

Nagysebességű kommunikációs folyamatok nem-stacionárius datagram forgalmának időbeli mintázata

On the time patterns of the nonstationary datagram based fast communication processes

Dr. GÁL Zoltán¹, Dr. TERDIK György¹

¹Debreceni Egyetem Informatikai Kar
4028 Debrecen, Kassai út 26., Magyarország, <http://www.inf.unideb.hu>

Abstract

Nowadays expectations against modern communication services involve not just Quality of Service (QoS) enhancement for real-time applications but also increased transmission rate between the storing and processing of Big Data nodes. Transmission Control Protocol (TCP) has strict flow control of the data stream providing automatic adaptation to the path load of the process-to-process communication. User Datagram Protocol (UDP) based solutions are proposed to settle the communication efficiency. In this paper, we analyse the effect of three independent communication parameters on the efficiency of looped UDP communication: the size of the Maximum Transfer Unit (MTU), the bandwidth of the end-to-end session, and the segment size of the UDP protocol data unit. The usage of nonstationary multi-resolution methods helps to identify three characteristic patterns offering identification of the objective qualitative features of the UDP-based looped datagram communication services.

Keywords: datagram, high-speed network, nonstationarity, Fourier transform, nonlinearity, Wavelet transform, Empirical Mode Decomposition, Variational Mode Decomposition.

Kivonat

Napjainkban a modern kommunikációs rendszerekkel szembeni elvárások nem csupán a szolgáltatás minőségének (QoS) javítását jelentik a valós idejű alkalmazások számára, hanem a Big Data hubok tárolása és feldolgozása közötti megnövelt átviteli sebességet is. A Transmission Control Protocol (TCP) szigorúan szabályozza az adatfolyamot, így automatikusan alkalmazkodik a folyamatok közötti kommunikáció által használt útvonal terheléséhez. A tartalom átviteli hatékonyságának javítása céljából újabban az User Datagram Protocol (UDP) alapú megoldások javasoltak. Ebben a cikkben három független kommunikációs paraméter hatását elemezzük a visszacsatolós UDP kommunikáció hatékonyságára vonatkozóan: a maximális átviteli egység (MTU) mérete, a végpontok közötti munkamenet sávszélessége és az UDP szegmens mérete. A nem stacionárius többfelbontású módszerek alkalmazása lehetővé tette három jellegzetes forgalom-mintázat megkülönböztetését, amely a visszacsatolós datagram kommunikációs szolgáltatások objektív minőségi jellemzőinek beazonosítását szolgálja.

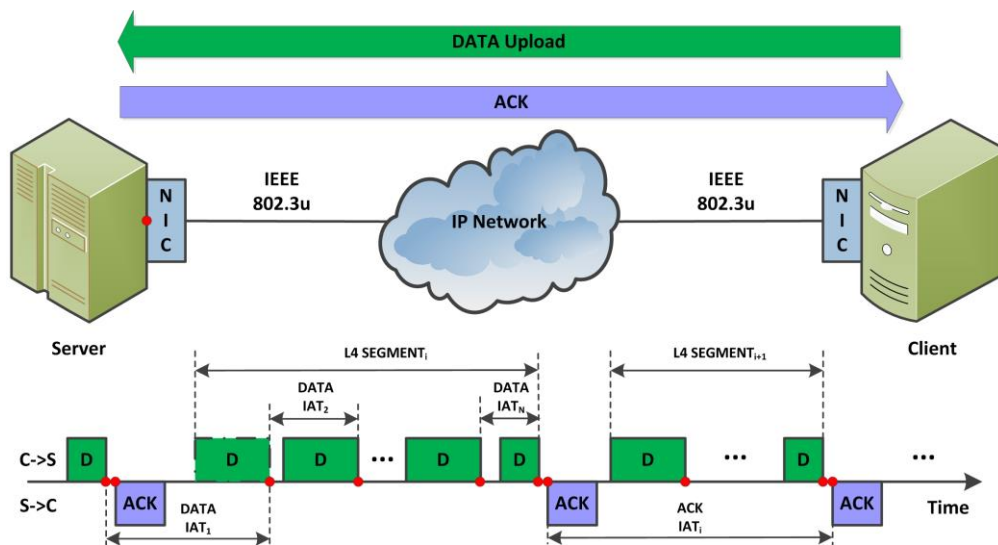
Kulcsszavak: datagram, nagysebességű hálózat, nem-stacionaritás, nem-linearitás, Fourier-transzformáció, Wavelet-transzformáció, Empirical Mode Decomposition, Variational Mode Decomposition.

1. BEVEZETÉS

A TCP szállítási réteg mechanizmusának intenzív adatfolyam szabályozása a folyamatról folyamatra való váltás funkcióját alacsony szinten tartja a kommunikációs útvonalak erőforrásainak (közvetítő csomópontok linkjei, pufferei és processzorai) mérsékelt hatékonyságú használatával [9]. E szigorú kapcsolatorientált mechanizmus indokoltságának hiánya technológiai átgondolást vont maga után a gyakorlatban. A TCP szegmens fejlécének 40 bájtos szigorú határa, az extrém többutas TCP evolúció összetettsége, az új kommunikációs verem lassú adaptációs módszere, és a hiányzó TCP-fejléc-titkosítás miatt ez a közel fél évszázados mechanizmus nem optimális a következő időszakra. Ezek a problémák megbízható szolgáltatásokkal kiegészített fejlesztési lépéseket nyitottak meg az egyszerűbb és gyorsabb UDP

mechanizmus felé [7]. Az UDP egy kapcsolat nélküli szállítási réteg mechanizmus és úgy működik, mint bármely más datagram típusú szolgáltatást a kommunikációs verem többi rétegében: IEEE 802.3 a adatkapcsolati rétegben, Internet Protokoll (IP) a hálózati rétegben és Simple Network Management Protocol (SNMP) az alkalmazási rétegben. Mindezen mechanizmusok előzetes beállítás vagy a megfelelő későbbi nyugtázási fázis nélkül végzik el a protokoll-adatelem (PDU) átvitelét. A PDU visszaigazolás hiánya gyorsabbá teszi az átvitelt, de a rétegszolgáltatást adatfolyam ellenőrzési lehetőség nélkül hagyja. Bármely datagram szolgáltatáson alapuló kommunikációs mechanizmus extra intelligenciát igényel a megbízhatóság és biztonság érdekében. A komplexitás és ezen extra megoldások erőforrás-felhasználásának intenzitása határozza meg a hatékonyságot és az alkalmazások viselkedését. Az UDP teljesítményét kutatók munkái egy új szállítási rétegmechanizmus szükségességét fogalmazták meg, amely egyesíti a TCP és az UDP hálózat előnyeit. A lehetséges megoldást a QUIC biztosítja, amely az Internet Engineering Task Force (IETF), az Internet technológia fejlesztéséért felelős szervezet Request for Comments (RFC) szabványai szerint került nyilvánosságra az utóbbi években.

