híradástechnika

1945 VOLUME LXXVIII. 2023

hírközlés ▪ informatika

1

HTE Infokom 2022

A Hírközlési és Informatikai Tudományos Egyesület folyóirata

Tartalom / Contents

Hírközlési és Informatikai Tudományos Egyesület
www.hte.hu
Elnök: Vágujhelyi Ferenc
H-1051 Budapest, Bajcsy-Zsilinszky út 12., 5. em./502.
Tel.: 353-1027 • e-mail: info@hte.hu

Főszerkesztő
SZABÓ CSABA ATTILA (BME, Hálózati Rendszerek és Szolgáltatások Tanszék)

Felelős kiadó: NAGY PÉTER • HU ISSN 0018-2028
Layout: MATT DTP Bt. • Nyomda: FOM Media

A folyóirat támogatói

Nemzeti Együttműködési Alap
Miniszterelnökség
Bethlen Gábor Alapkezelő Zrt.
Emberi Erőforrások
Petőfi Kulturális Ügynökség

www.hiradastechnika.hu

HTE Infokom 2022

2022. november 8-9-én huszonnegyedik alkalommal került megrendezésre a Hírközlési és Informatikai Tudományos Egyesület szervezésében az Infokommunikációs Hálózatok és Alkalmazások Konferencia és Kiállítás; a *HTE Infokom*. A helyszín ezúttal a Four Points by Sheraton Kecskemét Hotel és Konferenciaközpont volt.

Számunk cikkeit az Infokom 2022 előadásaiból válogattuk össze. A cikkek sorrendje követi a konferencia szekcióinak sorrendjét.

A multicast MPEG-TS-alapú műsorszórás lassan háttérbe szorul az adaptív streaming egyre jelentősebb előnyei miatt, ami azonban határozott kihívást támaszt a felhasználószámmal arányosan növekvő sávszélességigényt illetően. *Jursonovics Tamás* (Deutsche Telekom) „Médiaszolgáltatás és CDN egy távközlési szolgáltató szemszögéből” című cikkében bemutatja, hogyan kínál hatékony megoldást ezekre a kihívásokra a Telekom által épített tartalomszétosztó hálózat.

Ulelay Emília (Nemzeti Média- és Hírközlési Hatóság) „6 GHz: Wi-Fi vagy MFCN? – Szabályozói elképzelések a 6 GHz-es frekvenciasáv felső részének hasznosítására” írása a címben jelzett sávrész, a 6425–7125 MHz közötti tartomány jövőjével foglalkozik és felvázolja az ebben a kérdésben folyó nemzetközi és hazai szabályozói tevékenységet.

„Mikor válhat időszerűvé a 2G-mobilszolgáltatás kivezetése Magyarországon?” – kérdezi *Kollár Péter* (Nemzeti Média- és Hírközlési Hatóság) cikkének címével. Válasza: a hazai adottságok és a nemzetközi tapasztalatok alapján az látszik, hogy a 2G kivezetésére Magyarországon 2025-2030 között kerülhet sor, ha a 2G M2M felhasználásának migrációját sikerül megoldani.

Földes Gábor (Vodafone Intelligent Solutions) „Az önálló, ‘igazi’ 5G-mobilhálózatok pénzügyi mozgatórugói” címmel összefoglalja az önálló 5G-hálózatok kiépülésének technológiai-gazdasági mozgatórugóit, és rámutat arra, milyen üzleti tervek mutathatnak pozitív megtérülést.

Kocsis Imre (BME) „Digitális identitáskezelés átalakulóban: önrendelkezésű identitások” cikkében áttekinti az önrendelkezésű identitások (Self-Sovereign Identity, SSI) filozófiáját és legfőbb támogató W3C-szabványait az elosztott identitásokra (W3C) és az ellenőrizhető tanúsítványokra (VC). Demonstrálja az SSI működési modelljét, és bemutatja a legfontosabb, már ma is működő támogató hálózatokat.

A mesterséges intelligencia robbanásszerű fejlődése szinte minden területre hatást gyakorol, így az infokommunikációs hálózatokra is. *Farkas Károly* (BME, Gloster Networks) „Mesterséges intelligencián alapuló eljárások alkalmazása infokommunikációs hálózatokban” című írásában áttekinti ezen eljárások alkalmazásának főbb lehetőségeit az infokommunikációs hálózatok menedzselése kapcsán, valamint bepillantást nyújt a hálózati anomáliák detekciója terén elért saját kutatási eredményeibe is.

Dévai Gergely (Ericsson) és *Recse Ákos* (Ericsson, ELTE) „Szintetikus adatgenerálás – az adathozzáférés szent grálja?” címmel áttekintik a szintetikus adatgenerálás felhasználási területeit és a módszer elméleti alapjait, valamint bemutatják kereskedelmi és szabad szoftvereszközökkel kapcsolatos tapasztalataikat. Cikkük következtetése, hogy egy megfelelően megválasztott szintetikus adatgeneráló eszközzel olyan minőségű szintetikus adat állítható elő, amely akár gépi tanuló modellek tanítására is alkalmas.

Máray Tamás és *Szeberényi Imre* (BME) cikkének – „Neumann Jánostól a HPC-ig” – már a címe jelzi, hogy a szerzők egy fejlődési ívet vázolnak fel; a szuperszámítógépekét. Írásukban áttekintik a szuperszámítógépek (High Performance Computing, HPC) legfontosabb jellemzőit, speciális műszaki megoldásait, a jellegzetes felhasználási módokat és alkalmazási területeket. Szó esik a nemzetközi trendekről, és arról, hogy hol áll Magyarország ezen a területen.

A HTE a *Diplomaterv és Szakdolgozat Pályázat* kategória-nyerteseinek idén is felkínálta a lehetőséget, hogy munkájukról egyoldalas ismertetőt és bemutatkozást készítsenek. Ezekből az összefoglalókból almanach készült, amelyet a jelen különszámba is beszerkesztettünk. A munkákat a pályázat koordinátora, Farkas Károly előszava mutatja be.

Szabó Csaba Attila
főszerkesztő

Médiaszolgáltatás és CDN egy távközlési szolgáltató szemszögéből

JURSONOVICS TAMÁS
Deutsche Telekom
tamas.jursonovics@telekom.de

Kulcsszavak: CDN, streaming, automatizáció, TLS, IPTV

Az over-the-top streaming platformok terjedése új technológiai lehetőségeket teremt az internetszolgáltatók számára TV-rendszereik újragondolásához. A multicast MPEG-TS-alapú műsorszórás lassan háttérbe szorul az adaptív streaming egyre jelentősebb előnyei, az egységes kliensfejlesztés, a célzott reklámok, a WLAN-támogatás mellett, ami azonban határozott kihívást támaszt a felhasználószámmal arányosan növekvő sávszélességigény által. A Telekom által épített CDN bemutatása ezen kihívásokra kínál egy hatékony megoldást nyílt forráskódú szoftverek, konténertechnológia és teljes automatizáció formájában, mely egy extrém skálázható, hibatűrő, költséghatékony és flexibilis CDN-rendszert eredményez.

1. Bevezetés

Az over-the-top (OTT) szolgáltatások térnyerésével párosult innováció előtérbe helyezi az IPTV-szolgáltatók által használt műszaki megoldások időszerűségének és hatékonyságának kérdését. A közvetlen verseny az igény szerinti videó (VoD) területén és az élő műsorszolgáltatás lassú összefonódása a közösségi médiával arra ösztönzi az IPTV-szolgáltatókat, hogy folyamatosan keressék a megújulás lehetőségét helyzetük és szerepük megőrzése érdekében. Ezen törekvés egyrészről üzleti oldalon mutatkozik meg új előfizetői ajánlatok, értéknövelt szolgáltatások és szövetségek formájában, másrészről az IPTV műszaki újragondolásában. Tekintsünk csak a belső és külső hálózatokban használt különböző műsorszóró protokollok, az MPEG-TS multicast és az adaptív streaming által támasztott eltérő követelményekre, mint a lejátszó alkalmazások heterogén fejlesztési környezetére, a nem egységes *personal video recorder* megvalósításokra, a többféle hardverplatformra, és így tovább.

Ahhoz, hogy egy IPTV-szolgáltató felvehesse a versenyt az OTT-megoldásokkal, erőforrásait egy irányba kell összpontosítania. Ezt felismerve, a Telekom a jövőben felhagy a multicast alapú műsorszórással és kizárólag a HTTP-streaming használatára koncentrál, ami nagyban leegyszerűsíti az IPTV felépítését.

Ez azonban azt is jelenti, hogy a multicast által biztosított hatékonyságot nem lehet többé kihasználni, az előfizetőszámmal arányosan nő a műsorterjesztés hálózati erőforrásigénye. Ez a folyamatosan bővülő optikai hozzáférésekkel, illetve egy hatékony, saját tartalomelosztó hálózattal (CDN) ellensúlyozható, melynek mindenképpen ki kell szolgálnia az alábbi szempontokat: skálázhatóság több tíz terabit per másodperces tartományban, alacsony költségek, minimális fejlesztési erőforrások, flexibilitás, könnyű üzemeltethetőség.

A soron következő szakaszok ennek a CDN-nek a tervezése, építése során felmerült kihívásokat és az ezekre adott válaszainkat mutatja be: hogyan alkalmazkodtunk a belső és külső hálózatok eltérő igényeihez, miként tartottuk a fejlesztési erőforrásokat minimálisan, és használunk konténerizációt az egyszerű üzemeltethetőség érdekében, mely architektúra biztosít nagyfokú skálázhatóságot, miként választottuk meg az automatizálás mértékét, milyen lehetőségeink vannak a rugalmas skálázhatóság eléréséhez. Ezek után szó esik a jelenleg tapasztalt problémákról és ezek jövőbe mutató megoldásainak lehetőségeiről.

2. Külső hálózatok

A HTTP-streamingszolgáltatások alapvetően más igényeket támasztanak a tartalomelosztó hálózatokkal szemben, mint a webböngészés. Míg egy weblap megnyitása során az oldalon szereplő képek és scriptek letöltési sebessége nem kritikus tényező, azaz a felhasználók elfogadják, ha egy oldal néha lassabban jelenik meg az esetlegesen ingadozó elérhető sávszélesség miatt, addig tévénézés közben a legkisebb kiesés sem tolerált [1]. Az átviteli technológiától függetlenül az előfizetők elvárják az évek alatt megszokott kábeles/műholdas szolgáltatásminőséget. Hasonlóan, egy igény szerinti videó (VoD) késleltetett indulása még tűrhető, de a filmnézés közben egy esetleges képmegállás elfogadhatatlan.

Mindezek elkerülésére két megoldás terjedt el. Az első esetben – hasonlóan a multicast alapú műsorszóráshoz –, a videóátvitel IP-kapcsolatához rendelt szolgáltatásprioritási szint (DiffServ) biztosítja az elsőbbségi továbbítást, így a saját hálózat megfelelő méretezése mellett a szolgáltatás minősége garantálható. Mobil és Wi-Fi hozzáférés esetén a sávszélesség nem tervezhető, illetve külső hálózat esetén a szolgáltatásprioritás nem elérhető, ezért egyetlen járható út marad: kompromisszum keresése a felhasználói élmény tekintetében [2]. A videó felbontásának csökkenése, vagy tömörítési mér-

tékének növelése még mindig elfogadhatóbb a felhasználók szempontjából, mint a rövid szünetek és megállások, ezért a bevett gyakorlat szerint a multimédiás tartalmak több profilon (más sávszélességeken) is elérhetőek, melyek közül a lejátszó alkalmazás választ az aktuális sávszélesség alapján, ezzel minimalizálva egy esetleges képmegállás vagy újratöltés valószínűségét.

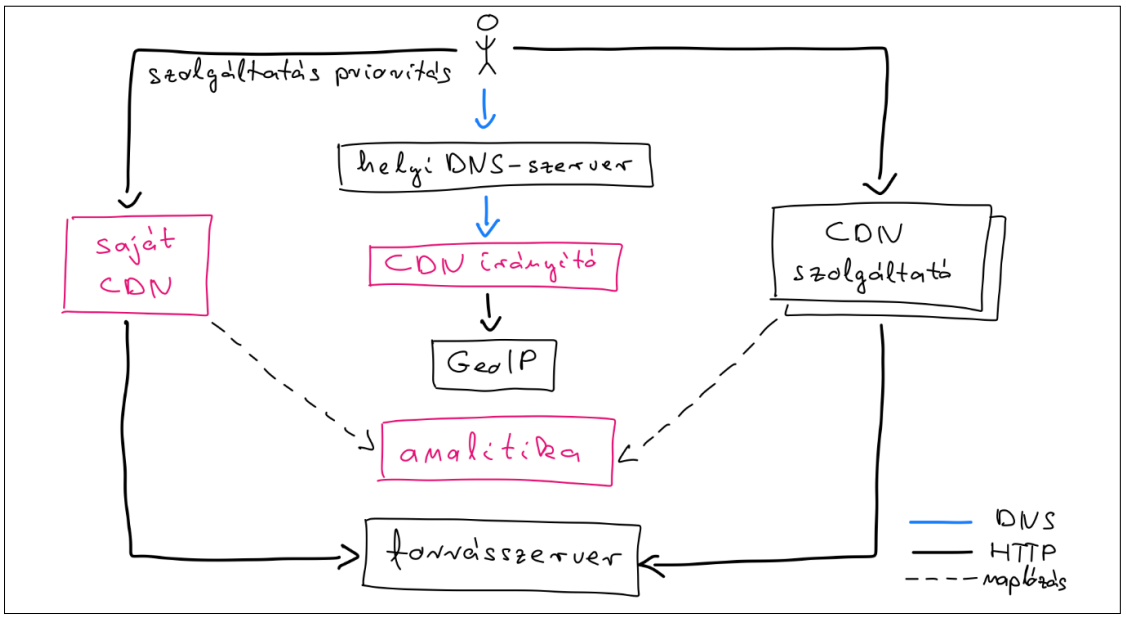
Mint látható, egy internetszolgáltatónak több megoldás is rendelkezésére áll a saját hálózatában a kívánt felhasználói élmény eléréséhez, azonban a külső hálózatok esetében a meghatározó tényező a sávszélesség. Ennek biztosítására az alábbi lehetőségek adódnak: peering kapcsolatok kiépítése vagy bővítése, azonban ez stratégiailag megfontolandó döntés, hiszen ezen kapcsolatok kétirányúak, melyek közvetlen, szabad utat nyitnak a szolgáltató hálózatába harmadik felek számára a peering partneren keresztül. Alternatívaként szóba jöhet CDN-kapacitás telepítése idegen szolgáltatóknál, de a felmerülő komplex üzemeltetési, biztonsági és versenytársi kérdéseket is figyelembe kell venni. Végső és egyben neutrális megoldásként a harmadik fél által biztosított CDN-szolgáltatások igénybevétele (Akamai, Cloudfront) említhető. Utóbbi előnye a tervezési és üzemeltetési feladatok kiszervezése, hátránya azonban a relatív magas díjak, melyek csökkentése érdekében több CDN-szolgáltatóra célszerű támaszkodni a köztük lévő versenyhelyzet kihasználása érdekében.

A Telekom IPTV-szolgáltatásának biztosításához a belső hálózatban IP-priorizálás, a külsőben két CDN-szolgáltató mellett döntöttünk *(lásd az 1. ábrát)*.

Ebből három további feltétel következett. Először lehetővé kellett tennünk, hogy a felhasználók egységesen kommunikálhassanak mindhárom CDN-rendszerrel. Itt elsősorban nem a HTTP-protokoll alkalmazására gondolunk, hanem a CDN-partnerek által előszeretettel ajánlott (és erősen reklámozott) értéknövelt szolgáltatásokra, mint a hitelesítés, jogosultságkezelés, címezhető TV-hirdetés és vízjelezés. Ezek első benyomásra megkönnyítik a szolgáltatás létesítését, azonban – szabadalmaztatott technológiáik révén – erős függőséget alakíthatnak ki a szolgáltató irányába. Ennek elkerülése érdekében a CDN-rendszer tervezése során az alapvető funkciókra koncentráltunk, és minden olyan szerepet, mely nem valósítható meg a HTTP-protokoll által biztosított tágabb keretek között, a CDN-en kívül realizáltuk. Szerencsére ez számunkra, egy IPTV-szolgáltató számára könnyű feladat, hiszen mind a kliensfejlesztés, mind a háttér-infrastruktúra a saját hatáskörünkben van, így a CDN-integrációt szabadon meghatározhattuk.

Másodszor a felhasználókat a megfelelő CDN felé kell irányítani. Ehhez igénybe lehet venni külső megoldásokat (Cedexis Openmix), vagy ahogy a Telekom esetében, egy saját fejlesztésű hiteles névszerverrel (authoritative nameserver) és egy GeoIP-szolgáltatással ez a funkció egyszerűen megvalósítható. A lényeg mindkét esetben, hogy az irányítás DNS-alapon történjen (CNAME mapping), mivel egy extra HTTP-átirányítás és tipikusan a vele járó TLS-kézfogás minden egyes csatornaváltáshoz hozzáadott késleltetése rontaná a felhasználói élményt. Megjegyzem, hogy míg HTTP-átirányítás során a felhasználó internetszolgáltatója (azaz az őt kiszolgáló optimális CDN) jó biztonsággal megállapítható a forrás IP-cím és egy GeoIP-adatbázis (Maxmind, Neustar) segítségével, addig a DNS-alapú irányításnál ez az információ csak akkor érhető el, ha a felhasználó által igénybe vett helyi DNS-szerver (resolver) támogatja az EDNS *client subnet extension*-t [3], mely lehetőséget biztosít a felhasználó IP-címének a CDN irányítója számára történő továbbítására. Természetesen az ISP a helyi DNS-szervereiben ezt egyszerűen aktiválhatja, illetve a Google által szabadon elérhető DNS-szerverek ezt automatikusan támogatják [4]. Mindezek ellenére számolni kell egy alacsony hibafaktorral, hiszen biztosan lesznek olyan felhasználók, akik hátrányukra változtatják meg a DNS-beállításaikat, így a számukra ideális CDN nem meghatározható.

Harmadszor pedig elengedhetetlen, hogy több CDN között dinamikusan irányított felhasználók esetén valós idejű naplózási és analitikai rendszert építsünk. Ehhez



1. ábra
A Telekom CDN-portfóliója

nem támaszkodhatunk az egyes CDN-ek saját megoldásaira, mert egy átfogó, korrelált képet kell alkotnunk az összes felhasználó streaming minőségéről, ezért egy saját rendszer létrehozása mellett döntöttünk. Ehhez jó megoldást biztosít az *ELK Stack/Search Guard,* melyek segítségével képesek vagyunk 60 másodperces késleltetéssel másodpercenként több millió naplóbejegyzés feldolgozására, keresésére, illetve elemzések és grafikonok biztosítására.

3. Fejlesztés

Egy ISP számára a saját hálózatán belüli CDN-szolgáltatás létesítéséhez a lehetőségek széles tárháza áll rendelkezésre [5]. Globális platformszolgáltatóktól kezdve (Akamai, Fastly, Cloudfront), külső fél által telepített és menedzselt CDN-rendszereken át (Qwilt, open caching), CDN-szoftverbeszállítókon keresztül (Synamedia), egészen a nyílt forráskódú rendszerekig (Apache Traffic Control). Természetesen egy teljesen új fejlesztésnek a lehetősége is nyitva áll. A megfelelő megoldás kiválasztása összetett kérdés, azonban a Telekom szempontjait figyelembe véve az *1. táblázat* szerinti értékelést vettük alapul.

Az *unicast TV* alapvető igénye az extrém skálázhatóság több tíz terabites tartományban. Míg ez a szolgáltatási kategóriában probléma nélkül elérhető peering kapcsolatokon vagy közvetlenül a szolgáltató hálózatába telepített infrastruktúrával, addig a menedzselt, beszállítói és nyílt forráskódú CDN-megoldásoknak tipikus problémája a központi, monolitikus kérésirányítás és -felügyelet, mely nehezen alkalmazkodik több milliós felhasználószámhoz. Ez számunkra határozottan az új fejlesztés irányába mutatott.

Egy CDN-szolgáltatás versenyképes alkalomszerű multimédia-továbbítás esetén, mint pl. labdarúgó VB vagy az olimpiai játékok, ahol nagy sávszélességigény csak rövid ideig jelentkezik, azonban a televíziózás során minden este 8 órakor és minden hétvégén megjelenik a sávszélességcsúcs, amely költséghatékonyan csak saját vagy osztott infrastruktúrával biztosítható. Így a szolgáltatói CDN-megoldásokat a saját hálózatban kizártuk.

Fontos szempont továbbá a rendszer flexibilitása, igény szerinti módosítása. Ez külső beszállítók esetében nem kívánt függőséget jelent. Mindig is alapvető és kritikus fontosságúnak tartottuk, hogy a megfelelő kompetencia és továbbfejlesztési lehetőség egy ISP oldalán is rendelkezésre álljon, ezért ebben a kérdésben is a beszállítói megoldásoktól távolodva, a nyílt forráskódú lehetőségek irányba billent a mérleg nyelve.

Egy CDN-rendszer semmiből történő kifejlesztése túlzott befektetést igényelne, ezért a számunkra optimális megoldást végül a nyílt forráskódú rendszerek és a teljes fejlesztés között találtuk meg: nem érdemes mindent nulláról kezdeni, de a nyílt forrás kódú CDN-projektek sem biztosítják a megfelelő fokú szabadságot, ezért köztes megoldásként a Telekom más szolgáltatásaiban már jelen lévő, nyílt forráskódú komponensekre támaszkodtunk. Saját fejlesztéssel ezeket egy CDN-rendszerbe szerveztük, így az erőforrásigényeket minimálisan tartottuk, mindemellett a flexibilitást maximálisan megőriztük (ld. következő szakaszt), nagyfokú automatizáció alkalmazásával pedig a menedzsmentkomponensek fejlesztését elkerültük (5. szakasz).

4. Architektúra

Az architektúrát alapvetően meghatározó szempontok az extrém skálázhatóság, illetve a magas rendelkezésreállás és hibatűrés. Ezek szerepét legegyszerűbben a tipikus CDN-ek problémáin keresztül tudjuk bemutatni. Ahogy azt a *2. ábra bal oldala* mutatja, egy CDN-rendszer fő komponensei közé tartozik a

• *request router,* mely a teljes CDN-infrastruktúra állapotát felügyeli, és DNS A/AAAA-választ adó bejegyzések (1), vagy HTTP 302/307 átirányítás (2) segítségével a felhasználók kérését egy optimális *edge cache*-hez irányítja (3). A cache találati arány növelése érdekében az egy helyre telepített cache-ek tárterületét *consistent hashing* segítségével [6] egy virtuális tárterületté szervezi.

• *Edge cache,* mely magát a tartalom közbülső tárolását és kiszolgálást biztosítja több hálózati helyszínen (point of presence, PoP), ideális esetben közel a felhasználókhoz.

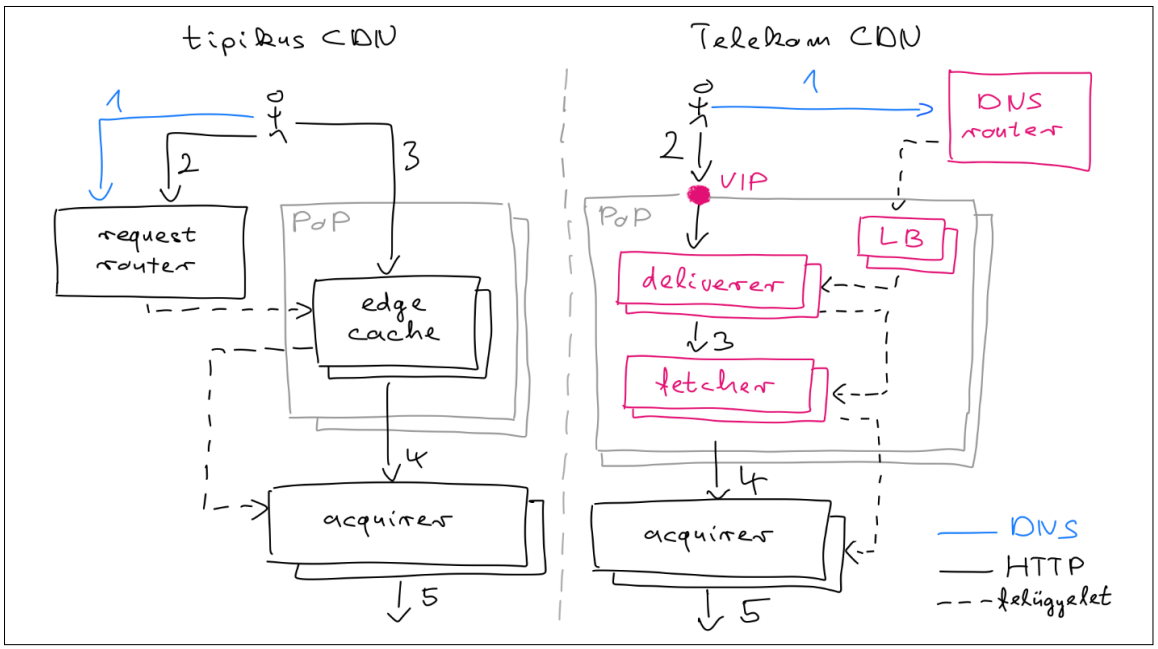
• *Acquirer,* mely feladata egy közbenső tárolási réteg kialakítása az *edge cache*-ek és a forrásszerver *(origin)* között (4, 5), utóbbi további tehermentesítése érdekében.

A legfőbb problémát a request router összetett és központi szerepköre, mely esetleges meghibásodása a teljes CDN-rendszerre hatást gyakorol, illetve ezen szerepkör mind a felhasználók, mind az edge cache-ek számával történő arányos skálázódása jelenti. Ahhoz, hogy a CDN-rendszer extrém és megbízható módon bővíthető legyen, a request router drasztikus egyszerűsítésére volt szükség.

A *2. ábra jobb oldalának* megfelelően, önfelügyelő és hibatűrő PoP-ok létrehozásával a központi felügyeletet elosztottuk a rendszeren. Az egy PoP-hoz tartozó komponensek önmagukban képesek a redundancia és hibajavítás biztosítására minden külső funkció nélkül, en-

1. táblázat
CDN létesítési alternatívák

	szolgáltatás	menedzselt	beszállítói	nyílt forráskódú	fejlesztés
skálázhatóság	+++	++	++	++	+++
költségek	+++	++	++	+	+
flexibilitás	+	+	++	+++	+++
tudás	+	+	++	++	+++
erőforrásigény	+	+	++	++	+++
üzemeltetés	+++	+++	++	+	



2. ábra
Architektúra

nek köszönhetően a request router már nem egyedileg felügyeli az egyes edge cache-eket, elegendő számára csupán a PoP-okra, mint logikai egységekre tekinteni. Így a request router méretezése függetlenedik a CDN méretétől.

Másfelől az edge cache-ek eléréséhez *direct routing* alapú terhelés elosztást [7] alkalmaztunk, mely lehetővé teszi az egyetlen virtuális IP-címről történő szolgáltatásnyújtást, függetlenül a PoP-ba telepített szerverek számától. Ez mindamellett, hogy nagyfokú skálázhatóságot biztosít, a request router szerepkörét tovább csökkenti, így a request router-nek már csak egy alapvető DNS-irányító feladatkört kell ellátnia, mely során a felhasználók kéréseihez (1) egyszerűen a közelükben lévő PoP-ok virtuális IP-címéhez rendeli. A HTTP-átirányítás szerepét veszti. Köszönhetően a DNS-infrastruktúrában alkalmazott köztes tárolási funkciónak, a request router (immár DNS-router), a felhasználók számával sem skálázódik.

Az előbbiek lehetővé teszik a több ezer edge cache, vagy akár a több száz terrabites tartományba való skálázhatóságot, azonban így egy fontos előny is elveszik, mégpedig a consistent hashing által biztosított virtuális tárhely lehetősége. A DNS-irányítás során a felhasználó által elérni kívánt URL, azaz tartalom nem ismert, ezért a DNS-router-nek nincs lehetősége az egyes tartalmak edge cache-ekhez történő rendelésére. Ez egy nélkülözhetetlen funkció a CDN költségeinek alacsonyan tartásához, ezért ennek biztosítása érdekében a Telekom architektúrájában az edge cache-ek szerepét, hasonlóan a Fastly által szervezett PoP-okhoz [8], ketté kellett választani.

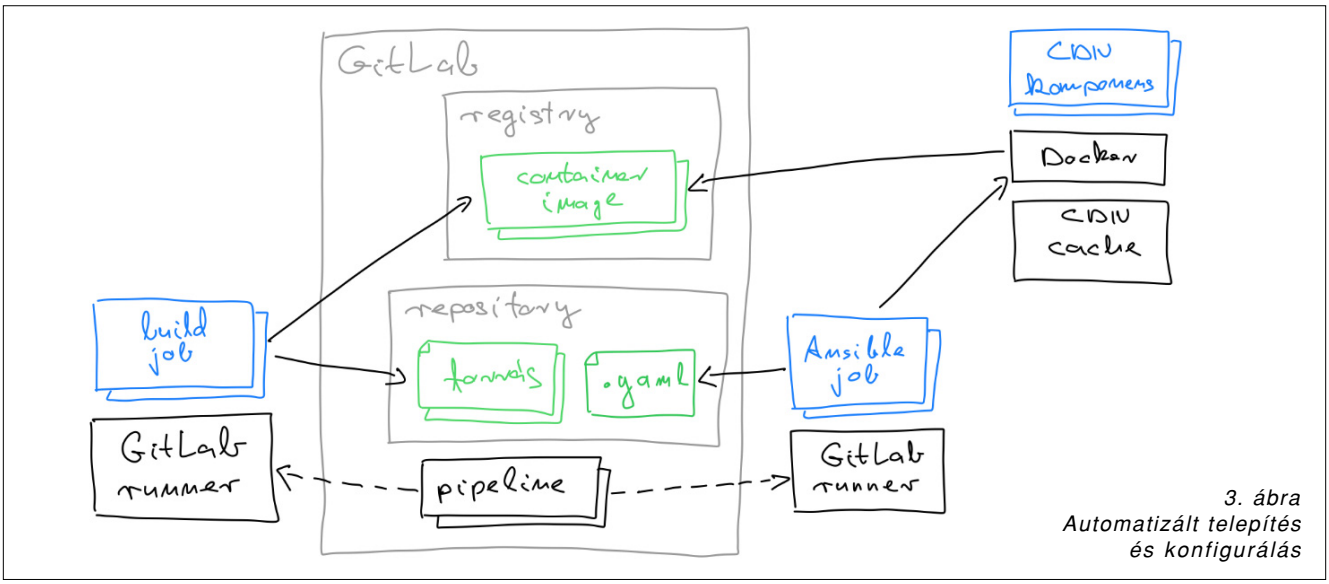
A felhasználókat közvetlenül a *deliverer*-ek szolgálják ki (2), melyek a top tartalmak közül is a top-ot tárolják. Mivel a kéréseket véletlenszerűen rendeljük az egyes deliverer-ekhez, ezért ezen tartalmak többszörösen, minden deliverer-en tárolásra kerülnek, így az erre allokált tárhely alacsonyan van meghatározva. Viszont a deliverer-ek már értelmezik a HTTP-kéréseket, így számukra lehetséges a *consistent hashing* alkalmazása a *fetcher*-ek irányába (3), ahol az egyedi tárhelyek már egy közös virtuális tárhellyé vannak szervezve a magas találati arány fenntartása érdekében. Természetesen e két funkciókhoz nem szükséges külön szervereket fenntartani, egy komponensben megvalósítva, logikailag szétválaszthatóak.

A kompromisszum, ami ezzel az architektúrával jár, az a kérések egy részének kettős kezelése, ami azonban megfelelő *deliverer/fetcher* méretezéssel az IPTV-szolgáltatások esetén elfogadható szinten tartható. A fetcher-ek szerepe és feladatköre nem változott, továbbra is a forrásszerverek tehermentesítését végzik (4, 5).

A rendszer felépítéséhez a következő komponenseket választottuk:

- Operációs rendszer: Ubuntu Linux.
- DNS-router:
  - dnsdist: DNS-terheléselosztás, túlterheléses támadás elleni védelem,
  - PowerDNS: hiteles névszerver,
  - FastAPI: RESt APi szerver a kérésirányítási logika megvalósításához.
- Deliverer, fetcher, acquirer:
  - Varnish Cache: nyílt forráskódú, közbenső tároló (reverse caching proxy),
  - HAProxy: TLS-végződtetés vagy kezdeményezés,
  - keepalived: PoP-terheléselosztás,
  - Python-alapú erőforrás felügyelet.

Az üzemeltetés további egyszerűsítése érdekében az összes alkalmazás Docker alapon, konténerekben fut. Célunk nem egy edge-felhő létrehozása volt, hiszen ez jóval túlmutatna a CDN-rendszer keretein, hanem egy HW/OS-absztrakció megvalósítása. Ez lehetővé teszi a fizikai komponensektől és a gazda operációs rendszertől független patch-menedzsmentet, és a függőségek konténeren belüli egyedi és flexibilis kezelését, mely a rendszer fejlesztését, tesztelését és automatizációját nagyban megkönnyíti.



3. ábra
Automatizált telepítés és konfigurálás

5. Automatizáció

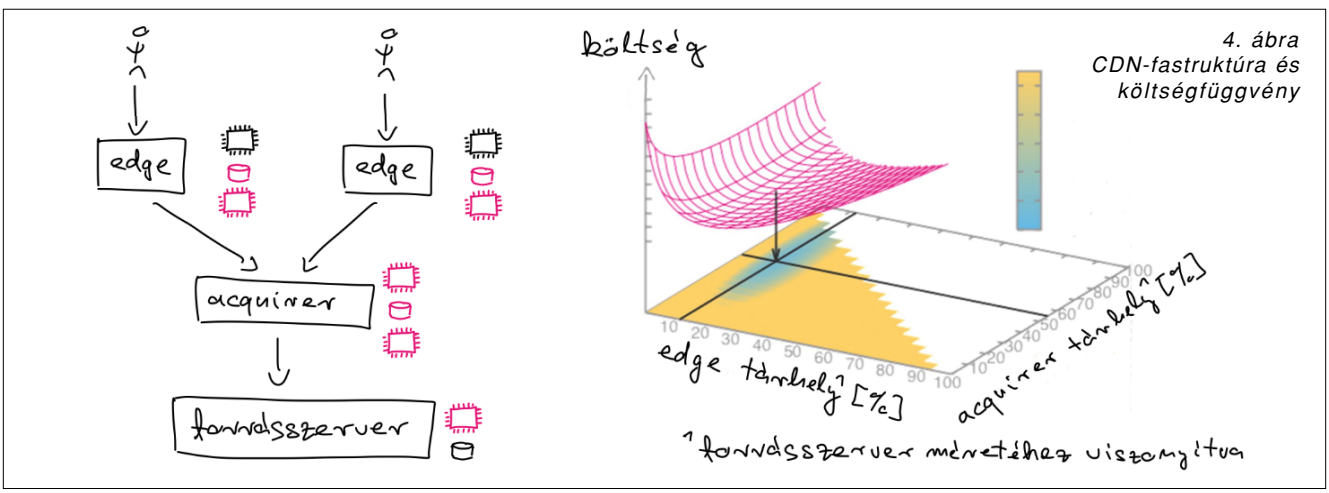
Az *unicast streaming* által igényelt jelentős kapacitás biztosításához nemcsak egy erősen skálázható és robusztus rendszerre van szükség, hanem sok szerverre is. A Telekom CDN-rendszerét több tucat telephelyen több száz szerver alkotja. Ahhoz, hogy egy ilyen méretű rendszer könnyen üzemeltethető, illetve a fejlesztési erőforrásigény alacsonyan tartható maradjon, egy új szemléletre volt szükség a rendszermenedzsment területén: teljes telepítési és konfigurálási automatizációra, illetve a kezelői felület hiányára. Tekintsük át ezeket részletesebben.

Ha egy CDN-rendszer már meglévő szoftverkomponensek integrációjával épül, akkor a legnagyobb fejlesztési erőforrást ezen komponensek távoli konfigurálása és adminisztrációs felülete igényli. Ha jól megnézzük ezeket a feladatköröket, észrevehetjük, hogy az adminisztrációs felületnek a feladata nem más, mint a rendszer tulajdonságainak egy absztrakt leírása, a távoli konfiguráció pedig ezen leírás átfordítása a komponensek egyedi vezérlésére. Erre már léteznek szabadon elérhető, kiváló megoldások, mint Ansible, Puppet, Chef, Salt stack.

A Telekom CDN tervezése során a választásunk a *yaml* formátumra esett, mely könnyen érthető és kényelmes absztrakciót biztosít. Az automatizációra számításba vettük a *Puppet* környezetet, azonban ennek nem túl elterjedt programozási környezete miatt végül az *Ansible* segítségével valósítottuk meg. Ez kiválóan integrálható a Telekomnál rendszeresített *GitLab*-környezetbe és a *Molecule* [9] keretrendszer segítségével teszt automatizációt is lehetővé tesz *(3. ábra)*.

6. Méretezés és rugalmas skálázhatóság

CDN-rendszer tervezése során alapvető kérdés a méretezés és költségoptimalizálás. Dedikált hardver-infrastruktúrát feltételezve ezt csúcsterhelésre kell számolni a cache-ek erőforrásigényei alapján processzor, tárhely (ideértve a memóriát is) és hálózati interfészek mentén. Alacsony igényei miatt a CDN-menedzsment és kérésirányítás, illetve a belső hálózaton olcsón és nagy sávszélességen elérhető hálózati kapcsolatok miatt a hálózati interfészek is elhanyagolhatóak, így a méretezést alapvetően 2 ortogonális dimenzión kell elvégezni a cache-ek processzor- és tárhelyigényei alapján.



4. ábra
CDN-fastruktúra és költségfüggvény

Figyeljük meg a CDN fastruktúrájú felépítését a *4. ábra bal oldalán!* Könnyen megérthető, hogy az *edge processzor* erőforrásigénye egyenes arányban áll az éppen szolgáltatást igénybe vevő felhasználók számával, hiszen egy streaming kapcsolat elsősorban a processzort terheli (TLS-kézfogás, titkosítás, TCP-stack, adatok mozgatása a memóriából/térhelyről a hálókártya irányába). Csúcsterhelésre ez egy jól becsülhető, fix érték. Az *edge cache* tárhely méretezése viszont már egy függő változó, mely elsősorban a cache találati arányt, ezen keresztül pedig a tartalom beszerzéséhez szükséges processzor igényt és az *acquirer*-en megjelenő kapcsolatok számát, azaz az acquirer processzorigényét határozza meg. Hasonló módon belátható, hogy az acquirer tárhely megválasztása az acquirer és a forrásszerver processzorigényét határozza meg. Magától értetődően a forrásszerver tárhelymérete a teljes tartalommal megegyező, mely szintén előre, jól becsülhető, fix érték.

Ebből következően a CDN méretezése során feladatunk ezen szabad paraméterek optimális megválasztása, mely nem triviális feladat, eléréséhez két dolog szükséges. Egyfelől meg kell határozni a cache találati arány értékét a cache tárhely méretének függvényében. Ez a tartalomfogyasztási szokásoktól és a cache-ben implementált *eviction*-algoritmustól függ, azonban az esetek nagy többségében ez nem írható le zárt algebrai formában. Ezen dilemma feloldása érdekében a CDN méretezése során egy parametrizálható cache-modellt alkottunk Python-alapokon, mely alkalmas a tartalomfogyasztási statisztikák alapján a bejövő forgalom becslésére. Ezen egyedi cache-modellek hálózatba szervezésével a teljes CDN-rendszer szimulálhatóvá válik, így már meg lehetett becsülni a szükséges erőforrásigényeket, azaz a CDN költségét egy választott szabad paramétercsoportra.

Másfelől a CDN költségének minimalizálásához optimálisan kell megválasztani ezen paramétereket. Ezt egyszerűen egy Monte-Carlo-szimulációval oldottuk meg, mely segítségével minimalizálható a CDN költségfüggvénye *(lásd a 4. ábra jobb oldalán)*. Erről leolvasható a teljes költség különböző *edge* és *acquirer* tárhely tükrében. Az ábra által szemléltetett példa-CDN-konfigurációról jól megállapítható, hogy a CDN költségének minimalizálása érdekében az edge tárhely méretét 12%-ban, az acquirer-t pedig 54%-ban kell választani, valamit a költség érzékenyebb az edge tárhely változására, mint az acquirer-ére.

A fenti gondolatmenetet folytatva érdekes kérdés e költségmodell időbeli alakulása. A fent vázolt megoldás egy pillanatnyi optimumot ad, azonban a felhasználók tartalomfogyasztása, így a CDN-rendszer erőforrásigénye is időfüggő *(5. ábra)*. Az eddigi méretezés jól alkalmazható dedikált infrastruktúra esetén, azonban, ha egy edge felhő által adott lehetőségeket szeretnénk kihasználni, vagy a CDN-rendszer energiafogyasztását is számításba szeretnénk venni, akkor további költségcsökkentést érhetünk el az egy időpontban szükséges erőforrások igény szerinti felszabadításával vagy lekötésével. Ennek modellezésére további tényezők figyelembevételére van szükség. A tárhely felszabadítása során a cache találati arány csökkenni fog, ami megnöveli a relatív tartalombeszerzési sávszélességet, azaz az alsó rétegek terhelése nő, illetve tárhelyfoglalás esetén az újonnan kapott tárhely „üres”, ennek tartalommal való feltöltése ideiglenesen szintén rontja a találati arányt. Ezek függvényében a fent bemutatott cache modell átdolgozására van szükség, melynek képesnek kell lennie a felhasználók tartalomfogyasztási szokásainak előrejelzésére.

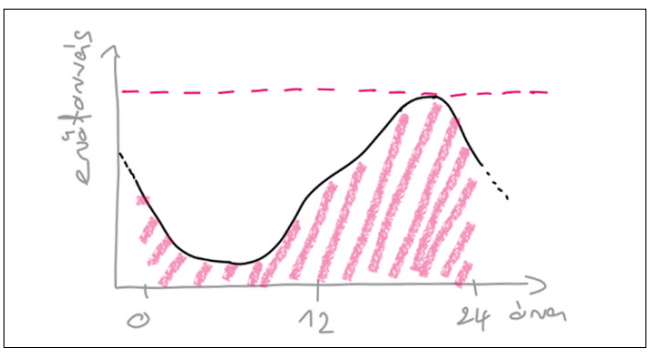
7. Továbbfejlesztés

A nyílt forráskódú komponensekre épülő rendszereknek folyamatosan alkalmazkodniuk kell a változó hardver- és szoftverkörnyezethez. Ez a Telekom CDN-rendszerével is így van, a jelenlegi funkcionalitás az elmúlt öt év során alakult ki az újonnan megnyíló (és rég elfeledett, de újra felfedezett) lehetőségek mentén. A cikk írása időpontjában két fontos lehetőséget említünk meg, melyek kiaknázását fontosnak tartjuk.

Elsőnek a *Varnish Cache* legfőbb hiányosságát, a tartós (nem felejtő) tárhely használatának hiányát említem. A Varnish Cache elsősorban egy kiváló RAM-cache, azonban a memória túl drága nagy méretű videótár hatékony tárolásához. Szerencsére, ebben a témakörben előre tudtunk lépni, és jelenleg tesztelés alatt áll a nemrég nyilvánosságra hozott „buddy” és „fellow” *storage plugin* [10], mely az *io_uring* könyvtár [11] alkalmazásával ezt a problémát orvosolja.

Másodiknak a TLS implementálásának hiányát emeljük ki. Ez, ahogy az architektúrából is látszik, könnyen áthidalható egy TLS-terminátor segítségével, azonban ebben az esetben elveszik az IP-réteg közvetlen és dinamikus kezelésének a lehetősége, azaz az edge cache nem látja közvetlenül a TCP-kapcsolatokat és ezek paramétereit, statisztikáit. Valamint nem tudja közvetlenül az L3/L4-es protokollokat vezérelni, úgymint dinamikus szolgáltatásosztályhoz (ToS) rendelést, vagy a TCP *congestion control* meghatározását. Célul tűztük ki, hogy a nyílt forráskódú Varnish Cache is képes legyen a TLS natív támogatására és terminálisára, ami egyben megnyitja a lehetőséget a kernel TLS használatához. Ezzel a technológiával reményeink szerint a processzor szignifikánsan tehermentesíthető [12].

5. ábra CDN időbeli erőforrásigénye



Mindezek mellett egy ellenpéldát is szeretnénk felhozni. Az elmúlt évek során a hardverplatform-kapacitás lendületes fejlődésének lehettünk tanúi: Intel Scalable Gen 4, PCIe5, 400GE hálózati kártyák, a több mint 100 magos processzorok is könnyen elérhetővé váltak. Sorra dőlnek meg az egy szerverrel kiszolgálható sávszélesség-rekordok [13], ugyanakkor egy CDN esetében nem biztos, hogy érdemes ezt a trendet követni, és a lehető legtöbb kapacitást egy cache-ben összpontosítani. Gondoljunk csak arra, hogy egyetlen, 800 Gbit/s-ot kiszolgáló szerver kiesése, még UHD-videó sávszélességét feltételezve is, egyidejűleg 50.000 felhasználóra lenne hatással. IPTV-szolgáltatás nyújtása során maximálisan törekedni kell a felhasználói élmény megőrzésére, ezért – meglátásunk szerint – az új technológiák megjelenése inkább a kompaktabb szerverek felé fogja terelni a CDN-infrastruktúrát. Az egy szerverrel kiszolgált felhasználók száma ezzel tetőzni fog.

8. Összefoglalás

Az előző szakaszokban bemutattunk egy nyílt forráskódú komponensekre épülő, skálázható, flexibilis és könnyen üzemeltethető tartalomelosztó hálózatot, mely moduláris felépítése és nagyfokú automatizációja révén költséghatékony megoldást biztosít streamingszolgáltatások nyújtásához. Azonban cikkünk igazi lényege nem ez, hiszen a felhasznált komponensek szabadon elérhetők, az alkalmazott technológiák zöme pedig több mint húszéves múltra tekint vissza, – azt is mondhatnánk; nincs itt semmi különös látnivaló!

Amit át szeretnénk adni az olvasónak, az egyfelől a nyílt forráskódú szoftverekben rejlő potenciál felismerése. Meglátásunk szerint ma már elértük azt a szintet, hogy egy közösség által készített és támogatott szoftverekkel igenis lehet nagy megbízhatóságú rendszereket építeni és akár kritikus szolgáltatásokat biztosítani, ha nyitottak vagyunk e közösségekben önzetlenül részt venni és a közös célt sajátunknak tekinteni.

Másfelől, – de nem utolsósorban –, a telko-világban a kreativitás fontosságát és szerepét szeretnénk kiemelni. Sajnos, az eddigiekben sok olyan helyzettel szembesültünk, amikor a szaktudás és kockázatvállalás hiánya miatt egy adott beszállítóra túlzott módon hagyatkoznunk kellett, ami nehezen feloldható függőséget és további tudásvesztést okozott. El kell kerüljük az ilyen helyzeteket, lehetőséget és teret kell engednünk a tanulásra, valamint vállalnunk kell néhány hibánk következményét, mert csak így fejlődhetünk és teremthetünk olyan megoldásokat, melyek maradéktalanul szolgálják saját céljainkat. Egy beszállító valamiyen szempontból mindig a saját érdekét fogja szem előtt tartani...

„Mérnökök vagyunk,
merjünk szabadon alkotni és építeni!”

Hivatkozások

[1] K. Kerpez, D. Waring, G. Lapiotis, J.B. Lyles and R. Vaidyanathan, “IPTV service assurance,” IEEE Communications Magazine, Bd.44, Nr.9, pp.166–172, 2006.

[2] S. Michael, S. Egger, M. Slanina, T. Zinner, T. Hoßfeld and P. Tran-Gia, “A survey on quality of experience of HTTP adaptive streaming,” IEEE Communications Surveys & Tutorials, Bd.17, Nr.1, 2014.

[3] C. Contavalli, W.v.D. Gaast, D.C. Lawrence, W. Kumari, RFC 7871 – Client Subnet in DNS Queries, 2016.

[4] Google, “EDNS Client Subnet (ECS) Guidelines,” [Online]. *https://developers.google.com/speed/public-dns/docs/ecs*

[5] D. Rayburn, “Updated List Of CDN Vendors and History of All CDNs To Date,” 16.01.2023. [Online]. *https://www.streamingmediablog.com/2023/01/cdn-list.html*

[6] D. Karger, E. Lehman, T. Leighton, R. Panigrahy, M. Levine and D. Lewin, “Consistent hashing and random trees: Distributed caching protocols for relieving hot spots on the world wide web,” 1997.

[7] W. Zhang, “Linux virtual server for scalable network services,” in Ottawa Linux Symposium, 2000.

[8] Fastly, “Clustering,” [Online]. *https://developer.fastly.com/learning/vcl/clustering*

[9] Red Hat, Inc., “Ansible Molecule,” [Online]. *https://molecule.readthedocs.io/en/latest/*

[10] N. Goroll, “Two new storage engines for Varnish-Cache,” 06.02.2023. [Online]. *https://varnish-cache.org/lists/pipermail/varnish-announce/2023-February/000757.html*

[11] “Efficient IO with io_uring,” [Online]. *https://kernel.dk/io_uring.pdf*

[12] S. Shwartsman, T. Jursonovics, “Kernel TLS and TLS hardware offload,” [Performance]. EuroBSDCon, 2019.

[13] D. Gallatin, T. Jursonovics, “The ‘other’ FreeBSD optimizations used by Netflix to serve video at 800Gb/s from a single server,” [Performance]. EuroBSDCon, 2022.

A szerzőről

JURSONOVICS TAMÁS a Budapesti Műszaki Egyetem Villamosmérnöki és Informatikai Karán szerezte diplomáját (2003) és védte meg doktori disszertációját (2014) IPTV-rendszerek témakörben, illetve a Hochschule Darmstadt – University of Applied Sciences-en szerzett MBA-fokozatot (2015). 2003-ban csatlakozott a Deutsche Telekom csoporthoz, kezdetben a mobil TV-rendszer kiépítésén dolgozott a Westel900/T-Mobile Magyarországnál, majd több, a Magyar Telekomnál vezetett nemzetközi IPTV integrációs projekt után 2011-től Németországban, a Deutsche Telekom-nál fejleszti a Magenta TV szolgáltatást támogató CDN-rendszert. Számos médiatartalom-átvitellel kapcsolatos publikáció szerzője, az ELTE Adattudományi és Adattechnológiai Tanszékével közös CDN-optimalizációs kutatások aktív résztvevője, az MTA köztestületi tagja, elkötelezett híve és szószólója a mérnöki tudás és tapasztalat népszerűsítésének.

6 GHz: Wi-Fi vagy MFCN?
Szabályozói elképzelések a 6 GHz-es frekvenciasáv felső részének hasznosítására

ULELAY EMÍLIA
Nemzeti Média- és Hírközlési Hatóság
ulelay.emilia@nmhh.hu

Kulcsszavak: ITU, CEPT, rádiószabályzat, 6 GHz, Wi-Fi, 5G, MFCN, IMT, WRC-23

A 6 GHz-es frekvenciasáv használatának lehetőségéért verseny indult, több rádiószolgálat, technológia is számít, pályázik rá. Különösen a felső sávrész (6425–7125 MHz) jövője kérdéses. A frekvenciasáv alsó felében (5945–6425 MHz) vezeték nélküli hozzáférési rendszerek – többek között rádiós helyi hálózatok (WAS/RLAN-ok) – megvalósítására nyílt lehetőség a közelmúltban, de a felső sávrész sorsa bizonytalan. Egyben talán biztosak lehetünk; a jövő mobilkészüléke, melyet pár év múlva kezünkbe vehetünk, a 6 GHz-es frekvenciasávot is elérhetővé teszi így vagy úgy.

1. Bevezetés

A Nemzetközi Távközlési Egyesület (International Telecommunication Union, ITU) közelgő 2023. évi Rádiótávközlési Világértekezlete (WRC-23) napirendjén szerepel a 6425–7025 MHz és 7025–7125 MHz frekvenciasávok – melyeket együtt a 6 GHz-es frekvenciasáv felső részének emlegetünk – nemzetközi mozgó távközlés (IMT) célú azonosítása. Ez a közelgő esemény, a globális harmonizációt szolgáló döntéshozatali fórum előkészítése felszínre hozta, sőt kiélezte két, a szélessávú adatátviteli kapacitás területén egymással versengő műszaki terület igényeinek ütközését. Egy biztos: mind a WAS/RLAN-eszközök, mind a cellás mobilrendszerek gyártói, hálózatüzemeltetői és nagy felhasználói csoportjai igyekeznek a végfelhasználók adatátviteli igényeinek kielégítésében többlet-erőforrásra szert tenni. Elcsépeltnek tűnik a kifejezés, miszerint a korlátos erőforrás jelen esetben nem csupán korlátos, hanem szűkös, de ez most valódi fejfájást okoz a rádióspektrum szabályozásával foglakozóknak, jelenleg elsősorban nemzetközi színtéren.

A WRC-23 döntésének a globális harmonizáció a tétje, hisz regionálisan vagy akár országonként is lehetséges egyedi szabályozási környezetet kialakítani, míg a szomszédokat nem zavarja az országon belüli használat, de méretgazdaságossági szempontból mindenképpen a lehető legszélesebb harmonizált szabályozás a cél.

Az ITU-szintű, azaz a világszintű azonosítás nem elégséges, további műszaki harmonizációs szabályok alkotására lesz szükség. Elsősorban CEPT-, valamint uniós szabályozási tevékenység indult, azaz először ezekben a nemzetközi szervezetekben kell majd a tagállamoknak egyetérteniük. A CEPT és EU szabályalkotása esetén a magyar jogszabályok megfelelő módosítását követően kerülhetnek a hazai piacra az új, szélessávú adatátviteli igényeket kielégítő, hozzáférést biztosító eszközök, melyekben már a 6 GHz-es frekvenciasáv felső része is támogatja majd az adatátviteli szolgáltatásokat.

Mindenesetre ez az eset tökéletes példája annak, hogy a döntéshozatal miként történhet, ha egymással versengő technológiák jelennek meg olyan frekvenciasáv használatára, ahol mindkét használati módra nem áll rendelkezésre elégséges frekvencia. E cikk célja – a döntéshozatali folyamatok áttekintése mellett azt is megvizsgálni –, hogy a 6 GHz-es frekvenciasáv felső részében a WAS/RLAN-eszközök, vagy a cellás mobilrendszerek lesznek elérhetők a szabályozási folyamat eredményeként a publikum számára.

2. A WRC-23 1.2 napirendi pontja

A rádióspektrum-használatra vonatkozó nemzetközi szabályokat az ITU Rádiószabályzata[1] határozza meg, amely az ITU igazgatási szabályainak egyike, így valamennyi tagállam köteles azt betartani, alkalmazni. A Rádiószabályzat határozza meg, hogy az egyes ITU-körzetekben mely frekvenciát mely rádiószolgálat használhatja, és a használat alapvető közös szabályait is rögzíti. A szabályzat alapvetően a tagállamok egymás közötti viszonyára hat ki. A Rádiószabályzat módosítására kizárólagosan a Rádiótávközlési Világértekezlet jogosult.

