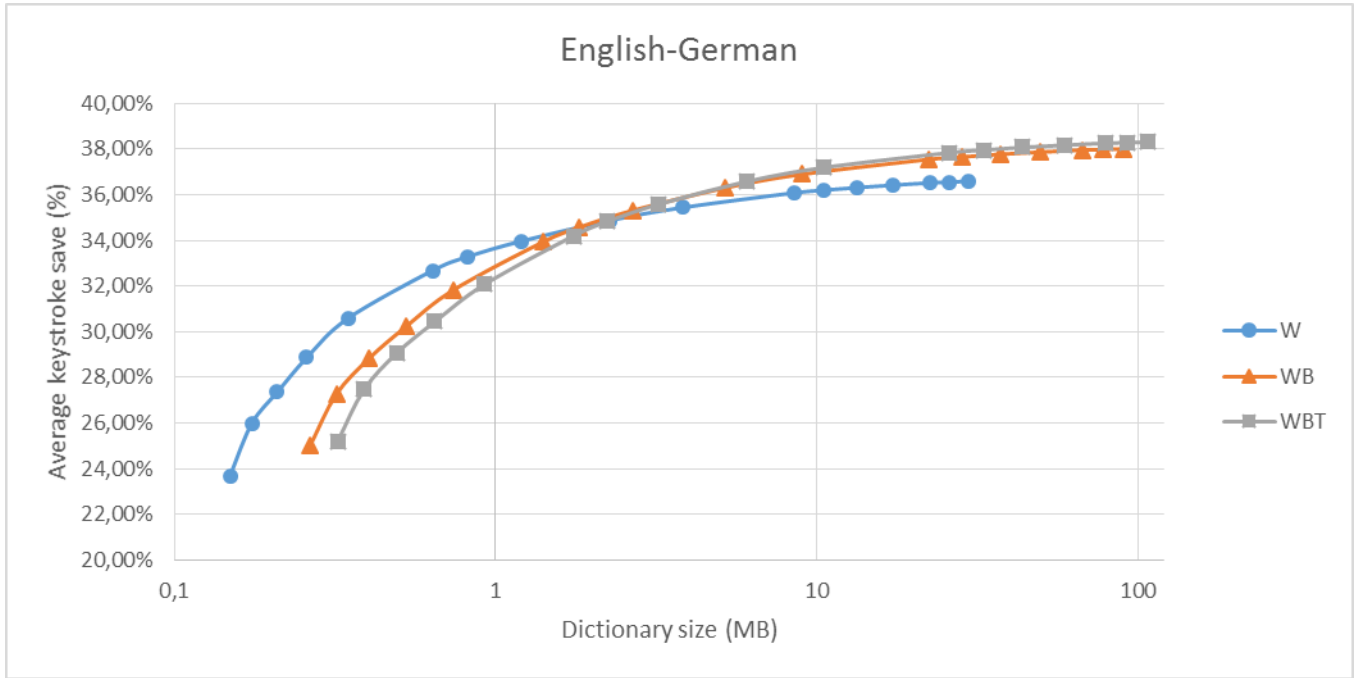


Figure 2. Keystroke saving results for English-German language pair, based on the physical size of the dictionary. For small dictionary sizes the word-only approach is more efficient, but including bigrams and trigrams result in greater keystroke saving beyond 3 megabytes.

The bilingual corpus were extracted from the proceedings of the European Parliament, downloaded from the OPUS project [8]. English-German and English-Hungarian language pairs were chosen. 200k sentence pairs were used for dictionary creation and 20k for keystroke saving measurement. The thresholds used: $o_{min} = 1e-3$, $s_{min} = 0.3, 0.25, 0.2, 0.15, 0.1, 5e-2, 4e-2, 3e-2, 2e-2, 1.5e-2, 1e-2, 9e-3, 8e-3, 7e-3, 6e-3$,

$5.5e-3, 5e-3$. Thresholds were chosen so that dictionary size is between a few hundred kilobytes and a hundred megabytes.

IV. RESULTS

Measurement results can be seen on Fig. 1, Fig 2 and Fig 3. The dictionary sizes on the horizontal axes are displayed with a logarithmic scale for better readability. The letters on the

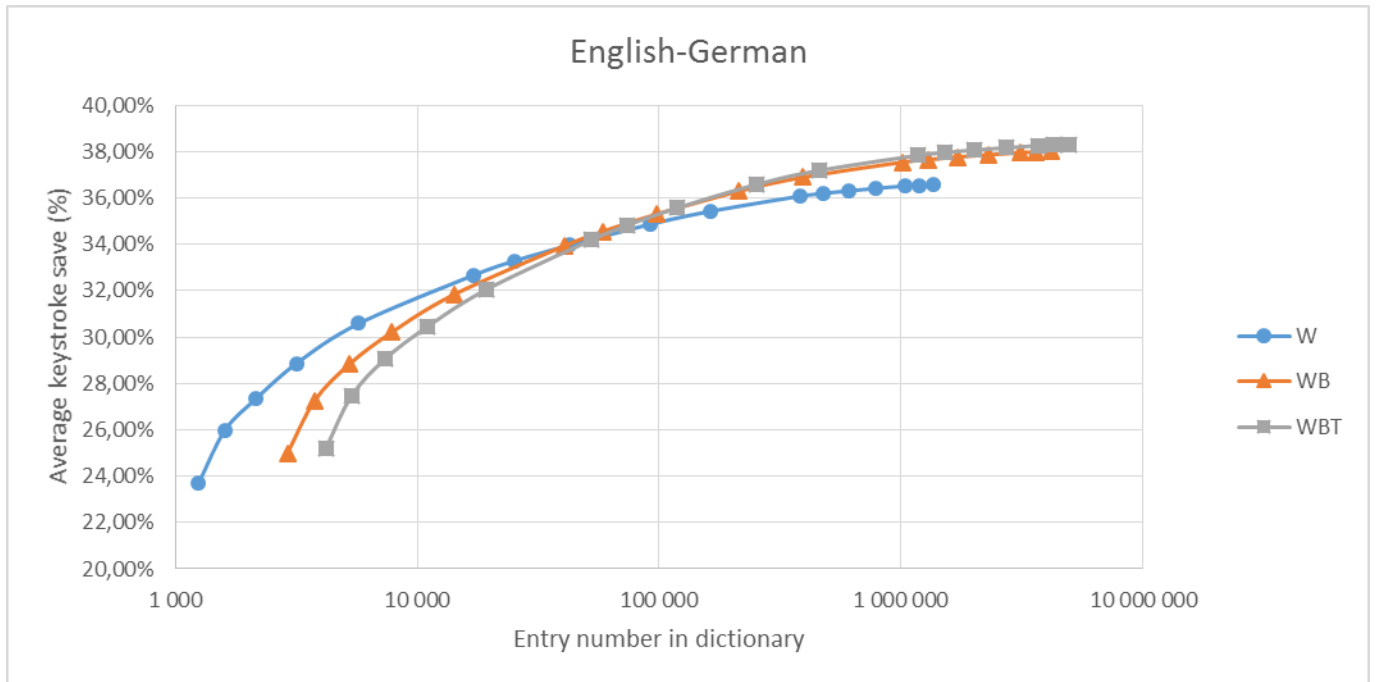


Figure 1. Keystroke saving results for English-German language pair, based on the entry number in the dictionary. For small dictionary sizes the word-only approach is more efficient, but including bigrams and trigrams result in greater keystroke saving beyond 50 000 entries.

legend stands for: W – words; WB – words and bigrams; WBT – words, bigrams and trigrams.

First of all, the curves and their relative positions are similar in the two figures. This means that using physical size of the dictionary gives basically the same results as the number of entries, they are interchangeable.

Secondly, it can be concluded that keystroke saving increases with dictionary size, but its marginal keystroke saving decreases, i.e. it reaches a plateau. The maximum keystroke saving value is about 36.6% for the only-word case, 38.0% for the word-bigram case, and it is around 38.3% for the word-bigram-trigram case for the English-German language pair.

Keystroke saving values are lower for the English-Hungarian language pair, the maximums are around 31-32%. The reason behind this is the fact that the Hungarian language is agglutinative, therefore the number of frequently used text units are higher, making it more difficult to predict the one being typed.

Dictionaries using only words performs better at small sizes (the other curves seems to be shifted right in that interval), but including bigrams and trigrams result in greater keystroke saving after 3 MB or 50k entries. Hence, using only words is better for small, compact dictionaries. On the other hand, state of the art computers can handle a 100 MB dictionary easily, so it's worth including longer text units in the dictionary.

Including trigrams in addition to bigrams changes only a little. This could be caused by the o_{min} threshold being too high, but lowering that threshold causes insufficient memory problems in the pairing and scoring step.

V. SUMMARY

We presented a dictionary based approach for predictive translation typing in this paper. We investigated one aspect of dictionary creation: how does including multi-word text units in the dictionary affect the dictionary size and the keystroke

saving. Measurements were carried out on English-German and English-Hungarian language pairs.

It was stated that including bigrams and trigrams along words in a dictionary makes its size bigger, but comes with greater keystroke saving. As state of the art computers can handle even the bigger dictionaries easily, the use of multi-word text units is advisable.

As future study, other kinds of text units could be investigated. For example, the definition could be loosened, and not just consecutive words could be allowed, or the order of the words could be changed.

This dictionary based approach could be compared to monolingual prediction and to the TransType solutions. The sensitivity and optimal interval of the thresholds could also be measured.

REFERENCES

- [1] Barrachina, Sergio, et al. "Statistical approaches to computer-assisted translation." Computational Linguistics 35.1 (2009): 3-28.
- [2] Foster, George, Pierre Isabelle, and Pierre Plamondon. "Target-text mediated interactive machine translation." Machine Translation 12.1-2 (1997): 175-194.
- [3] Foster, George, Philippe Langlais, and Guy Lapalme. "TransType: text prediction for translators." Proceedings of the second international conference on Human Language Technology Research. Morgan Kaufmann Publishers Inc., 2002.
- [4] How, Yijue, and Min-Yen Kan. "Optimizing predictive text entry for short message service on mobile phones." Proceedings of HCI. Vol. 5. 2005.
- [5] Masui, Toshiyuki. "An efficient text input method for pen-based computers." Proceedings of the SIGCHI conference on Human factors in computing systems. ACM Press/Addison-Wesley Publishing Co., 1998.
- [6] McInnes, Bridget T. Extending the log likelihood measure to improve collocation identification. Diss. UNIVERSITY OF MINNESOTA, 2004.
- [7] Silfverberg, Miika, I. Scott MacKenzie, and Panu Korhonen. "Predicting text entry speed on mobile phones." Proceedings of the SIGCHI conference on Human factors in computing systems. ACM, 2000.
- [8] Tiedemann, Jörg. "Parallel Data, Tools and Interfaces in OPUS." LREC. 2012.

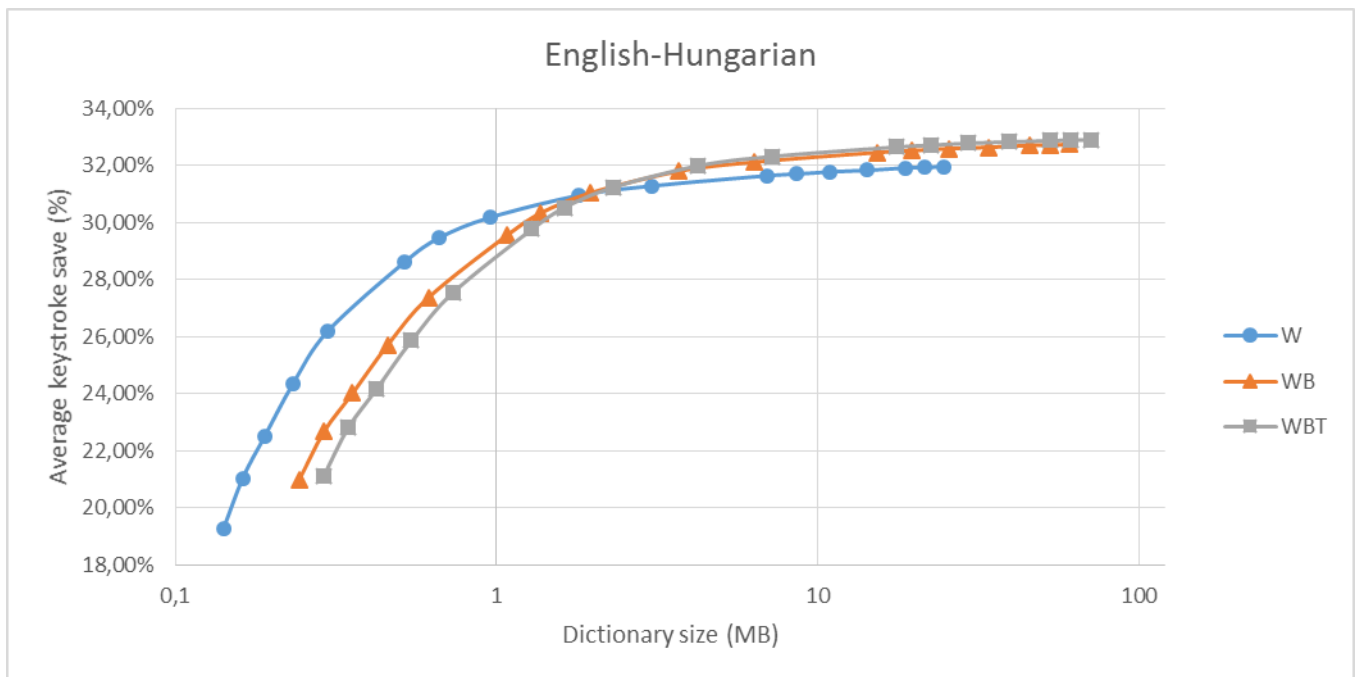


Figure 3. Keystroke saving results for English-Hungarian language pair, based on the physical size of the dictionary. For small dictionary sizes the word-only approach is more efficient, but including bigrams and trigrams result in greater keystroke saving beyond 2 megabytes.

Automated abstraction in model-driven engineering

Csaba Debreceeni

Budapest University of Technology and Economy,
Department of Measurement and Information Systems,
1117 Budapest, Magyar tudósok krt. 2.
debreceenis@mit.bme.hu

Abstract—In software development, the model-driven engineering (MDE) paradigm envisions a process starting from high-level system models and through several well-defined abstraction and refinement steps automatically derive source code, documentation or configuration artifacts.

In MDE the concepts of view-models are widely used to provide specific focus for the designers by abstracting away unnecessary details of the underlying system model to highlight relevant domain specific information (e.g., reliability model, networking architecture, etc.). However, maintaining these abstract views upon changes in the system model is cumbersome, especially in order to avoid the recalculation of the complete view and provide instantaneous view-model updates.

In this paper, I propose an approach to incrementally synchronize abstract view-models defined by declarative model queries. These view-models are automatically populated with derived objects in accordance with the lifecycle of regular model elements. The approach is evaluated on an ongoing avionics research project built on top of the Eclipse Modeling Framework (EMF).

Keywords: *model-driven engineering, derived object, view-models, abstraction*

I. INTRODUCTION

Modeling is considered as one of the most elementary discipline in engineering. Its purpose is to overcome complexity, increase understandability or ease design. In software development, the model-driven engineering paradigm follows this concept by envisioning a process starting from high-level system models and through several well-defined abstraction and refinement steps automatically derive source code, documentation or configuration artifacts.

In Model-driven Engineering (MDE), the concept of view-models (e.g., reliability model, hardware architecture etc.) automatically derived from detailed system models to highlight relevant, domain specific information is a widely used and accepted approach (model abstraction). Synchronization and maintenance between these system (source) models and their different view-models are usually done by batch transformations. Unfortunately, as these transformations lack of incremental execution whenever a change happens in the corresponding system model the view-models have to be cleared and then all the transformations have to be executed to rebuild them

In this paper, I propose an approach that can incrementally synchronize abstract view-models defined by a set of queries. These view-models consist of derived objects automatically materialized from the source model tightly following its lifecycle, resulting in full featured (view-)model.

An initial prototype framework is built upon the Eclipse Modeling Framework (EMF) [12] considered as the de facto industry standard for modeling, and EMF-IncQuery [11], an incremental graph pattern matching framework with declarative pattern definition capabilities. Finally, the approach is evaluated on an ongoing avionics research project.

The rest of the paper is structured as follows. Section II provides a motivating example captured from avionics domain. Section III gives an overview of the approach where the main challenges are also mentioned. Section IV describes the methodology to define view-models with model queries. The concrete working mechanism is presented in Section V while in Section VI, related works are assessed. Finally, Section VII concludes this paper.

II. A MOTIVATING SCENARIO: MODEL-ABSTRACTION IN AN AVIONICS SYSTEM

The motivating example of this paper is extracted from an industrial avionics research project where the main purpose was to define a complex, semi-automated development chain to allocate software functions onto different hardware architectures. The tool takes Matlab Simulink (SIM) block diagram models as input, and it automatically abstracts them into a functional architecture model (FAM), which later serves as an input for the allocation step. The example is simplified by considering only subsystems collected into functions and complex block-port-signal interconnections into information links.

A sample Simulink model depicted in Figure 1., consists of five blocks, out of which *Navigation*, *Engine Management System* (EMS) and *Flight Management System* (FMS) represent functions (blue rectangles tagged with “func”) while SubS1 and SubS2 represent blocks which are irrelevant for allocation (green rectangles without a “func” tag). The abstracted FAM model contains only equivalents of the three function blocks and it also abstracts from complex block-port-signal interconnections, representing them as (logical) links between functions where the provider function sends data to the consumer.

This work was partially supported by the CERTIMOT (ERC HU-09-01-2010-0003), MONDO (EU ICT-611125), TÁMOP-4.2.2.C-11/1/KONV-2012-0001 and a collaborative project with Embraer. The TÁMOP-4.2.2.C-11/1/KONV-2012-0001 project has been supported by the European Union, co-financed by the European Social Fund.

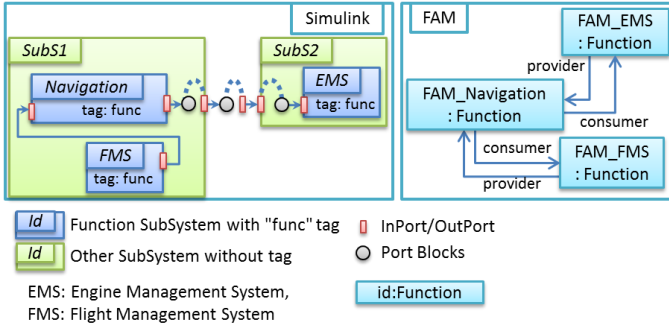


Figure 1. Simulink model example and its proper FAM representation

As only the SIM models are allowed to be modified, the abstract FAM is a view model which always reflects the current structure of the underlying SIM model.

The main concepts of these modeling languages are captured by corresponding metamodels in EMF presented in Figure 2. The main *EClass* (represented by a rectangle) of the Simulink metamodel is the *SubSystem* that includes *InPort* and *OutPort* elements and *PortBlocks*. It has a tag *EAttribute* which stores the role of the element. Subsystems may contain *Blocks* along the subBlocks *EReference*. A *Signal* represents a communication link between one *OutPort* and some *InPorts* of different *SubSystems*. The (simplified) FAM metamodel contains *Function* elements, where a function can be connected to another function as a representation of information links. A generic *Traceability* metamodel is used to define directed links between a specific set of *EObject* instances (common supertype for all EMF classes) from the source and target metamodels using *Trace* elements.

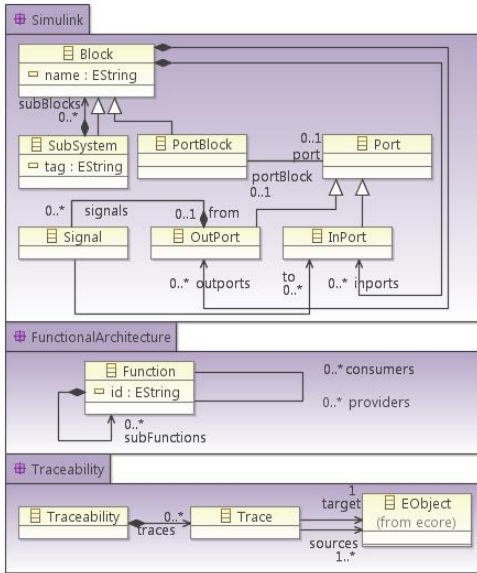


Figure 2. Simulink and Functional Architecture metamodels

III. OVERVIEW OF THE APPROACH

Main purpose of this approach is to create regular models where the elements of them are synchronized immediately and incrementally when the corresponding source model is modified. As a process (depicted on Figure 3.), the approach

uses *source model*, *target metamodel* and *derivation rules* defined by model queries (*item*, *edge*, *attribute*) as input whereas the output is a *target model* that realizes the *target metamodel* with the previously defined properties.

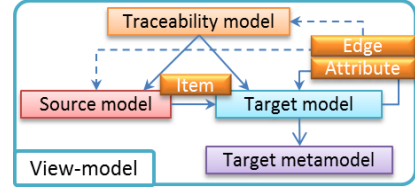


Figure 3. Overview of the approach

Referring to the motivating example, Simulink model is interpreted as *source model*, FAM metamodel represents the *target metamodel*, while the concrete FAM model will be the *target model* (derivation rules are explained in Section IV).

An initial prototype implementation is built upon the EMF-IncQuery and its Viewers [10] technology. EMF-IncQuery is an incremental model query framework that allows to define queries via its pattern language and incrementally get notifications when a new match appears for a match or an old disappears. The Viewers technology provides an API to display query results for example in a tree-view or a graphical Zest view. It works with a similar terminology as we need such as *item* or *edge*. With some modification, I also extended this layer to be able to use the “attribute” expression. Theoretically, the prototype is a Viewers contribution.

During the implementation, I faced with the following challenges: *How to define derivation rules and view-models by derivation rules?* Derivation rules, in this solution, can be given by model queries where a match can be interpreted as a precondition for a rule and the execution is defined as additional information to the specified pattern. *How to trace the created elements during the synchronization?* An important task is to get the ability to trace all view-model elements from the source objects. It allows the approach to notify only the necessary part of the view-models about a change from the source model. *How to execute these derivation rules?* At the same time, more matches can appear during the modification of the source model, but the approach has to solve the conflicts between the rules and in all cases, the output has to be deterministic.

IV. DEFINITION OF VIEWMODELS

View models are defined by using a fully declarative, rule-based formalism. Preconditions of rules are defined by model queries (building on the language of the EMF-IncQuery framework [1]), which identify parts of interest in the source model. Derivation rules then use the result set of a query to define elements of the view model. Informally, when a new match appears in the result set of a query then the corresponding derivation rule is fired to create elements of the target view model. When an existing match disappears in the result set of a query, the inverse of the rule is fired to delete the corresponding view model elements.

The capabilities of the EMF-IncQuery language is demonstrated in Listing 1. using a few simple patterns. Pattern

function enumerates all SubSystems that are tagged as “func”. The trace pattern lists all possible pairs of elements that are connected via a Trace model element. Note that the types of the parameters are undefined which enables to use the same pattern for arbitrary model elements connected by the traceability metamodel. The functionId pattern pairs the name of the source Block and the corresponding Function - for this reason, it reuses the previously defined patterns using the find keyword. Pattern inPortToInPort finds all inport pairs connected to each other. The connectedBlocks lists the block pairs linked through several inports. Finally, pattern link select block pairs that have to be linked in the FAM model.

```

pattern trace(src, trg) {
    Trace.sources(trace, src);
    Trace.target(trace, trg); }

@Item(eClass = "Function")
pattern function(subsys : SubSystem) {
    SubSystem.tag(subsys, "func"); }

@Attribute(source = func, target = id, eAttribute = "id")
pattern functionId(func : Function, id : EString) {
    find functionSubsystem(subsys);
    Block.name(subsys, id);
    find trace(subsys, func); }

pattern inPortToInPortConn(src : InPort, trg : InPort) {
    Port.portBlock.outputs.signals.to(src, trg); } or {
    Port.portBlock.inports(outer, src);
    OutPort.signals.to(outer, trg); }

pattern connectedBlocks(prov : Block, cons : Block)
    Block.outputs.signals.to(prov, consumerInPort);
    Block.inports(cons, consumerInPort); } or {
    Block.outputs.signals.to(prov, innerIP);
    find inPortToInPortConnection+(innerIP, consumerInPort);
    Block.inports(cons, consumerInPort); }

@Edge(source = prov, target = cons, eReference = "consumers")
pattern link(prov : Function, cons : Function) {
    find functionSubsystem(p);
    find connectedBlocks(p, c);
    find functionSubsystem(c);
    find trace(p, provider); find trace(c, consumer); }

```

Listing 1. Model queries and derivation rules

The presented pattern language has the ability to add additional information to patterns with different annotations appearing on the top of specific patterns started with “@” character. Instead of relying on a full-fledged model transformation language, I extended the EMF-IncQuery framework to interpret the annotations as derivation rules. As a result, we obtain a declarative formalism compliant with the execution semantics of incremental and non-deleting graph transformations while the language engineering environment of EMF-IncQuery requires very little further adaptation. A derivation rule is obtained when a query is extended by one or more of the following annotations:

- a) **@Item(eClass = "ClassName"):**
The Item annotation is used to add a new view object of type “ClassName” from target metamodel.
- b) **@Edge(source = src, target = trg, eReference = “ref”):**
The Edge annotation is used to set a reference of source to the value of target.
- c) **@Attribute(target = trg, value = val, eAttribute = “attr”):**
The Attribute annotation is used to set an attribute of target to the value of val.