A klasszikus statisztikai időszerelemzés módszerei képesek értékelni az adatkommunikációs szolgáltatások tulajdonságait, figyelembe véve az adatok helyhez kötött sajátosságait. A gyors kommunikációs folyamat során létrejött valós adatátvitel méréséből származó idősorok gyakran nem-stacionaritást mutatnak. Ezt a tulajdonságot a hálózati igények kiszámíthatatlanul és gyorsan változó adottságai okozzák. Ez számos tényezőnek köszönhető például a hálózati torlódások, a változó forgalmi terhelések és a topológiában bekövetkezett változások. A nagysebességű hálózatokban a forgalom nagysága jelentősen változhat pillanatról pillanatra, ami a hálózati torlódások és a csomagok veszteségarányának változásához vezet. Hasonlóképpen változhat a hálózati topológia állapota, illetve a hálózati csomópontok vagy összeköttetések terhelése, ami az útválasztási nyomvonalak módosulását és késleltetését okozzák. Ezenkívül a nagysebességű hálózatok gyakran kifinomult hálózati protokollokat használnak és olyan mechanizmusokat, mint a torlódásszabályozás és az adatfolyam szabályozás az adatok biztonságos időbeni kézbesítése céljából. Ezek az algoritmusok nem-stacionaritás jellegű változékonyságot okoznak a hálózat dinamikájában.



1. ábra. Visszacatolós UDP forgalom mérésének szemléltetése

A nagysebességű hálózatok nem-lineárisak, mert adattovábbítási képességük nem arányos a bemeneti jel vagy a forgalmi terhelés mennyiségével. A hálózat erősen nem-lineáris jellege azt jelenti, hogy a bemeneti jel kis változásai nagy változásokat eredményezhetnek a hálózat viselkedésében. A nem-linearitás egyik fő oka a nagysebességű hálózatokban a torlódás. Torlódásos hálózatban a csomagok eldobhatók vagy késleltethetők, ami nem-lineáris változásokhoz vezet. Egy másik ok a használt hálózati protokollok komplexitása, amely az erőforrások korlátai miatt áramlásszabályozás alkalmazását igényel. Összességében, a nagysebességű hálózatok nem-stacionárius és nem-lineáris jellege jelentős kihívást jelent a hálózattervezők és a szolgáltatók számára. Emiatt fejlett technikák és algoritmusok használata szükséges a megbízható és hatékony kommunikáció számára. Fejlett modellezési és szimulációs technikákra van szükség a hálózat viselkedésének minél pontosabb előrejelzéséhez, valamint a teljesítmény optimalizálásához. A dolgozat legfontosabb eredményei a következők:

- A csomagkapcsolt datagram általános, nem-lineáris és nem-stacionárius tulajdonságainak igazolása különböző logikai rétegekhez tartozó mechanizmusok esetén (L2-adatkapcsolat: Ethernet; L3-hálózat: IP; L4-transzport: UDP).
- Áttekintést adunk az elemzéshez használt nemlineáris és nemstacionárius módszerekről a hálózati forgalom idősoraira vonatkozóan.
- A felhasználói datagramon alapuló visszacsatolós kommunikációs mechanizmus tulajdonságait ismertetjük valós hálózati forgalomból mért idősorok segítségével.
- A visszacsatolós UDP forgalom klaszterezése olyan független paraméterek alapján történik, mint a maximális adatelem hossz (MTU), az átviteli ráta és a szállítási réteg szegmens mérete.

A dolgozat további részének felépítése a következő: kapcsolódó munkák a szállítóréteg-mechanizmusok hatékonyságáról a második részben található. Ezt követi a nem-stacionárius és nem-lineáris idősorok elemzését lehetővé tevő módszerek bemutatása. Az elemzés eredményeit és értelmezését a negyedik rész tartalmazza. Végül a következtetéseket vonjuk le és megadjuk a kapcsolódó problémák lehetséges folytatási irányát.

2. AZ UDP MECHANIZMUS TELJESÍTMÉNYÉT ELEMZŐ FONTOSABB MUNKÁK ÁTTEKINTÉSE

Az UDP egyszerűsége miatt, kimondottan ezt a protokollt elemző tudományos dolgozatok száma jellemzően nem sok. Ehhez képest nagyszámú olyan elemzés létezik, amely a TCP és az UDP összehasonlítását végzi [16]. Az UDP meghaladja a TCP átviteli hatékonyságát rövid fájlok esetén. A csomagvesztés kompenzációja miatt, azonban a TCP előnyösebb nagyon hosszú fájlok továbbításánál. Erőteljes hálózati terhelés esetén a TCP nagy késleltetéseket okoz, míg az UDP minden esetben adatot képes veszteni. Az adatátviteli szolgáltatás minősége (QoS) érthető okok miatt a két esetben különbözik egymástól. A QoS mindkét szállítási mechanizmust befolyásolja, ám ennek számszerű kimutatása a hálózati alkalmazás típusától függ.

Az átviteli teljesítmény, a késleltetési idő és a csomagvesztési arány szerinti összehasonlítást különböző esetekre ns-2 szimulátor segítségével is elvégezték [4]. Az esetek egymástól a hálózat terhelési szintje, illetve a topológia szempontok alapján különböztek. A TCP és az UDP hatékonyságát a torlódási algoritmus (csak TCP esetén), illetve a pufferek mérete (mindkét mechanizmust) jelentős mértékben befolyásolja. A különböző algoritmusok különböző forgatókönyvek esetén egymástól eltérő módon viselkednek.

IEEE 802,11 vezeték nélküli környezetben UDP felett IP feletti hangátvitelt (VoIP) teszteltek [10]. A rádiós csatorna hatótávolsága, a csomagméret és a csatorna minőségének hatását elemezték VoIP forgalomnál. Azt találták, hogy az UDP teljesítménye és a VoIP minősége erőteljesen függ a szolgáltatói hálózat minőségétől. Minél hosszabb a rádiós csatorna hossza, annál alacsonyabb az UDP teljesítménye, míg a VoIP viszonylag stabil marad.

Nagysebességű vonalakon nagymennyiségű adat továbbításához egy új protokollt javasoltak PA-UDP (Performance Adaptive UDP) néven [6]. Ez a mechanizmus a hálózati adottságokhoz igazodik azáltal, hogy dinamikusan változtatja a csomagok méretét és az átviteli rátát a fogadótól visszaérkező nyugta alapján. Azt találták, hogy nagy terhelésnél a PA-UDP teljesítménye és késleltetése jobb, mint akár a TCP, akár az UDP esetén.

Egy másik datagram alapú továbbfejlesztési javaslatot (UDT: UDP-based Data Transfer) nagy kiterjedésű és nagysebességű környezetre dolgoztak ki [13]. A dolgozat a TCP nagy kiterjedésű hálózatokon tapasztalt korlátait nem kellő mértékben tárgyaló tudományos dolgozatok hiányát pótolja. Az UDT WAN környezetben nagysebességű átvitelt biztosít torlódásvezérléssel, hibaérzékeléssel és visszaállítással, valamint adaptív adatfolyam szabályozással. Az algoritmus visszacsatolást alkalmaz, amely során a fogadó értesíti a küldőt a hálózat állapotáról, így a forrás entitás ennek függvényében szabályozza a küldési rátát.

Az RBUDP (Reliable Blast UDP) protokoll javaslatot ugyancsak nagysebességű WAN környezetbe javasolták további szerzők [14]. A TCP és az UDP mechanizmusok hiányosságának bemutatása után a BPUDP előnyeit tárgyalják: megbízható és méretezhető nagysebességű forgalom biztosítása nagyméretű fájlok esetén. Adaptív torlódásvezérlés segítségével az átviteli rátát szabályozzák a célállomástól visszaérkező nyugta alapján. A csomag-fluxus (packet-blasting) technika segítségével a küldő több csomagot is elküld anélkül, hogy nyugtára várna.