A Rádiótávközlési Világértekezlet a rádiótávközlés globális és azon belül európai fejlődését évekre meghatározó, kiemelkedő jelentőségű eseménye. Az ott elfogadott döntések utat nyitnak egyrészt a meglévő rádiószolgálatok kibővítéséhez, zavartalan működésének biztosításához, másrészt új technológiák és alkalmazások kifejlesztéséhez és bevezetéséhez. Az ITU Alapokmánya alapján a Rádiótávközlési Világértekezletre 2-3 évente kerülne sor, a gyakorlatban négyévente gyűlnek össze a világ spektrumgazdálkodói, hogy globális szinten rögzítsék a rádióspektrum és a műholdas erőforrások használatának szabályait. A rádióspektrum-gazdálkodásban, a szakmai vizsgálatokban ez a négy éves ciklus meghatározó. A rádióspektrum-gazdálkodás területén vala-

1 Radio Regulations, RR, azaz Rádiószabályzat, hazai szabályozókban: Nemzetközi Rádiószabályzat.

mennyi nemzetközi szabályozással foglakozó szervezetben, így a CEPT és az EU megfelelő formációiban is kialakításra kerül a Világértekezlethez igazodó szervezeti struktúra is, és az egyéb munkafolyamatok mellett a feladatok ütemezésére is befolyással van a Rádiótávközlési Világértekezlet. Az adott Világértekezlet fogadja el a következő Világértekezlet napirendjének tervezetét, amely terjedelmében is tekintélyes, nem is beszélve annak tartalmáról és a döntésekhez szükséges előkészítő munkákról. Egy-egy Világértekezleten hozzávetőlegesen 180 tagállam 3000 szakértője vitatkozik egymással négy héten keresztül.

A soron következő Rádiótávközlési Világértekezlet 2023-ban kerül megrendezésre Dubaiban (Egyesült Arab Emirátus), 2023. november 20. és december 15. között. A szakmai előkészítő munka az előző, azaz a WRC-19 eredményeként elfogadott, WRC-23 napirendjére vonatkozó döntésére (811. (WRC-19) Határozat) alapozva kezdődött a WRC-19 utolsó napját követően szinte azonnal, a Konferencia Előkészítő Értekezlet első ülésén (Conference Preparatory Meeting, CPM-1). A végleges napirendet az ITU Tanácsa 1399. számú határozatában állapította meg. A napirend a hozzá kapcsolódó döntésekkel az ITU hat hivatalos nyelvén bárki által elérhető[2].

A WRC-23 napirendjéből az egyik legvitatottabb elem a felső 6 GHz-es frekvenciasávval is foglakozó 1.2 napirendi pont a 245. (WRC-19) Határozattal összhangban, amelynek csak egy szelete a 6425–7025 MHz és 7025–7125 MHz sávok nemzetközi mozgó távközlés (IMT) célú azonosítása, beleértve az esetleges járulékos felosztást elsődleges jelleggel a mozgószolgálat számára. A döntéshez szükséges megosztási és összeférhetőségi vizsgálatokat az ITU megfelelő munkacsoportjai végzik.

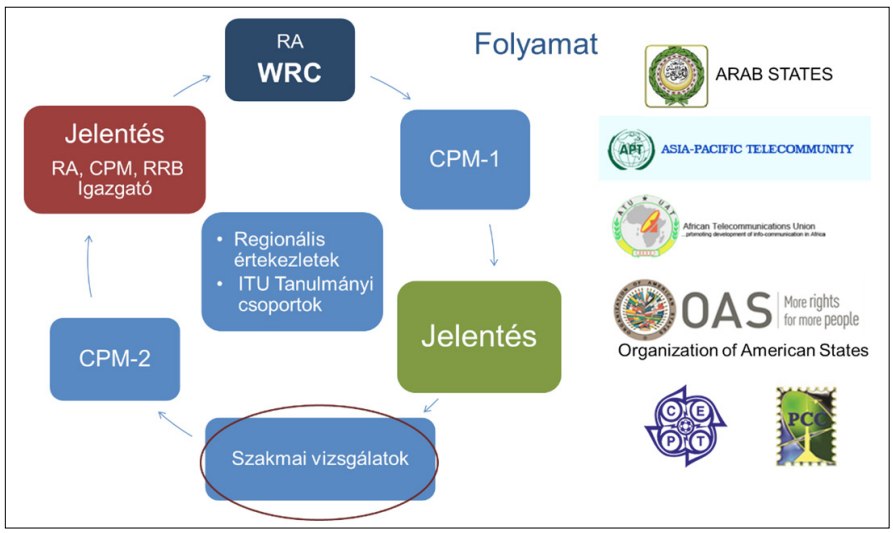
A WRC-khez szorosan kötődnek a Konferencia Előkészítő Értekezletek (CPM), melyből az elsőt közvetlenül a WRC után tartják. Ezek a CPM-1-ek azt célozzák, hogy a WRC-én elfogadott következő WRC napirendjének tervezete alapján felállítsák azt a szervezeti struktúrát, amiben a napirendi pontokban szükséges döntéseket megalapozó vizsgálatok folynak majd. Ezzel indul a négyéves vizsgálati ciklus.

A CPM-2-t a soron következő WRC előtt legalább 6 hónappal tartják. A CPM-2 tárgyalja azt a jelentéstervezetet, amiben valamennyi WRC napirendi pont áttekintése megtörténik, benne az előzmények mellett a lefolytatott vizsgálatok ismertetésével és a lehetséges megközelítésekkel (döntési opciókkal). Ezek mentén alakulnak később ki a WRC-23 döntései. A CPM23-2 ülésére Genfben kerül sor 2023. március 27.– április 6. között.[3] A CPM23-2-re készült Jelentéstervezet jelenleg, mint közbenső, döntéselőkészítő anyag csak az igazgatások képviselői számára férhető hozzá. Az azonban nem titok, hogy az 1.2 napirendnél a 6 GHz felső sávrészére sokféle megközelítés szerepel a tervezetben. A „ne változzon semmi”-től, a „majd 2030-ban változtatunk”-on át, a „legyen a sáv IMT, és határozzunk meg minden műszaki feltételt a WRC-n határozatban”-ig több lehetőség felmerült. Ez is mutatja, hogy világviszonylatban nincs nyugvóponton a kérdés.

2.1. CEPT előkészületek

A Rádiótávközlési Világértekezletre az európai országok a CEPT (Postai és Távközlési Igazgatások Európai Értekezlete) keretein belül közösen is készülnek a CPG (Conference Preparatory Group, Konferencia előkészítő csoport) elnevezésű formáció keretében és együttesen végzik a WRC napirendi pontjainak vizsgálatát, törekedve közös álláspont kialakítására. Az egyes napirendi pontok áttekintéséről összefoglaló dokumentumokat, ún. CEPT Brief-eket[4] készítenek, melyben az adott napirendi pont kapcsán nemcsak a vizsgálati eredményeket és a lehetséges CEPT-álláspontot, hanem más régiók véleményét is megismerhetjük. A CEPT is megerősítette, hogy a 6425–7125 MHz-es frekvenciasávban különböző vezeték nélküli szélessávú felhasználások képzelhetők el, köztük kiemelésre került az IMT vagy WAS/RLAN, illetve olyan megosztott keretrendszer, amiben mind az IMT, mind a WAS/RLAN együtt élhet. Az 1.2 napirendi pont kapcsán azonban még nem alakult ki Európai Közös Javaslat (European Common Proposal, ECP). A viták még zajlanak.

1. ábra A WRC-k négyéves ciklusának áttekintése



2.2. Uniós felkészülés

Az ágazatunkat meghatározó uniós jogi aktus, a Kódex[5] több cikkében, így a 28., 39. és a 45. cikkben is deklarálja a tagállamok kötelezettségét, miszerint tiszteletben tartják

2 *https://www.itu.int/wrc-23/booklet-wrc-23/*
3 *Az Infokom 2022 novemberi konferenciáján elhangzott előadás nyomán készült cikk lezárása még az idézett CPM23-2 genfi ülés előtt történt, így az arra készült Jelentéstervezet pontos tartalmát, illetve az ott született dokumentumokat még nem ismerhettük.*
4 *A CEPT-összefoglalók legfrissebb változatát, köztük az 1.2 napirenddel foglalkozót is a következő linken lehet elérni: https://www.cept.org/ecc/groups/ecc/cpg/client/meeting-documents/?flid=31135*
5 *Az EP és a Tanács (EU) 2018/1972 irányelve (2018.12.11.) az Európai Elektronikus Hírközlési Kódex létrehozásáról (HL L 321., 2018.12.17., 36.o.)*

a vonatkozó nemzetközi egyezményeket, beleértve az ITU Rádiószabályzatát és az ITU keretében elfogadott más, rádióspektrumra vonatkozó egyezményeket. Az EU tehát elismeri az ITU szabályait, ösztönzi azok alkalmazását, együttműködik vele, viszont nem tagja az ITU-nak. Az EU az ITU szektortagja, szavazati vagy felszólalási joga nincsen. Ugyanakkor az Európai Unió szorosan együttműködik a CEPT-tel és természetesen a tagállamokkal, és a CEPT-eredményekre alapozva tanácsi határozatban rögzíti az uniós tárgyalási álláspontot a WRC-kre.

Az adott WRC-n a tagállamok által képviselendő álláspontokról szóló határozattervezetet az Európai Bizottság készíti elő, de ebben támaszkodik a Rádióspektrum Politikai Csoport (Radio Spectrum Policy Group, a továbbiakban: RSPG[6]) jelentéseire, szakvéleményeire. Az RSPG-ét létrehozó határozat alapján az RSPG segíti az Európai Bizottságot azokban az Európai Unió működéséről szóló szerződés 218. cikkének (9) bekezdésével összhangban végzett feladatokban, melyekkel az Európai Tanács részére készít javaslatokat az illetékes nemzetközi szervezetekben az Unió nevében elfogadandó álláspontokat meghatározó határozatok elfogadására irányuló javaslatokkal kapcsolatos előkészítő munkájában. Az RSPG egyértelműen rögzítette WRC Szakvéleményében[7], hogy a szubszidiaritás elve értelmében, azokban a témakörökben, amelyek nem tartoznak az Unió kizárólagos hatáskörébe, csak akkor járhat el, ha a javasolt intézkedés céljait a tagállamok nem tudják kielégítően megvalósítani, és a javasolt intézkedés mértéke vagy hatásai miatt jobban megvalósítható uniós szinten. A WRC-23 előkészítési folyamata még nem jutott odáig, hogy a Bizottság határozattervezetet nyújtott volna be a Bizottságnak.

Az RSPG – az uniós jogi szabályozási keretrendszer figyelembe vételével – áttekintette a WRC-23 napirendjét, vizsgálva az uniós szakpolitikai célokat és hatályos uniós harmonizációs döntéseket. Az RSPG emellett azokra a napirendi pontokra is figyelemmel volt, melyek keretében hozott döntések kihathatnak az uniós szakpolitikákra. A beazonosított napirendipontok esetében az RSPG ajánlásokat fogalmazott meg a legmegfelelőbb cselekvésre. Az RSPG WRC-23-re vonatkozó szakvéleménye hosszasan foglakozik az 1.2. napirendi ponttal.

A Rádiószabályzat az ITU-tagállamok közötti kapcsolatokat rendezi. Alapvetően annak érdekében szabályoz, hogy a tagállamok közötti káros zavarok ne fordulhassanak elő. Az Rádiószabályzat egy keretrendszer, mely mozgásteret biztosít a tagállamok számára, melynek keretében az egyes frekvenciák használatával szinte bármilyen rádiószolgálat keretébe tartozó alkalmazást megvalósítsanak, ha az annak keretében működő rádióállomás nem okoz káros interferenciát, és nem tart igényt védelemre a többi működő állomással szemben. A Rádiószabályzat egyetlen rendelkezése sem érintheti az Európai Unió és tagállamai azon jogát, hogy bármely harmonizált műszaki feltételt megvalósítsanak. Az uniós műszaki harmonizációs szabályok betartása kötelező, hazai jogrendbe építése alól csak kivételes esetekben és korlátok között lehet felmentést kapni.

3. A 6 GHz-es frekvenciasáv jelenlegi használata

A rádióspektrum használatának „bibliája” az NFFF, azaz a nemzeti frekvenciafelosztásról, valamint a frekvenciasávok felhasználási szabályairól szóló 7/2015. (XI. 13.) NMHH rendelet. Az NFFF, mint címe is árulkodik róla, szól a frekvenciasávok felosztásáról, ami a frekvenciasávok rádiószolgálatok részére, valamint polgári, nem polgári és együttes célú használatra történő felosztását jelenti. A felosztás, azaz a rádiószolgálatok keretei között megvalósítható alkalmazások, alkalmazáscsoportok is az NFFF-ben találhatók, ahogyan a használat részletes műszaki – esetenként harmonizált – feltételei is. Az NMHH hatályos szövegét adatbázis-formátumban feldolgozó informatikai rendszer elérhető az NMHH honlapján. A támogató rendszer neve: Spektrumgazdálkodást Támogató Információs Rendszer, mely a következő linken érhető el: *https://stir.nmhh.hu/publicview*.

Az alsó 6 GHz-es sáv (5925–6425 GHz frekvenciasáv) használatára vonatkozó uniós harmonizált szabályokat az (EU) 2021/1067 bizottsági végrehajtási határozat[8] rögzíti. Az implementációs kötelezettségének Magyarország 2022-ben tett eleget az NFFF módosításával. A jogszabályi előírásoknak megfelelően az alsó 6 GHz-es frekvenciasávban hétköznapi nyelven Wi-Fi üzemelhet. (A sávhasználati feltételeket az NFFF 3. számú mellékletében, a 4.13. pontban „WAS/RLAN rendszerek az 5945–6425 MHz sávban” címen találhatják meg az érdeklődők.)

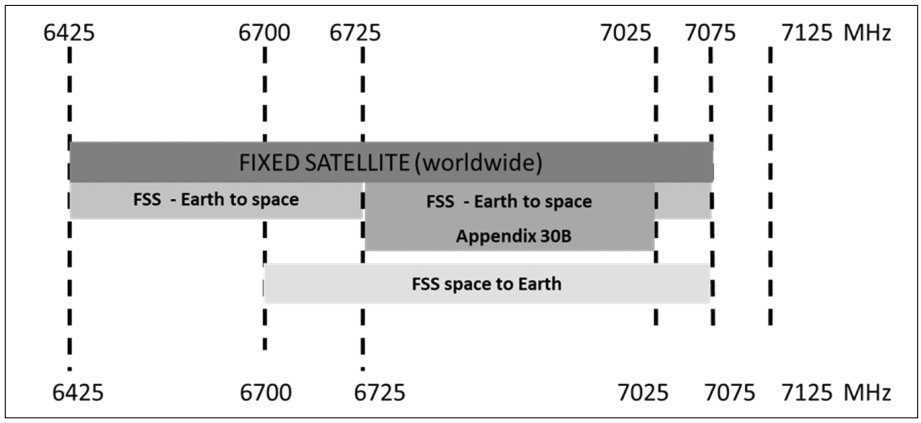
A felső 6 GHz-es sáv jelenleg még nem használható sem 5G-re, sem WAS/RLAN-célokra Európában, így Magyarországon sem. A legtöbb európai országban állandóhelyű rendszereket üzemeltetnek ebben a frekvenciasávban, és nagytávolságú pont-pont összeköttetéseket valósítanak meg a használatával. Számolni kell azonban az FSS (Fixed-Satellite Service, Műholdas állandóhelyű szolgálat), a Műholdas Föld-kutatás (EESS, másodlagos, RR 5.458), a RAS (Radar Acoustic Sounding System, Rádióakusztikus szondázó rendszer) (6650–6675,2 MHz, RR 5.149), és az EU Kopernikusz program (6425–7250 MHz) jelenlétével.

A műholdhasználattal és a projektekkel kapcsolatos információk az *2. ábrán* tekinthetők át. Számos műholdprojekt van folyamatban, és hosszú távú tervezésre van szükség az iparág igényeinek támogatására.

6 RSPG: Radio Spectrum Policy Group – Rádióspektrum Politikai Csoport (A rádióspektrum-politikával foglalkozó csoport felállításáról és a 2002/622/EK határozat hatályon kívül helyezéséről szóló 2019. június 11-i 2019/C 196/08 bizottsági határozat alapján működő Rádióspektrum Politikai Csoport (RSPG) tanácsadói minőségben a rádióspektrum európai stratégiai kérdéseivel foglalkozik.)

7 RSPG szakvélemény a WRC-23-ról, RSPG22-040 FINAL számú Brüsszelben 2022. december 9-én elfogadott döntés, https://radio-spectrum-policy-group.ec.europa.eu/system/files/2023-01/RSPG22-040final-RSPG_Final_Opinion_on_WRC23.pdf

8 A rádióspektrum 5945–6425 MHz-es frekvenciasávjának vezeték nélküli hozzáférési rendszerek – többek között rádiós helyi hálózatok (WAS/RLAN-ok) – megvalósítására történő harmonizált használatáról szóló, 2021. június 17-i (EU) 2021/1067 bizottsági végrehajtási határozat.



2. ábra
A felső 6 GHz-es frekvenciasáv FSS használata. A CEPT CPG 1.2. napirendi pontjára vonatkozó összefoglalóból

Mintegy 16 CEPT-adminisztráció jelzett ezidáig valamilyen típusú műholdas használatot a sávban, míg 8 tagállam arról számolt be, hogy nincs műholdas használat náluk. Emellett 10 tagállamban nem adtak ki engedélyt földi állomás üzemeltetésére.

A 6425–7250 MHz frekvenciasáv globálisan a Kopernikusz képalkotó mikrohullámú érzékelők (Copernicus Imaging Microwave Radiometer, CIMR) használatára tervezett, ami a Kopernikusz program hat kiemelten fontos jelöltküldetésének egyike[9]. Ezzel a programelemmel az Unió Európai zöld megállapodásának (Green Deal-nek)[10] megvalósítását is célozzák. Az EU 2021/696 Rendelet 27. Cikk (4) bekezdésen alapuló kötelezettsége a tagállamoknak, hogy meghoznak minden szükséges intézkedést a program zökkenőmentes működésének biztosítása érdekében, beleértve a programhoz szükséges frekvenciák megfelelő szintű védelméhez való hozzájárulást.

Magyarországon jelenleg nem polgári célra használt a felső 6 GHz-es frekvenciasáv nagy része. A 6425–7075 MHz frekvenciasáv elsődleges jelleggel a rádiólokáció és a légi rádionavigáció szolgálatok számára is felosztott.

A hazai felosztást a *3. ábra* szemlélteti.

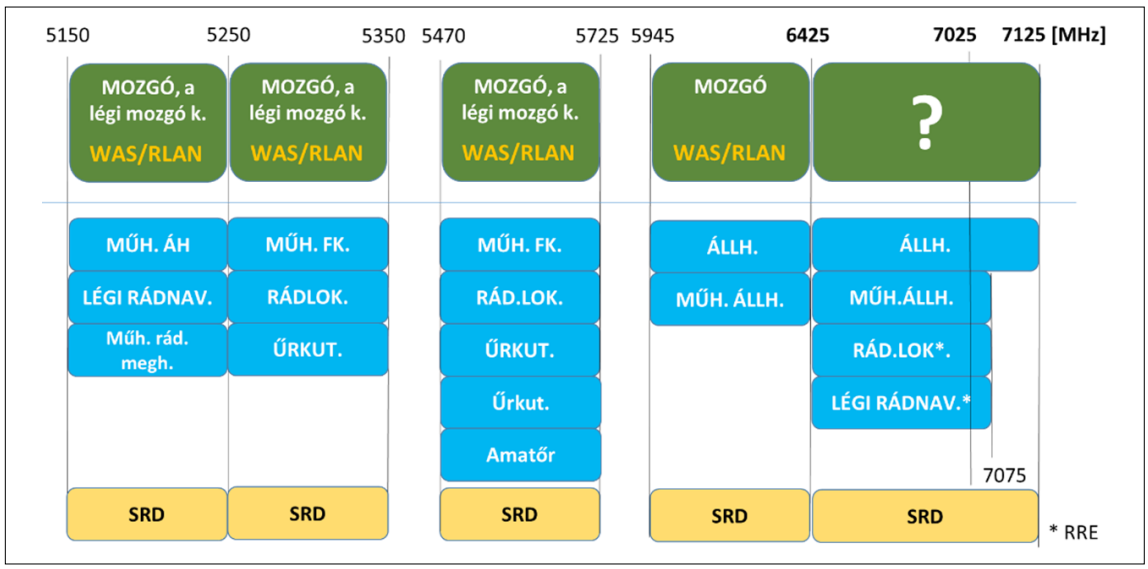
4. Két lehetséges versengő használati mód

Két versengő alkalmazást azonosítottak be a 6425–7125 MHz-es frekvenciasáv jövőbeni vezeték nélküli szélessávú használata érdekében; az IMT-t, azaz hétköznapi nyelven mobilhasználatot, valamint a WAS/RLAN-t, azaz a Wi-Fi-típusú használatot. Természetesen a műszaki szabályokat úgy kell kialakítani, hogy a meglevő használat védelme biztosított legyen, azaz az inkumbens szolgálatokkal való együttélés biztosított legyen. Itt jegyezzük meg, hogy a Rádiószabályzat szerint egyaránt mozgó szolgálatra felosztott frekvenciasávokban valósulnak meg a WAS/RLAN-alkalmazások és az IMT-alkalmazások is.

Nézzük meg, hogy az egyes alkalmazásokra eddig mely frekvenciasávok felosztása, elosztása történt meg.

4.1. Az IMT jelenlegi használata Magyarországon

Az uniós műszaki harmonizációs folyamatok eredményeként jelenleg Magyarországon az összes olyan frekvenciasáv elérhető, mely az unió bármelyik másik tagállamában is technológia semleges módon széles-



3. ábra
A 6 GHz-es frekvenciasáv magyarországi felosztásáról

9 *Az uniós űrprogram és az Európai Unió Űrprogramügynökségének a létrehozásáról, valamint a 912/2010/EU, az 1285/2013/EU és a 377/2014/EU rendelet és az 541/2014/EU határozat hatályon kívül helyezéséről szóló (EU) 2021/696 számú, 2021. április 28-ai parlamenti és tanácsi rendelet, https://eur-lex.europa.eu/legal-content/HU/TXT/PDF/?uri=CELEX:32021R0696&from=EN*

10 *https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/european-green-deal_hu*

1. táblázat A magyarországi mobil rádiótávközlési célokra elosztott frekvenciákról

	700 MHz	800 MHz	900 MHz	1500 MHz	1800 MHz	2100 MHz	2600 MHz	3600 MHz	Összesen
Rendelkezésre áll	60,0 MHz	60,0 MHz	70,0 MHz	90,0 MHz	150,0 MHz	120,0 MHz	180,0 MHz	390,0 MHz	**1 120,0 MHz**
Elosztott	50,0 MHz	60,0 MHz	70,0 MHz		150,0 MHz	120,0 MHz	165,0 MHz	390,0 MHz	**1005,0 MHz**
Telekom	20,0 MHz	20,0 MHz	20,0 MHz		60,0 MHz	50,0 MHz	60,0 MHz	120,0 MHz	350,0 MHz
Yettel	10,0 MHz	20,0 MHz	30,0 MHz		40,0 MHz	30,0 MHz	40,0 MHz	140,0 MHz	310,0 MHz
Vodafone	20,0 MHz	20,0 MHz	20,0 MHz		40,0 MHz	40,0 MHz	65,0 MHz	110,0 MHz	315,0 MHz
DIGI	0,0 MHz	0,0 MHz	0,0 MHz		10,0 MHz	0,0 MHz	0,0 MHz	20,0 MHz	30,0 MHz

sávú vezeték nélküli elektronikus hírközlési szolgáltatásokra azonosításra került. Ez a stádium, az uniós határozatokban, mint „designate" kifejezés jelenik meg, ami a hazai jogszabályban, az NFFF-ben tervezett vagy kijelölt státuszt jelent, ez az első lépés a rádiós rendszerek megvalósításához. Ettől eggyel közelebb visz a tényleges használathoz a „make available", ami versenyeztetésre jelölt sávok esetén legalább a piaci igény felmérését jelenti, azaz ha jelentkezik piaci igény, az NMHH hivatalból indítja az elosztási eljárást, azaz jelen kategóriában tipikusan (ezidáig kizárólagosan) a versenyeztetési eljárást.

Az 1500 MHz-es frekvenciasáv iránti igényt már több alkalommal vizsgálta az NMHH nyilvános meghallgatásain, de ezidáig indokolt igény nem jelentkezett, így az elosztás érdekében nem indult el az eljárás. De nézzük meg, mennyi frekvenciát használnak jelenleg Magyarországon mobil célokra. Erről az *1. táblázatban* található áttekintés.

4.2. A WAS/RLAN jelenlegi használata Magyarországon

A WAS/RLAN-rendszerek használatát az IMT-hez hasonlóan az uniós műszaki harmonizáció jellemzi. A rádiós helyi hálózatokhoz való hozzáférés általános szabályait a Kódex[11] 56. Cikke szabályozza. A rádiós helyi hálózat – a Kódex meghatározása szerint – olyan kis teljesítményű, vezeték nélküli hozzáférési rendszer, amely kis hatótávolságon belül üzemel, úgy, hogy a más felhasználók által a közelben használt hasonló rendszerek zavarásának a kockázata csekély, és amely nem kizárólagos jelleggel harmonizált rádióspektrumot használ. Jellemző az engedélymentes használat.

A harmonizált műszaki feltételeket a Bizottság a Rádióspektrum Határozat[12] alapján alkotta. A 2400 MHz-es frekvenciasáv harmonizált műszaki feltételeire a 2006/771/EK, valamint az (EU) 2022/180 bizottsági végrehajtási határozat[13] vonatkozik. A 2005/513/EK bizottsági határozat[14] harmonizálta a rádióspektrum 5 GHz-es sávjának (5150–5350 MHz és 5470–5725 MHz) vezeték nélküli hozzáférési rendszerek, többek között rádiós helyi hálózatok céljára történő használatát. A *2. táblázat* nyújt áttekintést a WAS/RLAN alkalmazás körébe tartozó rendszer céljára használható frekvenciákról.

A WAS/RLAN-rendszerek a 10 GHz alatt (beltéri használatra) összesen 1168,5 MHz-et érhetnek el.

2. táblázat A WAS/RLAN-szabályozás áttekintése Magyarországon

Frekvenciasáv	Teljesítmény	Kültér	Beltér	Légijármű	Vonat	Járművek (személyautó, teherautó, buszok)	UAS
2400 – 2483,5 MHz	100 mW	I	I	I	I	I	I
5150 – 5250 MHz	200 mW	N	I	I	40 mW	40 mW	5170 – 5250 MHz
5250 – 5350 MHz	200 mW	N	I	N	N	N	N
5470 – 5725 MHz	1000 mW	I	I	N	N	N	N
5725 – 5875 MHz	25 mW	I	I	I	I	N	N
5945 – 6425 MHz	LPI – 23 dBm	N	I	I	I	N	N
	VLP – 14 dBm	Helyhez kötött telepítés nem megengedett	I	N	N	N	N
57 – 71 GHz	40 dBm	Helyhez kötött telepítés nem megengedett		N	N	N	N

11 *Az Európai Parlament és a Tanács (EU) 2018/1972 irányelve (2018. december 11.) az Európai Elektronikus Hírközlési Kódex létrehozásáról (HL L 321., 2018.12.17., 36.o.)*

12 *Az Európai Parlament és a Tanács 676/2002/EK határozata (2002. március 7.) az Európai Közösség rádióspektrum-politikájának keretszabályozásáról (rádióspektrum-határozat) (HL L 108., 2002.4.24., 1.o.)*

13 *A Bizottság (EU) 2022/180 végrehajtási határozata (2022. február 8.) a 2006/771/EK határozatnak a kis hatótávolságú eszközök általi rádióspektrum-használattal kapcsolatos harmonizált műszaki feltételek naprakésszé tétele tekintetében történő módosításáról (az értesítés a C(2022) 644. számú dokumentummal történt) (EGT-vonatkozású szöveg), C/2022/644, (HL L 29, 2022.02.10., 17–39.old.)*

14 *A Bizottság 2005/513/EK határozata (2005. július 11.) a rádióspektrum 5 GHz-es frekvenciasávjának a vezeték nélküli hozzáférési rendszerek, beleértve a rádiós helyi hálózatokat is (WAS/RLAN), megvalósítására történő harmonizált felhasználásáról (HL L 187., 2005.7.19., 22.o.)*

5. Ki dönt és mi alapján, hogy ki férjen hozzá a rádióspektrumhoz?

Az infokommunikációs technológiák (IKT), mint feltörekvő technológiák fontos szerepet játszanak a társadalmi-gazdasági fejlődés támogatásában. A technológiák, köztük az IMT-2020 alkalmazása növeli a termelékenységet, új lehetőségeket teremt és új szolgáltatásokat generál. Az IMT-rendszerek különféle használati forgatókönyveket képesek támogatni, beleértve a továbbfejlesztett mobil szélessávot (eMBB), a masszív géptípusú kommunikációt (mMTC), valamint az ultra-megbízható, alacsony késleltetésű kommunikációt (URLLC).

Mivel az IMT-alkalmazások iránti kereslet folyamatosan növekszik, további rádióspektrum azonosításának lehetőségét vizsgálja az összes, a cikkben említett nemzetközi szervezet és a hazai szabályozó, az NMHH is. Ennek a vizsgálatnak a keretei között a középső frekvenciasávokban, a kapacitás bővítése érdekében a 6 GHz-es sáv jelentős lehetőséggel kecsegtet. A WRC-23 1.2. napirendi pontja hozzájárulhat ahhoz, hogy az ITU tagállamai nagyobb rugalmasságot biztosítsanak a tagállamok számára az IMT megvalósításához megfelelő frekvenciasávok azonosításban. A WAS/RLAN-rendszerek számára is páratlan lehetőséget jelent a 6 GHz-es frekvenciasáv, hiszen itt nyílna lehetőség globálisan nagyobb, egybefüggő frekvenciamennyiség engedélymentes használatára.

A szabályozók a jelöltek műszaki alkalmasságának, azaz a minimumfeltételek teljesítésének megállapítását követően vizsgálják, hogy a jövőbeli adatmennyiségi igények kielégítésére mely rendszer lehetne hatékonyabb, melyikkel lehetne inkább javítani a lefedettséget. Mérlegelnek méretgazdaságossági, társadalmi, gazdaságossági szempontokat, valamint a spektrumhatékonyság oldaláról is vizsgálódnak. A vizsgálatok még nem zárultak le a felső 6 GHz-es frekvenciasávra. Az ITU mozgószolgálati azonosítása pedig még biztosít mozgásteret a későbbeikben a részletesebb harmonizációs döntés meghozatalára.

A CEPT napirenddel foglalkozó összefoglaló tervezetében elismerte, hogy egyes, a CEPT-en kívüli országok és/vagy régiók javasolhatnak IMT-azonosítást a 6425–7125 MHz frekvenciasávban, és bár nem szorgalmazza vagy proaktívan támogatja azt, a CEPT fontolóra veszi az IMT-azonosítás esetleges elfogadásának feltételeit ebben a sávban. A CEPT azt is hangsúlyozza, hogy az esetleges IMT-azonosítás nem zárja ki a felső 6 GHz-es frekvenciasáv használatát más, esetlegesen nem elsődleges alkalmazás bevezetését. A kiegészítő rendelkezések között egyértelműen szerepelnie kell egyéb mozgószolgálati alkalmazásoknak, mint például a WAS/RLAN, ami az IMT-hez hasonlóan szélessávú adatátvitelt tesz lehetővé.

Az RSPG azon a véleményen van, hogy egy jövőbeli európai politikai stratégia, valamint harmonizációs határozat a 6425–7125 MHz frekvenciasávra alaposan kidolgozott kell legyen. Az Európára vonatkozó döntés szempontjai között a következőket emeli ki a szakvélemény:

– az IMT (beleértve az IMT-2030-at) és a WAS/RLAN spektrumigényét középtávon;
– lehetséges funkciók, amelyek lehetővé teszik a rádiótávközlési megoldások közötti megosztási lehetőségeket (szolgálatok és alkalmazások együttélése a sávban), beleértve az IMT és a WAS/RLAN közötti alkalmazásokat;
– a középső sávú lehetséges spektrumhasználati esetek a 7125 MHz és 20 GHz között, ami alternatívát jelenthet IMT-re.

Egy biztos: jelenleg Magyarországon a meglévő nem polgári használatra tekintettel nem valósulhatna meg egyik rendszer sem.

6. Összefoglalás

Összességében megállapíthatjuk: a rádióspektrum-gazdálkodók a műszaki alkalmasság megállapításán túl, amennyiben több igény van, mint amennyi a rendelkezésre álló rádióspektrum mennyisége, nem szeretik maguk eldönteni, hogy mely lehetséges technológia legyen megengedett. Ennek egyik ékes példája a WRC 1.2 napirend alapján terítékre került 6 GHz-es frekvenciasáv. A technológiasemlegesség elve, a befektetésvédelmi célok, az innováció elősegítése… mind azt ösztönzik, hogy ne a szabályozó hozzon döntést, hanem a legjobb azt a piacra bízni. Ráadásul mindkét igénylő társadalmi hatékonyság szempontjából ugyanazt a digitalizációs célt támogatja, hisz a vezeték nélküli szélessávú adatátvitelnek kiváló megvalósítási lehetőségei vannak. A cél mindkét alkalmazás esetében azonos: a digitalizáció terjedéséhez stabil, szélessávú adatátvitelt lehetővé tevő infrastruktúra biztosítása. Társadalmi hasznosság szempontjából így igen nehezen lenne eldönthető a kérdés. A spektrumhatékonyság lehetne egy másik fokmérő, de az egyik rendszer az egyik, a másik pedig a másik szempontból jobb egy hajszállal.

Egy biztos, még további műszaki vizsgálatokra van szükség, és egy esetleges WRC-döntés sem jelenti a folyamat végét, hisz a műszaki harmonizáció európai, valamint uniós szinten csak ezt követően indul. A műszaki szabályozók bíznak a mérnöki leleményességben, amivel a megosztott használat előtt megnyílna a lehetőség. Végső esetben az alternatívák vizsgálatával foglalkozunk majd.

Sem a témakörrel foglakozó előadás, sem a cikk megírása nem eredményezte a megvilágosodást, és a kettő között eltelt idő sem volt elégséges a verseny eldöntésére. Itt is láthatjuk, hogy a spektrumgazdálkodás nem sprint, hanem maraton, terepakadályokkal és nehéz időjárási körülményekkel.

A szerzőről

ULELAY EMÍLIA a rádióspektrum-gazdálkodás szabályozási kérdéseivel több mint két évtizede foglakozik. Szakmai pályafutását az NMHH jogelődjénél, a HIF-nél kezdte, majd az ágazati szabályozó minisztériumoknál gondoskodott a rádióspektrum szakpolitikai feladatairól, végezte a hazai és a nemzetközi szabályozás alakítását, ellátta a nemzetközi képviseleti feladatokat.

Mikor válhat időszerűvé a 2G-mobilszolgáltatás kivezetése Magyarországon?

KOLLÁR PÉTER
Nemzeti Média- és Hírközlési Hatóság
Kollar.Peter@nmhh.hu

Kulcsszavak: mobiltechnológiák, 2G-, 3G-, 3G-kivezetés, NetreFel program

A mobiltechnológiák természetes evolúciója, hogy a nagyobb sávszélességet és funkcionalitást kínáló újabb generációk előbb-utóbb kiszorítják a korábbiakat. Magyarországon 2022 végére egy szolgáltató már teljesen, kettő pedig részlegesen kivezette a 3G-szolgáltatást. Ebben nagy szerepet játszottak a 3G lekapcsolását támogató NetreFel program kommunikációs, koordinációs és támogatási eszközei. A program tapasztalatai a 2G-lekapcsolás koncepcionális előkészítésében és a kivezetés gyorsításában egyaránt komoly segítséget jelenthetnek. A hazai adottságok és a nemzetközi tapasztalatok alapján az látszik, hogy a 2G kivezetésére Magyarországon 2025–2030 között kerülhet sor, ha a 2G M2M felhasználásának migrációját sikerül megoldani.

1. Bevezetés

A 2019-ben indult első komolyabb szakmai konzultációk után, hosszú és gondos szakmai előkészítést követően az NMHH 2022-ben jelentős kommunikációs, támogatási és koordinációs erőfeszítéseket tett a *NetreFel* program keretében arra, hogy a 3G-hálózatok lekapcsolása és a 3G-szolgáltatás kivezetése zökkenőmentesen, az érintettek tudatos felkészítésével történjen meg. Ennek eredményeként 2022 végéig a Magyar Telekom már teljesen, a másik két 3G-t üzemeltető szolgáltató (Yettel, Vodafone)[1] pedig részlegesen – lényegi lakossági és vállalati felhasználói panaszok nélkül – már kivezette a 3G szolgáltatását Magyarországon. Ezzel hazánk a 3G-lekapcsolás nemzetközi élmezőnyébe került, ami azért különösen fontos, mert így a felszabaduló korlátos frekvencia és az egyéb erőforrások jövőálló, innovatív technológiák (4G, 5G) és szolgáltatások bevezetésére fordíthatók. A program tapasztalatai a 2G kivezetésére vonatkozó előkészületek kapcsán is hasznosíthatók. A konkrét lehetőségek vizsgálatához, a koncepcionális előkészületek megkezdéséhez először azt az időtávot érdemes megbecsülni, amelyen a 2G kivezetésére reális esély nyílik.

A cikk először a 3G kivezetésének legfontosabb tapasztalatait és eredményeit tekinti át, ezt követően pedig – a lehetséges 2G lekapcsolási céldátum meghatározása érdekében – a hazai felhasználói trendeket (2G-készülékállomány, 2G-hívás- és adatforgalom, 2G-hálózatot használó M2M-alkalmazások) és a releváns nemzetközi tapasztalatokat elemzi.

2. A 3G kivezetésének tapasztalatai és eredményei

Magyarországon 2022 elején mintegy 560 ezer lakossági felhasználónak volt 3G-s készüléke, közülük 180 ezren bonyolítottak mobil adatforgalmat. A 3G-hálózatok lekapcsolása közvetlenül ez utóbbi csoportot érinti, mivel a szolgáltatás leállításával a 3G-készülékek hangforgalmát a 2G-hálózatok átveszik, ugyanakkor mobil internetre ezek a készülékek szinte alkalmatlanná válnak. (Bár a 2,5G-nek nevezett EDGE-adatszolgáltatás továbbra is elérhető marad, ez egy átlagos felhasználó számára nem nyújt valós internetezési alternatívát.) Közvetetten persze ezek a 3G-készüléktulajdonosok is kárvallottjai lehetnek a lekapcsolásnak, hiszen esetükben a mobilnet előfizetés lehetősége szűnik meg. Lényegében ez a fogyasztóvédelmi megfontolás hívta életre 2021 végén, 2022 elején a *NetreFel programot (https://netrefel.hu)*, amely később a mobilinternet általános népszerűsítésére és a mobilnetet nem használó 2G-készüléktulajdonosokra is kiterjesztésre került, így egy átfogó digitalizációs fejlesztéspolitikai kezdeményezéssé vált.

Az NMHH által indított és koordinált program hármas célrendszert fogalmazott meg:

- a mobilinternetet még nem használó lakossági célcsoportok digitális társadalomba történő bevonása (kommunikáció, edukáció, motiváció, készüléktámogatás);
- a lakossági 3G-felhasználóknak a lekapcsolásra történő felkészítése (kommunikáció), illetve körükben a 4G/5G-szolgáltatások igénybevételére nem alkalmas készülékek cseréjének támogatása (készülékcsere-támogatási program);
- az üzleti és közületi 3G-használók (M2M, IoT) felkészítése a várható lekapcsolásra (konferenciák és tájékoztatók, workshopok keretében).

A NetreFel kommunikációs program változatos eszközökkel (folyamatosan frissített tematikus web- és Facebook-oldal, fizetett médiamegjelenések, PR-tevékenység, konferenciák, podcast, médiaszereplések stb.) hívta fel az érintettek figyelmét a digitális eszközök és szolgáltatások előnyeire, valamint a 2G- és 3G-képes telefonok lecserélésének fontosságára.

1 A Digi mobilszolgáltatása nem alkalmaz 3G-s technológiát.

A NetreFel készülékcsere program *(https://mobilcsere.nffku.hu/)* 2022. február–2023. március között 20-40 ezer forintos támogatást biztosított, ha egy lakossági felhasználó 3G- vagy 2G- készülékét leadva új legalább 4G-képes telefon vásárlása mellett döntött.

A 3G M2M- és IoT-szolgáltatók és felhasználók felkészítése érdekében az NMHH nyilvános meghallgatást, konferenciát és számos szolgáltatói egyeztetést tartott, felvette a kapcsolatot a kormányzati intézményekkel, önkormányzatokkal és kezdeményezte az ITM-nél, hogy az általános vállalkozásfejlesztési célú pályázatok keretében az elavult M2M- és IoT-eszközök cseréje a támogatható tevékenységek, illetve az elszámolható költségek között szerepeljen.

A NetreFel program a 3G/2G-készülékcsere-program 2023. március végi lezárásával fontos mérföldkőhöz érkezett. A program részletes értékelését az NMHH 2023 első félévében fogja elvégezni, de az már ma is egyértelmű, hogy a NetreFel program jelentősen hozzájárult ahhoz, hogy 2022 folyamán a három vezető mobilszolgáltató közül a Telekom teljesen, a Yettel és Vodafone pedig részlegesen úgy kapcsolhatta le 3G-hálózatát, hogy egyáltalán nem voltak a 3G-szolgáltatás kiesésével kapcsolatos fogyasztói panaszok. Ennek fő okai az alábbiak:

- A NetreFel program *kommunikációs eszközei* a felhasználók széles körében tudatosították, hogy a 3G-hálózatok lekapcsolása milyen kihívásokat és lehetőségeket jelent.
- A NetreFel *készülékcsere program* keretében kiépült és mindenki számára könnyen hozzáférhetővé, igénybevehetővé vált az a 3G lekapcsolásához szükséges „védőháló", amely jogi, eljárásrendi és informatikai értelemben egyaránt problémamentesen, országos lefedettséggel működött. *Így a leginkább rászorultak szinte ingyen juthattak belépő szintű 4G-képes készülékhez.*
- A készülékcsere program előkészítése során
  - a szolgáltatók az NMHH módszertani iránymutatása és minőségbiztosítása mellett minden egyes 3G-lekapcsolással érintett fogyasztójukat beazonosították, a lehetséges fogyasztói panaszok kezelésére felkészültek;
  - a programba való bekapcsolódásuk révén a független készülékforgalmazók (elektronikai áruházláncok) számára is világossá vált, hogy mit jelent a 3G lekapcsolása, így ezeket az információkat vevőik számára át tudták adni. Ezek az információk nem utolsósorban befolyásolják az elektronikai áruházláncok beszerzési politikáját is, azaz kevésbé lesz preferálva az elavult technológiát használó készülékek árusítása.

A NetreFel program *további járulékos eredményeként* azonosíthatóak az alábbiak:

- Magyarország Európában élenjár a 3G-hálózatok lekapcsolásában, ez a szolgáltatók és a szabályozó számára nemzetközi elismerést jelent. A 3G-készülékcsere program eljárásrendje, módszertana, informatikai infrastruktúrája más országokban is alkalmazható, nemzetközi érdeklődésre tarthat számot.
- A készülékcsere program előkészítésében és lebonyolításában megvalósított példaértékű együttműködés a mobilszolgáltatók és az NMHH között más digitális fejlesztési programok esetében is jó alap lehet.
- Egyértelművé vált, hogy az NMHH képes professzionális szinten előkészíteni és lebonyolítani, államigazgatási és piaci szereplők együttes részvételével megvalósuló, tömegeket érintő digitális fejlesztési, támogatási programokat.

A 3G-hálózatok lekapcsolása (energiamegtakarítás) és a csere során leadott korszerűtlen készülékek kivonása a forgalomból környezeti, fenntarthatósági szempontból is előnyös (zöldítés). *Végül és nem utolsósorban a 3G-lekapcsolás módszertani, kommunikációs, koordinációs, jogi, informatikai tapasztalatai a 2G-lekapcsolás előkészítése során is komoly segítséget jelenthetnek.*

3. Mikor vezethetjük ki Magyarországon a 2G-mobilszolgáltatást?

A 2G-hálózatok lehetséges lekapcsolásának időtávját három közelítésben vizsgáljuk:

- a hazai 2G-telefonszegmens adottságai,
- a hazai 2G M2M-szegmens adottságai,
- a nemzetközi 2G-lekapcsolási tapasztalatok és előkészületek tapasztalatai.

3.1. A hazai 2G-telefonszegmens adottságai

Magyarországon a jelenleg használatban lévő mobilkészülékek mintegy ötöde kizárólag 2G-képes[2]. Mint az *1. ábra* mutatja, ezen készülékek szerepe a hangforgalom bonyolításában folyamatosan csökkent az elmúlt években: 2015-ben a hívásforgalomnak még 40 százaléka, ma már viszont csak kevesebb, mint 10 százaléka zajlik ezeken a hálózatokon.

	2G		3G		4G	
	hívás	*internet*	*hívás*	*internet*	*hívás*	*internet*
2015	40,4%	1,8%	59,4%	43,5%	0,2%	54,7%
2016	30,8%	1,6%	68,9%	22,2%	0,3%	76,2%
2017	22,2%	1,3%	71,0%	9,9%	6,8%	88,8%
2018	17,4%	1,1%	70,5%	5,5%	12,2%	93,4%
2019	16,1%	0,8%	58,7%	3,2%	25,2%	96,0%
2020	16,7%	1,1%	39,1%	3,3%	44,2%	95,6%
2021	15,3%	0,1%	22,3%	4,1%	62,4%	95,9%
2022 II. n.év	8,4%	0,0%	19,5%	1,9%	72,1%	98,1%

1. ábra
Hívás- és internetforgalom megoszlása hálózattípus szerint
(Forrás: NMHH Mobilpiaci jelentés 2022/Q2)

Az internetforgalom a 2G-hálózatokon – a relatíve alacsony sávszélesség miatt – sohasem volt jelentős és a 3G-hálózatok megjelenésével, a 4G-hálózatok kiépülésével relatíve gyorsan marginálissá vált: 2015-ben az internetforgalom mintegy 2 százaléka bonyolódott ezeken a hálózatokon, mára viszont a 2G-eszközök adatforgalma már alig mérhető (2021: 0,1%). Természetesen ez a szám napjainkban sem nulla, hiszen történik adatforgalom (némelyik kritikusnak is nevezhető) ezeken a hálózatokon, csak a növekvő adatforgalom miatt, az arányszám egyre alacsonyabb lesz.

Középtávon a 2G-készülékek hangforgalmának és a 2G-készülékek számának további csökkenésére számíthatunk. Ezt a várakozásunkat támasztják alá azok a nemzetközi előrejelzések, melyek szerint jelentős növekedés az okostelefonok gyártása és értékesítése terén rövid- és középtávon már nem várható, a gyártók, forgalmazók számára sokkal inkább a meglévő készülékek cseréje/upgrade-je jelenthet üzleti lehetőséget[3].

Jelenleg a világon megközelítőleg 1,37 Mrd 2G-, 1,63 Mrd 3G- és 4,67 Mrd 4G- LTE-előfizetés létezik. Az 5G-hálózatok felfutása még csak most indult, 2022-ben mintegy 220 millió aktív készülék tartozott ebbe a kategóriába. Így a nemzetközi várakozások alapján a következő 3-5 éves időszak meghatározó tendenciája a 2G- és a 3G-készülékek számának erőteljes csökkenése és a 4G- valamint 5G-képes készülékek számának jelentős növekedése lesz. Ennek egyik további oka, hogy a 2G- és 3G-telefonok és a legalább 4G-képes okostelefonok közötti árkülönbség számottevően lecsökkent az elmúlt 3-4 évben.

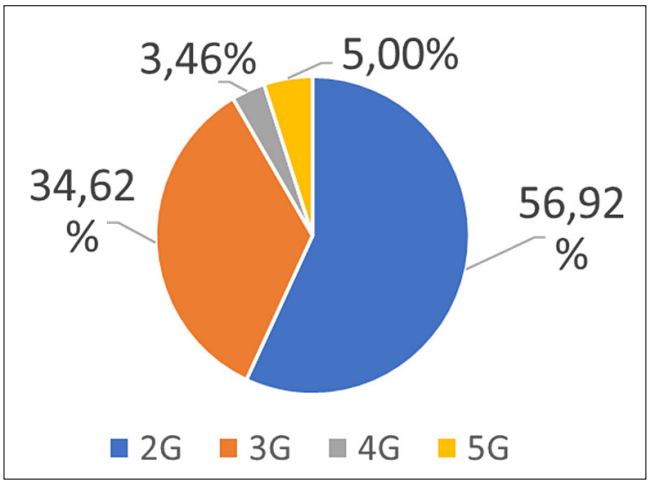
A készülékállomány természetes megújulásának üteme szempontjából meghatározó, hogy a mobiltelefonok csereciklusa 1-7 év között változik országonként. Becslésünk alapján Magyarországon a készülékek átlagos élettartama mintegy 50 hónapra becsülhető. Ez azt jelenti, hogy évente, minden egyéb tényező változatlansága esetén a készülékek mintegy negyedét cserélik le. Feltételezve, hogy a cserék intenzitása az elavultabb régebbi készülékszegmensben magasabb, azt gondoljuk, hogy megfelelő szabályozási[4], kommunikációs és támogatási eszközökkel *a hazai csak 2G-képes készülékeket 3-5 év alatt korszerűbb telefonokra lehetne cserélni.*

3.2. A hazai 2G M2M szegmens adottságai

A telefonszegmensnél nagyobb kihívás körvonalazódik az M2M-eszközök esetében a 2G-lekapcsolás tekintetében.

Mint a *2. ábra* mutatja, az M2M SIM-kártyák többsége Magyarországon 2G-típusú. Ráadásul ebben a szegmensben a technológiaváltás üteme sokkal lassabb lesz, mint a mobiltelefonok esetében:

- A 3G-hálózatok lekapcsolásával, a 3G SIM-ek adatforgalmát is a 2G-hálózatok veszik át.



2. ábra
M2M SIM kártyák eloszlása hálózattípus szerint, 2022/Q1
(Forrás: Becslés szolgáltatói adatszolgáltatás alapján)

- Középtávon alig azonosítható olyan fogyasztói igény, amely a 2G SIM-ek lecserélésének irányába hatna: az NMHH mobilpiaci jelentése alapján (2021/Q1) az M2M SIM-ek közül csak minden tizedik – jellemzően 4G-képes – kártya bonyolít jelentősebb, havi 10 Mbyte-ot meghaladó adatforgalmat (pl. videós felügyeleti rendszerek).

Az összes M2M-szolgáltatást használó eszköz közül – a 3G kivezetéséhez hasonlóan – a pénztárgépek tekinthetők a leginkább kritikus felhasználói körnek a 2G lekapcsolása szempontjából. *A hazai 3G SIM-kártyák döntő többségét adó, mintegy 220-230 ezer kártya pénztárgépekben működött 2021 végén*[5]. A 3G lekapcsolása során alapvető elvárás volt, hogy pénztárgépek semmilyen működési, adóügyi kockázatot ne jelentsenek. Ugyanez a probléma fog jelentkezni a 2G-technológia kivezetésének vizsgálata kapcsán.

A 2G- és 3G-hálózaton működő pénztárgépek esetében szinte egyáltalán nem számíthatunk – a telefonszegmenshez hasonló – organikus technológiai korszerűsítésre, mivel

- jelen pillanatban a hatályos jogszabályok még támogatják a 4G-adatátvitelnél alacsonyabb technológia használatát is,
- egy pénztárgép akár 15-20 évig működőképes lehet,
- egy 4G-képes pénztárgép, semmilyen extra szolgáltatást nem kínál a 2G/3G-képeshez viszonyítva a felhasználók számára,
- a pénztárgépbe megbonthatatlanul került beépítésre az adóügyi egység, ezért csak kizárólag egyben lehet cserélni.

Így az M2M-szegmens esetében egy teljes (minden szolgáltatót érintő) 2G-lekapcsolás időpontját jelenleg még nagy bizonytalansággal sem lehet megbecsülni.

2 *A NetreFel készülékcsere-program szolgáltatói adatai alapján becsült szám.*
3 *https://www.statista.com/statistics/263441/global-smartphone-shipments-forecast/*
4 *Vietnam például olyan módon segíti elő az okostelefonok elterjedését, illetve digitális törekvései teljesülését, hogy 2021 januárjában rendeletet adott ki, amely július elejétől megtiltotta a 2G- és 3G- készülékek (nem-okostelefonok) gyártását és importját.*
5 *https://www.origo.hu/gazdasag/20190901-izer-norbert-ot-eves-a-koltsegvetesnek-mar-tobb-szazmilliard-forintot-hozo-online-penztargep.html*

3. ábra
Azok a szolgáltatók, amelyek már kivezették a 2G szolgáltatást (Forrás: 2G and 3G sunsets and how to prepare, Onomondo 2022; letölthető: https://onomondo.com/resource-hub/2g-3g-sunset/#why-sunset)

3.3. Nemzetközi tapasztalatok

Ha a nemzetközi tapasztalatok alapján próbáljuk meghatározni a magyarországi 2G-hálózatok lekapcsolásának várható időtávját, akkor azt látjuk, hogy a 2G-technológiák kivezetésére már a 2010-es években is találunk példákat. Ekkor jellemzően azon technológiák lekapcsolásáról döntöttek az egyes szolgáltatók, amelyek a WCDMA-szabványoktól eltértek és nem biztosították a hosszútávú fejlődési lehetőséget a szolgáltatók és a mobilpiac egésze számára. Ilyen technológiai zsákutca volt például Japánban a PDC-, míg más országokban – elsősorban USA, Kanada, Indonézia, India, Oroszország – a CDMA-technológia.

A második 2G-lekapcsolási hullámot, 2015 után, a WCDMA 2G-s technológia kivezetése jelentette. Ezekben az esetekben már az egy technológiacsaládon belüli helyettesítési, spektrum-, valamint költséghatékonysági megfontolások indokolták a legrégebbi technológia kivezetését. Ugyanakkor a 2G lekapcsolása eddig nem vált „tömegessé”, 2022-ig világszerte csak mintegy 40 szolgáltató vezette ki hálózatán a 2G-szolgáltatást. A folyamat Észak-Amerikában és Ázsiában haladt leginkább előre.

Ha az európai 2G-lekapcsolási terveket vizsgáljuk, akkor az alábbiak láthatók:

- Európában eddig az EU „magországok” (Németország, Franciaország, Olaszország, Egyesült Királyság, Luxemburg, Belgium, Hollandia), valamint Norvégia és Lengyelország szolgáltatói hoztak ezidáig nyilvánosságra lekapcsolási terveket.
- A 2G lekapcsolása jellemzően globális stratégia részeként jelenik meg. Az eddig nyilvánosságra hozott 16 lekapcsolási időpont közül 5 Vodafone, 5 pedig Orange leányvállalatokhoz kapcsolható.
- A 2G lekapcsolásának első hulláma 2025 körül várható Európában, a többi szolgáltató pedig 5-10 éves távlatban tervezi ezt a lépést.

4. Összefoglalás

A magyarországi 2G-hálózatok lekapcsolása fontos lökést adhat a korszerű mobiltechnológiák és alkalmazások terjedésének, a frekvenciakészlet hatékonyabb felhasználásának.

A 2G-mobilkészülékek állománya, a 2G-hálózatok hívás és internethasználati mintái, valamint a nemzetközi tapasztalatok egyaránt azt mutatják, hogy a 2G lekapcsolása 2025–2030 között reális üzlet- és szakpolitikai cél lehet. Ehhez azonban az M2M-szegmensben jelenleg fontos szerepet játszó 2G-szolgáltatások sikeres migrációjára van szükség, különös tekintettel a kritikus fontosságú és érzékenységű pénztárgépszegmensre.