The patterns from Listing 1. have derivation annotations defined that are to be read as follows. The derivation rule of function prescribes the derivation of a view object of type Function in the FAM together with a traceability link. The functionId sets the id attribute of the created object to the name of the corresponding Simulink block (passed by the pattern parameter identifier). Finally, the edge annotation of link pattern connects two function elements in the view-model. Applying these rules on the SIM model of Figure 1. , we obtain all the Function nodes of the FAM model and the appropriate consumers references between them. (Because of the EMF-specific *EOpposite* reference, the providers reference is also set automatically.)

V. POPULATE VIEWMODELS WITH DERIVED OBJECTS

The key to update any target model synchronously is incremental query evaluation, where query result changes caused by model manipulation are also provided without complete reevaluation. We use the EMF-IncQuery framework [1] that supports incremental query evaluation over EMF instance models by constructing a Rete rule network [9] that processes change notifications sent by the EMF notification API. It caches the current query results and propagates changes of them to the Viewers layer of the framework.

Even though the construction and storage of Rete-network imposes a slight performance overhead on model manipulation, and a memory cost proportional to the cache size (approx. the size of match sets), EMF-IncQuery can evaluate very complex queries over large instance models very efficiently.

The framework also provides an interface to notify any observer object when a new match appears or an old one disappears. With this functionality, our query-based approach has the ability to react on these events by firing all the required derivation rules.

In this approach, the previously mentioned derivation rules are interpreted as follows when they are fired if a new match of the annotated patterns is appeared or disappeared:

- a) **@Item:**
Appeared: an EObject is created with the specified type defined in the annotation and Trace object is also appeared where its target is the new object and its sources are the parameters of the match.
Disappeared: when a match disappears that previously defined an object in the view-model, it is possible to find the target object through the traceability model and delete it from the view-model.
- b) **@Edge:**
Appeared: the annotation defines which parameters of the match act as source and target objects. In this case, the framework decides whether the multiplicity of reference is many or single and according to this, adds or sets the target of the reference.
Disappeared: after a match disappears, the framework decides again whether the multiplicity of the reference is many or single and according to this, removes the target to

from the reference or sets its value to the default (which usually means the “null” value).

c) *@Attribute*:

Appeared: the annotation defines which parameters of the match act as target object and value. In this case, the framework tries to convert the value to the corresponding type and then it sets to the specified attribute.

Disappeared: after a match disappears, the framework sets the value of the specified attribute to the default one.

Because the edge and attribute typed derivation rules refer to derived objects in view-models, the item typed derivation rules are fired at the highest priority. If two or more item typed or edge/attribute typed derivation rule can be fired, arbitrary order can be defined. This means that two derivation rules that can be fired at the same time are never in conflict.

In our running example, the view-model is built up in the following sequence (depicted in Figure 4.): (1) a new match appears for the function pattern. (2) A new Trace object is created in traceability model and (3) a new Function element is added to view-model. When the necessary functions exist in the view-model (4) new matches appear for the functionId pattern and (5) the corresponding id values of the specified functions are set according to the matches.

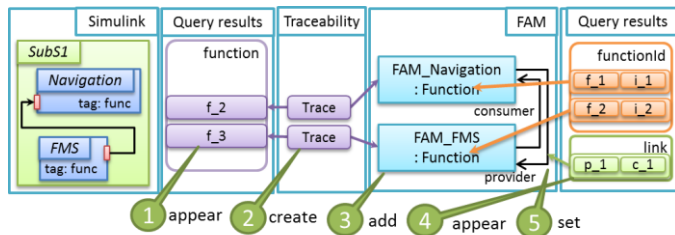


Figure 4. Internal traceability and working mechanism

VI. RELATED WORK

Defining view-models and synchronizing them with the sources is a common problem in MDE. Using derived classes and model queries captured in OCL was first proposed in [2] but without support for incremental view calculation. This solution can be further developed using recent results in incremental support of OCL topic as in [3][4].

View maintenance by incremental and live QVT transformations is used in [5] to define views from runtime models. The proposed algorithm operates in two phase, starting in check-only mode before an enforcement run, but its scalability is demonstrated only on models up to 1000 elements.

VirtualEMF [6] allows the composition of multiple EMF models into a virtual model based on a composition metamodel, and provides both a model virtualization API and a linking API to manage these models. The approach is also able to add virtual links based on composition rules. In [7] an ATL-based method is presented while in [8] correspondences between models are handled by matching rules defined in

ECL, where the guards use queries similarly to our approach, although incremental derivation is not discussed.

I believed that this contribution is unique in incremental synchronization of view-models that can be used to create automatic abstraction from detailed models.

VII. CONCLUSION AND FUTURE WORK

In the paper, an approach is presented for deriving and incrementally synchronizing non persistent view model. To achieve this functionality with the help of the EMF-IncQuery team, I extended the framework to (i) support derivation rule definitions by annotations, (ii) provide an implicit traceability between the source and view model elements and (iii) manage derivation rule execution based on the changes of their match set (disappear, appear).

A main direction for future work, I am planning to support (i) the chaining of view-models along derivation rules and (ii) improve the semantics of derivation rules to define more complex queries. Finally, to provide a deeper evaluation of the scalability of the approach as currently it was only tested on models with no more than 1000 elements.

REFERENCES

- [1] Ujhelyi, Z., Bergmann, G., Hegedűs, A., Horváth, A., Izsó, B., Rath, I., Szatmári, Z., Varró, D.: EMF-IncQuery: An integrated development environment for livemodel queries. *Science of Computer Programming* (2014) In press.
- [2] Balsters, H.: Modelling database views with derived classes in the UML/OCL-framework. In Stevens, P., Whittle, J., Booch, G., eds.: *UML 2003 - The Unified Modeling Language. Modeling Languages and Applications*. Volume 2863 of LNCS. Springer Berlin / Heidelberg (2003) 295–309
- [3] Cabot, J., Teniente, E.: Incremental integrity checking of UML/OCL conceptual schemas. *J. Syst. Softw.* 82(9) (2009) 1459–1478
- [4] Groher, I., Reder, A., Egyed, A.: Incremental consistency checking of dynamic constraints. In: *FASE 2009*. Volume 6013 of LNCS., Springer (2010)
- [5] Song, H., Huang, G., Chauvel, F., Zhang, W., Sun, Y., Shao, W., Mei, H.: Instant and incremental QVT transformation for runtime models. In Whittle, J., Clark, T., Khne, T., eds.: *Model Driven Engineering Languages and Systems*. Volume 6981 of LNCS., Springer Berlin Heidelberg (2011) 273–288
- [6] Clasen, C., Jouault, F., Cabot, J.: Virtual Composition of EMF Models. In: *7thèmes Journées sur l'Ingénierie Dirigée par les Modèles (IDM 2011)*, Lille, France (2011)
- [7] Xiong, Y., Liu, D., Hu, Z., Zhao, H., Takeichi, M., Mei, H.: Towards automatic model synchronization from model transformations. In: *Proceedings of the twenty-second IEEE/ACM international conference on Automated software engineering. ASE '07*, New York, NY, USA, ACM (2007) 164–173
- [8] Kolovos, D.S.: Establishing correspondences between models with the epsilon com-parison language. In: *Proceedings of the 5th European Conference on Model Driven Architecture - Foundations and Applications. ECMDA-FA '09*, Berlin, Heidelberg, Springer-Verlag (2009) 146–157
- [9] Charles L. Forgy. A network match routine for production systems. Working paper, 1974.
- [10] Zoltán Ujhelyi, Tamás Szabó, István Ráth, and Dániel Varró. Developing and visualizing live model queries. In *Proceedings of the First Workshop on the Analysis of Model Transformations (AMT '12)*. ACM (2012.), New York, NY, USA, 35–40.
- [11] EMF-IncQuery, <https://www.eclipse.org/incquery/>
- [12] Eclipse Modeling Framework, <https://www.eclipse.org/modeling/emf>

TDM- és WDM-PON hálózatok analitikus költségelemzése

Czékus József, Mitcsenkov Attila

Budapesti Műszaki és Gazdaságtudományi Egyetem

Villamosmérnöki és Informatikai Kar, Távközlési és Médiainformatikai Tanszék

Budapest, Magyarország

czekus@tmit.bme.hu, mitcsenkov@tmit.bme.hu

Kivonat — A jövő PON hálózatai a hasznos optikai sáv szélesség növelése érdekében vagy kiterjesztik az aktuális TDM megoldások képességeit vagy WDM technológiát alkalmaznak. Munkám során megbecsültem a TDM és WDM-PON hálózatok alkalmazása közötti várható beruházási költségkülönbséget, modellezve a várható rövid- és középtávú változásokat a hardverek költségében. A vizsgálatok során figyelembe vettem a TDM rendszerekben a végpontok forgalmának statisztikus multiplexálásból adódó kapacitásbeli nyereséget is. A forgalmi modellezéshez Pareto-eloszlást alkalmaztam, az osztott közeg statisztikus multiplexálásának számítására. Végül az elkészített költségmodellek és a becsült felhasználószámok hányadosaként kapott, egy felhasználóra jutó beruházási költségek alapján összehasonlítottam a TDM és WDM-PON technológiák kilátásait.

Keywords: TDM-PON, WDM-PON, statisztikus multiplexálás, költségmodell

I. BEVEZETÉS

Napjainkban egyre több internetszolgáltató telepít passzív optikai hálózatot (PON - Passive Optical Network) az előfizetői kiszolgálásához. Jelenleg az időosztáson alapuló TDM-PON (Time Division Multiplexing PON) hozzáférési hálózatok a legelterjedtebbek, azonban a hullámhosszosztásos (WDM - Wavelength Division Multiplexing) hálózati architektúrák várhatóan jobban kihasználják az optikai átvitel nyújtotta nagy sáv szélességet. Ebben a tanulmányban egy olyan módszert szeretnék bemutatni, amellyel meghatározható, milyen sáv szélességigények mellett fordul meg a két technológia költségaránya, azaz mikor éri el a technológiai váltás ideje. A beruházási költségek meghatározása az egyes alkatrészek és komponensek árának változásán alapul.

A pont-multipont rendszerek egy felhasználóra vetített beruházási költségének meghatározásához elsőként meghatároztam az adott kapacitású TDM, illetve WDM-PON rendszerek segítségével kiszámálható végpontok számát. A TDM-PON-okban a nem aktív felhasználók sáv szélessége az aktív felhasználók között megosztható. Ez a statisztikus multiplexálásból származó nyereség jelentős a TDM hálózatokban, ezért fontos volt figyelembe vennem. A második fejezetben bemutatom ezt az általam használt forgalmi modellt.

A fejlődő technológiák költség alapú összehasonlításához költségmodellek készítésére volt szükségem. E modelleket a hálózati komponensek árának összegzésével állítottam elő. A harmadik fejezetben mutatom be az így kapott, a hálózati

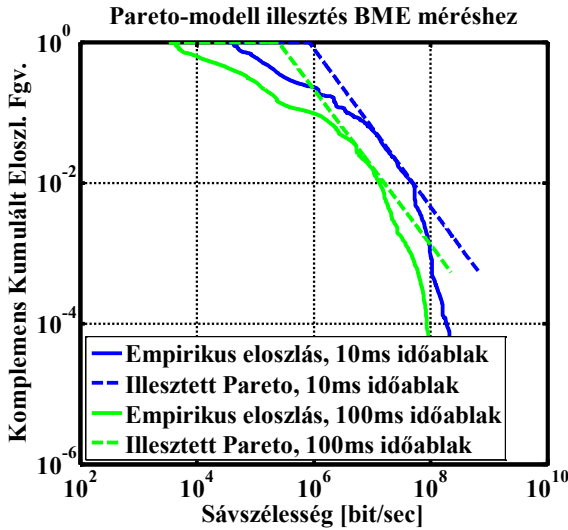
elemek fizikai tulajdonságain alapuló költségmodelleket. Végül a negyedik fejezetben bemutatom a statisztikus multiplexálás és az alkalmazott költségmodellek használatával kapott eredményeket, továbbá összehasonlítom a TDM és WDM-PON hálózatokat egy főre jutó beruházási költségeik alapján.

II. FORGALMI MODELLEZÉS ÉS STATISZTIKUS MULTIPLEXÁLÁS

A hálózati szolgáltatók érdekelték a költségeik csökkentésében, így statisztikus multiplexálást is alkalmaznak a hálózataikban. Ennek köszönhetően az összegzett hálózati forgalmak nem egyenlők a maximális sáv szélességek összegével, mivel nem minden előfizető használja egyszerre a hálózatot csomagok továbbítására. A nagy átviteli sebességet nyújtó optikai hálózatok elterjedése és az [1],[2] és [3] tanulmányokban bemutatott internethasználati szokások megváltozása miatt új forgalmi modellt használtam. Az következő jelenségek az [1],[2] és [3] tanulmányokban találhatók:

- A) Az egyes felhasználók által generált forgalom inhomogén, azaz általában az előfizetők kis része generálja a teljes forgalom jelentős részét. Ez igazolja a felhasználók Pareto-modell szerinti viselkedését.
- B) A felhasználók ritkán használják ki a szolgáltatótól megvásárolt sáv szélességüket főként a TCP protokoll korlátai miatt.
- C) Az egy felhasználóra jutó sáv szélesség átlaga havi szinten nagyon alacsony.

Az általam használt modellben a felhasználók Pareto-eloszlás szerint használják ki a nekik szánt sáv szélességet. Ennek a modellnek a létjogosultságát az 1. ábra alátámasztja, ahol a felhasználói sáv szélesség értékek komplementer összesített eloszlás függvénye került ábrázolásra. Az összegzéshez használt időablak 10ms és 100ms nagyságú volt. A BME mérési adatokat mutatja az 1. ábra, ami a Budapesti Műszaki és Gazdaságtudományi Egyetem hálózatán mért értékekből készült. Továbbá a CAIDA [4] mérési adatokkal (egy nagysebességű internet gerinchálózati kapcsolaton a CAIDA szervezet által mért értékek) összehasonlítva is hasonló eredmény adódott. Annak ellenére, hogy ezek a mérések nem hozzáférési hálózatokból származnak, az eredmények biztatóak a sáv szélesség kihasználtság Pareto-eloszlással történő modellezésével kapcsolatban.



1. ábra: BME mérési adatok és a Pareto-eloszlás

A felhasznált modell, a Pareto-eloszlás módosított változata (1), ahol S_m jelöli az eloszlás által felvehető maximális értéket. Ez az érték jelképezi a maximálisan megengedett sávszélességet az előfizetőknek, amelyet nem képesek átlépni. Ez a módosítás lehetővé teszi $\alpha < 1$ meredekség használatát. A módosított Pareto-eloszlás várható értéke és szórása, ebben a sorrendben, a (3) és (4) alapján számolható. Alkalmazva a centrális határeloszlás tételét egymástól független és azonos eloszlású változók összegzésére, (2) alapján kiszámíthatjuk a szükséges aggregált sávszélességet N felhasználónak a megadott ε elérhetőség mellett (ha N nagy).

$$\hat{F}_x(x, S_m, X_m, \alpha) = \begin{cases} 0 & \text{if } x \leq X_m \\ 1 - \left(\frac{X_m}{x}\right)^\alpha & \text{if } X_m < x < S_m \\ 1 & \text{if } x \geq S_m \end{cases} \quad (1)$$

$$C_R = N \cdot M + \sqrt{2 \cdot N} \cdot \sigma \cdot \text{Erfc}_{inv}(2 \cdot \varepsilon) \quad (2)$$

$$M = \int_{X_m}^{S_m} f_x(x) \cdot x \cdot dx + (1 - \hat{F}_x(x, S_m, X_m, \alpha)) \cdot S_m \quad (3)$$

$$= \frac{\alpha}{\alpha - 1} \cdot X_m - \frac{\alpha}{\alpha - 1} \cdot \frac{X_m^\alpha}{S_m^{\alpha-1}} + \left(\frac{X_m}{S_m}\right)^\alpha \cdot S_m$$

$$\sigma^2 = \int_{X_m}^{S_m} f_x(x) \cdot x^2 \cdot dx + (1 - \hat{F}_x(x, S_m, X_m, \alpha)) \cdot S_m^2 - M^2 \quad (4)$$

$$= \frac{\alpha}{\alpha - 2} \cdot X_m^2 - \frac{\alpha}{\alpha - 2} \cdot \frac{X_m^\alpha}{S_m^{\alpha-2}} + \left(\frac{X_m}{S_m}\right)^\alpha \cdot S_m^2 - M^2$$

III. KÖLTSÉGMODELLEK

A munkám során meg kellett becsülnöm az optikai hozzáférési hálózatok beruházási költségét és előrejelzést kellett adnom a ma még nem elérhető TDM és WDM-PON

rendszerek bővített kapacitású változatainak beruházási költségére is. Így költségmodellek készítésére volt szükségem. Ehhez megvizsgáltam az említett rendszerek felépítését, a szükséges alkatrészek fizikai felépítését és működési elvét [5]-[7]. Ezek után elkészítettem a központi vezérlőegység (OLT - Optical Line Termination), a felhasználói végberendezés (ONU - Optical Network Unit) és az elosztó berendezés (teljesítmény osztó vagy hullámhossz alapú multiplexer/demultiplexer) költségmodelljeit. Az optikai szálak költsége a TDM és WDM-PON rendszerek összehasonlítása során figyelmen kívül hagyható feltételezve, hogy ugyanaz a kábelhálózat mindkét megoldást képes kiszolgálni, így a kábelhálózat nem képezi az alkalmazott költségmodell részét.

A továbbiakban néhány jelölést vezetek be a leírás tömörsége és áttekinthetősége érdekében:

- $totalbw$: az osztott közeg sávszélessége,
- $userbw$: a felhasználónkénti sávszélesség,
- $users$: az osztott közeget használó felhasználók száma,
- $lambdas$: a rendszerben alkalmazott hullámhosszak száma.

A. TDM költségfüggvények

A TDM-PON hálózati eszközök fizikai tulajdonságainak tanulmányozása után azt tapasztaltam, hogy a hálózati eszközök ára az egy felhasználóra jutó sávszélesség értékétől és a csatlakoztatott felhasználók számától függ. Egy időosztásos hálózatban minden optikai adó- és vevőegységnek támogatnia kell az osztott közeg teljes sávszélességén a kommunikációt, különben a felhasználói végberendezések nem képesek kommunikálni a szolgáltatói központi egységgel. A teljes TDM-PON beruházási költséget a következőkben bemutatott eszközök költségösszege adja. A tapasztalatok alapján a hálózati építőelem által kezelt sávszélesség közelítőleg gyökös összefüggésben áll az építőelem árával. Így, az OLT költségfüggvénye a következő:

$$\text{COST}_{\text{TDM}}^{\text{OLT}} = k_{\text{TDM}}^{\text{OLT}} \times \sqrt{totalbw}. \quad (5)$$

Hasonlóan az OLT-hez, az ONU-k ára is a teljes osztott közeg kapacitásától függ. Továbbá, mivel minden egyes előfizetőnek szüksége van egy ONU-ra, így a felhasználók számával lineárisan nő a költség:

$$\text{COST}_{\text{TDM}}^{\text{ONU}} = k_{\text{TDM}}^{\text{ONU}} \times users \times \sqrt{totalbw}. \quad (6)$$

A teljesítményosztót megvizsgálva azt a következtetést vonhatjuk le, hogy alapvetően minden ilyen jellegű osztó 1:2 osztású építőelemekből épül fel. Tehát az ára lineárisan függ a rá köthető felhasználók, vagyis a kimenetei számától, mivel egy 1:2N arányú osztó feltételezhetően 2-szer annyiban kerül, mint egy 1:N osztásarányúval rendelkező. Ezt a belsejében található építőelemek száma indokolja. Így az osztó költségét a következő egyenlet adja meg:

$$\text{COST}_{\text{TDM}}^{\text{SPLITTER}} = k_{\text{TDM}}^{\text{SPLITTER}} \times users. \quad (7)$$

B. WDM költségfüggvény

A WDM-PON rendszerek költségmodellje a TDM-PON rendszerek költségfüggvényeihez hasonlóan lett kialakítva. Mivel az OLT minden hullámhosszt feldolgoz, így a használt

hullámhosszak számától is függ az ára. A hullámhosszak számától való függés hasonlóan az átviteli kapacitáshoz, gyökös összefüggésre adódott. A WDM költségfüggvényt a következő egyenlet adja meg:

$$COST_{WDM}^{OLT} = k_{WDM}^{OLT} \times \sqrt{userbw} \times \sqrt{\lambda_{mbdas}}. \quad (8)$$

A WDM-PON hálózatban teljesítményosztó helyett hullámhossz szelektív elosztóegység, AWG (Arrayed Waveguide Grating) található a hullámhosszak szétosztására. Ennek az eszköznek az ára négyzetes összefüggésben áll az optikai csatornák számával, hiszen minden újabb hullámhossz implementálásával több fizikai építőelemre van szükségünk az AWG-ben. Mindemellett precízebben kell megvalósítani a gyártónak az eszközt, mivel a kommunikációra használt optikai sáv sávszélessége rögzített és ezáltal egy újabb hullámhossz bevezetésével csökken a hullámhosszak közti védőtávolság. A csatornaszámok növekedésével így a belső építőelemek minőségileg és mennyiségileg is változnak, így indokolt a négyzetes összefüggés. Az AWG költsége:

$$COST_{WDM}^{AWG} = k_{WDM}^{AWG} \times users^2. \quad (9)$$

A WDM ONU költsége hasonlóra adódik, mint a TDM ONU-é, azonban csak az egy felhasználóra jutó sávszélességgel kell számolnunk, mivel az adó-vevő nem az osztott közeg sávszélességével tartja a kapcsolatot az OLT-vel. A hozzá tartozó költségfüggvény:

$$COST_{WDM}^{ONU} = k_{WDM}^{ONU} \times users \times \sqrt{userbw}. \quad (10)$$

IV. TDM vs. WDM

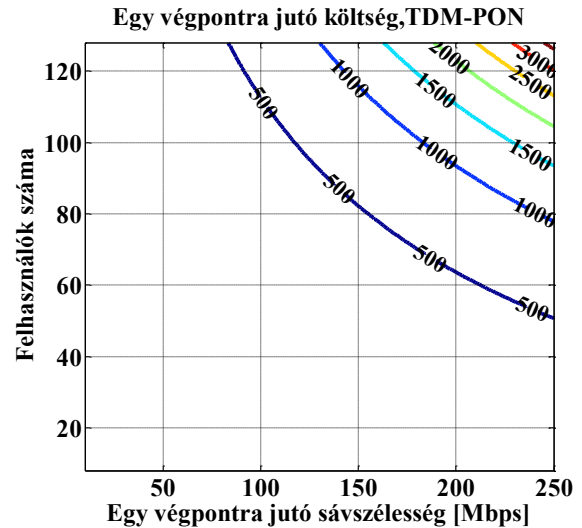
A költségmodellek megalkotása után a megfelelő paraméterértékek megválasztására volt szükség, amire az ipari partnertől kapott példaadatok alapján került sor. A súlyozó tényező értékeit a modellből számolt és a kapott költségek közti négyzetes hiba minimalizálásával határoztam meg.

A. Költségfüggvények

Az elkészített TDM-PON költségfüggvényét mutatja a 2. ábra szintvonalakkal. Az ábra az egy előfizetőre jutó költségeket mutatja felhasználószám és felhasználóknak szánt sávszélesség függvényében. A szintvonalak a megegyező költségű pontokat kötik össze, és a hozzájuk rendelt számok az egy felhasználóra jutó költségeket jelölik. Látható, hogy a növekvő felhasználószám és nagy felhasználónkénti sávszélesség mellett a költség meredeken növekedni kezd, ami az ONU-k számának és az osztott közeg sávszélességének növekedése miatt várható volt.

A 3. ábra a költségek minimumértékeit és az optimális felhasználószámokat mutatja be, a sávszélességek függvényében. A várakozásainknak megfelelően a sávszélesség növelésével csökken az optimális felhasználószám. Ez hasznos információ egy TDM-PON hálózat beruházási költségeinek tervezésénél. A WDM-PON-ok költségábráját mutatja a 4. ábra. A diagram jól mutatja, hogy hullámhosszosztásos hálózatot építeni nem éri meg kevés felhasználóra a költségek gyors emelkedése miatt, ami egybeesik a gyakorlati tapasztalatokkal, mert ilyen esetekben általában pont-pont hálózatot alakítanak ki, így a hálózat egyszerűsíthető és az

AWG ára megspórolható. Azt is megfigyelhetjük ezen az ábrán, hogy nagyobb felhasználószámok esetén, az egy felhasználóra jutó sávszélesség növelése nem emeli meg drasztikusan a költségeket.



