

Vremenske transformacije struktura zavisnosti multivarijabilnih signala

Dragana Bajić, Tamara Škorić, Marko Mozetić

Univerzitet u Novom Sadu, Departman za energetiku, elektroniku i telekomunikacije
Fakultet tehničkih nauka
Novi Sad, Srbija
tamara.ceranic@gmail.com, dragana.bajic@gmail.com,
markomozetic95@gmail.com

Nina Japundžić-Žigon

Univerzitet u Beogradu, Institut za farmakologiju, kliničku farmakologiju i toksikologiju
Medicinski fakultet
Beograd, Srbija
nina.japundzic@gmail.com

Sažetak— U radu se analizira transformacija struktura zavisnosti multivarijabilnih signala iz probabilističkog u vremenski domen. Strukture zavisnosti su višedimenzionalne gustine verovatnoće koje prikazuju jačinu povezanosti posmatranih signala. Izražavaju se statičkim modelom poznatim kao ‘gustina kopule’. Dinamički prikaz jačine povezanosti je nov i originalan vremenski niz koji omogućava analizu fluktuacija jačine međusobne povezanosti više signala tokom vremena. Takva analiza je značajna za medicinske aplikacije, stoga su kao ilustrativni primer analizirani signali sistolnog pritiska i pulsog intervala laboratorijskih životinja. Ključni element transformacije je estimacija gustine kopule pa je rad fokusiran na poređenje različitih metoda transformacije.

Ključne reči—strukture zavisnosti; gustina kopule; procena gustine; kardiovaskularni vremenski nizovi; (key words)

I. UVOD

Strukture zavisnosti povezanih vremenskih nizova vizualizuju se putem funkcije gustine raspodele poznate kao gustina kopule. Sama kopula je kumulativna funkcija raspodele, definisana krajem šeste decenije dvadesetog veka [1]. Kopula parametar θ procenjuje jačinu uzajamne povezanosti posmatranih vremenskih nizova. Ovaj parametar može da se izvede iz klasičnih statističkih alata kao što su Kendalov i Spirmanov koeficijent [2], ali jedino kopula omogućava vizualizaciju jačine povezanosti.

Gustina kopule je definisana u apstraktnom domenu kopula verovatnoća pa ne omogućava prikaz fluktuacija međusobne povezanosti signala tokom vremena. U [3] je uvedena metoda transformacije jačine povezanosti proizvoljnog broja signalâ u jedan vremenski niz. Time se generiše jedan signal koji prikazuje fluktuacije intenziteta sprege između izvornih signalâ od odmerka do odmerka (odnosno, ako se radi o kardiovaskularnim signalima, za svaki otkucaj srca). Time je omogućena primena alata razvijenih za jednodimenzionalne signale. To je značajno jer postoji niz tehnika dizajniranih za analizu jednodimenzionalnih, eventualno dvodimenzionalnih signala (na primer aproksimativna i uzorkovana entropija, *ApEn* i *SampEn*) koje ne mogu da se primene za analizu jačine povezanosti većeg broja signala tokom vremena.

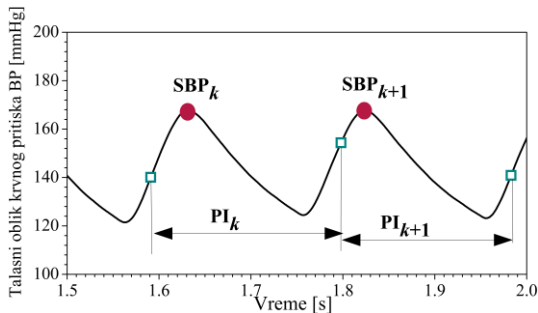
Za transformaciju je potrebno proceniti gustinu raspodele koja odgovara svakoj tački u višedimenzionalnom kopula domenu. Procene gustine raspodele se često trivijalizuju, ali problem je kompleksan i posvećene su mu čitave knjige (na primer [4]). U ovom radu porede se klasične i savremene metode procene gustine verovatnoće povezanosti bivarijantnih kardiovaskularnih vremenskih nizova. Signali koji se koriste opisani su u poglavlju II a radi se o sistolnom pritisku (SBP) i pulsnom intervalu (PI) laboratorijskih pacova kojima su date različite doze vazopresinskih antagonista. Poglavlje III ukratko prikazuje matematičke osnove kopule i opisuje klasične i savremene metode procene gustine. Rezultati uporedne analize prikazani su u poglavlju IV nakon kojeg slede zaključna razmatranja.

II. EKSPERIMENTALNI PODACI

Signali su snimani na odraslim laboratorijskim Wistar pacovima, mužjacima, težine 300–350 g. Tokom eksperimenta pacovi su držani u kontrolisanim laboratorijskim uslovima: temperatura $22 \pm 2^\circ\text{C}$; relativna vlažnost vazduha 60 – 70%; ciklusi dana i noći smenjivani na po 12h. Eksperimen je ispoštovao Direktivu 2010/63/EU, nacionalni Zakon o dobrobiti životinja 2009/6/RS i Pravilnik 2010/RS. Protokol je odobrio Etički odbor Univerziteta u Beogradu (licenca n°323-07-10519/2013-05/2).

Signal talasnog oblika krvnog pritiska (BP) siman je tako što je u abdominalnu aortu pacova implantiran senzor pritiska TL11M2-PA-C50-PX-DSI. Pacovima su ubrizgavani vazopresinski antagonisti V1a i V2 putem kanile prema protokolu iz [5].