A QUIC (Quick UDP Internet Connections) protokoll biztonsági és teljesítmény elemzését szemléletesen a [19] dolgozatban végezték el annak érdekében, hogy az új protokoll viselkedése és korlátai

megismerhető legyen. A titkosítás, a multiplexelés és a torlódásvezérlés jellemzőit vizsgálták mélyrehatóan. A QUIC biztonság képességét úgy elemezték, hogy különböző támadásoknak tették ki hálózati rétegben, kriptográfiai aspektusban és protokoll szinten. Összehasonlították a QUIC, a TCP és a TLS (Transport Layer Security) mechanizmusokat e tulajdonságok mentén. A QUIC jobban teljesített, mint a TCP, illetve TLS. Kevésbé befolyásolható a teljesítménye a külső hatásokkal és a biztonsági képessége is stabilabb. Ugyanakkor a QUIC adott forgatókönyvek esetén nem garantálja a pártatlan (fairness) kommunikációt és bizonyos forgalmakat előnyben részesít másokhoz képest.

Jelen dolgozat újszerűsége abban áll, hogy nem-lineáris és nem-stacionárius empirikus elemzési módszerekkel tárgyaljuk a datagram kommunikációs veremszerkezet feletti visszacsatolásos adatátvitelt. Értelmezzük Fourier-transzformáció segítségével a jelfelbontásból származó nulla-metszési pontok számát.

3. ALKALMAZOTT MÓDSZER

Mivel az UDP alapú kommunikációs mechanizmusok nem-lineáris és nem-stacionárius folyamatok, ezért áttekintjük a kapcsolódó legfontosabb statisztikai módszereket. Ehhez a Diszkrét Fourier-transzformáció (DFT – Discrete Fourier Transform), a rövid időtartamú Fourier-transzformáció (STFT – Short Time Fourier Transform), a Diszkrét Wavelet-transzformáció (DWT – Discrete Wavelet Transform), az Empirikus Módus-felbontás (EMD – Empirical Mode Decomposition) és a Variációs Módus-felbontás (VMD – Variational Mode Decomposition) módszereket tekintjük ét. Minden esetben az $x(t) \in \mathbb{R}$ idősor N darab egymástól azonos periódusra lévő időpontokban mintavételezett értékét az $x[k]$, $k = [0, 1, \dots, N - 1]$ jelöli. A felsorolt módszerek közös jellemzője, hogy az UDP alapú adatforgalom eseményei közötti, $x(t)$ -vel jelölt időköz (IAT – Inter Arrival Time) idősort egymásra ortogonális, vagy közel ortogonális módusok, illetve maradvány összegére bontják fel. A módusok pontos vagy közelítő ortogonalitása a módszertől függ. Pontos ortogonalitás esetén a maradvány nulla. A módusok k darabszáma ugyancsak a módszertől függ, amelyet zárt formula, vagy közelítő algoritmus segítségével lehet meghatározni. A felbontási összefüggést az alábbi egyenlet adja:

$$x(t) = \sum_{i=1}^k mode_i(t) + res(t). \quad (1)$$

Ezek a módusok diszjunkt, vagy közel diszjunkt frekvencia sávokhoz tartoznak és a felbontási módszertől függenek [2, 18]. A tárgyalt módszerek legfontosabb tulajdonságait az 1. Táblázat hasonlítja össze.

A felbontási módszerek legfontosabb tulajdonságai

1. táblázat

Tulajdonság	DFT	STFT	DWT	EMD	VMD
Időtartomány aspektus	Nincs	Van	Van	Van	Van
Frekvencia-tartomány aspektus	Van	Van	Van	Van	Van
Szűrési aspektus	Globális	Lineáris	Diádikus	Diádikus	Lineáris

A DTF kivételével mindegyik felsorolt módszer idő- és frekvencia-tartományban is dolgozik. Mindez ellenére a DFT-nek vitathatatlanul kiemelkedő jelentősége van a stacionárius módusok jellemzésénél. Az $x(t)$ jel diszkrét Fourier-transzformációját az alábbiak szerint kapjuk [3]:

$$\mathcal{F}\{x[n]\} = X[k] = \sum_{n=0}^{N-1} x[n] e^{\frac{i2\pi kn}{N}}, \quad k = 0, 1, \dots, N - 1, \quad (2)$$

Az $X[k]$ szekvencia inverz diszkrét Fourier-transzformáltja az alábbi:

$$\mathcal{F}^{-1}\{X[k]\} = x[n] = \frac{1}{N} \sum_{k=0}^{N-1} X[k] e^{\frac{i2\pi kn}{N}}, \quad n = 0, 1, \dots, N - 1, \quad (3)$$

Az $X[k]$ abszolútértékének négyzetét, $|X[k]|^2$ periodogramnak nevezzük, amely az $x(t)$ jel energiájának spektrális sűrűsége, spektruma. A $W_N = e^{i2\pi/N} \in \mathbb{C}$ tagok az N -edik komplex egységgyökök, ahol $i = \sqrt{-1}$. Az FFT a DFT gyors algoritmus változata 2^r ($r \in \mathbb{N}$) típusú pontokra olyan szekvenciáknál,

amelyek hossza $N = 2^r$ alakú. A DTF sok fontos tulajdonsága közül itt csak az $x_1[n]$ és $x_2[n]$ szekvenciákra vonatkozó konvolúciós tételt emeljük ki:

$$\mathcal{F}\{x_1[n] * x_2[n]\} = \mathcal{F}\{x_1[n]\} \cdot \mathcal{F}\{x_2[n]\}. \quad (4)$$

Az STFT olyan jelfeldolgozási módszer, amely a jel frekvenciáját időben elemzi [12]. A jelet kisméretű, egymással fedésben lévő intervallumokra bontja, és ezek mindegyikében számolja a Fourier-transzformált értékét. Az ablakfüggvény $w[n]$, amelynek időtartama rövid, azaz értéke 1 a $[0, N_w - 1]$ intervallumon, máshol pedig 0. Az $x(t)$ jel STFT transzformáltját az alábbi formula adja:

$$X[n, k] = \sum_{m=n-(N_w-1)}^n w[n-m] x[n] e^{-i2\pi km/N}. \quad (5)$$

Az STFT a jel frekvencia tartalmát figyeli az idő függvényében, ami lehetővé teszi a jel frekvenciája változásának érzékelését az idő mentén.