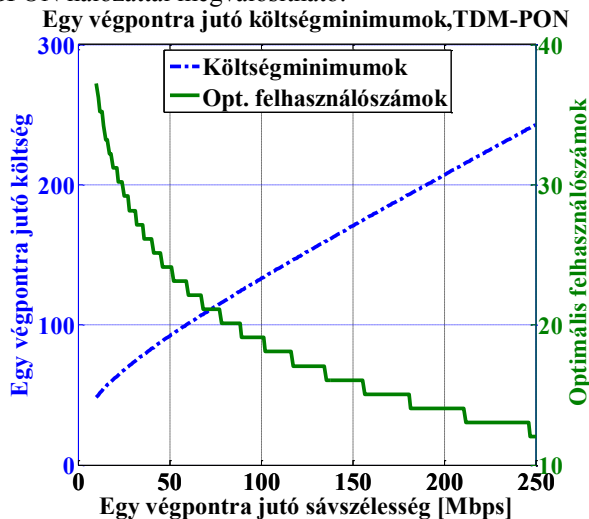
2. ábra: Egy előfizetőre jutó költségek, TDM-PON esetében, statisztikus multiplexálás mellett

B. Technológiák összehasonlítása

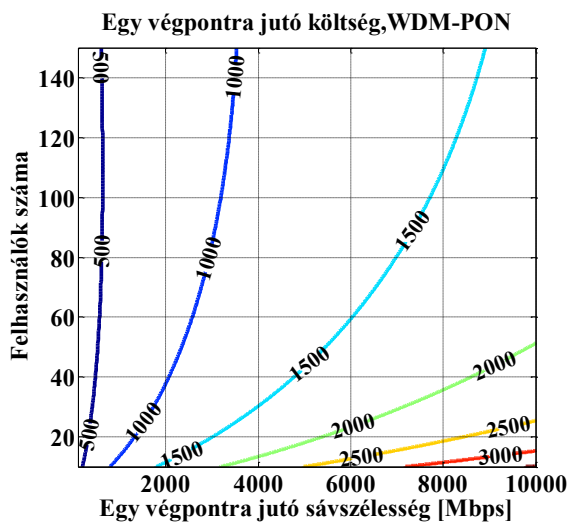
Az alkalmazott és implementált költségmodellek használatával már lehetőségünk van az osztott közeg hozzáférési módjában különböző optikai hálózatok összehasonlítására. Dolgozatomban az idő- és hullámhosszosztásos hálózatokat hasonlítottam össze egymással. A különféle TDM és WDM-PON rendszereket elsődlegesen az egy felhasználóra eső fajlagos költségek révén hasonlítjuk össze: ez talán a legfontosabb mérőszám, amely egy szolgáltató számára a technológiai döntést meghatározhatja. Az 5. ábra ezt az összehasonlítást segíti: a vízszintes tengelyen a felhasználó számára nyújtott szolgáltatás elsődleges mércéjeként a garantált sávszélességet mutatja. A függőleges tengely, illetve a TDM és WDM görbék egyes pontjai pedig azt a minimális költséget jelölik, amely az egy végpontra eső adott sávszélesség mellett elérhető optimális paraméter-beállítás, azaz az osztott közeghez tartozó sávszélesség és felhasználószám mellett.

Vegyük észre, hogy a két görbe metszi egymást, nagyjából 600 Mb/s egy felhasználóra jutó sávszélesség értéknél. Ez azt jelenti, hogy a TDM-PON hálózati technológiákat a sávszélesség határig érdemes használni, onnantól a hullámhosszosztásos rendszerek jelentenek jóval kedvezőbb megoldást. A kapott eredmény egybeesik a várakozásaimmal, mivel a TDM-PON hálózatok esetén az osztott közeg sávszélességét jelentősen növeli az előfizetőkénti sávszélesség növekedése, amely nagysebességű és drága adóvevőket igényel az OLT-ben és az ONU-ban is. A WDM-PON technológia esetében ezzel szemben az ONU adóvevő sebességének csak a felhasználónak szánt sávszélességet kell támogatnia, így várható, hogy a sávszélesség növelésével egy pont után már nem éri meg TDM-PON technológiát telepíteni. Egy nemrég véget ért EU kutatási projekt [8] eredményeihez

hasnoln az ily kapott vltási pont 600 Mb/s-os rtke a felhasználnknknti sávszélességre teljesen reális, célszerűen 10GPON hálózattal megvalósítható.



3. ábra: TDM-PON minimális felhasználóra jutó költség metszet és optimális felhasználószámok

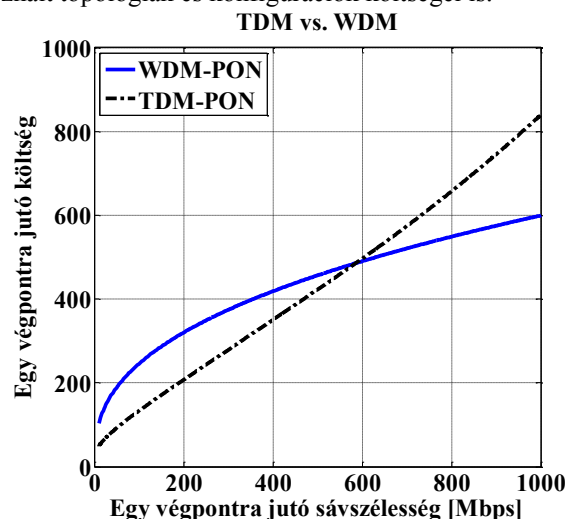


4. ábra: Egy előfizetőre jutó költségek, WDM-PON esetében

V. ÖSSZEFOGLALÁS

A munkám során a cél egy olyan módszer kidolgozása volt, amely alkalmas optimális PON konfigurációk alapján származtatott fajlagos költség alapú összehasonlításra. Ennek első lépéseként megvizsgáltam a napjainkban is elérhető optikai hozzáférési hálózati technológiákat, illetve a jövőben várhatóan elterjedő architektúrákat. A dolgozatban bemutatam egy hatékony módszert a hálózati végpontok forgalmának multiplexálásával elérhető megtakarítás becslésére, melynek révén nagy mértékben javítható a hálózat kihasználtsága, és csökkenthetők annak költségei. A továbbiakban bemutatásra kerültek a költség-alapú összehasonlításához nélkülözhetetlen költségmodellek. Ezek kialakításakor fontos szempont volt, hogy segítségükkel

egyaránt jól becsülhetőek legyenek a jelenleg is elérhető hálózatépítéshez szükséges eszközök árai és a ma még nem használt topológiák és konfigurációk költségei is.



5. ábra: TDM és WDM-PON összehasonlítása

Az implementáció során meghatároztam az egy felhasználóra eső költségeket a felhasználószám és a sávszélesség függvényében TDM-PON és WDM-PON hálózatok esetére is. Ez alapján megkaptam az összehasonlításához szükséges minimális felhasználónkénti költségeket mindkét rendszerre és képes voltam az összehasonlításuk elvégzésére. A jelenlegi technológiai korlátok mellett 600 Mb/s-ra adódott az a sávszélesség, amelytől kezdve érdemes a TDM-PON hálózatok használatáról WDM-PON hálózatok használatára áttérni.

IRODALOMJEGYZÉK

- [1] G. Maier, et al.: "On dominant characteristics of residential broadband internet traffic", in Proc. of the 9th ACM SIGCOMM conference on Internet measurement (IMC '09), Chicago, Illinois, USA, Nov. 2009.
- [2] S. Sundaresan, et al.: Broadband internet performance: a view from the gateway, in Proc. SIGCOMM '11, Toronto, Ontario, Canada, Aug. 2011.
- [3] P. Megyesi, S. Molnár: "Analysis of elephant users in broadband network traffic", in *Advances in Communication Networking*. Lecture Notes in Computer Science, Springer, 2013, vol. 8115, pp. 37–45.
- [4] The Cooperative Association for Internet Data Analysis, <http://www.caida.org>
- [5] OASE Project, Survey of Next-Generation Optical Access System Concepts, 29.10. 2010.
- [6] Vetter, Peter. "Next generation optical access technologies." *European Conference and Exhibition on Optical Communication*. Optical Society of America, 2012.
- [7] Y. Luo, et al.: "Time and wavelength division multiplexed passive optical network (TWDM-PON) for next generation PON stage 2 (NG-PON2)," *IEEE J. of Lightwave Technology*, Vol. 31, No. 4, February 15, 2013.
- [8] Ralf Hülsermann, Klaus Grobe, Dirk Breuer: "Results from EU FP7 Project OASE on Next-Generation Optical Access", in *Photonic Networks, 14. 2013 ITG Symposium. Proceedings*, May 2013.

Obszervatóriumi műszer fejlesztése és megvalósítása tellurikus áramok méréséhez

Gorócz Vilmos

Budapesti Műszaki és Gazdaságtudományi Egyetem
Szélessávú Hírközlés és Villamosságtan Tanszék
BME-HVT
Magyarország, Budapest
gorocz@mht.bme.hu

Csurgai-Horváth László

Budapesti Műszaki és Gazdaságtudományi Egyetem
Szélessávú Hírközlés és Villamosságtan Tanszék
BME-HVT
Magyarország, Budapest
csurgai@mht.bme.hu

Absztrakt— A BME-HVT Űrkutató Csoportja és az MTA Geodéziai és Geofizikai Intézet (MTA-GGI) együttműködésével 2011-ben megkezdődött egy tellurikus áramok változásait regisztrálni képes műszer fejlesztése. A tellurikus áramok változásait a Széchenyi István Geofizikai Obszervatóriumban már alapításától (1957) folyamatosan regisztrálják. Az obszervatóriumban két, egymástól 500 méterre a talajban elhelyezett ólom elektródapár segítségével mérik a tellurikus áramok által létrehozott elektromos feszültséget. A korszerűsítés igénye azért merült fel, hogy a mérés frekvenciatartományát kiterjesszük, illetve jobb minőségű regisztrálást valósítsunk meg. A cikk részletesen bemutatja a fejlesztés fontosabb lépéseit és tervezési megfontolásait beleértve az elektronikai, konstrukciós és szoftvertervezés legfontosabb állomásait is.

Kulcsszavak; tellurikus áramok, műszerfejlesztés, analóg és digitális szűrők

I. BEVEZETÉS

A föld felszíne alatt, a talajban folyó áramokat tellurikus áramoknak nevezzük. Ha a föld belsejében változó mágneses tér van jelen, akkor ez térerősséget indukál és a talaj véges vezetőképessége miatt – a differenciális ohm törvény értelmében – áram folyik a talajban.

A tellurikus áramoknak két nagy forrását különböztetjük meg. Léteznek terresztikus források, amikor a tellurikus áramok forrása Föld légkörében található, mint például a zivatar jelenségek és a villámok. Az extraterresztikus források leginkább az alacsonyabb frekvenciákon jelennek meg, így ezek periódusideje több órára is lehet. Az általuk generált tellurikus áramok szintje a mérsékelt övezetben az 1-20mV/km tartományba esik. A terresztikus források ezzel szemben magasabb frekvenciákon (1Hz felett) jelennek meg [1].

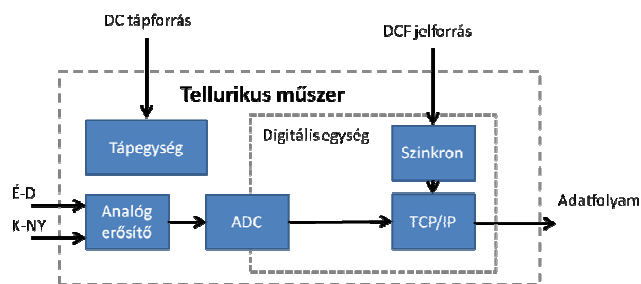
A tellurikus áramok mérésére földbe ásvó ólom elektródák szolgálnak. Az elektródák közötti feszültség arányos a tellurikus áramokkal, ha a talaj vezetőképessége állandó. Ezt jogosan feltételezhetjük, mivel az elektródák mélységében ez a paraméter csak igen lassan változik. Összesen 4 elektróda van elhelyezve az égtájak irányában, így valósul meg az észak-déli, illetve a kelet-nyugati irányú áramok regisztrálása. Az elektródák a Föld felszíne alatt 2 méterre helyezkednek el. Ez a mélység már elég nagy ahhoz, hogy a talaj vezetőképességét ne befolyásolják gyors időjárási jelenségek (például egy heves

zápor). Az elektródák egymástól 500 méterre vannak elhelyezve, tehát 1mV/m térerősség 500μV-nak mérhető, ha feltesszük, hogy a tellurikus áramok által gerjesztett térerősség 500 méteren belül nem változik jelentősen.

II. ÚJ MÉRŐMŰSZER ÖTLETE

A fentebb tárgyalt, tellurikus jelenségek méréséhez szükség volt egy olyan műszerre, amely képes megfelelő minőségben regisztrálni a jelet. A műszer tulajdonképpen nem áramot mér, hanem az elektródák közötti potenciálkülönbséget, melyből térerősséget, illetve áramot lehet számolni az adatfeldolgozás során. A műszer megvalósítására a következő paramétereket tűztük ki célul:

- Frekvenciasáv: DC-100Hz
- Csatornák száma: 2 db, azonos paraméterekkel (észak-dél, kelet-nyugat)
- Minimális amplitúdó: 5μV (10μV/km)
- Maximális amplitúdó: 5V (10V/km)
- Az új műszer a jelenleg üzemelővel párhuzamosan is működtethető legyen.
- A műszernek szinkronizálnia kell a mérési adatokat a DCF77 szabványú időjel segítségével.
- Maximális közös módusú jel: 18V_{pp} (DC-100Hz)



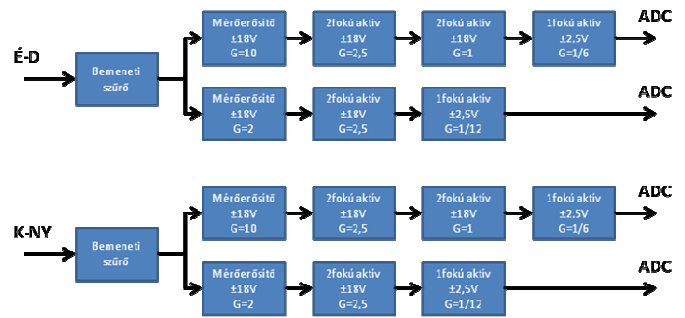
I. ábra A tellurikus műszer rendszerterve

Az elektródákról (É-D Észak-dél; és K-NY Kelet-Nyugat) érkező jelet egy analóg egység kondicionálja, szűri és illeszti az AD (Analog to Digital) konverterhez. Az analóg szűrést úgy kell megtervezni, hogy az AD konverzió során ne történjen átlapolódás, illetve ügyelni kell a lineáris fázismentesítésére, mivel a jel időtartománybeli viselkedése is fontos információkat hordoz. A specifikációban mérendő

frekvenciatartományként a DC-100Hz szerepel, tehát a mintavételi tétel értelmében legalább 200Hz mintavételezési frekvenciát kell alkalmazni. Azért, hogy az analóg szűrőket könnyebben meg tudjuk valósítani, az úgynevezett túlmintavételezési módszert alkalmazzuk, amelynek lényege, hogy a jelet jóval nagyobb mintavételezési sebességgel mérjük, mint a fent említett elméleti, minimális 200Hz frekvencia, majd ezután decimálást hajtunk végre. A mintavételezési frekvenciát 3,6kHz-re választottuk, így az analóg erősítőnek 3,5kHz (=3,6kHz-100Hz) frekvencia felett kell elnyomni a nem kívánt frekvencia összetevőket. A mérendő legkisebb jelszint viszonylag kicsi ($5\mu V_p$), így a tervezés során biztosítani kell a jó minőségű regisztrálást, hogy a zaj ne befolyásolja túlságosan nagy mértékben a mérési adatokat. A legnagyobb problémát a mérési frekvenciatartományban megjelenő differenciális zajok okozhatják, amiket akár a földben folyó mesterségesen keltett áramok is okozhatnak, főként galvanikusan csatolva. Ezen zajokhoz elvileg sem lehet olyan műszert készíteni, amely képes lenne ezeket szűrni, ezért ebben az esetben a fizikusoknak kell eldönteni az adatok kiértékelése során, hogy mi tekinthető elméletekkel alátámasztható jelenségnek a mérési adatsorok közül. Tekintve, hogy a mérőrendszert hosszú jelvezetékek alkotják (500m-es elektródátávolságok), ezért könnyen megjelenhetnek közös módusú jelek az elektródák vezetékén inductív vagy kapacitív csatolással, elsősorban az energiahálózatról, de problémát okozhatnak például a középhullámú adók is, ahol inkább sugárzás útján csatolódhat a jel a mérőrendszerre. Ezen zajok ellen már fel lehet készíteni a műszert és - ahogyan a specifikációban is olvasható - $18V_{pp}$ nagyságú közös módusú jelszintig hibamentesen folyik a regisztráció. A közös módusú zajok a mérőrendszer megfelelő kialakításával is csökkenthetőek. Beszélhetünk még a műszer által keltett zajokról, de ezek a műszer megfelelő kialakításával minimálisra csökkenthetőek. A műszer a keletkező adatfolyamot Ethernet hálózaton keresztül továbbítja a számítógép felé. A mért adatok időbélyegét egy DCF77-es modul segítségével állítjuk elő, amely a műszeren kívül kap helyet. A szoftver egyik feladata, hogy feldolgozza a DCF (Deutschland Long wave signal from Frankfurt at 77.5kHz frequency) adatfolyamot is. A műszer DC tápellátást igényel, az egyes egységeknek szükséges tápfeszültségeket lineáris tápegységek segítségével állítjuk elő az eszköz dobozán belül. Kapcsolóüzemű tápegység használata nem indokolt, mivel a teljesítmények alacsonyak, illetve a lineáris tápegységek kevesebb zajt termelnek, amely nagy előnyt jelent egy ilyen típusú mérőműszerben.

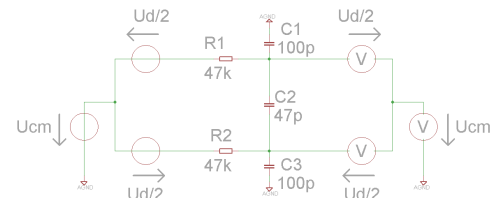
III. ANALÓG EGYSÉG

A mérendő maximális jelszint 5V, így az erősítőnek legalább 120dB dinamikatartománnyal kell rendelkeznie, ha $5\mu V$ minimális amplitúdójú jeleket is szeretnénk regisztrálni. Ez elvileg mérhető egy 24 bites AD konverterrel, azonban a kis jelszintek várhatóan zajszint alá kerülnének, valamint nem is lehetne olyan jó minőségben regisztrálni azokat, mivel ezek esetében kisebb a kvantálási tartomány is. Ezért a mérendő feszültségtartomány regisztrálását 2 független erősítőlánccal valósítottuk meg (2. ábra).



2. ábra Az analóg egység blokkvázlata

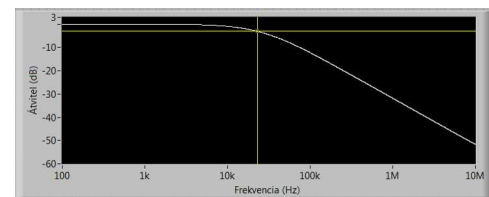
A bemeneti szűrő előtt alkalmazunk túlfeszültség védelmet szikraköz (surge arrester), illetve zener dióda segítségével. Ezek a védelmek elsősorban a villámcsapás esetén bekövetkező túlfeszültség fellépését igyekeznek meggátolni a műszer bemenetén. Bár ezt a funkciót nehéz tesztelni, de biztosan nő a műszer túlfeszültségtűrő képessége, hiszen már az is jelentős feszültségnövekedést eredményezhet, ha a mérőrendszer környezetét éri a villám. A bemeneti elsőfokú passzív szűrő (4. ábra) egy RC hálózat, melynek feladata a nagyfrekvenciás komponensek elnyomása. A műszer mérési frekvenciatartományában erősítése 1, fázistolása elhanyagolható.



3. ábra Bemeneti szűrő elvi rajza

A szűrő közös módusú átvitele könnyen meghatározható, mivel ez esetben a kapcsolás egy egyszerű, elsőfokú RC aluláteresztő tagra egyszerűsödik, ahol $R=R_1 \times R_2$ és $C=C_1+C_3$, így a határfrekvencia 33,86kHz-re adódik. A differenciál módusú átvitel ennél lényegesen bonyolultabb $R=R_1=R_2$ és $C_1=C_3$ mellett az (1) szerint áll elő.

$$H_D(s) = \frac{sR(C_1 + C_2) + 1}{s^2R^2(2C_1C_2 + C_1^2) + sR(C_1 + C_2) + 1} \quad (1)$$



4. ábra Bemeneti szűrő differenciál módusú átvitele

A szimuláció alapján a határfrekvencia közelítőleg 23kHz. Az így meghatározott törésponti frekvenciák a valóságban kisebbre adódnak, mert a bemeneti szűrőt 2 mérőerősítő terheli párhuzamosan. A mérőerősítők (INA115 típusú) 6pF bemeneti kapacitással rendelkeznek közös és differenciál módusban egyaránt, tehát a törésponti frekvenciák várhatóan 10-20%-kal kisebbek lesznek, de ez nem okoz problémát, mivel a szűrővel

szemben csak az a követelmény, hogy 100Hz-en ne okozzon jelentős amplitúdó változást, illetve fázistolást, valamint nyomja el a nagyfrekvenciás (MHz feletti) komponenseket.

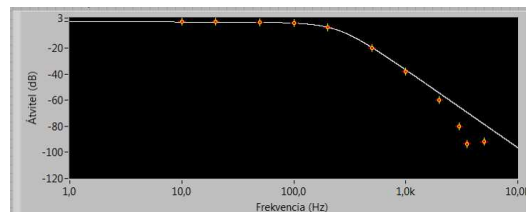
A bemeneti szűrő után egy mérőerősítő fogadja a differenciális jelet, amely jó közös módusú elnyomást is biztosít. A közös módusú elnyomást a mérőerősítő ellenállásainak a szimmetriája és az erősítés nagysága határozza meg. Az erősítő ellenállásainak a szimmetriáját az által maximalizáltam, hogy egy olyan áramkört választottam (INA115), amely integrált, lézerrel beállított ellenállásokat tartalmaz, így tipikusan 96dB elnyomást biztosít. Ha a mérőerősítőben 1-nél nagyobb erősítést alkalmazunk, akkor ez azt eredményezi, hogy a differenciál módusú jel szintje nő, míg a közös módusú állandó értéken marad. Ez segít az elnyomás növelésében, de ez a növekedés kb. 10-es erősítés után telítődik. Túl nagy erősítést pedig nem célszerű választani, mert akkor az erősítő könnyen telítésbe mehet a differenciál módusú jeltől, a tápfeszültséget pedig nem lehet korlátlanul növelni. Azért, hogy a kis és a nagy jelszintek egyaránt mérhetőek legyenek, mindkét csatornán alkalmaztunk egy kisebb és egy nagyobb erősítéssel rendelkező fokozatot. Az egyes csatornák közös módusú elnyomását az alábbi táblázat tartalmazza. A mérést úgy végeztem, hogy 50Hz-hez közeli frekvenciájú közös módusú szinusz jelet kapcsoltam az áramkörre és a műszer adatfeldolgozó szoftverének segítségével előállítottam a spektrumot, melyből ki tudtam számolni a jel amplitúdóját. Azért választottam 50Hz-hez közeli, de eltérő frekvenciát, hogy erre a frekvenciára jellemző közös módusú elnyomást tudjam mérni. Ha pontosan 50Hz-et választok, akkor a mérést nagy mértékben zavarhatják az egyéb 50Hz-es berendezések. A mérés folyamán figyelembe vettem a zajt is, melyet úgy küszöböltem ki, hogy a mérési frekvencia közelében megvizsgáltam, hogy mekkora átlagos zaj terheli a mérést és ezt a zajszintet feltételeztem a mérési frekvencián is, ezzel korrigáltam a mérési eredményt. A kisebb erősítéssel rendelkező csatornák (NSN és EWN) lehetővé teszik a nagyobb jelszintek (>1V/km) mérését, azonban ezek nagyon ritka jelenségek, melyeket a fizikai célok miatt gyakorlatilag lehetetlen előre jelezni, ezért nem tartottuk célszerűnek egy AGC áramkör használatát. A további fokozatok erősítését úgy választottuk meg, hogy a mérhető tartományt leképezze az AD konverter mérési tartományára, amely $\pm 2,5V$.

CMRR	dB
NSN	109,74
NSG	115,74
EWN	108,23
EWG	112,46

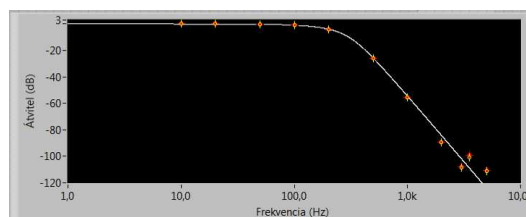
1. táblázat Közös módusú elnyomás

A mérőerősítőt az antialiasing szűrők követik, melyeket aktív első-, illetve másodfokú szűrőtagokkal állítunk elő. Az antialiasing szűrők követelménye, hogy 100Hz alatt kicsi, 3,5kHz felett pedig nagy elnyomással rendelkezzenek, illetve a jel tartományában (100Hz alatt) lineáris fázismenetük legyen. Ez utóbbira azért van szükség, mert a tellurikus áramok változásainak regisztrálásakor az időtartománybeli jelenségek is fontosak és egy nemlineáris fáziskarakterisztikával rendelkező szűrő – a diszperzió miatt – eltorzítaná ezeket a jelenségeket, ezért választottuk a Bessel típusú szűrőt [2]. A sávon kívüli elnyomás mértékét az határozza meg, hogy

mekkora jelszinteket kívánunk mérni, hiszen elméletileg átlapolódás mindig fellép a mérés során, csak azt kell biztosítani, hogy az átlapolódott jelek oly mértékben csillapodjanak, hogy ne zavarják a mérendő jelet. Ha szűrőként 5. fokú Bessel szűrőt alkalmazunk, akkor a szimulációs eredményekből látható (6. ábra fehér görbe), hogy 3,5kHz-en 120dB elnyomást érhetünk el, amely biztosítja, hogy a mérési tartományon belüli nagyfrekvenciás jelek szintje a legkisebb mérendő jelszint alá csillapodik. Az 5. és 3. fokú Bessel szűrő képleteiből kiindulva megterveztük, szimuláltuk és megmértük a szűrő karakterisztikáit, amely az 5-6. ábrákon látható. A mérést a teljes rendszeren végeztük, de a karakterisztikát (ebben a tartományban) csak az antialiasing szűrők határozzák meg. Mivel a mintavételi frekvencia 3,6kHz, ezért az 1,8kHz-nél nagyobb mérési pontokat úgy vettük fel, hogy figyelembe vettük az átlapolódás jelenségét, vagyis a 2kHz-es jel 200Hz-en jelent meg a spektrumban és ebből számoltunk jelszintet. A mérést diszkrét frekvenciákon (vörös és sárga pontok) végeztük. Jól látható, hogy a vörös és sárga pontok precízen együtt futnak, ami jól jellemzi az erősítőfokozatok egyformaságát, illetve azt, hogy a törésponti frekvencia megválasztása is jól sikerült.



