

## FARIMA MODELI SAOBRAĆAJA KOMUNIKACIONIH MREŽA

Miodrag Bakmaz, Bojan Bakmaz  
Saobraćajni fakultet u Beogradu

**Sadržaj:** *Proučavanjem visokokvalitetnih saobraćajnih merenja ustanovljeno je da saobraćaj u mrežama velikih brzina pokazuje sebi-sličnost, odnosno zavisnost u dugom opsegu, pa se ne može tretirati klasičnim modelima. Novija merenja realnog saobraćaja su pokazala postojanje zavisnosti u dugom i kratkom opsegu kod saobraćajnih zapisa. Zbog toga se od modela očekuje da opišu obe zavisnosti istovremeno. Ovde je razmatran FARIMA( $p, d, q$ ) model kao jedan od boljih modela sa tom sposobnošću.*

**Ključne reči:** *FARIMA, mrežni saobraćaj, multimedijalni saobraćaj, komunikacione mreže, sebi-sličnost, LRD, SRD, NGN*

### 1. Uvod

Saobraćajni modeli igraju značajnu ulogu u karakterizaciji i analizi realnog mrežnog saobraćaja i mrežnih performansi. Značajnim pomeranjem ka mrežama velikih brzina, koje omogućuju integrisane servise sa pouzdanim kvalitetom servisa (QoS), uloga preciznih saobraćajnih modela u dizajniranju i simuliranju postaje sve važnija. U osnovi su stacionarni ili nestacionarni. Stacionarni modeli mogu se razmatrati kroz dve klase opsega zavisnosti: kratku i dugu (*Short (Long) Range Dependence*, SRD, LRD). U modele zavisnosti u kratkom opsegu spadaju Markovi procesi, konvencionalne raspodele, Pareto raspodela i regresioni modeli, a karakterišu se korelacionom strukturom koja je značajna za relativno malo realizacija. Saobraćajni modeli zavisnosti u dužem opsegu imaju značajnu korelaciju za više realizacija, a najpoznatiji su: frakcionalna ARIMA (*Fractional Autoregressive Integrated Moving Average*, FARIMA), frakcionalni Gausov šum (FGN), frakcionalno Braunovo kretanje (FBM), diskretno Braunovo kretanje, agregacija ON – OFF izvora visoke varijabilnosti [1].

Savremene studije, bazirane na vrlo kvalitetnom merenju podataka, ukazuju na postojanje *self similar* (sebi sličnog) procesa, kao realističnog matematičkog sredstva, koje omogućava karakterisanje statističkog ponašanja saobraćaja u mrežama za prenos podataka (LAN, MAN, WAN), ISDN signalizacionom D kanalu, kod TCP/IP i WWW saobraćaja, u VBR kodiranim video sekvencama. Statistička analiza prikupljenih podataka pokazala je drastičnu razliku svojstava saobraćaja paketskih mreža u odnosu na postojeće klasične modele. Rezultujući saobraćaj, pri agregaciji raznih individualnih

izvora i nižem mrežnom opterećenju, nije postajao “izgladeniji”, kako se očekivalo, nego je odražavao svojstva *bursty* (burnog, praskavog) saobraćaja. Ova opservacija se pokazala vrlo važnom, jer je uticala na preispitivanje nekih ključnih pretpostavki koje su činjene u procesu planiranja mrežnih performansi. *Burst*-evi se mogu opisivati statistički, koristeći *self similarity*, svojstvo povezano sa fraktalnošću, osobinom pojavljivanja istog objekta, nezavisno od skale na kojoj se vrši posmatranje.

Osnovno svojstvo *self similar* procesa je LRD (zavisnost u dugom opsegu), koja se može okarakterisati parametrom Hurst-a. Osim egzaktnog LRD procesa, zavisnosti u obliku frakcionalnog Gausovog šuma, razmatraju se i asimptotski *self similar* procesi, kao što su: “haotična mapa”, *heavy tailed* ON-OFF modeli, GARIMA, FARIMA. Poslednji od modela je generalno i fleksibilno rešenje, koje omogućava fitovanje za SRD i LRD autokorelacione funkcije. Direktni konkurenti su mu WIG (*Wavelet Independent Gaussian*) model, koji je efektivniji u proračunavanju, ali daje lošije rezultate, i MWM (*Multifractal Wavelet Model*), za koji se smatra da je pogodniji po većini svojstava [11].

U nastavku rada izložene su osnovne osobine autoregresivnih modela, da bi posebna pažnja bila usmerena na svojstva *self similar* procesa i Frakcionalni ARIMA model, koji su poslednjih godina predmet većeg broja istraživanja.