A hatóság számára ez a kihívás annál is inkább fontos, mivel az NMHH 2023–2026-os stratégiájának „Proaktivitásra” és „Alkalmazkodókészségre” vonatkozó értékválasztása szerint[6] a hatóság a jövőben is aktívan támogatni kívánja a technológiaváltást és az innovációt a távközlési piacon:

- *„Az NMHH a szabályozási és társadalmi kihívások időben történő felismerésére és megoldására törekszik a digitális világban, és ennek érdekében, hatáskörében eljárva önállóan megteszi vagy kezdeményezi a szükséges lépéseket.”*
- *„...Az NMHH nyitott és innovatív hozzáállásával követi a piaci, társadalmi és a technológiai változásokat.”*

A 2G-lekapcsolás koncepcionális előkészítése ennek a stratégiai orientációnak egy fontos eleme lehet a következő 2-3 évben.

A szerzőről

KOLLÁR PÉTER több, mint huszonöt éves tapasztalattal rendelkezik a távközlési és IT-területen. Dolgozott hagyományos sodrott rézérpáras hálózatok építésénél, de részt vett a modern kábeltévé hálózatok kiépítésében és digitalizációjában is. Ismeri mind a B2C-, mind a B2B-piacra fókuszáló szolgáltatásokkal kapcsolatos elvárásokat, üzleti megoldásokat. Jelenleg a Nemzeti Média- és Hírközlési Hatóság frekvencia- és azonosítógazdálkodással kapcsolatos munkáját irányítja igazgatóként.

6 Szabadság és biztonság a digitalizáció világában – A Nemzeti Média- és Hírközlési Hatóság Stratégiája 2023–2026, 3. old., letölthető:https://nmhh.hu/dokumentum/235259/NMHH_strategia_2023_2026.pdf

Az önálló, „igazi” 5G-mobilhálózatok pénzügyi mozgatórugói

FÖLDES GÁBOR
Vodafone Intelligent Solutions
gfoldes80@gmail.com

Kulcsszavak: önálló 5G, hálózatszeletelés, üzleti terv, beruházások megtérülése

Az 5G-hálózatok teljesítménye a 2019 óta jellemzően 4G-hálózatokra épülő, nem önálló 5G-hálózatoknál elmaradt a várakozásoktól, amelyek nagy kapacitású, nagy megbízhatóságú, valamint kis késleltetésű és sok eszközt kiszolgáló hálózatokra vonatkoztak. 2022-re ugyanakkor megérett a keresleti és technológiai kínálati oldal is arra, hogy az önálló 5G-hálózatokra végre megtérülő üzleti tervek születhessenek. A kutatási kérdés arra irányul, hogy melyek a technológiai-gazdasági mozgatórugói az önálló 5G-hálózatok kiépülésének és milyen üzleti tervek mutathatnak pozitív megtérülést.
A várt kutatási eredmények szerint elsősorban az üzleti ügyfelek keresleti igénye specializálódik, amelyek a 4G-alapú, nem önálló 5G-hálózatok standard megoldásaival nem elégíthetők ki. Az önálló 5G-hálózat a specializált teljesítményű hálózatszeletek kialakításával tudja biztosítani az üzleti ügyféligények kielégítését, majd ezen logika mentén lakossági igények kielégítése is megtörténhet. Az üzleti terv megtérülését az internetkapcsolat biztosítását meghaladó 5G-adatmonetizáció, az önálló 5G-hálózatokkal elérhető hálózatszeletelésre épülő egyedi árazás, valamint a költséghatékonyság biztosíthatja.

1. Bevezetés

A mobil szélessávú internet-hozzáféréseknél a 4G áttörő szerepét követően jelentős szakmai várakozások övezték az 5G bevezetését. A szakirodalomban az 5G víziója három pilléren alapult: a kapacitás kiterjesztésén (eMBB – enhanced Mobile Broadband), nagy megbízhatóságú és alacsony késleltetésű kommunikáción (URLLC – Ultra-Reliable Low-Latency Communications) és a sok készülék egyidejű kezelésén (mMTC: massive Machine-Type Communication) [1]. Az első évek legtöbb kereskedelmi 5G-hálózata azonban messze elmaradt ezen közép- és hosszú távú vízióktól, és a gyakorlatban az ügyfelek csak nagyobb le- és feltöltési sebességet észlelhettek. Az 5G „ígérete” és „valósága” közti különbségnek nem utolsósorban gazdasági okai is vannak.

A hálózatok méretezésének csúcsidei volumentervezését már nem a hang-, hanem az adatforgalom határozza meg, miközben a volumen jelentős összetétel-változását a hang- és adatbevétel összetétel-változása jóval lassabban követi, amikor a jövedelmező mobil hangforgalom után már az adatforgalom monetizációs nehézségekről beszélünk. A mobilszolgáltatók emiatt pénzügyileg sokkal tudatosabbá és óvatosabbá váltak, így inkább a kevésbé kockázatos 4G-alapú, nem önálló 5G-hálózatok (5G NSA – Non Standalone) kiépítése mellett döntöttek, amíg nem látják biztosítva a beruházás-intenzívebb önálló 5G (5G SA – Standalone) kiépítés megtérülését.

Ez a tanulmány öt szakaszból áll. A bevezető rész a műszaki beruházások és üzemeltetés pénzügyi és controlling döntéshozatalainak keretét vázolja fel a nyugat-európai szolgáltatókra kitekintve. Ezután az 5G 2019. évi megjelenésének gazdasági környezetét mutatja be a szolgáltatói oldalról, majd a harmadik rész a 2022-ben bekövetkezett szolgáltatói környezet változásának főbb tényezőit írja le, amelyek az önálló 5G-hálózatok kiépítéséhez vezethetnek. A következő szakasz az önálló 5G-hálózatok bevételi és megtérülési mozgatórugóit mutatja be, valamint az ennek értékelését szolgáló üzleti terv keretét. Az ötödik rész az önálló 5G-hálózatok kiépítésében élen járó amerikai és európai szolgáltatók stratégiáját vázolja fel, – mind a piacra újonnan belépő kihívó, mind a bentlévő inkumbens szolgáltatók oldaláról –, majd az ezt követő összefoglaló, utolsó szakasz a főbb következtetéseket fogalmazza meg az 5G iránti várakozások tekintetében.

Pénzügyi, kontrolling oldalról megközelítve a vállalati teljesítmény szempontjából a bevételek (sales), a beruházások (CAPEX – capital expenditure) és a működési költségek (OPEX – operational expenditure), valamint ezeknek időbeli trendje és egymáshoz viszonyított aránya (bevételarányos beruházások – CAPEX/Sales) a meghatározók.

A mobilkommunikáció (2G) megjelenése óta a nyugat-európai távközlési szolgáltatókra vonatkozóan az *1. táblázat* a bevételnövekedési, a *2. táblázat* a bevételarányos beruházási trendeket mutatja be főbb időszakonként *(a táblázatokat lásd a következő oldalon)*.

A hangalapú bevételek vezérelte magas bevételnövekedési időszakot lezáró 2009-es gazdasági válság időszakában csökkenő, majd az ezt követő konjunkturális időszakban is csak stagnáló bevételeket láthatunk. Az adatforgalom jelentős, az Ericsson Mobility Report által 2021–2026 között CAGR 25% körülire becsült növekedés [3] ellenére az adatmonetizációs nehézségek miatt összességben a jelen időszakban is csak mérsékelt bevételnövekedés várható.

1. táblázat Bevételnövekedési trend nyugat-európai távközlési szolgáltatóknál Forrás: saját szerkesztés [2] alapján

Szakasz	Időszak	Jellemzői	Bevétel növekedése (időszaki CAGR %)
I.	***1995–2008***	A 2G (GSM 900) megjelenésétől a gazdasági válságig	6%-os átlagos növekedés
II.	***2009–2014***	A 2009-es válságtól a kilábalásig	2,8%-os átlagos csökkenés
III.	***2015–2020***	2015-től kezdődött konjunkturális időszak	0-1% körüli stagnálás
IV.	***2020–2025***	A COVID-időszaktól és az ukrán válságtól, előrejelzés	0-2% körüli mérsékelt növekedés

2. táblázat Bevételarányos beruházási trend nyugat-európai távközlési szolgáltatóknál Forrás: saját szerkesztés [2] alapján

Szakasz	Időszak	Jellemzői	Bevételarányos beruházás (CAPEX/Sales %)
I.	***1995–2008***	2G (GSM 900) és 3G kiépítése	20-30%
II.	***2009–2012***	Hozzáférési hálózat (RAN) életciklusának modernizációja	10-15%
III.	***2013–2019***	4G kiépítése, konvergens szolgáltatóknál párhuzamosan FTTx-hálózat kiépítésének megkezdése	12-17%
IV.	***2020–2025***	Hozzáférési hálózat (RAN) életciklusának modernizációja, az 5G kiépítése, konvergens szolgáltatóknál párhuzamosan FTTx-hálózat kiépítése	15-20%

A bevételarányos beruházások a kezdeti alapinfrastruktúra kiépítésekor 20-30%-ot mutattak. A mobilhálózatok felállását követően, még a hozzáférési hálózat (RAN – Radio Access Network) modernizációjakor is elegendő volt az éves árbevételek 10-15%-át visszaforgatni a beruházásokba. Azonban az 5G kiépítése egybeeshet a következő hálózatmodernizációval, valamint az FTTx-beruházások felfutásával mind a háztartások lefedése, mind a mobil bázisállomás adatátviteli igényeinek kielégítésére, amely következtében újra 15-20% köré emelkedett a CAPEX/Sales arány.

Mindezek következtében a szolgáltatók stagnáló-enyhén emelkedő bevételek és magas bevételarányos beruházások miatt a legtöbb esetben még nem az „igazi”, hanem a 4G-alapú, nem önálló 5G-hálózatok kiépítése mellett döntöttek.

2. Az 5G 2019-es bevezetésének műszaki-gazdasági környezete a szolgáltatók oldaláról

A 5G bevezetésének döntése körüli tényezők az alábbi négy alapvető csoportba sorolhatók: (i) fizetőképes fogyasztói kereslet, (ii) kínálat, amely a szolgáltatók által elérhető technológián és költségszintjén alapszik, (iii) készülékek elérhetősége, árszintje, (iv) beruházási igények priorizálása.

A *keresleti oldali* igények 2019-ben nem látszódtak egyértelműen. Az adatforgalom monetizációja az 5G időszakában sem vált tisztábbá, habár már fontos alkalmazási területként merült fel a vezeték nélküli helyhez kötött internetszolgáltatás (FWA – Fixed Wireless Access). Az 5G-bevezetések kapcsán a szolgáltatói kommunikációban is a magasabb le- és feltöltési sebesség, esetleg a nagyobb kapacitás (eMBB) jelent csak meg.

A *kínálati oldalon* a szolgáltatók még csak korai fázisban, gyakran koncepciók szintjén találkozhattak az önálló 5G kiépítéséhez szükséges technológiákkal, mint amilyen a hálózatvirtualizáció, a diszaggregáció, a nyílt hozzáférési hálózat (open RAN) vagy a hálózatszeletelés (network slicing).

Az akkori *készülékkínálat* sem segítette elő az önálló 5G-hálózatok kiépítése melletti döntést, hiszen még az újonnan eladott készülékek fele sem volt 5G SA-képes [4].

A *beruházások* tekintetében az éves bevételek 15-20 százalékának beruházásokba való visszaforgatása már kifeszített keret volt, amelyet a mobil és vezetékes optikai fejlesztések között kellett megosztani, miközben csak mobil oldalon az újabb életciklus-hálózatmodernizáció és 5G-kiépítés szükséglete összeérhetett.

Mindezek következtében kijelenthető, hogy 2019-ben, az 5G indulása küszöbén egyik tényező sem segítette a mobil szolgáltatókat, hogy nagy arányban az önálló 5G-hálózatok kiépítése mellett döntsenek. Rövid- és középtávon teljesen racionális pénzügyi és műszaki döntés volt a 4G-alapú 5G NSA-hálózatok kiépítése köztes, áthidaló megoldásként, amely jóval kisebb pénzügyi kockázattal is járt.

3. Az 5G műszaki-gazdasági környezetének változása a szolgáltatók számára 2022-re

Az időszak értékelése hasonló keretrendszerben történhet, kisebb változtatással: a keresleti és készülékoldal változatlanul vizsgálat tárgya marad, a kínálati oldalt célszerű lehet műszaki hálózati egységek mentén szétbontani a hozzáférési hálózati (RAN) és maghálózati (Core) elemekre. A beruházási keret változatlan a maximalizált 15-20%-os bevételarányos beruházással, így erre a szempontra külön nem térünk ki.

A *keresleti oldalon,* elsősorban az üzleti szegmensben (B2B) az 5G NSA képességét meghaladó fizetőképes üzleti igények jelentkeznek. Az 5G „ígéretének” három alappillére a kapacitásra (eMBB: 500 km/h sebességű mobilitás támogatása), magas megbízhatóságra és kis késleltetésre (URLLC: késleltetés <1 ms), valamint a számos eszköz párhuzamos kezelésére (mMTC: több, mint 1 millió eszköz/km^2) vonatkozóan különféle súlyozású kombinációkban már üzleti probléma megoldásaként fogalmazódik meg a B2B-ügyfelek körében. Később fokozatosan lakossági (B2C – Business to Consumers) ügyfélkörben is jelentkezhetnek olyan igények, amelyek kielégítéséhez 5G SA-megoldás lesz szükséges.

A *készülékoldalon* a készülékgyártók is felismerték a „igazi” 5G-ben rejlő potenciált és 2022-ben már az új készülékek több mint kétharmada 5G SA-képes, emiatt a teljes használatban lévő készülékállományban is növekszik az arányuk [4].

A legnagyobb átalakuláson a *műszaki kínálati oldal* megy keresztül. A 4G-től független 5G SA-hozzáférési hálózatok előfeltétele egy ezt támogató új 5G SA-maghálózat kiépítés cloud-nativ szolgáltatásalapú architektúrával. Továbbá a virtualizáció a hálózati képességeket és funkciókat (VNF – Virtualized Network Functions) nyújtó maghálózatra is kiterjed. Ezek mind nagyon jelentős, technológiai ugrást jelentő változások.

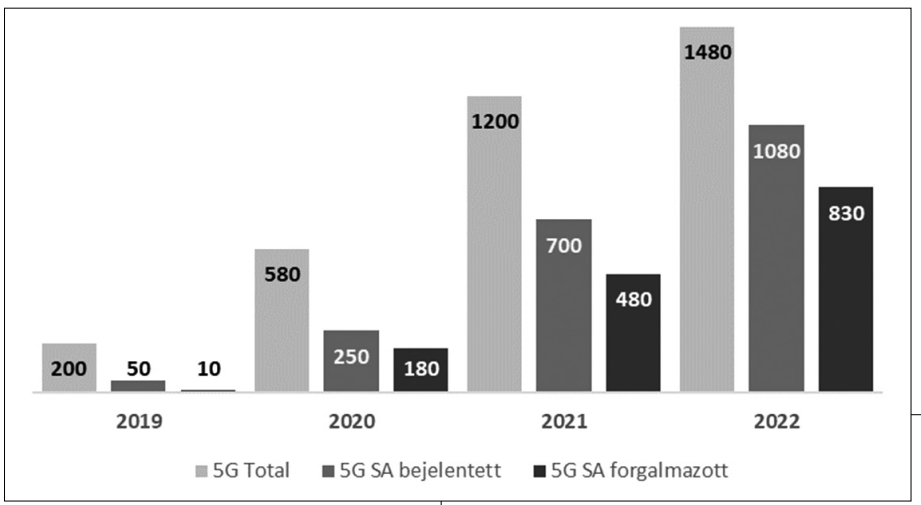
Ezzel párhuzamosan a hozzáférési hálózatban is megjelenhet a virtualizáció, diszaggregáció (hardver-szoftver és hálózati elemek funkciója között is), amellyel szoftveralapú hálózatok (SDN – Software Defined Networks) jöhetnek létre.

2022-re összességében letisztultabb technológiai fejlődési irányok, élénkülő fizetőképes kereslet és üzleti igények jelentek meg – elsősorban a B2B-szegmensből –, amelyek már közelebb visznek a pozitív, megtérülő szolgáltatói üzleti tervek (BC – Business Case) kialakításához.

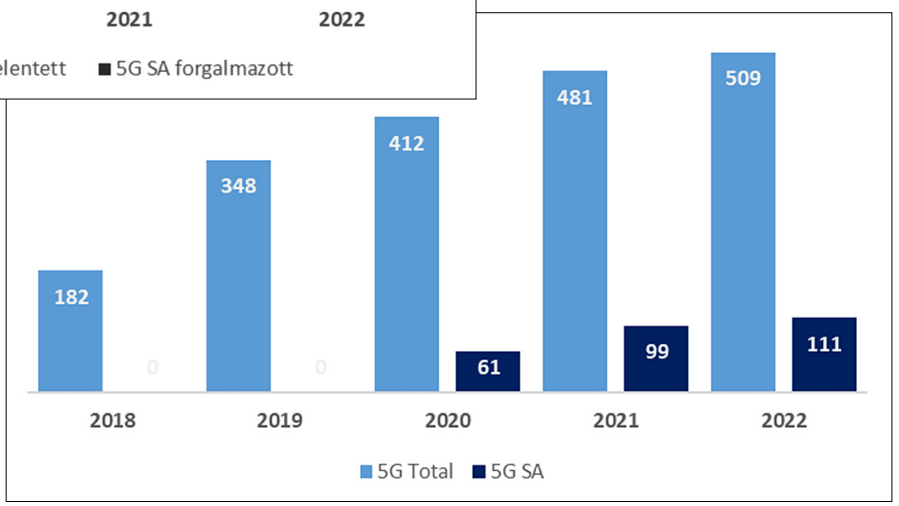
Jól látható, hogy a „valódi” 5G elérése nemcsak a hozzáférési hálózatokba való beruházásokat érinti, hanem a maghálózati beruházásokat is, adott esetben teljes maghálózati cserét végrehajtva, amelyek az életciklus cseréhez kapcsolva merülnek fel pénzügyileg reálisan. 2019-ben még a legtöbb szolgáltatónál ennek alacsony realitása volt, de a tisztábbá váló üzleti igények és hálózati életciklusban való előrehaladás miatt 2022-ben már egyre több szolgáltató kezdte meg SA-fókuszú maghálózat cseréjét.

A GSA reportja szerint 2020-ben 61 kereskedelmi 5G SA-hálózat üzemelt világszerte és 2022-re már 111-re emelkedett azon szolgáltatók száma, amelyek üzemeltetettek vagy beruháztak már az önálló 5G-be, így összességében az 5G-szolgáltatók 20%-a már elindult az SA irányába [5].

Az *1. ábra* az 5G SA-képes készülékek megjelenését, a *2. ábra* pedig az 5G-hálózatok számának növekedését mutatja.



1. ábra
5G SA-képes készülékek megjelenése a teljes 5G-képes készülékeken belül
Forrás:
saját szerkesztés [4] alapján



2. ábra
5G SA-képes hálózatok növekedése a teljes 5G-képes hálózatokon belül
Forrás:
saját szerkesztés [5] alapján

4. Az önálló 5G bevételi és megtérülési mozgatórugói, valamint az üzleti terv struktúrája

A bevételnövekedés fő mozgatórugója az üzleti, B2B-szegmens igényeinek egyedi megoldásokkal való kielégítése lehet. Műszakilag a standard szolgáltatási szinttől eltérő, az „igazi” 5G eMBB-, URLLC- és mMTC-képességek ügyfélre szabott kombinációt biztosítani képes hálózatszeletek révén valósíthatja meg a szolgáltató, aminek előfeltétele az SA-architektúrájú 5G.

4.1. Hálózatszeletelés jelentősége önálló 5G-bevételekben

A hálózatszeletek kialakítása az önálló 5G-hálózatok integráns része, amivel a szolgáltató azonos fizikai infrastruktúrán, szoftveralapon szétválasztva, ügyféligények szerint szeparált hálózatrészeken tud szimultán szolgáltatni. Azaz a hozzáférési (connectivity) típusú egységes szolgáltatás helyett egyéni ügyféligényre szabott szolgáltatásképesség-kombinációkat tud nyújtani például a teljesítmény, késleltetés, skálázhatóság, minőség, rendelkezésreállás, redundancia, biztonság dimenziói mentén. Mindez egyedi árazást, árdiszkriminációt tesz lehetővé árrugalmassági alapon, amely a bevételnövekedés fő motorja lehet.

A jelenlegi előrejelzések szerint a hálózatszeletek 90%-a üzleti, B2B-szegmenshez köthető, és a várt bevételnövekedések miatt a B2B-bevételek aránya 20% fölé emelkedhet az összes ügyfélszegmensen belül a nyugat-európai piacokon. Ugyanakkor a lakossági, B2C-szegmensben is lehetnek olyan felhasználói igények (például online játék, VR/AR, okosotthon), melyek kielégítésére felmerülhet a hálózatszeletelés alkalmazása [1].

Az Ericsson és az Arthur D. Little tanácsadó cég tanulmánya szerint [6] 400 darab 5G üzleti lehetőséget megvizsgálva legalább 30%-nál szükséges az 5G SA-hálózatszeletelés a megvalósításhoz. A hálózatszeleteléses megoldásokból 2025–2030 között átlagosan (CAGR) 25% feletti bevételnövekedés várható. Az esettanulmányok árbevétel-potenciálja alapvetően hat iparágból tevődik össze a *3. ábra* szerinti megoszlásban.

Az ábrázoltak alapján az előrejelzett bevételek kétharmada az alábbi TOP4-üzletágból származik: egészségügy, kormányzat, szállítás és energia. Az egészségügy területén a szolgáltatások mobil ICT- (Information and Communication Technology) eszközökkel való távoli nyújtása bír jelentős potenciállal. A kormányzati szolgáltatásoknál a közterület-felügyeleti, vészhelyzeti és kritikus szolgáltatások szerepe lesz jelentős. A szállítás és közlekedés terén a gyártás és üzemeltetés, valamint a gépjárműveknek nyújtott szolgáltatások lehetnek jelentősek. Az energiaiparból az előállítás, szállítás és finomítás területén juthat az „igazi” 5G nagyobb szerephez, jelentősebb bevételt generálva [5-6].

4.2. Üzleti modellek értékelési keretrendszere

A szolgáltatók az 5G SA bevezetését pénzügyi szempontból üzletiterv- (BC – Business Case) elemzésen keresztül validálják. Az üzleti tervben logikailag a bevétel növekedésének és a műszaki költséghatékonyság javulásának kell tudnia fedezni és megtéríteni legalább közép-hosszú távon a hozzáférési és maghálózati transzformáció költségeit.

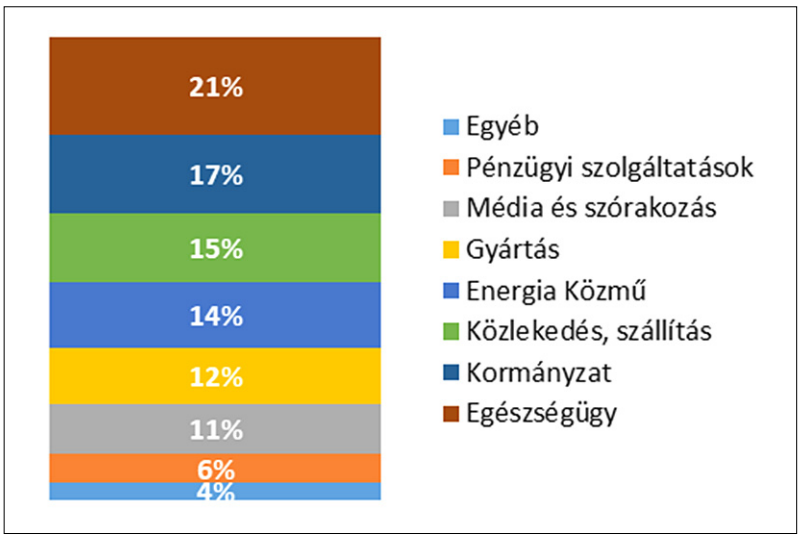
A *bevételek* növekedésének egyik része új ügyfelek megszerzéséből és új szolgáltatások bevezetéséből származhat, míg a másik része a meglévő ügyfelektől származó egységárbevételből (ARPU – Average Revenue Per User), amelynél fontos megjegyezni egyedi megoldások révén egyedi árdiszkrimináció, lefölöző és prémiumárazás lehetőségét is.

A *beruházási költségeknél* alapvetően három fő irányt kell jól megbecsülni és kontroll alatt is tartani. A maghálózat cseréjét, amely a hálózati képességekért felelős, a hozzáférési hálózat feljavítását SA-képességgel és mindkét hálózati elemnél a virtuális, cloud-alapú irányba való elmozdulás kialakításának költségeit kell számba venni.

Az *üzemeltetési költségeknél* az üzemeltetési és fenntartási költségek (O&M – Operational and Maintainance) virtualizációból várható költségmegtakarításával szemben a magasabb koordinációs költségek és az esetleges párhuzamos NSA-SA-hálózatüzemeltetésből származó költségtöbblet figyelembevétele szükséges. További figyelembevehető költségmegtakarítási lehetőség lehet az energiafelhasználás optimalizációja, valamint tisztán SA-hálózatra való áttérés után a hálózati architektúra egyszerűsödéséből származó megtakarítás.

Összességében a fenti dimenziók mentén összerakott szabad pénzáramlás jelenértékének tíz éves időtávon pozitívnak kellene lennie, azaz alapinfrastruktúra esetén is elvárható, hogy a mobilhálózatok tíz év körüli életciklusának legkésőbb a végére megtérülést mutasson az üzleti terv, annak alapvető érdekében, hogy pénzügyileg is racionális legyen a beruházás megvalósítása.

3. ábra
Az 5G előrejelzett bevételi potenciáljának üzletágak szerinti megoszlása
Forrás: saját szerkesztés [6] alapján



5. Szolgálatói stratégiák az önálló 5G bevezetésére az USA-ban és az EU-ban

Megvizsgálva az önálló 5G-hálózatok fejlesztéséről vagy már kereskedelmi indulásáról szóló bejelentéseket, egyértelműen kirajzolódik, hogy az adott piacon a kihívó szolgáltatóként megjelenő szereplők jártak élen a 5G SA korai bevezetése melletti döntésekben. Azaz azok a szolgáltatók, akik úgy gondolták, hogy az adott piacon akkor lehetnek sikeresek (ügyfélszerzésben, költséghatékonyságban), ha a méretgazdaságossági előnyökkel rendelkező piacvezető inkumbens szolgáltatótól nagyon eltérő, rövid távon nem lemásolható, differenciáló képességet tudnak felmutatni, amivel árban kedvezőbb vagy magasabb műszaki tartalmú ajánlatokkal megszerezhetik az inkumbens versenytárs ügyfeleit, és ráadásul költséghatékonyabb új technológia révén még profitábilisabb, fenntartható üzleti modell keretében. Mindez a stratégia különösképpen igaz az újonnan piacra lépő zöldmezős (greenfield) operátorokra.

Az 5G SA-hálózatok elterjedésében is az USA jelentősen megelőzi az EU-t, miután kettő jelentős 5G SA-hálózat épül ki.

Az egyik – a kábeltelevíziós háttérrel a mobilpiacra belépő – új szereplőhöz köthető, a Dish-hez, amely saját 4G-hálózattal nem is rendelkezik, és az 5G-hálózat kiépítésére egyből az SA-megoldást választotta teljesen racionális döntésként. Továbbá zöldmezős beruházóként lehetősége volt a legmodernebb architektúrát választania, azaz teljesen virtualizált, cloud-natív hálózatot épít AWS nyilvános felhő megoldást integrálva és természetesen önálló 5G-maghálózattal. A mobil szolgáltatást többszöri csúsztatást követően 2022-ben kezdte meg [8].

A másik jelentős önálló 5G-hálózat a T-Mobile USA-hoz köthető, amely szintén kihívó szereplőként már több mint egy évtizedes útkeresés után, majd versenytársa, a Sprint megvásárlását követően tudta stabilizálni a pozícióját, és az 5G bevezetéskor az SA mellett való döntés is hozzájárult a növekedési pályára állításhoz, amely eredményeként mára a vállalat piaci kapitalizációs értéke utolérte a piacvezető AT&T értékét [9].

Az EU piacán a két meghatározó szereplő, a Deutsche Telekom és a Vodafone hasonló stratégiát követ. Mindkét szolgáltatóra jellemző, hogy honos hazai piacon inkumbens hátterükkel sokkal inkább kiváró magatartást tanúsítottak, mert egy estleges elhamarkodott sikertelen hálózattranszformációnak súlyosabb pénzügyi hatásai lehettek volna.

A Deutsche Telekom – miközben az USA-ban az 5G SA egyik úttörője volt – Németországban 2022-ben csak 200 települést lefedő hálózatot épített ki [10].

A Vodafone is aktívabb szerepet vállalt a német piacon, ahol kihívóként volt jelen és már 2021-re 170 települést fedett le, de terveiben 2025-re a teljes lakosság lefedése szerepel. A honos angliai piacon ugyanakkor még csak 5G SA-hálózatszeletelés tesztjeinél tart.

6. Összefoglalás

Az 5G „ígérete” és „valósága” a kezdeti, 2019. évi időszakban a legtöbb szolgáltatónál még távol álltak egymástól, amelynek műszaki és pénzügyi okai is voltak. Az önálló 5G-hez kapcsolódó technológiák, mint az önálló 5G-maghálózat vagy hálózatszeletelés inkább elvi szinten szerepeltek csak, és a megfizethető 5G SA-képes készülékellátottság sem volt jellemző. Pénzügyileg a legtöbb szolgáltató a párhuzamos 5G és FTTx szélessávú hálózatfejlesztései miatt kifeszített beruházási keretekkel rendelkezett, amiben nem fért bele a barnamezős inkumbens szolgáltatóknál egy 5G SA hozzáférési és maghálózati transzformáció, főleg, hogy ha a 4G-hálózat életciklus modernizációja sem volt esedékes.

2022-re azonban egyértelmű fordulást láthatunk az önálló 5G-hálózatok irányába, mert fizetőképes kereslet mutatkozik az üzleti szegmensből hálózatszeleteléssel megvalósítható egyedi hálózatképesség (pl. kapacitás, késleltetés) kombinációkra. A készülékellátottság folyamatosan javul, a technológiai megoldások pedig széles körben elérhetőek, amellyel fokozatos maghálózati, majd a hozzáférési hálózati transzformáció is meg tud indulni. A szolgáltatói üzleti tervekben bevétel oldalról az üzleti szegmens igényei, költségoldalról az új technológiából és virtualizációból származó megtakarítások eredményezhetnek megtérülést a hálózat életciklusának végére.

Jól látható, hogy az egyes piacokon kihívó, esetleg zöldmezős új szereplők nyitottak először megkülönböztetés és költséghatékonyság miatt az új önálló 5G-hálózat irányába, amelyeket néhány év lemaradással követnek a barnamezős inkumbens szolgáltatók is, mert az önálló 5G-képesség hiánya hosszú távon már egyértelmű versenyhátrányt jelenthet.

Hivatkozások

[1] Kyllesbech, L.: 5G Standalone – Network Slicing, a Bigger Slice of the Value Pie (Part II), 2022. *https://www.linkedin.com/pulse/5g-standalone-network-slicing-slice-pie-part-ii-dr-kim/?trackingId=qPzp9akOIUdgBff%2F0D4s3w%3D%3D*

[2] Kyllesbech, L. The Nature of Telecom Capex, 2022. *https://techneconomyblog.com/2022/07/06/the-nature-of-telecom-capex/*

[3] Ericsson: Mobility report, 2022. *https://www.ericsson.com/49d3a0/assets/local/reports-papers/mobility-report/documents/2022/ericsson-mobility-report-june-2022.pdf*

[4] GSA report: 5G-Standalone August 2022 Summary Report, 2022. *https://gsacom.com/paper/5g-standalone-august-2022-summary-report/*

[5] GSA report: 5G-Standalone November 2022 Summary Report, 2022. *https://gsacom.com/paper/5g-standalone-november-2022-summary/*

[6] Ericsson – Arthur D. Little: Network slicing: a go-to-market guide to capture the high revenue potential, 2021. *https://www.ericsson.com/49c844/assets/local/networks-slicing/docs/network-slicing-value-potential.pdf*

[7] Ericsson: Network slicing: Top 10 use cases to target, 2021.
https://www.ericsson.com/49c6d0/assets/local/networks-slicing/docs/top-network-slicing-use-cases.pdf
[8] Fierce Wireless: Vodafone, Ericsson tout 5G network slicing trial as first in U.K., 2022.
https://www.fiercewireless.com/5g/vodafone-ericsson-tout-5g-network-slicing-trial-first-uk
[9] Opensignal: Understanding the mobile experience on T-Mobile's standalone 5G network, 2021.
https://www.opensignal.com/2021/02/18/understanding-the-mobile-experience-on-t-mobiles-standalone-5g-network
[10] Fierce Wireless: Deutsche Telekom touts 5G SA rollout in Germany, 2022.
https://www.fiercewireless.com/5g/deutsche-telekom-touts-5g-sa-rollout-germany

A szerzőről

FÖLDES GÁBOR távközlési gazdasági szakértő, 20 éves pénzügyi és szabályozási gyakorlattal a tudományos és üzleti kontrolling területeken. 2004-ben végzett a Budapesti Közgazdaságtudományi és Államigazgatási Egyetemen. Ezt követően 15 évet dolgozott a távközlésben, üzleti kontrolling területen a Deutsche Telekom, Magyar Telekom és Telenor Magyarország vállalatoknál. Később az NMHH-hoz csatlakozott a szabályozói pénzügyi modellek területére, és támogatta a BEREC költségszabályozási feladatait. Jelenleg kontrolling területen dolgozik a Vodafone Intelligent Solution-ban. 2020/22-ben PhD-tanulmányokat folytatott a Budapesti Műszaki és Gazdaságtudományi Egyetemen – költséghatékony 5G-mobilhálózatok kialakítása kutatási témában –, hálózatmegosztásokat, toronyvállalatokat, virtuális és nyílt hozzáférési hálózatokat vizsgálva.

Digitális identitáskezelés átalakulóban: önrendelkezésű identitások

KOCSIS IMRE
BME Méréstechnika és Információs Rendszerek Tanszék
kocsis.imre@vik.bme.hu

Kulcsszavak: önrendelkezésű identitás, decentralizált identitás, ellenőrizhető tanúsítványok, EBSI, blokklánc

A folyamatosan növekvő összekötöttséggel és digitalizációval az online és digitális azonosítás klasszikus modelljei jó ideje egyre nehézkesebben alkalmazhatóak. A változást az önrendelkezésű identitások (Self-Sovereign Identity) jelenleg folyamatban lévő szabványosítása és bevezetése hozza el. A cikk áttekinti az SSI filozófiáját és legfőbb támogató W3C-szabványait az elosztott identitásokra és az ellenőrizhető tanúsítványokra. A cikk fontos használati eseteken keresztül demonstrálja az SSI alapvető működési modelljét, és bemutatja a legfontosabb, már ma is működő támogatóhálózatokat, ezek között a European Blockchain Service Initiative-t is, továbbá a megjelenőben lévő ipari használati esetek is röviden áttekintésre kerülnek.

1. Bevezetés

Világunk és életünk egyre inkább hálózatokhoz kötötté és „kiberfizikaibbá" válásával a digitális és online identitáskezelés klasszikus megközelítései egyre kevésbé képesek az azonosítás és tulajdonság-meghatározás biztonságos, hatékony és sok fél között interoperábilis támogatására. Az önrendelkezésű identitások – angol terminológiában Self-Sovereign Identity (SSI) – filozófiája, szabványai és technológiái hosszú idő óta formálódnak viszonylag szűk szakmai körökben; az utóbbi években azonban terjedésük széles körben is megindult az újszerű műszaki képességek és részben újszerű vállalati, szabályozói és állami igények egymásra találásával.

Cikkünk áttekintő bevezetést nyújt az önrendelkezésű identitások kialakulóban lévő szakterületére és szemlélteti a jelenlegi főbb megvalósításokat, azok alkalmazásaira is rámutatva. Kitérünk az Európai Unión belüli digitális azonosítás és az SSI kapcsolatára, végül rámutatunk az SSI – jelenleg még részben felderítés alatt álló – ipari és telekommunikációs relevanciájára.

A cikk következő szakasza az önrendelkezésű identitások filozófiáját és alapfogalmait mutatja be. A harmadik szakasz a jelenlegi szabványtámogatást ismerteti, majd a jelenleg rendelkezésre álló főbb platformokat és hálózatokat veszi sorra. Az ötödik szakasz az önrendelkezésű identitások tágabb európai kontextusát szemlézi, végül a megjelenőben lévő ipari alkalmazásokra mutat rá.

2. Önrendelkezésű identitások

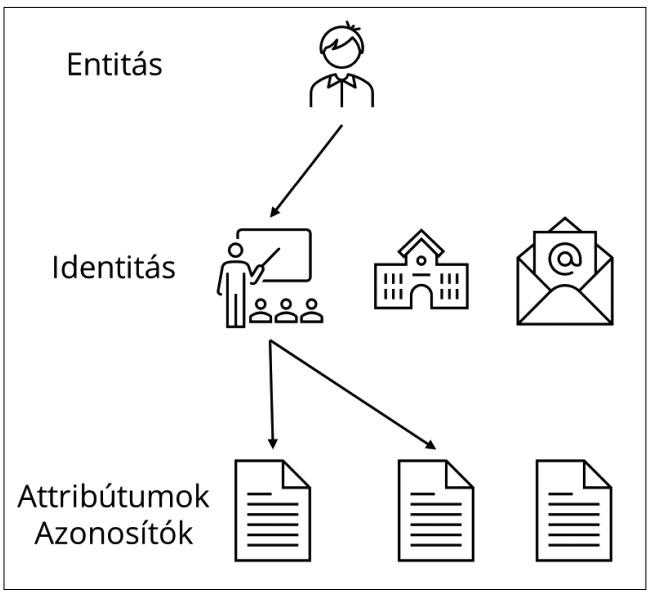
Általános értelmű fogalomként egy természetes személy identitása azt ragadja meg, hogy „ki ő"; mik azok a tulajdonságai, amik másoktól megkülönböztetik *(1. ábra)*.

A hétköznapi életben mindannyian számos identitással rendelkezünk: más-más jellemzőink relevánsak a munkahelyen, az állammal szemben és a magánéletben. Identitásunk részét képezik végzettségeink, a hivatalos személyazonosságunk, különböző jogosítványaink és felhatalmazásaink is. Ezek bizonyítására elfogadunk (visszavonható) „tanúsítványokat" különböző állami és nem állami forrásoktól, de az természetes igény marad, hogy az identitásainkkal magunk rendelkezhessünk és ne az identitásunkkal kapcsolatban tanúsítványokat kiállító felek.

2.1. Az online identitáskezelés klasszikus megközelítései

Ezt a természetes igényt ma nem elégítik ki sem az általános értelemben vett digitális azonosítás, sem pedig az online azonosítás széles körben elterjedt megoldásai, és ahogy életünk egyre inkább digitalizálttá válik, úgy egyre égetőbb kihívásként jelentkezik az identitások önrendelkezésűvé tételének megvalósítása.

1. ábra
A klasszikus identitásmodell: entitás, identitás, attribútumok



Az online világ meghatározó protokolljai, mint az IP és a TCP, a csomópontok és csomóponti szolgáltatások, és nem a felhasználók azonosítását támogatják – utóbbi nem is volt cél tervezésük során. Így azonban az internet egy, az alkalmazási réteg által is használható „bizalmi réteg" nélkül maradt.

Kezdetben az internetalapú alkalmazások külön-külön saját azonosítási megközelítéseket vezettek be – legtöbbünk mind a mai napig felhasználónév-jelszó párosok tucatjait kénytelen karbantartani. Ezen identitások felett azonban a szolgáltató rendelkezik; saját belátása szerint törölheti és módosíthatja őket. Érdemes megjegyeznünk, hogy mindez a szabványos, X.509 tanúsítvány alapú, jóval általánosabb használati körű digitális identitáskezelésre is igaz, – csak hierarchiába szervezett módon.

Emellett évtizedek óta probléma az is, hogy a szolgáltatótól az „identitás" ki is szivároghat, ezért is javasolt ragaszkodnunk az egyszer használatos jelszavakhoz. Végeredményben a klasszikus, centralizált modellben nagyszámú identitást hozunk létre az esetek többségében alapvetően egy cél érdekében; identitásazonosításra a szolgáltatás igénybevételéhez *(2. ábra)*.

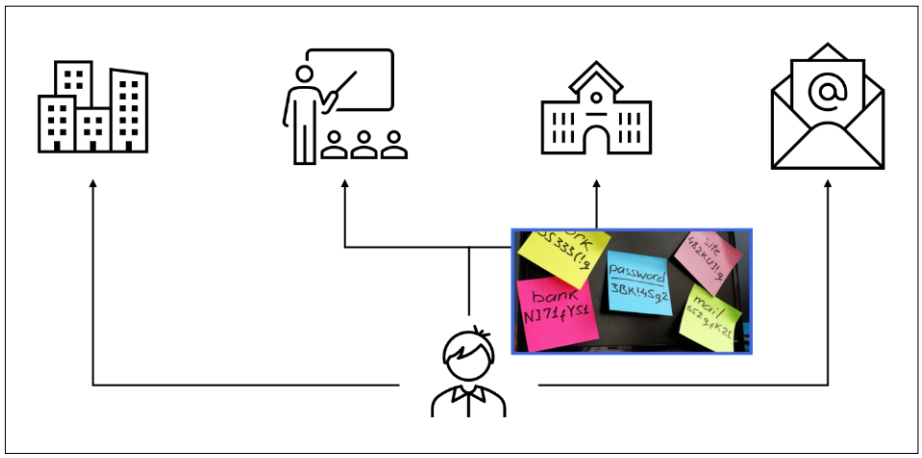
Ezt a töredezettséget próbálták a 2000-es években először néhány szervezetek közötti identitás-federációs megoldással, majd olyan „felhasználóközpontú" identitáskezelési sémákkal (pl. OpenID, OAuth) orvosolni, melyek felhasználói hozzájáruláshoz kötik az identitások szolgáltatások közötti megosztását. Végeredményben egy olyan online világ alakult ki – a technológiák teljes képességkészletétől függetlenül – ahol néhány nagy alkalmazásszolgáltató identitásszolgáltatás-federációt is végez. (Gondoljunk csak arra, hogy hány kisebb szolgáltatónál lehetséges a google-fiókunkkal vagy a facebook-fiókunkkal belépni *(3. ábra)*.

Bár így a töredezettség csökken és az identitások szolgáltatások közötti átjárhatósága nő, a végfelhasználó az identitáskezelés tekintetében még inkább kiszolgáltatottá válik egy olyan piaci (és nem állami) szereplő felé, amelyik ingyenes szolgáltatása felhasználási feltételeiben sokszor explicite nem vállal semmilyen érdemi garanciát.

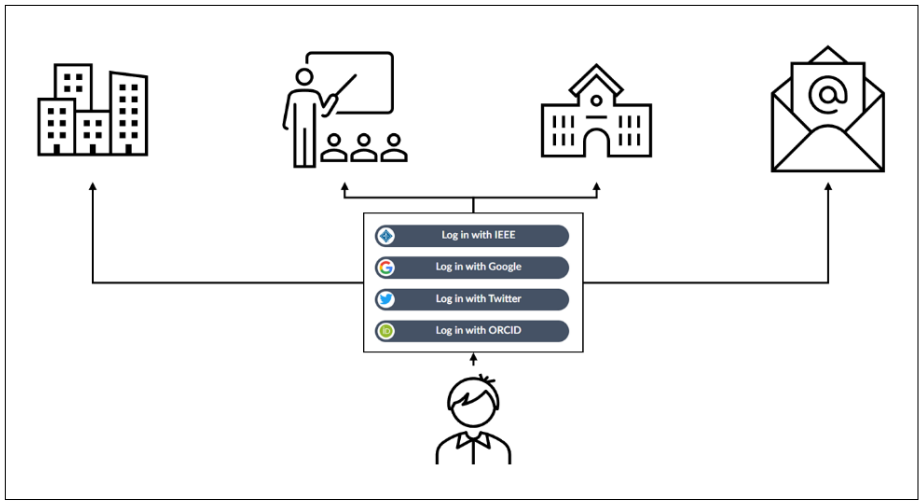
Mind a mai napig a centralizált és federált jellegű identitáskezelési megoldások jellemzik a széles körben alkalmazott digitális identitáskezelési megoldásokat (a hierarchikus tanúsítványkezelés mellett). Évek óta zajlik azonban a radikális változások előkészítése több terület összefogásával, amelyek között – az adatbiztonság identitáskezelés aldiszciplínája mellett – kiemelt szerepe van a blokklánc-technológiáknak és a közszférának is.

2.2. Az önrendelkezésű identitások jellemzői

A 2010-es évek közepén kezdtek körvonalazódni az „önrendelkezésű identitásokkal" kapcsolatos alapvető követelmények. Nagyhatású cikkében Christopher Allen



2. ábra
Klasszikus online identitáskezelés



3. ábra
Felhasználói hozzájárulással federált online identitáskezelés

kriptográfus ezeket tíz elvként definiálja [1], melyekből jelen cikk szerzője a következőket tartja a legfontosabbaknak:

– Kontroll – a felhasználó teljes mértékben rendelkezzen az identitásai felett.
– Hozzáférés – a felhasználó teljes mértékben férjen hozzá az identitásaival kapcsolatos adatokhoz.
– Perzisztencia – a felhasználó identitásai addig létezhessenek, ameddig csak a felhasználó azt szeretné (akár örökké).
– Interoperabilitás – az identitások széles körben használhatóak legyenek, geográfiailag és informatikailag is.
– Hozzájárulás – az identitás bármilyen használatához a felhasználó hozzájárulása legyen szükséges.
– Adatminimalizáció – az identitással kapcsolatos forgatókönyvekben legyen lehetőség csak a minimálisan szükséges információ közlésére.

2.3. A blokklánc-technológia hatása

Az SSI követelményeinek egyértelművé válásával párhuzamosan kialakultak és elterjedtek az üzemeltetés tekintetében szabad csatlakozású (unpermissioned) blokklánc-technológiák, mint a Bitcoin és az Ethereum [2], majd később a szervezetek jogosultságkezelt és diverz köre által üzemeltetett elosztott főkönyvön alapuló technológiai platformok (Distributed Ledger Technology, DLT) [3], mint amilyen például a Hyperledger Fabric [4]. A blokkláncok széles körű elterjedése az SSI szempontjából két alapvető elemmel gazdagította az informatikai rendszertervezés mintakészletét: a „többes hitelesítésű” adatregiszterekkel és az önrendelkezésű kriptográfiai azonosítással.

A blokkláncok koncepcionálisan „többes hitelesítésű” adatregiszterek: példájuk azt mutatja, hogy lehetséges létrehozni és hosszú távon sikeresen üzemeltetni olyan elosztott kvázi-adatbázisokat, ahol a módosítások befogadásáról az üzemeltető csomópontok valamilyen elégséges többsége dönt; a többség által az „adatbázis-protokoll” betartását vagy (kripto)pénzbeli ösztönző-mechanizmusok, vagy az üzemeltető közösség jogi megállapodásai szavatolják.

A blokklánc-technológiák felhasználói oldala hasonlóan fontos, a széles körű használat szempontjából új elem. A blokklánc-hálózattal szemben egy felhasználó a tranzakciók digitális aláírásával bizonyítja identitását. Ez az identitás azonban pusztán egy – a felhasználó által szabadon létrehozott aszimmetrikus kriptográfiai kulcspár publikus részéből levezethető – kriptográfiai álnév, melyhez a privát kulcsot az ún. „tárcájában” (wallet) tartja, mely egy hardveres vagy szoftveres megvalósítású céleszköz. A publikus blokklánc-hálózatokban egy adott publikus kulcsra (abból képzett „címre”) átvezetett kriptopénz csak a privát kulcs segítségével adható tovább az elosztott főkönyv által megvalósított könyvelésben. A felhasználó teljes mértékben maga rendelkezik a privát kulcshoz kötött identitása fölött a blokklánc-rendszerekben, és ezen identitásokból tetszőleges számút hozhat létre.

E két elemből származik a megjelenőben lévő SSI-szabványok és technológiai ökoszisztémák két alapgondolata: az identitások és a tanúsítványok erős szétcsatolása és a kriptográfiai alapú, önrendelkezésű identitáskezelés ellenőrizhető adatregiszterekben.

2.4. Decentralizált identitások

A centralizált és a legtöbb federált identitáskezelési megoldás alapvető gyengesége, hogy egységben kezeli az identitást – pontosabban: az identitás tulajdonlásának bizonyítását – és az identitásnak pusztán logikailag részét képező, időben változékony tanúsítványokat. Valójában ezek megfelelő kriptográfiai és adatregiszter-támogatással szétcsatolhatóak.

A séma egyik fele egy azonosító és az azonosító tulajdonlásának bizonyításához szükséges követelmények (tipikusan: publikus kulcs, aminek megfelelő digitális aláírást kell tudnia létrehozni), melyeket egy úgynevezett ellenőrizhető adatregiszterben (Verifiable Data Registry) tárolunk. Utóbbi lehet blokklánc, de akár centralizált, elégségesen megbízhatónak tartott – például állami üzemeltetésű – nyílt identitáskezelő szolgáltatás is. Ha az identitásokat decentralizált platformon kezeljük, akkor jellemzően decentralizált identitásként (decentralized identity, DID) hivatkozunk rájuk; egyébként általános értelemben vett (kriptográfiailag) verifikálható identitásként (verifiable identity, VID). Az explicit, tanúsítványoktól és jogosítványoktól függetlenített identitáskezelés lehetővé teszi az identitások felhasználó általi kezelésének számos forgatókönyvét, például számos identitás létrehozása, vagy az identitásokhoz kötött authentikációs követelmények dinamikus kezelése (egy alapvető példa a nyilvános kulcs frissítése).

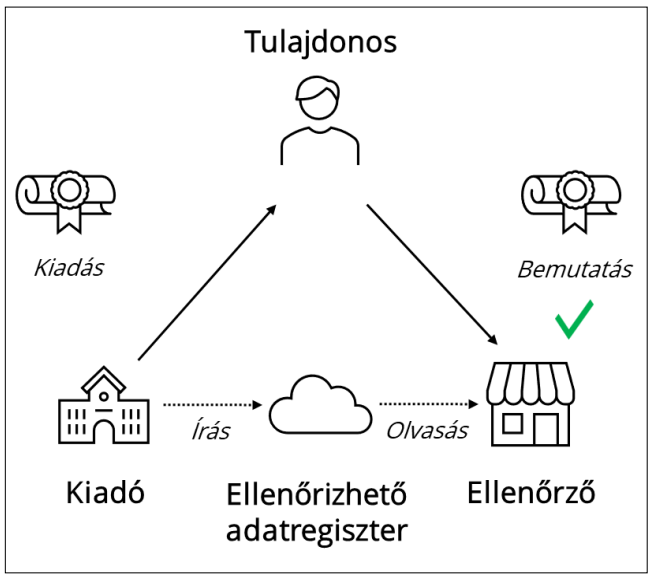
2.5. Ellenőrizhető tanúsítványok

A séma másik felét az úgynevezett ellenőrizhető tanúsítványok (verifiable credential, VC) adják. Megjegyzendő, hogy bár az angol terminológiában elválik a „credential” és a „certificate” fogalma, magyar fordításként mégis az „ellenőrizhető tanúsítvány” tűnik a legalkalmasabbnak, annak ellenére, hogy a tanúsítványt klasszikusan „certificate” értelemben használjuk.

Az ellenőrizhető tanúsítványok olyan digitális dokumentumok, melyeket egy kibocsájtó egy DID-re ad ki, a saját DID-jában meghatározott módon hitelesítve (digitálisan aláírva) ezeket. Ezek módosíthatóak és visszavonhatóak anélkül, hogy a visszavonás magát a DID-et érintené. Érett megoldások léteznek arra, hogy a tanúsítvány-ellenőrző felek egy tanúsítvány vissza nem vont voltát privacy-védő módon, a kibocsátók által nyilvánosan (például egy blokklánc felett) publikált és karbantartott, úgynevezett kriptográfiai akkumulátorokon keresztül végezzék. Maguk a tanúsítványok azonban nem kell, hogy ellenőrizhető adatregiszterben publikáltak és követettek legyenek; azokat tulajdonosuk tárolhatja kizárólag a tárcájában is.

A képet a blokklánc-technológiákban alkalmazásuk hatására széles körben ismertté vált tudásmentes bizonyítások (zero-knowledge proof, ZKP) [5] teszik teljessé:

megfelelően létrehozott VC-khez létrehozhatóak olyan ellenőrizhető matematikai bizonyítások, melyek csak a VC-be foglalt információk egy részét fedik. A szokásos iskolapélda a digitális személyi, mint ellenőrizhető tanúsítvány: míg a tanúsítványban számos érzékeny információ szerepel (pl. név, születési név, pontos születési idő és hely, anya leánykori neve), alkohol és dohánytermék vásárlásánál az ellenőrző fél számára ezek nagy része nem releváns – pusztán az, hogy a felmutatott tanúsítványt a megfelelő állami szerv adta-e ki, az érvényes-e jelenleg és hogy a tulajdonos 18 évnél idősebb-e. Ez utóbbi tulajdonság bizonyítható matematikailag ZKP-val úgy, hogy nemhogy a többi, tanúsítványba foglalt jellemző, de még a pontos születési idő sem kerül átadásra.



4. ábra Az SSI bizalmi háromszög

2.6. A bizalmi háromszög

Ezen elemekből áll össze a *4. ábrán* látható SSI „bizalmi háromszög” (trust triangle). A tanúsítványokat kiadó felek (Issuer) a tanúsítvány-tulajdonos (Holder) számára ellenőrizhető tanúsítványokat ad ki, az ellenőrzéshez szükséges adatokat – főként a saját identitását és a visszavont tanúsítványok kriptográfiai akkumulátorát – egy ellenőrizhető adatregiszterben kezelve; a tanúsítvány-ellenőrző fél a megfelelőségéről megbizonyosodhat a kiadóval való kommunikáció nélkül, pusztán az ellenőrizhető adatregiszterre támaszkodva.

3. Szabványtámogatás

Komoly technológiai-kísérleti előzmények és hosszú előkészítési folyamat után a World Wide Web Consortium (W3C) 2022 nyarán adta ki DID-ajánlásának 1.0 verzióját [6]; az erre épülő ellenőrizhető tanúsítvány adatmodell-ajánlás (Verifiable Credentials Data Model) már az 1.1 változatnál tart [7]. E két szabvány W3C-ajánlásként megjelenése alapvető fontosságú, annál is inkább, mert mögöttük egy sor, részben már ma is ipari érettségű implementáció áll.

A DID-szabvány kettéválasztja a DID „címzését” és tartalmát. Szigorúan véve egy DID egy *did:<metódus>:<metódus-specifikus azonosító>* szerkezetű karaktersorozat, amely egy úgynevezett DID-dokumentumra (DID Document) oldható fel *(5. ábra)*; a feloldás módját a metódus rész vezérli. A W3C-metódusok egy kanonikus regiszterét [8] is karban tartja; ezek között megtalálhatóak

– teljesen decentralizált megközelítések, mint az Ethereum nyilvános blokklánc feletti Ethereum Name Service-en (ENS) keresztül való DID-dokumentumfeloldás és -kezelés;
– szervezetek diverz, de jogosultságkezelt közössége által üzemeltetett blokkláncok feletti névfeloldás, melynek hozzáférése azonban nyílt: ilyen az EU tagállamai által üzemeltetett European Blockchain Services Infrastructure (EBSI); és
– klasszikus értelemben centralizált, illetve szervezetek közössége által üzemeltetett és csak számukra hozzáférhető blokkláncok.