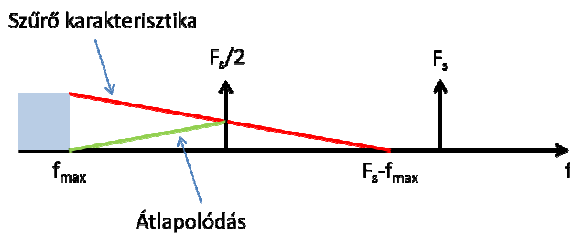
5. ábra 3. fokú Bessel szűrő szimulált (fehér) és mért (piros, sárga) karakterisztikája



6. ábra 5. fokú Bessel szűrő szimulált (fehér) és mért (piros, sárga) karakterisztikája

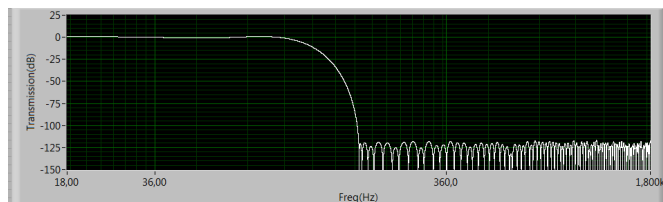
IV. MINTAVÉTELEZÉS, SZÜRÉS

Az analóg jelformálás után A/D konverziót alkalmazunk egy 24 bites $\Delta-\Sigma$ (Delta-Sigma) típusú AD konverter segítségével (ADS1255). A fentiekben már szerepelt, hogy a mérés során a túlmintavételezési eljárást alkalmaztuk. A mintavételi frekvenciát úgy választottuk meg, hogy ne kelljen túlságosan nagy fókuszú analóg szűrőket alkalmazni, de az adatsebesség se legyen túl nagy, mivel ez megnehezíti az adatok feldolgozását. A túlmintavételezés után a spektrum a 7. ábra szerint alakul. A hasznos jel $f_{\max}$ -ig terjed. Ha analóg erősítők segítségével $F_s-f_{\max}$ felett sikerül megfelelő mértékben elnyomni a jelet, akkor az F_s mintavételezési frekvencia hatására létrejövő átlapolódás már nem zavarja az $f_{\max}$ frekvenciánál kisebb összetevőket.



7. ábra Frekvencia spektrum

Az AD konverzió során jóval nagyobb adatsebesség keletkezik, ezért szükség van decimálásra, amely átlapolódás mentesen lecsökkenti az adatsebességet. Ez a funkció a műszeren kívül, egy számítógépből kap helyet. A jelet egy FIR szűrőn vezetjük át, majd a szűrés után minden 10. mintát el kell dobni, hogy a decimálás létrejöhessen. Azonban az eldobandó mintákat nem célszerű kiszámolni, ezért a FIR szűrést úgy hajtjuk végre, hogy minden 10. minta beérkezése után számolunk ki egy kimeneti értéket. Az így alkalmazott szűrő 180Hz felett minden összetevőt elnyom, így valóban átlapolódás mentesen fog megtörténni a decimálás [3].



8. ábra FIR szűrő átviteli karakterisztikája

Az AD konverter adatlapjában megtalálhatók a konverzió minőségét jelző ENOB értékek a mintavételi frekvencia függvényében. Az adatlap szerint 3,6kHz környékén 18 bitre lehet számítani, azonban a megválasztott mintavételi frekvencia ellenére nem ezt az értéket, hanem a 360Hz-hez tartozó értéket kell figyelembe venni, ami jelen esetben 20 bit körül van. Ennek oka, hogy az AD konverter egy delta-sigma típusú konverter, ezért a konverzió minőségét alapvetően a konverter analóg áramköreinek minősége és a konverter belső órajele határozza meg. Az alacsonyabb adatsebességek felé az ENOB értékek azért növekszenek, mert a konverter digitális áramkörei egyszerű átlagolást hajtanak végre, melynek hatására a zérus várható értékű, gauss eloszlású zaj teljesítménye csökken. Azonban, ezt a digitális átlagolást részben az adatfeldolgozó szoftverben levő decimáló FIR szűrő végzi, így a konverzió minőségét a 360Hz-hez rendelhető ENOB érték határozza meg.

A műszer zaját mérés is ellenőriztem olyan módon, hogy 10kΩ ellenállást (feltételezett, átlagos generátor impedancia) kapcsoltam az elektróda bemenetekre és megmértem a zajszintjét V_{pp} értékben. Ezt az értéket leosztottam egy közepes jelszinten mért erősítéssel, így megkaptam, azt a jelszintet, amit ideális esetben a bemenetre kellene kapcsolni ahhoz, hogy ekkora jelszint jelenjen meg a kimeneten. A mérési eredményekből látható, hogy az erősítéssel rendelkező csatornákon sikerült 5μV_p-nél kisebb zajszintet elérni, így ezek a jelkomponensek várhatóan nem vesznek el a zajban. Figyelembe kell venni azt, hogy itt a teljes,

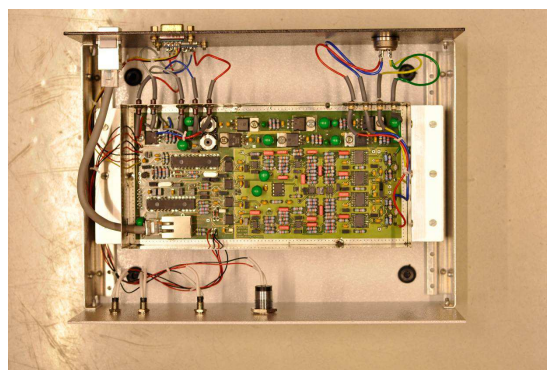
műszer által kezelt sávon történt a mérés, egy szűkebb frekvenciatartományban kisebb zaj várható.

Noise level	μV _{DD}
NSN	29,85
NSG	9,39
EWN	49
EWG	8,43

2. táblázat A műszer zajja

V. MEGVALÓSÍTÁS, TELEPÍTÉS

A műszer elektronikája egy ónozott vasdobozban kapott helyet, ami védi az áramkört a külső mágneses és elektromos zavaroktól. A külső alumíniumdoboz hátlapjára vannak szerelve a csatlakozók, illetve az előlapjára az állapot jelző LED-eket. A műszer telepítésére 2014. március 26-án került sor.



9. ábra Tellurikus műszer belseje összeszerelt állapotban

VI. ÖSSZEFOGLALÁS

A telepítés óta a regisztrálás folyamatosan üzemel. Bízunk benne, hogy a rögzített adatsorok között sok érdekes anomáliát találunk, melyet az MTA-GGI együttműködésével a megfelelő tudományos cikkekben vagy folyóiratokban publikálni tudunk a jövőben.

VII. KÖSZÖNETNYILVÁNÍTÁS

Köszönetünket fejezzük ki az MTA CSFK GGI részéről Dr.Bencze Pál konzulens professzor felé a munkánkhoz nyújtott szakmai segítségért és a téma felvetésért, Dr.Weszergom Viktor GGI és a Széchenyi István Observatórium igazgatójának a munka anyagi támogatásáért és a telephelyen a munkavégzés lehetőségének biztosításáért, Dr. Szabó József tanszéki munkatársunknak a megbeszélések során adott tanácsaiért, iránymutatásáért, valamint a Dr.Bánfalvi Antal vezette Űrtechnológia labor és a tanszéki műhely dolgozóinak, hogy a műszer megvalósulhatott.

VIII. REFERENCIÁK

- [1] Ádám Antal, Kis Árpád, Lemperger István, Novák Attila, Prácsér Ernő, Szarka László, Weszergom Viktor, "Geomágneses indukció, forrásvizsgálatok és az európai litoszférolemez elektromos modellje az indukciós kockázat becsléséhez", Magyar Tudomány 173. évf. (2012.) 12. sz. pp. 1449-1458.
- [2] Bessel Filter http://en.wikipedia.org/wiki/Bessel_filter
- [3] Dr. Gaál József, Híradástechnikai jelfeldolgozás (jegyzet) <http://www.hit.bme.hu/~gaal/hjf.htm>

Optimalizálási algoritmusok szenzorhálózat alapú pozicionálásra naprendszerbeli égitesten

Szeile Aliz

Hálózati Rendszerek és Szolgáltatások Tanszék

BME

Budapest, Magyarország

szeile@mcl.hu

A jövőben elképzelhető, hogy néhány nagyon drága űreszköz helyett nagyszámú olcsóbb mérőeszközt juttatnak el egy-egy bolygó felszínére, amelyek autonóm módon mozogva végeznek méréseket és juttatják el az eredményeiket egy földi irányító központba. Vizsgálataim tárgya, hogy egy szenzorokból álló hálózat mozgásának irányítása hogyan oldható meg a kapcsolat megszakadása nélkül.

Az adatok feldolgozásánál fontos szempont, hogy melyiket hol mértük. Kidolgoztam a több lépcsős helymeghatározási módszert, mely alkalmas a lefedett terület jelentős kiterjesztésére, azonban a fellépő mérési hibák torlódhatnak, így a helyzetbecslés egyre nagyobb eltérést mutathat a valós értékektől. Ez nagyobb mértékű adatvesztéssel járhat, így a vizsgálataim másik fő célja a több lépcsős helymeghatározás hatékonyságának növelése.

A több lépcsős pozicionálásra különböző optimalizáló algoritmusokat valósítottam meg egy C++ nyelven írt programban. A kapott eredmények részletes képet adnak az algoritmusok hatékonyságáról.

Kulcsszavak; szenzorok, szenzorhálózat, pozicionálás, naprendszerbeli égitest, optimalizáló algoritmusok

I. BEVEZETÉS

A szenzorok hatékonyságát eddig már sokféle témakörben vizsgálták és bizonyították. Ezekkel az apró méretű érzékelőkkel képesek vagyunk parkolóknak megállapítani az üres helyek számát, földtani vizsgálatokhoz adatokat gyűjteni, vagy akár biztonsági rendszereket fejleszteni. Jelen dolgozatomban egy a megszokottól eltérő területen vizsgáltam egy általam feltételezett szenzorhálózat működését.

A napjainkban használt űreszközök igen költségesek és bonyolult működtetésűek, így fennáll a lehetőség, hogy ezen eszközök mellett egyszerűbb és olcsóbb módszereket vetnek majd be naprendszerünk megismerésére. A NASA egyik nemrég megjelent cikkében szó volt arról, hogy a közeljövőben lehetséges lehet olyan módszer alkalmazása, miszerint egy naprendszeren kívüli kisbolygót robbantással közelebb hoznak és így elérhető közelségben lesz ahhoz, hogy megismerjük [1]. Ezen tervek megnyitják a szenzorhálózatok alkalmazhatósági területét az űrkutatásban. A dolgozatomban egy naprendszerbeli égitest felszínén mozgó szenzorhálózat

alkalmazását és annak hatékonyságát vizsgálom, ami sok újszerű, érdekes problémát vet fel.

A szenzorhálózat kialakításánál fontos szempont volt a szenzorok teljesítőképessége és mobilitása. Ezek alapján a hálózatomban kétféle szenzor is szerepel, illetve mobilitását tekintve egy mozgó hálózat. A szenzorok a felszínen gurulva igyekeznek eljutni egy célpontig.

Egy naprendszerbeli égitest (pl. Mars) felszínén mozgó szenzorhálózat többféle, számunkra érdekes adatot gyűjthet össze, mint pl. hőmérséklet, talajminta, légköri jellemző. Azonban ahhoz, hogy ezen adatok feldolgozása megfelelő legyen és hiteles képet kapjunk az égitestről, ismernünk kell a mérések pontos helyét, így az adatokat pozíciókhoz tudjuk kötni.

A pozicionálás megvalósításához háromszögeléses algoritmust alkalmaztam. A jelterjedést befolyásoló terepviszonyok miatt kialakul egy mérési hiba. Ez a hiba először még minimális, azonban többlépcsős helymeghatározásnál már akkora méretű is lehet, hogy a kommunikációt akadályozza, és ez adatvesztést eredményezhet. Ennek a hibának a visszaszorítása végett kidolgoztam két változatát az algoritmusnak, és ezeket a változatokat vizsgáltam a továbbiakban különböző szempontokból.

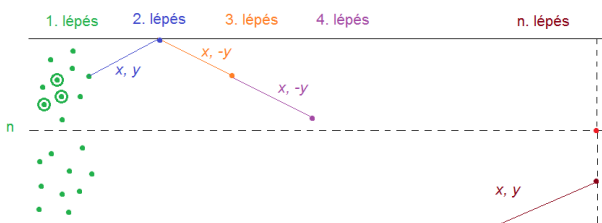
II. A JAVASOLT SENZORHÁLÓZAT FELÉPÍTÉSE ÉS MŰKÖDÉSE

A. Nagyobb teljesítményű szenzorok a hálózatban

A nagyobb adóteljesítményű szenzorokra innentől kezdve szuperszenzorokként fogok hivatkozni. Ha nincsenek szuperszenzorok a hálózatban, akkor a szenzorok energiaellátás szempontjából olcsóbbak, ugyanakkor mindegyik szenzornak képesnek kell lennie kommunikálni a földi irányítással, hogy a mérési adatok eljuthassanak a Földre. Ha vannak a hálózatban szuperszenzorok, akkor ugyan ezek energiaellátása drágább, de csak a szuperszenzornak kell képesnek lennie a nagy teljesítményű kommunikációra, ami jelentős költségcsökkenést eredményez. Ezért úgy döntöttem, hogy a szuperszenzorok lesznek a szenzorhálózat nyelői. Munkámban azzal a feltételezéssel éltem, hogy a szuperszenzorok pontosan ismerik a mindenkori pozíciójukat, így a többlépcsős helymeghatározás első lépésében referenciaként használhatók.

B. Szenzorok mozgása

Két fő csoportra osztottam a mozgás fajtáit. Egyik a véletlenszerű mozgás, ahol előre nem meghatározott pályán mozognak a szuperszenzorok, másik az adott pályán való mozgás. A két mozgásfajtának egy kombinációját vettem. Eszerint a szenzorok kezdetben egy véletlenszerű irányba kezdenek el mozogni és ezen a vonalon haladnak tovább egy előre meghatározott pont felé. Egy általam biztonsági sávnak nevezett sávot is megadtam (a szenzorok y koordinátájára egy felső és egy alsó korlát), amelyen belül a szenzorok mozoghatnak. A biztonsági sáv alkalmazásával nem lesz olyan szenzor, amelyik a sávot eltávolodna, és ezért kiesne a szenzorbolyból. Ha a szenzor elérte ennek a sávnak a határát, akkor irányt vált, az eddigi y koordinátája előjelet vált, így érve el, hogy a szenzorok a biztonsági sávon belül mozogjanak és ne távolodhassanak el túlságosan egymástól. A mozgás illusztrálása az 1. ábrán látható.



1. ábra, A szenzorok mozgása

C. Többugrásos pozicionálás és a háromszögelési módszer

A pozicionáláshoz a háromszögelés módszerét [2] alkalmaztam, mely a javasolt többugrásos helymeghatározás alapjául is szolgál. Ez alapján egy szenzor helyzetének meghatározásához szükség van másik három szenzor helyzetének ismeretére. Így kiválasztottam három szuperszenzort, amelyek helyzete kezdettől fogva ismert a földi irányítás miatt, kezdettől fogva referenciapontoknak számítanak. Ezekből megállapítható azon szenzoroknak a helyzete, amelyek beleesnek a szuperszenzorok hatótávolságaiba, azaz abba a távolságba, amelyen belül még látják egymást és képesek kommunikálni. A kiszámítás matematikai módja a következő alfejezetben írtak alapján történik. A többugrásos helymeghatározás során a kiszámolt helyzetű szenzor is referenciaponttá válik és így az ő segítségével is lehet tovább számolni a többi szenzor koordinátáit. Ha a szenzor helymeghatározásában szerepet játszik egy bizonyos hibaérték, akkor a belőle tovább számított szenzorok helymeghatározásában is szerepelni fog ez a hiba valamilyen súlyozással, azaz a hiba torlódik. A helymeghatározási folyamatot rekurzívan folytatva meghatározhatjuk az egész szenzorboly helyzetét.

III. POZICIOBECSLÉSI ALGORITMUSOK

A. Optimalizálatlan alap algoritmus

A kezdeti algoritmusom esetében még nincs optimalizáló eljárás. A kiszámítandó helyzetű szenzor egy egyszerű, bár kevésbé hatékony módszerrel választja ki a potenciális referencia szenzorok közül a három szülőjét. Az algoritmus abból áll, hogy a bolyban szereplő szenzorokat elkezd

vizsgálni, és ha referencia pont, akkor beveszi a szülők közé. Az első három találat után a vizsgálat megáll és adottak lesznek a használni kívánt referencia szenzorok. Tehát a bolyból az első három referencia szenzort választja ki.

A módszer azért nem hatékony megoldás, mert ha a lehetséges referencia szenzorok közül három olyat választ ki az algoritmus, akik már torlódott hibát hordoznak magukkal, akkor ő hibájukat az adott szenzorunk is hordozni fogja. Tehát a számítása valamilyen mértékben biztosan eltér majd a valós helyzetétől. Ezzel szemben, ha választhatná szülőknek a három alap szuperszenzort, akkor jóval kisebb hiba adódna a kiszámolt értékhez, mint az előbbi esetben, mert a szuperszenzorok pozíciója biztosra ismert. Erre az algoritmusra a továbbiakban BPA (*Basic position algorithm without optimization*) fogok hivatkozni.

B. Számítási hop számot figyelembe vevő algoritmus

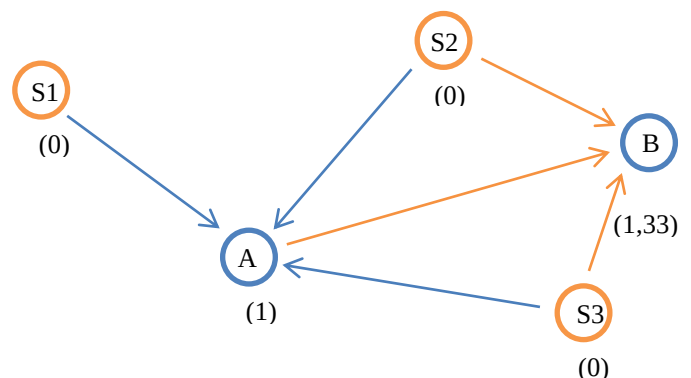
Ez az algoritmus leginkább abban különbözik az előzőtől, hogy minként választ magának egy szenzor szülőket. A szenzorok tárolnak egy olyan hop értéket, ami arra utal, hogy hány lépcsős pozicionálás eredménye a kiszámolt koordinátájuk. Minél későbbi lépcsőben kerül sor egy szenzor pozíciójának becslésére, a hibatorlódás annál jelentősebb. A cél ezért az, hogy olyan szenzorokat jelöljünk ki referenciaként, melyeknek koordinátáit a lehető legkevesebb lépcsőben számíthatók ki. A szuperszenzorok hop száma alaphoz az egész vizsgálat során 0. A szülők hopszáma átlagolódik, és ehhez hozzáadódik egy egységnyi érték, ezzel jelezve, hogy plusz egy számítás történt. Példaként vegyünk egy A szenzort, aki nem referencia pont és a szülői legyenek a szuperszenzorok, S_1 , S_2 , S_3 . Ekkor A hopszáma az alábbi lesz:

$$(1+1+1)/3 = 1 \quad (18)$$

Ha B szintén nem referencia szenzor szülői közé beletartozik ezután A szenzor és két szuperszenzor, akkor az ő hopszáma a következőképp alakul.

$$(1+1+2)/3 = 1,33 \quad (19)$$

A bemutatott példa a következő 2. ábrán látható.



2. ábra, PAHN algoritmus szemléltetése

Ebből kifolyólag a hopszám azt is mutatja, hogy mennyire pontos számítást tudunk végezni a pozicionálás során. Minél kisebb egy szenzor hopszáma, annál pontosabb számítási eredményt kapunk. Pont ezért egy szenzor az alapján választ magának szülőket, hogy minél kisebb legyen a hopszáma. Így

mindig a lehető legpontosabb koordinátákkal tud majd számolni. Erre az algoritmusra a továbbiakban PAHN (*Position Algorithm with Hop Number*) fogok hivatkozni.

C. Átlagolást alkalmazó algoritmus

A következő algoritmus is a szülők választásában tér el az előzőektől. Ebben az esetben nem csak egy lehetőséget vizsgál a módszer, hanem minden lehetséges hármas csoportot kiválaszt, akik szülők lehetnek és ezek alapján kiszámolja mindegyikre a koordinátákat. Végül pedig ezen értékek átlagát veszi. Erre az algoritmusra a továbbiakban EBPA (*Extended basic positioning algorithm*) fogok hivatkozni.

IV. JELTERJEDÉSI MODELLEK KEZELÉSE

A szimulációm során két jelterjedési modellt is alkalmaztam. Amikor egy szenzor adatot küld át egy másiknak, akkor a jel egy valós terepviszonyokkal rendelkező területen halad keresztül. A terepviszonyok befolyásolják a jel csillapítását, ezért ehhez a folyamathoz a Deygout modellt[3] használtam fel. Ezután a vevő szenzor megkapja a jelet, de nem ismeri a befolyásoló terepviszonyokat, ezért ő a szabadtéri jelterjedési modell[4][5] szerint számolja ki a küldő szenzortól való távolságát. Mivel a jel valóságban megtett távolsága eltér a szabadtéri modell szerinti kiszámolt távolságtól, így a szenzor valamelyest arrébb helyezi magát, mint ahogy valójában van. Ekkor történik az a mérési eltérés, ami a későbbiekben hibatorlódáshoz vezet.

Minden pozicionálásnál a két modell által kiszámolt távolságok közötti eltérés miatt a számolt koordináták eltérnek a valós koordinátáktól. Ez az eltérés kezdetben még minimális, ugyanakkor többlépcsős helymeghatározásnál már olyan nagymértékű lehet, hogy adatvesztés léphet fel. Ez abból adódik, hogy a szenzor úgy érzékeli, hogy már nincs olyan közelségben a többiekhez, hogy továbbküldje az adatát és így az adatátvitel nem valósul meg, ellenben a valóságban még beleférne a hatótávolságba. Másik jelentős probléma, amikor a szenzor úgy érzékeli, hogy még nem lépte át a biztonsági sávhatárt és nem változtat irányt, a valóságban azonban már túllépte, ennek következtében leszakadhat a bolytól. Mindez azt jelenti, hogy a hibatorlódásból adódó következmények súlyosak lehetnek, ezért tartom fontosnak foglalkozni a továbbiakban a problémával. Ezzel kapcsolatban végeztem vizsgálatokat, amikre majd a következő fejezetben térek ki részletesen.

V. Eredmények

Azért, hogy az előző fejezetben bemutatott szenzorhálózatot megfelelően tudjam vizsgálni, egy szimulációs programot írtam C++ nyelven. A szimulációs programot úgy írtam, hogy minél inkább közelítse egy valós szenzorhálózat működését. Különböző paramétereket értékét lehet változtatni a programban a vizsgálandó eseteknek megfelelően. Ezek egy olyan kimenetet generálnak, amelyekből számszerű értékek kaphatók. Ezeket ábrázoltam grafikonokon és elemeztem őket. A programomat úgy írtam, hogy a vizsgálatoktól függően más és más paraméterek összefüggéseit elemezni tudjam. Ezért azokat a paramétereket, amelyeket később vizsgáltam, globális paraméterként vettem fel, így az értékük könnyen állítható,

változtatható. Ez azért szükséges, hogy a szimuláció ne csak egy értékre fusson le, hanem egy egész mintasorozatra is. Így kézzel tudtam az adott mintaértéket beállítani. Ezen kívül fontos, hogy ezek a globálisként megadott paraméterek jellemzik a szimulációs környezetet is. Azok a paraméterek, amelyek értékét az adott vizsgálat során nem változtattam, azok egy-egy környezeti jellemzőként viselkednek. Fontos, hogy egy szimulációs vizsgálat alatt több mintaérték esetén a környezeti tényezők ne változzanak, ezzel biztosítottam, hogy csak a vizsgált értéktől függjön az eredmény. A globális paramétereim az alábbi táblázatban található. Ezeket az értékeket használtam környezeti jellemzőkként, az éppen aktuálisan vizsgált paramétertől eltekintve, ezek az értékek nem változtak a vizsgálatok során.