## 2. Autoregresivni modeli

Modeli regresije eksplicitno definišu sledeću slučajnu promenljivu u sekvenci, preko prethodne, unutar specificiranog vremenskog prozora. Osnovni je linearni autoregresivni model reda  $p$ , sa uobičajenom oznakom  $AR(p)$ , kod koga se slučajna veličina u vremenu izražava linearnom kombinacijom ranijih ekvidistantnih opservacija. Model ima sledeći oblik:

$$X_t = \sum_{i=1}^p \phi_i X_{t-i} + \varepsilon_t \quad (1)$$

gde je  $\varepsilon_t$  Gausov beli šum normalne raspodele i srednje vrednosti nula (te slučajne promenljive odražavaju grešku i često se nazivaju reziduali),  $\phi_i$  su realni brojevi, a  $X_t$  određene korelisane slučajne promenljive, a ponekad se u model uključuje i konstanta. Od značaja je operator zaostajanja (realizacije, pomeranja unazad)  $B$  (*lag, backshift operator*), definisan preko  $X_{t-1} = BX_t$ , tako da se može formirati polinom  $\Phi(B)$  i dobiti pregledniji oblik  $AR(p)$  procesa

$$\Phi(B) = 1 - \sum_{i=1}^p \phi_i B^i, \quad \Phi(B)X_t = \varepsilon_t. \quad (2)$$

Potrebna su neka ograničenja vrednosti parametara da bi model ostao stacionaran. Tako za  $AR(1)$  model, kod koga je proces izražen sa  $X_t = \Phi X_{t-1} + \varepsilon_t$ , postoji stacionarnost kovarijanse za  $|\Phi| < 1$ .

Estimacija parametara može se vršiti korišćenjem regresione analize. Autokorelaciona funkcija (ACF) predstavlja odnos kovarijanse (korelacionog momenta) i varijanse,  $R_k = E[(X_{t-k} - \mu)(X_t - \mu)]/E[(X_t - \mu)(X_t - \mu)] = cov(X_{t-k}, X_t)/v(X_t)$ , pa se množenjem jednačine (1) sa  $X_{t-k}$ , usrednjavanjem i delenjem sa varijansom dobija

$$R_k = \sum_{i=1}^p \phi_i R_{k-i} + \sigma_\varepsilon^2 \delta_k, \quad (3)$$

gde je  $k = 0, \dots, p$ , što obezbeđuje  $p + 1$  jednačinu.  $\sigma_\varepsilon$  je standardna devijacija šuma, a  $\delta_k$  Kronekerova delta funkcija. Ovo je sistem Yule-Walker-ovih jednačina, pomoću kojih se određuju  $\phi_i$ , kao i  $\sigma_\varepsilon$ . Odabir prikladnog reda  $p$  može se automatizovati, recimo korišćenjem AIC (*Akaike Information Criterion*), a u praksi je najčešće dovoljno  $p \leq 3$ .

Model pokretnih srednjih vrednosti reda  $q$ , MA( $q$ ) (*Moving Average*), je oblika

$$X_t = \varepsilon_t - \sum_{j=1}^q \theta_j \varepsilon_{t-j}, \quad (4)$$

gde su  $\theta_j$  parametri modela, a  $\varepsilon_{t-j}$  beli šum u trenucima  $t - j$ .

Autoregresivni model pokretnih srednjih vrednosti reda ( $p, q$ ), ARMA( $p, q$ ) (*Autoregressive Moving Average*), je sa stacionarnim podacima i oblika

$$X_t = \sum_{i=1}^p \phi_i X_{t-i} + \varepsilon_t - \sum_{j=1}^q \theta_j \varepsilon_{t-j}, \quad (5)$$

koji se, uz pomoć operatora realizacije (pomeranja unazad)  $B$ , može ekvivalentno predstaviti kao  $\Phi(B)X_t = \theta(B)\varepsilon_t$ . Ovde AR deo modelira rekorelacioni efekt a MA služi za fitovanje korelacije od drugih realizacija, odnosno ograničenje uticaja prethodnih slučajnih ispada (*shock*) na budućnost. Estimacija parametara ARMA modela je komplikovanija nego kod AR modela, pošto estimacija MA parametara zahteva rešavanje sistema nelinearnih jednačina, ili korišćenje spektralne faktorizacione tehnike.

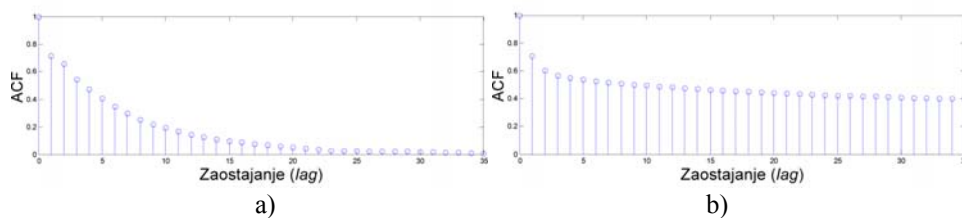
Autoregresioni integrisani model pokretnih sredina, ARIMA( $p, d, q$ ), predstavlja proširenje ARMA modela, a dobija se uzimanjem u obzir činjenice da polinom  $\Phi(B)$  ima  $d$  korenova jednakih jedinici, dok ostali leže van jediničnog kruga. Konkretnije,  $d$  predstavlja nivo diferenciranja kojim se transformišu nestacionarni podaci u stacionarne serije i ARIMA svodi na ARMA model.