Blokklánc alapon vagy sem, a DID-dokumentumokat kezelő ellenőrizhető adattár célja, hogy egy olyan megbízható platformot adjon, ahol a felhasználók biztosak lehetnek identitásaiknak az SSI-elveknek megfelelő kezelésében, és amit a DID-ekre épülő VC-ket ellenőrző felek is elégségesen megbízhatónak tartanak.

A VC-k adattartalmát a szabvány nem határozza meg, lehetséges szemantikájuk köre igen széles, az elektronikus személyitől a diplomákon és különböző felhatalmazásokon keresztül a kereskedelmi pontgyűjtőprogramok támogatásáig. Így az ellenőrizhető adattár egy másik fontos funkciója, hogy a VC-sémákat kezelje, és sokszor itt történik az ellenőrizhető tanúsítványok kiadásához kapcsolódó meghatalmazási lánc kezelése is (például egy központi oktatási hivatal által diplomák kiadására feljogosított szervezetek identitásainak kezelése).

A DID- és VC-szabványpár alapvetően adatmodell-jellegű; tényleges alkalmazásukhoz a következő protokollhierarchia szükséges:

- Az egyfajta „közszolgáltatásként” üzemeltetett ellenőrizhető adatregiszterek kezelési és elérési protokollja.
- A kommunikáló felek (illetve adattárcáik) közötti közvetlen kapcsolatfelépítés és közvetlen DID-csere protokollja.
- Adatcsere, kiadási és ellenőrzési protokollok az ellenőrizhető tanúsítványok szintjén (pl. kihívás-válasz azonosításra).
- Alkalmazási szintű protokollok és bizalmi ökoszisztémák (pl. oltási tanúsítványok és elfogadásuk szabályai).

Ezen protokollrétegeknek a szabványosítási törekvései részben folyamatosan zajlanak, részben pedig már ma is több lehetőség áll rendelkezésre megvalósításukra. Kiemelendő, hogy a műszaki interoperábilitás mellett mind a négy rétegben többszereplős irányítási problémák (az angol governance szó értelmében) is megoldandóak:

- Az ellenőrizhető adatregisztereknek megbízhatónak és pártatlannak kell lenniük: tartalmukat vagy egy

többszereplős konszenzusnak, vagy (állami) szabályozásnak kell szavatolnia.
- Meg kell teremteni az adattárcák és a végfelhasználók kommunikációját végző ágensek szabványosítási és adatvédelmi megfelelőségének, valamint megfelelőség-ellenőrzésének kereteit.
- Az ellenőrizhető tanúsítványok használatának szintjén a tanúsítványokkal kapcsolatos bizalom megteremtése nem műszaki kérdés; például a VC-ként kiadott diplomák hitelességét javallott egy „államilag regisztrált kibocsátó” mechanizmussal támogatni.
- Végül az ökoszisztémák szintjén a széles körű bizalom megteremtéséhez is szükséges a governance; a műszakilag természetesen adódó bizalmi integrációs lehetőségek (például oltási igazolványok sokszereplős elfogadása, vagy a kölcsönös diplomaelfogadás az EU országaiban) gyakorlati bizalmi aspektusai közel sem maguktól értetődőek.

A technológia és a governance rétegeit, valamint ezek még előttünk álló kihívásait független szereplőként elsődlegesen a Linux-alapítvány által létrehozott „Trust over IP” (ToIP) alapítvány vizsgálja [9]; a ToIP-alapítvány munkája részeként egy referencia-architektúrát is definiált.

4. Platformok és hálózatok

Az SSI-elvek és kialakuló szabványtámogatás első ipari minőségű és máig meghatározó platformját a Sovrin-alapítvány [10] hozta létre. A Sovrin-hálózat decentralizált, de üzemeltetésében nem szabad csatlakozású; üzleti modelljének alapgondolata az, hogy az adatregiszter írási operációiért kell fizetni. Bár önmagában a hálózat létrejötte is fontos SSI-mérföldkő volt (a „MainNet” 2017 óta üzemel), de talán még fontosabb, hogy az alapítvány a kódbázist a Hyperledger alapítvány égisze alatt nyílt forráskódúvá tette a Hyperledger Indy projektben (egyes komponensek később kiszervezésre kerültek a Hyperledger Aries projektbe).

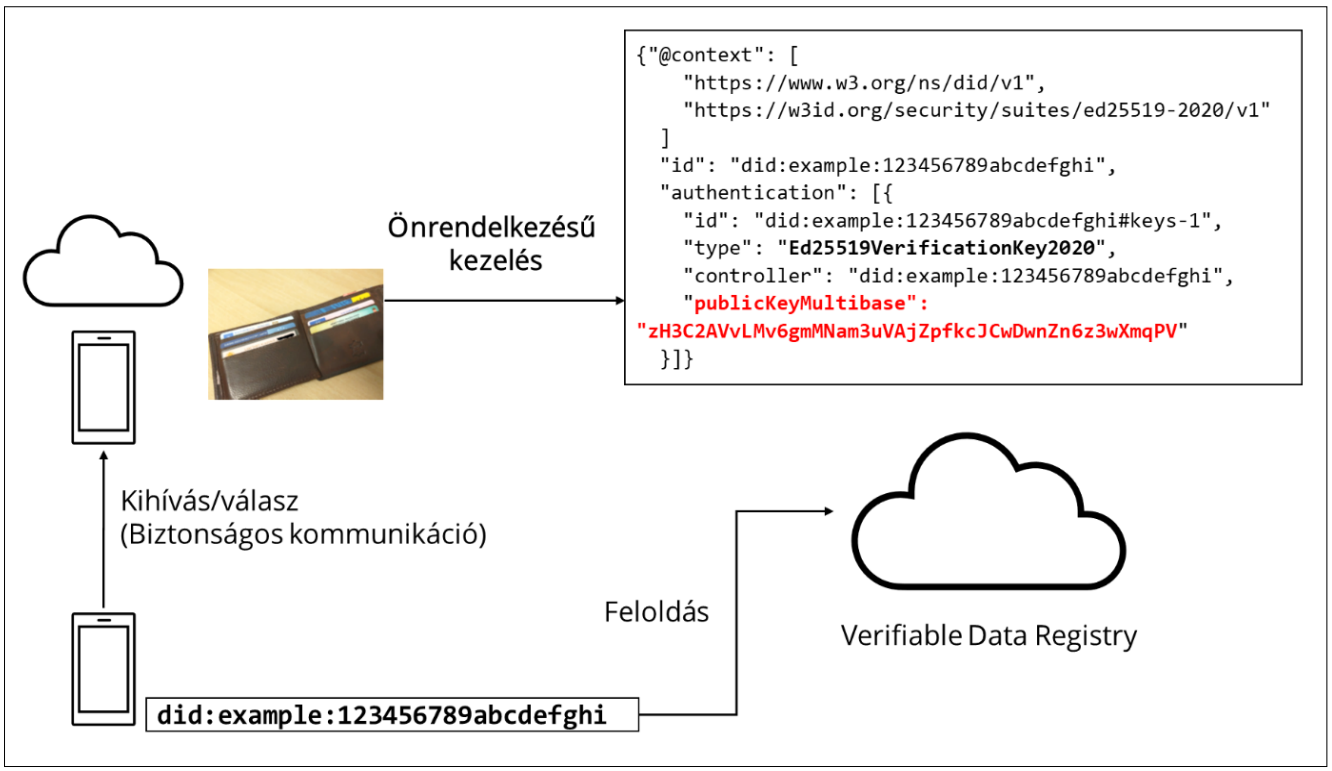
Az önrendelkezésű identitáskezelés és a felette megvalósítható ellenőrizhető tanúsítvány-interoperabilitás mind a digitális, mind az online identitáskezelés területén olyan transzformációs potenciállal rendelkezik, mely természetessé teszi azt, hogy bevezetésüket állami szereplők szorgalmazzák. Ma már több állami jellegű vagy erősen állami támogatású, illetve államközi, SSI-jellegű identitásszolgáltató hálózat üzemel, vagy áll kialakítás alatt. Ezek közül a talán legmeghatározóbbak az EBSI, az Alastria, a LACChain és az IDunion.

4.1. European Blockchain Services Infrastructure

Az EBSI-projektet [11] 2018-ban indította az Európai Bizottság és a European Blockchain Partnership (EBP), a tagállamok mellett Norvégia és Liechtenstein részvételével. A projekt egy olyan, a tagállamok kijelölt szervezetei által közösen üzemeltetett blokkláncplatformot épít ki, mely határokon átívelő, (tág értelemben vett) bizalmi szolgáltatásokat kínál.

Elsődleges célja jelenleg egy SSI-jellegű identitáskezelési réteg (European Self Sovereign Identity Framework, eSSIF) segítségével a közszolgáltatások támogatása, mint például a diplomák és egyéb felsőoktatási tanúsítványok határokon átnyúló ellenőrzése; távlatilag a rendszert az európai üzleti szféra számára is elérhetővé

5. ábra W3C DID és feloldása DI-dokumentumra



tervezik tenni (például adathitelesítésre, dokumentumkezelésre és eszközkövetésre). A kísérleti hálózat és használati eseteinek kidolgozása után az EBSI éles hálózatának kialakítása jelenleg is zajlik. Hazánk jelenleg három EBSI-csomópontjának egyikét a Budapesti Műszaki és Gazdaságtudományi Egyetem üzemelteti, a szerző szakmai vezetésével.

Az Alastria és a LACChain az EBSI-hoz nagyban hasonló kezdeményezések, előbbi az Ibériai-félszigeten, utóbbi Latin-Amerikában.

4.2. IDunion

Az IDunion [12] egy elsődlegesen német szövetségi támogatásból létrejött, de szervezetileg 2022 nyara óta európai szövetkezetként (Societas Cooperativa Europaea, SCE) üzemelő szervezet és hálózat. Sajátossága, hogy a természetes személyek mellett nagy hangsúlyt fektet a szervezetek és a „dolgok" (things) identitáskezelésének támogatására is – utóbbit deklarált módon az ipari IoT és az Ipar 4.0 egy fontos elemének tekintve.

A hálózat nem csak a szervezetek (ideértve a cégeket is) és az állampolgárok közötti megbízható azonosítást támogatja, de a szervezetek közötti biztonságos adatmegosztást is. A hálózat éles szolgáltatása várhatóan 2023 második felében indul.

4.3. KILT

Az elvi kiindulópontnak tekinthető, korlátozott, ám decentralizált és nem állami üzemeltetésű Sovrin-hálózathoz képest nem csak a „jobban szabályozott", illetve „zártabb", de a „teljesen decentralizált" irányban is számos komoly fejlesztés folyik. Ezek közül mindenképp kiemelésre érdemes a KILT-protokoll és -blokklánc, mely irányításában is teljesen decentralizált, mivel az a blokklánc kriptotoken-gazdaságához kötött. A KILT-blokkláncon kriptoeszközben való fizetéssel lehet DID- és VC regiszter-szolgáltatásokat igénybe venni és kifejezetten támogatja (digitális) eszközökhöz DID-k létrehozását és ezen DID-k más DID-khoz kötését, például tulajdonviszonyok követése céljából.

A KILT-blokklánc a Polkadot „blokkláncok blokkláncá" ökoszisztéma része, így távlatilag annak is egy modelljét adja, hogy az SSI hogyan integrálható egy elosztott főkönyvi ökoszisztémába és annak okosszerződéseibe.

5. SSI és Európa

Az EBSI-hálózat, mint kezdeményezés illeszkedik az Európai Bizottság biztonságos és megbízható európai digitálisidentitás-kezeléssel kapcsolatos törekvéseinek sorába.

Ennek hivatalos keretét a 910/2014/EU eIDAS (a belső piacon történő elektronikus tranzakciókhoz kapcsolódó elektronikus azonosításról és bizalmi szolgáltatásokról) [13] rendelet felülvizsgálata és kiegészítése adja (az „eIDAS 2.0"), amelyhez kapcsolódóan megindult az európai digitális identitástárcák (European Digital Identity Wallet) referencia-architektúráinak [14] és kezdeti implementációinak kialakítása. Az identitástárcák várhatóan előírt képességként kell majd, hogy támogassák a W3C ellenőrizhető tanúsítványok kezelését (az ISO/IEC 18013-5 alapú digitális azonosítók mellett).

Így bár ma még nyílt kérdés, hogy specifikusan az EBSI, mint megoldás szerepe pontosan mi lesz az európai identitáskezelésben, de a nemzeti identitástárcák várhatóan képesek lesznek az EBSI-hálózat használatára is. Továbbá megjegyzendő, hogy az EBSI már ma is rendelkezik saját tárca-megvalósításokkal és tárca-megfelelőségi követelményrendszerrel, és eIDAS-kompatibilis azonosítást tesz lehetővé.

6. Ipari használati esetek

A decentralizált azonosítóknak és ellenőrizhető tanúsítványoknak számtalan magától értetődő alkalmazása adódik az állampolgárok, az állam és a szervezetek különböző azonosítási, tanúsítvány-, és jogosítvány-ellenőrzési viszonylataiban. A meglévő viszonylatokon túl már számos innovatív kezdeményezés is egyre inkább természetesen integrálja az SSI-megoldásokat, a blokklánc alapú élelmiszer-eredetkövetéstől a közeljövő digitális jegybankpénzekig (Central Bank Digital Currencies, CBDC).

Az ipar és a telekommunikáció számára is új lehetőségek nyílnak meg, különösen, ha az okos eszközök is önrendelkezésű identitásokat kapnak. A jelenleg elérhető (nyílt) végponti szoftver-megvalósítások körülbelül a közismert Raspberry Pi eszközök számítási teljesítményét igénylik; néhány folyamatban lévő fejlesztés eredményeként azonban már a közeljövőben várható, hogy az általános célú önrendelkezésű identitások natívan kezelhetővé váljanak jóval gyengébb SoC-megoldásokban is. Ipari alkalmazásként már ma is ismert az SSI bevezethetősége és előnyei

- a fizikai eszközök cégeken átnyúló, kollaboratív állapotmonitorozásban;
- követésükben, karbantartásukban és életciklus-kezelésükben, valamint
- a csalások elkerülésében.

7. Összefoglalás

Az identitásokat és jellemzőik bizonyítását szétcsatoló, önrendelkezésű identitástechnológiák széles körű bevezetésük és elterjedésük küszöbén állnak; cikkünk egy bevezető jellegű áttekintést adott a terület alapmegoldásairól, szabványairól és fejlődéséről.

Az önrendelkezésű identitások alkalmazásai ki fognak terjedni a digitális és online identitáskezelés klaszszikus forgatókönyveire, de egyúttal jóval túl is mutatnak azokon; várható, hogy az SSI-nek kulcsszerepe lesz az Ipar 4.0 és az ipari IoT szervezeti együttműködéseinek és dinamikusan kezelt infrastruktúráinak biztonságossá és megbízhatóvá tételében is.

Hivatkozások

[1] A. Christopher, “The Path to Self-Sovereign Identity”, *http://www.lifewithalacrity.com/2016/04/the-path-to-self-sovereign-identity.html*

[2] D. Vujičić, D. Jagodić and S. Randić, “Blockchain technology, bitcoin and Ethereum: A brief overview,” 17th International Symposium INFOTEH-JAHORINA, East Sarajevo, Bosnia and Herzegovina, 2018, pp.1–6., doi: 10.1109/INFOTEH.2018.8345547

[3] M. Rauchs et al., “Distributed Ledger Technology Systems: A Conceptual Framework,” 2018, SSRN-en: *https://ssrn.com/abstract=3230013,* doi: 10.2139/ssrn.3230013

[4] E. Androulaki et al., “Hyperledger fabric: a distributed operating system for permissioned blockchains,” In Proceedings of the Thirteenth EuroSys Conference (EuroSys ‘18), ACM, 2018, Article 30, pp.1–15. doi: 10.1145/3190508.3190538

[5] ZKProof Community, “ZKProof Community Reference,” *https://docs.zkproof.org/reference.pdf*

[6] World Wide Web Consortium, “Decentralized Identifiers (DIDs) v1.0”, *https://www.w3.org/TR/did-core/*

[7] World Wide Web Consortium, “Verifiable Credentials Data Model v1.1”, *https://www.w3.org/TR/vc-data-model/*

[8] World Wide Web Consortium, “DID Specification Registries – DID Methods”, *https://w3c.github.io/did-spec-registries/#did-methods*

[9] Trust over IP alapítvány, “The Trust over IP model”, *https://trustoverip.org/wp-content/toip-model/*

[10] A Sovrin-alapítvány honlapja, *https://sovrin.org/*

[11] Európai Bizottság, a European Blockchain Services Infrastructure honlapja, *https://ec.europa.eu/digital-building-blocks/wikis/display/EBSI/Home*

[12] Az IDunion európai szövetkezet honlapja, *https://idunion.org/*

[13] Az Európai Parlament és Tanács 910/2014/EU rendelete (2014. július 23.) a belső piacon történő elektronikus tranzakciókhoz kapcsolódó elektronikus azonosításról és bizalmi szolgáltatásokról, *https://eur-lex.europa.eu/legal-content/HU/ALL/?uri=CELEX:32014R0910*

[14] Európai Bizottság, “The European Digital Identity Wallet Architecture and Reference Framework,” *https://digital-strategy.ec.europa.eu/en/library/european-digital-identity-wallet-architecture-and-reference-framework*

[15] F. Marquart, S. Bin. Shams, “Bringing Trustworthiness in Industrial Device Lifecycle using Verifiable Credentials,” *https://hgf22.sched.com/event/14H61*

A szerzőről

KOCSIS IMRE adjunktus, vezető blokklánc-kutató a Budapesti Műszaki és Gazdaságtudományi Egyetem (BME) Méréstechnika és Információs Rendszerek Tanszékén. Műszaki informatikus diplomáját, majd PhD-fokozatát is a BME-n nyerte el. Specializációja a blokklánc-technológiák ipari és szervezetközi alkalmazásai, valamint a blokklánc-alapú megoldások szolgáltatásbiztonságra és teljesítményre tervezése és validációja. A Hyperledger alapítvány felé a BME fő koordinátora, valamint az egyetem EBSI-csomópontjához kapcsolódó tevékenységek vezetője.

Mesterséges intelligencián alapuló eljárások alkalmazása infokommunikációs hálózatokban

FARKAS KÁROLY

BME VIK Hálózati Rendszerek és Szolgáltatások Tanszék / Gloster Networks Kft.
farkas.karoly@vik.bme.hu / farkas.karoly@gloster.hu

Kulcsszavak: mesterséges intelligencia, SDN, NFV, VNF, hálózatmenedzsment, anomáliadetekció, AREP

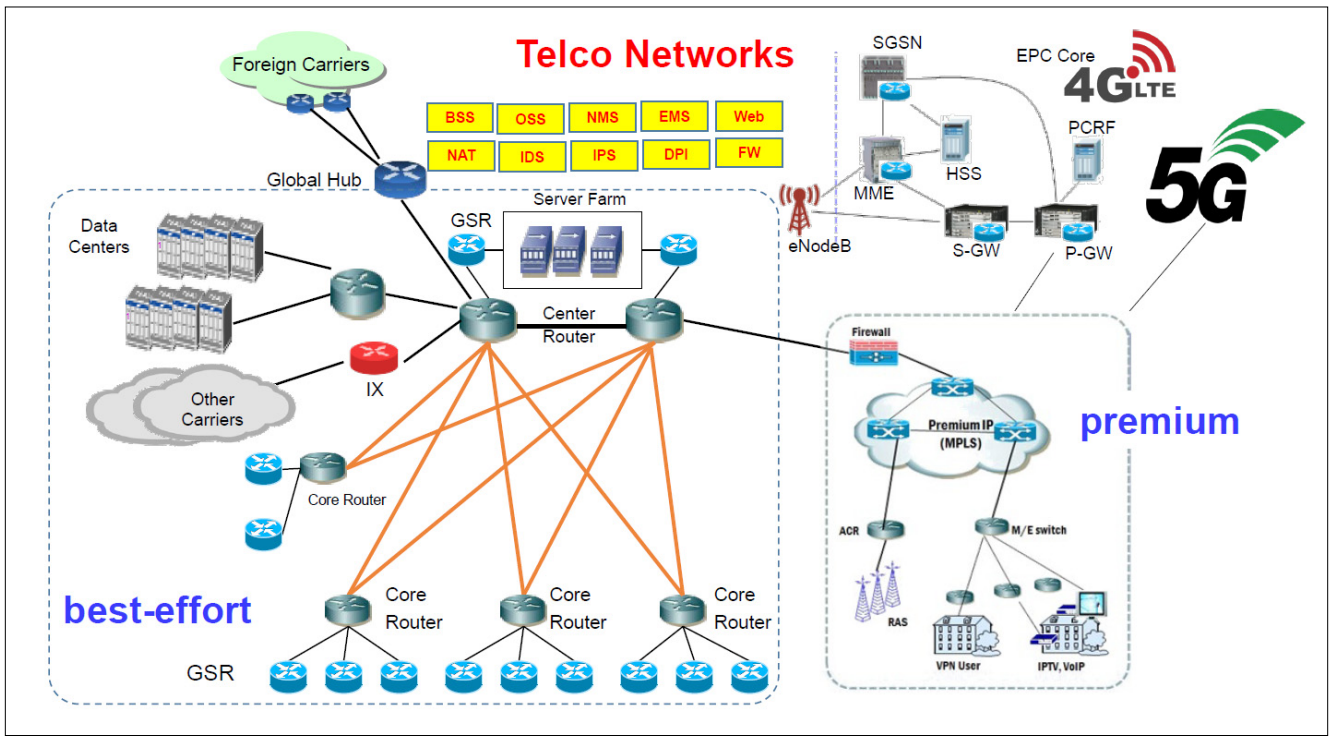
A mesterséges intelligencia robbanásszerű fejlődése szinte minden területre hatást gyakorol, így ez alól az infokommunikációs hálózatok sem jelentenek kivételt. A gépi tanuláson alapuló eljárások jelentős segítséget nyújthatnak az egyre komplexebb feladatok megoldásában, illetve bizonyos funkciók automatizálásában, ezáltal hatékonyabbá téve a hálózatok tervezését, kialakítását, működtetését és felügyeletét. A cikkben röviden áttekintjük az MI-alapú eljárások alkalmazásának főbb lehetőségeit az infokommunikációs hálózatok menedzselése kapcsán, valamint bepillantást nyújtunk a hálózati anomáliadetekció terén elért saját kutatási eredményeinkbe.

1. Bevezetés

Az infokommunikációs hálózatok hagyományosan rengeteg, adott funkcióra dedikált eszközből (pl. útvonalválasztók, kapcsolók, tűzfalak, mobilhálózati bázisállomások, tornyok és antennák) épülnek fel, ahogyan ezt az *1. ábra* szemlélteti. Ez a statikus, rugalmatlan szerkezet nehézkessé és költségessé teszi ezen hálózatok üzemeltetését az operátorok számára. Ebben lényeges változást az egyre nagyobb teret nyerő, szoftveralapon megvalósított virtualizált hálózatok és szolgáltatások használata és elterjedése hozhat. Azonban az ilyen SDN[1]- /NFV[2]- és VNF[3]- [2,3,4] alapú megoldások új kihívások elé is állítják az operátorokat az így kialakított hálózat menedzselésének megnövekvő komplexitása, vagy a gyors és flexibilis átkonfigurálhatóságból adódó hibázási lehetőségek tekintetében.

Ezen kihívások hatékony kezelésére a mesterséges intelligencián (MI) [5] alapuló eljárások robbanásszerű fejlődése szolgálhat. Az MI-kutatások felgyorsulása szin-

1. ábra Infokommunikációs hálózatok hagyományos felépítése [1]



1 *Software Defined Networking – Szoftveralapú hálózatok*
2 *Network Functions Virtualization – Hálózati funkciók virtualizálása*
3 *Virtualized Network Function – Virtualizált hálózati funkció*

te minden területre hatást gyakorol; a képfelismeréstől kezdve a nagy számosságú adathalmazok kezelésén és feldolgozásán keresztül a komplex feladatok automatizálásáig az alkalmazási lehetőségek tárháza szinte végtelen. Nem meglepő, hogy az infokommunikációs hálózatok tervezése, telepítése, működtetése és monitorozása kapcsán is egyre több MI-alapú eljárással találkozhatunk.

E cikkben ismertetjük a mesterséges intelligencia használatának néhány, a tématerület szempontjából releváns szcenárióját. Megnézzük, hogyan tudják támogatni az MI-alapú eljárások többek között a hálózatmenedzsmentet, azon belül is az NFV életciklus-menedzsmentet, a Traffic Engineering[4]-et vagy a hálózati anomáliadetekciót. Ez utóbbi alkalmazási terület kapcsán rövid bepillantást nyújtunk a kutatócsapatunk által elért kutatási eredményekbe is. A továbbiakban először áttekintjük az SDN/NFV- és VNF-alapú hálózati megoldások és szolgáltatások által kínált előnyöket, valamint rendszerezzük a mesterséges intelligenciával kapcsolatban leggyakrabban használt és népszerű fogalmakat. Ezt követően ismertetjük az MI fentebb említett releváns alkalmazási szcenárióit, közülük az egyiket saját kutatási eredményekkel illusztrálva, végül rövid összegzéssel zárjuk a témát.

2. A hálózatokban megjelenő korszerű paradigmák

Az infokommunikációs hálózatokban tradicionálisan alkalmazott, adott funkciókra dedikált eszközökből és eljárásokból álló megközelítést egyre inkább kezdi felváltani a szoftveralapon megvalósított és általános hardvereken futtatható, könnyen migrálható virtualizált funkciókból felépített SDN/NFV-hálózatok és VNF-alapú szolgáltatások használata. Ráadásul a mesterséges intelligencián alapuló eljárások alkalmazása az utóbbi időben robbanásszerűen terjed szinte minden területen, így az infokommunikációs hálózatokban is. Az alábbiakban röviden áttekintjük ezen paradigmák legfontosabb jellemzőit.

2.1. SDN

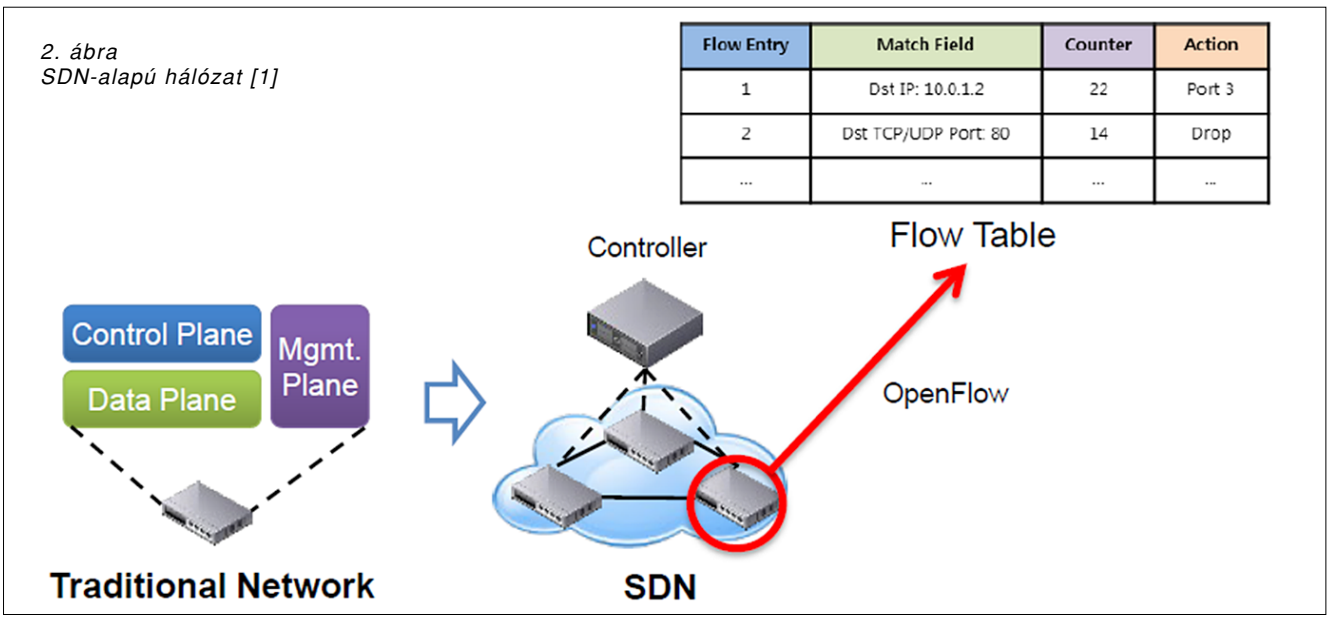
A hálózati eszközök hagyományosan egy monolitikus, zárt egységet alkotnak, amelyet telepítésük és konfigurálásuk után csak egy adott célra – például útvonalválasztásra vagy lokális kommunikációban adatcsomagok gyors továbbítására – lehet használni, alapvető működésüket nem lehet megváltoztatni. Ráadásul minden hálózati eszközön külön kell implementálni az adat-, vezérlő- és menedzsmentsíkot.

Az SDN- [2] alapú megközelítés ezzel szemben az infokommunikációs hálózatok működését szoftveres alapokra helyezi, lehetővé téve a kialakított funkciók megváltoztatását, átprogramozását. Az adat- és a vezérlősík szétválasztásra kerül, így a vezérlési funkciók és a döntési intelligencia kikerül az eszközökből egy központi vezérlőbe, mely eszközök ezáltal egyszerű csomagtovábbító komponensekké válnak *(2. ábra)*.

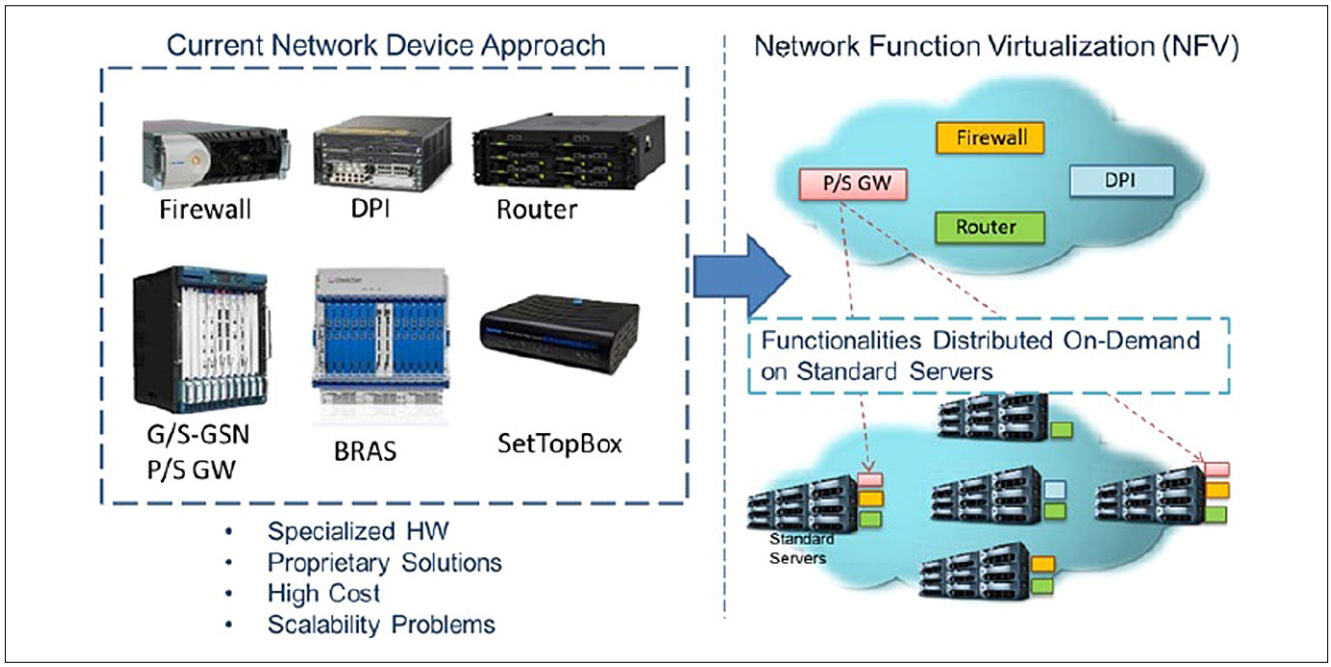
Itt van például az SDN-hálózatokban előszeretettel alkalmazott OpenFlow protokoll [6], amelynél egy központi kontrolleralkalmazás vezérli az útvonalválasztók működését, de a tényleges csomagtovábbítást továbbra is a rendszerint adott funkcióra optimalizált hardverrel ellátott útvonalválasztó végzi. Ezáltal maga az útvonalválasztó eszköz kialakítása egyszerűbb és olcsóbb, a központi vezérlő szoftverben pedig gyorsabban és hatékonyabban lehet az újonnan megjelenő megoldásokat és eljárásokat bevezetni.

Ezen megközelítés előnye még a programozhatóság, gyorsaság, rugalmasság mellett, hogy jóval egyszerűbbé válik gyártófüggetlen hálózatok kialakítása, illetve jelentős megtakarítás érhető el mind a beruházási (CAPEX[5]), mind a működési (OPEX[6]) költségek tekintetében.

2. ábra
SDN-alapú hálózat [1]



4 *Traffic Engineering – Hálózati forgalom elhelyezése, elvezetése*
5 *Capital expenditure – Tőkebefektetés, beruházási költség*
6 *Operational expenditure – Működési költség*



3. ábra NFV-alapú hálózat [1]

2.2. NFV

Eljárások, alkalmazások virtualizált környezetben való futtatása nem újkeletű dolog, az informatika számos területén megjelenik. Ma már könnyen lehet szinte bármilyen operációs rendszert vagy tetszőleges alkalmazást telepíteni rendszerint felhőben futó, virtualizált környezetbe.

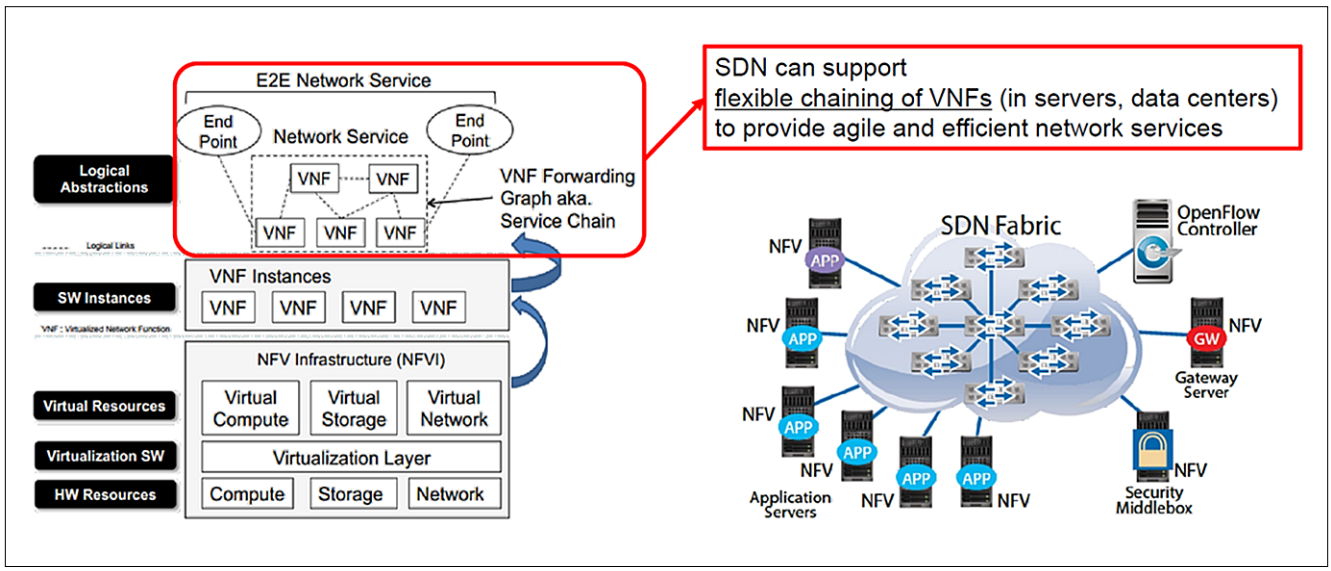
A hálózatok világában sincs ez másként a hálózati funkciók virtualizációjának köszönhetően. Az NFV [3] egy, az SDN-től független, de azt jól kiegészítő hálózati paradigma, amely leválasztja a szoftvert a hardverről, így biztosítva a rugalmas hálózattelepítést és a dinamikus működést *(3. ábra)*. Ebben a megközelítésben a hálózati csomópont-funkciók IT virtualizációs technológiák használata segítségével virtualizált alapvető összetevőkből épülnek fel. A virtualizált funkciók nem igényelnek dedikált vagy célorientált hardvert, ezért rendszerint általános jellegű, de nagy teljesítményű szervereken futó virtuális gépekre vagy konténerekbe telepítjük őket.

Ezen megközelítés legfőbb előnyei: a virtualizált funkciók szinte tetszőleges elhelyezéséből, vagy esetleges migrálásából adódó rugalmasság; új funkciók vagy szolgáltatások gyors bevezetésének lehetősége; nagyfokú skálázhatóság a felmerülő igények függvényében történő erőforrás-allokációnak köszönhetően; illetve a CAPEX/OPEX-költségekben jelentkező megtakarítás.

2.3. VNF

NFV-környezetben futó virtuális hálózati funkciók segítségével, ezen VNF-ek [4] összefűzésével vagy láncolásával könnyen kialakíthatók szoftveralapú, rugalmas, dinamikusan átkonfigurálható, végpontok számára nyújtott átfogó hálózati szolgáltatások *(4. ábra)*.

4. ábra VNF koncepció [1]

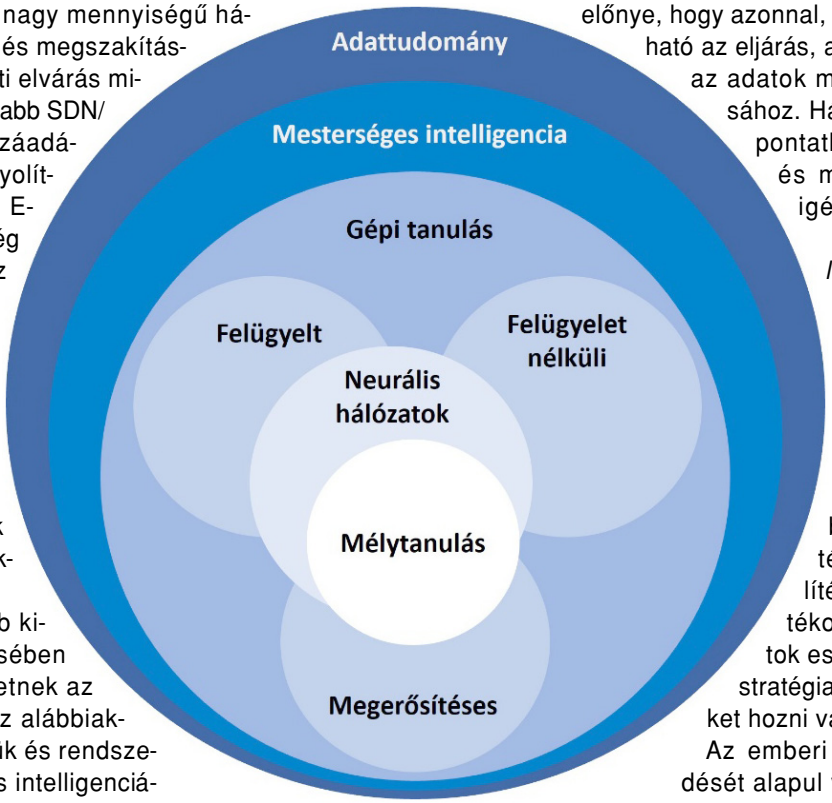


A VNF-ek megvalósíthatók önálló entitásként rendszerint meglévő hálózati paradigmák használatával. Azonban az SDN-alapú megoldások számos előnyt nyújtanak az NFV-környezetek kialakításához, különösen akkor, ha nagy kiterjedésű, földrajzilag különböző elhelyezkedésű NFV-infrastruktúrát kell kezelni és összehangolni. Ezért nem meglepő, hogy számos olyan szolgáltatói platform létezik, amelyik egységes ökoszisztémába foglalja az SDN- és NFV-alapú megoldásokat.

A VNF-ek segítségével létrehozott hálózati szolgáltatások legfőbb előnyei a gyors és rugalmas kialakítás mellett a központosított menedzsment; optimális konfiguráció kialakítása a hálózat aktuális állapotának függvényében; gyors reagálás a hálózati környezet változásaira; valamint gyártófüggetlen infrastruktúra kialakításának lehetősége.

Azonban az előnyök mellett sajnos számos limitációval is számolni kell. Így a hálózatmenedzsment egyre komplexebbé válik a nagy mennyiségű hálózati forgalom gyors és megszakításmentes kezelése iránti elvárás miatt, valamint minden újabb SDN/NFV-komponens hozzáadása a rendszerhez bonyolítja a hálózatkezelést. Ezen túl a sebezhetőség is növekszik, mivel az SDN/NFV-megoldások ugyan lehetővé teszik az operátorok számára a hálózati környezet gyors és rugalmas konfigurálását, de az esetleges hibás konfigurációk hálózati fennakadásokhoz vezethetnek.

Ezeknek és egyéb kihívásoknak a kezelésében is segítségünkre lehetnek az MI-alapú eljárások. Az alábbiakban röviden áttekintjük és rendszerezzük a mesterséges intelligenciával kapcsolatos fogalmakat.



5. ábra
Mesterséges intelligencia és a gyakoribb, kapcsolódó fogalmak viszonya

2.4. Mesterséges intelligencia

Az MI [5] mesterségesen létrehozott tudat általi intelligenciát jelent, amely képes önállóan viselkedni, tanulni, döntéseket hozni. Az MI, mint tématerület, a számítástechnika, számítógép-tudomány egyik ágát képviseli, de manapság már szinte minden területen találkozhatunk vele. Sokszor rokon értelemben használjuk a gépi tanulás kifejezéssel, bár annál szélesebb tartományt fed le.

A hétköznapokban egyre sűrűbben találkozunk az MI-vel, illetve a gépi tanulással kapcsolatos fogalmakkal. A gyakoribb fogalmakat és ezek egymáshoz való viszonyát az *5. ábra* szemlélteti[7].

Adatok kezelését, feldolgozását, belőlük hasznos információ kinyerését rendszerint hagyományos matematikai módszerek – például statisztikai analízis –, vagy manapság egyre inkább mesterséges intelligencián alapuló eljárások segítségével végezzük. A gépi tanulás az MI egyik, ha nem a legnagyobb válfaja. Ezen belül megkülönböztetünk felügyelt, felügyelet nélküli, illetve megerősítéses tanulást alkalmazó módszereket.

Felügyelt tanulás esetén az eljárásunkat először egy tanuló adathalmazon tanítjuk, majd az így létrehozott modellt alkalmazzuk problémák megoldására. Ennek előnye, hogy állandó mintázattal rendelkező adatokon nagy pontossággal és gyorsan működő modellt lehet előállítani. Hátránya a tanítás jelentős idő- és erőforrásigénye, valamint, ha változik az adatok mintázata, akkor újra kell tanítani a modellt.

Felügyelet nélküli tanulás esetén elmarad a különálló tanulási fázis, az eljárás működés közben tanul. Ennek előnye, hogy azonnal, valós időben használható az eljárás, ami képes adaptálódni az adatok mintázatának változásához. Hátránya, hogy sokszor pontatlanabb eredményt ad és működése erőforrásigényes.

Megerősítéses tanulás esetén az eljárás döntések sorozatán keresztül próbál egy stratégiát megtanulni, amelyre a folyamat végén kap csak visszajelzést, és a visszajelzések tükrében változtatja a stratégiát. Ez a megközelítés jól alkalmazható játékok, vagy olyan feladatok esetén, ahol valamilyen stratégia mentén kell döntéseket hozni vagy megoldást találni.

Az emberi idegrendszer működését alapul vevő *neurális hálózatok* használata az egyik olyan eljárás, amely mindhárom tanulási módszer esetében alkalmazható és amely egyre nagyobb népszerűségnek örvend manapság. Az ilyen hálózatok több rétegbe szervezett, egyszerű műveleteket végző neuronokból állnak, melyek egymáshoz kapcsolódva és együttműködve hozzák létre a kívánt eredményt. Több változatuk létezik az egyszerűbb előrecsatolt hálózatoktól kezdve a visszacsatolt hálózatokon át a komplexebb, memóriával is rendelkező hálózatokig. Ide lehet sorolni a *mélytanulást* alkalmazó módszereket is – bár sokan ezt bővebb kategóriának tekintik –, ahol a *mély* jelző arra utal, hogy a tanuló eljárás architektúrája több rétegből áll.

7 Az egyes kategóriák határait illetően nincs egységesen elfogadott álláspont, ezért az irodalomban találkozhatunk az ábrán feltüntetettől eltérő kategorizálással is.

A mesterséges intelligencián alapuló eljárások számos feladat megoldásában nyújthatnak segítséget a hálózatok területén is, amit az alábbiakban néhány példán keresztül szemléltetünk.

3. Mesterséges intelligencia alkalmazási szcenáriók hálózatokban

Az MI-alapú eljárások széleskörűen alkalmazhatók az infokommunikációs hálózatokban is. Az egyik kézenfekvő alkalmazási terület a hálózatok menedzselésének támogatása, az egyre komplexebbé váló menedzsment feladatok elvégzésének elősegítése. Például MI-alapú módszerek segítségével automatizált módon azonosíthatók a hálózati események közötti korrelációk; megjósolható a hálózat jövőbeli viselkedése; vagy akár önmenedzselő hálózatok hozhatók létre, amelyek képesek saját magukat szervezni, optimalizálni, hibából helyreállni az üzemeltető beavatkozása nélkül.

MI használatával támogatható konkrét menedzsment-feladatok többek között az NFV életciklus-menedzsment; Traffic Engineering; hálózatianomália-detekció; gyökérhiba analízise; meghibásodás előrejelzése; hálózati erőforrások kihasználtságának előrejelzése; a hálózatot ért támadás, illetve behatolás detektálása. Az alábbiakban ezen feladatok közül néhány kapcsán röviden áttekintjük, hogyan is lehetnek az MI-alapú eljárások a segítségünkre.

3.1. NFV életciklus-menedzsmentje

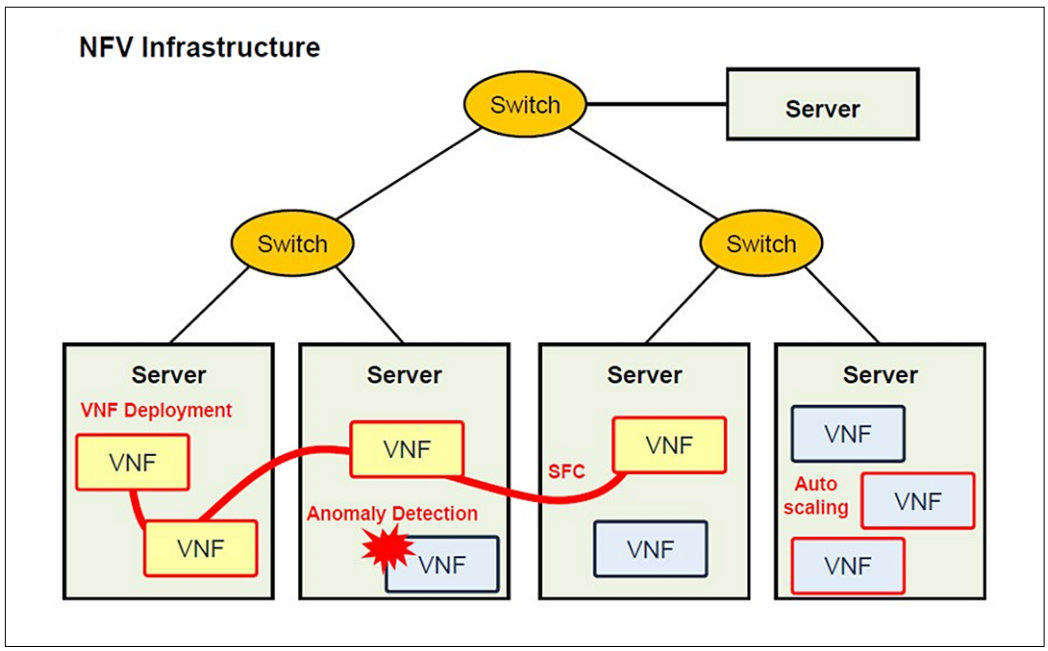
NFV-infrastruktúrán kialakított hálózatokban az NFV életciklus-menedzsmentje magába foglalja a VNF-ek létesítését, telepítését; a szolgáltatásfunkciók láncolását; az automatikus skálázást; illetve az anomáliadetekciót *(6. ábra)*.

A működési költségek optimalizálása céljából a VNF-ek fizikai szervereken való létesítését, telepítését érdemes a forgalom változásához igazítani, így az adott időszakra az adott VNF-eket a forgalmi igényhelyhez közeli, vagy kevésbé terhelt szervereken létrehozni. A VNF-ek dinamikus létrehozásában, áthelyezésében nyújthatnak támogatást az MI-alapú eljárások. Például a hálózati forgalom alakulásának megtanulására előrecsatolt, míg a hálózati topológia megtanulására gráfalapú neurális hálózatok alkalmazhatók, melyeket egy megfelelő modellben kombinálva előre megjósolható, hova érdemes az adott VNF-eket telepíteni vagy áthelyezni [7].

Összetettebb hálózati szolgáltatásokat az alapvető szolgáltatásfunkciókat biztosító VNF-ek egymás utáni láncolásával lehet biztosítani. Azonban azt meghatározni, hogy az adott funkciókat megvalósító, különböző szervereken futó VNF-ek kiválasztásának és láncolásának mi az optimális módja, közel sem triviális feladat. Ebben lehet segítségünkre például az MI-alapú, megerősítéses tanulást alkalmazó Q-learning[8] nevezetű eljárás [8].

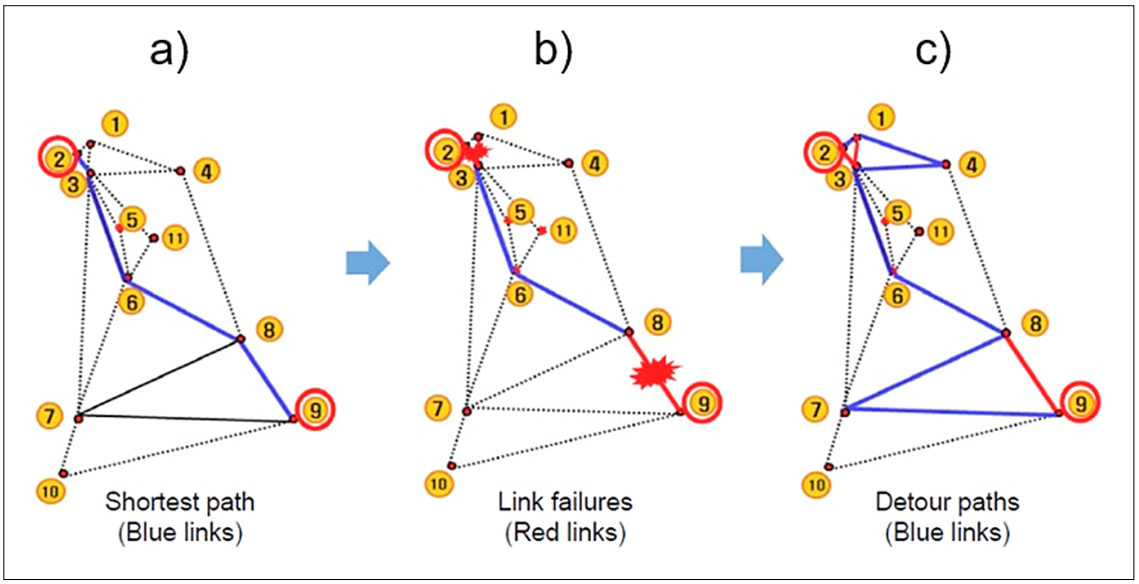
A szolgáltatások megfelelő skálázása szintén olyan feladat, amiben segítségül tudjuk hívni az MI-t az erőforrások optimális kihasználása céljából. Így például adott funkciót ellátó VNF-példányok számának a felhasználói igények függvényében történő automatikus növelésére vagy csökkentésére hatékonyan alkalmazható az ugyancsak megerősítéses tanulást használó, deep Q-network[9] nevezetű eljáráson alapuló automata skálázó eszköz [9].

Az anomáliák – normálistól eltérő viselkedés – detektálása nem csak a hagyományos hálózatokban, hanem az NFV-környezetben is fontos feladat, hiszen ezek sokszor hibás működésre utaló jelek. Például az adott VNF-példányok abnormális állapotának detektálásával és a szükséges beavatkozással megelőzhetők vagy gyorsan orvosolhatók a szolgáltatás nyújtására vonatkozó követelmények megsértéséből adódó problémák. Erre al-



6. ábra NFV életciklus-menedzsmentje [1]

8 Q-learning – Q-tanulás
9 Deep Q-network – Mély Q-hálózat



7. ábra Kerülő útvonal meghatározásra link meghibásodás esetén [1]

kalmazható többek között a felügyelt tanuláson alapuló XGBoost-eljárás, amely a monitorozott modul normális működését megtanulva jelzi, ha ettől eltérő viselkedést tapasztal [10].

3.2. Traffic Engineering

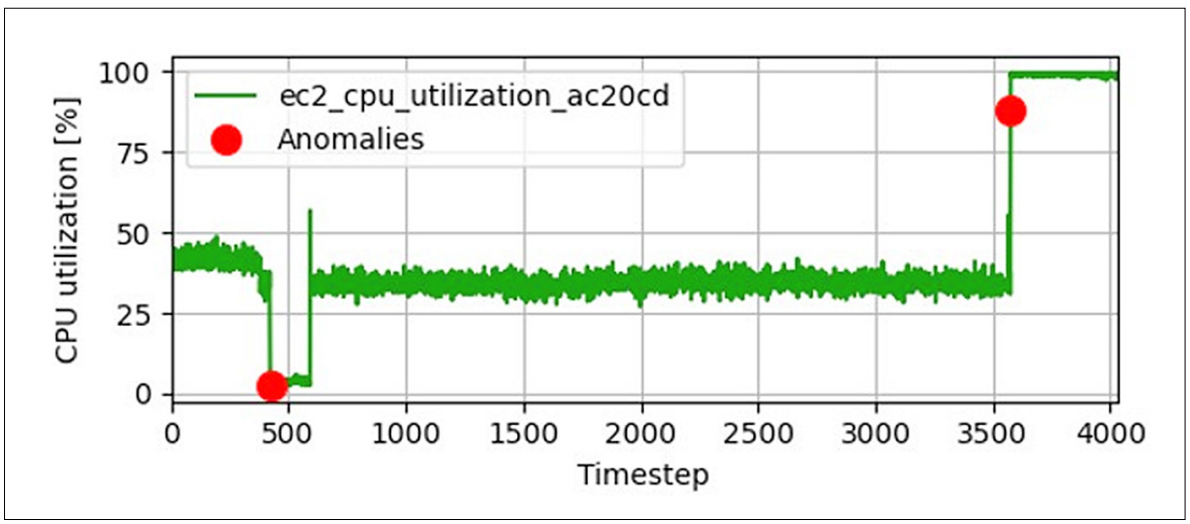
A hálózattervezés egyik kiemelt tématerülete olyan redundáns erőforrásokkal rendelkező hálózatok tervezése, amelyek hibatűrők, így meghibásodás esetén is képesek ellátni a feladatukat. Az ilyen redundáns hálózatokon történő forgalomelhelyezést/elvezetést hívjuk Traffic Engineering-nek, amely egyrészt a terhelés elosztásával biztosítja az erőforrások hatékonyabb kihasználását, másrészt az elvárt működést még meghibásodás esetén is.