A DWT a jelet skálázással különböző hosszúságú frekvencia sávokba tördeli fel és wavelet függvények segítségével az idő és frekvencia tartományban egyszerre vizsgálja [11]. A DWT egy $\Psi(t) \in L^2(\mathbb{F})$ anya waveletre épül, amely az alábbi feltételt teljesíti:

$$\int_0^\infty |\mathcal{F}\{\Psi(k)\}|^2 \frac{dk}{k} < \infty, \quad (6)$$

ahol $\mathcal{F}\{\cdot\}$ a Fourier-transzformáltat jelöli. További követelmény az anya wavelet-tel szemben, hogy a magasabb rendű momentumok értéke nulla legyen, vagyis $M \geq 1$ esetén:

$$\int_{-\infty}^\infty t^m \Psi(t) dt = 0, \quad m = 0, 1, \dots, M. \quad (7)$$

Az elemi wavelet-ek az anya wavelet-ből állíthatók elő az alábbi módon:

$$\psi_{j,k}(t) = \frac{1}{\sqrt{2^j}} \Psi\left(\frac{t}{2^j} - k\right), \quad (8)$$

ahol j , illetve k neve a skálafaktor, illetve az időbeli eltolás. Az $x(t)$ jel DWT transzformáltját az alábbi formula adja:

$$DWT\{x(t)[j, k] = \int_{-\infty}^\infty x(t) \psi_{j,k}(t) dt. \quad (9)$$

Az ebben a formában létrehozott ortogonális wavelet bázis-függvények tipikusan rögzítettek idő- és frekvencia-tartományban, ami ezeket előnyösen teszi alkalmazhatóvá az $x(t)$ jel többfelbontásos szempontok szerinti elemzéséhez.

Az EMD egy komplex jel háttér rezgés-komponenseinek megjelenítését teszi lehetővé [15]. Az $x(t)$ jelet lényeges módus függvények (IMF – Intrinsic Mode Function) halmazára bontja fel. Ezek a jel különböző frekvenciájú rezgő módusai, diszjunkt frekvenciasávok mentén. Az EMD algoritmus az elemzett jel lokális maximum és minimum értékeit azonosítja be, amelyekre harmadrendű spline-ok segítségével burkoló görbét illeszt, amit IMF „részletes” komponensnek hív. Az eredeti jel és a burkológörbe különbsége képezi a „maradvány” függvényt. Az algoritmus következő ciklusában a „maradvány” függvényre ismétli a burkológörbe illesztését, amiből újabb IMF „részletes”, illetve „maradvány” függvényeket állít elő. A folyamat addig ismétlődik, amíg a „maradvány” monotonná válik, ezáltal nem bontható tovább. Az így létrehozott IMF-ek az eredeti $x(t)$ jel csökkenő frekvencia tartományában lévő módus komponensei (ld. 3.1 képlet). Az $x(t)$ jel Hilbert-transzformáltját az alábbi összefüggés alapján határozzuk meg:

$$\mathcal{H}\{x(t)\} = y(t) = \frac{1}{\pi} P \int_{-\infty}^\infty \frac{x(\tau)}{t-\tau} d\tau, \quad (10)$$

ahol P a Cauchy főérték. Minden valós és véges függvénynek, amely az L^p osztályhoz tartozik, véges a Hilbert-transzformáltja. Az $x(t)$ és $y(t)$ jelekből egy analitikus, komplex $z(t) \in \mathbb{C}$ jelet állítunk elő:

$$z(t) = x(t) + iy(t) = a(t)e^{i\theta(t)}, \quad (11)$$

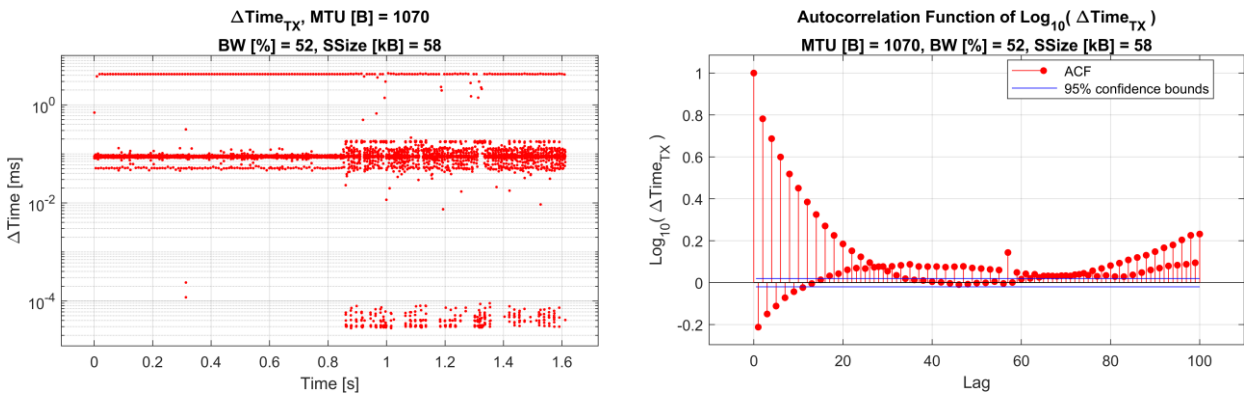
ahol $a(t)$ és $\theta(t)$ rendre az amplitúdó, illetve a fázis függvény:

$$a(t) = [x^2(t) + y^2(t)]^{1/2}, \theta(t) = \arctan\left(\frac{y(t)}{x(t)}\right). \quad (12)$$

Az $x(t)$ jel pillanatnyi körfrekvenciája, $\omega(t)$ az alábbi formula alapján határozható meg:

$$\omega(t) = \frac{d\theta(t)}{dt}. \quad (13)$$

A felbontásból származó módusok k száma jellemzően kicsi ($k \leq 20$) és a módusok csak közelítően ortogonálisak, vagyis a maradvány nem nulla, ugyanakkor monoton függvény. Az EMD módszer eredménye az adatoktól függ, és nem garantál előre definiált bázis függvényeket, ami miatt az algoritmus jól alkalmazható az olyan nem-stacionárius és nem-lineáris jelek elemzéséhez, mint az adathálózatok, orvosi biológiai rendszerek, pénzügyi folyamatok, stb. (ld. 2. ábra).



2. ábra. a) Adat IAT idősor, $x(t)$; b) Az $x(t)$ jel autokorrelációs függvénye.
Fájlviteli eset: $(MTU, BW, SSize) = (170 B, 52\%, 58 kB)$

A VMD az EMD-re hasonlító, olyan jelfeldolgozási technika, amely komplex jel háttérében létező rezgések megjelenítését teszi lehetővé [5]. Az $x(t)$ jelet stacionárius függvények halmazára bontja fel azáltal, hogy optimalizálásra visszavezethető probléma megoldásaként komplex értékű módusokat azonosít be. A ritkasági feltételt az alábbi formula adja:

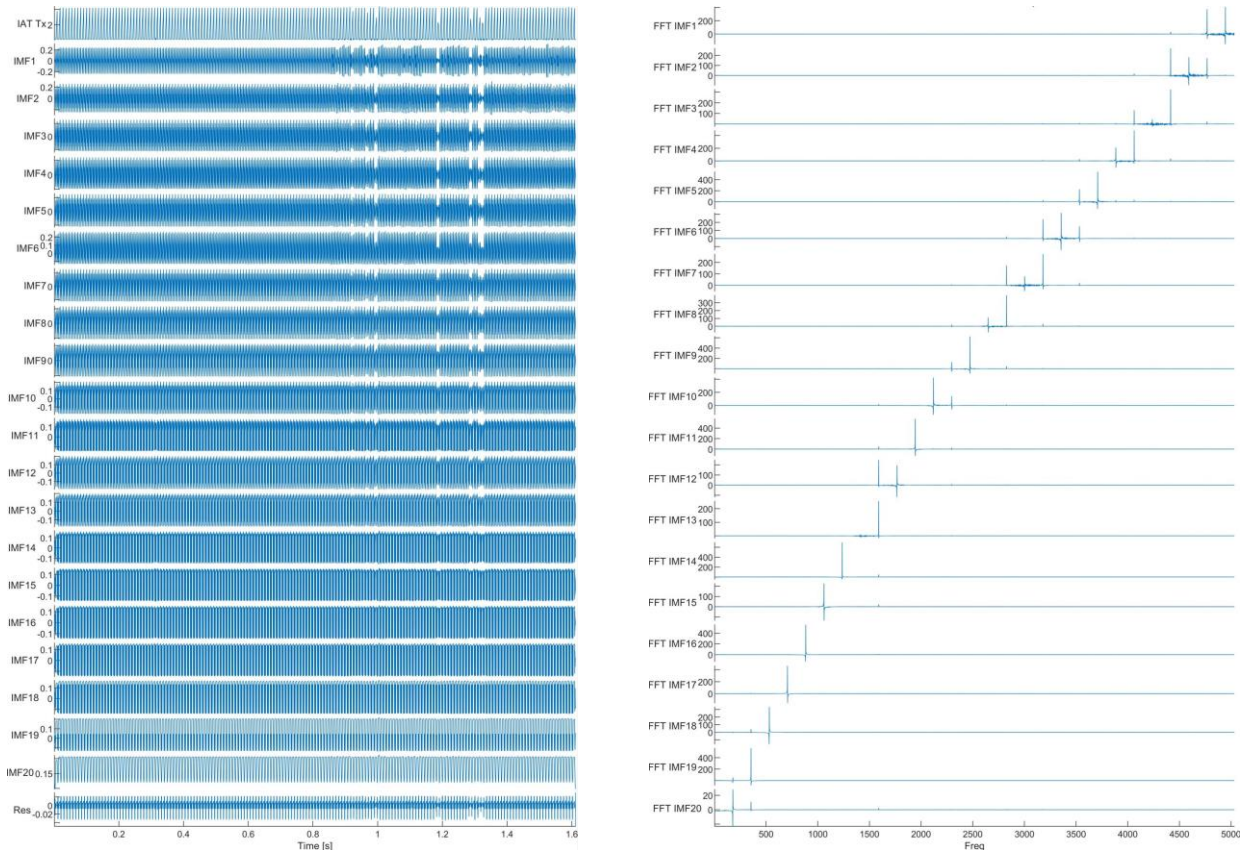
$$\min_{\{mode_k\}, \{\omega_k\}} \left\{ \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{i}{\pi t} \right) * mode_k(t) \right] e^{-i\omega_k t} \right\|_2^2 \right\} \quad (14)$$

ahol ω_k a pillanatnyi körfrekvenciák, ∂_t az idő szerinti parciális derivált jele, $\delta(t)$ a Dirac delta függvény és $\|\cdot\|_2^2$ a Frobenius második norma négyzete. Az optimalizációs feladat megoldása iteratív algoritmus alapján történik, amely a módus függvényeket és a ritkasági feltételt aktualizálja. Az eredő VMD módusok a frekvencia tartalom szerint, it is csökkenő sorrendben állítódnak elő. A 3.1 egyenlet alapján létrehozott módusok ortogonálisak és a maradvány nulla. A VMD módusokat sikeresen lehet alkalmazni olyan nem-stacionárius és nem-lineáris jelek elemzésénél, mint a hang, képek és orvosi biológiai jelek. A VMD az EMD-hez képes jobb zajelnyomási képességgel és gyorsabb konvergenciával rendelkezik.

4. VISSZACSTOLT UDP ALAPÚ ADATFORGALOM MÉRÉSE ÉS TULAJDONSÁGAI

Az UDP protokoll által visszacsatolásos (hurkolt) formában végzett méréseket 1080 alkalommal, ugyanazon adatfájlnak egy kliens és egy szerver gép közötti, különböző környezeti feltételek melletti átvitele során végeztük. A független paraméter-hármas elemei az alábbiak: i) a kliens gép hálózati interfész kártyájának maximális keretmérete (MTU – Maximum Transfer Unit); ii) az alkalmazás szintű átviteli ráta (Bw – Bandwidth); iii) és az UDP szegmens mérete (SSize – Segment Size). Az UDP visszacsatolás úgy valósult meg, hogy a fogadó oldalon lévő szerver az érkező adat szegmensek alapján nyugtát küld vissza a kliensnek. Ez a nyugta állapot-információ két mezőt tartalmaz: a) az UDP szegmenshez tartozó IP csomagok sikeres vagy sikertelen átvétele fogadó szerver oldalon, bináris jelöléssel (0 – IP csomag nem érkezett meg, 1

– IP csomag megérkezett); b) a nyugta küldésének időpillanata a szerver oldalon. Az állapotinformáció a hiányzó IP adatsomagok újraküldéséhez szükséges. Az időbélyeg az alkalmazás szintű átviteli ráta figyeléséhez és vezérléséhez került felhasználásra.



3. ábra. a) Az adat IAT idősorok EMD felbontása ([ms] skálán)
b) Az adat IAT idősorok spektruma ([Hz] skálán).

Az átvitel feltételei: (MTU, BW, SSize) = (1070 B, 52 %, 58 kB)

Az adatfájl mérete 10 MB, a hálózati paraméterek lehetséges értékei pedig a következők:

$$MTU[i] = 510 + i \cdot 80 [B], \quad i = 1, 2, \dots, 12,$$

$$Bw[j] = 37 + j \cdot 6 [\%], \quad j = 1, 2, \dots, 9,$$

$$SSize[k] = -2 + k \cdot 6 [kB], \quad k = 1, 2, \dots, 10.$$

A gyors fájlátviteli idősorok (ld. 2a ábra) nemstacionárius jelleggel rendelkeznek (ld. 2b ábra). Az események közötti időköz (IAT) adatsorokat módusokra felbontva vizsgáltuk több frekvencia skálán és kerestük az időbeni jellemző mintázatokat. Az IAT folyamatok idő-frekvencia tartományban való jellemzéséhez a módusok Hilbert-transzformációját alkalmaztuk. A DC komponens kinyeréséhez a Hilbert-transzformáció következő tulajdonságát használtuk: a $\mathcal{H}\{x(t)\}$ nem más, mint az eredeti $x(t)$ jel $\pi/2$ fázissal való eltolása. Így az eredeti jel és annak kétszer vett Hilbert-transzformáltja összege a DC nélküli, szűrt AC jelet adja. A maradék AC komponensek lehetővé teszik a jel spektrumának és pontos mintázatainak előállítását. Az EMD és a VMD módszereket használtuk az adat IAT és a nyugta IAT idősorok elemzéséhez. A módusok száma korlátos, $k \leq 20$, amelynek megfelelő VMD felbontások a 3a és 4a ábrákon láthatók.