A paraméter neve	Alapértelmezett érték [mértékegység]
Szenzorok száma	200 [db]
Terület	400 [m ²]
Elmozdulás maximális értéke	4 [m]
Szenzorok hatótávolsága	30 [m]
D pont távolsága	400 [m]

A *szenzorok száma* (n) adja meg, hogy összesen hány szenzor alkotja a szenzorhálózatot. Ez a szám magában foglalja a szuperszenzorokat és a mérésre szánt szenzorokat is. Ebből a 200 szenzorból 3 mindig szuperszenzor. Azért választottam ennyi szuperszenzort, mert 3-mal már helymeghatározás végezhető el. Tehát az lenne az ideális, ha 197 sikeres helymeghatározási algoritmus futna le minden lépésben.

A *terület* (w) azt a felszíni részt foglalja magában, amelyen a szenzorok elhelyezkednek, és amely területet vizsgálnak a naprendszerbeli égitesten. Ezt a területet kell bejárni a szenzoroknak, az itt mérhető adatokra van szükség. Ezzel az értékkel jelöltem ki azt a már fentebb említett sávot is, amely meghatározza, hogy a szenzorok milyen térrészen mozoghatnak még anélkül, hogy könnyen leszakadjanak a bolytól.

Az *elmozdulás maximális értékével* (m) azt a számot határoztam meg, amely egy felső korlátja a mozgásvektor értékének. Azaz egy vizsgált lépésben maximálisan ekkora távot tehet meg szenzor.

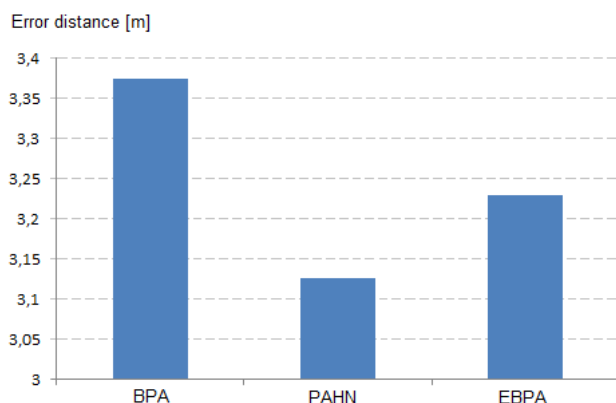
A *szenzorok hatótávolsága* (r) az a távolság, amelyen belül egy másik szenzort még lehet érzékelni és így képes a vele való kommunikációra. Ez a hatótávolság minden szenzorra ugyanúgy jellemző, és ugyanazt az értéket kapja. Tehát minden szenzor (a szuperszenzorok is) hatótávolsága megegyezik.

A *D pont távolsága* (d) pedig azt a távot határozza meg, amelyet el kell érnie legalább egy szenzornak, ahhoz, hogy a szimuláció véget érjen. Ha egy szenzor is elérte azt az x koordináta értéket, ami megegyezik ennek a D pontnak az x koordinátájával, akkor a szimulációt befejezettnek tekintem.

A. Hibatávolságok vizsgálata a módszerek függvényében

Ebben az esetben azt vizsgáltam, hogy a három algoritmusnál hogyan alakulnak a hibatávolságok. A hiba itt a valós koordináták és a helymeghatározó algoritmus által kiszámolt koordináták közötti eltérést jelöli méterben kifejezve. A

hibákat a szenzorokra és a megtett lépésekre nézve is átlagoltam, így egy átlátható képet kaptam az eredményekről.



3. ábra, Hibatávolságok a módszerek függvényében

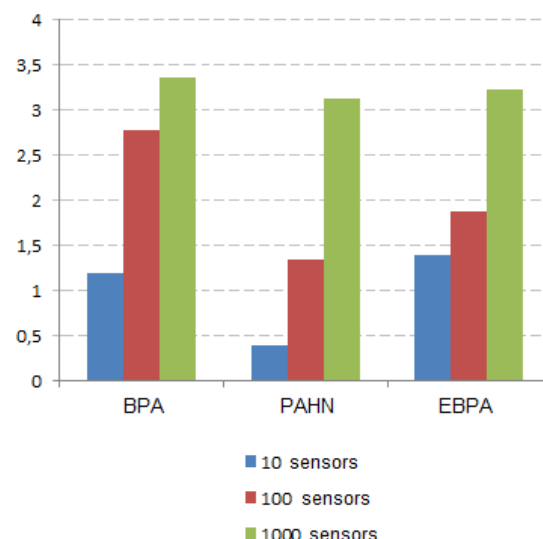
A 3. ábrán látható, hogy a BPA algoritmus átlagban majdnem 3,4 méteres hibával dolgozik. Ehhez képest a PAHN algoritmus jóval kisebb, ~3,1 méteres hibákat generál, míg az EBPA algoritmus a kettő közé lő be ~3,2 méteres hibával. Leolvasható, hogy mindkét javított módszer érzékelhetően jobb eredményt hoz. A kisebb hibatávolság kevesebb leszakadó szenzorral is jár, hiszen kevesebb eséllyel lépik túl észrevétlenül a biztonság sávot, illetve az adatvesztés lehetősége is csökken. A leghatékonyabb algoritmus ebből a vizsgálati szempontból a PAHN. Ez azzal magyarázható, hogy ebben az esetben mindig a lehető legpontosabb értékekkel történik a számítás, mindig a legbiztosabb helyzetű szenzorok segítségével történik a pozicionálás.

B. A hibatávolság változása a szenzorok számának függvényében

Ez a vizsgálatom abból a célból készült, hogy megfigyelhető legyen, hogyan alakul a hibatávolság átlagos értéke akkor, ha a szenzorok száma változik. Ez az eredmény mindhárom algoritmusnál értelmezhető, így egy grafikonon ábrázoltam a kapott értékeket az átláthatóság miatt.

A szimulációs eredményeket a 4. ábra mutatja. A BPA algoritmus esetében, ha növekszik a szenzorok száma, akkor egyre kisebb valószínűséggel lesz benne a három szuperszenzor az elsőnek kiválasztott szülők között. Így a kiszámolt koordináták is pontatlanabbá válnak. Az EBPA algoritmus értéke kevés (10) szenzorra a BPA értéke felett van. Ez azt jelenti, hogy kevés szenzorszám esetén nagyobb átlagos hibát kapunk, ez viszont nagyobb szenzorszám esetében megfordul. Ennek az oka, hogy az EBPA esetében a kiugróan magas eltérések elhúzzák az átlagolt koordináta értékeket rossz irányba, míg az BPA esetében kevés szenzornál nagyobb a valószínűsége, hogy a szülők maguk a szuperszenzorok lesznek. Itt is látható, hogy a PAHN algoritmus a hatékonyabb a három közül.

Error distance [m]



4. ábra, Hibatávolság a szenzorok számának függvényében

VI. Összegzés

Munkám során kialakítottam több olyan algoritmust, amelyet követve a szenzorhálózat tagjai képesek egymáshoz viszonyított helyzetüket meghatározni. A becslési eljárások azonban nem adnak pontos eredményeket, hiszen mérési és számítási hibákkal mindig kell számolnunk. A szimulációk során a hibák meglétére és torlódására fókuszáltam, ami azt eredményezte, hogy meg tudtam állapítani egyes összefüggéseket az alkalmazás hatékonyságának terén, illetőleg alapot adott az algoritmusok összehasonlítására. A hibatorlódás többlépcsős helymeghatározás esetén egyre nagyobb mértékben van jelen és egyre inkább eltérő eredményeket ad a valós helyhez viszonyítva. Ez olyan problémákat vet fel, amelyek veszteségekkel is járhatnak. A szimulációs mérések azt bizonyították, hogy a javasolt többlépcsős helymeghatározás során a 2. algoritmus bizonyult a leghatékonyabbnak. Dolgozatomban igyekeztem figyelni arra, hogy aktuális problémát közelítsek meg és ennek a problémának minél részletesebb és hatékonyabb vizsgálatára törekedjek.

IRODALOMJEGYZÉK

- [1] John Broph et al.: "Asteroid Retrieval Feasibility Study". Keck Institute for Space Studies, California Institute of Technology, Jet Propulsion Laboratory, 2013.
- [2] Z. Butler and D. Rus: *Event-based Motion Control for Mobile Sensor Networks*, IEEE Pervasive Computing, Vol. 2, No. 4, pp. 34–42, Oct.-Nov. 2003.
- [3] Haslett, C.: *Essentials of Radio Wave Propagation*. Cambridge (UK): Cambridge University, Press, 2008.
- [4] Horváth Zoltán: *A Terepdomborzat hatása a kis és közepes magasságon feladatokat végrehajtó pilóta nélküli repülőgép kommunikációs csatornájának stabilitására*, Hadmérnök: katonai műszaki tudományok online (2006-) IV. évfolyam 3. szám 2009. szeptember
- [5] Bori Zoltán: *Beltéri helymeghatározó rendszerek optimalizálása*, Budapesti Műszaki és Gazdaságtudományi Egyetem, 2011, Budapest

SSR based monopulse tracking system

Dóra Kiss¹, Levente Dudás², Ádám Szüllő³, Rudolf Seller⁴

Budapest University of Technology and Economics

Department of Broadband Infocommunications and Electromagnetic Theory

Budapest, Hungary

¹dkisshun@gmail.com, ²dudas@mht.bme.hu, ³szullo@mht.bme.hu, ⁴seller@mht.bme.hu

Abstract— SSR (secondary surveillance radar) systems are obtaining duplex communications between the transponders on the aircrafts and the interrogators on the ground stations. Transponders are communicating in either mode A/C or S, using different up- and downlink frequencies. Monopulse is a special beam-forming method which ensures a more accurate detection of the tracks than the conventional technologies. With these technological backgrounds the aim is to develop a tracking system that detects the tracks and rotates the antenna-matrix exactly to the direction of the aircraft. First of all, this project requires the design, simulation, production and measurement of the antenna array and the hybrid system. Next step is to design the hardware and software of the signal-processing PCB (printed circuit board), and the control system of the rotator. The final step of production is to test the system under proper circumstances.

Keywords- Secondary surveillance radar, monopulse, PCB design, microstrip hybrid.

I. INTRODUCTION

This individual tracking system includes the advantages of the SSR systems and the monopulse beam-forming technology as well. Further advantages are the small size, portability and relatively cheap manufacturability.

The antenna array consists of 4 rectangular patch antennas resulting 3 channels, such as one sum (Σ) and two differentials (Δ). The high frequency signal processing part consists of 4 quadrature microstrip hybrids [1]. Both of these above are manufactured of FR4 printed circuit board disks. These channels are connected to the PCB by SMA connectors then after filtering the signal, a mixer converts it to IF (Intermediate frequency) so it can be digitalized by an ADC (Analog-digital converter). Finally the digital signal is going to be processed by an FPGA (Field programmable gate array).

The development started with producing the antennas and the high-frequency signal processing unit. As for the antennas microstrip patch radiators were chosen, since there are many advantages of this type of antennas such as planar, simple in structure, low in cost and easy to be fabricated. However it suffers from an inherent limitation of narrow bandwidth.

II. THEORETICAL BACKGROUND

A. Basics of SSR systems

The efficiency of this technology lies in the usage of different up- and downlink frequencies. As for uplink, 1030

MHz is used, for downlink 1090 MHz so there is no interference between the channels. Another advantage comparing to the primer radar technology is that unlike the R^4 propagation loss in the primer case, the loss is decreased till R^2 , where R is the distance between the ground station and the track. Also contrary to primer systems the detection does not depend on the shape and size of the object. However a minimal input signal level is necessary for the transponder to answer, which is called ERP (Effective radiated power). According to the ICAO (International Civil Aviation Organization) regulations, the ERP is 52.5 dBW, which limits the range to approximately 200 NM.

B. Monopulse

In case of monopulse technology the receiver beams are simultaneously overlapping each other [2]. If a target enters the beams, it is sensed by multiple receivers so amplitude – and phase comparison can be done [3]. This makes possible to detect the aircraft's angle parameters, radial velocity and distance more accurately. The name itself expresses the advantage that only one pulse is enough to get precise angle information about the tracks.

To get this beam form, it is needed to obtain a Σ and two Δ signals out of the high frequency signals that come from the antennas. To achieve this and to preserve phase stability, these summarizing and subtracting circuits are being produced also on RF.

As the first step of signal processing the PSD (Phase Sensitive Detector) creates an error signal e which can be calculated as it is shown in (1).

$$e = \frac{\Delta}{|\Sigma|} \cos\theta \quad (1)$$

If we plot this e as a function of $\Delta\Phi$, where the latter is the azimuth or elevation deviation from the main direction, it can easily be seen that the Δ varies faster than the Σ , giving more exact estimation. The gradient of the curve depends on the beams' degree of the overlap.

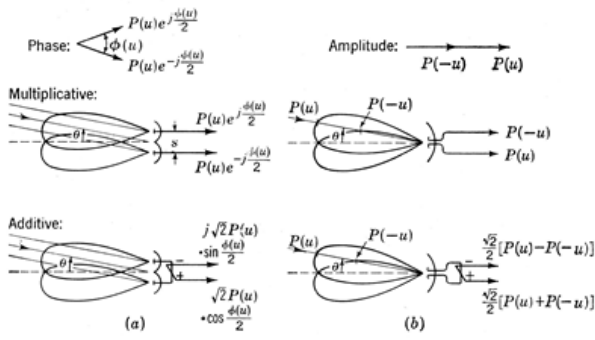


Figure 1 Graphical sketch of the Amplitude – and phase comparison, showing the overlap of the beams

III. RECTANGULAR MICROSTRIP PATCH ANTENNA-MATRIX

Microstrip patch antennas are simple radiating planar structures situated on one side of a dielectric substrate which has a grounded plane on the other side. It has many advantages and widely used not just for civilian applications but in military as well. They can be described as cavity resonators [5], since it can be modeled with a shortcut on the top and bottom planes with open circuits on both other sides. The evolving dominant propagation mode will be quasi-transversal electromagnetic (qTEM), because it's radiating not just in air but also in the dielectric substrate [4]. This fringing around the periphery of the edge will alter the original dielectric constant (ϵ_r) and results an 'effective dielectric constant' as follows.

$$\epsilon_{reff} = \frac{\epsilon_r + 1}{2} + \frac{\epsilon_r - 1}{2} \left[1 + 12 \frac{h}{W} \right]^{-1/2} \quad (2)$$

Where h is the height (35 μm) and W is the width of the radiator.

1) Geometry

The radiation pattern are affected by the following parameters

- L (Length)
- W (Width)
- Dielectric height (h)
- Patch height (t)

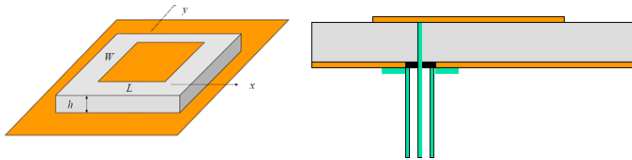


Figure 2 Geometric parameters and coaxial cable feeding of the rectangular microstrip patch antenna

The length of the patch should be less than $\lambda/2$ hence both the TM_{10} and TM_{01} can propagate [6]. The tangential components of the electric field intensity are being summarized in phase, creating the radiating field that is perpendicular to the

patch. This created field lengthens the original L parameter such as

$$L_{eff} = L + 2\Delta L \quad (3)$$

where

$$L_{eff} = \frac{c}{2f_0 \sqrt{\epsilon_{reff}}} \quad (4)$$

And

$$W = \frac{c}{2f \sqrt{\frac{\epsilon_r + 1}{2}}} \quad (5)$$

is how the width of the patch can be calculated.

2) Feeding

Feeding can be done two ways, either microstrip line connection or probe feed by coaxial cable [4]. The latter will be used in this project.

The location of the feeding will determine the polarization of the rectangular patch. To obtain circular polarization the feeding should be on the diagonal, so both TM_{10} and TM_{01} can be created at the given frequency. This way the phase shift of the modes is 90° so they will compose orthonormal basis. The modes will be summarized in far field, achieving the circular polarization.

3) Simulation

After calculating the W and L parameters, the patch can be simulated using various finite element methods. The other parameters like h and t are already given, since it has been previously decided to use FR4 printed circuit board sheet, which means $t = 35 \mu\text{m}$ and $h = 1.6 \text{ mm}$. To simulate the final radiation pattern Ansoft HFSS is used.

However, before simulating in HFSS, estimations can also be given to determine the distances of the patch radiators to achieve the best pattern. It's obvious that they should be placed equidistant but the radiation pattern highly depends on this distance which can be explained by wave-propagation qualities. To make calculations easier, distance is always being expressed in λ .

Let us place the centers of the patch antennas to the corners of a $\lambda/2$ -sided square, while the length and width of the patch is approximately $\lambda/4$. In this construction the radiation pattern of the Σ and Δ channels can be simulated by LabView. The simulation returned with the following result.

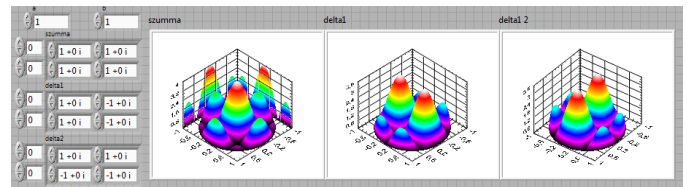


Figure 3 LabView simulation result for the radiation patterns of the sum and delta channels respectively. On the left side of the picture the matrices can be seen, such as $\Sigma = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ in case of the sum.

Figure 3 shows the antenna matrix of a $\lambda/4$ sized 4-patch system in a λ distance from each other on both sides. This proves a satisfying result so using this information a final-pattern simulation in HFSS was made which resulted an approximately 10 dB of Gain. The radiation pattern was also proper, so **one** patch antenna was manufactured. It was necessary to see if the HFSS simulation works properly in practice. Obviously the final antenna array will have slightly different equivalent radiation pattern but the operating frequency and the reflection coefficient can already be measured at an individual patch which validates the operation of the array later to be realized. Since S_{11} was around -15dB at the required 1090 MHz of frequency, we can state that shaped and sized like this patch (using the same FR4) can be used to obtain the whole array.

IV. QUADRATURE MICROSTRIP HYBRID SYSTEM

To obtain the Δ and Σ channels the choice was the creation of a quadrature microstrip hybrid system that consists of 3 hybrids [7].

Main parameters of the hybrids can be seen below.

$$f_0 = 1090 \text{ MHz} \quad (6)$$

$$Z_0 = 50 \Omega \quad (7)$$

$$Z_0/\sqrt{2} = 35 \Omega \quad (8)$$

$$W_{Z_0} = 3.000 \text{ mm} \quad (9)$$

$$L_{Z_0} = 37.95 \text{ mm} \quad (10)$$

$$W_{Z_0/\sqrt{2}} = 5.199 \text{ mm} \quad (11)$$

$$L_{Z_0/\sqrt{2}} = 36.98 \text{ mm} \quad (12)$$

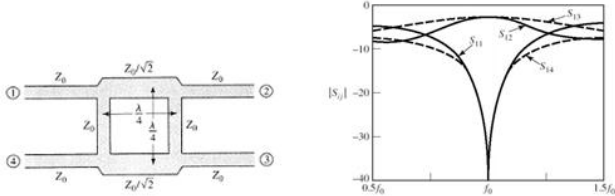


Figure 4 Sketch of a quadrature microstrip hybrid and its ideal S-parameters

As it can be seen on Figure 4, these hybrids are 4-port directional couplers that are designed for 3 dB power split. The width of the lines for the Z_0 and $Z_0/\sqrt{2}$ lines are 3 mm

and 5.2 mm respectively [8].

Assuming that we have two input signals like vectors of magnitude A and B , the sum and a difference output can be easily obtained by adding a 90° section to one of the ports. Lengthening input A by a $\lambda/4$ using Z_0 impedance results a sum – and a difference output at the end [9]. Using vector notations the channels can be described as follows:

$$\Sigma = \frac{1}{\sqrt{2}}(A\angle -180^\circ)(B\angle -180^\circ) \quad (13)$$

$$\Delta = \frac{1}{\sqrt{2}}(A\angle -270^\circ)(B\angle -90^\circ) \quad (14)$$

The material is an FR4 board just like in the patch antenna system's case. It gives the conducting layer (copper) height such as the dielectric constant that is typically 4.3.

The simulated results are pretty much similar to the expected (ideal) ones so using those parameters it has been produced. Using the laboratory's S-parameter vector analyzer all the ports has been measured to get the actual values of S_{11} , S_{12} , S_{13} and S_{14} . The measurements showed quite good results on the given 1090 MHz.

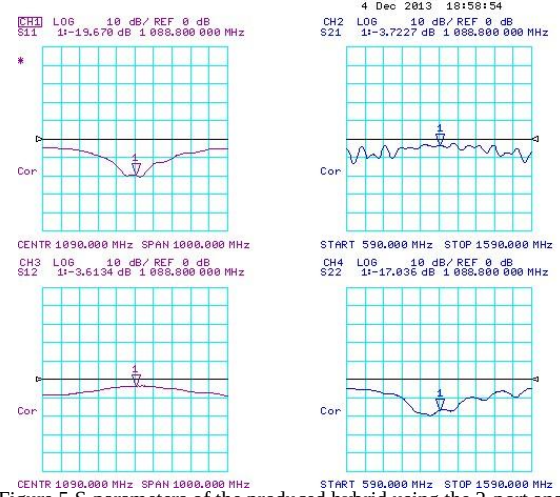


Figure 5 S-parameters of the produced hybrid using the 2-port analyzer for measurement. The return loss is -19 dB, the S_{12} and S_{21} parameters are close to -3 dB which are satisfying results.

Furthermore to prove that the hybrid operates properly the phase difference between two output ports can also be measured. For this measurement two port of the hybrid had been measured one by one. Using the results a plot was easy to obtain, which can be seen on Figure 6.

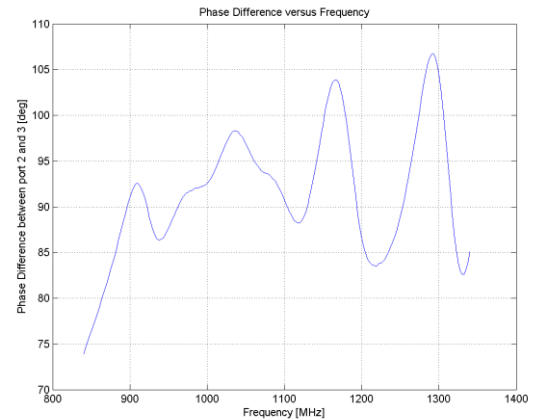


Figure 6 Plot of the phase difference between the two output ports. X-axis shows the frequency in MHz units while Y-axis is the phase difference in degrees. The phase shift at the 1090 MHz center frequency is quite close to 90° which allows us to state that the hybrid operates properly.

According to these measurements the hybrids can be used to create the aforementioned Δ and Σ channels.

V. PROSPECTS AND CONCLUSIONS

Summarizing the results that have been described through the chapters above, we can state that the microstrip patches can be manufactured, the three channels can be created using the quadrature hybrids with elongated benchmarks. As for the PCB design, all the components are chosen and all the connectors are defined. The schematic is ready, after the design of the layout it can be sent to production and the programming job can soon be done.

ACKNOWLEDGMENT

This work had been supported by the Radar Research and Remote Sensing Group Laboratory. The author would like to thank the help and all the education to the supervisors.

REFERENCES

- [1] S. Kingsley, Understanding Radar Systems. SciTech Publishing, 1st edition, 1999
- [2] Merrill Skolnik, Radar Handbook. McGraw-Hill Professional, 3rd edition, 2008
- [3] M. N. Jazi, Z. H. Firouzeh, H. Mirmohammad-Sadeghi, "Design and Implementation of Aperture Coupled Microstrip IFF Antenna" PIERS ONLINE, VOL. 4, NO. 1, 2008
- [4] K.L. Wong, Design of Nonplanar Microstrip Antennas and Transmission Lines, John Wiley & Sons, New York, 1999.
- [5] K.S. Yee, "Numerical Solution of Initial Boundary Value Problems Involving Maxwell's Equations in Isotropic Media", IEEE Trans. Antennas and Propagat., AP-14, pp. 302 - 307, 1966.
- [6] M. Amman, "Design of Rectangular Microstrip Patch Antennas for the 2.4 GHz Band", Applied Microwave & Wireless, pp. 24 - 34, November/December 1997.
- [7] J. Lange, "Interdigitated stripline Quadrature Hybrid", IEEE Transactions on Microwave Theory and Techniques, Vol. MTT-20, December 1969
- [8] Wen Pin Ou, "Design Equations for an Interdigitated Directional Coupler", IEEE Transactions on Microwave Theory and Techniques
- [9] Presser A, "Interdigitated Microstrip Coupler Design", IEEE Transactions on Microwave Theory and Techniques, vol. MTT- 26, 1978

DVB-T Alapú Passzív Radar

Pető Tamás

Mikrohullámú Távérzékelés Laboratórium
Szélessávú Hírközlés és Villamosságtan Tanszék
Budapesti Műszaki és Gazdaságtudományi Egyetem
Budapest, Magyarország
E-mail: petotamas.mk@gmail.com

Dudás Levente

Mikrohullámú Távérzékelés Laboratórium
Szélessávú Hírközlés és Villamosságtan Tanszék
Budapesti Műszaki és Gazdaságtudományi Egyetem
Budapest, Magyarország
E-mail: dudas@mht.bme.hu

Kivonat—A légtér ellenőrzéséhez használt radar rendszerek többsége radarjelek folyamatos sugárzásával képesek felderíteni a repülőgépek helyzetét. Ezek a radarjelek rendszerint olyan nagy energiájú impulzusok, amelyek előállításához rendkívül költséges nagy teljesítményű mikrohullámú erősítők szükségesek. Az adóegységek számottevő gyártási és üzemeltetési költsége mellett a kibocsátott nagy energiájú impulzusok alapján a radar könnyedén detektálható és ez által zavarható is. A passzív radar rendszerek, ezzel ellentétben nem tartalmaznak adóegységet, ezért szinte lehetetlen érzékelni őket. Most, hogy az eddig leküzdhetetlen akadályt állító technológia korlátok kezdenek eltűnni, az elméleti kutatások mellett kiemelt szerepet kap a rendelkezésre álló eszközökkel való realizálhatóság vizsgálata is. A cikk egy olyan passzív radar megvalósíthatósági kísérletét mutatja be, amely a digitális földfelszíni műsorszóró jelet (DVB-T) használja fel a céltárgyak észlelésére. A cikkben egyszerű modellből kiindulva szimulációs és mérési eredményeken keresztül kerül igazolásra a passzív radarok detektálási mechanizmusa.