Példaként tekintsük a *7. ábrán* látható hálózatot. Normál működés esetén a 2 és 9 csomópont közötti legrövidebb útvonalon, azaz a 2–3–6–8–9 csomópontok érintésével történik a forgalom elvezetése *(7.a ábra)*. Tegyük fel, hogy meghibásodnak az 1–3, 2–3 és a 8–9 csomópontok közötti linkek *(7.b ábra)*, ezért a szolgáltatási követelményeket kielégítő optimális kerülő útvonalat kell meghatároznunk a kiinduló- és célcsomópont között. Ez a kerülő útvonal esetünkben a 2–1–4–3–6–8–7–9 csomópontokat érinti *(7.c ábra)*.

Egyszerűbb hálózati topológia esetén nem tűnik túl bonyolult feladatnak egy ilyen kerülő útvonal meghatározása, de nagyobb, komplex topológiájú hálózatoknál bizony ez már egy nehéz, összetett feladat. Ebben lehet segítségünkre a fentebb már említett, megerősítéses tanuláson alapuló Q-learning eljárás, amely meghibásodás esetén is képes automatikusan megtalálni a szolgáltatási követelményeket kielégítő legrövidebb útvonalat.

3.3. Hálózati anomáliadetekció

Az infokommunikációs hálózatok üzemeltetése során az egyik kiemelt feladat a hálózati komponensek jellemzőinek, teljesítményének folyamatos monitorozása. Így az esetleges anomáliákat azonnal detektálni lehet, jobb esetben előre lehet jelezni. Ezen a területen a kutatócsoportunk is aktív kutatómunkát végez, melynek keretében olyan mesterséges intelligencián alapuló anomália-detektorok fejlesztésével foglalkozunk, amelyek képesek valós időben felismerni és jelezni az esetleges anomáliákat.



8. ábra Anomáliák hálózati eszközről gyűjtött telemetria adatokban [12]

A *8. ábrán (lásd az előző oldalon)* például egy hálózati eszközről gyűjtött, az eszköz processzorának kihasználtságát megjelenítő telemetriaadatok láthatók az idő függvényében. Ezek az adatok többnyire egy konstans érték körül ingadoznak. Viszont az egyik időszakban megfigyelhető egy kiugróan alacsony, egy másik időszakban pedig egy kiugróan magas érték, amelyek a normálistól eltérő működésre, így esetleges meghibásodásra, vagy akár a hálózatunkat ért támadásra is utalhatnak. Amennyiben – lehetőség szerint valós időben – detektálni tudjuk ezeket az anomáliákat, úgy szükség esetén gyorsan be is tudunk avatkozni, biztosítva ezáltal a normális működést.

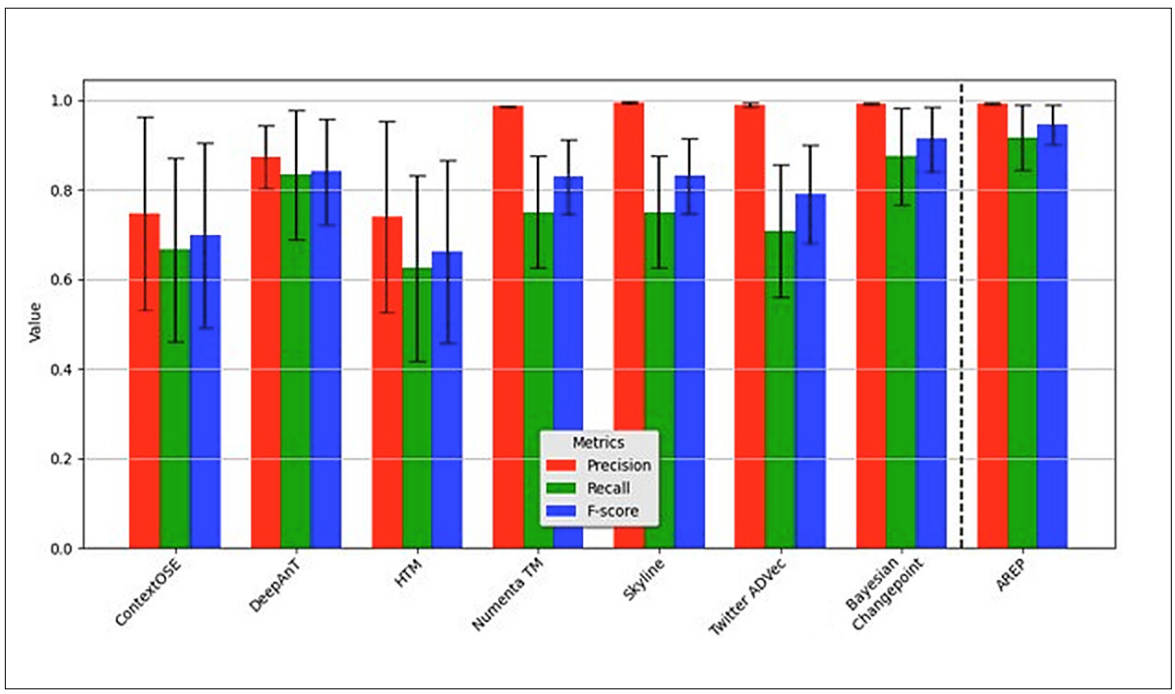
Kutatócsoportunknak sikerült egy LSTM[10]-alapú, memóriával rendelkező neurális hálózatot használó, AlterRe[2] névre keresztelt eljárást kidolgoznia [11], majd ezt továbbfejlesztenie az AREP nevezetű eljárássá [12]. Ezek az anomáliadetektorok a felügyelet nélkül tanuló algoritmusok csoportjába sorolhatók, így valós időben alkalmazhatók az anomáliák jelzésére és képesek adaptálódni a monitorozott adatok mintázatának változásaihoz.

Az AREP teljesítményét – egyéb vizsgálatok mellett – összevetettük néhány korszerű, az utóbbi időben publikált anomáliadetektorral is, ennek az eredménye látható a *9. ábrán*. Három különböző, ezen a területen általánosan használt metrika (Precision, Recall, F-score) mentén végeztük el az összehasonlítást. Minden esetben igaz, hogy minél nagyobb átlagos értéket kapunk, annál jobban teljesít az adott eljárás az adott metrika vonatkozásában. Az ábrán feltüntetésre került a kapott értékek szórása is. Az eredményekből kiolvasható, hogy az AREP az egyik metrika mentén (Precision) legalább olyan jól teljesít, mint a legjobban szereplő versenytársai, míg a másik két metrika esetén (Recall, F-score) túl is teljesíti azokat.

4. Összefoglalás

Mesterséges intelligencián alapuló eljárások alkalmazása manapság már szinte minden területen tetten érhető, így az infokommunikációs hálózatok esetében is. Ami nem meglepő, hiszen az egyre komplexebbé váló hálózati környezetben a bonyolult, összetett feladatok elvégzését nagy mértékben elősegíthetik az öntanuló, nagy mennyiségű adatokat gyorsan feldolgozni képes eljárások.

A cikkben röviden áttekintettük az utóbbi időszakban az infokommunikációs hálózatok terén egyre jobban terjedő korszerű paradigmákat, így az SDN/NFV/VNF-alapú hálózati megoldásokat és a mesterséges intelligencia kapcsán gyakran használt alapvető fogalmakat. Ezután felvázoltuk, hogy a hálózatmenedzsmenthez kapcsolódóan hogyan tudják segíteni az MI-alapú eljárások – többek között – az NFV életciklus-menedzsmentjét vagy a Traffic Engineering-et, valamint bepillantást nyújtottunk az MI által támogatott hálózati anomáliadetekció terén elért saját kutatási eredményeinkbe.



9. ábra AREP összevetése egyéb korszerű anomáliadetektor eljárásokkal [12]

10 *Long Short-Term Memory – Hosszú rövidtávú memória*

Hivatkozások

[1] Hong, J.W., “Artificial Intelligence-based Network Management,” Webinar, Nov. 2021., *http://dpnm.postech.ac.kr*

[2] Benzekki, K., El Fergougui, A., and Elbelrhiti Elalaoui, A., “Software-defined networking (SDN): A survey,” Security and Communication Networks, Vol. 9, no.18, pp.5803–5833, 2016. DOI: 10.1002/sec.1737

[3] ETSI, “Network Functions Virtualization – Introductory White Paper”, Oct. 2012. *https://docbox.etsi.org/isg/nfv/open/Publications_pdf/White%20Papers/NFV_White_Paper1_2012.pdf*

[4] Stalling, W., “Foundations of Modern Networking: SDN, NFV, QoE, IoT, and Cloud,” Pearson Education, 2016.

[5] Russell, S.J., and Norvig, P., “Artificial Intelligence: A Modern Approach (2nd ed.),” Prentice Hall, 2003., ISBN 0-13-790395-2

[6] McKeown, N. et al., “OpenFlow: Enabling innovation in campus networks,” ACM SIGCOMM Computer Communication Review, Vol. 38, no.2, pp.69–74, Apr. 2008. DOI: 10.1145/1355734.1355746

[7] Kim, H. et al., “Graph Neural Network-based Virtual Network Function Deployment Prediction,” 16th International Conference on Network and Service Management (CNSM), Nov. 2020.

[8] Lee, D. et al., “Q-learning Based Service Function Chaining Using VNF Resource-aware Reward Model,” The 21st Asia-Pacific Network Operations and Management Symp. (APNOMS 2020), Daegu, Korea, Sep. 2020.

[9] Lee, D. et al., “Deep Q-network-based auto scaling for service in a multi-access edge computing environment,” International Journal of Network Management (IJNM), Vol. 31, no.6, Nov. 2021. DOI: doi.org/10.1002/nem.2176

[10] Hong, J. et al., “Machine Learning based SLA-Aware VNF Anomaly Detection for Virtual Network Management,” 16th International Conference on Network and Service Management (CNSM 2020), Virtual Conference, Nov. 2020.

[11] Vajda, D., Pekar, A. and Farkas, K., “Towards machine learning-based anomaly detection on time-series data,” Infocommunications Journal, Vol. 13, no.1, pp.36–44, March 2021. DOI: https://doi.org/10.36244/ICJ.2021.1.5

[12] Farkas, K., “AREP: An adaptive, machine learning-based algorithm for real-time anomaly detection on network telemetry data,” Springer Neural Computing & Applications Journal, Vol. 35, no.8, pp.6079–6094, March 2023. DOI: 10.1007/s00521-022-08000-y

Köszönetnyilvánítás

A cikkben feldolgozott téma és a bemutatott eredmények a Nemzeti Kutatási Fejlesztési és Innovációs Alap Mesterséges Intelligencia Nemzeti Laboratórium program, valamint a Pro Progressió Alapítvány támogatásával, a BME VIK Hálózati Rendszerek és Szolgáltatások Tanszék és a Gloster Networks Kft. együttműködésében valósultak meg.

A szerzőről

FARKAS KÁROLY 1998-ban szerzett műszaki informatikus mérnök, majd 1999-ben bankinformatikus szakmérnök diplomát a BME Villamosmérnöki és Informatikai Karán. Doktori tanulmányait a BME-n, majd a Zürichi Műszaki Egyetemen (ETH Zürich) folytatta, ahol 2007-ben megszerezte a PhD-fokozatot. 2007–2012 között a Nyugat-magyarországi Egyetem (NymE) Informatikai és Gazdasági Intézetének, 2008-tól a BME Hálózati Rendszerek és Szolgáltatások Tanszékének docense, ahol 2017-ben habilitált. 2016 és 2022 között a NETvisor Zrt. K+F igazgatója volt. 2022 júliusától a Gloster Infokommunikációs Nyrt., majd 2023 januárjától a Gloster Networks Kft. oktatási vezetője. A 2011/2012-es tanévet az Universitaet Zürich-en töltötte vendégprofesszorként. Farkas Károly oktatási és kutatási tevékenységét elsősorban a kommunikációs hálózatok területén végzi autonóm, önszerveződő vezeték nélküli mobil hálózatok, IoT, Ipar 4.0 és mesterséges intelligencia témakörökben. Több mint 100 tudományos publikációval rendelkezik és rendszeresen szerepel reguláris vagy meghívott előadóként, valamint szervezőként különböző hazai és nemzetközi eseményeken, konferenciákon, tanfolyamokon. 2016-ban az MTA elismerő oklevélben részesítette az MTA Bolyai János Kutatási Ösztöndíj keretében végzett kiemelkedő kutatói munkájáért. Koordinátora a BME-n működő lokális Cisco Hálózati Akadémiának, a Cisco Oktató Laboratóriumnak, valamint kezdeményezője és főszervezője a BME–Pannon–HTE NetSkills Challenge országos tanulmányi versenynek. A Cisco Hálózati Akadémia képesített instruktora, ill. instruktor trénere, rendelkezik számos Cisco ipari vizsgával, 2017-ben Cisco Instructor Excellence Advanced Award díjban részesült. Farkas Károly tevékenységét a Hírközlési és Informatikai Tudományos Egyesület 2015-ben Arany Jelvénnyel és Pollák-Virág díjjal ismerte el. 2020-tól a HTE tudományos területért felelős elnökségi tagja.

Szintetikus adatgenerálás – az adathozzáférés szent grálja?

DÉVAI GERGELY, RECSE ÁKOS
Ericsson Magyarország, ELTE
{gergely.devai; akos.recse}@ericsson.com

Kulcsszavak: mesterséges intelligencia, gépi tanulás, adathozzáférés, szintetikus adat

A mesterséges intelligenciát, gépi tanulást alkalmazó projektek számára szükséges tanítóadatok sokszor nem hozzáférhetőek az adattudósok számára. Különösen fontos probléma ez személyes vagy üzletileg érzékeny adatok esetén. Ilyen jellegű problémákra jelenthet megoldást a szintetikus adatgenerálás. Ez a technológia automatizált módon tanít generatív modelleket az eredeti adatokon, majd a modell segítségével az eredeti adatok statisztikai jellemzőit és egyéb mintázatait hordozó, szintetikus adathalmazt generál. Mindezt úgy, hogy az eredeti adat egyes rekordjairól a generált adat nem „szivárogtat ki" információt. Elemzésünkben áttekintjük a szintetikus adatgenerálás felhasználási területeit, valamint a módszer elméleti alapjait. Bemutatjuk négy kereskedelmi és hat szabad szoftvereszközzel kapcsolatos tapasztalatainkat és mérési eredményeinket.

1. Bevezetés

A mesterséges intelligencia és gépi tanulás térhódításával olyan eszköz került a vállalatok kezébe, aminek segítségével könnyebben megérthetik, felmérhetik és megjósolhatják a felhasználók igényeit. Az ehhez szükséges adatok sokszor azonban nem vagy csak részben hozzáférhetőek, például, ha egészségügyi vagy személyes adatokról van szó, melyet az Általános Adatvédelmi Rendelet (GDPR) miatt nem lehet az MI-alkalmazást fejlesztő szakemberek, cégek számára elérhetővé tenni. Ilyen jellegű problémákra jelenthet megoldást a szintetikus adatgenerálás.

Kétféle megközelítés terjedt el új adatok előállítására. Egyrészt készíthetünk szimulátorokat, melyekkel megpróbáljuk minél pontosabban leírni az adott rendszer viselkedését, hogy aztán új adatokat tudjuk generálni velük. Ebben az esetben a valós folyamatok modelljét egy szakértő állítja elő, programozza le. Másrészt az eredeti, érzékeny adatokkal taníthatunk generatív modelleket, amelyek automatikusan felismerik a bemenő adatok statisztikai tulajdonságait és jellemző mintázatait, majd ez alapján képesek olyan új adatpontokat generálni, amik minél több tulajdonságban hasonlítanak az eredeti adathoz. Cél azonban, hogy a generált adathalmaz alapján ne lehessen az eredeti adat egyes rekordjaira vonatkozó megállapításokat tenni. A két lehetőség közül ezúttal az utóbbival foglalkozunk.

Cikkünkben azt vizsgáljuk, hogy a jelenleg elérhető szintetikus adatgenerátorok képesek-e olyan minőségű új adatot előállítani, hogy a tisztán ilyen adatokon tanított gépi tanulási algoritmusok pontossága nem marad el lényegesen az eredeti adatokon tanított algoritmusok teljesítményétől, továbbá mérésekkel mutatjuk be, hogy a szintetikus adatok mennyire megkülönböztethetőek vagy éppen hasonlóak az eredeti adathoz.

A 2. szakaszban a szintetikus adatok felhasználási területeivel, a 3. szakaszban a módszereivel és eszköztárával foglalkozunk. A 4. szakasz vizsgálataink módszertanát, az 5. az eredményeit ismerteti. Az összefoglalás a 6. szakaszban található.

2. Használati esetek

Ahogy azt a bevezetőben is érintettük, jogi, üzleti vagy adatvédelmi megfontolásból magas szintű adatvédelmi szabályokat és folyamatokat alkalmazhatnak a vállalatok, melyek sok esetben nem teszik lehetővé a felhasználói adatok széleskörű megosztását. Szintetikus adatgenerátor használatával azonban úgy juthatnak értékes adathoz a fejlesztők, hogy nem kell a valós felhasználói adatokkal dolgozni. Természetesen ez a módszer is körültekintő használatot igényel, hiszen egy túltanított adatgeneráló algoritmus előállíthat olyan adatot, ami megfeleltethető az eredeti adathalmaz valamely rekordjának, vagy akár azonos azzal. Ennek elkerülése érdekében a generált adat „hasznosságán" túl annak „biztonságosságát" is vizsgálni kell.

A fent említett eseten túlmutatóan egyéb környezetben is használhatunk szintetikus adatgeneráló algoritmusokat. Az egyik ilyen lehetőség az adathalmaz-bővítés: amennyiben kevés az elérhető adat egy pontos gépi tanulási algoritmus tanításához, lehetőség van olyan új adatpontokat generálni, melyekkel kiegészíthetjük az eredeti adathalmazt. Ez a módszer különösen akkor ígéretes, ha a generált adatokhoz egy szakértő plusz tudást ad hozzá, például szűrőfeltételek megadása segítségével.

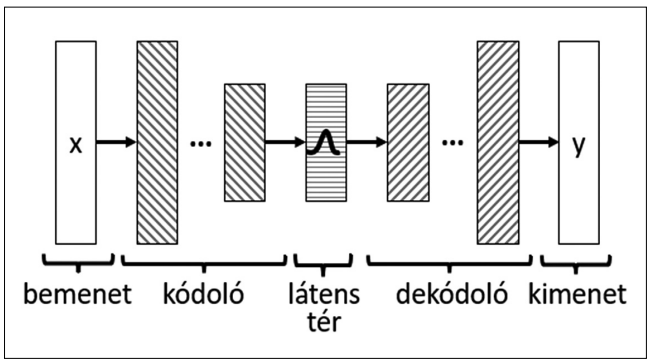
Példa erre a képekkel dolgozó modellek tanítóhalmazának bővítése egyszerű műveletekkel (forgatás, kivágás, skálázás stb.), melyek a képekhez rendelt címkét helybenhagyják [1]. Itt a szakértői tudás a megfelelő transz-

formációk kiválasztásában jelenik meg. Egy másik példa a ritka eseményekkel foglalkozó modellek (pl. anomália-felismerés) esete. Mivel ezekből a valós adathalmazban kevés szerepel, a tanítás hatékonysága érdekében a múltbeli ritka vagy szélsőséges eseményekből kiindulva új, hasonló adatpontokat kell generálni.

3. Módszerek és eszközök

Adatgenerálásra számos modellarchitektúra használható. A teljesség igénye nélkül ezek közül mutat be néhányat ez a szakasz. Kitérünk továbbá az általunk vizsgált kereskedelmi és szabad szoftvereszközökre.

A VAE (Variational Autoencoder) egy neuronháló-architektúra *(1. ábra)*, amelyben a bemenő vektorokat (x) a kódoló egy látens térbe transzformálja. Ez utóbbi a VAE esetén eloszlások kombinációja. A dekódoló a látens tér pontjait képezi le a bemenettel megegyező alakú kimenetté (y). A komponenseket úgy tanítják, hogy a be- és kimenet eltérése minimális legyen.



1. ábra A VAE neuronháló-architektúra

A VAE úgy használható adatgenerálásra, hogy a látens tér eloszlásaiból mintavételezünk, majd ezeket a dekódoló adatvektorokká alakítja.

A kopulák olyan matematikai objektumok, amelyeket többváltozós eloszlások modellezésére használhatunk. A többváltozós eloszlást két komponensre, az egyes változók marginális eloszlására és az egyes változók közötti összefüggésre bontjuk. A kopula feladata ez utóbbi modellezése. Adatgenerálás az érintett eloszlásokból történő mintavételezéssel történik.

A GAN- (Generative Adversarial Network) architektúra *(2. ábra)* esetén a Generátor neuronháló véletlen zajból, mint stimulusból állít elő adatvektorokat. A Diszkriminátor egy másik neuronháló, ami vagy egy valódi, vagy egy generált adatvektort kap inputként, és a feladata azt eldönteni, hogy milyen inputot kapott. A Generátor és Diszkriminátor modellek tehát „egymás ellen játszanak”: minél valószerűbb a generátor kimenete, annál nehezebb a Diszkriminátor dolga. Szintetikus adat előállítására a Generátor használható a tanítás befejeztével.

A GAN-architektúrának számos változata van, például feltételes és adatbázis-generáláshoz használt (cGAN, CTGAN) [6], nembináris diszkriminátorral rendelkező (WGAN) [14], vagy differenciális adatvédelemhez használt (DPGAN) [13] stb.

A fenti módszereknek számos kereskedelmi és szabad szoftveres implementációja létezik. Ezek közül kutatásunk során a következőket vizsgáltuk:

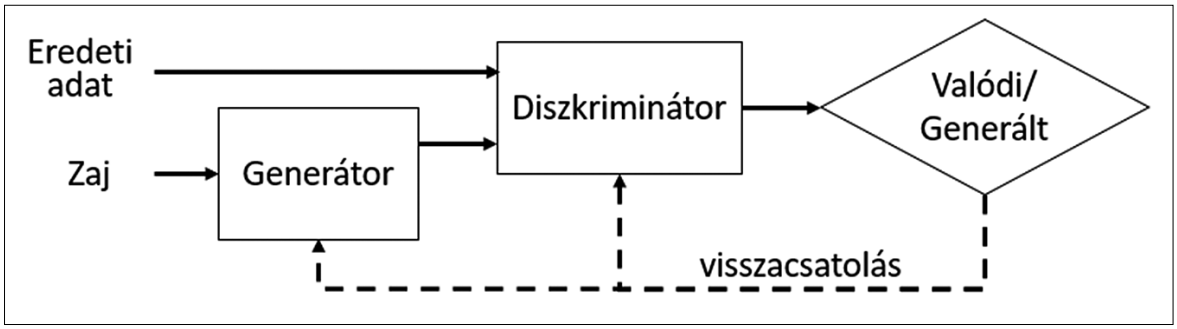
- DeltaPy (szabad, nyílt forráskódú) [10]
- Gretel (kereskedelmi) [7]
- Gretel nyílt forráskódú változata [8]
- Mostly.AI (kereskedelmi) [2]
- SDV (szabad, nyílt forráskódú) [5]
- Statice (kereskedelmi) [4]
- Synthpop (szabad, nyílt forráskódú) [11,12]
- Syndata (szabad, nyílt forráskódú) [15]
- YData (kereskedelmi) [3]
- YData nyílt forráskódú változata [9]

4. Módszertan

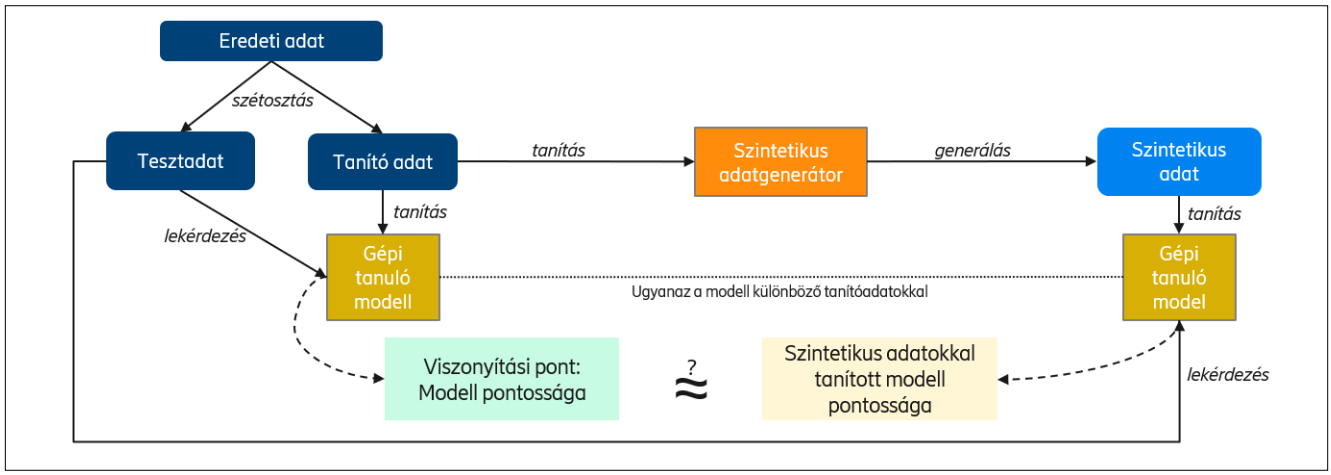
A kísérletekhez olyan adathalmazra volt szükség, amelyen mobilhálózatok szempontjából releváns gépi tanuló algoritmusokat lehet tanítani. Egy laborkísérlet mérési eredményeit használtuk: az adatokban szerepeltek a mobiltelefon rádiós környezetére vonatkozó információk, a maghálózatban mért adatforgalmi jellemzők, valamint egy videókonferencia-alkalmazás által feljegyzett adatok: a hang és videó minőségét leíró mérőszámok. Ez utóbbiak általában nem állnak a hálózatüzemeltetők rendelkezésre, ezért ezek becslését választottuk a kísérletünkben szereplő gépi tanuló modell feladatául.

A nyers adatrekordok sorrendezettek voltak és időbélyeget is tartalmaztak. Az első kísérlet arra irányult, hogy mennyiben tudják az egyes szintetikus adatgeneráló eszközök megőrizni egy ilyen adatsor tulajdonságait. Ehhez 6 idősoros tulajdonságot választottunk ki.

A második kísérlethez a nyers adatokat átalakítottuk időbélyeg és sorrendezés nélküli adathalmazzá. Ezen az adathalmazon egy Tensorflowban megvalósított, egysze-



2. ábra A GAN-architektúra



3. ábra A szintetikus adatgenerátorok kiértékelési módszertana

rű neurális hálót tanítottunk, amelynek feladata egy kiválasztott, alkalmazási rétegbeli mutató becslése volt a hálózati mérőszámok alapján.

Az adathalmazt teszt- és tanítóhalmazra bontottuk. A modell az utóbbi alapján tanult, míg az előbbi segítségével kiértékeltük a teljesítményét. Ez látható a *3. ábra bal oldalán*.

A modell tanítását megelőzően az adatok a gépi tanulásnál szokásos előfeldolgozáson mentek keresztül, például számos oszlopra – beleértve a céloszlop értékeit is – logaritmusfüggvényt alkalmaztunk. A modell becsléseinek átlagos hibájára a teszthalmazon mérve 0,2 adódott úgy, hogy a céloszlop logaritmusának értékei egy 4-es szélességű intervallumban szóródnak.

Megjegyzendő, hogy ennek a kísérletnek nem a választott mutató minél pontosabb előrejelzése volt a célja, hanem annak megállapítása, hogy mennyit romlik a modell teljesítménye, amikor valós adatok helyett szintetikus adatokon tanítjuk. Ennek a célnak a 0,2 átlagos becslési hibát adó modell megfelelt, ezért ennek javítására további erőfeszítéseket nem tettünk.

A tanító adatot ezután az egyes szintetikus adatgenerátorok tanítására használtuk, majd az eredeti tanítóhalmazzal azonos méretű szintetikus adathalmazt generáltunk az eszközökkel. Ezeken a szintetikus adatokon tanítottuk a korábban említett Tensorflow-modellt és megvizsgáltuk a kapott átlagos hibákat. Fontos, hogy a kiértékelés itt is az eredeti tesztadatokon történik. Ez látható a *3. ábra jobb oldalán*. Minél közelebb van ezen második modell teljesítménye az eredeti adaton tanított modell teljesítményéhez (azaz minél kevésbé nő az átlagos becslési hiba), annál jobb minőségű a szintetizált adat, illetve az azt létrehozó generátor eszköz.

A szintetikus adathalmazzal szemben azonban nemcsak az a követelmény, hogy minél hűebben megőrizze az eredeti adat statisztikai jellemzőit és a benne rejlő mintázatokat. Ugyanilyen fontos, hogy ne „szivárogtassa ki” az eredeti adat egyes rekordjainak információit, például ne legyen kikövetkeztethető a szintetikus adatból, hogy az eredeti adatban szereplő előfizetők mikor kinek telefonáltak, milyen szolgáltatásokat használtak a mobiljukkal és merre jártak. Egyes generátoreszközök automatikusan előállítanak egy dokumentumot a szintetizált adathalmazról, amelyben ezt az aspektust is elemzik, különféle „támadásokat” szimulálva.

Ezeken a generált dokumentumokon túl egy saját metrikával is vizsgáltuk, mennyire biztonságosak az egyes szintetikus adathalmazok. Megmértük, hogy milyen mértékben hasonló az eredeti adat teszt- és tanítóhalmaza egymáshoz („önhasonlóság”), majd azt is, hogy a szintetikus adat mennyire hasonlít az eredeti tanítóhalmazhoz („szintetikus adat hasonlósága az eredetihez”). Ha a szintetikus adat jobban hasonlít az eredetihez, mint amennyire az eredeti adat önhasonló, az arra utal, hogy a generátor eszköz nem valódi szintézist végez, hanem részben vagy egészben másolja az eredeti rekordokat.

5. Eredmények

A kereskedelmi szoftverekkel kapcsolatos eredmények bizalmasak, ezért az ebben a szakaszban a vizsgált négy eszközre SZ1-SZ4-ként hivatkozunk, véletlenszerű hozzárendeléssel.

Az első kísérletben idősoros adatokból kiindulva szerettünk volna új adatsorokat generálni, azonban erre egyik eszköz sem volt képes az elvárt minőségben. Az SZ2-szoftver tudott bizonyos metrikák alapján hasonló adatsort előállítani, mint a bemeneti halmaz, azonban a további vizsgálatok során kiderült, hogy bizonyos adatoszlopokat egy az egyben emelt át onnan, amely kizárja a szintetikus adatok használhatóságát az általunk érdekesnek vélt esetekben. A többi kereskedelmi eszköz a 4. szakaszban említett 6 idősoros tulajdonság közül csupán 2-4 tulajdonságot tudott megőrizni a szintetizált adatban. Hogy megértsük az idősoros adatok generálásának határait, készítettünk egy egyszerű, szinuszhullámot leíró adathalmazt, amit bemeneti adatként alkalmaztunk. Az SZ1 és SZ2 üzleti szoftverek képesek voltak lekövetni a mintázatot, és azt alapul véve új adatot tudtak generálni. A többi program esetében nem volt sikeres ez a kísérlet sem.

A 4. szakaszban ismertetett módszertant alkalmazva megmértük, hogy az egyes eszközök által generált szin-

tetikus adathalmazokkal tanított gépi tanulási algoritmus milyen hibával képes megbecsülni a céloszlop értékeit a mérés elején leválasztott teszthalmazon. Ennek eredményei az *1. és 2. táblázatban* láthatóak.

Összesen négyféle neurális hálón alapuló modellt használtunk különböző számú szintekkel és neuronokkal, ezek 1-4-ig vannak számozva a táblázatban. A táblázatokban az első sor az eredeti adatokon tanított és validált modell átlagos hibáját adja meg. A következő sorok az egyes szoftverek által előállított szintetikus adatokon tanított modell átlagos hibáját, valamint a referenciaértékhez képest számolt hibanövekedést írják le. A kereskedelmi szoftverek közül az SZ1 és SZ2 produkálta a legkisebb átlagos hibát a méréseink során. Az SZ2-program esetében azonban egyes kiugró értékeket ki kellett szűrnünk a teszthalmazból. Az SZ1-program adatain végzett mérés átlagosan 31%-kal mutatott rosszabb eredményt, mint a referenciamérés, míg az SZ2 esetén ez 13,75%. Jól látható, hogy nagy eltérés lehet a mérésekben a modell felépítésétől függően (SZ1 a legjobb: 18%, SZ1 a legrosszabb: 39%).

A nyílt forráskódú programcsomagok közül a Synthpop-szoftver ért el a kereskedelmi szoftverekhez hasonló eredményt. A többi ingyenes eszköz által generált adatok használatakor a modellek legalább 187%-kal rosszabbul teljesítettek, mint a referenciamérések során, a legrosszabb esetben ez a szám elérte a 385%-ot. A 2. táblázatban nem szereplő nyílt forráskódú eszközökkel a kísérlet sikertelen volt.

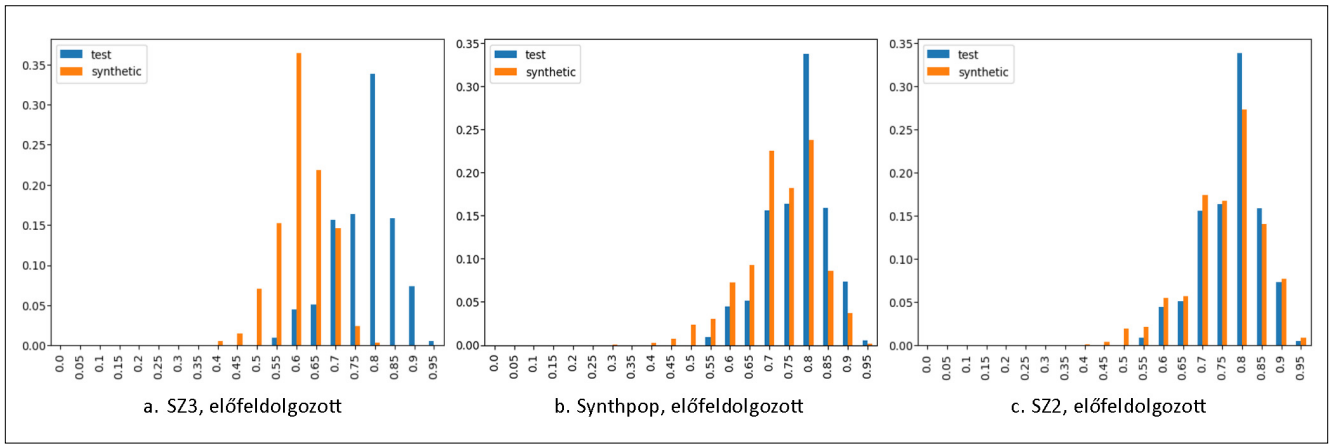
Egy másik kísérletben az eredeti és a szintetikus adatok közti hasonlóság azon dimenzióját szerettük volna leírni objektív metrikával, amely szerint az új adatpontokból nem tudunk visszakövetkeztetni az eredeti adatpontokra. Ennek mérésére a 4. szakaszban leírtaknak megfelelően hasonlósági mérőszámot alkalmaztunk egyrészt az eredeti adaton belül, másrészt az eredeti és a gene-

1. táblázat Kereskedelmi adatgenerátorok eredményei

	Időbélyeg nélküli adathalmaz				Előfeldogozott, időbélyeg nélküli adathalmaz			
	gépi tanulási modell:				gépi tanulási modell:			
	1	2	3	4	1	2	3	4
Referencia	0.2247	0.2087	0.2011	0.1951	0.2236	0.1953	0.2062	0.1943
SZ1	0.2767	0.2721	0.2752	0.2675	0.2649	0.2718	0.2666	0.2622
SZ1 hibatöbblet	23%	30%	37%	37%	18%	39%	29%	35%
SZ2	-	-	-	-	0.2426	0.2297	0.2323	0.2254
SZ2 hibatöbblet	-	-	-	-	8%	18%	13%	16%
SZ3	0.2967	0.2952	0.2918	0.2998	0.4517	0.5539	0.4792	0.6196
SZ3 hibatöbblet	32%	41%	45%	54%	102%	184%	132%	219%
SZ4	0.4793	0.5031	0.4804	0.4843	0.4848	0.4843	0.4756	0.488
SZ4 hibatöbblet	113%	141%	139%	148%	117%	148%	131%	151%

	Időbélyeg nélküli adathalmaz				Előfeldogozott, időbélyeg nélküli adathalmaz			
	gépi tanulási modell:				gépi tanulási modell:			
	1	2	3	4	1	2	3	4
Referencia	0.2247	0.2087	0.2011	0.1951	0.2236	0.1953	0.2062	0.1943
Gretel (nyílt)	-	-	-	-	0.6423	0.629	0.595	0.6506
Gretel hibatöbblet	-	-	-	-	187%	222%	189%	235%
Synthpop	0.2578	0.2355	0.254	0.2401	0.2563	0.2412	0.286	0.2446
Synthpop hibatöbblet	15%	13%	26%	23%	15%	24%	39%	26%
SDV	0.6261	0.56764	0.57009	0.58682	0.71699	0.75138	0.54922	0.69036
SDV hibatöbblet	279%	272%	283%	301%	321%	385%	266%	355%

2. táblázat Nyílt forráskódú adatgenerátorok eredményei



4. ábra
Adatpontok hasonlósága az eredeti adatokon belül (kék), valamint az eredeti és szintetikus adatpontok között (narancs)

rált adat között. A mérések során szinte az összes szintetikus adathalmaz megfelelőnek bizonyult, mivel a szintetikus adat nem hasonlított jobban az eredetihez, mint az eredeti adat önmagához. A *4.a. ábra* egy ilyen sikeres mérés eredményét mutatja.

Az SZ3-eszköz által generált szintetikus adatok elemenkénti hasonlósága az eredeti adathoz narancssárgával szerepel az ábrán, és mivel ez balra tolódva található a kékkel jelölt eredeti adathalmazon belüli hasonlósági hisztogramtól, sikeresnek tekinthetjük a mérést. Hasonló ábrákat kaptunk a többi eszköznél is, amelyek nem szerepelnek a fenti ábrák között. A Synthpop esetén *(4.b. ábra)* a hisztogram eltolódása jóval kevésbé szembetűnő, de még elfogadható. Az SZ2-eszköz diagramján pedig látható, hogy a kívánt tulajdonság a nagy hasonlóságú rekordok esetén sérül (*4.c. ábra,* lásd a hisztogramok jobb szélét). Ez az eredmény utalhat arra, hogy a szintetikus adathalmaz bizonyos pontjaiból nagyobb valószínűséggel lehet visszakövetkeztetni az eredeti adatpontokra, mint az eredeti adathalmaz tetszőleges adatpontjaiból a többire.

Végül az eszközök nem funkcionális képességeivel kapcsolatos tapasztalatokat foglaljuk össze:

• **Grafikus felhasználói felület:** A vizsgált nyílt forráskódú szoftverek egyike sem rendelkezik grafikus felhasználói felülettel. Az SZ1 és SZ3 kereskedelmi szoftverek teljes grafikus támogatással érkeznek, míg az SZ2- és SZ4-programok részlegesen vezérelhetők a grafikus felületen keresztül. Utóbbi esetekben néhány funkció csak a parancssoros felületen érhető el.

• **Dokumentáció:** Az SZ1-szoftver esetében teljeskörű dokumentációt kapunk. SZ2 esetében a dokumentáció példákon keresztül mutatja be a program használatát, míg az SZ3- és SZ4-programok használatához a gyártó munkatársainak utasításait követtük többségében. A nyílt forráskódú megoldások közül mindegyik dokumentációval érkezett, melyekben, bár lehetett hibát találni, alapvetően jól leírták a szoftvereket felhasználói szempontból.

• **Használhatóság:** Érdemes megjegyezni, hogy ez a pont inkább informatív jelleggel szerepel, hiszen részben a programok használata során szerzett szubjektív vélemények alapul. Az SZ1-, SZ3- és SVD-szoftvereket találtuk a legkönnyebben és legintuitívabban használhatónak. A Synthpop-szoftver az automatikus típusfelismerés hiányát leszámítva egy letisztult programozói interfészt biztosít. Az SZ2- és a nyílt forráskódú Gretel-programok limitáltan felhasználóbarátok (például olyan egyedi koncepciók alkalmazása, amik nincsenek megfelelően dokumentálva), míg a SZ4-program használata nagyobb nehézségek árán volt csak lehetséges.

• **Konfigurálhatóság:** A vizsgált programok többségében széles skáláját támogatták a beállítási lehetőségeknek. A Synthpop-szoftvernek a Python-alapú csomagját használtuk, azonban az eredeti csomag R-ben íródott és a konfigurációs paraméterek egy jelentős része nem érhető el a Python-alapú verzióban.

• **Robusztusság:** A tesztelt szoftverek általában egy-egy hibától eltekintve stabilan működtek. Az SZ1-program kiemelkedően stabilnak bizonyult. Az SZ2- és SZ4-programok azonban olyan hibákat produkáltak, melyekre nem számít a felhasználó és nincs is kész megoldás az orvoslásukra. (A használhatósági jellemzőkhöz hasonlóan ez a metrika is informatív jelleggel szerepel, mivel a programok használata során szerzett szubjektív tapasztalatokon alapul.)

• **Jelentésgenerálás:** A kereskedelmi szoftverek mindegyike képes felhasználónak szánt (például pdf formátumú) jelentést generálni az előállított adatok statisztikai

3. táblázat
Az adatgenerátorok használhatósága a nem funkcionális kritériumok alapján

Szoftver	Használhatóság
SZ1	6
SZ2	3.5
SZ3	5
SZ4	3
SDV	4
Synthpop	3
Gretel (nyílt)	2.5

jellemzőiről, valamint az eredeti adatokhoz való hasonlóságukról. Ezzel szemben a nyílt forráskódú programok nem nyújtottak ilyen jellegű funkciókat, kivéve a Gretel nyílt változatát, amely képes nagyon limitált jelentéseket generálni.

A *3. táblázatban* a korábban felsorolt nem funkcionális tulajdonságokat egyesével értékelve (0-1) majd azokat összeadva elért pontszámok láthatóak. A magasabb pontszám jobb használhatóságot jelent.

6. Összefoglalás

Cikkünkben bemutattuk, hogy a gépi tanulási algoritmusok fejlesztéséhez gyakran nem érhető el a szükséges adat, amelyre bizonyos esetekben megoldást jelenthet szintetikus adatgenerátorok alkalmazása. Megmutattuk, hogy személyes vagy érzékeny adatok, ritka és kevés adat esetén is érdemes megfontolni ezen eszközök alkalmazását, viszont ezek körültekintő használatot igényelnek. Megmutattuk, hogy milyen módszereken alapulnak a jelenleg elérhető szintetikus adatgeneráló algoritmusok. Bemutattuk az általunk használt mérési környezetet, melyben neurálisháló-alapú algoritmusok tanításának pontosságát értékeltük ki egy valós és különböző eszközök által generált szintetikus adathalmazokon.

Vizsgálatunk során négy kereskedelmi és hat nyílt forráskódú szoftvert használtunk. A kísérletek három nyílt forráskódú eszközzel sikertelenek voltak. Egy kereskedelmi eszközt (SZ2) pedig az adatvédelemmel kapcsolatos aggályok miatt zártunk ki. A maradék eszközök közül a legjobbak (a kereskedelmi SZ1 és a nyílt Synthpop) csupán 15-18%-os hibatöbbletet produkáltak a legkedvezőbb esetben és 23-31%-osat átlagosan. Láttuk azonban, hogy a különböző eszközök között nagy a különbség, illetve egy-egy programon belül is függ az eredmény a vizsgált neurális háló tulajdonságaitól.

Összegyűjtöttük továbbá a programok használatakor tapasztalt, nem funkcionális képességekkel kapcsolatos észrevételeinket. A kereskedelmi szoftverek átlagosan felülmúlták a nyílt forráskódú programokat mind funkcionális, mind nem funkcionális képességek tekintetében. A Synthpop nevű programcsomag kiemelkedett a nyílt forráskódú programok közül, azonban ez az eszköz egyáltalán nem támogatja az idősoros adatokat.

A mérések alapján láthatjuk, hogy a szintetikus adatgenerátorok gyakorlati alkalmazása lehetséges, azonban számítani kell a felmerülő többlethibára. Ez azonban elfogadható, különösen akkor, ha szintetikus adatgenerálás nélkül egyáltalán nem hozzáférhető az eredeti adat az adott projekt számára. Az idősoros adatok generálása kapcsán problémákba ütköztünk: ezen a területen az eszközök további fejlődésére számítunk.

Hivatkozások

[1] Perez, L., & Wang, J. (2017): The effectiveness of data augmentation in image classification using deep learning. arXiv preprint arXiv:1712.04621.
[2] Mostly.AI weboldala: *https://mostly.ai/synthetic-data-platform/*
[3] YData weboldala: *https://ydata.ai/*
[4] Statice weboldala: *https://www.statice.ai/*
[5] Synthetic Data Vault weboldala: *https://sdv.dev/*
[6] Lei Xu, Maria Skoularidou, Alfredo Cuesta-Infante, Kalyan Veeramachaneni: Modeling Tabular Data using Conditional GAN, Advances in Neural Information Processing Systems, Vol. 32, 2019.
[7] Gretel weboldala: *https://gretel.ai/*
[8] Gretel nyílt forráskódú verziója: *https://github.com/gretelai/gretel-synthetics*
[9] YData nyílt forráskódú verziója: *https://ydata.ai/products/synthesizer*
[10] Derek Snow: DeltaPy: Tabular Data Augmentation, 2020. *https://github.com/firmai/deltapy*
[11] Synthpop Python implementáció: *https://github.com/hazy/synthpop*
[12] Synthpop R implementáció: *https://www.synthpop.org.uk/*
[13] Liyang Xie et. al.: Differentially Private Generative Adversarial Network, CoRR, 2018.
[14] Martin Arjovsky et. al.: Wasserstein GAN, arXiv, 2017.
[15] Syndata implementáció: *https://github.com/LLNL/SYNDATA*
(weboldalak elérési dátuma: 2023.02.28.)

A szerzőkről

RECSE ÁKOS 2016-ban mérnökinformatikus szakon BSc-, 2018-ban pedig gazdaságinformatikus szakon MSc-fokozatot szerzett a Budapesti Műszaki és Gazdaságtudományi Egyetemen. 2018 óta az Eötvös Loránd Tudományegyetem Informatikai Doktori Iskolájának hallgatója. 2019-től a European Institute of Innovation & Technology (EIT) doktori iskolájának hallgatója. Doktori tanulmányai során főbb témái az elosztott, felhőalapú infrastruktúrák automatizált felügyelete és erőforrás-menedzselése, valamint a decentralizált adattárolási szolgáltatások vizsgálata. 2022 óta dolgozik az Ericsson Magyarországnál, System Architect pozícióban.

DÉVAI GERGELY jelenleg az Ericsson mobilhálózat-analitikai és hálózatautomatizálási termékeivel, valamint a kutatási eredmények termékesítési lehetőségeivel foglalkozik. Diplomát és doktori fokozatot az Eötvös Loránd Tudományegyetem Informatikai Karán szerzett, ahol azóta is oktat. Számos ipari-egyetemi kutatási projektet vezetett, főként szoftvertechnológiai témákban. Érdeklődése középpontjában az analitikai rendszerek praktikus kihívásai állnak: adathozzáférés, zártláncú automatizálás, az adatfeldolgozás erőforrásigényének csökkentése.

Digital Trust Services adathitelesítő platform – blockchain alapú digitális ujjlenyomat tárolással

SOLYMOS GYULA
4iG Nyrt.
gyula.solymos@4ig.hu

Kulcszavak: blockchain, adathitelesség, Digital Trust Services, AI, felhőszolgáltatás, ZERO TRUST, ellátási lánc, termékkövetés, drón-feketedoboz

Az adatalapú gazdaságnak hiteles adatra van szüksége, amelynek a forrása, változatlansága igazolt, hogy arra alapozva kialakulhassanak a széleskörű adatalapú együttműködések, szak- és államigazgatás, valamint a másodlagos adatpiacok. Különös figyelmet kell fordítani az AI-megoldások betanító adatai hitelességének megvalósítására, mert csak így biztosítható ezen megoldások átlátható és biztonságos működése.
A blockchain-technológiára épülő Digital Trust Services egy digitálisujjlenyomat-alapú adathitelesítési platform megoldás, amely segítségével a digitális adatok széles körének hitelesítése és hitelességének ellenőrzése – iparág- és vállalkozásmérettől függetlenül – mindenki számára elérhetővé válik.

1. Bevezetés

A blockchain-technológiát sokan a kriptovalutákon keresztül ismerték meg és sokan még ma is azt gondolják, hogy egy energiafaló technológia, amit alapvetően a fintech világ szereplői tudnak kamatoztatni. Ezen a szemléleten ideje változtatni, hiszen *a blockchain a megbízhatóság technológiája,* amely segítségével elérhetjük, hogy egy adat hitelessége ne a jogszabályok betartásán és az adminisztráció megbízhatóságán múljon.

Napjainkban egyre több adatot állítunk elő és továbbítunk, és a formálódó adatalapú társadalomban egyre nagyobb az igény adatalapú együttműködésekre, ugyanakkor rengeteg hamisított adat vesz minket körül, ezért sokan elvesztették bizalmukat a digitális adatok iránt. Kétségtelenül egyre nagyobb veszélyt jelent a manipulált adat, amely előállítása a mesterséges intelligencia-megoldások használatával egyre könnyebb [1]. Miközben adatalapú társadalmat építünk, a manipulált adatok egyre nagyobb veszélyt jelentenek az adatalapú együttműködésekre és ellátási láncokra, adat-piacterekre, valamint a szak- és államigazgatásra is [2].

A cikkben bemutatásra kerül a *blockchain-technológiára épülő Digital Trust Services (DTS) platform megoldás,* amely megvalósításán az EU Important Projects of Common European Interest (IPCEI-CIS) programjában dolgozunk, és amely az új eIDAS reguláció tervezetre [3] alapozva képes lesz újraértelmezni a hagyományos adat fogalmát és megteremteni a *megbízható, ellenőrizhető adatalapú digitális ökoszisztéma* hátterét.

A DTS egyszerűségének és a jövőben szabad szoftverként elérhető digitális ujjlenyomatot előállító kódnak köszönhetően *nagyfokú interoperabilitást* fog biztosítani. Könnyen beépíthető lesz az adatgyűjtő szenzorokba és IoT-eszközökbe, valamint a kommunikációs és felhőmegoldásokba, így az adathitelesség automatikusan beépülhet minden cég és szervezet IT-megoldásába és így a mindennapi munkájába. A megoldás képes lesz a meglévő adatvagyon hitelesítésére, valamint a hiteles adatfelhasználás követésére is, amely az adatalapú együttműködések és adatpiacterek elterjedésének alapja lehet.

A cikkben röviden bemutatásra kerül egy általános célú fájlhitelesítő, valamint két ipari, egy agrár és egy egészségügyi területen fejlesztés alatt lévő laborinformatikai alkalmazáshoz kapcsolódó – a DTS pilotjának tekinthető – adathitelesítő blockchain-megoldás működése.

2. Adathitelesítési igény

Egyre több adatot termelünk a mindennapi tevékenységünk során, ugyanakkor azt tapasztaljuk, hogy ezekben az adatokban egyre kevésbé bízhatunk meg. Ennek oka, hogy egyre több olyan esetre derül fény, amikor akár globálisan is ismert cégek a publikált műszaki, termékminőségi, statisztikai adataikat a saját érdekeik szerint „kozmetikázzák” (pl. gépjármű CO2-kibocsájtása), vagy akár a termék származási, vagy gyártási adatait „zöldre mossák” [4]. További problémát jelent az egyre szaporodó IoT-eszközök száma [5], amelyek többnyire könnyen feltörhető kommunikációs csatornákon, illetve publikus internetes továbbítással közlekednek, és sokszor elsőként felhőmegoldásokban kerülnek letárolásra, így azok hitelessége és változatlansága sok szempontból megkérdőjelezhető [6].

A bizalmatlanságot tovább fokozza a rengeteg megtévesztésre alkalmas manipulált kép-, hang- és videóanyag, amivel nap mint nap találkozhatunk, és amelyek előállítása az interneten elérhető AI-alkalmazásokkal napról napra könnyebbé válik, így manapság ezen manipulációkat már nem csak a nagytudású szakemberek, hanem laikusok is meg tudják valósítani [7]. De nem csak képeket, videókat, hanem szinte mindenféle adatot egy-

re könnyebb észrevétlenül manipulálni, ideértve az IoT-szenzorok mérési, egészségügyi, diagnosztikai, GPS-követési, mezőgazdasági és egyéb adatokat is.

A manipulált adatokkal a hackerek könnyen összezavarhatják az együttműködő IT-rendszereket (például az energiaszektorban) anélkül, hogy bejutnának az adott cég tűzfalán belülre. *Növekvő veszélyt jelent a meglévő – akár éveken át gyűjtött – szakterületi (pl. egészségügy, agrárium), illetve szak- és államigazgatási adatbázisok manipulációja,* mert ha ezen adatok alapján hozunk döntéseket, vagy tanítunk be a jövőben AI-megoldásokat, akkor az nem az általunk elvárt eredményt, hanem a manipuláló érdekeit fogja erősíteni.

Képzeljük el, hogy mi történhetne, ha egy hacker hozzáférne az EU stratégiai adatbázisainak valamelyikéhez, és a saját érdekei szerint manipulálná azt, és mivel jelenleg nincs olyan megoldásunk, amivel meg tudnánk különböztetni a manipulált adatot a hitelestől, a következő naptól ezzel az adatbázisra alapozva végeznénk a munkánkat és hoznánk döntéseket.

Miközben digitális gazdaságot és társadalmat építünk, amiben egyre nagyobb szerepe van az adatoknak, a mindennapos és egyre könnyebb adatmanipuláció révén elveszett a bizalom a digitális adatokban! Ez a bizalomvesztés gátolja az adatalapú együttműködések, a másodlagos adathasználat és az adatpiaciok kialakulását is.

2.1. Az adathitelesség iránti igény legfőbb területei

Az adathitelesség igénye egyre erősebben jelenik meg az alábbi területeken:

- Adatvezérelt üzleti működés
- Digitális adatalapú együttműködések
- Termékkövetés
- Adatvezérelt közigazgatás
- V2X- és M2M-kommunikáció
- Adatpiacterek
- ZERO TRUST-alapú rendszerek [8]

2.2. AI/MI betanító adatok hitelességének kritikus fontossága

A gépi tanulási megoldásokhoz *nagy mennyiségű megbízható adatra van szükség,* mivel ezen rendszerek azok alapján tanulnak és fognak működni [9]. Amennyiben egy ilyen betanító adatbázisba nem az adott eljáráshoz tartozó, vagy véletlen, illetve sérült adatok kerülnek, úgy az MI-alapú megoldás hibásan fog működni.