Oblik modela je  $\Phi(B)\Delta^d X_t = \theta(B)\varepsilon_t$ , gde su  $\Phi(B) = 1 - \sum \phi_i B^i$  i  $\theta(B) = 1 - \sum \theta_i B^i$ , polinomi reda  $p$  i  $q$  respektivno, a  $\Delta$  je operator diferenciranja definisan kao  $\Delta = 1 - B$ , odnosno  $\Delta X_t = X_t - X_{t-1} = (1 - B)X_t$ . Član  $\varepsilon_t$  predstavlja proces belog šuma za diskretne vrednosti  $t$  i ima konstantnu varijansu. Operator zaostajanja  $B$  ima svojstva da je  $B^{\pm s} X_t = X_{t \mp s}$  i  $B^s c = c$ , za  $c$  konstantno.

ARIMA se koristi za modeliranje homogenih nestacionarnih vremenskih serija. Ako vremenska serija ispoljava nestacionarnost u nivou, može se modelirati kao ARIMA( $p, 1, q$ ), a ako je nestacionarna u trendu, modelira se kao ARIMA( $p, 2, q$ ).

#### 4. Osobine zavisnost u dugom opsegu

Zavisnost u dugom (širokom) opsegu (LRD) je statistički fenomen koji pobuđuje mnogo pažnje u oblasti komunikacionog saobraćaja poslednjih petnaestak godina. Vremenska serija poseduje LRD ako ima autokorelacije koje su izražene preko cele vremenske skale, opadajući po hiperboličkom zakonu, za razliku od zavisnosti u kratkom opsegu (SRD), gde je opadanje autokorelacija eksponencijalno (slika 1). LRD utiče na vremena čekanja pri opsluzi paketa i najčešće se karakteriše  $H$  (Hurst) parametrom, koji ima vrednost u granicama od 0,5 do 1. Postoji veliki broj raznih statistika koje se mogu koristiti za estimaciju parametra  $H$ .



Slika 1. Autokorelacione funkcije a) SRD i b) LRD procesa

Za vremensku seriju  $X_t$ , koja je slabo stacionarna (*weakly*), odnosno ima konačnu srednju vrednost i kovarijansu zavisnu samo od operatora realizacije između dve tačke u seriji, kaže se da je LRD, ako  $\sum R_k$  divergira. Često se pretpostavlja specifična funkcionalna forma oblika  $R_k \sim C_\rho k^{-\beta}$ , gde je  $C_\rho$  veće od 0, a  $\beta$  između 0 i 1. Parametri  $\beta$  i  $H$  su u relaciji  $\beta = 2 - 2H$ .

LRD se može konstatovati za slabo stacionarnu vremensku seriju  $X_t$  u frekvencijskom domenu, ako se spektralna gustina ponaša kao  $f(\lambda) \sim C_f |\lambda|^{-1+\beta}$ , kada frekvencija  $\lambda \rightarrow 0$ , za neko  $C_f > 0$  i neko realno  $\beta$  između 0 i 1. LRD odražava prisustvo statističkog *self-similarity*, pa se pri analizi telekomunikacionog saobraćaja često koristi taj termin.

Spektralna gustina ACF  $R_k$  definiše se kao

$$f(\lambda) = \frac{\sigma^2}{2\pi} \sum_{k=-\infty}^{\infty} R_k e^{ik\lambda}, \quad (6)$$

gde je  $i$  imaginarna jedinica, a definicija spektralne gustine proizilazi iz Wiener-Kinchine teoreme.

LRD se u vremenskom domenu manifestuje visokim stepenom korelacije između tačaka podataka na određenim rastojanjima. U frekvencijskom domenu to se manifestuje kao značajan nivo snage na frekvencijama bliskim nuli. LRD je komplikovan za statističku obradu. U vremenskom domenu meri se samo na višim realizacijama, gde je dostupno nekoliko sponova i gde je greška merenja najveća. U frekvencijskom domenu merenje se vrši na frekvencijama bliskim nuli, takođe gde je najteže vršiti merenje. Dok se  $H$  parametar dobro matematički definiše, postoje velike teškoće pri merenju u realnim okolnostima.