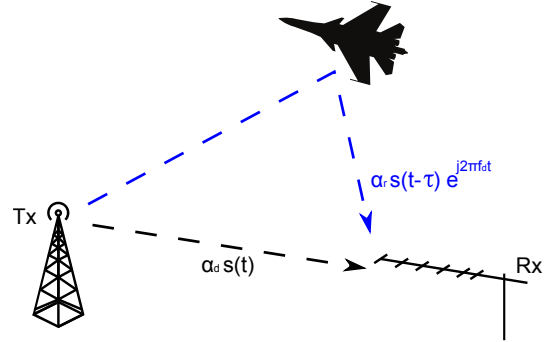
Index Terms—DVB-T passzív radar, légtér ellenőrzés, passzív radar kísérlet

I. BEVEZETŐ

A passzív radar rendszerek a környezetükben a már eleve meglévő lévő rádiófrekvenciás forrásoktól származó jeleket használják fel a céltárgyak észlelésére. A módszer alapötlete korántsem új keletű, azonban a gyakorlatba való átültetés csak a jelenlegi eszközökkel vált elérhetővé.

Az utóbbi időben ezért megnövekedett számú tudományos publikáció foglalkozik a passzív radarok működésével [5]. Számos ezek közül behatóan tanulmányozza a lehetséges megvilágító források képességeit [7],[5]. Az eddigiekben részletesen kivizsgált megvilágító források közül a DVB-T jel kiemelkedően jó korrelációs tulajdonságokkal rendelkezik, ezért alkalmas lehet a passzív elven működő radarok referencia forrásként való felhasználására.

A DVB-T jel alkalmazhatóságával kapcsolatban jelentős számú publikáció érhető el [3],[4],[6],[7],[8]. Az eddig bemutatott eredmények azonban csak igen kis számban alapoznak ténylegesen realizált rendszerek mérési eredményeire. A dokumentum célja, hogy az eddigiekben elméletben részletesen kivizsgált eljárásokat gyakorlati módszerek alkalmazásával is igazolja. A detektálási eljárások működésének verifikálása első közelítésben szimulációs technikákkal majd repülőtéri mérésekkel is bemutatásra kerülnek.



1. ábra. Passzív radar jelútjai egyszerűsített modell szerint

II. A MŰKÖDÉS ALAPELVE

A szakaszban röviden ismertetésre kerül a passzív radarok egy leegyszerűsített modellje, amely alapján megérthető a detektálás alapvető koncepciója. A fenti (1.) ábra a radar működését illusztrálja. Az adótoronyból kisugárzott műsorszóró jel a közvetlen terjedési úton kívül számos más reflexiók úton keresztül is eljuthat a radar vevőantennájához. A legegyszerűbb modell felépítéséhez azonban a különböző terjedési utak közül elegendő kizárólag két úttal számolni. Ezek a közvetlen jelút és egy az érzékelni kívánt repülőgépről reflektált jel útja.

A különböző utakon terjedő jelek egymástól függetlenül csillapítódnak. Jelöljük a direkt út csillapítását α_d -vel, valamint a reflektált jelút csillapítását α_r -el. A csillapítás mellett mindkét jel a véges terjedési sebességek miatt eltérő idő késleltetéssel érkezik meg a vevőhöz. Jelöljük a repülőgépről reflektált jel és a direkt úton érkező jel időkésleltetésének különbségét $\tau = t_r - t_d$ -vel. A mozgó repülőgépről visszavert jel a repülőgép sebességétől, irányától és a geometriai helyzetétől függően valamekkora Doppler-eltolódást is szenved. Ezt jelöljük f_d -vel. A radar vevőantennája tehát ennek a két jelnek az összegét fogja venni. Matematikailag ezek alapján az $r(t)$ vett jelet a következőképpen írhatjuk le:

$$r(t) = \alpha_d s(t) + \alpha_r s(t - \tau) e^{j2\pi f_d t} \quad (1)$$

A detektálás során célunk, hogy a vett jelben megkeressük a közvetlen úton érkező jel vagy más néven referencijel időben τ -val késleltetett és f_d Doppler-frekvenciával eltolt másolatát. Ez praktikus korreláció számítással tehető meg. A vett jel autokorrelációs függvényének kiszámításával, azokon a helyeken kapunk maximális jel-zaj viszony értékű csúcsoakat, ahol a vett jel a lehető legjobban egyezik meg önmaga időben késleltetett változatával. A direkt úton érkező (referencia) jelet a korreláció elvégzése előtt korrigálnunk kell ugyanazzal az f_d Doppler-frekvenciával, amivel a reflektált jel is eltolódik, hogy maximális kiemelést kapjunk. Az előzőek alapján a repülőgép detektálását az alábbi kétdimenziós Doppler-eltolás keresztkorrelációs függvény kiszámításával végezhetjük el:

$$\chi(\tau, f_d) = \int_{-\infty}^{\infty} r(t) r^*(t - \tau) e^{j2\pi f_d t} dt \quad (2)$$

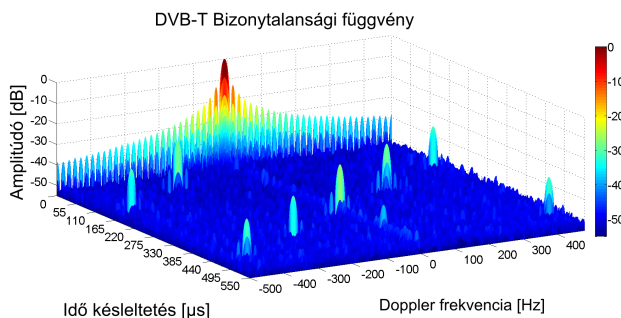
III. A DVB-T JEL BIZONYTALANSÁGI FÜGGVÉNYÉNEK ANALÍZISE

Mivel a jelfeldolgozó algoritmusokban korreláció számítását alkalmazunk, ezért a felhasznált megvilágító jel korrelációs tulajdonságainak vizsgálata meghatározó jelentőségű a passzív radar hatékony működésének szempontjából. A megvilágító jel alkalmazásában rejlő lehetőségek elemzéséhez a jel egyértelműségi függvényét kell kivizsgálnunk. Az egyértelműségi függvény kiszámítási módja jelen esetben megegyezik az első szakaszban közölt detektálás során számítandó Doppler-eltolás korrelációs függvénnyel (2).

A maximális Doppler-eltolódás 500 Hz-re lett megválasztva egy átlagos céltárgy várható sebességének megfelelően (UHF sávú adás, maximális 300 – 400 $\frac{\text{km}}{\text{h}}$ -ás céltárgy sebesség). Az egyértelműségi függvény kiszámítása során a koherens integrálási idő hozzávetőlegesen 60 ms a kellően részletes néhány Hz-es Doppler felbontás biztosításához.

A Digitális földfelszíni televíziós műsorszóró jel a felhasznált OFDM modulációs sémának és az alkalmazott csatorna és forrás kódolásoknak köszönhetően kellő képen véletlenszerűsített. OFDM moduláció használata során ugyanis az átvinni kívánt információt, egymáshoz relatíve közel lévő szinuszos vivők sokaságára ültetik rá.

A (2.) ábra a DVB-T jel egyértelműségi függvényét ábrázolja. Az ábra alapján a DVB-T jel Doppler-eltolás korrelációs függvénye 50 dB-es dinamikatarományal rendelke-



2. ábra. A DVB-T jel bizonytalansági függvénye

zik. Ez a dinamikataromány, (mint az a későbbiekben bemutatásra kerül majd) már elegendően nagy ahhoz, hogy a reflexiókból származó korrelációs csúcsoakat azonosítani lehessen. A függvényben megjelenő korrelációs kiemeléseket a jelben lévő folytonos és szórt pilotjelek okozzák. Ezek a korrelációs csúcsoak esetenként megnehezíthetik a céltárgyak detektálását, ezért az eltüntetésükkel érdemes foglalkozni. [3],[4],[6],[8]

IV. A DETEKTÁLÁSI ELJÁRÁSOK SZIMULÁCIÓJA

A szakaszban egy DVB-T jelet felhasználó passzív radar kísérleti megvalósítása kerül bemutatásra. A tesztelés helyszínéül célszerű olyan helyet választani, ahol a mozgó céltárgyak közel haladnak el a radar antennájához, ezáltal viszonylagosan nagy teljesítményszintűek a reflektált jelek. A mérés helyszínéül ezért a Budapesti Liszt Ferenc Nemzetközi repülőtérrel választottuk.

A működés koncepciója első közelítésben szimulációs módszerekkel került ellenőrzésre. A szimulációs program segítségével egy felszálló repülőgép detektálhatóságát vizsgáltuk. A szimulációs eljárás során előállítjuk a radar vevőantennájának talppontjában lévő jelek mintavételezett reprezentációja, majd ezután a szimulált jeleken végrehajtjuk a passzív radar jelfeldolgozási eljárásokat.

A program blokkvázlatos működését a lenti(3.) ábra szemlélteti. Az első lépésben a céltárgy geometriai paraméterei kerülnek kiszámításra az aktuális koordinátái, sebessége és gyorsulása alapján. A kapott távolság és szögparamétereket felhasználva a következő összefüggések segítségével a szimulátor kiszámítja a figyelembe vett jelutak paramétereit ($\alpha, \tau, f_{Doppler}$).

A következő egyenlet segítségével a referencijel vételi szintjét számíthatjuk ki.

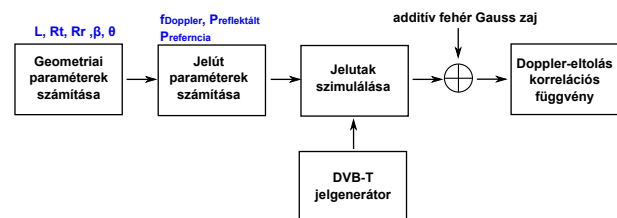
$$P_{referencia} = \frac{P_t G_t G_r \lambda^2}{(4\pi)^2 L^2} \quad (3)$$

A bisztatikus radaregyenletet felhasználva a céltárgyról reflektált jel teljesítményszintjét határozhatjuk meg.

$$P_{reflek} = \frac{P_t G_t \sigma G_r \lambda^2}{(4\pi)^3 R_t^2 R_r^2} \quad (4)$$

A céltárgyról reflektált jel Doppler eltolódásának mértékét pedig a következő egyenlettel számíthatjuk.

$$f_{Doppler} = \frac{2v}{\lambda} \cos(\theta) \cos\left(\frac{\beta}{2}\right) \quad (5)$$



3. ábra. Szimulációs program blokkdiagramja

A fenti összefüggések (3, 4, 5) paramétereinek jelentése, illetve a szimuláció során használt értékei az alábbiak:

- $P_t \rightarrow 100 \text{ kW}$ -az adó teljesítménye (Széchenyi hegyi DVB-T adótorony)
- G_t -adó antenna nyeresége
- R_t -az adó és a céltárgy közötti távolság
- R_r -a vevő és a céltárgy közötti távolság
- $\sigma_b \rightarrow 100 \text{ m}^2$ -bisztatikus radar hatásos keresztmetszet
- $L \rightarrow 26 \text{ km}$ -vevő és adótorony közti távolság (Széchenyi hegyi DVB-T adótorony)
- G_r -vevő antenna nyeresége
- $\lambda = 0.5 \text{ m}$ -hullámhossz
- θ, β az adó, a céltárgy és a radar elhelyezkedéséből adódó geometriai paraméterek.

[1],[2]

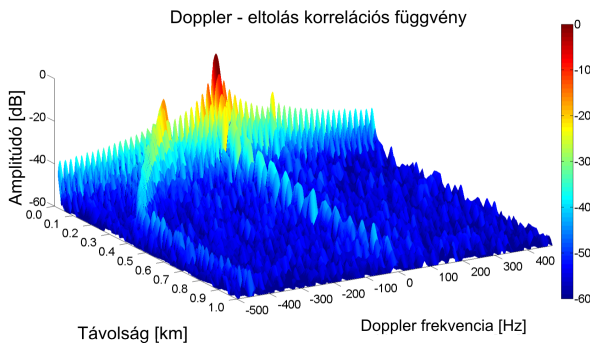
Ezt követően a DVB-T generátor által előállított forrásjelet $s(t)$ módosítva elkészül a végleges szimulációs jel. Az előállított jel matematikai alakját a következő formában írhatjuk le:

$$\sum_{i=1}^N \alpha_i s_i(t - \tau_i) e^{j2\pi f_{di}t} + n(t) \quad (6)$$

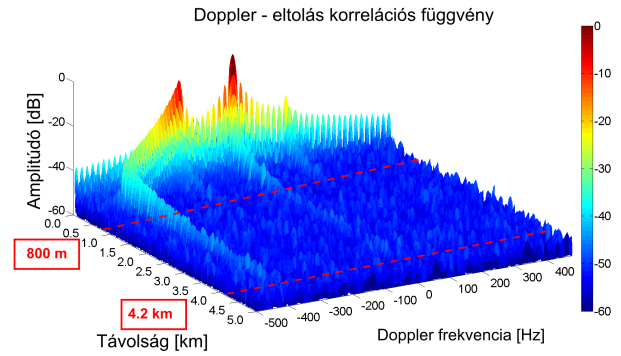
, ahol i az aktuálisan figyelembe vett jelút sorszámát, α a jelút csillapítását, τ az idő késleltetését, f_d pedig a Doppler frekvenciáját jelöli, $n(t)$ a jelet terhelő additív fehér Gauss zaj. A szimulációs jelet 10^6 mintával állítottuk elő (a mintavételi frekvencia 9 MHz).

A szimulációt iteratíván több egymást követő időpillanatra futtattuk le. Az eredmények összefűzésével a lenti (4) ábrán látható eredményt kaptuk. Az ábrán látható Doppler-eltolás korrelációs függvényen egyértelműen azonosítható a felszálló céltárgy nyomvonala.

Igen fontos megjegyezni, hogy a szimuláció elvégzése során izotróp vevő antennát feltételeztünk a radar vételi oldalán. Jól látható, hogy a céltárgy hozzávetőlegesen 800 m-es távolságnál alámerül a referenciajel korrelációs függvényének dinamikatarományába. A kísérleti mérés eredményességének növelése érdekében ezért a vevő oldalon mindenképp célszerű



4. ábra. Felszálló céltárgy nyomvonala a Doppler eltolás korrelációs függvényen izotróp vevőantenna használatával



5. ábra. Felszálló céltárgy nyomvonala a Doppler eltolás korrelációs függvényen irányított Yagi vevőantenna használatával

legalább irányított antenna használatával növelni a reflektált jel és a referencia jel közötti izolációt. A következő (5.) ábrán látható kép egy 15 dB előre-hátra viszonytal rendelkező Yagi antenna feltételezésével készült szimuláció eredményét mutatja be.

A radar várható hatótávolsága ezáltal 4.2 km -re növekedett meg az előbbi 800 m -es esethez képest. A passzív radar működési alapkoncepciójának kipróbálásához ez a hatótávolság már elegendően nagy lehet. Szintén fontos megemlíteni, hogy bemutatott szimulációk során a céltárgy hatásos radar keresztmetszetének változása, illetve a vevőantenna fő nyálábjában való tartózkodásának mértéke nem lett figyelembe véve.

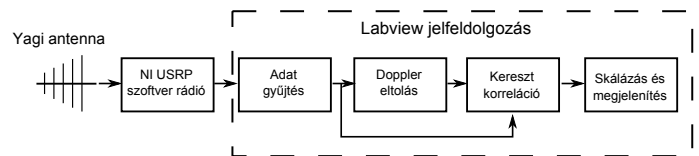
V. REPÜLŐTÉRI MÉRÉS

V-A. Mérési összeállítás

A rádiófrekvenciás jelek vétele és adatgyűjtése a National Instrument 2920-as univerzális szoftver rádiójával történt. Jelen esetben a tesztelés elvégzéséhez ez igen kézenfekvő választás, hiszen a radar elvi működésének kipróbálásához nem szükséges vevő egység tervezése és építése. A mérés elrendezését (6.) ábra szemlélteti. A mérések az előző szakaszban szimulált 15 dB-es előre-hátra viszonytal rendelkező Yagi antennával készültek. A méréshez tervezett Yagi antennáról érkező jeleket a szoftver rádió lekeveri, mintavételezi, majd egy, a méréshez elkészített Labview program számára az IQ mintákat gigabit Etherneten keresztül továbbítja. A mért adatok tárolása után a Doppler-eltolás korrelációs függvény számítása offline módon történt.

A szoftver rádió használata során beállított fontosabb paraméterek a következők voltak.

- Működési frekvencia $\rightarrow f = 610 \text{ MHz}$



6. ábra. Mérési összeállítás

- IQ Mintavételi frekvencia $\rightarrow f_s = 9 \text{ MHz}$
- Bitmélység $\rightarrow 16\text{bit}$
- Gyűjtött minták száma $\rightarrow 10^6$

V-B. Mérési eredmények

A minták begyűjtése után a 2. egyenlet alapján több egymást követő pillanatképre kiszámításra került a Doppler-eltolás keresztkorrelációs függvény. Egy pillanatkép kiszámításához 10^6 minta került felhasználásra, ezt átszámítva 110 ms-ot kapunk a koherens integrálási időre. A pillanatképek eredményeinek összevágásával a (7) és (8) ábrákon prezentált eredmények születtek. Az ábrákon láthatóan hozzávetőlegesen 300 Hz-es Doppler frekvencia körül megjelenik egy korrelációs csúcs, amely az idő elteltével mind távolságban, mind pedig Doppler frekvenciában elmozdul a kiszámított korrelációs függvényeken.

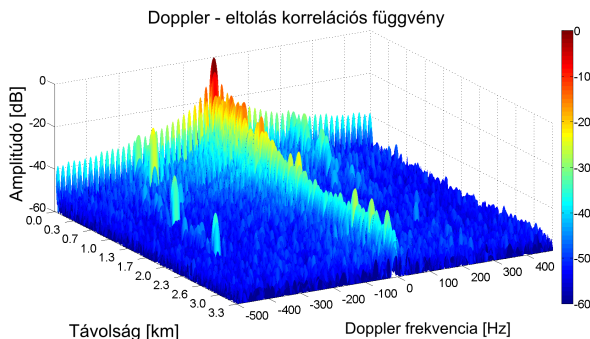
Az eredményeket összevetve, elmondhatjuk, hogy az előzetesen szimulált értékeknek megfelelő eredményeket kapunk. A radar mérési alapelve tehát helytállónak mondható.

A radar leglényegesebb korlátozó tényezője a nagy teljesítménnyel érkező direkt jel, amely az előzetesen becsülteknek megfelelően elfedte a céltárgyról távolabbról reflektálódó jeleket.

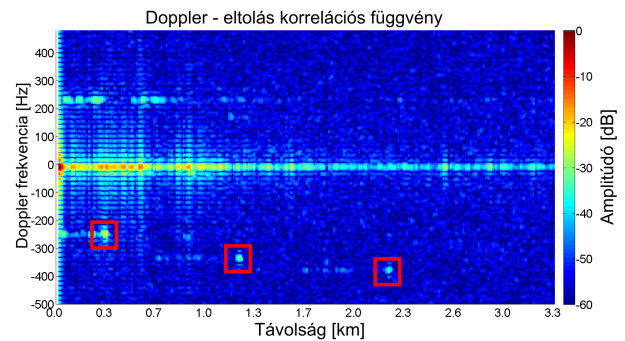
VI. ÖSSZEGZÉS

Napjainkban a passzív radarok kutatása kezd igen népszerűvé válni az alacsony gyártási és üzemeltetési költségeknek valamint a figyelemre méltó katonai jelentőségének köszönhetően. A radar elterjedése elsősorban az újonnan elérhető nagy teljesítményű jelfeldolgozó rendszereknek köszönhető.

A működés teóriájával, valamint a lehetséges megvilágítási források alkalmazásáról számos tudományos publikáció jelent meg az utóbbi időben. Emellett viszonylag kevés figyelem fordul az elérhető technológiai eszközökkel való realizálhatóság vizsgálatára. Jelen dokumentumban egy DVB-T (digitális földfelszíni videó műsorszó) jelet felhasználó passzív radar megvalósíthatósági kísérlete került bemutatásra. A radar alapvető működését először szimulációs, majd gyakorlati méréseken keresztül sikerült igazolnunk. A mérési



7. ábra. Sikeresen detektált felszálló céltárgy a kétdimenziós korrelációs függvényen - perspektívikus nézet



8. ábra. Sikeresen detektált felszálló céltárgy a kétdimenziós korrelációs függvényen - felül nézet

eredmények alapján a passzív radarok céltárgy detektálási elve az első szakaszban felvázoltak szerint gyakorlatban is helytállónak bizonyult.

Az adótorony által kisugárzott referenciajel viszonylag magas teljesítményszinttel érkezik a passzív radar antennájához, ami a céltárgyakra reflektált jelek elfedését okozza. A hatótávolság növeléséhez ezért a vett jelben ki kell oltani a nagy szintű referenciajelet. A referenciajel kioltása praktikus segédantennák, illetve antennarendszer alkalmazásával oldható meg. [9].

KÖSZÖNETNYILVÁNÍTÁS

A kutatási projektet a nemzeti fejlesztési ügynökség támogatta (NFÜ, Magyarország) az alábbi projekt keretében: Kognitív Módszerű UAV Ütközés Elhárítás Milliméteres Hullámhosszú Radar Alkalmazásával, UWBSRR12, 2012-2015.

HIVATKOZÁSOK

- [1] Merrill I. Skolnik "Radar Handbook" McGraw-Hill Professional; 2 edition January 1, 1990
- [2] Nicholas J. Willis "Bistatic Radar" by SciTech Publishing Inc. , 2005
- [3] M. Radmard, M. Bastani, F. Behnia, M. M. Nayebi "Cross Ambiguity Function Analysis of the '8k-mode' DVB-T for Passive Radar Application", Radar Conference, 2010 IEEE Washington, DC, 10-14 May 2010.
- [4] H. Andrew Harms, Linda M. Davis, James Palmer "Understanding the signal structure in DVB-T signals for passive radar detection", Radar Conference, 2010 IEEE Washington, DC 10-14 May 2010
- [5] F. Berizzi, M. Martorella, D. Petri, M. Conti, A. Capria "USRP technology for multiband passive radar", Radar Conference, 2010 IEEE-Washington, DC 10-14 May 2010
- [6] Diego Langelotti "Impact of synchronization on the ambiguity function shape for PBR based on DVB-T signals", Radar Symposium (IRS), 2010 11th International Vilnius, Lithuania, 16-18 June 2010
- [7] M. Radmard, M. Bastani, F. Behnia, M. M. Nayebi "Advantages of the DVB-T Signal for Passive Radar Applications", Radar Symposium (IRS), 2010 11th International Vilnius, Lithuania, 16-18 June 2010
- [8] Zhiwen Gao, Ran Tao, Yongfeng Ma, Tao Shao "DVB-T Signal Cross-Ambiguity Functions Improvement for Passive Radar", Radar, 2006. CIE '06. International Conference Shanghai, 16-19 Oct. 2006
- [9] Tao R., Wu H.Z., Shan T. "Direct-path suppression by spatial filtering in digital television terrestrial broadcasting-based passive radar", Radar, Sonar & Navigation, IET (Volume:4, Issue: 6) December 2010

Nulla törésmutatójú metaanyag tervezése tölcsérantennák irányítottságának növeléséhez

Farkas Pál, Szabó Zsolt

Szélessávú Hírközlés és Villamosságtan Tanszék

Budapesti Műszaki és Gazdaságtudományi Egyetem

1111 Budapest, Egry József utca 18. V. emelet

Absztrakt— A nulla törésmutatójú metaanyagok alkalmazhatók síkfelületű lencsék létrehozására, melyekkel növelhető egy antenna elektromágneses mérete, illetve csökkenthető egy adott nyereséghez szükséges apertúra méret. Olyan FishNet típusú metaanyag lencsét tervezünk, amelyik a 5,2 GHz-es tartományban működik. Bemutatjuk a tölcsérantenna-lencse alkalmazásra megtervezett metaanyag tulajdonságait.

Kulcsszavak: metaanyag, nulla törésmutató, távközlési antennák

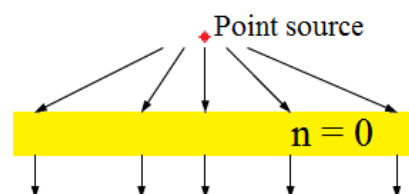
I. BEVEZETÉS

Az anyagok és az elektromágneses tér kölcsönhatása függ az anyag kémiai összetételétől, de ebben a tekintetben még inkább meghatározó a strukturáltság. Metaanyagoknak nevezzük az olyan hullámhossz alatti struktúrákat, melyek homogén anyagként viselkednek. Segítségükkel teljesen új, természetes anyagok esetén szokatlan vagy nem létező elektromágneses tulajdonságok hozhatók létre. A gyártási technológia fejlődése lehetőséget ad az ilyen struktúrák létrehozására, akár az optikai tartományra is.