Még jelentősebb problémát okozhat, ha az MI-megoldásokat rosszindulatúan (hackerek, vagy az adatokhoz hozzáférő belső emberek), valamely érdek szerint szisztematikusan manipulált adatokkal tanítják be [10]. Ezzel könnyen elérhető, hogy az AI/MI-alkalmazások észrevétlenül az érdekeknek megfelelően működjenek, például egy adott termék választását, vagy manipulatív üzleti döntést javasoljanak.

A gépi tanulási megoldások a betanítás után egyfajta fekete dobozként működnek és a hagyományos algoritmusokhoz hasonlóan a működésük taxatív módon nem ellenőrizhető, – ugyanakkor egyre kritikusabb döntések megalapozásához hívjuk őket segítségül –, ezért csak akkor bízhatunk meg bennük, ha biztosítjuk a betanító adatok hitelességét és manipuláció-mentességét, így ezek *ellenőrizhetőségének és hitelességének megteremtésére, valamint jogszabályi előírására is kritikusan nagy figyelmet kell fordítani*.

2.3. A digitális adat, mint termék hitelessége

A hackerek ma már nem csak eltulajdonítani, vagy blokkolni akarják az adathozzáférést, hanem egy annál sokkal veszélyesebb elkövetési mód; *az adatmanipuláció, mint hacker-módszer* kezd kibontakozni, ami azokra a cégekre jelent még nagyobb veszélyt, amelyeknek a terméke maga a digitális adat.

Egy tengerentúli egészségügyi szolgáltatót már ért a közelmúltban ilyen incidens, melynek során a támadók a több tízezer elért adatból csupán néhány tucat képalkotó diagnosztikai leletet manipuláltak AI segítségével, de mivel nem tudták ellenőrizni, melyik kép hiteles, vagy melyik hamis, kénytelenek voltak a hackerekkel megállapodni.

3. Adathitelesítés: Global State of the Art

Digitális dokumentumok hitelesítésére a cégek és az igazgatás szereplői jellemzően a „digitális aláírást" használják [11], amely bevett gyakorlat, de nem alkalmas az egyre nagyobb tömegben előállított olyan digitális adatok hitelesítésére és hitelességének ellenőrzésre, mint az IoT-adatok, szerverlogok vagy AI-betanításhoz használt adatok, illetve a meglévő adatbázisok, amelyeket a cégek meg szeretnének osztani egymással, vagy a formálódó adatpiacokon értékesíteni.

Vannak ugyan SSL- [12] és egyéb tanúsítványok az operációs rendszerekben és a böngészőkben, és sok helyen alkalmazzák a hash-képzési technológiát is [13], de ezek nem jelentenek egységes, mindenki számra elérhető olyan megoldást vagy eszközt, amelyekkel a felhasználó teljes bizonyossággal ellenőrizni tudná, hogy egy általa felhasznált adat a keletkezésétől, illetve a hitelesítéstől kezdődően változatlan.

A fentiek miatt egyre több cég keresi az adathitelesítés alternatív lehetőségeit, és egyre többen kezdik el használni a Blockchaint, mint hitelesítési technológiát. A ma épülő új adathitelesítési megoldások azonban teljesen zárt rendszerek, így csak egy-egy együttműködő cégháló igényeit képesek kielégíteni. Erre jó példa a PharmaLedger platform [14], amely a gyógyszeripar termékkövető és adatmegosztó megoldása, vagy – a Magyar Nemzeti Bank által indított – a Magyarországon működő bankok és biztosítók közt történő hiteles adatmegosztást biztosító blockchain-alapú rendszer. Mindkettő alapvetően egy zárt üzleti csoport igényei alapján lett kidolgozva, és nem alkalmas arra, hogy ez a két „siló" rendszer hitelesen osszon meg adatokat egymás közt, azaz abszurd módon nincs például lehetőség arra, hogy

egy gyógyszergyár a digitális adatait hitelesen megosszsza a bankjával, hiába van mindkettőnek adathitelesítést szolgáló megoldása.

Kijelenthető, hogy a technika mai állása szerint a cégek (ill. iparágak) egyedi, zárt adathitelesítési megoldásokat fejlesztenek, melyek egymás közt nem átjárhatóak, ami így gátolja az iparágakon átívelő adatalapú együttműködéseket, az adatpiacok kialakulását.

4. Blockchain-adathitelesítés és problémái

A blockchain-adathitelesítés az adatok összeláncolásán túl a blockchain „validátor" csomóponti szerverek konszenzus mechanizmusára épít. Ennek következtében a láncba írt adatok minden csomóponti szerverhez hozzáférő szervezet, illetve annak felhasználója által elérhetővé válnak, amely üzleti alkalmazás esetén több kérdést is felvet.

Publikus blockchain-hitelesítés

Egy adat vagy dokumentum hitelességét úgy biztosíthatjuk a legegyszerűbben, ha azt egy publikus blockchain-láncra (pl. Ethereum [15]) feltöltjük. Ez az adat bármikor később kiolvasható a publikus blockchainből, és mivel a működéséből adódóan az nem megváltoztatható, ezért biztosak lehetünk benne, hogy a letöltéskor a feltöltött adat vagy dokumentum azonos lesz a feltöltött anyaggal.

Mivel azonban a publikus blokkláncok teljes tartalmát minden validátor-csomópont (akár több százezer ismeretlen szereplő) eléri, azok üzleti adatok kezelésére csak erősen korlátozottan használhatóak, mivel az adott adat vagy dokumentum „mindenki" számára elérhetően jelenik meg.

A publikus láncok további korlátja, hogy azok működésének díjfizetése az adott lánc saját fizetőeszközéhez kötött (Ether, Bitcoin stb.), így ezen rendszerek használatának nagy a pénzügyi kitettsége és volatilitása. További probléma, hogy a különösen a kisebb láncok esetében a gyakorlatban nem elképzelhetetlen a validátorok 51%-os rosszindulatú összejátszása, ami pont a tárolt adatok integritásába vethető bizalmat kérdőjelezi meg.

Enterprise, „zárt" blockchain és az adattárolás problémája

A publikus blokkláncok korlátozott üzleti felhasználhatóságát az enterprise blokkláncmegoldások segítségével kiküszöbölhetjük azáltal, hogy egy ilyen megoldásban *minden validátor-csomópontot üzemeltető szereplő ismert*. Ezt a megoldást előszeretettel használják az azonos iparágban együttműködő cégek és ellátási láncok a termékeikhez kapcsolódó digitális adatcseréjük hitelesítésére (pl. PharmaLedger).

Amíg egy termékhez tartozó adatot hitelesítünk és osztunk meg másokkal, addig ez a megoldás kiválóan működik, viszont, ha egy olyan rendszert szeretnénk felépíteni, amiben konkurens cégek, illetve üzleti szereplők, vagy azokat felügyelő szervezetek is részt vesznek, és *nem szeretnénk, hogy a hitelesítő csomópontot futtató szervezetek láthassák mindenki blokkláncba feltöltött üzleti adatát, akkor elértük ezen megoldások korlátját.*

Természetesen titkosítva is feltölthetjük a blokkláncba az adatokat, de annak kezelése jelentősen megnöveli a kiépítendő rendszerünk komplexitását, erőforrásigényét, árát, és nem utolsó sorban jelentősen lassítani fogja annak működését.

Szintén problémát jelent, hogy az együttműködő cégek iparáganként egyedi adattárolási metódusokat és megoldásokat fejlesztenek. Ezekben hiába vannak blockchain-megoldásra építve hitelesen tárolva az adatok, ha egy másik iparág vagy a rendszerhez csatlakozó külső szereplő számára egy adott adat hitelessége nem igazolható. *Ezek a rendszerek ezért csak siló-szerűen, más iparágak vagy szereplők felé történő interoperábilitás lehetősége nélkül működnek.*

5. Digital Tust Services (DTS) platform és működése

5.1. A DTS célja és eredménye

A DTS célja olyan *blockchain-technológiára épülő, iparágtól és adattípustól független, interoperábilis hitelesítő és hitelességellenőrző felhőszolgáltatás létrehozása,* amely megteremti az adatok tömeges hitelesítésének platformját.

A projekt fő célja az adatok (IoT, szenzor, kép, videó stb.) keletkezésekor, illetve első kommunikációs csatornákba kerülésekor történő azonnali hitelesítése, de a megoldásunk lehetőséget biztosít a már meglévő adatvagyon és speciális adatok (pl. security, programkód) hitelesítésére is, megteremtve ezzel a hiteles adatok hasznosítására, illetve egymás elektronikus dokumentumainak elfogadására épülő együttműködéseket, valamint az adatcsere és az adathozzáférések hiteles visszakövethetőségét is.

A projekt eredményeként egy olyan, több országban is alkalmazható adathitelesítési platform jön létre, amely alkalmas nagymennyiségű adat hitelesítésére és annak felhasználása előtti hitelesség, változatlanság és teljeskörűség ellenőrzésére. A platformhoz, annak interfészei segítségével, bármely meglévő szoftvermegoldás csatlakozható, így annak hitelesítése a meglévő iparági és közigazgatási, valamint adatcsere- és piactér-alkalmazásokba is könnyen beépíthető lesz.

A blockchain- (distributed ledger) technológia hitelessége az EU elektronikus azonosításról és bizalmi szolgáltatásokról szóló *eIDAS 2.0 rendeletének* új szövegtervében is szerepel, így annak hitelessége magas szinten is elismert, ami a *DTS-megoldás jogi hátterét* biztosítja.

Projektünk célja, hogy az Európai Uniós megbízható adatterek kialakítását és adatfelhasználási regulációk és célok elérését támogassa a *GAIA-X, Data Governance Act, AI Act, EU Trust Mark, New Industrial Startegy, SME Strategy* törekvésekhez illeszkedve.

5.2. A DTS működésének rövid műszaki összefoglalása

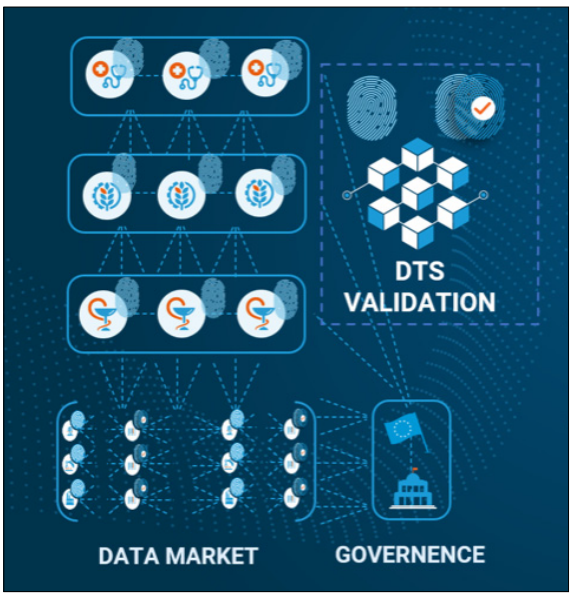
A DTS-platform adathitelesítése (ideértve a dokumentumokat, média- és egyéb elektronikus anyagokat is) az adatok digitális ujjlenyomatának, avagy #HASH-értékének előállításán és egy olyan zárt, enterprise blockchain megoldást alkalmazó tárolásán alapul, amely hitelesítő szervezetek által üzemeltetett független csomópontszerverek elosztott adattárolásával *szavatolja a hitelesítő #HASH megváltoztathatatlan tárolását és hiteles visszakereshetőségét.*

A #HASH egy olyan kriptográfiai módszerekkel előállított digitális leképezése az adatrekordnak, dokumentumnak, vagy bármilyen fájl tartalmának, amely az eredeti adat birtokában bármikor megismételhető, de ha az eredeti adat tartalmában a legkisebb módosulás is történt (például a dokumentumba belekerült egy „.") annak értéke már nem fog megegyezni az eredetivel. A #HASH ugyan a nyers adatból készül, de abból az adat maga semmilyen formában nem rekonstruálható.

A DTS-rendszerhez csatlakozók egy egységes #HASH MAKER modult kapnak, amelyet a saját hálózatukon belül telepítenek, és interfészek segítségével az adatgyűjtő, illetve kezelő megoldásaikkal összekapcsolnak, így az adatok hitelesítése (annak beállításainak megfelelően) automatikusan meg tud valósulni. A hitelesítő modulon az adat áthalad, de *a DTS-platform számára kizárólag csak a #HASH értékek és annak metaadatai (partnerazonosító, időbélyeg stb.) kerülnek továbbításra,* így a cég adatai házon belül maradnak.

Ezen működésnek köszönhetően a megoldás megfelel a GDPR követelményeinek, amely így számos üzleti területnek megoldást jelenthet, ahol most ezen szabályok betartása nehézségbe ütközik.

A digitális ujjlenyomat ugyan a már ismert megoldásra épül, de nem egy egyszerű #HASH, hanem egy olyan innovatív megoldás, ami a metaadatok segítségével *az adott adattípushoz tartozó egyedi munkafolyamait és üzleti elszámolási funkciókat is ki tudja szolgálni,* ugyanakkor nagyfokú interoperábilitást biztosít a különféle iparágak és ellátási láncok között.



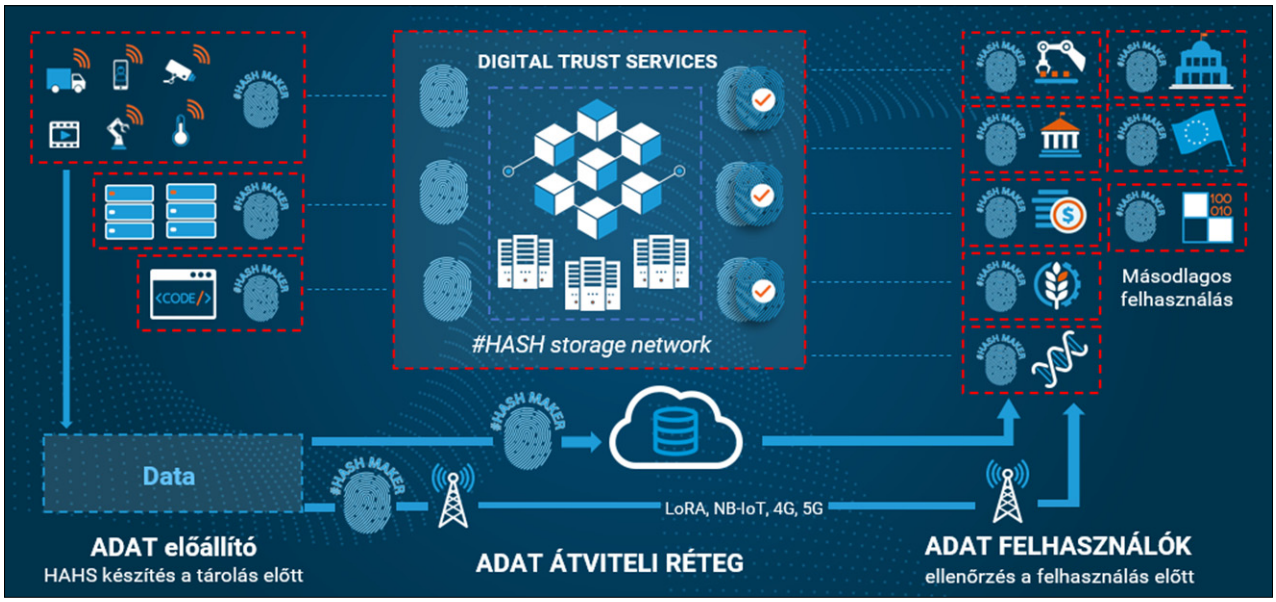
Amennyiben az adatujjlenyomat-előállítás az adatot előállító eszközben nem biztosítható, úgy a #HASH-képzés az azokat továbbító (pl. 5G, NB-IOT, LoRaWan műhold stb.) szolgáltató hálózatába érkezve – egyfajta értéknövelt szolgáltatásként – abba beépítve valósítható meg.

A felhőszolgáltatók szintén integrálhatják a #HASH MAKER megoldást a rendszereikbe, amely segítségével a felhőben tárolt, majd onnan felhasznált adatok változatlanságát az ügyfeleik ellenőrizni tudják, ezzel *növelve a bizalmat a felhőalapú tárolási megoldásokban.* A rendszer ugyancsak alkalmas a ZERO TURST elvárás szerint a felhőben és platformokon futó szoftvermegoldások és AI-algoritmusok kódjainak módosítás- és manipulálás-mentességét igazolni, ami napjainkban egyre fontosabb.

ADATMEGOSZTÁS – HITELESSÉGELLENŐRZÉS

Az adatgazda az adatot továbbra is a megszokott módon (direkt kapcsolat, API, e-mail stb.) oszthatja meg a partnereivel vagy harmadik féllel, akik a megkapott adatról a DTS #HASH MAKER megoldása segítségével előállítják annak digitális ujjlenyomatát. Ennek DTS-rendszerbe történő beküldésével és ellenőrzésével *az adatot fogadó partner, üzleti, vagy közigazgatási szereplő ellenőrizni tudja az adat tartalmának sértetlenségét, teljeskörűségét, a hitelesítés keletkezésének időpontját, illetve az azt előállító entitás (természetes/jogi személy vagy eszköz) adatait is.*

Mivel a *DTS #HASH MAKER szabad szoftverként elérhető megoldás* lesz, így annak kódját bárki elérheti, beépítheti a meglévő alkalmazásába, weboldalába, szolgáltatásába, vagy végfelhasználóként applikációként letöltheti és így könnyen ellenőrizheti bármely általa (levélben, weboldalon, interfész stb. keresztül) megkapott adat, kép, dokumentum stb. hitelességét és változatlanságát.



DTS – Hitelesítés- és hitelesség-ellenőrzés digitális ujjlenyomat segítségével

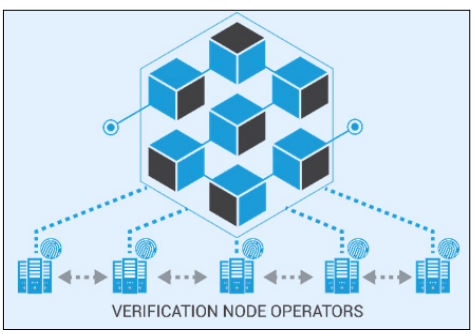
Amennyiben az adattulajdonos az *adat-nyomkövetési funkciót* is bekapcsolja, úgy a DTS-platform blockchain-okosszerződés (smart contract) segítségével egyedileg engedélyezi/tiltja az adathitelesség ellenőrizhetőségét, így a rendszerünk az illetéktelenül megszerzett adatok hitelesként való felhasználását meg tudja akadályozni. Mivel ez esetben a DTS logolja a hitelesség ellenőrzők azonosítóit, így az adathasználók köre is visszakövethetővé válik.

Az európai országokon átívelő együttműködésekben érintett adatok megosztásakor azok digitális ujjlenyomatát (#HASHértékeket) az érintett országok DTS-platformjai mellett az *EBSI Európai Blockchain Szolgáltatásban* is tervezzük letárolni, illetve hitelesíteni.

5.3. Többcsatornás adathitelesítés – az igényekre szabott bizalom garanciája

A blockchain-rendszer az adatok összeláncolásán túl, azok hitelességét a DLT (Distributed Ledger Technology), az elosztott főkönyvi rendszer [16] garantálja. Ezen főkönyvek – azaz a DTS-blockchainben tárolt digitális ujjlenyomatok a verifikációs node-ot üzemeltetők szerverein is megjelennek.

A hitelesség konszenzusos igazolásához azonban minimálisan elegendő egy lánc három csomóponton való tárolása. Erre építve a DTS egy úgynevezett többcsatornás adathitelesítési modellt épít fel, ahol az *egyes iparágak vagy együttműködő szereplők maguk választhatják ki és dönthetik el, hogy mely szereplők futtassák közülük a hitelesítésben résztvevő csomóponti szervereket.* Ezáltal egy ellátási láncban nem szükséges minden szereplőnek validációs csomópontot fenntartania, elég, ha azon szereplők közül három üzemeltet ilyen szervert, akikben a többiek megbíznak.



A hitelesítő szereplők ugyanakkor nem csak az adatelőállításban résztvevő szereplők lehetnek, hanem akár közigazgatási, kamarai, minisztériumi szervezetek is, akit maga az iparág vagy az állam nominál a feladatra. Fontos ugyanakkor ismételten megerősíteni, hogy mivel a DTS adatujjlenyomat-alapú működéséből következően nem kerül adat a blockchainba, így *a hitelesítésben résztvevő szereplők csak az adathoz tartozó #HASH- és metaadatokat tárolják, magát az adatot nem!*

Annak, hogy a DTS-ben a hitelesítés csatornákra osztható, a legfőbb előnye, hogy minden csatorna a szereplői által megbízhatónak ítélt szereplőkkel validáltathatja az adatot. A hitelesítő szereplők közt átfedések is lehetnek, azaz lehet olyan szereplő, aki több csatorna hitelesítési konszenzusában is részt vesz. Ilyenek lehetnek a fent is említett állami, kamarai, vagy szakigazgatási szereplők, amelynek köszönhetően a hitelesített adatokat *az állami folyamatokban is elfogadják, ami az adatalapú közigazgatás megteremtésének alapfeltétele.*

5.4. Alkalmazási példák

- IoT- és szenzoros adatgyűjtés adatainak hitelesítése.
- Ipar 4.0 során keletkező adatok hitelesítése – megbízható minőségi és műszaki adatok/garancia.
- Logisztikai folyamatok, termék- és járműkövetés, átadás-átvétel, felhasználáshitelesítés.
- Élelmiszer-előállítás (vegyszerhasználat, származás) adatainak hitelesítése.
- Egészségügyi adatok, leletek, képi diagnosztika hitelesítése/manipulációmentesség igazolása.

- Laborfolyamatok (humán, állatorvosi, élelmiszer, ipar) minőségbiztosítása.
- Energetika: Smart Grid/okos mérők és elszámolások hitelesítése.
- Ügyvitel digitális aláírás nélkül: dokumentumok, fájlok tartalmi hitelesítése és továbbítása, átvételkövetés.
- Médiaanyagok (hangrögzítés, videó, kép, hírforrás stb.) hitelesítése.
- IT Security: LOG-hitelesítés, router konfiguráció ellenőrzés, szoftverkód-hitelesítés stb.
- Mesterségesintelligencia-betanító adatok hitelesítése.
- XML adatok, alkalmazások közti adatcsere érkezés- és fogadáshitelesítés.
- Adatalapú jogügyletek és szakigazgatási bevallások alapjának megteremtése.
- EU Green Deal/ESG célok adatalapú bevallása/ellenőrzése.

A fent felsoroltak csak példák az alkalmazhatóságra, ezen felül számos további üzleti és közigazgatási alkalmazási terület azonosítható és valósítható meg.

5.5. Adathitelesítési pilotprojektek rövid bemutatása

Az alábbiakban olyan projekteket szeretnék röviden bemutatni, amelyek mögött a DTS-koncepciót alkalmazó pilotmegoldás működik, vagy hamarosan működni fog.

5.5.1. Ipari termék minőségellenőrzése

A Széchenyi Egyetemmel karöltve egy egészségügyi implantátumokat gyártó cégnél egy gépilátás-alapú ipari termék minőségét ellenőrző rendszer került kidolgozásra. A termékek minőségét igazoló adatok hitelesítésére a DTS digitális ujjlenyomat-alapú megoldásának pilotváltozatát építettük be, amely az alábbi adatokat hitelesíti:

- Kamera által készített tárgykép (egyedi azonosítószám felismeréssel),
  - gyártási hely, operátor, anyag származása stb.
- Gépitanulás-alapú gépilátás-szoftvermegoldás verzió és minőségellenőrzési paraméterek.
- A gépilátás-megoldás által feldolgozott eredménykép, minőségi adatok,
  - nem megfelelő minőség esetén a javítás, korrekció utáni újabb adatok összefűzése.

A projekt eredménye: *a gyártó termékeinek minőségét a nemzetközi forgalmazók, az azt felhasználó kórházak, illetve maguk a betegek is leellenőrizhetik* egy, a termékről – akár egyszerű e-mailben megküldött – a minőségellenőrzés során készített fénykép és minőségi bizonyítvány alapján. Mivel ezen dokumentumok a DTS-sel hitelesítettek, azok digitális ujjlenyomatának ellenőrzésével azok eredetisége, manipulációmentessége könnyen leellenőrizhető.

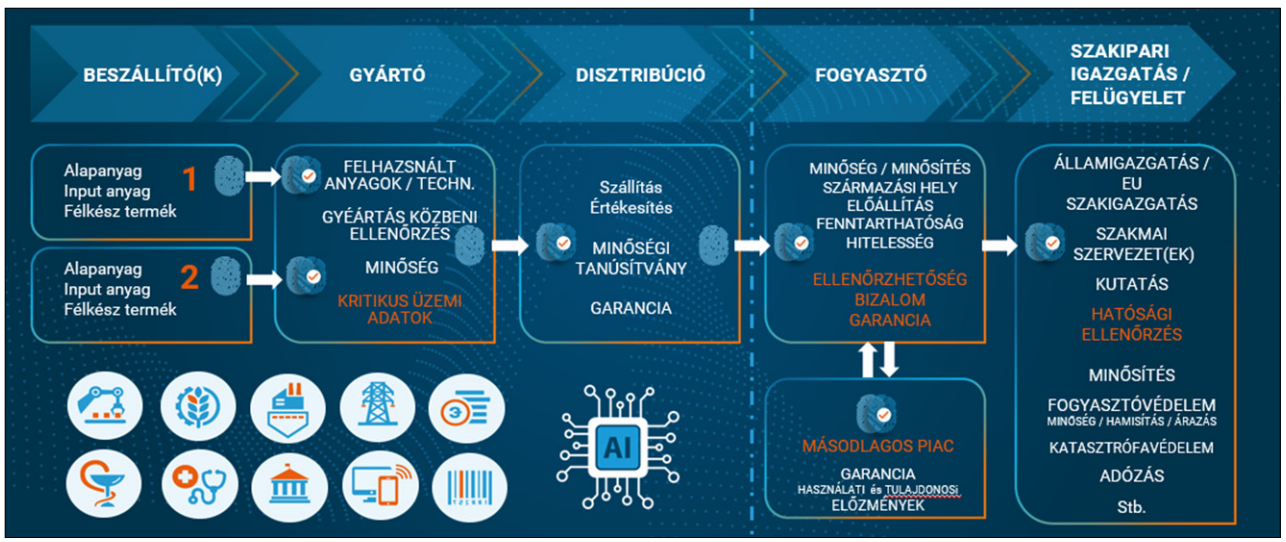
5.5.2. Ipari karbantartási és szerviztevékenység hitelesítése

A Széchenyi Egyetemmel együttműködve egy egészségügyi implantátumokat gyártó cégnél megvalósuló 3D/VR „Digital Twin" megoldás részeként elkészülő karbantartást támogató és tevékenységet naplózó rendszerbe integrálásra kerül a DTS digitálisujjlenyomat-alapú megoldásának pilotváltozata, amely az alábbi adatokat hitelesíti:

- Karbantartási/szerviztevékenység naplóbejegyzései:
  - munkavégző személy(ek),
  - tevékenység megnevezése, leírása,
  - felhasznált/kicserélt alkatrészek/kellékanyagok listája.
- Karbantartást/kellékanyagcserét igazoló fénykép, jegyzőkönyv, dokumentáció.

A projekt eredménye: *létrejön egy hiteles adatokra épített karbantartási és szerviztevékenységet naplózó megoldás.* Az ebben tárolt adatokban minden érdekelt szereplő, a gépszállítótól kezdve a karbantartásért felelős és a cégvezetők is megbízhatnak. A hiteles karbantartási/igazolt minőségű kellékanyag használati adatokra garanciaérvényesítési és üzemeltetésellenőrzési/karbantartási felhasználási esetek építhetőek.

Adathitelesség biztosítása a teljes ellátási láncban



5.5.3. Drónrepülési és permetezési feketedoboz (flight recorder)

A drónhasználat és a permetezés (azaz vegyszerkijuttatás) veszélyes üzem, azért a drón repülését és a permetlé elsodródását okozó meteorológiai körülmények hiteles igazolása a drónhasználat elterjedésével egyre fontosabb megoldandó kérdés.

A projekt során kidolgozott, a szél elsodródási hatását a drón útvonalának folyamatos módosításával kompenzáló Precíziós Drón Permetezési Platform háttérszolgáltatásaként beillesztettük a DTS pilotverzióját, amely *a drón repülési, telemetriai és permetezési munkavégzési (szemcseméret, nyomás, vegyszer stb.), valamint időjárási adatait digitális ujjlenyomat technikával hitelesíti.*

A projekt eredménye: a platform háttérszolgáltatásaként egy *drónrepülési feketedoboz jön létre,* amely segítségével vitás esetekben a drón útvonala és a kijutatott vegyszer, a repülés során mért meteorológiai adatok – és az abból következő vegyszer-elsodródás és -kompenzáció adatai – *hitelesen kinyerhetőek és vizualizálhatóak.* A pilotprojekt hosszú távon egy általános célú, minden – nem hobbi-, illetve afeletti kategóriájú – drón repülési adatát hitelesítő, a polgári repülésben használt megoldáshoz hasonlóan hiteles feketedoboz-megoldás megalapozását is hivatott előkészíteni.

5.5.4. #HASH – Fájlhitelesítő és hitelességellenőrző megoldás

#HASH munkanéven elkészítettünk egy webböngészőben elérhető pilotszolgáltatást, amelynek célja, hogy bármilyen fájl (word, xls, kép, videó stb.) egyszerű módon hitelesíthető legyen. A hitelesítő felhasználó által feltöltött fájlokról (a megadott metaadatokkal kiegészítve) a megoldásunk elkészíti annak „digitális ujjlenyomatát", és elküldi a DTS blockchain-hitelesítő rendszer számára.

A felhasználó a hitelesített fájlokat *bármilyen szokványos (email, felhő, social platformok, chat pendrive stb.) módon megoszthatja a partner(ei)vel.* Ezen partnerek a #HASH-szolgáltatás felületére hitelességellenőrző felhasználóként belépve, a megkapott fájlok feltöltésével, annak digitális ujjlenyomatának elkészítésével és blockchainben tárolt változatával történő összehasonlításával ellenőrizheti a fájlok hitelességét és változatlanságát.

Amennyiben a fájl hiteles, a rendszer a fájlhoz tartozó metaadatokat is kiírja az ellenőrző felhasználó számára, így az meggyőződhet róla, hogy a fájlt, dokumentumot mikor és ki hitelesítette, azóta változott-e annak verziója, illetve (például termék esetén) milyen további kapcsolódó (garancia, minőségellenőrzés stb.) dokumentumok állnak rendelkezésre.

A projekt eredménye: egy olyan webböngésző-alapú, a DTS működési elvéhez igazodó megoldás, amely fájltípustól függetlenül képes előállítani azok digitális ujjlenyomatát, ami alapján az adott fájl (annak ismételt feltöltésével) annak változatlansága és metaadatai (hitelesítő szervezet, felhasználó, időpont stb.) is ellenőrizhetővé válik. *A megoldás kizárólag csak a digitális ujjlenyomat előállításának erejéig futtatja át a fájlt a rendszeren, azaz semmilyen formában nem tárolja azt.*

5.5.5. Laborinformatikai munkafolyamat követés

Egy K+F projekt keretében kifejlesztett orvosi célú laborinformatikai szoftvermegoldás munkafolyamat- és adatkezelő rendszerében tárolt adatok hitelesítésére tervezzük a DTS-megoldásunkat beépíteni.

Ennek segítségével egy olyan egyedülállóan hiteles adatokra építő, a laborfolyamatokat hitelesítő megoldás jön létre, amely hitelesíti a következőket:

- A labor-mintavételezés helye, a mintavevő személye, végzettsége.
- A laborminta szállításának átadás-átvétele, szállítási körülmények.
- A labormintát feldolgozó laboráns felhasználó azonosítása, végzettsége.
- *Minden, a laborálás során történő mintatovábbítás, átadás-átvétel munkafolyamat és személy hitelesítése.*
- Laboráló gép működési adatai: kalibrálás, kötelező karbantartás, reagenshasználat.
- Laborkiértékelést hitelesítő azonosítása, végzettsége.
- Kiállított laboreredmény-dokumentum hitelesítése,
  – a hitelesítéshez kapcsolódóan a háttérfolyamatok hitelessége is igazolható.

Elvárt eredmény: A laborálás, az ahhoz használt gépi erőforrások pontossága, hitelessége, az emberi erőforrások szakértelme és az annak eredményeképpen álló lelet hitelessége igazolható a teljes labor mentén. A felhasználó, a folyamatban érintett személyek azonosságát és megfelelő diplomával való rendelkezését hosszabb távon az EBSI (European Blockchain Sevices Infrastructure) megvalósító megoldásai által szeretnénk ellenőrizhetővé és igazolhatóvá tenni.

6. Összefoglalás

A digitális adatalapú gazdaság és közigazgatás hatalmas előnyt nyújt az üzleti és a civil szereplők számára is, ugyanakkor a nem hiteles adatokra épített adatalapú együttműködések, adatpiacok, szakértői, illetve mesterséges intelligencia- és gépi tanulási megoldások kritikus biztonsági kockázatot jelentenek a teljes társadalomra nézve.

Ennek kivédésére a digitális gazdaság működésének alapját *az ADAT-ot a keletkezési helyéhez lehető legközelebb hitelesíteni kell,* amihez létre kell hozni egy olyan infrastruktúrát, amely iparágtól és cég méretétől függetlenül elérhetővé teszi az adatok tömeges hitelesítését, valamint annak interoperábilis ellenőrzésének lehetőségét.

Az adathitelesítésre a blockchain-technológia alkalmazása olyan megoldást kínál, ahol a hitelességet nem a jogszabályok betartása, vagy a rendszergazdákban való bizalom, hanem maga a technológia és az abban tárolt adatok megváltoztathatatlansága garantálja.

A cikkben bemutatott Digital Trust Services (DTS) megoldás többcsatornás hitelesítő csomópont felépítését biztosítja, hogy az adathitelességet olyan szereplők konszenzusa biztosítsa, akikben az adathitelesítők maguk megbíznak. A hitelesítésben szerepet vállalhat a szak- és államigazgatás is, annak érdekében, hogy az adat az adatalapú kormányzás során is felhasználható legyen.

A hiteles, ellenőrizhető adat létrejötte erősíti a bizalmat a teljes ellátási láncban, a vásárlókban és az igazgatási szereplőkben, míg annak ZERO TURST-megközelítéses alkalmazása jelentősen megnehezíti a hackerek és az adatmanipuláció lehetőségeit.

Nagyon fontos, hogy *a cégek a lehető leghamarabb elkezdjék beépíteni adatgyűjtési és feldolgozási megoldásaikba az ADATHITELESSÉGET,* mert ma még ez csak extra igénynek tűnik, de hamarosan már alapelvárás lesz a fogyasztóktól, és akinek nem lesz hiteles adata a termékéről, illetve annak előállításáról, az lemarad a versenyben a nemzetközi cégek mögött! *Ez különösen fontos azon cégek esetében, akiknek a terméke valamilyen digitálisan tárolt adat (pl. CT-felvétel), vagy AI/Ml-megoldás,* mert ha nem tudja igazolni ezek hátterében lévő adatok hitelességét és manipulációmentességét, akkor rövidesen nem fogja tudni értékesíteni a terméket, illetve megoldást.

Ha megteremtjük a megbízható adat fogalmát, tovább léphetünk a blockchain következő lehetőségének kiaknázásához: az aláírás nélkül is hiteles, automatikusan teljesülő *Digitális Szerződések (Smart Contract) és Token gazdaság* irányába, ugyanakkor az ebben rejlő lehetőségek kiaknázásához a jogszabályi háttér hazai kidolgozása is szükséges.

A DTS-projekt az Európai Unió IPCEI-CIS programjában megvalósítani kívánt rendszerként jelentős szerepet kíván vállalni a hiteles adatokra épített Európai Cloud ökoszisztéma hiteles adatokra való építésében és működésében.

Hivatkozások

[1] A. Chadha, V. Kumar, S. Kashyap, and M. Gupta, "Deepfake: An Overview", in Proc. of 2nd International Conf. on Computing, Communications and Cyber-Security, P.K. Singh, S.T. Wierzchoń, S. Tanwar, M. Ganzha, and J.J.P.C. Rodrigues, Ed., in Lecture Notes in Networks and Systems. Singapore: Springer, pp.557–566., 2021. doi: 10.1007/978-981-16-0733-2_39.

[2] P.B.K. Payne and P.H. Wu, ICCWS 2020 15th International Conference on Cyber Warfare and Security. Academic Conferences and publishing limited, 2020.

[3] European Commission, "Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL amending Regulation (EU) No 910/2014 as regards establishing a framework for a European Digital Identity", *https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52021PC0281&from=EN*

[4] M. Lukinovic and L. Jovanovic, "Greenwashing – fake green/environmental marketing", Fundamental and applied researches in practice of leading scientific schools, Vol. 33, pp.15–17., 2019. doi: 10.33531/farplss.2019.3.04.

[5] A. Castillo és A. Thierer, "Economic Perspectives Projecting the Growth and Economic Impact of the Internet os Things", 2015. doi: 10.2139/ssrn.2618794.

[6] M. Sain, Y.J. Kang, and H.J. Lee, "Survey on security in Internet of Things: State of the art and challenges", in 19th International Conf. on Advanced Communication Technology (ICACT), pp.699–704., 2017. doi: 10.23919/ICACT.2017.7890183.

[7] S. Karnouskos, "Artificial Intelligence in Digital Media: The Era of Deepfakes", IEEE Transactions on Technology and Society, Vol. 1., nr.3, pp.138–147., 2020. doi: 10.1109/TTS.2020.3001312.

[8] S. Rose, O. Borchert, S. Mitchell, and S. Connelly, "Zero Trust Architecture", National Institute of Standards and Technology (NIST), Special Publication, 800-207, 2020. doi: 10.6028/NIST.SP.800-207.

[9] Z. Ghahramani, "Probabilistic machine learning and artificial intelligence", Nature, Vol. 521., nr.7553., 2015. doi: 10.1038/nature14541.

[10] M.A. Ramirez et al., "Poisoning Attacks and Defenses on Artificial Intelligence: A Survey", arXiv, 2022. doi: 10.48550/arXiv.2202.10276.

[11] J. Katz, "Digital Signatures", Boston, MA: Springer US, 2010. doi: 10.1007/978-0-387-27712-7.

[12] "The SSL Protocol". *http://www.webstart.com/jed/papers/HRM/references/ssl.html*

[13] B. Preneel, "Cryptographic hash functions", European Transactions on Telecommunications, Vol. 5, nr.4., pp.431–448., 1994. doi: 10.1002/ett.4460050406.

[14] Pharmaledger, online: *https://pharmaledger.eu/*

[15] D.G. Wood, "Ethereum: A Secure Decentralised Generalised Transaction Ledger", *https://files.gitter.im/ethereum/yellowpaper/VIyt/Paper.pdf*

[16] M. Rauchs et al., "Distributed Ledger Technology Systems: A Conceptual Framework". Rochester, NY, 2018. doi: 10.2139/ssrn.3230013.

A szerzőről

SOLYMOS GYULA villamosmérnök, vállalkozás és minőségbiztosítási szakdiplomás menedzser. Több mint 20 éve dolgozik az IT világában, ahol széleskörű tapasztalatot szerzett a modern technológiákra épülő innovatív megoldások fejlesztésében. 2019-ben csatlakozott a 4iG Nyrt.-hez, ahol mesterségesintelligencia-, 5G-, és blockchain-technológia-alapú megoldásokkal és üzletfejlesztési projektekkel foglalkozik. Fő célja, hogy az innovatív ötletekből valós üzleti problémákra választ adó megoldásokat kovácsoljon. Közel 3 éve foglalkozik a blockchain-technológiával, amely alkalmazásáról már számtalan előadást tartott. A 2022 márciusában megalakult a Blockchain Koalíció Ellátási Lánc munkacsoportjának vezetőjeként célja, hogy a technológia szerepét minél több iparágban megtalálja és ezáltal a hazai ökoszisztéma kialakulását és új alkalmazások megvalósulását katalizálja. A DTS platform ötlet- és projektgazdája, amit az Európai Unió számára is bemutatott és jelenleg annak Important Projects of Common European Interest (IPCEI-CIS) program keretében történő megvalósításának előkészítésén dolgozik.

Neumann Jánostól a HPC-ig

MÁRAY TAMÁS, SZEBERÉNYI IMRE
Budapesti Műszaki és Gazdaságtudományi Egyetem
tamasmaray@gmail.com, szeberenyi.imre@vik.bme.hu

Kulcsszavak: szuperszámítógép, szuperszámítástechnika, HPC, felhő, EuroHPC, PRACE, szupergépek fejlődési trendjei, Komondor

Mostanában gyakran hallunk a szuperszámítástechnikáról. Szuperszámítógépek azonban nagyon régóta léteznek. Történetük már a kezdetektől fogva követi a számítástechnika fejlődését.
De mi különbözteti meg a szuperszámítógépeket a hagyományos számítógépektől? És a „felhőtől"? Miért van szükség a használatukhoz speciális ismeretekre? És miért kell mostanában beszélnünk róluk? Mi az oka annak, hogy az utóbbi években világszerte mindenütt előtérbe kerültek és fejlődésük, térhódításuk a különféle válságok és nehézségek közepette is töretlen, sőt gyorsulást mutat? Ezeket a kérdéseket feszegeti cikkünk, miközben áttekintjük a szuperszámítógépek legfontosabb jellemzőit, speciális műszaki megoldásait, a jellegzetes felhasználási módokat és alkalmazási területeiket.
Szó esik természetesen a nemzetközi trendekről, a kapcsolódó észak-amerikai, ázsiai és európai stratégiákról.
Végül egy pillantást vetünk arra, hogy Magyarország hol áll ezen a területen – milyen eredmények és milyen hiányosságok látszanak –, és a fejlődés iránya merre mutat.

1. Bevezetés

Az informatika felhasználása mindennapjaink részévé vált. Közhelynek számít, hogy informatika nélkül már el sem tudnánk képzelni életünket. A számítógép, a tablet, a mobiltelefon használata teljesen hétköznapi. De mi az a HPC? Mik azok a szuperszámítógépek és miért kellenek, kik használják? Mitől „szuperek"? Mi a kapcsolatuk a felhőkkel és gridekkel? Cikkünk alapvetően ezekre a kérdésekre próbál választ adni, röviden áttekintve a számítógépek fejlődését, a trendeket és a fejlődés hajtóerőit Neumann korától napjainkig. Kezdjük rövid definíciószerű válaszokkal, melyeket később bővebben kifejtünk.

HPC (high-performance computing): A nagy teljesítményű számítógépeket használó/gyártó iparág, szuperszámítástechnika, amivel nagy és összetett feladatokat lehet megoldani rövid idő alatt.

Szupergép vagy nagy teljesítményű számítógép: Olyan számítógép (rendszer), ami a komplex feladatokat az adott technológiai szinten az átlagtól sokkal gyorsabban – az eredmény szempontjából releváns idő alatt – képes megoldani.

Felhő: Számítógépes erőforrások (szerver, tároló, hálózat stb.) és szolgáltatások igény szerinti, kényelmes, helyfüggetlen elérése és önkiszolgáló létrehozása oly módon, hogy a szolgáltatónak az csak minimális erőfeszítésébe kerül. Van olyan felhő- vagy webszolgáltató, amely szupergépet használ a szolgáltatásához. De van olyan felhő is, amely alkalmas bizonyos szupergépes feladatok megoldására is.

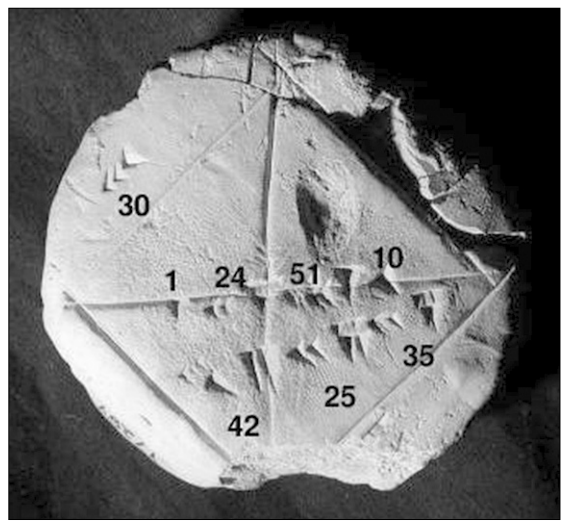
Grid: Alapkoncepciója azonos a felhővel. A megvalósítás és az üzleti modell merőben más. Történetileg az akadémiai szférában alakult ki és a felhő elődjének tekinthető.

2. Történeti áttekintés

A számítógépek létrejöttét alapvetően az alkalmazott matematika inspirálta. A numerikus módszerek és azok alkalmazása (pl. terület és űrtartalom kiszámítása) több ezer éve foglalkoztatják az emberiséget, amit olyan tárgyi bizonyítékok is bizonyítanak, mint a Yale Egyetem babilóniai gyűjteményében található agyagtáblák, melynek egyikén a √2 értékének becslésére látunk módszert *(1. ábra)*.

Az első elektronikus számítógépek felhasználási területe a numerikus analízis, modellezés és a rejtjelezés volt. De mi inspirálta, illetve mi inspirálja a mai napig a mérnököket az egyre nagyobb teljesítményű számítógé-

1.ábra Babilóniai agyagtábla √2 becslésére (i.e.1800) (Forrás: Bill Casselman, upload.wikimedia.org/wikipedia/commons/0/0b/Ybc7289-bw.jpg)



pek készítésére? A válasz egyszerű: a nagyobb, részletesebb felbontás, pontosabb és gyorsabb eredmények igénye. A nagyméretű számítási problémák megjelenése már az elektronikus számítógépek megszületését is megelőzte: 1922-ben Lewis Fry Richardson olyan numerikus módszert publikált [1], ami a különböző meteorológiai paraméterek (hőmérséklet, szélerősség stb.) figyelembevételével alkalmas globális időjárás-előrejelzésre. A probléma mérete miatt keletkező nagymennyiségű számítási feladat elvégzésére Richardson becslése szerint 64 ezer ember folyamatos munkájára (a számítások kézi elvégzésére) lett volna szükség. Richardson numerikus módszere a saját korában megfelelő számítógép nélkül kivitelezhetetlen volt, de ötletével nem csak a mai meteorológiai számítási módszerek alapját teremtette meg, hanem felvázolta a számítási teljesítmény növelésének egyik módszerét is: a párhuzamosítást.

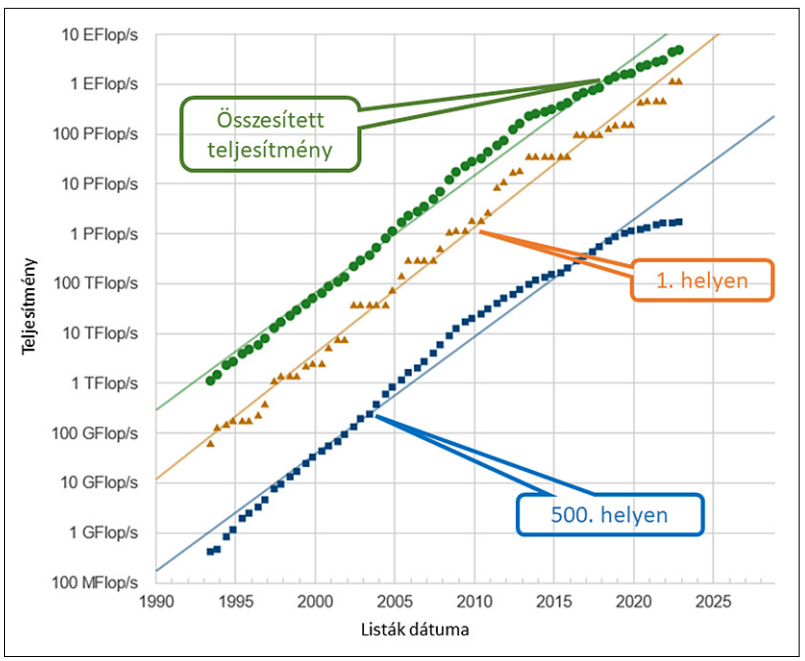
3. Számítási teljesítmény növelése

Az elektronikus számítógépek elemi matematikai és tárolási műveletek sorozatát végzik a másodperc töredéke alatt. A feladatok nagyságának, bonyolultságának növekedése egyre gyorsabb gépek létrehozását igényli. Ez a kielégíthetetlen igény a tudományos számítások körében a korai számítógépektől napjainkig jelen van. A teljesítménynövelés egyik útja a magasabb órajelfrekvencia alkalmazása, a másik pedig a feladatok párhuzamosítása. Az órajelfrekvencia azonban nem növelhető határtalanul a fizika törvényei miatt, ezért ma már szinte kizárólag a párhuzamosítás a járható út.

Egy anekdota szerint Daniel Slotnick a párhuzamosítás ötletét Neumann Jánosnak is felvetette 1952-ben, aki az ötletre szimplán csak ennyit mondott: „too many tubes” (túl sok elektroncső) [2]. Az elektroncsöves számítógépek korában ugyanis a csövek meghibásodási gyakorisága határozta meg alapvetően az üzembiztonságot. Neumann válaszával arra utalt, hogy ha ezek számát növelik, akkor a meghibásodási gyakoriság elviselhetetlenül nagy lesz. Megjegyezzük, hogy az első igazi masszívan párhuzamos számítógép, az ILLIAC IV Daniel Slotnick ötlete alapján épült meg tranzisztoros technológiával 1970-ben, de beüzemeléséhez 1973-ig kellett várni.

A tárolási kapacitás mérése/kifejezése már beépült napjainkba: a tárolt adatok mennyiségét mérjük. Hogyan fejezhető ki a teljesítmény? Definíció szerint az adott időegység alatt végrehajtott (munkát) műveletek számát mérjük. Abban az esetben, ha az egy másodperc alatt elvégzett műveletek számára (instruction per second) koncentrálunk, akkor kIPS, MIPS, ... mértékegységet használunk, abban az esetben, ha valós számokkal végrehajtott matematikai műveletek számát (floating point operations per second) figyeljük, akkor kFlop/s, MFlop/s, ... mértértékegységet használunk. Mindkét adat meghatározható méréssel, azaz benchmark-programokkal, illetve elméleti úton a gép órajelének és felépítésének ismeretében.

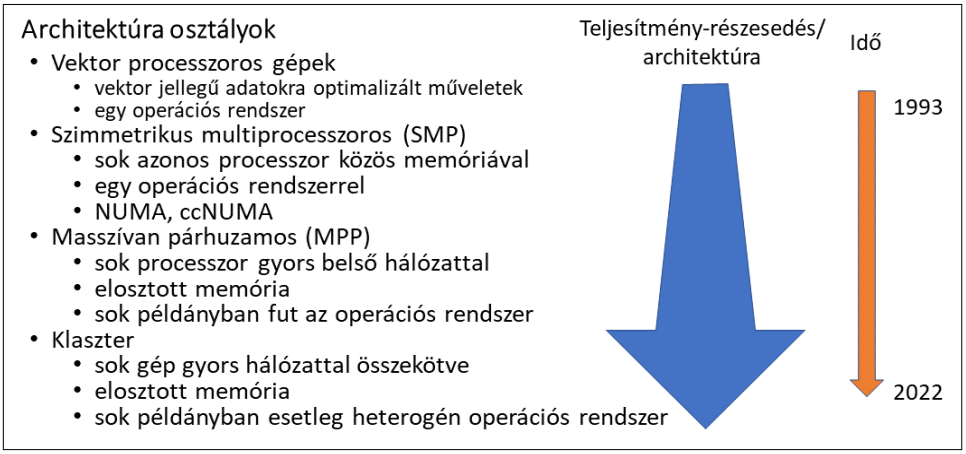
2. ábra
Teljesítménynövekedés időbeli alakulása és trendje a TOP 500-as listán (Forrás: szerző, TOP 500 alapján)



4. Szuperszámítógépek

Ahogy a bevezetőben említettük, az elnevezés nem időtálló és nem egy abszolút skálán értendő. A mindenkori szokásos számítási teljesítménynél nagyobb számítási és/vagy tárolási kapacitással rendelkező számítógépet, vagy számítógéprendszert nevezzük így. Ha megnézzük 90-es évek szuperszámítógépeinek teljesítményét/kapacitását, akkor azt tapasztaljuk, hogy a zsebünkben levő mobiltelefon bőven lekörözi azt. A nagyvilágban üzembe helyezett nagy teljesítményű számítógépek rangsorolására 1993-ban létrejött a *top500.org* lista, ami önkéntes alapon tárolja, rangsorolja és publikálja az 500 legnagyobb teljesítményű gép főbb jellemzőit, mint például az architektúrát, felhasználásának jellegét, operációs rendszerét, benchmark programmal mért teljesítményét (R_{max}), valamint az elméleti teljesítményét (R_{peak}). A 2022 novemberében publikált adatok szerint az első helyezett gép teljesítménye 1,1 EFlop/s ($1{,}1*10^{18}$ Flop/s), az 500. helyen pedig egy 1,7 PFlop/s teljesítményű gép áll. Beszédes adat, hogy 1993-ban az első helyezett számítógép teljesítménye nem érte el a 60 GFlop/s értéket.

A top 500-as lista számos lekérdezési és keresési lehetősége közül talán a legérdekesebb a teljesítmények időbeli alakulását mutató grafikon (statistics/perfdevel) [3], melyről az olvasható le, hogy a teljesítménynövekedés töretlen *(2. ábra)*. A növekedési ütemre jellemző, hogy négyévenként a tudományos számítógépek teljesítménye majdnem 14-szeresére nő.



3. ábra
Jellemző architektúraosztályok a top500-as listán.
(Forrás: szerző)

5. A fejlődés háttere

Mi áll a fejlődés hátterében? Alapvetően azok a technológiai megoldások, melyek egyre nagyobb elemsűrűséget tesznek lehetővé a félvezetőgyártásban. Míg 1971-ben 10 µm volt a sorozatgyártásban elérhető sűrűség, addig ma már ez 3 nm. Ezzel együtt a hálózati technológia és a memóriák írási/olvasási sebessége is rohamléptekben fejlődik. Mindezek lehetővé tették a speciális gyorsító eszközök (FPGA, Phi, GPGPU) megjelenését és fejlődését is. A top500-as listát elemezve észrevehető, hogy a szupergépek architektúraosztályai mára jelentősen eltolódtak a klaszterarchitektúra felé, ami a jobb skálázhatósággal magyarázható. A *3. ábra* röviden összefoglalja a különböző architektúraosztályok jellemzőit.

6. Mire használják?

A szuperszámítógépeket leginkább a tudományos számítások és kutatás céljaira használják, de jelentős szerepük van a nagy és komplex rendszerek modellezésében. A komplex, sok objektumból álló, mint az időjárási, geodéziai, űrkutatási területen használt modellek pontosabb, részletgazdagabb feldolgozásához nagy számítási és tárolási kapacitásra van szükség. Az anyagkutatás ma már molekula/atomi szinten modellezi az anyagot ahhoz, hogy pontosabb képet kapjon annak jellemzőiről. A gyógyszerkutatásban például tipikus feladat az úgynevezett fehérjedokkolási probléma megoldása, ami szintén számításigényes. A különböző szimulációk is nagy erőforrásokat igényelnek; a gyártás és tervezés során, a részecskefizikai kutatásban, vagy éppen az univerzum kutatásában. Fontos megemlíteni, hogy a különböző IT-szolgáltatások megvalósításához is nagy erőforrásra van szükség, így gyakran ezek hátterében is szupergép áll (pl. Facebook, DeepL).