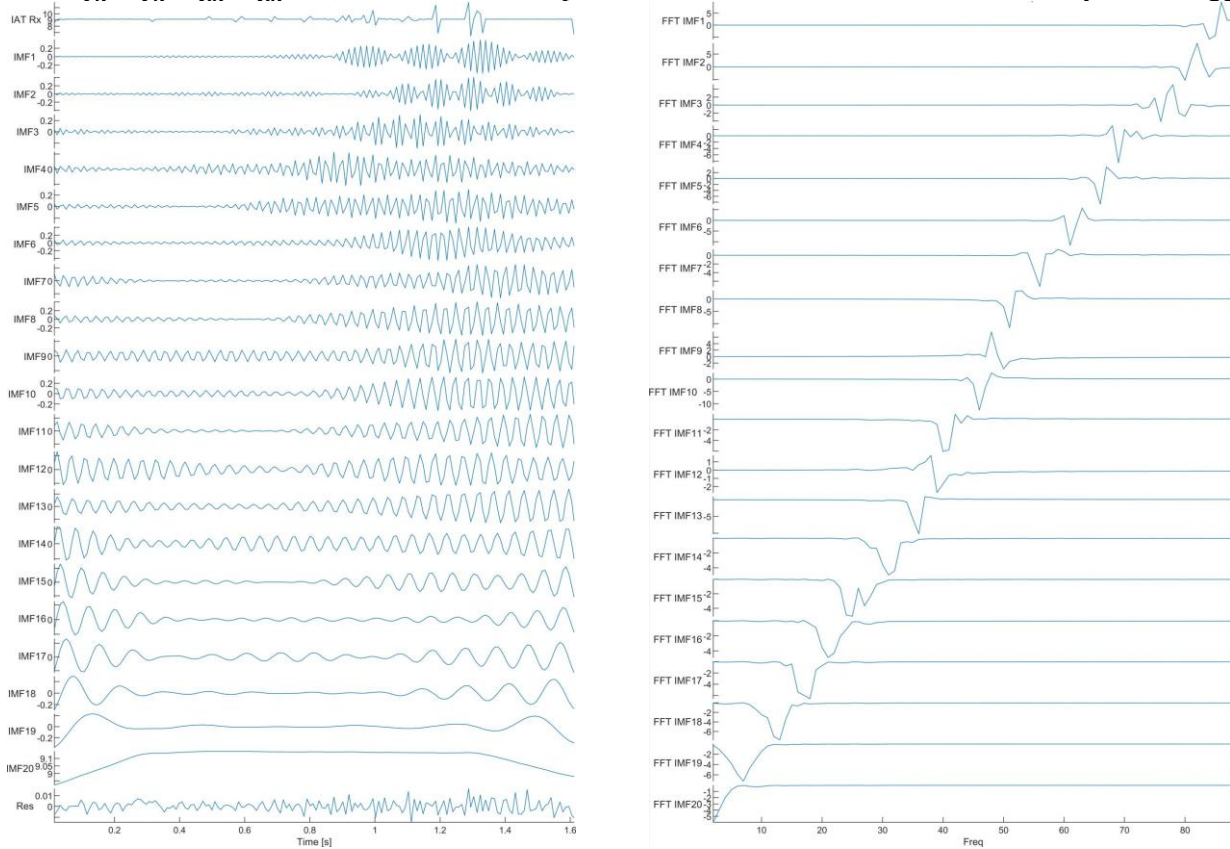
Fontos tulajdonsága a VMD-nek, hogy a módusok Fourier-transzformáltja a frekvenciák szerint diszjunkt intenzitásokat mutat és a függés közel lineáris (ld. 3b és 4b ábrák). Az adatok és a nyugta idősorok EMD módusai nem-lineárisan függenek az időtől és a frekvenciától. A legtöbb tudományos dolgozat ezt a függést diadikus szűrőnek tekinti [1][8][17][20].

Mivel az EMD és a VMD módusok szimmetrikus, független és stacionárius idő függvények, a nullametszéspontok száma a frekvenciákat jól méri. E függést az 5a és 5b ábrák szemléltetik. Megállapítható, hogy úgy az adat IAT, mint a nyugta IAT esetén az EMD, illetve VMD módusok exponenciális, illetve lineáris formában függenek az IMF-ek darabszámától:

$$EMDRate_{Tx}(k) = a_{Tx} \cdot e^{-k \cdot b_{Tx}}, \quad EMDRate_{Rx}(k) = a_{Rx} \cdot e^{-k \cdot b_{Rx}} \quad (15)$$

$$VMDRate_{Tx} = a_{Tx} \cdot k + b_{Tx}, \quad VMDRate_{Rx} = a_{Rx} \cdot k + b_{Rx} \quad (16)$$

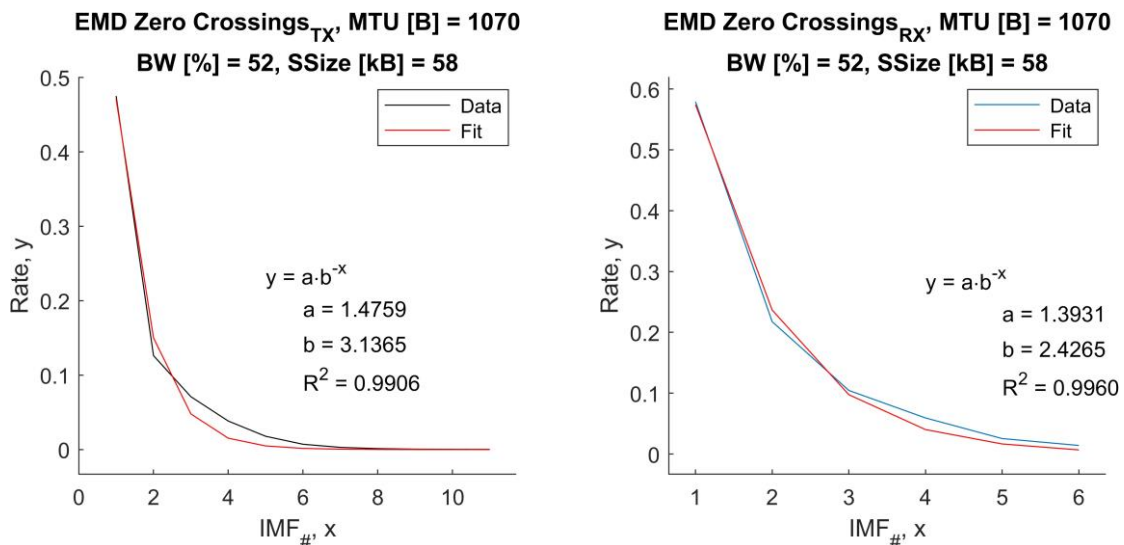
ahol a_{Tx} , b_{Tx} , a_{Rx} , b_{Rx} az (MTU, BW, SSize) paraméter-hármas által definiált hálózati környezettől függ.



4. ábra. a) A nyugta IAT idősorok VMD felbontása (ms skálán)
 b) A nyugta IAT idősorok spektruma (Hz skálán).