Az elmúlt tíz év során világszerte nagyon intenzív metaanyag alapú kutatás folyik a mikrohullámú frekvencia tartományban. Számos olyan metaanyag alapú eszközt demonstráltak, amelyeknek a tulajdonságai sokkal jobbak vagy méretei lényegesen kisebbek a hagyományos mikrohullámú eszközöknél. A mikrohullámú tartományon használatos metaanyagokat alkalmazásuk szerint három fő csoportba lehet sorolni:

- hullámvezető komponensek kialakítására szolgáló metaanyagok, amelyekkel több sávú, speciális sáv szélességű miniaturizált hullámvezetők, kisméretű szélessávú hullámcsatolók, kompakt rezonátorok, szűrők, egyenletes teljesítmény összegzők és osztók készíthetők[1][2],
- mikrohullámú refrakciót megvalósító metaanyagok, amelyekkel fókuszáló metaanyag rétegek, metaanyag prizmák, antireflexiós metaanyagok, szuper felbontású képalkotó rendszerek, transzformációs optikával tervezett elrejtő köpenyek[3] készíthetők,
- antennák és antennarendszerek kialakítására szolgáló metaanyagok, amelyekkel passzív vagy aktív, egysávú vagy többsávú, egy-, vagy többdimenziós leaky-wave [4] vagy rezonáns antennák és visszaverők készíthetők.

A nulla törésmutatójú anyagok jellemzője, ahogy azt az 1. ábra is szemlélteti, hogy a belőlük kilépő hullámok merőlegesek az anyag felületére [5]. Ez a tulajdonság Descartes törvényével is alátámasztható. Ezt a jelenséget kihasználva mikrohullámon olyan sík lencsék készíthetők, melyekkel növelhető az antennák irányítottsága. Több, manapság publikált cikk is foglalkozik hasonló lencsékkel [6], [7] melyekben patch-, illetve tölcsérantennához tervezett lencsákat vizsgálnak.



1. ábra: Nulla törésmutatójú anyag viselkedése pontforrás általi megvilágítás esetén

Ezen cikk témája egy C sávú, nulla törésmutatójú metaanyag lencse tervezése egy optimális (piramidális) tölcsérantennához. Részletezzük az effektív anyagparaméterek kiszámításának módját, majd egy analitikus modellt ismertetünk, amellyel a vizsgált struktúrák elektromágneses jellemzői meghatározhatók. Bemutatjuk a vizsgált struktúrákat, és meghatározzuk a hozzájuk rendelhető effektív anyagparamétereket.

II. A C- SÁVBAN MŰKÖDŐ ANTENNA PARAMÉTEREI

Az antenna nyeresége 19 dBi, a tölcsér bennfoglaló méretei (apertúra méretek, magasság): 254mm, 199 mm, 277mm. A tölcsérantennák nyeresége méretük növelésével elviekben akármeddig növelhető, azonban a nyereség növekedése mellett az egy dB nyereségnövekményhez szükséges hossznövekmény is egyre nagyobb lesz [9]. Például a jelenlegivel megegyező üzemi frekvencián működő 20, illetve 21 dB-es optimális tölcsérek méretei rendre: 284mm, 223mm, 355mm, illetve 318mm, 251 mm, 455mm. Tehát nagy előírt nyereség esetén érdemes alternatív módszereket alkalmazni az irányítottság növelése érdekében. A dielektrikum lencsékhez képest a nulla törésmutatójú lencse előnye, hogy sík felületű, és C sávban a nyomtatott huzalozású lemezek gyártási technológiájával legyártható.

Az antennát tápláló csótápvonal geometriája miatt az antennában csak az alapl módus van jelen. A tölcserben közelítőleg gömbhullámok terjednek. Ezért az elektromos térerősség az apertúrán vertikálisan polarizált, abszolútértékben az alapl módusnak megfelelően súlyozott. Ezért az alkalmazott metaanyag lencse vizsgálatakor a gerjesztés vertikálisan polarizált síkhullámmal modellezhető.

III. METAANYAG STRUKTÚRÁK HOMOGENIZÁLÁSA

A. Effektív anyagparaméterek kiszámítása transzmissziós és reflexiós adatokból

Egy sík felületű homogén anyag elektromágneses jellemzői: a komplex törésmutató (N_{eff}) és a hullámimpedancia (Z_{eff}) az anyagra merőlegesen beeső, síkhullám gerjesztés esetén meghatározhatóak az anyag effektív vastagsága (d_{eff}) a reflexiós együttható (S_{11}) és a transzmissziós együttható (S_{21}) ismeretében az alábbi összefüggésekből

$$Z_{\text{eff}} = \pm \sqrt{\frac{(1 + S_{11})^2 - S_{21}^2}{(1 - S_{11})^2 - S_{21}^2}} \quad (1)$$

$$e^{iN_{\text{eff}}k_0d_{\text{eff}}} = \frac{S_{21}}{1 - S_{11} * R_{01}} \quad (2)$$

$$N_{\text{eff}} = \frac{1}{k_0d_{\text{eff}}} \left[\Im \left\{ \ln \left(e^{iN_{\text{eff}}k_0d_{\text{eff}}} \right) \right\} + \frac{2m\pi}{k_0d_{\text{eff}}} - i \Re \left\{ \ln \left(e^{iN_{\text{eff}}k_0d_{\text{eff}}} \right) \right\} \right] \quad (3)$$

A további paraméterek: k_0 : a vákuumbeli terjedési tényező, m egész szám. Ezen összefüggéseket egy metaanyag struktúra paramétereire alkalmazva a kérdéses struktúra effektív paramétereit adódnak. Látszik, hogy a törésmutató ($n_{\text{eff}} = \text{Re}\{N_{\text{eff}}\}$) a fenti egyenletekből nem határozható meg egyértelműen. A Kramers- König egyenlet segítségével viszont N_{eff} képzetes része alapján (κ_{eff}) kiválasztható a megfelelő megoldás. A módszerről és megvalósításáról [8]-ban pontos leírás található.

A struktúra effektív vastagsága a geometriából adódik, jelen esetben megegyezik a vastagsággal. Az S_{11} , S_{21} értékek elektromágneses tér szimulációs szoftverrel számíthatók. Erre a célra az időtartománybeli megoldók alkalmazása kézenfekvő, mivel egy futással széles frekvenciatartományra kiszámíthatók a keresett paraméterek.

A publikált eredményekhez elektromágneses tér szimulációs program, illetve az effektív paraméterek kiszámítására a [8]-ban megadott, online elérhető MATLAB program került felhasználásra.

B. Analitikus közelítés

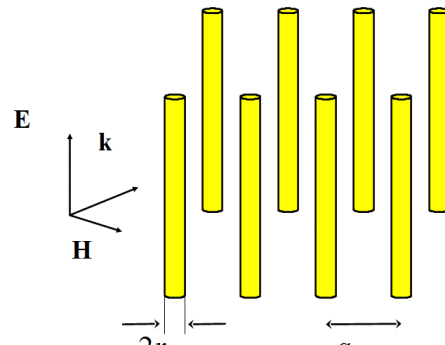
A Drude modellt felhasználva analitikus becslés adható az effektív paraméterekre: az elektronok mozgását egy csillapított oszcillátorral modellezve kiszámítható a vezetőképesség (σ) (4), illetve az anyag komplex permittivitása (ϵ_{eff}) (5).

$$\sigma = \frac{\eta q^2 \tau}{m} \frac{1}{1 + i\tau\omega} \quad (4)$$

$$\epsilon_r(\omega) = 1 + \frac{\omega_p^2}{-\omega^2 + i\Gamma\omega} \quad (5)$$

$$\omega_p = \frac{\eta q^2}{m} \quad (6)$$

η az elektron sűrűség, τ : az elektronok átlagos relaxációs ideje, q : az elektron töltése, m az elektron tömege, $\Gamma = 1/(m*\tau)$. Az ω_p pedig az anyagra jellemző, úgynevezett plazma frekvencia. Az $\omega \ll \omega_p$ esetben az anyag vezetőként, $\omega \gg \omega_p$ esetén pedig dielektrikumként viselkedik, míg $\omega \approx \omega_p$ esetén $\epsilon_{\text{eff}} \approx 0$, ezen a tartományon az anyag nulla törésmutatójú tulajdonságokat mutat. Valódi fémek estén ω_p a látható tartományban van, de megfelelő struktúra alkalmazásával a töltéssűrűség csökkenthető, mely révén ω_p a mikrohullámú tartományba helyezhető át. Erre példa a 2. ábrán látható elrendezés. Ez r sugarú, párhuzamos, egymástól a távolságra lévő hengeres vezetőkből áll, melyek egy négyzetrács sarokpontjaiban állnak. Itt a plazma frekvencia az r/a aránnyal kontrollálható.



2. ábra: Párhuzamos vezető rudakból álló egyszerű metaanyag struktúra, mely tulajdonságai a Drude modell alapján számíthatók

A metaanyag struktúra paramétereinek analitikus becslésére további módszer az elemi cella adott módusú gerjesztés esetére vonatkozó koncentrált paraméterű helyettesítőképének alkalmazása. A 2. ábrán látható elrendezés egy egyszerű soros R-L taggal modellezhető, melyek az anyag ellenállását és a huzal öninduktivitását jellemzik. Ezzel a modellel az egymástól r távolságra lévő párhuzamos w szélességű vezetősávok effektív permeabilitása is becsülhető az $r = w/4$ helyettesítéssel.

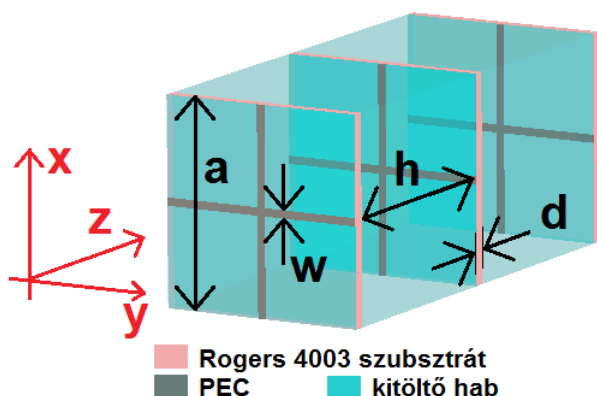
IV. C-SÁVÚ NULLA TÖRÉSMUTATÓJÚ METAANYAG STRUKTÚRÁK TERVEZÉSE

A tervezés célja egy 5,2 GHz sávközépi frekvenciájú, metaanyag lencse paramétereinek meghatározása, illetve tulajdonságainak vizsgálata.

A szimuláció tárgya minden esetben egy elemi cella vizsgálata. Az elrendezést a 3. ábra mutatja. Kettő, illetve háromrétegű lencsék szimulációjára került sor. A lencsék közötti tér kitöltése 1,08 relatív dielektromos állandójú közeg. Külön szimulációban lett megvizsgálva a Rogers 4003-as hordozók hatása. A cellasíkok a z tengelyre merőlegesek, peremfeltételként az x tengelyre merőleges felületen a tökéletes mágneses vezető (PMC), az y tengelyre merőleges felületen

tökéletes elektromos vezető (PEC) lett előírva. A z tengelyre merőleges két port lett a gerjesztés, illetve a nyelő.

A háromrétegű lencse geometriája a 3. ábrán látható, a paraméterek értékei: $h=15,7$ mm, $a=11,5$ mm, $w=0,54$ mm, a szubsztrát vastagsága 0,8 mm a fémzés vastagsága 0,05 mm.

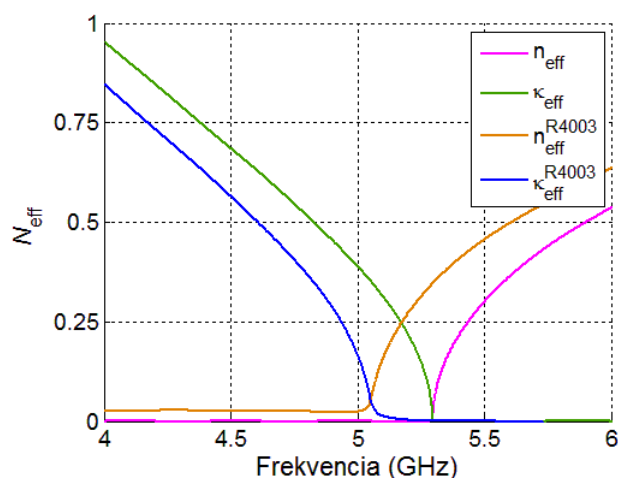


3. ábra: A síkhullámmal gerjesztett, háromrétegű elemi cella geometriája

A 4., 5. és 6. ábrákon a háromrétegű, Rogers 4003 szubsztráttal, illetve az a nélkül szimulált metaanyag jellemzői láthatók. A hordozó elhangolja a metaanyagot, viszont a szubsztrátot is tartalmazó lencse kedvezőbb transzmissziós tulajdonságokkal rendelkezik.

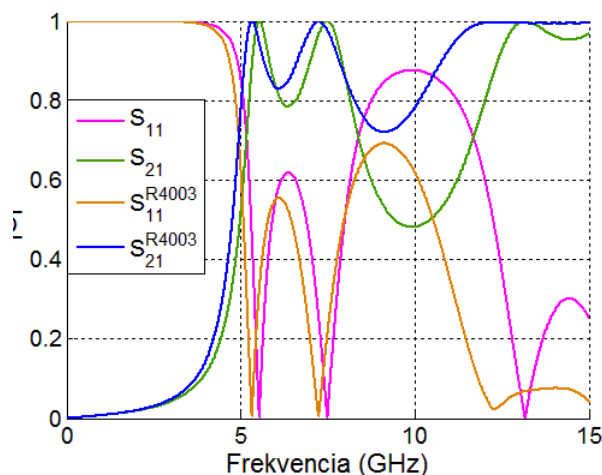
A lencse sáv szélességét egyrészt a transzmissziós tényező határozza meg, másrészt a törésmutató. Ez alapján adható meg az a frekvencia tartomány, ahol a lencse valóban fókuszál. Az S_{11} , és S_{21} paraméterek a 4. ábrán, a metaanyag lencse törésmutatója a 5. ábrán látható. A meghatározott üzemi frekvenciatartományokat az alábbi táblázat tartalmazza, mely alapján az üzemi frekvenciatartomány az 5-5,6 GHz-es sáv.

Rogers 4003 szubsztrátos lencse sáv szélessége	
Előírások	Üzemi frekvencia tartomány [GHz]
$S_{21} < 0,8$ (≈ -2 dB)	$4,8 < f$
$n_{\text{eff}} < 0,5$ & $\kappa_{\text{eff}} < 0,2$	$4,98 < f < 5,6$

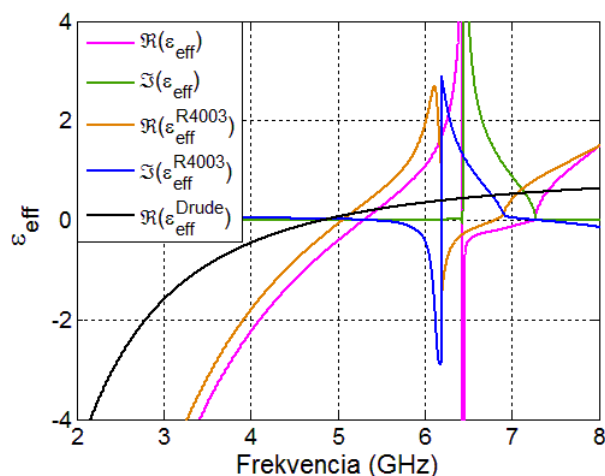


4. ábra: A Rogers4003 szubsztráttal, illetve az a nélkül szimulált (vesztégmentes) háromrétegű metaanyag lencse effektív komplex törésmutatója

A kiszámolt effektív permittivitás és a Drude modell alapján számítható becslés a 6. ábrán látható. A Drude modellel láthatólag jó közelítő becslést lehet adni a lencse plazmafrekvenciájára. A helyettesítőkép elemeinek értékei: $L=10,4$ nH, $R=0,0035$ Ohm - ra adódnak ($5.8e7$ S/m vezetőképességű fémzés esetén).



5. ábra: A Rogers4003 szubsztráttal, illetve az a nélkül szimulált (vesztégmentes) háromrétegű metaanyag lencse transzmissziós és reflexiós együtthatói



6. ábra: A Rogers4003 szubsztráttal, illetve az a nélkül szimulált (vesztégmentes) háromrétegű metaanyag lencse effektív permittivitása

A három-, és kétrétegű, szubsztráttal szimulált metaanyag lencsék tulajdonságainak összevetése a 7., 8. és 9. ábrákon láthatók. Utóbbi paraméterei: $a=12,8$ mm, $h=17,5$ mm, $w=0,6$ mm.

A törésmutatóban nem látszik lényegi különbség, a kétrétegű lencse transzmissziós tényező alapján definiált sáv szélessége viszont kedvezőbb. A lencse sáv szélességét a törésmutató ugyan ebben az esetben is korlátozza, viszont a kedvezőbb transzmisszió révén az elért effektív nyereségnövekedés a kétrétegű lencse esetén nagyobb lehet.

V. KONKLÚZIÓ

Egy metaanyag struktúra került behangolásra, antennás alkalmazásra. A hangolás során figyelembe lett véve a veszteséges hordozó hatása is. Ezen szimulációk alapján 0,6 GHz sávzsélességűre becsülhető a lencse.

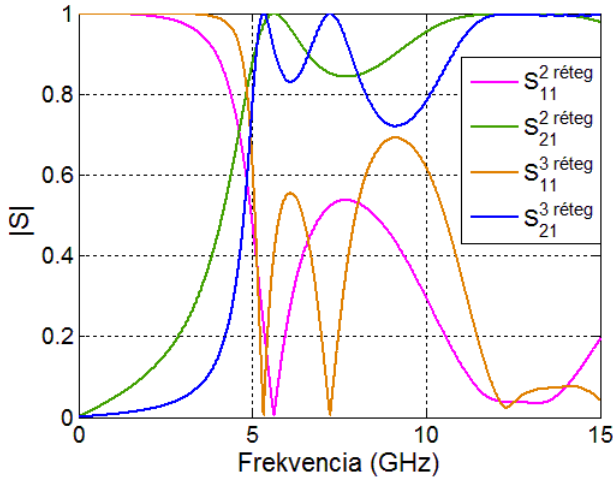
Az itt meghatározott geometriai méretek előreláthatólag jó alapot adnak a végleges paraméterek meghatározására.. Az ezzel kapcsolatos munka folyamatban van, azonban a lencse véges kiterjedése, illetve az antenna maga, a vártnál nagyobb mértékben befolyásolják az eredményeket. E miatt a lencsék tényleges környezetben adódó sávzsélessége nem kielégítő. [6] és [7] eredményeit figyelembe véve az itt tárgyalt elemi cellával optimalizáció, segítségével 5% -10% közötti relatív sávzsélesség lesz elérhető. A későbbiekben, akár más elemi cellát alkalmazva, a 10-20% sávzsélességű, 2 dB nyereségnövekményt biztosító metaanyag lencse geometriájának meghatározása van célul kitűzve.

VI. KÖSZÖNETNYILVÁNÍTÁS

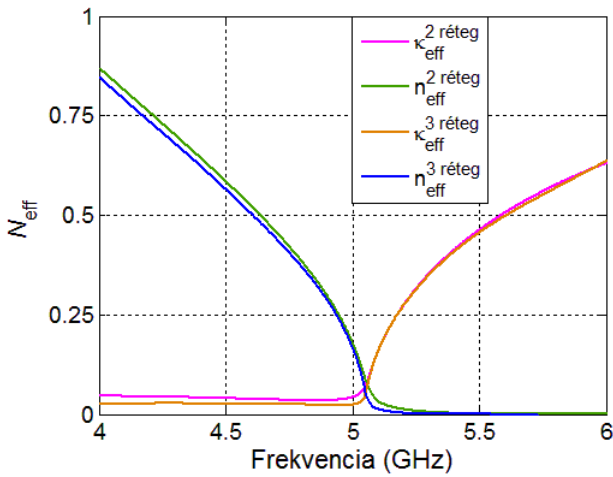
Köszönjük a Magyar Tudományos Akadémia Bolyai János Kutatói Ösztöndíj és az Eureka Metafer projekt támogatását

REFERENCES

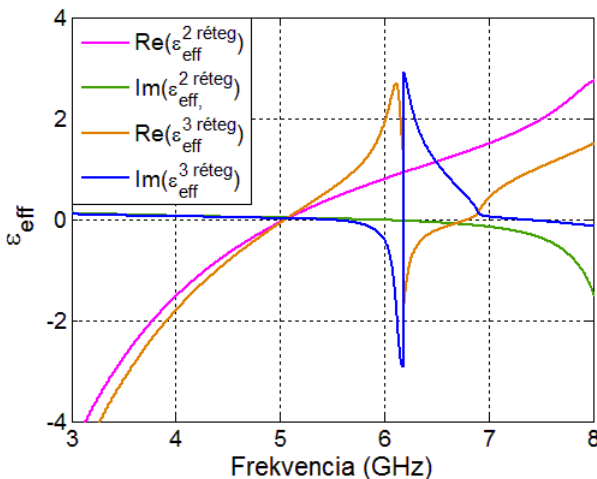
- [1] F.-Y. Meng, Q.Wu, Y. Liang, K. Zhang, and Le-Wei Li "Design, Fabrication, and Testing of Three-Dimensional Miniaturized Rectangular Cavity Resonator Based on Metamaterial" IEEE Transactions On Magnetics, Vol. 45, NO. 10, OCTOBER 2009
- [2] Marco Caiazzo, Student Member, IEEE, Stefano Maci, Fellow, IEEE, and Nader Engheta, Fellow, IEEE "A Metamaterial Surface for Compact Cavity Resonators", IEEE Antennas And Wireless Propagation Letters, Vol. 3, 2004
- [3] Schurig, David, et al. "Metamaterial electromagnetic cloak at microwave frequencies." Science 314.5801 (2006), pp. 977-980.
- [4] Sungjoon Lim, Christophe Caloz, and Tatsuo Itoh "Electronically-Controlled Metamaterial-Based Transmission Line as a Continuous-Scanning Leaky-Wave Antenna" Microwave Symposium Digest, 2004 IEEE MTT-S International (Volume:1)
- [5] Ziolkowski, Richard W. "Propagation in and scattering from a matched metamaterial having a zero index of refraction." Physical Review E 70.4(2004): 046608.
- [6] Dongying Li, Szabó Zsolt, Xianming Qing, Er-Ping-Li, and Zhi Ning Chen, "A High Gain Antenna Wit an Optimized Metamaterial Inspired Substrate" IEEE Transactions on Antennas and Propagaion, vol. 60, no. 12, pp: 6018-6022, Dec. 2012
- [7] Q.Wu , P.Pan, F.-Y. Meng L.-W. Li, J. Wu, "A novel flat lens horn antenna designed based on zero refraction principle of metamaterials" Applied Physics A 87, pp. 151-156, 2007
- [8] Z. Szabó,G.-H. Park, R. Hedge, and E.-P.-Li,"A unique extraction of metamaterial parameters based on Kramers-ronig relationship," IEEE Trans. Microwave Theory Tech. Vol. 58, no.10, pp2046-2653,Oct. 2010.
- [9] John L. Volakis: Antenna Engineering Handbook, McGrawHill, 2007



7. ábra: A két- és háromrétegű metaanyag lencsék transzmissziós és reflexiós tényezői



8. ábra: A két- és háromrétegű metaanyag lencse effektív komplex törésmutatója



9. ábra: A két- és háromrétegű metaanyag lencse effektív permittivitása

Radardetektor tervezése olcsó ISM rádió IC segítségével

Simon Dániel

Silicon Laboratories Hungary Kft.

Szélessávú Hírközlés és Villamosságtan Tanszék
Budapesti Műszaki és Gazdaságtudományi Egyetem

Budapest, Magyarország

simon.dani@mailbox.hu

Zólmay Attila

Silicon Laboratories Hungary Kft.

Szélessávú Hírközlés és Villamosságtan Tanszék
Budapesti Műszaki és Gazdaságtudományi Egyetem

Budapest, Magyarország

zolomy@mht.bme.hu

Kivonat — A közúti közlekedés során rengeteg alkalommal lehet találkozni úgynevezett traffipax-os sebességmérőkkel, melyeknek két fajtája létezik: radaros és optikai elven működő eszköz. A berendezés hatótávolságát megelőzően egy hasonló érzékenységű vevővel (detektorral) lehetséges a mérés érzékelése még a bemérési pont előtt. A radar elven működő sebességmérőknek Magyarországon 3 típusa létezik, ezek 13,45 GHz, 24,125 GHz illetve 34,36 GHz-es frekvencián működnek [5]. A cikk egy 13,45 GHz-es frekvencián működő radar sebességmérőhöz készített költségvetésként detektor felépítését, működését és mérési eredményeit mutatja be. A detektor egy szuperheterodin vevő, mely a radar 13,45 GHz-es jelét lekeveri 920 MHz-es közép-frekvenciára. Amennyiben érzékeli a jelet, úgy a rádiós vevő az RSSI (Received signal strength indication – Vételi jelerősség) érték figyelésével jelzi azt a kimeneten. A lokál jel előállítását egy a vevőhöz hasonló rádiós modulnál történik, amelyben a kimeneti teljesítményerősítő kapcsoló üzemben működik, a kimeneti impedancia megfelelő illesztése esetén így négyszögjel jön létre az illesztőhálózatot követően. Ennek az 5. harmonikusát kihasználva, frekvenciaháromszorozó segítségével áll elő a lokál oszcillátor jele.