Pojam *self similar* uveden je sedamdesetih godina prošlog veka, pri opisu statističkih svojstava u raznim naučnim oblastima, dok među prve značajnije radove, koji sagledavaju taj fenomen na primeru Eternet saobraćaja, spada rad [2]. U odnosu na Puasonov proces koji, ako se posmatra na finijoj vremenskoj skali, izgleda *bursty*, a agregiran na grubljoj skali deluje izglađeno, *self similar* (fraktalni) proces, kada se agregira na grublju vremensku skalnu, održava *bursty* karakter.

*Self-similar* stohastički proces može se u osnovi definišati na sledeći način:

Kontinualni stohastički proces  $X$  je *self similar*, sa parametrom  $H$ , ako je, za neki pozitivni faktor širenja  $c$ , proces na novoj vremenskoj skali  $ct$ ,  $c^{-H}X_c$ , jednak po raspodeli originalnom procesu  $X$ .

Za diskretno vreme, neka je  $X = (X_t, t = 0, 1, 2, \dots)$  stacionarni slučajni proces u širem smislu, zavisan u dugom opsegu, sa srednjom vrednošću  $\mu$ , varijansom  $\sigma^2$  i

autokorelacionom funkcijom  $R_k$ . Novi proces  $X^{(m)}$  dobija se usrednjavanjem originalne serije  $X$  po nepreklapajućim blokovima veličine  $m$ . Svaki  $X^{(m)}$  je u širem smislu stacionarni slučajni proces sa autokorelacionom funkcijom  $R^{(m)}(k)$ . Znači,  $X$  je *self similar* (egzaktno drugog reda), sa parametrom  $H = 1 - \beta/2$ , ako agregirani procesi imaju istu autokorelacionu strukturu kao  $X$ , odnosno,  $R^{(m)}(k) = R_k$ , za  $m = 1, 2, \dots$ , dok za varijansu važi  $v(X^{(m)}) \sim m^{-\beta}\sigma^2$ ,  $0 < \beta < 1$ , za  $m \rightarrow \infty$  (varijansa opada sporije od  $m^{-1}$ , što je karakteristika klasičnih saobraćajnih modela). Autokorelaciona funkcija je oblika

$$R^{(m)}(k) = R_k = \frac{1}{2}((k+1)^{2H} - 2k^{2H} + |k-1|^{2H}), \quad k \geq 0. \quad (7)$$

Kraće, proces  $X$  je *self similar* ako se agregirani procesi  $X^{(m)}$  ne razlikuju od  $X$  u pogledu svojstava prvog i drugog reda. Proces je asimptotski *self similar* sa parametrom  $H = 1 - \beta/2$ , ako  $R^{(m)}(k) \rightarrow R_k$ , kada  $m \rightarrow \infty$ .

*Self similar* svojstvo se manifestuje preko nekoliko ekvivalentnih oblika: sporog opadanja varijanse, zavisnosti u dužem opsegu, nedegenerativne autokorelacije, Hurst-ovog efekta. Autokorelaciona funkcija sledi hiperboličnu zakonitost (zakon opadanja snage), koja opada sporije od eksponencijalne i za ovaj proces, kao LRD, ukupna suma autokorelacija je beskonačna.

Aktuelni mrežni saobraćaj je agregacija raznih saobraćajnih tokova različitih karakteristika i odražava *self similar* svojstva. Ako su *self similar* parametri dva *self similar* procesa isti, njihova agregacija biće egzaktno *self similar*, a ako su različiti, agregirani proces biće asimptotski *self similar*, sa  $H$  parametrom jednakim maksimalnom od parametara individualnih izvora.

Hurst-ov efekt se karakteriše preko Hurst-ovog parametra, već navedenog  $H$ , koji je formiran tako da obuhvati stepen *self-similarity* u datom empirijskom zapisu. Za empirijsku vremensku seriju  $X_t$  ( $t = 1, \dots, n$ ), sa srednjom vrednošću uzorka  $X(n)$  i varijansom  $S^2(n)$ , može se definisati  $R/S$  statistika (*Rescaled Adjusted Range*), data sa  $R(n)/S(n)$ , pri čemu važi

$$R(n) = \max \left\{ \sum_{i=1}^k (X_i - X(n)), 1 \leq k \leq n \right\} - \min \left\{ \sum_{i=1}^k (X_i - X(n)), 1 \leq k \leq n \right\}. \quad (8)$$