Az átvitel feltételei: (MTU, BW, SSize) = (1070 B, 52 %, 58 kB)

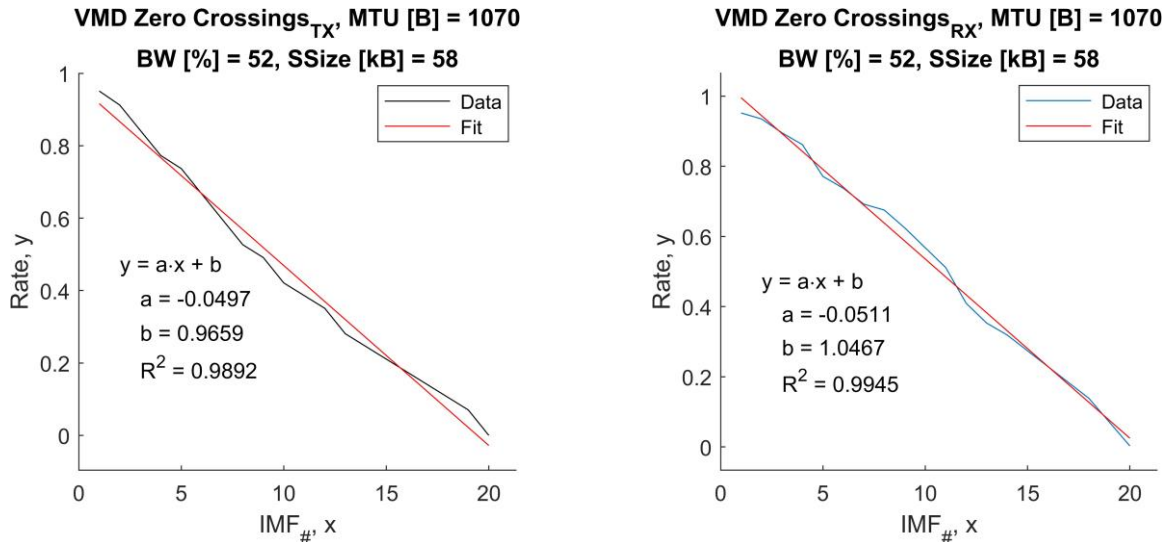
Az MTU, BW és SSize paraméterek tartományának az alsó, illetve felső felét minden esetben L, illetve H indexekkel jelöltük. A három független paraméter ilyen formában $2^3 = 8$ különböző régiót határoz meg a paraméter-hármasok terében: LLL, LLH, LHL, LHH, HLL, HLH, HHL, HHH (ld. 7a és 7b ábrák).



5. ábra. a) Az adat IAT EMD módusainak nulla-metszéspnt száma;
 b) A nyugta IAT EMD módusainak nulla-metszéspont száma.

Az átvitel feltételei: (MTU, BW, SSize) = (1070 B, 52 %, 58 kB)

Fontos kiemelnünk egy megfigyelt adottságot: az EMD segítségével előállított nulla-metszési pontok rátája ugyan exponenciális függvény, de az alapja nem 2 (ld. 7a ábra). A legtöbb esetben az exponens alapja a küldött adatok esetén $I_{Tx} = (1,7 \dots 3.5)$ intervallumban van, átlagosan 2,6 értékkel, míg a nyugták esetén $I_{Rx} = (2 \dots 2,8)$ intervallumban van, átlagosan 2,3 értékkel. Ez módosítja az EMD módszer diádikus szűrőnek tudományosan kinyilvánított hatását. A VMD lineáris illesztéssel kapott egyenesének meredeksége $a^* = 0,05 = \frac{1}{20} = \frac{1}{k_{max}}$, amelyet a VMD módusok maximális száma, $k \leq 20$ okozott (ld. 7b ábra).



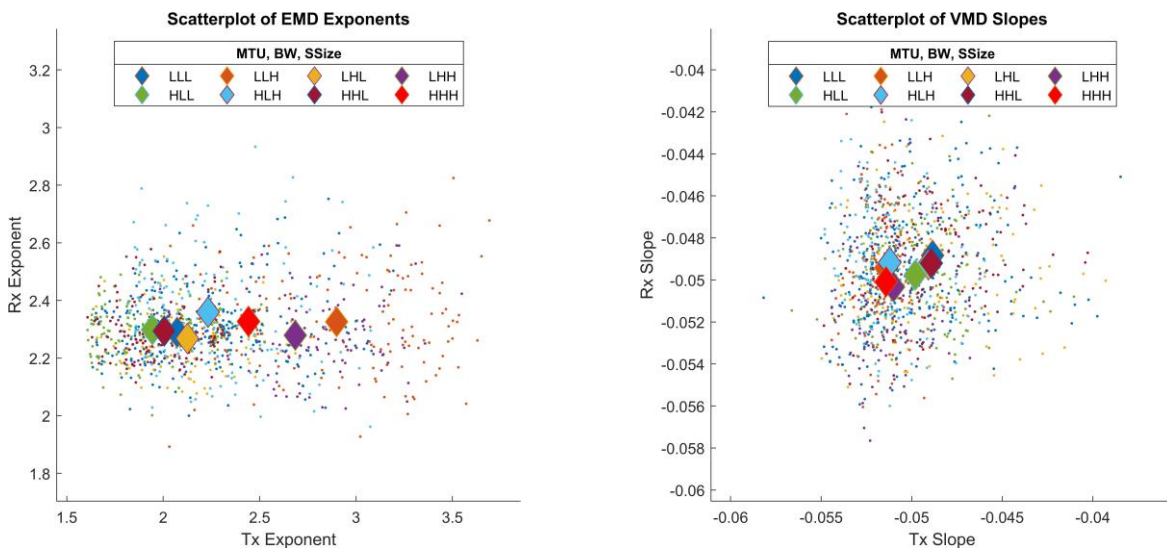
6. ábra. a) Az adat IAT VMD módusainak nulla-metszéspnt száma;
b) A nyugta IAT VMD módusainak nulla-metszéspont száma.
Az átvitel feltételei: (MTU, BW, SSsize) = (1070 B, 52 %, 58 kB)

A paraméter alterek súlypontjának helyzete alapján három mintázat-klaszter alakult ki:

$$[C_1, C_2, C_3] = [xxL, LxH, HxH], \quad (17)$$

ahol az „x” karakter az adott paraméter semleges hatását jelenti (pl. $xxL = LLL \vee LHL \vee HLL \vee HHL$, és \vee a VAGY logikai operátort jelöli). A nulla-metszési ráta illesztési paramétereinek klaszterezési tulajdonsága alapján kijelenthetjük, hogy visszacsatolós UDP adatátvitel esetén három forgalmi mintázat-csoport különböztethető meg a maximális keretméret, átviteli ráta és szegmensméret alapján:

C1) Alacsony szegmensméret esetén a másik két paraméternek nincs hatása a forgalomra. Az IP csomagok darabolásának gyakorisága elenyésző az adatküldőnél. A forgalom sztochasztikus jellegű és az átviteli ráta közel van alkalmazás szintű Bw paraméternél beállított értékhez.



7. ábra. a) Az adat és nyugta EMD exponensek szórásdiagramja;
b) Az adat és nyugta VMD meredekségek szórásdiagramja.

C2) Az MTU alacsony, és az UDP szegmensméret magas értéke esetén a B_w paraméternek elhanyagolható hatása van a forgalomra. Ilyenkor erőteljes darabolás történik az IP rétegben, mivel nagyméretű adatszegmensek küldése történik a rövid Ethernet keretekben. Az UDP forgalom jellegét a darabolás határozza meg. Az adat-, illetve nyugtakeretek számának aránya ebben az esetben a legnagyobb.

C3) A maximális keretméret és a szegmensméret magas értéke esetén a B_w paraméternek nincs befolyása a forgalomra. A darabolás intenzitása az IP rétegben mérsékelt, az átviteli ráta maximális érték alatti ingadozása ugyanakkor elhanyagolható.