7. Globális verseny

Bemutattuk, hogy a szuperszámítógépek története tulajdonképpen egyidős az elektronikus számítógépek történetével. A szuperszámítástechnika, mint a számítástechnika egy speciális részterülete, mégis az utóbbi két-három évtizedben kapott egészen kiemelt figyelmet, és ezzel párhuzamosan a jelentősége ugrásszerűen megnőtt. Számos oka van ennek, az adatforradalomtól kezdve a digitalizáció terjedésén át a tudományos módszerek, valamint high-tech iparágak szédületes fejlődéséig. Mindenesetre az széles körben nyilvánvalóvá vált, hogy az innovációs és versenyképesség szintje és a szuperszámítástechnika alkalmazásának mértéke között direkt összefüggés van. Nagyon sok fejlett gazdaságban felismerték ezt, és ennek nyomán megindult az egyre gyorsuló verseny a világ vezető országai között a szuperszámítógépek fejlesztése és alkalmazása terén. *„To compete, you must compute!”* hangzik a jól ismert szlogen, melynek érvényessége ma már nem kérdés.

Bár a szuperszámítógépek mind a mai napig elsősorban nélkülözhetetlen kutatási infrastruktúrát jelentenek, ma már a legkülönfélébb ipari és gazdasági termelőfolyamatok és szolgáltatások produkciós fázisaiban is jelen vannak. Nemzetközileg is versenyképes tudományos/kutatási tevékenység a legtöbb területen alig képzelhető el szuperszámítógépes háttér nélkül, hiszen az egyre pontosabb és összetettebb szimulációk, adatbányászat vagy akár mesterséges intelligencia alapú eljárások más módon nem kezelhetők, főképp nem gyorsan és költséghatékony módon. A szuperszámítógépek segítségével számtalan olyan kérdés vizsgálható – elsősorban a tudomány területén – amelyre más módon nem lenne lehetőség. Természetesen, aki a legújabb tudományos eredményeket birtokolja, az az innovációt és fejlődést is meghatározza és így versenyelőnyre tesz szert. Ezért olyan komoly az erőfeszítés a világ összes fejlett régiójában a szuperszámítástechnika területén.

7.1. Egyesült Államok

A vezető szerep hagyományosan az Amerikai Egyesült Államoké, amely több száz szuperszámítógéppel, a világ teljes HPC-kapacitásának közel a felével rendelkezik. Az USA a hardver- és szoftvertechnológia fejlesztésében, a piacvezető HPC-gyártók tekintetében, a kiépített infrastruktúrában és a HPC alkalmazásának mértékében is az élen jár. A Fehér Ház által még 2015-ben elfogadott NSCI (National Strategic Computing Initiative) [4, 5] új lendületet adott a fejlődésnek, és kitűzte az ak-

kor még szinte utópisztikusnak tűnő célt: az exaflop határ áttörését. (1 Eflop/s = 10^{18} matematikai művelet másodpercenként.) Nehéz felfogni, hogy ez milyen óriási szám. Akkora, hogy a Földön élő 8 milliárd ember mindegyikére másodpercenként több, mint 137 millió számítási művelet jut. A legnagyobb technológiai kihívást az jelentette, hogyan lehet a fajlagos energiaigényt olyan mértékben leszorítani, hogy még egy ekkora gép is táplálható és hűthető maradjon. Bár az üzemterv szerint az exaflop-határt 2020-ra kellett volna elérni, végül csak 2022-re sikerült, amikor üzembe helyezték az Oak Ridge National Laboratory-ban a világ első 1 EFlop/s feletti teljesítményű szuperszámítógépét, a *Frontier*-t, ami Cray-technológiát használ és közvetlen folyadékhűtésű *(4. ábra)*. A gép legfontosabb paraméterei az *1. táblázatban* láthatók.

A Frontier az amerikai energiaügyi tárca egyik kiemelt kutatóintézetében kapott helyet, de felhasználása széleskörű és általános célú, a tudományos kutatás minden területét támogatja (anyagtudomány, űrkutatás, biotechnológia, klímakutatás, szoftvertechnológia, mesterséges intelligencia, lézerfizika, plazmafizika stb.)

A Frontier teljesítménye elképesztő, de ez sem elegendő a tudományos kutatásoknak. Így Amerikában hamarosan (még idén) átadásra kerül további két, még nagyobb szuperszámítógép: az *Aurora* és az *El Capitan,* melyek külön-külön is meg fogják haladni a 2 EFlop/s teljesítményt. Mindeközben a fejlesztés és a gyártás erősen konszolidálódott, a gyártók száma lecsökkent. Az IBM kiszállt a hagyományos szuperszámítógépek piacáról, az SGI és a Cray a HPE-be olvadt.

7.2. Ázsia

Ázsia két vezető szuperszámítógép hatalma Japán és Kína. Természetesen más országokban (pl. Dél-Korea vagy India) is gyors a fejlődés, de a globális versenyben e két állam a meghatározó. Japánban nagy a hagyománya a szuperszámítógépek fejlesztésének és használatának, már többször is elfoglalták a TOP500-lista első helyezését saját gyártmányú rendszereikkel. Jelenleg a vezető japán szuperszámítógép, a Fujitsu által épített *Fugaku* (442 PFlop/s) második helyezett a listán, de mögötte további 30 japán szuperszámítógép is található. A Fujitsu mellett az NEC is jelentős gyártónak számít.

Kína csak az elmúlt két évtizedben kapcsolódott a versenybe, de a fejlődés hihetetlenül erős. Érdekesség: ugyanabban az évben (2000-ben) került Kína először a listára (egy IBM géppel), amikor Magyarország is. Mára Kína birtokolja a TOP500-as szuperszámítógépek számának 1/3-át (többet, mint az USA), ami az összes kapacitás 10%-át teszi ki. Kína nemcsak alkalmazza a HPC-technológiát, hanem fejleszti is, így a kínai szuperszámítógép-gyártók (Inspur, Sugon, Lenovo, Huawei) az ameri-

1. táblázat
A Frontier technikai paraméterei

Elméleti teljesítmény	R_{peak} > 1,5 EFlop/s
Mért teljesítmény	R_{max} = 1,102 Eflop/s
CPU	9472 db AMD Epyc Milan (606,708 core)
GPU	37888 db Radeon Instinct MI250X (8,335,360 core)
Belső hálózat	Slingshot
Memória	9,2 Pbyte
Háttértár	700 Pbyte
Operációs rendszer	HPE Cray OS (Linux)
Fizikai helyszükséglet	680 m^2
Energiaigény	21 MW (Green500 no.1)

4. ábra Frontier
(Forrás: Oak Ridge National Laboratory, U.S. Dept. of Energy)

kaiak legerősebb versenytársai. Legnagyobb szuperszámítógépe, a *Sunway TaihuLight* a legfrissebb TOP500-listán jelenleg a 7. helyet foglalja el a 93 PFlop/s teljesítménnyel. A gyártás évében (2016), majd azt követően másfél évig az első helyen állt. Érdekessége, hogy kínai fejlesztésű processzorra épül. Operációs rendszere Linux-alapú. Természetesen Kína is törekszik az exafloptartomány elérésére, és várható, hogy ez hamarosan meg is történik.

7.3. Európa

Nem meglepő módon az európai országok között is a gazdaságilag legerősebbek a legaktívabbak a szuperszámítástechnika alkalmazásában. Németország, az Egyesült Királyság, Franciaország, Olaszország, Spanyolország, Svájc vezetik a listát. Bár az 1980-90-es években Európa nem állt rosszul a globális versenyben, az ezredforduló után azonban fokozatosan lemaradt. Míg Észak-Amerikában és Ázsiában idejekorán felismerték a szuperszámítástechnika növekvő fontosságát, mindez Európában csak késve következett be, ráadásul az európai országok külön-külön akkora invesztíciót sem tudtak végrehajtani.

Az EU az első jelentősebb, európai összefogásra építő HPC-programot (DEISA – Distributed European Infrastructures for Supercomputing Applications) 2005-ben indította, majd ezt követte 2010-től 25 ország együttműködésével a PRACE (Partnership for Advanced Computing in Europe) [6], amelyekhez kapcsolódva számos HPC-vonatkozású nemzetközi K+F-projekt is társult. A programok és projektek sikeresek voltak, – sok eredményt hoztak és megállították Európa lemaradásának növekedését –, azonban az egyre gyorsuló globális versenytársakhoz való felzárkózáshoz nem bizonyultak elegendőnek.

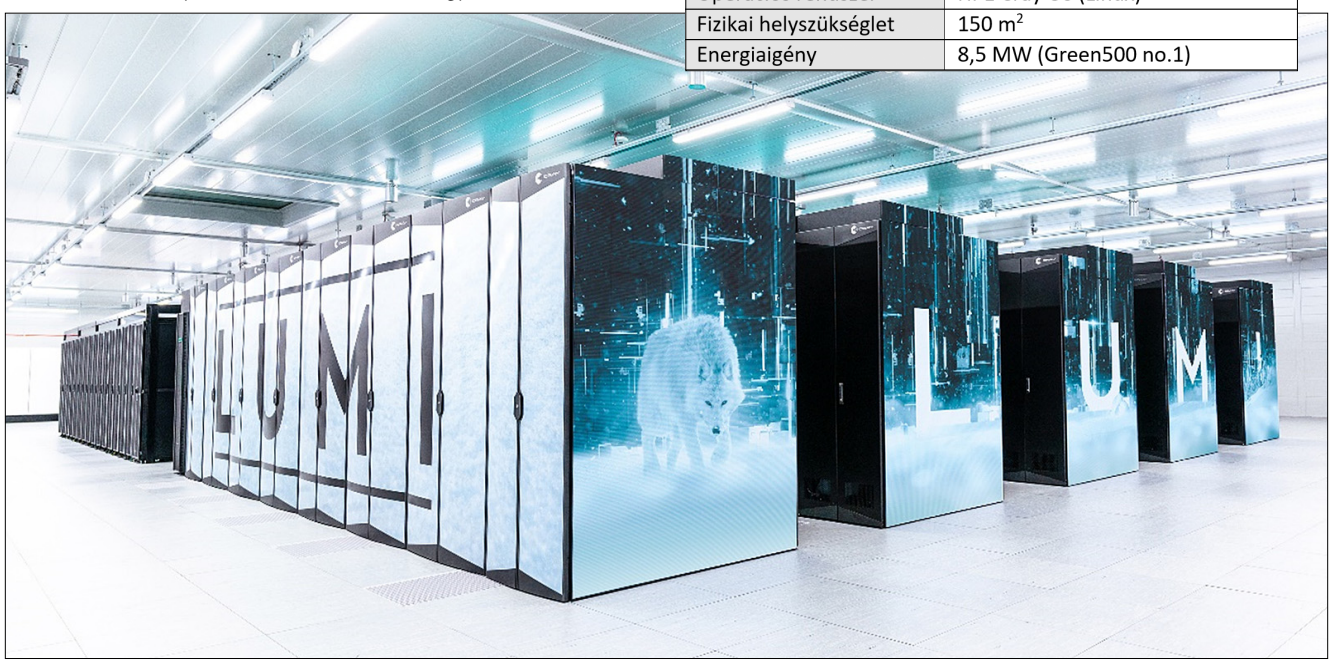
A tanulságokat leszűrve az EU 2019-ben stratégiai fontosságú fejlesztési területnek nyilvánította a szuperszámítástechnikát [7] és 1 milliárd euro induló forrással létrehozta az EuroHPC Joint Undertaking (EuroHPC Közös Vállalkozás) nevű programot [8]. Ez deklaráltan egy felzárkóztatási program, széles körű európai összefogással, a tagországok saját forrásainak mobilizálásával, a teljes HPC-ökoszisztéma felölelésével (infrastruktúra beruházások, hardver- és szoftverfejlesztés, alkalmazások, szolgáltatások, képzés), az európai HPC-"ipar" megerősítésével. Az EuroHPC is az exaflop tartomány elérését tűzte ki célul, de az ehhez vezető úton először petaflop és pre-exa kategóriájú gépeket épít. Az EuroHPC-hez eddig 28 ország csatlakozott, köztük Magyarország is.

A program első petaflop kategóriájú szuperszámítógépei Szlovéniában, Luxemburgban, Csehországban, Bulgáriában és Portugáliában épültek. E gépek teljesítménye 4,5-13 Pflop/s közé esik, így a TOP500 felső ötödében találhatók. Az első – és egyben legnagyobb – EuroHPC pre-exa kategóriájú szuperszámítógép a *Lumi,* Finnországban épült meg tíz ország összefogásával. A Lumit 2022-ben adták át, és jelenleg a világ 3. legerősebb szuperszámítógépe az aktuális 308 PFlop/s teljesítményével, amely végső konfigurációjában el fogja érni a 375 Pfops/s-ot *(5. ábra)*. A Lumi szintén HPE/Cray gyártmányú, technológiája szinte teljes egészében megegyezik a Frontierével. A gép messze északon, olyan környezetben került telepítésre, ahol egész évben megújuló energiával és természetes hűtéssel (free cooling) működtethető. A Lumi paramétereit a *2. táblázat* tartalmazza.

2. táblázat
A Lumi technikai paraméterei

Elméleti teljesítmény	R_{peak} > 420 PFlop/s
Mért teljesítmény	R_{max} = 309,1 PFlop/s
CPU	5632 db AMD Epyc
GPU	10240 db Radeon Instinct MI250X
Belső hálózat	Slingshot
Háttértár	117 Pbyte
Operációs rendszer	HPE Cray OS (Linux)
Fizikai helyszükséglet	150 m^2
Energiaigény	8,5 MW (Green500 no.1)

5. ábra A Lumi (Forrás: CSC, Finnország)



A Lumit követően még 2022-ben átadásra került az olaszországi Bolognában az EuroHPC második pre-exa szuperszámítógépe is, a 174 PFlop/s teljesítményű *Leonardo*. A gépet az európai Atos/Bull cég építette, Intel-processzorokat és Nvidia GPU-gyorsítókártyákat tartalmaz. A Leonardo jelenleg a világ 4. legerősebb szuperszámítógépe. A kihasználására létrejött hat országból álló konzorciumnak Magyarország is tagja. A Leonardo után a harmadik európai pre-exa szuperszámítógép a várhatóan 2023-ban Barcelonában épülő *Mare Nostrum 5* lesz.

Az új tervezési ciklusban már 7 Mrd euro forrásból gazdálkodó program időközben az első európai exaflop kategóriájú szuperszámítógép *(Jupiter)* megépítését is bejelentette, amely Németországban a Jülichi Szuperszámítógép Központban (JSC – Jülich Supercomputer Centre) valósul meg 2024-re. Emellett további legalább 20 PFlop/s teljesítményű, kisebb EuroHPC-rendszerek épülnek Görögországban, Írországban és Lengyelországban is.

A fejlődés tehát Európában is rendkívüli módon felgyorsult, és ez az előretörés a TOP500-listán is szembeötlő (hozzávetőlegesen 12% kapacitást képvisel). Fontos kiemelni, hogy az EuroHPC nemcsak az infrastruktúra fejlődésében hoz látványos eredményeket, hanem a szuperszámítógépek felhasználása, az alkalmazás- és algoritmusfejlesztések, illetve a szakemberképzés területén is jelentős az előrelépés. Külön figyelmet érdemel, hogy Európa saját processzor fejlesztésébe is kezdett, hogy csökkentse technológiai függőségét. A 2015-ben indult EPI (European Processor Initiative) elsődleges célkitűzése, hogy olyan processzort fejlesszen ki, amely a jövő európai számítógépeinek motorja lehet. A fejlesztés ARM- és RISC-alapokra épít, és a megvalósításra létrehozott vállalkozás (SiPearl) éppen 2023-ra ígéri az első működőképes változat megjelenését.

6. ábra Az első hazai szuperszámítógép (forrás: a szerző)

7.4. Magyarország

Magyarországon az első valódi, tudományos célú szuperszámítógépet 2001-ben adták át az NIIF (Nemzeti Információs Infrastruktúra Fejlesztési) program keretében *(6. ábra)*. 60 GFlop/s teljesítményével rögtön a TOP500-listára is felkerült, és hamar nagyon népszerűvé vált a kutatók között. A következő 15 évben az infrastruktúra az NIIF-Program keretében több lépcsőben megújult és technológiában, kapacitásban egyaránt sokat fejlődött. A gépet és utódait az egyetemeken és akadémiai intézetekben működő több száz tudományos kutatási projektben használták kiemelkedő sikerrel, nemzetközileg elismert eredményeket produkálva.

2015-ben az NIIF már 8 szuperszámítógépet működtetett közel 0,5 PFlop/s aggregált kapacitással, köztük az első GPU-gyorsítókkal ellátott *Leo* nevű gépet, amely a Debreceni Egyetem campusán direkt a számára épített új, korszerű HPC-központban kapott elhelyezést és 2022 év végéig a legnagyobb hazai szuperszámítógép volt.

Az NIIF Intézet jogutódja 2016-tól a KIFÜ (Kormányzati Informatikai Fejlesztési Ügynökség) lett. A kiöregedő Leo-t és a többi kisebb elavuló szuperszámítógépet 2022-ben váltotta fel az új, szintén a debreceni központban elhelyezett *Komondor* nevű szuperszámítógép, amely 3,09 PFlop/s teljesítménnyel került fel a TOP500-lista 199. helyére. Bár e gép teljesítményét tekintve a világ legnagyobb szuperszámítógépének alig 1/300-ada, technológiájában hasonlít ahhoz. Szintén a HPE/Cray építette, korszerű, melegvizes hűtésű, AMD Epyc Milan processzorokat és 216 db Nvidia A100 GPU-t, valamint Slingshot belső hálózatot tartalmaz. Elmondható tehát, hogy a legkorszerűbb technológiát képviseli.

Magyarország 2011-ben csatlakozott a PRACE-hez, az európai szuperszámítógépes szakmai együttműködéshez, és induláskor belépett az EuroHPC Közös Vállalkozásba is. Mindezek eredményeképp hazánk számos európai HPC-vonatkozású projektben vett részt az elmúlt években, ami a nemzetközi kapcsolatok és partnerség fejlődésén túl sok-sok szakmai tapasztalat átvételét is jelentette. Ahogy már említettük, Magyarország a Leonardo pre-exa konzorciumnak is tagja, így a hazai felhasználók számára könnyített hozzáférés nyílik a világ 4. legerősebb szuperszámítógépéhez is. Természetesen az itt felhasználható kapacitás korlátozott, ahogy a Komondor sem lesz képes több éves távlatban kielégíteni a gyorsan növekvő igényeket. Eddig gyakorlatilag csak a kutatói közösség volt aktív felhasználója a szuperszámítástechnikának, ami nem véletlen, hiszen a HPC alapvetően kutatási infrastruktúra. A közeljövőben azonban várható, hogy egyre több innovatív, kutatást-fejlesztést végző ipari szereplőnél is fellép az igény.

Elég jól prognosztizálható, hogy 2025-re a nemzeti HPC-infrastruktúrával szemben támasztott hazai igények összességében bőven meg fogják haladni a 20 PFlop/s

teljesítményt. Kérdés, hogy ez miként lesz kiszolgálható, hiszen ehhez már most el kellene indulnia a következő fejlesztési lépcsőnek. 2022-ben hazánk elnyerte az Euro HPC 35%-os pénzügyi támogatását egy új, 20 PFlop/s nagyságú szuperszámítógép építésére és üzemeltetésére, a projekt azonban nem kezdődött el. E nélkül azonban a hazai tudományos-kutatási tevékenység nemzetközi szinten versenyhátrányba kerül, és Magyarország kimarad sok olyan projektből és lehetőségből – köztük például a kvantumszámítástechnika korai alkalmazásából – amelyeket az EU jelentős forrásokkal támogat, és amelyek a HPC területén előttünk járó környező országok (pl. Csehország, Lengyelország, Szlovénia, de ma már Bulgária is) számára nyitva állnak. Mindez a magyar innovációs és versenyképességet hátrányosan érinti.

8. Összefoglalás

Az extrém számítási teljesítményre képes szuperszámítógépek a modern kor nélkülözhetetlen eszközei. A mindennapi életben közvetlenül ritkán találkozunk velük, mégis életünk számtalan területére komoly hatással vannak. A szuperszámítógépek a digitális korszak laboratóriumai, a tudományos kísérletek és felfedezések – szinte bármely területről legyen is szó – ma ezek alkalmazásával folynak. Az eredmények pedig előbb vagy utóbb a minket körülvevő környezetben, mindennapi használati cikkeinkben, az árukban és szolgáltatásokban, gyakorlatilag mindenben megjelennek. Az ultranedvszívóképességű pelenkától kezdve az időjós jelentésén át a lázcsillapítóig. De az emberiség nagy, globális kihívásainak megoldására – a klímaváltozástól kezdve a fenntartható élelmiszer termelésen át a járványok leküzdéséig – is csak akkor van vajmi esély, ha szuperszámítógépek segítenek bennünket. Ezért fontos, hogy ismerjük, értsük, megvalósítsuk és minél intenzívebben alkalmazzuk a szuperszámítástechnikát.

Hivatkozások

[1] Richardson, Lewis Fry, 1922: Numerical Prediction by Numerical Process, Cambridge Univ. Press. 2nd Edition (corrected version of the original with a new Foreward by Peter Lynch), Cambridge University Press 2007. ISBN: 978-0-521-68044-8.

[2] Slotnick, Daniel, 1982: The Conception and Development of Parallel Processors – A Personal Memoir. Annals of the History of Computing. 4 (1): 20-30. doi:10.1109/mahc.1982.10003.

[3] Szupergépek teljesítményének időbeli alakulása a top 500-as listán, *https://www.top500.org/statistics/perfdevel/*

[4] Creating a National Strategic Computing Initiative, *https://www.whitehouse.gov/the-press-office/2015/07/29/executive-order-creating-national-strategic-computing-initiative*

[5] Advancing U.S. Leadership in High-Performance Computing, *https://www.whitehouse.gov/blog/2015/07/29/advancing-us-leadership-high-performance-computing*

[6] Partnership for advanced computing in Europe, *https://prace-ri.eu/*

[7] The European strategy for High Performance Computing, *https://digital-strategy.ec.europa.eu/en/library/european-strategy-high-performance-computing*

[8] The European High Performance Computing Joint Undertaking, *https://eurohpc-ju.europa.eu/index_en*

(weboldalak elérési dátuma: 2023.02.28.)

A szerzőkről

MÁRAY TAMÁS egyetemi tanulmányait a BME Villamosmérnöki Karán végezte. Itt szerzett diplomát informatikából és doktori címét is itt védte meg. Tanított a BME-n, majd 16 éven át a magyar felsőoktatási és kutatói hálózatot működtető NIIF Intézet műszaki igazgatója volt. Ő hozta létre Magyarország első webszerverét 1993-ban, valamint meghatározó szerepe volt az Internet és alkalmazásainak hazai fejlesztésében és elterjesztésében. Szakmai irányításával került megvalósításra az első, TOP500-kategóriájú, kutatási célú szuperszámítógép Magyarországon, 2001-ben. Több nemzetközi HPC szakmai szervezetben delegált. Vezetésével alakult meg a HPC Kompetencia Központ (HPC@hu) a KIFÜ-ben.

SZEBERÉNYI IMRE a BME Villamosmérnöki Karán végzett, mint villamosmérnök 1983-ban. Diploma után az BME Folyamatszabályozási Tanszékén (jelenleg Irányítástechnika és Informatika Tanszék) kezdett dolgozni. Munkájával hozzájárult a UNIX-kultúra és az egyetemi számítógépes hálózat kialakulásához. Elsőként teremtette meg az elektronikus üzenetküldés lehetőségét a BME és más külföldi egyetemek között, ami új kaput nyitott a BME oktatói és kutatói számára a nemzetközi kapcsolattartáshoz. Számos olyan kutatásfejlesztési projektben vett részt, melynek fő célja a nagy számítási kapacitások kialakítása és felhasználása. Informatikai doktori fokozatát is ezzel kapcsolatos témában szerezte. A NIIF (KIFÜ) szakértőjeként részt vett több szuperszámítógép specifikációjában és tesztelésében. Szakmai irányításával került beszerzésre és beüzemelésre a BME 360 magos szupergépe 2012-ben. Jelenlegi kutatási és oktatási tématerülete elsősorban a számítási felhők, a HPC és elosztott rendszerek. Szakmai vezetésével és számos egyetemi hallgató munkája eredményeként jött létre a BME VIK oktatási felhőrendszere, a CIRCLE.

HTE Diplomaterv és Szakdolgozat Pályázat, Almanach 2022

A HTE 2022-ben a hagyományokhoz híven ismét meghirdette diplomaterv és szakdolgozat pályázatát, amelyre három kategóriában – a mesterszakos (MSc), az alapszakos (BSc) és az üzemmérnök alapszakos (BProf) végzősök – pályázhattak.

A pályaművek benyújtása és a bírálati folyamat teljes mértékben elektronikusan, transzparens módon, az Easy Chair elnevezésű konferenciamenedzsment-rendszerben történt. A pályázat bírálatának lebonyolítását a HTE Tudományos Bizottsága felügyelte, és ez a Bizottság tett javaslatot a díjazottakra a kialakult bírálati eredmények alapján.

A pályázatra idén 25 pályamű érkezett, ebből 12 az MSc, 10 a BSc és 3 a BProf kategóriában. A pályázók között négy felsőoktatási intézmény hallgatói képviseltették magukat; a Budapesti Műszaki és Gazdaságtudományi Egyetem mellett a Dunaújvárosi Egyetem, a Nemzeti Közszolgálati Egyetem és a Széchenyi István Egyetem hallgatói.

A pályaművek bírálatát egy felsőoktatási és ipari szakemberekből álló bíráló bizottság végezte. A bírálóknak értékelniük kellett a pályaműveket a témaválasztás korszerűsége, a kapcsolódó irodalom feldolgozása, saját munka mennyisége és színvonala, az elért eredmények, valamint a szerkesztés és formai elemek alapján. Ezen szempontokra adott bírálói értékelések összesítését követően állt elő a díjazottak sorrendje.

Ezzel párhuzamosan az IEEE Communication Societyvel és a Hungarian Joint ComSoc/MTT/AP/ED/EMC Chapterrel közösen az angol nyelven megírt pályaművek között egy kategóriafüggetlen „HTE – IEEE ComSoc Thesis Award" különdíj került kiírásra, a korábbi évek hagyományait követve. A különdíjat Dr. Simon Csaba (BME) vezetésével egy szakemberekből álló nemzetközi bizottság ítélte oda.

A tavalyi év után a HTE idén is küldetésének tekinti a díjazott pályaművek egyoldalas kivonatainak megjelentetését almanach formájában, hogy az utókor is bepillantást nyerhessen ezekbe az értékes és nívós munkákba. Ugyanakkor nem titkolt célunk, hogy ez a gyűjtemény érdekes és izgalmas korrajzként is szolgáljon, felvillantva az infokommunikációs területen a végzős hallgatókat megragadó, az adott időszakban korszerűnek számító témákat.

A jelen összeállítás a 2022-es *HTE Diplomaterv és Szakdolgozat Pályázat* díjazott pályaműveiből készült kivonatoknak és a szerzők rövid szakmai életútjának gyűjteménye.

Minden díjazottnak gratulálunk és sok sikert kívánunk további pályafutásukhoz!

Farkas Károly
HTE Tudományos területért felelős elnökségi tag, a pályázat koordinátora

További információk:

https://www.hte.hu/hte-diplomaterv-szakdolgozat-palyazat

TARTALOM

Véletlenszerű környezetek mély megerősítéses tanulásban önvezetéshez

BÉRES ANDRÁS

BME, Távközlési és Médiainformatikai Tanszék
beres.andris@gmail.com

Konzulensek: Dr. Gyires-Tóth Bálint, Moni Róbert (BME, Távközlési és Médiainformatikai Tanszék)

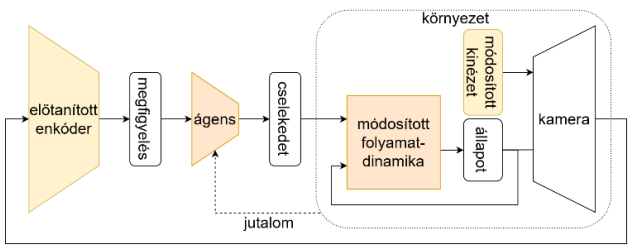
Kulcsszavak: megerősítéses tanulás, mélytanulás, neurális hálózat, önvezető autó, véletlenszerű környezet

A megerősítéses tanulás és a mélytanulás két olyan tudományterület, melyek együttesen az önvezető járművek kihívásának megoldására kínálnak lehetőséget, azonban van két megoldásra váró probléma a valós alkalmazásuk útjában, ezekkel foglalkozom dolgozatomban.

Az egyik ilyen probléma a megerősítéses tanulás adatéhsége, aminek egy lehetséges megoldása szimulátorok használata, és az ezekben történő tanítás utáni alkalmazás a valóságban. A másik említett nehézség az, hogy ezek az algoritmusok általában nem robusztusak környezetük változásaival szemben, ami a valóságban történő alkalmazásukat megnehezíti.

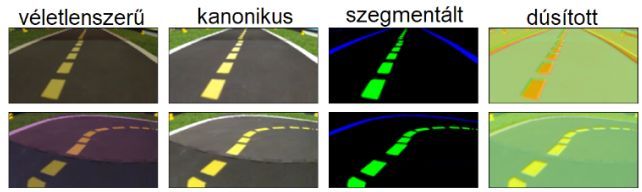
Dolgozatomban a véletlenszerű kinézetű és dinamikájú környezetek módszereit vizsgálom, ezzel párhuzamosan reprezentációtanulást alkalmazok felügyelt vagy felügyelet nélküli módon, hogy a bemeneti képeket tömörebben reprezentálva a megerősítéses tanulást hatékonyabbá tegyem. Az ilyen módon feldolgozott képeket használom aztán fel megerősítéses tanulásra, hogy ezáltal az ágenseket robosztusabbá tegyem, és a valóságban történő alkalmazhatóságot elősegítsem.

Munkám célja a véletlenszerű környezetek módszereinek implementációja és összehasonlítása volt. Dolgozatomban öt véletlenszerű kinézeti és egy véletlenszerű dinamikájú megoldást vizsgáltam meg, azokat modulárisan implementálva a lenti ábrán látható módon, hogy a véletlenszerű dinamika és a véletlenszerű kinézet hatásait egymástól el tudjam választani.



A bemutatott módszereket a Duckietown önvezető környezetben alkalmaztam sávkövetésre, egy differenciális meghajtású jármű irányításával, mindössze egy kamera képe alapján. Az algoritmusokat a PyTorch nyílt forráskódú mélytanuló programkönyvtár segítségével valósítottam meg.

A feladat megoldásához egy 200.000 elemű adatbázist gyűjtöttem össze a szimulátor segítségével, melyben minden elem tartalmazza a kamera által aktuálisan látott képet 3 különböző szimulátor-kinézet szerint renderelve a második ábrán látható módon, és az aktuális állapothoz tartozó fizikai jellemzőket. Gyűjtöttem egy 19.000 képet tartalmazó adatbázist is, mely a valós tesztelés során a kamera által felvett képeket tartalmazza.



Ezeket az adatbázisokat felhasználva a képfeldolgozó hálózatok reprezentáció-tanulásához, összesen 13 ágenst tanítottam megerősítéses tanulással, hogy az eredményességüket a szimulátorban összehasonlíthassam.

Megvizsgáltam a sávkövetés problémáját szabályozástechnikai szempontból is, és megvalósítottam egy új szimulációs környezetet, ami pusztán a szabályozástechnikai feladatot modellezi, elválasztva azt a képfeldolgozástól. Ebben vizsgáltam meg a véletlenszerű dinamikák hatását, és további 4 ágenst tanítottam megerősítéses tanulással.

A megvalósított algoritmusokat nem csak szimulációban, hanem a valóságban is kiértékeltem, összesen 25 valós tesztet és mérést elvégezve. Az eredményeim azt mutatták, hogy a véletlenszerű kinézetek használata a megerősítéses tanulás során hatékony és szükséges is a megfelelő valós működés érdekében. Ezzel szemben a véletlenszerű dinamikák használata – bár a vonalkövetési viselkedést javítja –, hatása kisebb, és a tanítás folyamatát is lassítja. Használata javíthatja a teljesítményt, de a dolgozatom eredményei szerint nem mindenképpen szükséges a kameraalapú sávkövetéshez.

Dolgozatomban a szimulált és valós mérések eredményeit kiértékeltem és összegeztem, bemutattam a vizsgált módszerek előnyeit és lehetséges hátrányait is, a szimulátorban tanított hálózatok valós alkalmazhatósága szempontjából.

A szerzőről

BÉRES ANDRÁS tanulmányait a Budapesti Műszaki és Gazdaságtudományi Egyetemen végezte, villamosmérnöki alap- és mesterszakon. Jelenleg Budapesten a Continental AI Development Center-ben dolgozik deep learning mérnökként.

Automatizált kalibrációs eszköz fejlesztése elektromos kormányrendszer nyomatékszenzorához

BÁNYAI KLAUDIA

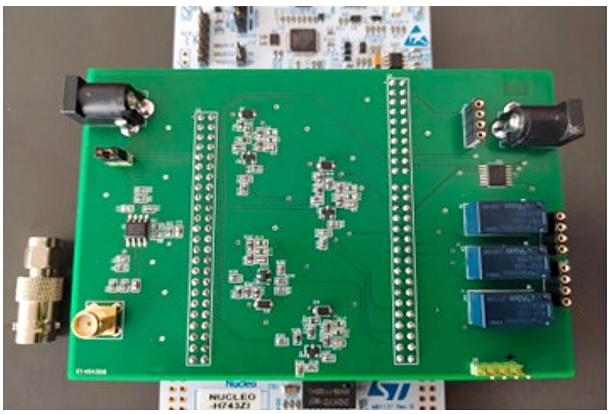
BME, Távközlési és Médiainformatikai Tanszék
banyai.klaudia5@gmail.com

Konzulensek: Paragi László (thyssenkrupp Components Technology Hungary Kft.),
Dr. Varga Pál (BME, Távközlési és Médiainformatikai Tanszék)

Kulcsszavak: nyomatékszenzor, kalibráció, automatizálás, kormányrendszer

Az autók egyre komplexebb funkciói miatt jelentősen megnőtt az alkalmazott szenzorok száma. Növekvő számuk miatt több megfigyelési ponttal szolgálnak, és nagyobb működési biztonságra adnak lehetőséget. Egyik ilyen biztonságkritikus pont a kormányoszlopon fellépő nyomaték ismerete.

A nyomatékszenzorok kalibrációja a pontosabb működés miatt elengedhetetlen. A korábbi kézi kalibráció hozzáértő szakembernek közel egy órát vett igénybe. Munkám keretein belül egy olyan eszközrendszer létrehozása volt a célom, ahol a folyamat egyszerűen elvégezhető és nem igényel szakembert. Ez az eszköz emellett rövidebb mérési időt és pontosabb kalibrációt is lehetővé tesz. Ennek megvalósítására egy mikrokontrollert használtam, amihez plusz nyomtatott áramkört kellett terveznem az egyes jelek csatlakozásához, ahogy az az alábbi képen látható.



Az év során megtörtént a részegységek hardveres tervezése és mérése, valamint ezek összeállítása egy nyomtatott áramkörön. Ezután a mikrokontrolleres programozással ismerkedtem, majd kódot írtam, illetve módosítottam a kiolvasás, kommunikáció és az analóg-digitális átalakítás megvalósítására az *Atollic* programban. Ezen programok megírásához az SPC- és az UART-protokollok ismeretére volt szükség. Részleteiben megismerkedtem a használt nyomatékszenzor-EEPROM tartalmával, majd Python-kódot írtam annak kiolvasására és átírására.

Ezután a felhasználói felületet alkottam meg a Python *tkinter* könyvtárával. A mérés és az ablak párhuzamos megjelenítéséhez betekintettem a több szálon futó programok működésébe, majd alkalmaztam a saját programomban.

Amikor az írás-olvasás a szenzor és a számítógép között megvalósult, rátértem a kalibrációhoz szükséges gain és offset értékek számolására. Ehhez segítségül szolgált a szenzor felhasználói kézikönyve. Amint a hardveres és szoftveres működés megvalósult, a verifikációhoz mérési összeállítást készítettem. Ehhez egy kalibrálatlan szervoegységet használtam egy mérőpadon.

Az automatizált kalibráció a kézi kalibrációhoz képest nagyobb pontossággal és rövidebb idő alatt végbement. A mérési idő 50 perc helyett ellenőrzéssel együtt 1,5 perc alatt elvégezhetővé vált. A felhasználói felület eleget tesz a kívánt funkcióknak. A tervezett eszköz megfelel a célkitűzésnek mind hardveres, mind szoftveres oldalról.

A szerzőről

BÁNYAI KLAUDIA villamosmérnöki tanulmányait a BME-n végezte alap-, és mesterszakon is (2016–22). Szakmai tevékenységei az egyetem nagyfeszültségű laboratóriumában kezdődtek, ahol alapképzése során fél évig demonstrátorként tevékenykedett. Mesterszakon a 'Vezetéknélküli rendszerek és alkalmazások' főspecializáció mellett az 'Okos városok' mellékspecializáción tanult. Az MSc alatt a Karlsruhe Műszaki Egyetemen hallgatott fél évet, ahol megismerkedett a német precizitással. Szakmai gyakorlatát a thyssenkrupp autóipari cégnél végezte, ahol már 2019 januárjától dolgozott szenzorfejlesztői pozícióban részmunkaidőben a tanulmányok mellett, nyaranta teljes munkaidőben. 2022 júliusában a SiliconLabs-nál kezdett Hardware Engineer pozícióban, ahol jelenleg is teljes munkaidőben dolgozik, témában közelebb az MSc-s főspecializációjához.

Kubernetes alapú skálázható automatizált folyamatmenedzsment keretrendszer nagyléptékű 5G-V2X szimulációkhoz

TÓTH VINCE

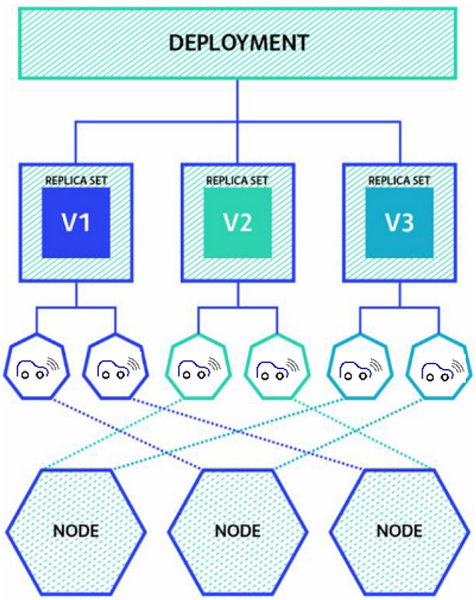
BME, Hálózati Rendszerek és Szolgáltatások Tanszék
vincetot97@gmail.com

Konzulensek: Dr. Bokor László (BME, Hálózati Rendszerek és Szolgáltatások Tanszék), Csányi Zsolt (Datatronic Kft.)

Kulcsszavak: C-ITS, szimuláció, Kubernetes, felhőrendszerek, 5G-V2X

A modern közlekedés egyik új technológiai kihívása a járművek közötti kommunikáció, és annak kihasználása a forgalom működésének javításához. Az új szolgáltatások, amelyeket a kooperatív intelligens közlekedési rendszerek és a vezetést segítő alkalmazások nyújtanak, hasznosak lehetnek a közlekedési balesetek számának csökkentésében, a károsanyag kibocsátásának csökkentésében, vagy a közlekedők kényelmének javításában.

A fejlesztések segítésére és támogatására hálózati és forgalmi szimulátorokat használnak a fejlesztők, amik a valóságot próbálják hűen utánozni annak érdekében, hogy az új alkalmazások hatékonyságáról információkat kapjanak. Ez a módszer igazán költséghatékony és gyors. Nem kíván valós beavatkozást a forgalomba, és növeli a forgalmi tesztelésbe kerülő alkalmazások biztonságát a virtuális tesztelések segítségével. A szimulációk infrastruktúráját a felhőben szinte a végtelenségig lehet erőforrással támogatni. A modern támogató funkciókkal pedig a fejlesztők munkáját lehet könnyebbé tenni.



A munkám során megterveztem és implementáltam egy olyan rendszert, ami modern felhőalapú infrastruktúrába ültet egy összetett járműkommunikációt megvalósító szimulációs rendszert. A rendszer elemei együtt erős eszközöket szolgáltatnak a valósághű szimulációhoz. A forgalomszimulátort a *SUMO* szolgáltatja, a hálózat szimulálásáért pedig az *Artery* keretrendszer a felelős. Felhőrendszernek a *Kubernetes* megoldását választottam, amit több nagy cég szolgáltatásai között megtalálhatunk, így a technológia korszerűségéről megbizonyosodhatunk. A rendszer létrehozásához egyedi virtuális gépeket használtam, amiket klaszterbe csatlakoztatva egy saját Kubernetes rendszert hoztam létre. Ennek feladata a skálázható elosztott tárhely és hálózati megoldás szolgáltatása, ami a szimulációk erőforrásait biztosítja majd.

A szimulátort a felhőben alkalmazható formában kellett integrálni. Ezt a *Docker* technológia segítségével értem el, magát a rendszert konténerizáltam. Ezáltal sokszorosítható és hordozható megoldást kaptam az amúgy összetett és bonyolult program helyett.

A felhőrendszer sajátos feladatainak megoldására egy központi alkalmazást fejlesztettem, ami a szimulátor natív bemenete alapján párhuzamosított konfigurációkat hoz létre, majd ezeket használja fel, hogy a felhőben is végrehajtható szimuláció-konténereket hozzon létre, ahogyan az ábrán is látható. A szimulációk ilyen formában több fizikai csomóponton végrehajthatók. Ezek a konténerek együttesen gyorsabb eredménnyel szolgálnak a szimulált programok viselkedéséről, mint az eredeti szimulátor, és automatikusan skálázhatóak az igényelt erőforrások szerint.

A kapott eredményeket egy másik alkalmazás fejlesztésével automatikusan vizualizálni és értékelni tudtam. Ez az alkalmazás a szimulátorból származó nyers adatok feldolgozásával percek alatt látványos és értékes információval szolgált a tesztelt alkalmazás hatékonyságáról.

A két alkalmazást magát is felhő-kompatibilissé téve, egy skálázható megoldást kaptunk, ami más DevOps technológiákkal kombinálva a fejlesztés szinte minden részét gyorsítani és egyszerűsíteni tudja, ezáltal segítve a modern alkalmazások megalkotását.

A szerzőről

TÓTH VINCE tanulmányait a Budapesti Műszaki és Gazdaságtudományi Egyetem mérnökinformatikus MSc-szakán végezte. Az egyetemi munkája során a kooperatív intelligens közlekedési rendszerekkel foglalkozott átfogóan, ez később az iparban szerzett tapasztalataival vegyült a felhőrendszerek fejlesztése terén, ami a dolgozat motivációját is szolgáltatta.

Szigetelés vizsgáló műszer kapcsolóüzemű tápegysége

LORÁNT GÁBOR DÁNIEL

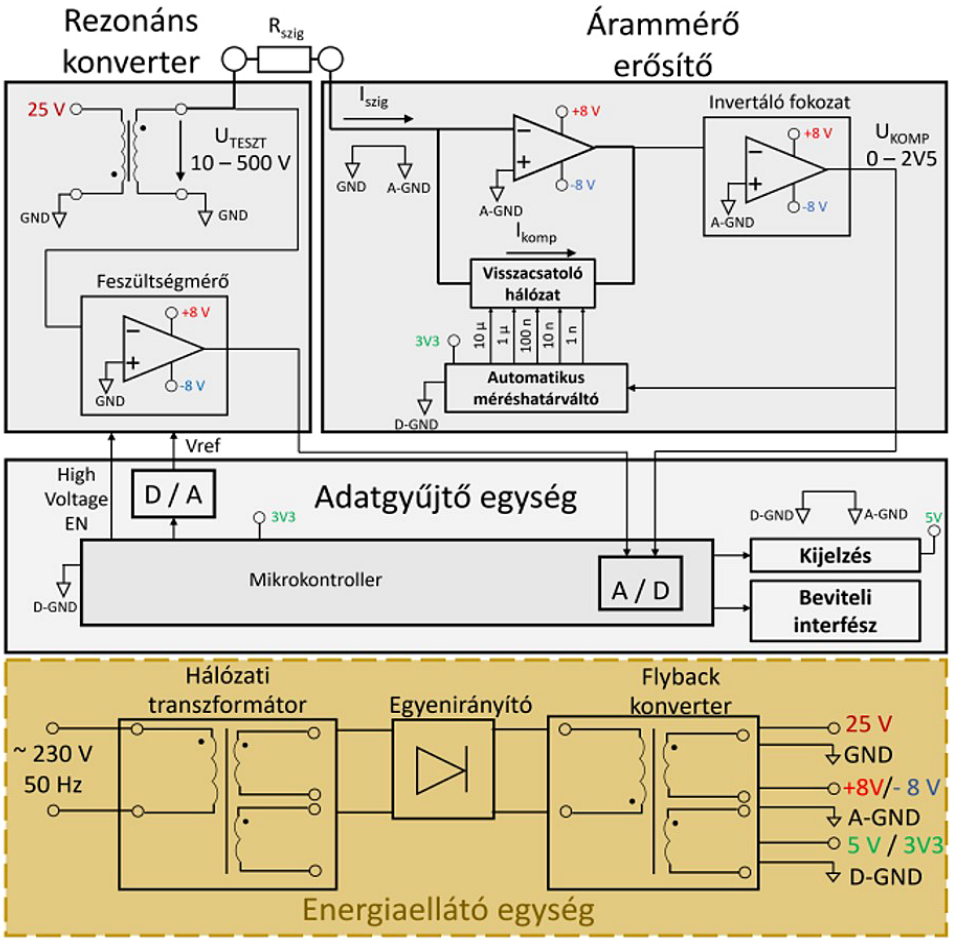
BME, Szélessávú Hírközlés és Villamosságtan Tanszék
lgd199819@gmail.com

Konzulens: Dr. Szabó József (BME, Szélessávú Hírközlés és Villamosságtan Tanszék)

Kulcsszavak: kapcsolóüzemű tápegység, Flyback konverter, szigetelésvizsgáló műszer, szabályozás

Diplomamunkám során egy szigetelésvizsgáló műszer kapcsolóüzemű tápegységét valósítottam meg. Az ábrán látható a szigetelésvizsgáló felépítése, az én feladatom a sárgával jelölt rész megvalósítása volt. Mind a tervezést, mind az építést, mind pedig a bemérést egyaránt a HVT Űrtechnológia Laboratóriumában végeztem.

Első lépésben a szigetelésvizsgáló műszer szükségességét, valamint működését tárgyaltam rövidebben. A kapcsolóüzemű tápegységek csoportosítása után a tápegység típusának választását részleteztem, ezt követően felvázoltam a tápegység blokkvázlatát, majd a részáramkörök működési elvét írtam le. Az áramkörön belül minden részáramkört három nagyobb részegységhez lehet sorolni; vannak az energiaátviteli lánc elemei, a vezérlőkör elemei, valamint a védelmi áramkörök.



Ezek után ismertettem a Flyback konverter transzformátorának méretezését és konstrukcióját. A transzformátor bemérésére a megfelelő menetszámok és a csatolások ellenőrzése végett volt szükség. A modelláramkör megépítése után az élesztés és a nem megfelelően működő részegységek javítása következett. Ezek után a tápegység funkcióit ellenőrző minősítő mérések leírását foglaltam össze. Végül pedig a minősítő mérések eredményeit ismertettem.

A szerzőről

LORÁNT GÁBOR DÁNIEL tanulmányait a BME Villamosmérnöki és Informatikai Karán végezte, azon belül a Szélessávú Hírközlés és Villamosságtan tanszéken. Egyetemi pályafutása alatt több munkahelyen is megfordult, legjelentősebb ezek közül a Bonn Hungary Electronics Ltd.-nél töltött ideje, ahol első teljes állását is megkezdte. Egyetem alatt a legjobban a nagyfrekvenciás áramkörök viselkedése foglalkoztatta, melyet munkahelyén is hasznosítani tud.

UWB alapú beltéri helymeghatározó rendszer elemeinek fejlesztése

KOVÁCS ZOLTÁN MÁRK

BME, Hálózati Rendszerek és Szolgáltatások Tanszék
kovacszoltanmark@gmail.com

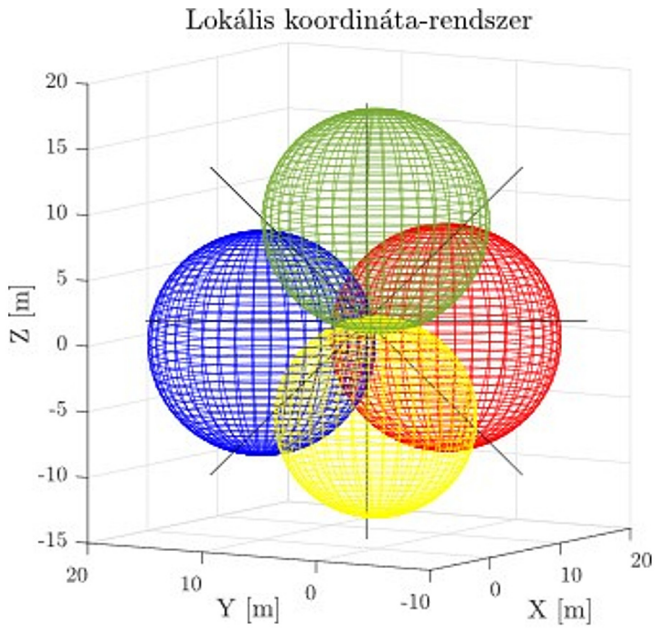
Konzulens: Dr. Matolcsy Balázs (BME, Hálózati Rendszerek és Szolgáltatások Tanszék)

Kulcsszavak: UWB, beltéri helymeghatározás, pozicionálás, adatbáziskezelés, hardvertervezés

Az ultraszélessávú (Ultra-Wideband, a továbbiakban: UWB) technológia hatékonyan alkalmazható beltéri helymeghatározás esetén. Ez a beltéri helymeghatározás pontossággal szembeni igényére vezethető vissza, amely igényt a globális helymeghatározó rendszerek általában nem tudják kellő megbízhatósággal kielégíteni. Ezt többek között a beltéri jelterjedési jelenségek, valamint a beltérben tapasztalható jelerősség-csökkenés okozza. Ezzel szemben az UWB-technológián alapuló módszerekkel képesek vagyunk 5-15 cm pontossággal beltéri pozíciót becsülni, ráadásul ezt adott esetben alacsony késleltetéssel. UWB esetén ezek a módszerek alapvetően beérkezési idő, illetve beérkezési időkülönbség alapján becsülik a keresett eszköz pozícióját.

A diplomaterv témáját képezi az UWB-technológia alapjainak ismertetése, a kapcsolódó szabvány által specifikált paraméterek grafikus illusztrálása és összehasonlítása más, helymeghatározásra alkalmas technológiával. Bemutatásra kerül, hogy milyen pontosságot befolyásoló tényezők léphetnek fel az UWB-technológia alkalmazása során, melyek közül egy jellemző hibafaktor hatása grafikus megjelenítést kap több üzenetváltási sémára is.

A vizsgamunka magában foglalja továbbá egy UWB ‘tag’ eszköz, illetve több UWB ‘anchor’ eszköz megtervezését és megépítését, amelyekkel utána demonstrálható a helymeghatározási algoritmus működése. A pozicionálás követését egy PC-re készített vizualizációs felület segíti, amelyen megjelenítésre kerül a ‘tag’, illetve az összes ‘anchor’. A vizualizációs felület egy MATLAB-applikációt foglal magába, amely felhasználóbarát módon végigvezet egy adatbázisba való bejelentkezésen, az ‘anchor’ eszközök koordinátáinak megadásán, illetve a pozicionálási algoritmus paramétereinek beállításán. Utóbbiakat tekintve lehetőség van analitikus vagy numerikus módszeren alapuló algoritmus beállítására.



Az ábra a numerikus algoritmus mechanizmusát szemlélteti, amikor is négy adott ‘anchor’ koordináta körül az egyes mért/számított ‘tag’–’anchor’ távolságoknak megfelelő sugarú gömbök feszülnek ki a térben, majd a ‘tag’ pozíciójának becslése az a térbeli pont lesz, amely az összes gömbfelülethez a legközelebb van. A becslési pontot egy nyújtott ágú fekete csillag szemlélteti. Egyéb beállítási lehetőség a megjelenítés módja, valamint egy utólagos adatszűrést megvalósító Kálmán-szűrő paraméterezése, illetve a becslési pontok átlagolása.

A szerzőről

KOVÁCS ZOLTÁN MÁRK tanulmányait a BME Villamosmérnök és Informatikai Kar Villamosmérnök szak alapképzésén és mesterképzésén folytatta 2015 és 2022 között. Érdeklődési körei főként a rádiófrekvenciás tervezés, általános és célhardver-tervezés, valamint az optikai tartományban történő adattovábbítás. Szakdolgozatát látható fényű kommunikáció témakörében írta, amelynek eredménye két modellautó, amelyek egymás között – fénnyel való adattovábbítással – egyirányú, haladó mozgás közbeni távolságtartást valósítanak meg. Diplomatervéhez fűződő témája a teljes mesterképzést átfogja, ezáltal elmélyült ismeretekre tett szert az UWB-alapú beltéri helymeghatározás témakörében a hardver- és szoftveroldalon egyaránt. 2022 őszétől rádiófrekvenciás hardverfejlesztő mérnökként dolgozik a Mediso Medical Imaging Systems Kft. elektromos fejlesztés csoportjában.

Anomáliadetekció az Ipar 4.0-ban valószínűség alapú eljárásokkal

BICSKI BÁLINT

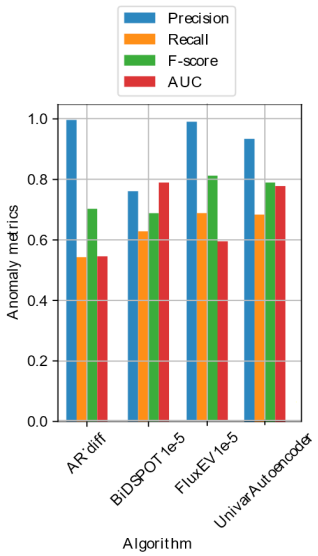
BME, Hálózati Rendszerek és Szolgáltatások Tanszék
balint.bicski@edu.bme.hu

Konzulensek: Dr. Pekár Adrián (BME, Hálózati Rendszerek és Szolgáltatások Tanszék),
Dr. Farkas Károly (Gloster Networks Kft. / BME, Hálózati Rendszerek és Szolgáltatások Tanszék)

Kulcsszavak: anomáliadetekció, idősor-analízis, Ipar4.0, statisztikai algoritmus, gépi tanulás

Az ipari rendszerek állapot-felügyeletének egyik fontos célja, hogy a rendszer folyamataiban a rendellenes viselkedést felismerje, illetve megpróbálja ezeket megelőzni. Az előbbit anomáliadetekciónak, az utóbbit anomáliaprevenciónak nevezzük. Ez a két folyamat óriási jelentőséggel bír, mivel a rendellenes viselkedés a kritikus berendezések meghibásodásának, vagy akár a rendszer elleni szándékos kibertámadásnak a jele is lehet. Szakdolgozatomban az anomáliák detektálására koncentráltam.