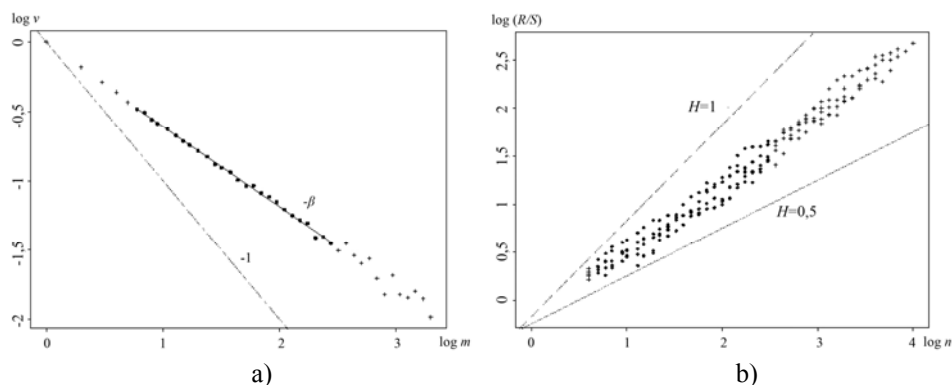
Empirijski je ustanovljeno da mnoge prirodne vremenske serije zadovoljavaju relaciju za matematičko očekivanje  $E[R(n)/S(n)] \sim cn^H$ , za veliko  $n$ , gde je  $c$  konstanta.  $H$  uzima vrednosti od 0,5 do 1. Mnogi realni procesi imaju vrednost  $H$  oko 0,73, za čisto slučajan proces je  $H = 0,5$ , dok je kod egzaktno *self similar* procesa  $H = 1$ .

Razni postupci se koriste za testiranje *self-similarity* i Hurst-ovog efekta. Dijagrami “varijansa – vreme” (*V-T plot*), odnosno  $\log v(X^{(m)})$  prema  $\log m$ , baziraju se na činjenici da varijansa agregiranog procesa, fitovana metodom najmanjih kvadrata, opada linearno u odnosu na  $m$ , pri log-log skalama i gradijentu  $\beta$  (slika 2a).

Dijagram  $\log R/S$  u odnosu na  $\log n$  ( $R/S$  plot), kao grafička analiza empirijskih podataka, jasno potvrđuje *self similar* prirodu generisanih sekvenci (slika 2b).

Periodogram odražava opadanje spektra snage serije sa približavanjem frekvencije nuli i ima gradijent niske frekvencije  $\gamma$ , pri čemu je  $H = (1 + \gamma)/2$  [14].

Koriste se i dijagrami varijanse za razne nivoe agregacije. Tehnika analize “talasića” (*wavelet*) i lokalni Whittle-ov estimator su novije tehnike koje dobro funkcionišu u komparativnim studijama i imaju jaku teorijsku podršku.



Slika 2. Dijagrami a) “varijansa – vreme” i b)  $\log R/S$

Postoje stohastički modeli koji poseduju *self similar* karakteristike, poput frakcionalnog Gausovog šuma, (FGN, *Fractional Gaussian Noise*), agregacije više ON/OFF izvora sa sporijim opadanjem raspodele (*heavy tails*), FARIMA, GARMA. Problem je generisanje saobraćajnih trasa direktno po modelima, za šta je razvijeno više algoritama, kao što su: FFT (*Fast Fourier Transform*), kojom se generiše FGN, RMD (*Random Midpoint Displacement*) algoritam, koji rekurzivno proširuje generisane sekvence, SRA (*Successive Random Addition*), sličan prethodnom, sa povećanom stabilnošću generisane sekvence i Hosking-ov, zasnovan na FARIMA(0,  $d$ , 0). FGN koristi tri parametra (srednja vrednost, varijansa,  $H$  parametar), što je dovoljno za uspešno modeliranje LAN saobraćaja, dok je FARIMA proces sa četiri ili pet parametara pogodan je za opisivanje VBR video saobraćaja sa prihvatljivom preciznošću. Anderson-Darling-ov test valjanosti fitovanja pokazuje raspodelu sekvenci uzoraka i pogodniji je od testa Kolmogorov-Smirnov-a, kada testira u odnosu na specificiranu normalnu raspodelu.

## 5. Frakcionalna ARIMA

Frakcionalna ARIMA (FARIMA) je generalizacija standardnog ARIMA procesa preko necelobrojnog parametra diferenciranja  $d$ ,  $-0,5 < d < 0,5$ , a za predstavljanje svojstava modela stacionarnog procesa kao LRD, odnosno asimptotskog *self similar* procesa, koristi pozitivne vrednosti za  $d$ . Operator  $\Delta$  može se izraziti kao

$$\Delta^d = (1 - B)^d = \sum_{k=0}^{\infty} \binom{d}{k} (-B)^k, \quad \binom{d}{k} = \frac{d!}{k!(d-k)!} = \frac{\Gamma(d+1)}{\Gamma(k+1)\Gamma(d-k+1)}, \quad (9)$$

gde  $\Gamma(x)$  predstavlja Gama funkciju. Prosesi kod kojih je  $d$  negativno, do  $-0,5$ , nazivaju se srednji memorijski procesi, čija praktična implikacija još nije razjašnjena.