5. ÖSSZEFOGLALÁS

10 MB méretű adatfájl szerverre való feltöltését végeztük el visszacsatolós UDP mechanizmus segítségével Ethernet lokális hálózatban. Három független környezeti változót használtunk: maximális keretméret, alkalmazás szintű átviteli ráta és szállítási rétegbeli szegmens méret. Átfogó áttekintést adtunk a nem-lineáris és nem-stacionárius elemzési módszerekről, amelyek a felöltött adat, valamint visszaküldött nyugta elemek beérkezési időközeiből kialakított idősorok felbontását végzik. Megállapítottuk, hogy az Empirikus Módus-felbontás, valamint a Variációs Módus-felbontás által létrehozott módusok különböző frekvencia tartományokhoz tartoznak, lehetővé téve ezen stacionárius függvények elemzését a nulla-metszéspontok száma segítségével. A nulla-metszéspont ráta Empirikus Módus-felbontás esetén exponenciális, míg Variációs Módus-felbontásnál lineáris eloszlású. Az EMD a b -alapú szűrők általánosított esete, ahol $b > 2$. A feltöltési adatforgalomnak UDP alapú visszacsatolás esetén három forgalom mintázata van az $(MTU, B_w, SSize)$ paraméter-hármas függvényében. A visszafele küldött nyugták forgalma b -alapú szűrő, ahol $b^* = 2,3$. Ez ellentétes a sok tudományos dolgozatban az EMD-vel kapcsolatosan eddig közölt diádikus szűrő hatással, ahol $b = 2,0$. További elemzések szükségesek a b -alapú szűrők b értékének függését illetően az Empirikus Módus-felbontásnál. Hasonló módon, a Variációs Módus-felbontás meredekségének a értékét is tovább célszerű értelmezni.

KÖSZÖNETNYILVÁNÍTÁS

Ezt a munkát a Debreceni Egyetem QoS-HPC-IoT Laboratórium és a TKP2021-NKTA-34 projekt támogatta. A TKP2021-NKTA-34 számú projekt a Kulturális és Innovációs Minisztérium Nemzeti Kutatási Fejlesztési és Innovációs Alapból nyújtott támogatásával, a TKP2021-NKTA pályázati program finanszírozásában valósult meg.”

IRODALMI HIVATKOZÁSOK

- [1] M. B. Abd-el-Malek, S. S. Hanna: Using filter bank property to simplify the calculations of Empirical Mode Decomposition, *Communications in Nonlinear Science and Numerical Simulation* 62 (2018), pp. 429–444, issn: 1007-5704, doi: <https://doi.org/10.1016/j.cnsns.2018.02.035>.
- [2] R. Bazi, T. Benkedjough, H. Habbouche, S. Rechak, N. Zerhouni: A hybrid CNNBiLSTM approach-based variational mode decomposition for tool wear monitoring, *The International Journal of Advanced Manufacturing Technology* 5 (2022), pp. 3803–3817, doi: 10.1007/s00170-021-08448-7.
- [3] P. Bloomfield: *Fourier analysis of time series: an introduction*, John Wiley & Sons, 2004.
- [4] F. T. AL-Dhief, N. Sabri, N. A. Latiff, M. Abbas, A. Albader, M. A. Mohammed, R. N. AL-Haddad, Y. D. Salman, M. Khanapi, et al.: Performance comparison between TCP and UDP protocols in different simulation scenarios, *International Journal of Engineering & Technology* 7.4.36 (2018), pp. 172–176.
- [5] K. Dragomiretskiy, D. Zosso: Variational Mode Decomposition, *IEEE Transactions on Signal Processing* 62.3 (2014), pp. 531–544, doi: 10.1109/TSP.2013.2288675.
- [6] B. Eckart, X. He, Q. Wu: Performance adaptive UDP for high-speed bulk data transfer over dedicated links, in: *2008 IEEE International Symposium on Parallel and Distributed Processing*, IEEE, 2008, pp. 1–10.
- [7] Fast transport layer protocol: QUIC, Official web site of IETF QUIC Working Group, url: <https://quicwg.org/>.
- [8] P. Flandrin, G. Rilling, P. Goncalves: Empirical mode decomposition as a filter bank, *IEEE Signal Processing Letters* 11.2 (2004), pp. 112–114, doi: 10.1109/LSP.2003.821662.
- [9] Z. Gál, G. Kocsis, T. Tajti, R. Tornai: Performance evaluation of massively parallel and high speed connectionless vs. connection oriented communication sessions, *Advances in Engineering Software* 157-158 (2021), p. 103010, issn: 0965-9978, doi: <https://doi.org/10.1016/j.advengsoft.2021.103010>.
- [10] S. Garg, M. Kappes: An experimental study of throughput for UDP and VoIP traffic in IEEE 802.11 b networks, in: *2003 IEEE Wireless Communications and Networking*, 2003. WCNC 2003. Vol. 3, IEEE, 2003, pp. 1748–1753.

- [11] J. Gilles: Empirical Wavelet Transform, *IEEE Transactions on Signal Processing* 61.16 (2013), pp. 3999–4010, doi: 10.1109/TSP.2013.2265222.
- [12] D. Griffin, J. Lim: Signal estimation from modified short-time Fourier transform, *IEEE Transactions on Acoustics, Speech, and Signal Processing* 32.2 (1984), pp. 236–243, doi: 10.1109/TASSP.1984.1164317.
- [13] Y. Gu, R. L. Grossman: UDT: UDP-based data transfer for high-speed wide area networks, *Computer Networks* 51.7 (2007), pp. 1777–1799.
- [14] E. He, J. Leigh, O. Yu, T. A. DeFanti: Reliable blast UDP: Predictable high performance bulk data transfer, in: *Proceedings. IEEE International Conference on Cluster Computing, IEEE, 2002*, pp. 317–324.
- [15] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, H. H. Liu: The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, *Proceedings of the Royal Society of London A: mathematical, physical and engineering sciences* 454.1971 (1998), pp. 903–995.
- [16] H. K. Rath, A. Karandikar: Performance analysis of TCP and UDP-based applications in a IEEE 802.16 deployed network, in: *2011 The 14th International Symposium on Wireless Personal Multimedia Communications (WPMC)*, 2011, pp. 1–5.
- [17] G. Rilling, P. Flandrin: One or Two Frequencies? The Empirical Mode Decomposition Answers, *IEEE Transactions on Signal Processing* 56.1 (2008), pp. 85–95, doi: 10.1109/TSP.2007.906771.
- [18] R. C. Sharpley, V. Vatchev: Analysis of the Intrinsic Mode Functions, *Constructive Approximation* 24 (2006), pp. 17–47, doi: 10.1007/s00365-005-0603-z.
- [19] M. Soni, B. S. Rajput: Security and performance evaluations of QUIC protocol, in: *Data Science and Intelligent Applications: Proceedings of ICDSIA 2020*, Springer, 2021, pp. 457–462.
- [20] Y. Yang, J. Deng, D. Kang: An improved empirical mode decomposition by using dyadic masking signals, *Signal, Image and Video Processing* 9.6 (2015), pp. 1259–1263, issn: 1863-1711, doi: 10.1007/s11760-013-0566-7.