A rendszermonitorozás során folytonos jelleggel, több forrásból is érkeznek ún. többdimenziós adatok, amely óriási mennyiségű feldolgozandó adathalmazt generál. A probléma megoldására használt megközelítések legnépszerűbb területeként a gépi tanulást azonosítottam, számos algoritmus létezik kifejezetten az anomáliadetektálásra szabottan ezen a területen. Ezek a megközelítések képesek felismerni az adatokban az összetett mintázatokat, és ez alapján készítik el modelljüket. Azonban mivel az adatokat kívánatos minél kisebb késleltetéssel feldolgozni, – hogy az esetleges anomáliák hatását minimalizálni tudjuk –, olyan gyors algoritmusra van szükség, amely képes az adatokat online módon elemezni, vagyis az adatdarabok beérkezésekor, nem pedig akkor, amikor a teljes adatkészlet elérhetővé válik. Továbbá fontos, hogy a változó viselkedésre is adaptívan reagáljon.



A gépi tanuló algoritmusok azonban gyakran nem felelnek meg ennek a követelménynek, ezért célszerű lehet egyszerűbb detektorokat használni, feltéve, hogy azok jó teljesítményt nyújtanak. Kutatásom során több olyan tanulmányt azonosítottam, mely azt állítja, hogy számos idősoros probléma nem igényel összetett gépi tanulási modelleket, mivel a következő eseményre vonatkozó összes releváns információt egy kis időablakban található néhány közelmúltbeli esemény közvetíti, egyszerűbb algoritmus használatát is lehetővé téve. Munkámban ezt a hipotézist teszteltem, egy korszerű gépi tanuló megközelítést (autoencoder) összevetve egyszerűbb valószínűségi megközelítésekkel.

A SPOT-algoritmus azt a feltételezést használja ki, hogy egy rendszer viselkedésében a normál mintázatok dominálnak az anomáliák felett. Az algoritmus egy valószínűségi modellt épít az adatokra, valamint támogatja a normál viselkedés változását a modell folyamatos dinamikus újrakalibrációjával. A SPOT csak olyan anomáliákat detektál, melyek az automatikus tűréshatárt meghaladják. Ennek továbbfejlesztése a FluxEV, melynek célja, hogy többlépéses valós idejű előfeldolgozással már az amplitúdóban normálnak számító, de mintázatra abnormális viselkedést is képes legyen detektálni. A SPOT és a FluxEV mellett egy autoregressziós statisztikai modellt is teszteltem több beállítás mellett.

Az egydimenziós adatsorokon működő valószínűségi és statisztikai algoritmusokat adaptáltam egy többdimenziós adatok tesztelésére alkalmas keretrendszerben való futtatáshoz, figyelembe véve a valós idejű működés követelményét. Az algoritmusok teljesítményének kiértékelésére teljesítmény- és idő-komplexitási méréseket végeztem. A teljesítményt többfajta, többdimenziós aggregációs és tűréshatárt számító megközelítés mellett is vizsgáltam, emellett javaslatokat tettem arra, hogy az algoritmusok különböző párosításokban mekkora mértékben egészítik ki egymás működését.

Az ábra egy víztisztító üzemben mért adatokon végzett anomáliadetekciós mérés eredményeit mutatja a legjobb modellbeállítások mellett. Az *F-score* metrikában a FluxEV, *AUC* metrikában pedig a SPOT múlta felül az autoencoder teljesítményét, mely időbeliségben is alulmaradt, futási ideje a többi algoritmus 210-560-szorosa volt.

A kiértékelés eredményei alapján a feltevést sikeresen tudtam igazolni; az egyszerűbb valószínűségi megközelítések fel tudják venni a versenyt a mélytanuló algoritmusokkal többdimenziós ipari idősorok feldolgozása során is.

A szerzőről

BICSKI BÁLINT 2022-ben szerzett BSc-diplomát. Tanulmányait a Budapesti Műszaki és Gazdaságtudományi Egyetem Hálózati Rendszerek és Szolgáltatások tanszékén mérnök-informatikus szakon folytatja, jelenleg tudományos munkatárs és MSc-hallgató. Hálózati folyamok elemzésével, gépi tanuláson alapuló feldolgozásával foglalkozik.

Hálózati feszítőfák automatizált átalakítása

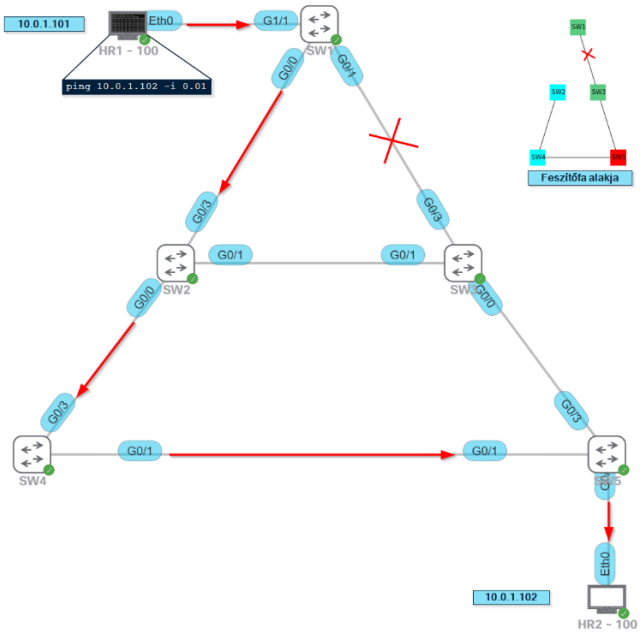
SZEPESI-NAGY ISTVÁN

BME, VIK, Hálózati Rendszerek és Szolgáltatások Tanszék
sznistvan@mensa.hu

Konzulens: Dr. Zsóka Zoltán (BME, Hálózati Rendszerek és Szolgáltatások Tanszék)

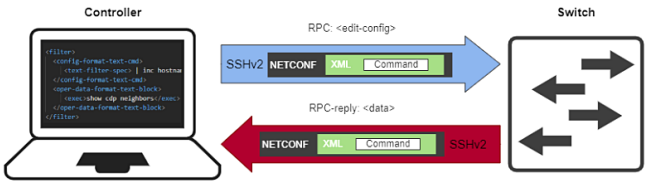
Kulcsszavak: automatizálás, switch, NETCONF, STP

A második rétegbeli eszközök között a Spanning Tree Protocol (STP) biztosítja a hálózati feszítőfák kialakítását. Statikusan összekötött végpontok esetében az STP-beállítások állandóak, azonban egy dinamikusan változó hálózatban szükséges lehet a kialakult feszítőfát átrendezni. Az átrendezés motivációi között lehet a teljesítmény növelése vagy a terheléselosztás.



Egy komplex felépítésű L2-hálózatban elengedhetetlen az STP alkalmazása, azonban folytonos átalakítás mellett a fa kialakítása elmozdulhat az optimumtól. Erre a problémára találtam ki és implementáltam egy módszert Python programozási nyelven, ami a NETCONF-protokoll felhasználásával valósul meg. A program automatizált rendszerben képes a hálózati topológia és az STP-beállítások alapján új feszítőfákat felépíteni, az egyes VLAN-okra nézve. A megvalósítás során alkalmaztam egy Kruskal-algoritmuson alapuló maximális súlyú feszítőfa-algoritmust, az új fa kialakítása érdekében pedig a switch-eken lévő port-költségeket és gyökér-prioritásokat módosítottam.

Az alkalmazott NETCONF-protokoll egy IETF által definiált szabvány, ami különböző hálózati eszközök konfigurációinak telepítéséért, változtatásáért és törléséért felelős. A kialakított NETCONF-protokoll már emberi nyelven értelmezhető segítséget nyújt (az SNMP-vel szemben) a hálózati eszközök menedzselésében. A NETCONF négy rétegre van felbontva, amelyek lehetővé teszik a biztonságos és stabil kommunikációt az eszköz és a kontroller között. A NETCONF-kapcsolat egyszerű és programozott megoldására az *ncclient* nevű Python-könyvtár metódusait használtam. Az *ncclient* minden olyan tulajdonságot biztosít, ami szükséges a megfelelő kapcsolat felépítéséhez.



A megvalósítás során a Cisco CML (Cisco Modelling Labs) virtuális hálózati környezetet használtam. Itt valósítottam meg azt a példahálózatot, amit valós eszközökön is létre lehetne hozni.

Dolgozatomban egy olyan módszert implementáltam, amely a hálózati konvergencia idejének mérése alapján effektíven és automatizált módon képes volt átalakítani és érvényre juttatni egy új feszítőfát az aktuális hálózati igénynek megfelelően.

A szerzőről

SZEPESI-NAGY ISTVÁN mérnök-informatikus alapdiplomáját a Budapesti Műszaki és Gazdaságtudományi Egyetemen szerezte meg 2022 januárjában. Alapszakos tanulmányai mellett a 2019-es BME TMIT IoT-versenyen első helyezést ért el az Innovációs Track kategóriában, 2020 februárjában pedig a Cisco CCNA vizsgáját teljesítette sikerrel. 2022-től a Pázmány Péter Katolikus Egyetem info-bionika mérnöki mesterszakán hallgató, rendszerbiológia specializáción. Érdeklődési körébe tartoznak a bioinformatikai megoldások, ezen belül a génexpressziós szintek vizsgálata különböző Machine Learning módszerekkel.

Optoelektronikus oszcillátorok vizsgálata

IVÁNYI LÁSZLÓ MÁTÉ

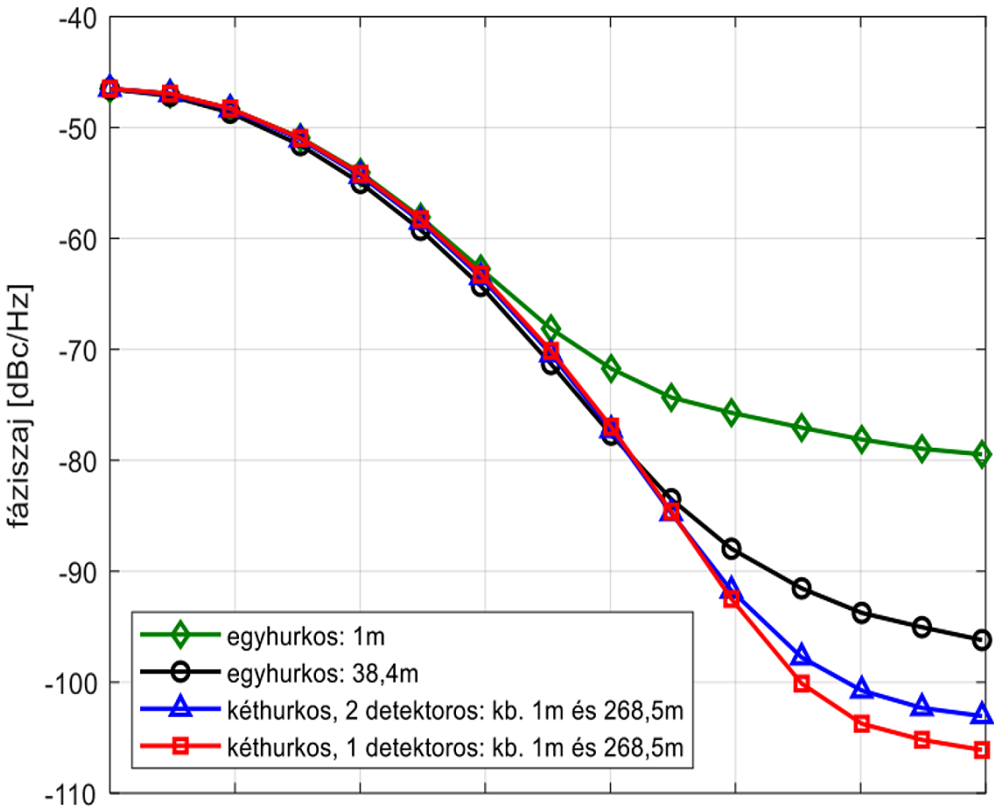
BME, Hálózati Rendszerek és Szolgáltatások Tanszék
mateivanyi@gmail.com

Konzulens: Gerhátné Dr. Udvary Eszter (BME, Hálózati Rendszerek és Szolgáltatások Tanszék)

Kulcsszavak: optoelektronikus oszcillátor, fáziszaj, 5G, SOA, EAM

A modern mobilhálózatokban az 5G megjelenésével szükség van a magas vivőfrekvenciás (>20 GHz), alacsony fáziszajú mikrohullámú jelek előállítására. A tendencia pedig azt mutatja, hogy a jövőben az egyre magasabb frekvenciák elérése lesz a cél, amelyre kedvező megoldást nyújt az optoelektronikus oszcillátor (OEO), amely a hagyományos elektronikus oszcillátorokhoz képest frekvenciafüggetlen fáziszajjal rendelkezik.

Szimulációkkal és mérésekkel vizsgáltam az optoelektronikus oszcillátort. Összeállítottam az egyhurkos, a kéthurkos–egy detektoros és a kéthurkos–két detektoros OEO-elrendezést, melyek elektromos kimenetein mért fáziszajt és módustávolságot összehasonlítottam egymással és a szimulációs eredményekkel, több különböző optikai szálhossz érték mellett.



Mindent összevetve, – amint az ábra is mutatja –, a három vizsgált elrendezés közül a kéthurkos, egy detektoros megoldás adta a legkedvezőbb spektrumképet. Meg kell említeni azonban, hogy a kéthurkos, egy detektoros elrendezés hátránya, hogy polarizáció-kontrollert is tartalmaz, amely a helyigény szempontjából kritikus kérdés.

Munkám során három különböző optikai intenzitásmodulátor-típust vizsgáltam meg, amelyek közül az elektroabszorpciós modulátort (EAM) találtam optimális választásnak az OEO elrendezésében. Módustávolság és fáziszaj tekintetében mindhárom modulátor alkalmazása közel azonos értékeket adott, azonban az EAM-lézerrel való integráltsága kis beiktatási csillapítást eredményezett, továbbá ez a modulátor biztosította a legstabilabb működést. A Mach-Zehnder modulátor (MZM) és a félvezető optikai erősítő (SOA) munkapontjának hőmérsékletfüggése miatt biztosítani kell a modulátorok hőfokstabilizálását. Továbbá a SOA munkapontbeállítása bonyolult feladat, amely a nem kontrollálható modulációs mélységnek köszönhető. Ennek köszönhető a SOA telítésbe vezérlődése és az OEO kimeneti jelének nemlinearitása.

A szerzőről

IVÁNYI LÁSZLÓ MÁTÉ egyetemi tanulmányait 2018-ban kezdte a BME Villamosmérnöki és Informatikai Karának villamosmérnöki szakán, ahol 2021-ben TDK második helyezést, 2022-ben pedig kitüntetéses diplomát szerzett. Jelenleg ugyanott MSc-s hallgató.

Apertúracsatolt MSA UHF sávra

VARGA DÁNIEL

BME, Szélessávú Hírközlés és Villamosságtan Tanszék
vargad9797@edu.bme.hu

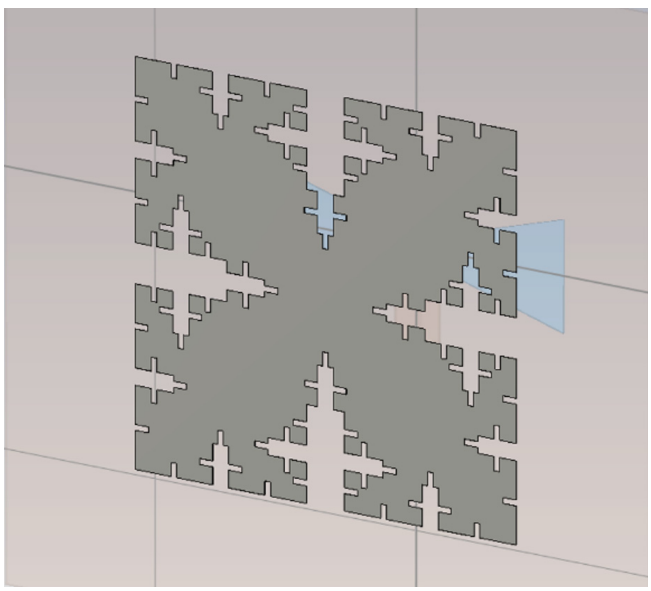
Konzulensek: Dr. Lénárt Ferenc (BME, Szélessávú Hírközlés és Villamosságtan Tanszék), Baranyai Nándor Zsolt (Pro Patria Electronics Kft.)

Kulcsszavak: mikrosztripantenna, fraktálantenna, passzív radar, apertúracsatolt patch, szélessávú

A szakdolgozat keretében egy UHF-sávú passzív radar antennarendszerébe tervezett elemi sugárzót készítettem el, melyhez egy apertúracsatolt mikrosztripantennát választottunk. Ezt az indokolta, hogy ez a struktúra számos lehetőséget kínál a működési sávszélesség kiterjesztésére, amelyek közül néhányat én is megvizsgáltam.

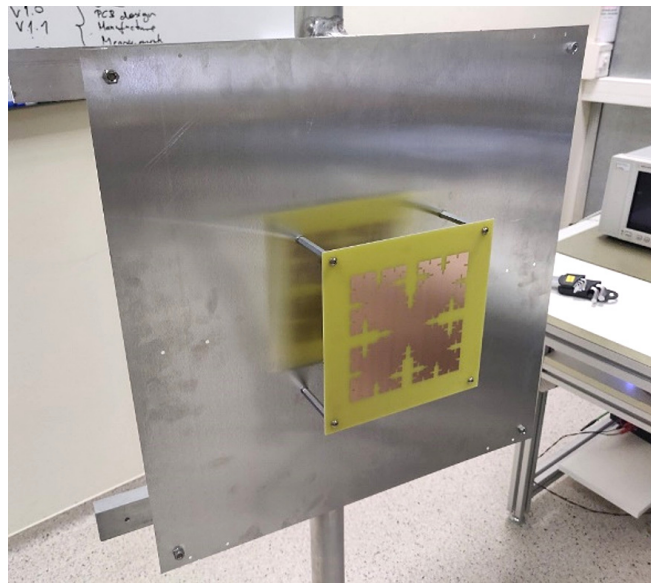
A dolgozatom első felében bemutatásra kerültek az antennákat leíró legfontosabb jellemzők, majd a működés alapjait fejtettem ki. Az így bemutatott antennaparaméterek figyelembevételével zajlott a tervezés folyamata. Ezt követően bemutattam többféle antennatípust és antennarendszert, majd részletesebben a mikrosztripantennákat is.

Ezek után ismertettem a tervezés folyamatát, ahol kitértem a specifikációban foglaltaknak – mint például a szélessávú működésnek, a patch méretcsökkentésének és az előre-hátra viszony javításának – a megvalósítására. Ezen belül felmértem, milyen kihatással vannak a különböző apertúra-alakzatok az antenna működésére, ezen belül leginkább a működési sávszélességükre. Megvizsgáltam, hogy hogyan lehet egy négyzet alakú patchet fraktállá alakítva megtartani annak elektromos jellemzőit, miközben így a patch mérete lényegesen lecsökken.



Bemutattam az antenna elkészítésének folyamatát, és hogy milyen, a megvalósításhoz szükséges módosításokat kellett figyelembe venni, melyek az antenna működésére hatással lehetnek.

Ismertettem a bemérés folyamatát, valamint az elkészült eszköz megmért tulajdonságait a szimulációs eredményekhez hasonlítva. Végül értékeltem a mérésssel is igazolt eredményeket és röviden kifejtettem az antennaelem további fejlesztésének és rendszerbe integrálásának lehetőségeit, valamint az utóbbiakból fakadó lehetséges jövőbeli problémákat.



A szerzőről

VARGA DÁNIEL tanulmányait a gimnázium után a BME Villamosmérnöki és Informatikai Karának villamosmérnöki szakán folytatta, ahol a szakdolgozatírást megelőző félévben kezdett antennákkal foglalkozni egyetemi konzulense segítségével. Ezután került a radarfejlesztéssel és gyártással foglalkozó Pro Patria Electronics Kft.-hez gyakornokként az Antennafejlesztő Csoportba, ahol végül a szakdolgozatát is írta.

Gépi tanulás-vezérelt képfelismerő iOS alkalmazás fejlesztése

SZABÓ BENEDEK ÁRON

Dunaújvárosi Egyetem, Informatika Intézet
szabo.benedek.aron@gmail.com

Konzulensek: Dr. Katona József (Dunaújvárosi Egyetem, Informatika Intézet), Szabó Sándor (Eötvös Loránd Tudományegyetem)

Kulcsszavak: iOS, CoreML, gépi tanulás, transfer learning, képfelismerés

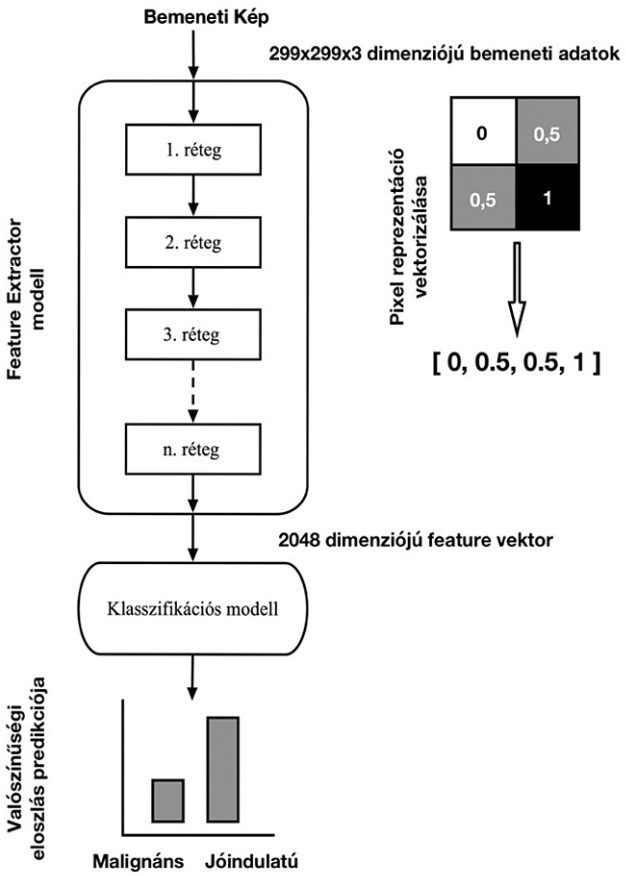
Szakdolgozatom célkitűzése egy olyan gépitanulás-alapú iOS- mobilalkalmazás elkészítése, illetve az ennek alapjául szolgáló bináris klasszifikációs modell betanítása, amely alkalmas a bemenetén megjelenő bőrelváltozások képeinek osztályozására aszerint, hogy az azokon látható elváltozás rosszindulatú-e vagy sem. Az implementáció olyan megközelítéseket és módszereket követelt, melyek alapján a mobilos környezetben ismeretes erőforrás-korlátok mellett is megbízható pontosságú és hatékonyságú applikációt tudtam készíteni.

A modell minőségét, így a klasszifikáció várható pontosságát a mintaadatok begyűjtése valamint felcímkézése, majd az elkészített mintahalmaz adott problémakör kontextusában meghatározott adatpontjai szerint történő kiválogatása és technikai előkészítése útján biztosítottam, figyelembe véve a rendelkezésre álló adatok mennyiségi és minőségi limitációit.

A modelltanításra kiválasztott Create ML segítségével egy, a célplatformra optimalizált konvolúciós modell képességeit feature extractor szerepkörben hasznosítva, és azt egy logisztikai regressziós modellel kaszkádosítva olyan transfer learning pipeline-t alkalmaztam, mellyel a mobilos eszközkörnyezet ismert erőforrás-korlátai mellett is elfogadható pontosságú és hatékonyságú megoldást tudtam generálni. A pipeline szervezését és vázlatos felépítését az ábrán láthatjuk.

A végleges modellverzió eléréséhez számos validációs és verifikációs modellmetrikát vettem figyelembe, melyek mentén a kívánt pontosságot modell- és osztályszinten egyaránt tudtam finomítani. Az overfitting elkerülése érdekében augmentációs segédtechnikákat is alkalmaztam, ügyelve arra, hogy ezzel ne torzuljanak a predikciók alapját adó sarkalatos adatpontok.

Az elkészített modellt a Core ML és Vision framework programozói interfészével integráltam, mindezt az alkalmazáskomponensek feladatkörének szeparációját, így a skálázhatóságot is biztosító MVVM-architektúra keretein belül. Az elkészített alkalmazás működését manuális tesztesetek alapján ellenőriztem és értékeltem ki.



A szerzőről

SZABÓ BENEDEK ÁRON tanulmányait a Dunaújvárosi Egyetem Mérnökinformatikus BSc-szakán végezte. Több éves szoftverfejlesztői, valamint csapatvezetői tapasztalattal rendelkezik hazai és nemzetközi projekteken egyaránt. Jelenleg projektvezető fejlesztőmérnökként dolgozik.

Kriptográfiai műveletek teljesítményértékelése autóipari mikrokontrolleren

PÁLICS MARCELL

BME, Hálózati Rendszerek és Szolgáltatások Tanszék
marcell.palics@icloud.com

Konzulensek: Kiss Miklós (thyssenkrupp Components Technology Hungary Kft.), Bajor Péter (BME, Hálózati Rendszerek és Szolgáltatások Tanszék)

Kulcsszavak: autóipar, beágyazott rendszerek, kiberbiztonság, kriptográfia

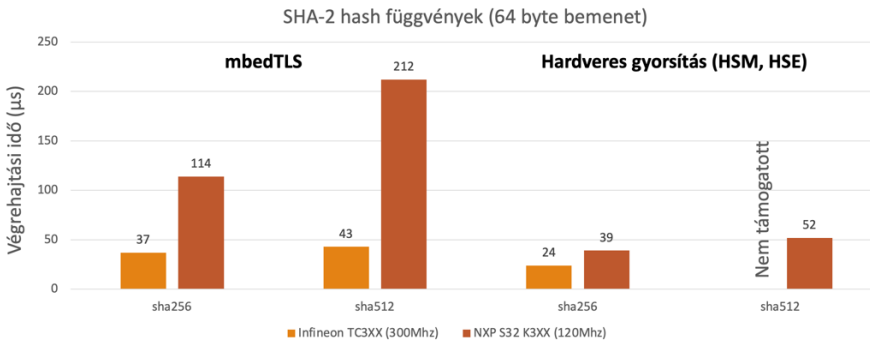
Az elmúlt évtizedben egyértelművé vált, hogy a modern járművekben használt elektronikus vezérlőegységek már rendelkeznek olyan szoftveres komplexitással, mely elkerülhetetlenné teszi információvédelmi szempontból történő vizsgálatukat, különösen biztonságkritikus alegységek, mint pl. fék- vagy kormányrendszerek esetén.

Az ilyen rendszerek üzembiztosságát *(safety)* célzó fejlesztéseket kiforrott szabványcsaládok támogatják, azonban az információbiztonság *(security)* megjelenése újabb kihívásokat jelent az autóipar számára. Ezek közé tartozik azon mikrokontrollerek kiválasztása, melyek a kedvező technikai paraméterek és ár arányának tartása mellett képesek olyan kriptográfiai műveletek futtatására, melyek támogatni tudják az olyan mechanizmusokat, mint a biztonságos rendszerindítás, a szoftverfrissítések validációja, vagy a szellemi termék (programkód) védelme. A safety követelmények és az üzleti céloknak való megfeleléshez – a fent említett mechanizmusokhoz – használt nagyobb számítási igényű kriptográfiai műveletek nem akadályozhatják a meglévő funkciók elérhetőségét. Sok esetben már számolni kell olyan megrendelői igényekkel és nemzetközi szabályozásokkal, melyek hosszú távon biztonságosnak tekinthető algoritmusok és kulcsméretek használatát követelik, akár a posztkvantum kriptográfia jövőképének figyelembe vételével.

A dolgozat témája egy olyan szoftver-keretrendszer létrehozása és bemutatása, mely minimális függőségekkel és a megfelelő hardverspecifikus alacsony szintű műveletek implementálásával felhasználható autóipari mikrokontrollerek kriptográfiai teljesítményének mérésére. A keretrendszer többrétegű architektúrája lehetővé teszi kriptográfiai meghajtók cseréjét, mely lehet egy beágyazott rendszerekre méretezett függvénykönyvtár (pl. mbedTLS, libLithium), vagy egy adott mikrokontroller kriptográfiai számításokat támogató társprocesszorának kezelése is.

A mérések elvégzéséhez szükséges tesztprogramok fejlesztése, valamint minden hardverfüggetlen és -függő komponens és eszközmeghajtó külön szoftverrétegekben biztosított. A rétegzett szeparáció lehetővé teszi, hogy a tesztek fejlesztése hardverfüggetlen maradjon, illetve a chip konfigurációja vagy a kriptográfiai meghajtók módosítása, frissítése sem függ a mérési programlogikától.

A dolgozat írásának idején a keretrendszerbe két különböző gyártó jelenleg is támogatott, komplex 32 bites mikrokontrollerének támogatása került implementálásra, melyek rendelkeztek kriptográfiai társprocesszorokkal. Az alacsony szintű eszközmeghajtókat, chipkonfigurációt mindkét esetben a publikusan elérhető gyártói csomagok és példaprogramok biztosították, ezek szolgáltak a legalacsonyabb szintű hardver-absztrakciós réteg alapjaként.



Az elkészült tesztprogramok közül futtatásra került kiválasztott kriptográfiai primitív függvények (pl. hash függvények), illetve komplex műveletek (pl. digitális aláírás ellenőrzése) végrehajtási idejének mérése is. Az eltelt idő meghatározása minden esetben a mikrokontrollerben található utasítás végrehajtástól függetlenül léptetett, belső számlálók segítségével történt. Az eredmények lehetővé teszik a vezérlők teljesítményének összehasonlítását szoftveres számítási és hardveres gyorsítási lehetőségeik használata esetén, emellett egymással is összevethetők.

A szerzőről

PÁLICS MARCELL tanulmányait a Budapesti Műszaki és Gazdaságtudományi Egyetem Villamosmérnöki és Informatikai Karán végezte üzemmérnök-informatikus alapszakon, Hálózat és biztonság specializáción. Jelenleg a thyssenkrupp Components Technology Hungary Kft.-nél dolgozik kiberbiztonsági tesztmérnökként.

„Cloud native” ChatOps fejlesztés

ADAMEK ÁDÁM

BME, VIK, Hálózati Rendszerek és Szolgáltatások Tanszék
adamek.adam@outlook.com

Konzulensek: Szabó Gergely (Origoss Megoldások Kft.),
Dr. Farkas Károly (Gloster Networks Kft. / BME, Hálózati Rendszerek és Szolgáltatások Tanszék)

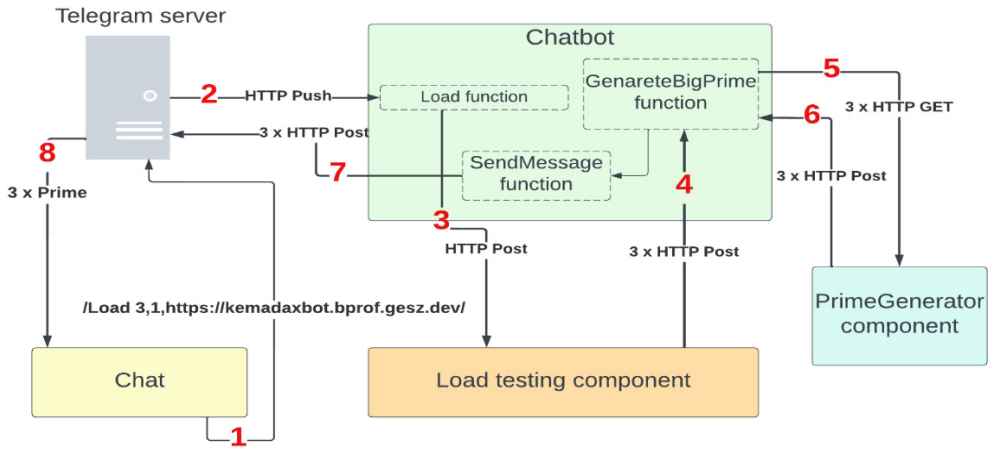
Kulcsszavak: Chatbot, DevOps, felhőalapú alkalmazásfejlesztés, Kubernetes, Cloud Native

A felhőalapú számítástechnika nagyléptékű fejlődése következtében mára már mindennapivá váltak azok az eszközök, amelyek lehetővé teszik a felhőalapú számítási kapacitások széles körben való felhasználását. Ezen kapacitások használatának növekedése többek között olyan előnyeinek köszönhető, mint az igény szerinti erőforrás biztosítása vagy a magas arányú rendelkezésreállási idő. Az ebben rejlő szinte végtelen lehetőségeket az egyik legjobban kihasználó koncepció a „cloud native”.

Egy „cloud native” alkalmazást a konténerizáció, a magas rendelkezésreállás és könnyű, gyors skálázhatóság jellemez. A meglévő kódbázis felhőkörnyezetbe költöztetése előtt érdemes azonban alaposan kielemezni a szolgáltató által ajánlott feltételeket, és felkészíteni a váltásra a már meglévő ökoszisztémát, hogy az minél zökkenőmentesebb legyen, és a cég ne essen áldozatul a „vendor lock-in” jelenségének, amikor adott gyártóhoz van kötve.

Napjainkban egyre inkább elterjedt a *ChatOps* koncepció, melynek lényege, hogy a különböző platformokhoz köthető munkafolyamatokat képes becsatornázni egy *chat* applikációba egy vagy több *chatbot* felhasználásával, ezzel megkönnyítve a kommunikációt és növelve a munkafolyamatok hatékonyságát. Egy chatbot pedig jól adaptálható „cloud native” módszerekkel.

Szakdolgozatom célja bemutatni azokat a technológiákat, illetve technikákat, amelyek lehetővé teszik, valamint megkönnyítik a felhőkörnyezetben való fejlesztési, üzemeltetési feladatokat. A szakdolgozat első része magához a felhőkörnyezethez logikailag legszorosabban kapcsolódó eszközöket, különösképpen a Kubernetest, illetve a „vendor lock-in” jelenséget mutatja be. Az ezt követő rész a fontosabb kód-integrációs, illetve applikáció-üzemeltetési technikákat és eszközöket foglalja össze. A harmadik rész a „cloud native” koncepció elemeit, valamint felépítését prezentálja. A negyedik rész pedig az előzőekben bemutatott technológiák és eszközök működését mutatja be egy olyan chatbot alkalmazáson keresztül, amely ChatOps műveleteket is végez.



Az ábra a szakdolgozathoz készült alkalmazás komponenseit és a komponensek közötti kommunikációt illusztrálja egy, a ChatOps koncepcióba illeszkedő parancs kiadását követően. A „Load” elnevezésű parancs kiadása után a tesztelésért felelős komponens több egymás utáni HTTP-kérés küldésével terheléses tesztnek veti alá a prímszámok generálásáért felelős komponenst. A terheléses teszt futását követően a „cloud native” monitorozási lehetőségek kihasználásával jól becsülhetővé válnak a paraméterek az automatikus skálázási módszerek implementálásához.

A szerzőről

ADAMEK ÁDÁM tanulmányait a Budapesti Műszaki és Gazdaságtudományi Egyetem Villamosmérnöki és Informatikai Karán, üzemmérnök-informatikus szakon végezte. Jelenleg az Ericsson-nál dolgozik szoftverfejlesztőként.

Szeparált információbiztonsági zóna létrehozása biztonsági tesztek elvégzésére

CSERNA LEVENTE

BME, Hálózati Rendszerek és Szolgáltatások Tanszék
cserna.levente@gmail.com

Konzulensek: Kiss Miklós (thyssenkrupp Components Technology Hungary Kft.), Bajor Péter (BME, Hálózati Rendszerek és Szolgáltatások Tanszék)

Kulcsszavak: autóipar, kiberbiztonság, információbiztonság, ISO/IEC 27001

Az elmúlt évtizedek során az autóipar jelentős változásokon ment keresztül. A gépjárműveket egyre inkább a *„négy keréken guruló számítógép”* jelzővel illetik, hiszen egyre összetettebb és bonyolultabb számítógépvezérelt rendszerekkel szerelik fel őket. A növekvő számítási teljesítmény és az elektronikus műszaki innovációk lehetővé teszik olyan elektromos és elektronikus (Electrical and Electronic, E/E) rendszerek implementálását és integrálását, amelyek tisztán mechanikus és hidraulikus technológiákkal megvalósíthatatlanok lennének.

A kiberbiztonság szakterülete egyre hangsúlyosabb szerephez jut az autóipari szektorban, hiszen a beágyazott rendszerek fejlődésével és számítógépvezérelt funkciók térnyerésével egyre nő a kiberbűnözők számára elérhető támadási lehetőségek száma. Ezek kihasználásának kockázata a gyakorlatban is bebizonyosodott, ezért a különböző szabványügyi szervezetek kifejezetten autóipar-specifikus kiberbiztonsági előírásokat tettek közzé (pl. ISO/SAE 21434, UNECE R155 és R156). Ezen szabályozásokban kiemelt szerepet kapott az E/E rendszerek kiberbiztonsági tesztelése, illetve a beágyazott rendszereken futó szoftverentitások bizalmasságának, sértetlenségének és rendelkezésreállásának megőrzése.

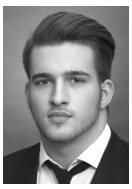
Az üzleti partnerek és a szabályozások elvárásainak való megfelelés olyan komplikált eljárások és eszközök alkalmazását teszi szükségessé, amelyek jelentősen megnövelik a vállalat jelenlegi információs technológiai infrastruktúrájának kiberfenyegetettségét. A szervezet kiberbiztonsági kockázatainak mérséklése, valamint a tesztelési és kutatási folyamatok során keletkező kritikus információk védelmének biztosítása érdekében egy szigorúan védett és elszigetelt zónára van szükség.

A szakdolgozat témája egy olyan szeparált információbiztonsági zóna kialakítása, amely lehetővé teszi a kritikus információk védelmét azáltal, hogy szisztematikus megközelítést biztosít az információbiztonság hatékony kezelésére az ISO/IEC 27k szabványcsalád követelményeinek és ajánlásainak alkalmazásával. Mindemellett a zóna által keretbe foglalt hardveres infrastruktúra biztosítja a beágyazott rendszereken futó szoftverek digitális aláírását, a kriptográfiai kulcsok biztonságos kezelését, illetve a hardveresen gyorsított titkosítást és valódi véletlenszám-generálást.

Az ISO/IEC 27k szabványcsalád követelményeinek és folyamatainak vállalati környezetbe való integrálása egy olyan információbiztonsági zónát eredményezett, amely többek között:

- a vállalati hálózati infrastruktúrától elszigetelt környezetet biztosít a tesztelési és kutatási folyamatok számára, ezáltal minimálisra csökkenti annak kiberfenyegetettségét;
- biztosítja a kritikus információk és információs eszközök bizalmasságát, sértetlenségét és rendelkezésre állását;
- hardveres biztonsági modulok (Hardware Security Module, HSM) segítségével biztosítja a szoftverek digitális aláírását és a kriptográfiai kulcsok megfelelő kezelését;
- teljeskörű kockázatkezelési folyamatot biztosít az ISO/IEC 27005-ben ismertetett módszertan szerint a kockázatok hatékony azonosítására és kezelésére;
- átfogó megközelítést alkalmaz az információbiztonság kezelésére, beleértve az incidensek felderítését és megelőzését, illetve a bevezetett ellenintézkedések folyamatos nyomon követését, felülvizsgálatát és fejlesztését.

A szerzőről



CSERNA LEVENTE tanulmányait a Budapesti Műszaki és Gazdaságtudományi Egyetem Villamosmérnöki és Informatikai Karán végezte üzemmérnök-informatikus alapszakon, Hálózat és biztonság specializáción. Jelenleg a thyssenkrupp Components Technology Hungary Kft.-nél dolgozik Product Cybersecurity Architect pozícióban.

Analóg jelkioltó DVB-T sávú passzív radarhoz

Analog signal suppressor for DVB-T band passive radar

SZABÓ ESZTER

BME, Szélessávú Hírközlés és Villamosságtan Tanszék
szabo.esztiiii@gmail.com

Konzulens: Herman Tibor (BME, Szélessávú Hírközlés és Villamosságtan Tanszék)

Kulcsszavak: passzív radar, referencia-jelkioltás, analóg jelkioltó, DVB-T, nagyfrekvenciás áramkör

Napjainkban egyre nagyobb figyelmet kapnak a passzív radaros technológiák, hiszen számos előnnyel rendelkeznek a klasszikus aktív radarokkal szemben. Azáltal, hogy nem sugároznak ki elektromágneses energiát, hanem csak vevő irányban működnek, nem felderíthetőek.

A Budapesti Műszaki és Gazdaságtudományi Egyetem Mikrohullámú Távérzékelés Laboratóriuma fejleszt egy DVB-T-sávban működő passzív radart. Szakdolgozati félévem során ennek a radarnak a teljesítményét és hatótávolságát javítottam egy külső részegység beiktatásával. Dolgozatom ennek az áramkörnek a megtervezését, realizálását, kalibrálását és mérését részletezi.

A passzív radarok esetében a hatótávolság korlátozásában komoly szerepet játszik a felderítő csatornákba beszivárgó referenciajel. Ennek oka, hogy a megfigyelő csatornák is veszik a direkt jelet, ami ráadásul sokkal nagyobb teljesítménysűrűségű az adott helyen a visszavert jelnél.

Ezen beszivárgás detekcióra gyakorolt hatása a jelfeldolgozás során különböző módszerekkel csökkenthető, ám teljesen nem szüntethető meg. Ennek kiküszöbölése érdekében terveztem egy referenciajel-kioltót, mely a felderítő csatornák jeléből kivonja a referenciajelet. Az analóg referenciajel-kioltó feladata, hogy még a jelfeldolgozási lánc előtt analóg módon csökkentse a felderítő csatornákba beszivárgó jel mértékét, javítva ezzel a radar érzékenységét.

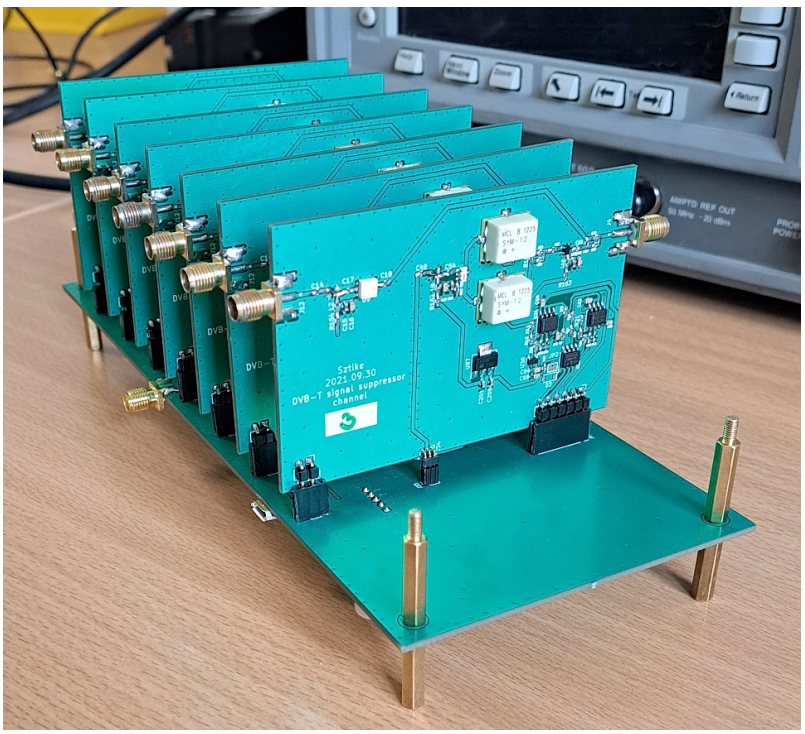
A jelkioltás IQ-moduláció segítségével történik, mely végén a két jelet ellenfázisban összegezve megkapjuk a referenciajel-mentes felderítő csatorna jelét.

Egy DVB-T-alapú passzív radar 8 elemű antennarendszeréhez kellett terveznem a jelkioltót, ahol a radar 1 referencia és 7 felderítő csatornából áll. A cél egy legalább 10 dB-es jelelnyomó képesség elérése volt.

Dolgozatomban kifejtettem a jelkioltás folyamatának alapgondolatát és a tervezés különböző lépéseit. Bemutattam a tervezés során végzett szimulációkat és méréseket, az egyes részegységek működését.

Részleteztem a prototípus során felmerült problémákat és ezen problémák megoldását a későbbi áramkörben. Indokoltam a nyomtatott huzalozású lemezek elkészítése során alkalmazott szempontokat és technikákat.

Dolgozatom a jelkioltóval végzett mérésekkel zárul, ahol megemlítettem a végzett mérések eredményét és az ezt segítő kalibrációs eljárás menetét, majd megjelenítettem és értékeltem a kapott adatokat.



A mellékelt képen az elkészült áramkör látható, ahol egy közös alaplaphoz csatlakoznak a különböző csatornák.

A szerzőről

SZABÓ ESZTER tanulmányait a Budapesti Műszaki és Gazdaságtudományi Egyetem villamosmérnöki alapképzésén végezte, ahol nagyfrekvenciás rendszerek és alkalmazások területen specializálódott. Egyetemi tanulmányai mellett Magyarország első rakétafejlesztéssel foglalkozó versenycsapatának, a BME Suborbitalsnak aktív tagja volt. Jelenleg a stockholmi KTH Royal Institute of Technology egyetem mesterszakos hallgatója.

A HTE-ről dióhéjban

A **Hírközlési és Informatikai Tudományos Egyesület** tekintélyes, 74 éves múltjával hazánk infokommunikációs szektorának legpatinásabb szervezete. Munkássága a távközlési és az informatikai iparágtól a hagyományos postai szolgáltatásokon át az internetig és a médiavilágig terjed. Az egyesület közel 70 jogi tagot, valamint majdnem 800 magánszemélyt számlál. Berkein belül 17 szakmai közösség munkálkodik meghatározott célok érdekében, a legjobb tudása szerint.

A HTE amellett, hogy kiemelt, véleményformáló szerepet tölt be a magyar infokommunikációs szakterület szabályozásában, működésében és fejlődésében, hazai és nemzetközi rendezvények elismert szervezőjeként is jegyzett. A HTE naptárában nem telik el hét szakmai találkozók, előadások, kerekasztal beszélgetések és egyéb szervezett események nélkül, de házigazdája többek között olyan *országos szakmai nagyrendezvények*nek, mint

- a projektmenedzserek találkozóját és információcseréjét segítő **Projektmenedzsment Fórum**,
- a **HTE Infokom**, amely az infokommunikációs hálózatok és alkalmazások legfrissebb piaci, műszaki és szabályozási kérdéseinek egyik legelismertebb rendezvénye.

Az egyesület az utóbbi években számos **nemzetközi tudományos konferencia és workshop** magyarországi megvalósítását nyerte el pályázati úton (bővebb információ a HTE nagyrendezvényeiről: *http://www.hte.hu/nagyrendezvenyek*).

Az egyesület szerteágazó kapcsolatrendszerének, valamint szakmai taglistájának köszönhetően teret ad a különböző vélemény- és információcseréknek, professzionális párbeszédeknek, a közös cél érdekében folytatott kollektív munkának. Lehetőséget biztosít arra, hogy az iparág valamennyi képviselője – akár a mindennapi konkurálást is háttérbe szorítva – kommunikáljon fontos kérdéskörökben. Mindeközben az egyének szakmai kiemelkedésének, véleményformálásának is utat enged. A HTE megkérdőjelezhetetlen erőssége, és így több évtizedes elismertségének egyik alapja, hogy mind munkásságában, mind véleményformálásában **szakszerű, kiegyensúlyozott és független**, felülemelkedik az esetleges politikai, vagy éppen vállalati érdekeken.

Bő hét évtizedet és két tucat szakterületet pár mondatban összefoglalni szinte lehetetlen lenne, ezért inkább vázlatpontokban szemléltetjük, mi minden érinti manapság **az egyesület tevékenységi körét:**

- hazai és nemzetközi konferenciák szervezése
- szakmai fórumok az internet technológiák és szolgáltatások elterjesztéséhez
- műsorelosztási és tartalomipari technológiák népszerűsítése
- rádiótávközlési technológiák ismeretterjesztése
- projektmenedzsment módszerek terjesztése
- szakmai folyóiratok, nyomtatványok kiadása
- nemzetközi kapcsolatok: együttműködés az IEEE-vel és más társszervezetekkel
- szakmai díjak odaítélése (pl. Puskás Tivadar-díj, Kempelen Farkas-díj, Pollák Virág-díj, HTE Fekete László-díj)
- kiegyensúlyozott szakmapolitikai, szakmai véleményalkotás (minisztériumok felkéréseire és társadalmi egyeztetésekben)
- főiskolák és egyetemek képzési és kimeneti követelményeinek véleményezése
- külföldi infokommunikációs eredmények honosításának segítése
- kutatásfejlesztési tevékenység támogatása

A HTE ars poeticájában tradicionálisan hangsúlyos szerepet kap a fiatalokkal való kapcsolat ápolása, a **szakmailag széles látókörű és tudományos érdeklődéssel bíró diákok tanulásának és pályaindításának segítése**. Az egyesület amellett, hogy az arra érdemes hallgatóknak biztosítja a különböző rendezvényeken, konferenciákon és fórumokon való kedvezményes részvételt, különböző pályázatokat hirdet, amelyek díjazottai anyagi elismerésben is részesülnek. A HTE Diplomaterv és Szakdolgozat Pályázatára például az ország 15 egyeteméről nevezhetnek.

A HTE időszakos kiadványai mellett hagyományosan két folyóiratot jegyez: a **Híradástechnikát** és az **Infocommunications Journalt.** Míg előbbi a hazai szakmai élet történéseit, kérdéseit és működését elemzi, az Infocommunications Journal mára nemzetközi elismertségre is szert tett, olyan nemzetközi adatbázisokban szerepelve, mint a Scopus, a Compendex és az Inspec. A lap kiadását a Nemzeti Média és Hírközlési Hatóság támogatja, szerkesztősége folyamatosan szervez keresztpublikációkat nemzetközi társlapokkal.

További információk:
www.hte.hu,
kérdéseit, észrevételeit pedig
az *info@hte.hu*
címre várjuk!

Summaries • of the papers published in this special issue

This Special Issue is compiled from the papers of the 24nd HTE Infokom 2022, the Infocommunications Networks and Application Conference, organized by the Scientific Association for Infocommunications (HTE).

Media distribution and content delivery from an ISP point of view

Keywords: CDN, streaming, automation, IPTV

The legacy, multicast based MPEG-TS is going to be slowly replaced by the adaptive bitrate streaming along its obvious superiority in single client development, AD insertion, and WLAN support, however this raises a clear challenge for the content delivery infrastructure in the bandwidth requirements linearly coupled with the number of customers. The CDN system, built by Telekom, describes a solution to meet this challenge by the utilization of open-source software, container technology and full automation, which results in a highly scalable, fault tolerant, cost effective and flexible system.

6 GHz: Wi-Fi or MFCN? – Regulatory ideas related to the utilization of the upper 6 GHz frequency band

Keywords: WRC-23, CEPT, 6 GHz, Wi-Fi, 5G, MFCN, IMT

A competition has started for the possibility of using the 6 GHz frequency band, several radio services and technology relies on it, applies for it. Especially the future of the upper part of the band is questionable. In the lower part of the frequency band (5945–6425 MHz), there is an opportunity to implement wireless access systems, including radio local area networks (WAS/RLANs). At the same time, we can be sure that the mobile device of the future, which we can get into our hands in a few years, will also make the 6 GHz frequency band available one way or another.

How to define a realistic timeframe for 2G switch off in Hungary?

Keywords: 2G, 3G, 3G sunset, "NetreFel" Program

In Hungary up to the end of 2022, one operator completely and two partially have switched off 3G services. Experiences of the "NetreFel" Program can effectively contribute to the conceptual foundation and the acceleration of the 2G Switch off process, which could take place between 2025–2030 based on the analyses of the current Hungarian situation and the review of international best practices, if a successful migration of 2G services can be completed by that time.

The standalone, "real" 5G mobile networks' financial drivers

Keywords: standalone 5G, network slicing, Business Case

The performance of 5G networks, since 2019 dominantly built on 4G based non-standalone 5G networks, lag behind expectations of enhanced, ultrareliable, low latency and massive machine communication enabling network visions. For 2022 the market become more matured both from demand and technology supply sides thus enabling positive return on investments for 5G standalone (SA) business cases. The research question seeks to answer what are the main techno-economic drivers of SA 5G widespread and which business cases might show positive return.

The coming change in digital identity management: Self-Sovereign Identities

Keywords: SSI, decentralized identity, EBSI, blockchain

With ever increasing connectivity and digitization, the classic models of online and digital identity management have been needing a significant revamp for some time; and with the ongoing standardization and implementation of so-called Self-Sovereign Identities (SSI), this is happening now. The paper reviews the philosophical stance of SSI and the key underlying W3C standards for Disitributed Identities (DID) and Verifiable Credentials (VC). The basic SSI operational model is demonstrated on important use cases – as personal identification and the sharing of certificates and diplomas – and a review of the most important, already existing SSI-supporting networks is provided, including the European Blockchain Service Initiative (EBSI).

Application of AI-based procedures in infocom networks

Keywords: Artificial Intelligence, SDN, NFV, VNF, network management, anomaly detection, AREP

The explosive development of artificial intelligence impacts almost all areas, so infocommunication networks are no exception. Procedures based on machine learning can significantly help solve increasingly complex tasks and automate certain functions, thus making network planning, design, operation, and supervision more efficient. In this article, we briefly review the main possibilities of applying artificial intelligence-based procedures in the management of infocommunication networks and provide insight into our research results in network anomaly detection.

Synthetic data generation – is it the holy grail of data access?

Keywords: machine learning, data access, synthetic data

Data scientists working on AI/ML projects often cannot have access to the required training data. Personal data and confidential business data are even more problematic. Synthetic data generation is a potential solution for these kind of problems. This technology uses generative models, trained automatically on the original data, then uses the model to generate synthetic data which mimics the statistical properties and other patterns of the original data. In addition, it does not "leak" information on the individual original records. In this paper we describe the use cases and foundational methods of synthetic data generation. We share our experience and measurement results about 4 commercial and 6 FOSS software tools.

Digital Trust Services data authentication platform – with blockchain-based digital fingerprint storage

Keywords: blockchain, DTS, ZERO TRUST, supply chain, AI, product follow up, drone blackbox

The data-based economy needs authentic data, the source and immutability of which is verified, so that extensive data-based collaborations, data-based professional and state administration, and secondary data markets can be formed based on it. The article presents Digital Trust Services (DTS) based on blockchain technology. It is an interoperable digital fingerprint-based data authentication platform solution, which enables the authentication and verification of the authenticity of a wide range of digital data, regardless of industry and business size, by everyone.

From John von Neumann to HPC

Keywords: supercomputer, High Performance Computing, cloud, EuroHPC, PRACE, Komondor

We often hear about supercomputers (HPC) these days. But what makes supercomputers different from traditional computers or from the "cloud"? What is the reason for their worldwide prominence in recent years, and why, in the time of various crises and difficulties, their development and expansion have continued unabated and even accelerated? These questions are addressed in this article, which gives an overview of the main characteristics of supercomputers, their specific technical solutions, their typical uses and applications.

Almanach 2022

A special section of this issue contains one-page abstracts written by winners of HTE awards for best diploma theses in 2022.