Za slučaj stacionarnog FARIMA(0,  $d$ , 0) procesa autokorelaciona funkcija ima oblik

$$R_k = \frac{\Gamma(1-d)\Gamma(k+d)}{\Gamma(d)\Gamma(k+1-d)}, \quad R_k \approx \frac{\Gamma(1-d)}{\Gamma(d)} k^{2d-1}, \quad k \rightarrow \infty. \quad (10)$$

Od posebnog značaja je veza parametra diferenciranja sa Hurst-ovim parametrom,  $H = d + 0,5$  i gradijentom  $\beta = 1 - 2d$ .

Ako se koristi Gausov beli šum, tada FARIMA ima Gausovu raspodelu. Ovo ograničava njenu pogodnost na modeliranje procesa koji imaju aproksimativno Gausovu raspodelu. Gausov beli šum se koristi zato što je suma više Gausovih slučajnih promenljivih takođe Gausova slučajna promenljiva, što takođe važi za opštije Stabilne (*stable*) slučajne promenljive.

FARIMA(0,  $d$ , 0) proces, kao najprostiji i fundamentalan, predstavlja model frakcionalnog diferencirajućeg šuma (FDN), sa osobinama sličnim frakcionalnom gausovom šumu (FGN), pri čemu, umesto  $H$ , parametar  $d$  odražava stepen LRD.

FARIMA( $p$ ,  $d$ ,  $q$ ) proces može se posmatrati kao ARMA( $p$ ,  $q$ ) proces aktiviran sa FDN. Relacija za FARIMA proces može se napisati kao  $X_t = \Phi^{-1}(B)\theta(B)\Delta^{-d}\varepsilon_t = W_t Y_t$ , gde  $Y_t = \Delta^{-d}\varepsilon_t$  predstavlja FDN proces.

FARIMA proces pokriva i LRD i SRD korelacionu strukturu, s tim što poseduje nedostatak da je proračun parametara modela vrlo spor i komplikovan. Postoje problemi sintetičkog generisanja zapisa i ako se parametri ne odaberu pažljivo, model se ne prilagođava realnim saobraćajni situacijama. Koristi se za modeliranje VBR video saobraćaja, generisanje dolaznih tokova, simulacionu analizu modela čekanja, kao i predikciju multimedijalnog saobraćaja.

## 6. Fitovanje aktuelnog saobraćaja

Radi dobijanja odgovarajućeg modela za dati saobraćajni zapis, potrebno je iterativna primena tri osnovna koraka, dok se ne dobije zadovoljavajuće rešenje. Prvo, vrši se identifikacija modela, odnosno određivanje prikladnih vrednosti parametara:  $d$  - nivoa diferenciranja,  $p$  - autoregresije,  $q$  - pokretne sredine, drugo, vrši se estimacija parametara, treće, sprovodi se dijagnostička provera.

Opšta procedura za fitovanje FARIMA modela u saobraćajne trase sastojala bi se od sledećih koraka:

1. Preprocesiranje merenih saobraćajnih zapisa radi dobijanja vremenske serije sa srednjom vrednošću nula.

2. Određivanje aproksimativne vrednosti  $d = H - 0,5$ . Parametar  $H$  može se dobiti nizom dijagrama i postupaka, od kojih su neki prethodno navedeni.

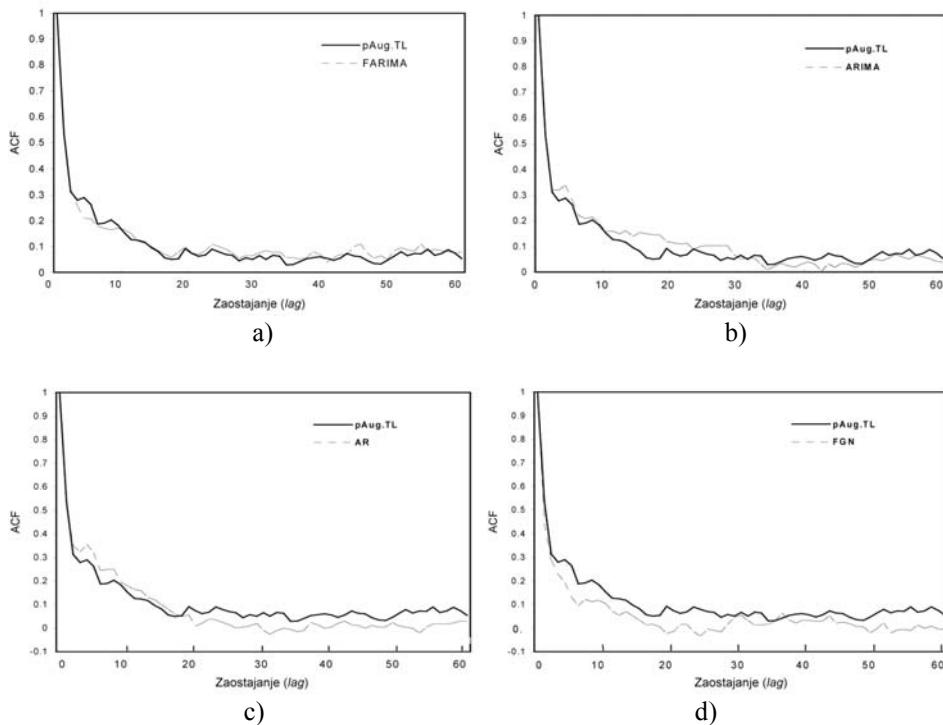
3. Vršiti se frakcionalno diferenciranje  $X_t$  i FARIMA problem se svodi na ARMA model, odnosno seriju  $W_t = \Delta^d \varepsilon_t$ .