Kulcsszavak—Radardetektor, szuperheterodin, ISM, rádiós IC, sebességmérő, szűrő

I. BEVEZETÉS

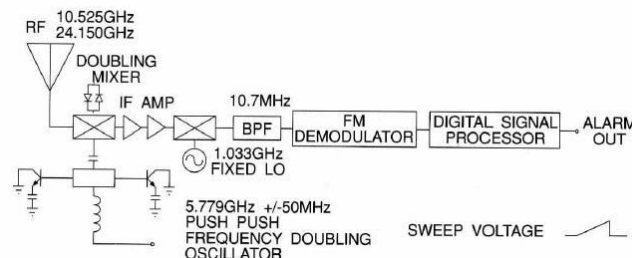
A cikk célja egy egyszerű, olcsó eszközökből álló radardetektor felépítésének és mérési eredményeinek bemutatása. A detektor elsősorban tanulmányi célból készült, a későbbiekben természetesen fel lehet használni a közutakon is.

A 175/2003 kormányrendelet újradefiniálta a közbiztonságra különösen veszélyes eszközöket, és ebből kimaradtak a radar/lézerdetektorok, ezért használatuk 2003 dec. 1-től nem minősül szabálysértésnek, a használója nem büntethető, a detektorok nem kizárhatók el [11].

A radaros sebességmérő a következő elven működik [10]: a berendezés adó egysége kibocsát egy konstans frekvenciájú modulálatlan jelet, mely a jármű felületéről visszaverődve a vevő készülékbe jut. Amennyiben megfelelő erősségű jel érkezik be, a jelfeldolgozó egység a frekvencia Doppler-eltolódása alapján megállapítja a jármű sebességét. A közegben terjedő jel folyamatosan csillapodik, és egy adott távolságon túl a műszer már nem képes a sebesség megállapítására a kis jelszint miatt. Ezen hatótávolságon kívül egy hasonló érzékenységű vevő berendezéssel (detektorral) lehetséges a mérés érzékelése még a bemérési pont előtt. A detektálást megkönnyíti, hogy a radar által kibocsátott jelnek a radar-jármű távolságot oda-vissza meg kell tennie. További előnyt jelent,

hogy a radar folyamatosan bocsát ki magából rádiójeleket, amely sokkal kevésbé irányított, mint lézer esetében. A berendezés hatótávolsága kb. 60 m, ahonnan már hiteles fénykép is készíthető a rendszámról. Kb. 150 m-ről történő detektálás esetén már elegendő idő adódik arra, hogy a bemérési pont előtt lelassítsunk a megengedett sebességre.

Az 1. ábrán látható egy manapság használatos radardetektor blokkvázlata [8]:



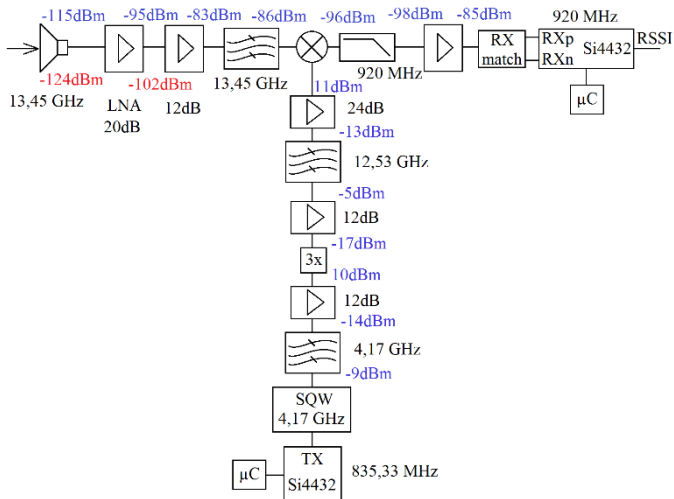
1. ábra: Általános radardetektor blokkvázlat

Az 1. ábra az Amerikában használatos frekvenciasávokra tervezett detektor felépítését mutatja, de természetesen a működési elv ugyanaz más frekvenciákra is. A beérkező radar jelét lekeverik egy 1 GHz körüli frekvenciára, mégpedig úgy, hogy a lokál oszcillátornak az alap- és felharmonikusát is felhasználják a keveréshez. Ezt követően egy fix lokál jel segítségével 10 MHz-es KF (középfrekvenciás) jel áll elő, ami alapján tud jelezni a detektor. A megfelelő működéshez elengedhetetlen a kis veszteségű, drága hordozó alkalmazása.

Az általam tervezett detektor különlegessége egyrészt abban rejlik, hogy ameddig lehetséges, addig olcsó, FR4-es NYÁK lemezzel van megvalósítva az áramkör. A magasabb frekvenciákon csak jelentős csillapítással lehetne megvalósítani szűrőket, ezért ott már kis veszteségű, duroid lemez lett alkalmazva. További különbség az 1. ábrán látható detektorhoz képest, hogy olcsó, szubGHz-es rádiófrekvenciás IC-k felhasználásával épül fel az áramkör, valamint a detektorban a mikrohullámú áramkört elemek száma minimális. Az 1. ábrától eltérően egyszeres keverés történik egy fix lokál jel segítségével, melyet frekvencia tizenötszörözést követően állítunk elő.

II. RENDSZERTERV

A 2. ábrán látható a saját készítésű radardetektor rendszerterve:



2. ábra: Radardetektor rendszerterv

A rendszerterv egy szuperheterodin vevő felépítését ábrázolja. A kézzel ábrázolt értékek a jel, a pirossal ábrázoltak pedig a zaj becslt átlagos teljesítményét adják meg az adott helyen. A 13,45 GHz frekvenciájú radar jelét egy antenna alakítja vezetett hullámmá, melynek átlagos zajteljesítménye 293K szobahőmérsékleten:

$$N = kBT_0 = 1,38 \times 10^{-23} \frac{J}{K} \times 100kHz \times 293K$$

$$= 4,04 \times 10^{-16}W \quad (1)$$

ahol $k = 1,38 \cdot 10^{-23} J/K$ a Boltzmann-állandó, B a vételi sáv szélesség, T_0 pedig a hőmérséklet.

Az átlagos zajteljesítmény dBm-ben kifejezve:

$$N^{dBm} = 10 \times \log N \approx -154 dBW = -124 dBm \quad (2)$$

A vevő áramkört vezérlő mikrokontroller úgy van felprogramozva, hogy a maximális zajszint felett 2 dB-lel lévő küszöbszintet meghaladó jelszint esetén adjon jelzést a vevő a kimeneten. Ez a maximális zajszint az átlagos zajszint felett van kb. 5 dB-lel. Az antennát egy 2 dB zajtényezővel rendelkező LNA (alacsony zajú erősítő) követi annak érdekében, hogy a jel-zaj viszony csak kis mértékben csökkenjen az antennát követő értékhez képest. Az előerősítő maximális kimeneti zajteljesítményére felírható a következő egyenlet [9]:

$$N_{ki}^{dBm} = F^{dB} + G^{dB} + N_{be}^{dBm} =$$

$$= 2dB + 20dB + (-124dBm + 5dBm) =$$

$$= -97dBm \quad (3)$$

ahol F a zajtényező, G az erősítés.

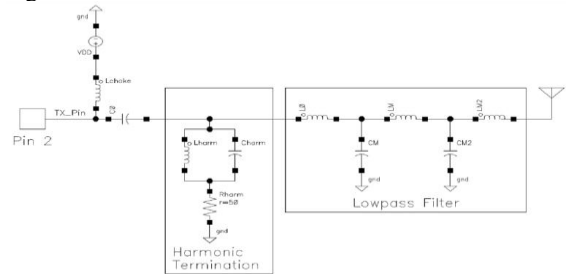
A jelszint az erősítésnek megfelelően 20dB-lel emelkedik, a zajszint pedig (3) alapján 22 dB-lel. Sorba kapcsolt blokkok eredő zajtényezője [7]:

$$F_z = F_1 + \frac{F_2 - 1}{G_1} + \frac{F_3 - 1}{G_1 G_2} + \dots + \frac{F_N - 1}{G_1 G_2 \dots G_{N-1}} \quad (4)$$

ahol F_i és G_i rendre az i . blokk zajtényezője és erősítése. (4)-ből adódik, hogy az eredő zajtényezőt döntően az első blokk befolyásolja, így a további fokozatoknál nem szükséges zajteljesítményt számolni, a jel-zaj viszony az LNA-t követő érték marad. Ha tehát az LNA után a maximális zajszint felett 2 dB-lel van az érzékelési küszöb, azaz ennyi a minimális jel-zaj viszony, akkor az antennát követően a zajtényező miatt 4 dB lesz ez az érték. Az átlagos és maximális zajszint közötti 5 dB-es viszonyt figyelembe véve a vezetett bemeneti érzékenység -115dBm lesz.

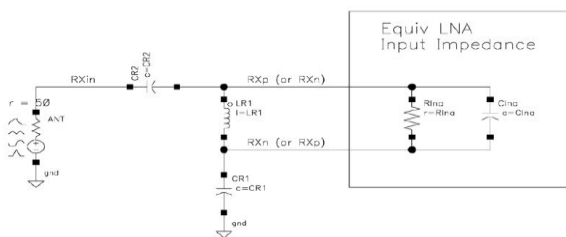
Az LNA után ismét egy erősítő fokozat következik, majd egy 13,45 GHz középponti frekvenciájú sávszűrő, mely a tükörfrekvenciát hivatott kiszűrni. A szűrő kimenő jele kerül rá egy keverő egyik bemenetére, mely a radar vett jelét lekeveri 920 MHz-es középfrekvenciára.

A keverő másik bemenetét egy 12,53 GHz frekvenciájú lokál oszcillátor szolgáltatja, melynek előállítása olcsó ISM (Industrial, scientific and medical– ipari, tudományos és orvosi frekvenciasávok) rádiós chippel történik. Az Si4432 típusú rádiós adó-vevő 240-960 MHz-es frekvenciatartományban üzemel alapesetben. Az IC-ben található teljesítményerősítő kapcsoló üzemben működik, így ezt követően egy a 3. ábrán látható harmonikus illesztő hálózattal a kimeneti hullámforma közel négyszögjel lesz [3]. A Pin2 jelzi az IC adó kimeneti lábát, ahova DC előfeszítést követően a harmonikus illesztő hálózat, majd egy aluláteresztő szűrő kapcsolódik. Ideális esetben a négyszögjel Fourier-sora csak páratlan felharmonikusokat tartalmaz, ezeket nyomja el az illesztő hálózatot követő aluláteresztő szűrő. Ahhoz, hogy maximális harmonikusok legyenek a kimeneten, szűrő nélkül kell vizsgálni a kimenetet. A 835,33 MHz alapharmonikussal rendelkező négyszögjel 5. harmonikusát használjuk ki a lokál jel előállításánál. Ezt a 4,176 GHz frekvenciájú jelet szűrjük, majd erősítjük, egy frekvenciaháromszorozó IC segítségével pedig 12,53 GHz-es jelet állítunk elő belőle. A frekvencia sokszorozása különböző intermodulációs termékeket is előállít a kimenetén, ezért ismét szűrés és erősítés szükséges.



3. ábra: Adóoldali illesztő hálózat

A keverés nemlineáris hatása miatt a lekevert jelben nem csak a 920 MHz-es jel jelenik meg, hanem a két bemeneti jel frekvenciájának összege is, emellett valamikor izolációval a 13,45 és 12,53 GHz-es komponens is továbbítódik. Ezen jelek elnyomására szolgál a kb. 2 GHz törésponti frekvenciájú aluláteresztő szűrő. Ezután egy erősítő, majd egy a 4. ábrán látható vevő illesztő fokozat következik [4], mely az 50 Ω-os tápvonalat az IC két 300 Ω-os bemeneti pontjához illeszti.



4. ábra: Vevőoldali illesztő hálózat

A vevő áramkört vezérlő mikrokontroller úgy van felprogramozva, hogy amennyiben érzékeli a beállított küszöbszintet meghaladó jelet, úgy egy LED és egy piezzo hangszóró jelzi azt.

Mivel a detektor mozgó járműben van elhelyezve, és adott esetben a radar maga is mozoghat, ezért számolni kell azzal, hogy a Doppler hatás miatt nem pontosan 13,45 GHz a beérkező jel frekvenciája. A Doppler frekvencia kHz nagyságrendű, ami a vevő 100 kHz-es sávszélességéhez képest elenyésző, viszont figyelembe kell venni a radar adó berendezésének frekvenciahibáit, valamint a lokál oszcillátor frekvencia stabilitását is. Ezek együttes hatása eredményezheti azt, hogy a 100 kHz-es vételi sávban nincs benne a lekevert jel. Ezért a detektor egy 50 MHz-es sávot fog lefedni, ezt a vételre használt Si4432 frekvenciájának folyamatos változtatásával (sweep) érhetjük el. Azonban figyelembe kell venni, hogy a sweepelés miatt a feldolgozási idő megnő, és szélsőséges esetben eredményezheti azt, hogy a detektor jelzéséig a rendelkezésre állóval összemérhető ideig tart az érzékelés. A vételhez szükséges megadni egy adatsebességet, legyen ez 40 kbps, ami $1/40\text{kbps} = 25\text{ }\mu\text{s}$ bitidőt jelent. Az 50 MHz-es sáv lefedéséhez 100 kHz-es lépésekkel 500-szor kell végrehajtani az RSSI érték kiértékelését, azaz $500 \cdot 25\text{ }\mu\text{s} = 12,5\text{ ms}$ ideig tart a bitek kiolvasása. Az első RSSI kiértékelése 6 bitidőt követően jelenik meg a kimeneten, tehát 75ms időt vesz igénybe a beérkező jel feldolgozása, ami elhanyagolhatóan kicsi a detektálás és a bemérés pont között lévő út megtételéhez szükséges időhöz képest.

III. DETEKTOR HATÓTÁVOLSÁG SZÁMÍTÁSA

A detektor hatótávolságának kiszámításához szükséges tudni a 13,45 GHz-es frekvencián üzemelő radar adóteljesítményét. [5] alapján ez a teljesítmény 13 dBm, a feltételezett antennanyereség pedig legyen 5 dB. A detektor antennájának nyereségét vegyük 2 dB-nek. A vezetett bemeneti érzékenység -115dBm, ebből kiszámítható a jel csillapodása a radartól a detektorig:

$$a_{sz} = P_{ad\acute{o}} - P_{vett} = 13\text{dBm} - (-115\text{dBm}) = 128\text{dB} \quad (5)$$

A szakaszcsillapítás felírható ideális hullámterjedés esetére:

$$a_{sz} = 10 \lg \left(\frac{4\pi}{\lambda} \right)^2 R^2 - G_A^{dB} - G_V^{dB} \quad (6)$$

ahol λ az üzemi hullámhossz, R a távolság, G_A és G_V pedig az adó- és vevőantenna nyeresége.

Valós hullámterjedés esetén azonban [6] alapján a távolság exponense 3 és 4 között változhat a terjedési viszonyoktól függően. R^3 távolságfüggéssel számolva a hatótávolság kb. 463 m, R^4 -nel pedig kb. 100 m. A minimális 100 m hatótávolsággal számolva kiszámítható, mennyi időnk van a bemérési pontig:

$$s = v_0 \times t - \Delta v \cdot t/2 \quad (7)$$

ahol $s = 100\text{ m}$, v_0 a kezdősebességet, Δv pedig a sebességváltozást jelöli. Vegyünk $\Delta v = 30\text{ km/h-s}$ sebességtüllépést. A sebességmérő radarok hatótávolsága 60 m körül van, ezért feltételezhetjük, hogy főleg lakott területen belül alkalmazzák ezeket, míg a km hatótávolságú lézerek inkább autópályákon használatosak. A (7) képlet alapján adódik, hogy lakott területen belül 80 km/h-s sebességről 50 km/h-ra történő lassításra még 5,53 másodpercünk van. Ha figyelembe vesszük az emberi reakcióidőt, akkor is marad elegendő idő a lassításra.

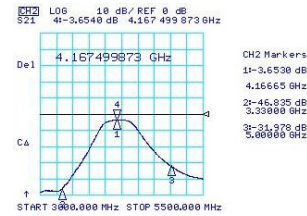
IV. MÉRÉSI EREDMÉNYEK

A. Szűrők mérése

1) 4,17 GHz-es szűrő

A 835,33 MHz-es adó frekvenciára felprogramozott rádiós chip kimenete egy négyszögjel. Ennek az 5. harmonikusát engedi át a 4,17 GHz középponti frekvenciájú csatolt

mikrosztríp tápvonalakból álló szűrő. Az 5. ábrán látható az S_{21} átviteli paraméter:

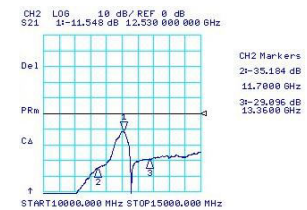


5. ábra: 4,17 GHz-es szűrő átvitele

A mindkét irányban 835,33 MHz távolságra lévő összetevőket láthatóan 30-40 dB-lel elnyomja a szűrő, míg a 4,17 GHz-es komponens kb. 3,6 dB beiktatási csillapítással engedi tovább.

2) 12,53 GHz-es szűrő

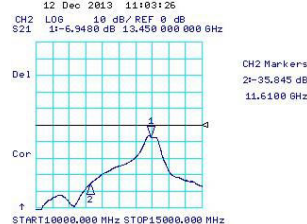
A frekvenciasokszorozó által előállított 12,53 GHz-es jel mellett különböző keverési termékek is megjelennek, ezeket nyomja el ez a szűrő. Felépítése hairpin típusú, a 4,17 GHz-es szűrőhöz hasonlóan szintén csatolt vonalakból épül fel. A 6. ábrán látható az S_{21} átviteli paraméter:



6. ábra: 12,53 GHz-es szűrő átvitele

A szűrő a zavaró összetevőket 30-40 dB-lel elnyomja, 12,53 GHz-en pedig -11,5 dB az átvitele. A mérést két SMA csatlakozóval végeztem, melyeknek méréssel ellenőrizve egyenként 2 dB a csillapítása. Ezt figyelembe véve a szűrő 12,53 GHz-en kb. 7,5 dB-t csillapít.

3) 13,45 GHz-es szűrő



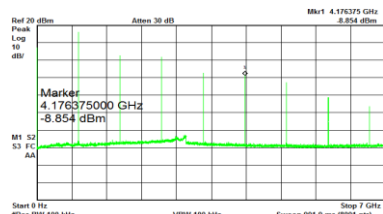
7. ábra: 13,45 GHz-es szűrő átvitele

A 13,45 GHz-es szűrő a radar beérkező jelét engedi át, a 11,61 GHz-en levő tükörfrekvenciát pedig elnyomja a 7. ábrának megfelelően kb. 36 dB-lel. A méréshez használt SMA csatlakozók csillapítását figyelembe véve a szűrő átvitele 13,45 GHz-en kb. -3 dB.

B. Lokál jel mérési eredményei

1) Harmonikus illesztő kimeneti spektruma

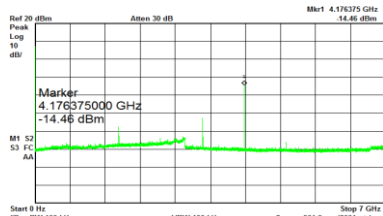
A lokál blokkon belüli Si4432 harmonikus illesztő hálózata úgy lett behangolva, hogy a 4,17 GHz-es harmonikus teljesítménye maximális legyen. Ennek eredménye látható az 8. ábrán.



8. ábra: Harmonikus illesztő kimeneti spektruma

A frekvencia növekedésével egyre meredekebb lefutás szükséges a négyszögletes jelalakhoz, ami 835,33 MHz esetében már csak közelítőleg teljesül. Éppen ezért a páros felharmonikusok is hasonló teljesítménnyel jelennek meg az illesztést követően, mint a páratlanok. A 4,17 GHz-es harmonikus jelszintje kb. -9 dBm, ezt kell egy sávszűrőn tovább engedni.

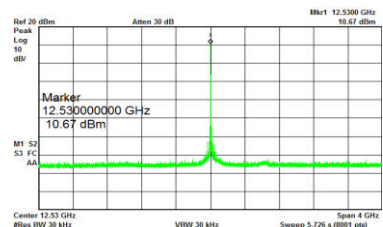
2) 4,17 GHz-es szűrő kimeneti spektruma



9. ábra: 4,17 GHz-es szűrő kimeneti spektruma

A szűrőt követően a 4,17 GHz-es komponens -14 dBm teljesítménnyel továbbítódik, a többit az 5. ábrának megfelelően elnyomja a szűrő.

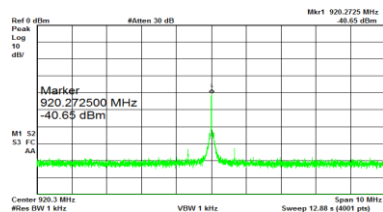
3) Lokál jel



10. ábra: Lokál jel spektruma

Az RMK-3-153+ típusú frekvenciasokszorozó kimenetén a korábbi harmonikusok frekvenciájának különböző lineáris kombinációi megjelennek. A 6. ábrának megfelelően a 12,53 GHz-es szűrő elnyomja a zavaró összetevőket, és előáll a lokál oszcillátor jele 11 dBm teljesítménnyel, ez látható a 10. ábrán.

C. KF jel mérési eredményei

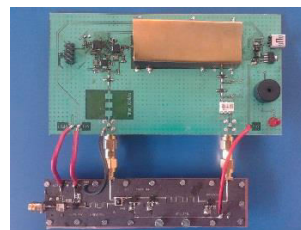


11. ábra: 920 MHz-es KF jel

A keverést követően előáll a 920 MHz frekvenciájú középfrekvenciás jel. Az LNA bemenetére -55dBm teljesítményű jelet vezettem, ennek megfelelően az SMA csatlakozók csillapítását figyelembe véve -40 dBm a lekevert jel teljesítménye.

D. Vezetett bemeneti érzékenység mérése

A részkomponensek mérését követően megvizsgáltam a vezetett bemeneti érzékenységet közvetlenül az LNA bemenetén. Azt tapasztaltam, hogy a rendszertervnek megfelelően kb. -115dBm jelszintet már érzékel a vevő, azaz a vártan megfelelő a működés.



12. Ábra: Az elkészült áramkör

V. ÖSSZEFOGLALÁS

A cikkben ismertetett radardetektort megépítettem, a mérési eredmények igazolják a működést. Egy jelgenerátor és két ismert nyereségű Ku sávban működő antenna segítségével meg lehet mérni a detektor hatótávolságát. A mérést követően le lehet vonni a következtetéseket, szükség esetén lehet növelni a hatótávolságot a vételi sáv szélesség, azaz a bemenetre redukált zajteljesítmény csökkentésével. Ezzel az érzékelési küszöb is lejjebb vihető, így a hatótávolság megnő. Azonban figyelembe kell venni, hogy a jelfeldolgozási idő megnő kisebb vételi sáv szélesség esetén, ezért ez utóbbit nem lehet akármeddig csökkenteni. Másik lehetőség a hatótávolság növelésére a minél nagyobb nyereségű antenna készítése, egyszerű patch antenna helyett antennasor alkalmazása.

További feladatként említhető a másik két frekvenciára egy detektor elkészítése. A nehézséget az okozza, hogy a frekvencia növekedésével egyre nehezebb megfelelő szűrőt készíteni, továbbá a felsőbb frekvenciákon működő áramkörti elemek költsége is jóval nagyobb, így a költséghatékony megoldást nehéz szem előtt tartani.

HIVATKOZÁSOK

- [1] David M. Pozar, "Microwave Engineering," 2nd ed, John Wiley & Sons, Inc., 1997
- [2] Si4432 adatlap: <https://www.silabs.com/Support%20Documents/TechnicalDocs/Si443x-short.pdf>
- [3] AN435 alkalmazási segédlet: <http://www.silabs.com/Support%20Documents/TechnicalDocs/AN435.pdf>
- [4] AN643 alkalmazási segédlet: <http://www.silabs.com/Support%20Documents/TechnicalDocs/AN643.pdf>
- [5] http://www.radar-tech.hu/index.php?option=com_content&view=article&id=67&Itemid=62
- [6] T.S. Rappaport, "Wireless Communications, Principles and Practice," 2nd ed, 2002
- [7] Bíró Viktor, "Mikrohullámú tranzisztors kisjajú erősítők tervezésének néhány kérdése," 1990 http://www.hiradastechnika.hu/data/upload/file/1990/09-10/1990_0910_02.PDF
- [8] Glen D. Martinson, Mark M. Burin, "Radar Detector Technology," BEL-Tronics Limited, Mississauga, Ontario, Canada, 1990 http://www.rfdh.com/ez/system/db/lib_jnl/upload/31/%5BAMW90su%5D_Radar_Detector_Technology.pdf
- [9] Principles of Radar: <http://www.newagepublishers.com/samplechapter/001875.pdf>
- [10] Merrill Skolnik, "Radar Handbook," 3rd ed, 2008
- [11] <http://www.ttkft.hu/jogszabaly.htm>