4. Određuju se  $p$  i  $q$  koristeći neki od poznatih postupaka za fitovanje ARMA modela i odabira njihova najbolja kombinacija, u saglasnosti sa identifikacijom modela i dijagnostičkom proverom.

5. Posle iterativnog određivanja  $p$  i  $q$ , estimiraju se parametri  $\Phi_i, \theta_i, \sigma^2$ .

Estimaciona procedura parametara FARIMA modela, radi fitovanja korelacionih osobina vremenskih serija, zasniva se na principima Hosking-ovog algoritma i može se naći u radu [6]. Parametar  $d$  se nalazi na osnovu  $H$  iz  $V-T$  ili drugog odgovarajućeg dijagrama. ARMA( $p, q$ ) komponenta iz FARIMA( $p, d, q$ ) procesa se dobija primenom Box-Jenkins algoritma za estimiranje parametara  $\Phi(B)$  i  $\theta(B)$ .

U [2] i [6] analiziran je, pored ostalog, *Bellcore* zapis pAug.TL i dobijeno je  $H = 0,8$ , odnosno  $d = 0,3$ , a posle frakcionalnog diferenciranja  $d$  je opalo na  $0,03$ , što znači da je LRD značajno umanjen. Na slici 3 prikazani su ACF zapisa, kao i ACF fitovani modelom FARIMA(1,  $d$ , 1), ARIMA(10, 1, 0), AR(8), i FGN. FGN model ne može dobro da fituje autokorelacionu funkciju sa aktuelnim zapisom u kratkom opsegu, dok AR model nema dobro fitovanje u dugom opsegu. FARIMA i ARIMA model omogućuju dobro fitovanje, pri čemu ARIMA zahteva mnogo više parametara.



Slika 2. ACF zapisa pAug.TL i fitovani modelima a) FARIMA, b) ARIMA, c) AR i d) FGN

Mada FARIMA modeli imaju niz prednosti, postoji problem velike proračunske kompleksnosti i dugotrajne procedure. Interesantno je da kada se fituje samo deo zapisa, parametri FARIMA modela ostaju isti [6]. Ova osobina omogućuje uprošćavanje



procedure fitovanja, redukuje vreme za izradu saobraćajnih modela, kao i modelovanje u realnom vremenu.

U radu [7] detaljno je izložena procedura za estimaciju parametara FARIMA modela i fitovanje statističkih karakteristika i ponašanja aktuelnog saobraćaja u redu čekanja. Za Ethernet LAN saobraćaj (Bellcore), posle filtriranja (diferenciranja) LRD komponente, primene Box-Jenkins algoritma za estimaciju pronađena je najveća značajnija vrednost parametra AR modela,  $p = 20$ . Za sekvencu videokonferencijske sesije ustanovljeno je najveće značajno  $p = 17$ . Analizirano je ponašanje ACF i V-T dijagram za razne slučajeve FARIMA( $p, d, q$ ), pri čemu je  $p$  uzimano do najveće značajne vrednosti,  $d$  je ustanovljeno iz zapisa, a za  $q$  je korišćena vrednost 0. Ustanovljen je značajan uticaj parametra  $p$  na izgled dijagrama.

Dalje, ponašanja raznih redova čekanja za sintetičke FARIMA zapise zavise samo od korelacionih svojstava, odnosa LRD/SRD, a ne i od marginalne raspodele, koja je Gausova za sve zapise, što nije slučaj kod aktuelnih saobraćaja. Analize pokazuju da pri manjim baferima AR(1) pokazuje veće gubitke od FARIMA(0,  $d$ , 0), za šta se razlozi mogu pronaći u V-T dijagramu, a sa povećanjem bafera LRD komponenta dolazi do izražaja i FARIMA(0,  $d$ , 0), kao i FARIMA(1,  $d$ , 0), produkuju gore performanse.

Kod aktuelnog saobraćaja, recimo LAN Ethernet, veličina bafera od 714 ćelija odgovara maksimalnom kašnjenju od 1 s, pri maksimalnom razmatranom opterećenju i sintetičke sekvence zadovoljavaju sa performansama i koef prostijeg FARIMA(0,  $d$ , 0) modela. S druge strane, za videokonferencijski saobraćaj maksimalno kašnjenje koje se toleriše je 100 ms i odgovara baferu veličine 510 ćelija. FARIMA(0,  $d$ , 0) sada ne zadovoljava u pogledu gubitaka ćelija, za razliku od FARIMA(17,  $d$ , 0). Uz veću toleranciju mogla bi se prihvatiti i FARIMA(1,  $d$ , 0).

## 7. Zaključak

Razne analize potvrđuju činjenicu da je FARIMA( $p, d, q$ ) proces pogodan saobraćajni model, koji je prihvatljiv za manje parametre  $p$ , pošto su SRD korelacije od bitnog uticaja na performanse čekanja. MA komponenta, odnosno parametar  $q$ , se najčešće smatra suvišnom. FARIMA model je u mogućnosti da obuhvati osobenosti realnog saobraćaja zahvaljujući svojstvu da istovremeno modelira LRD i SRD ponašanje vremenskih serija. Model ima primenu u dizajniranju mreža, menadžmentu, predikciji saobraćaja, kao i upravljanju mrežom baziranom na merenjima, a za očekivanje je da bude siguran alat NGN inženjeringa, zbog povoljnih iskustava u dosadašnjem korišćenju i broja parametara sa kojima se može manipulirati.

## Literatura

- [1] V. Frost and B. Melamed, "Traffic Modelling for Telecommunications Networks," *IEEE Commun. Mag.*, pp. 70-80, Mart 1994.
- [2] W. Leland, M. Taqqu, W. Willinger and D. Wilson, "On the Self-Similar Nature of Ethernet Traffic," *IEEE/ACM Trans. Networking*, pp. 1-15, Feb. 1994.
- [3] V. Paxson and S. Floyd, "Wide-Area Traffic: The Failure of Poisson Modeling," *IEEE/ACM Trans. Networking*, pp. 226-254, June 1995.

- [4] A. Adas, "Traffic Models in Broadband Networks," *IEEE Communications Magazine*, pp. 82-89, July 1997.
- [5] H. Michiel and K. Laevens, "Teletraffic Engineering in a Broad-Band Era," *Proceedings of the IEEE*, pp. 2007-2032, Dec. 1997.
- [6] F. Xue, J. Liu, Y. Shu, L. Zhang, O. W. W. Yang, "Traffic Modeling Based on Farima Models," *Proceedings of the 1999 IEEE Canadian Conference on Electrical and Computer Engineering*, <http://www.sfu.ca/~fxue/farima.ps>
- [7] M. Corradi, R. G. Garroppo, S. Giordano, M. Pagano, "Analysis of F-ARIMA Processes in the Modelling of Broadband Traffic," *Proc. of ICC 2001 - IEEE Intern. Conference on Communs.*, no. 1, pp. 964-968, Helsinki, Finland, June 11-14, 2001.
- [8] S. Floyd, V. Paxson, "Difficulties in Simulating the Internet," *IEEE/ACM Trans. on Networking*, no 4, pp. 392-403, Aug. 2001.
- [9] N. Sadek, A. Khotanzad, T. Chen, "ATM Dynamic Bandwidth Allocation Using F-ARIMA Prediction Model," *Proc. of ICCCN 2003, The 12th Intern. Conference on Computer Communications and Networks*, pp. 359 - 363, 20-22 Oct. 2003.
- [10] C. G. Dethe and D. G. Wakde, "On the Prediction of Packet process in Network Traffic Using FARIMA Time-Series Model," *J. Indian Inst. Sci.*, 84, pp. 31-39, Jan-Apr. 2004.
- [11] A. du Plessis, M. A. Van Wyk, "Modeling of Ethernet Traffic with Self-Similar Processes," <http://www.etd.rau.ac.za>.
- [12] M. Bakmaz, B. Bakmaz, "Mogućnosti autoregresionih modela telekomunikacionog saobraćaja," *Postel 2006*, decembar 2006.
- [13] O. I. Sheluhin, S. M. Smolskiy, A. V. Osin, *Self-Similar Processes In Telecommunications*, JohnWiley & Sons, 2007.
- [14] B. Bakmaz, *Kvalitet servisa u heterogenim bežičnim mrežama*, Zadužbina Andrejević, Beograd, 2008.

**Abstract:** *The studies of high quality traffic measurements have revealed that traffic in high-speed networks exhibits self similarity, i.e. long-range dependence that can't be captured by traditional models. Recent real traffic measurements found the co-existence of both long-range and short-range dependence in traffic traces. Therefore, models are required to describe both long-range and short-range dependence simultaneously. We consider FARIMA(p, d, q) model as one of good models with this capability.*

**Keywords:** *FARIMA, network traffic, multimedia traffic, communication networks, self similarity, LRD, SRD, NGN*

## **THE FARIMA TELETRAFFIC MODELS IN COMMUNICATION NETWORKS**

Miodrag Bakmaz and Bojan Bakmaz