Signali sistolnog pritiska (SBP) i pulsog intervala (PI) izvedeni su iz BP signala kao vremenski niz lokalnih maksimuma i vremenski niz interval između tačaka maksimalnog porasta BP (Slika 1). Artefakti su otklonjeni poluautomatski: jedan deo je otklonjen ADA filtrom dizajniranim za PI signale [6] a zaostale greške su otkrivene i ispravljene vizuelnim pregledom. Na osnovu doza koje su pacovi primili, formirano je 5 grupa signala: kontrolna grupa (bez leka), V1a-100 ng, V1a-500 ng, V2-100 ng i V2-500 ng.



Slika 1. Talasni oblik krvnog pritiska BP sa prikazanim SBP i PI odmercima

III. KOPULA I PROCENE GUSTINE

A. Empirijska kopula

Kopula je definisana Sklarovom teoremom [1]: Neka su x_1, \dots, x_D međusobno zavisni kontinualni signali (ukupno D signala, svaki dužine N) sa združenom funkcijom raspodele $F(x_1, \dots, x_D)$ i proizvoljnim marginalnim funkcijama raspodele $F_1(x_1), \dots, F_D(x_D)$. Signali se transformišu putem tzv. probabilističke integralne transformacije (PIT) u signale u_1, \dots, u_D čije su funkcije raspodele $U_1(u_1), \dots, U_D(u_D)$ uniformne i međusobno povezane novom združenom funkcijom raspodele $C(u_1, \dots, u_D)$, a to je kopula.

Probabilistička integralna transformacija (PIT) [7] konvertuje signale putem njihove funkcije raspodele: $u_i = F_i(x_i)$. Lako se dokazuje [7] da je rezultat uniformno raspodeljen na segmentu $[0, 1]$.

Kopula je definisana nad domenom $[0, 1]^D$ a formalno se definiše kao [1]:

$$C(u_1, \dots, u_D) = F(F_1^{-1}(u_1), \dots, F_D^{-1}(u_D))$$

$$C(F_1(x_1), \dots, F_D(x_D)) = F(x_1, \dots, x_D). \quad (1)$$

Dokazano je i farmakološki potvrđeno da je Frankova kopula najpogodnija za analizu međusobne zavisnosti kardiovaskularnih signala [8] i da je njena dinamička varijanta povezana sa barorefleksnim sekvencama i antisekvencama [9]. Stoga se Frankova kopula koristi i u ovom radu.

Kopula i njena gustina su teorijski koncepti. Za praktičnu analizu signala značajna je empirijska gustina kopule. Procenjuje se tako što se svaki od D izvornih signala konvertuje putem PIT, a zatim se od uniformnih signala dobijenih konverzijom formira združena gustina raspodele.

Sledeći korak je formiranje vremenskog signala koji prikazuje fluktuacije jačine zavisnosti posmatranih signala. Postiže se tako što se vremenskoj poziciji svake D -torki izvornih signala (u vremenskom domenu) pridruži odgovarajuća empirijska gustina kopule (iz probabilističkog domena). Iz toga sledi da je procena gustine kopule ključni korak u formiranju novog vremenskog niza.

B. Metode procene gustine kopule

Testirali smo četiri metode: histogram, lokalnu gustinu, Voronoi dekompoziciju i Markovljeve tranzicione matrice.

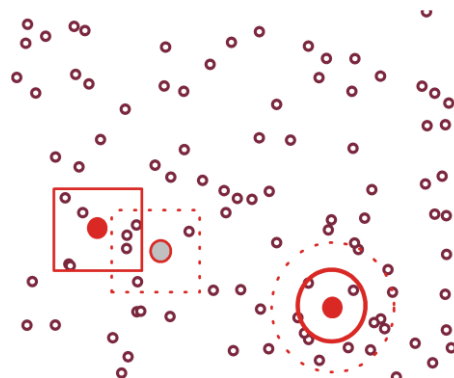
Histogram je jednostavna klasična mera procene gustine verovatnoće. Opseg $[0, 1]$ svakog od D uniformnih signala se podeli na K segmenata tako da postoji ukupno K^D ćelija u kojima se nalaze D -dimenzionalne tačke. Histogram predstavlja broj tačaka signala u svakoj od ćelija. Kada se podeli ukupnim brojem tačaka N , predstavlja procenjenju gustinu kopule.

Suviše veliki broj nivoa K dovodi do nepouzdanu procene verovatnoće zbog neispunjenja Jeruhimovih kriterijuma [10]; suviše mali broj smanjuje rezoluciju. Pored toga, sve tačke unutar jedne ćelije imaju istu gustinu, tako da je dobijeni vremenski niz gustina tačaka (tj. jačinâ njihovih međuzavisnosti) diskretan.

Lokalna gustina je takođe jednostavan klasičan metod. Gustina se procenjuje za svaku tačku posebno, prebrojavanjem tačaka koje su njenom neposrednom okruženju. Normalizovan broj susjednih tačaka predstavlja lokalnu gustinu. Ovakav pristup ima nedostatke. Pre svega, okruženja se preklapaju pa jedna tačka može da bude u susedstvu različitih tačaka, kao što je prikazano na slici 2. To unosi pristrasnost u procenu. Dodatno, procena verovatnoće i njene rezolucije zavisi od definisanog prečnika okruženja, što je takođe prikazano na slici 2. Konačno, suma tačaka za normalizaciju nije N (zbog preklapanja), pa procenjena mera, iako je srazmerna 'gustini', suštinski ne odgovara strožoj definiciji 'verovatnoće'.

Vrlo slična je metoda najbližih suseda, gde se precizira broj susjednih tačaka (takođe parametar!) i određuje rastojanje do najudaljenije [4]. Procena putem kernela ima još više stepena slobode jer svaki kernel ima svoje osobine a svakom kernelu može da se bira varijansa, tako da je rezultat pristrasan jer može da se podešava po potrebi [4].

Procena putem Markovljevih matrica spada u nove metode, koristi se i za potrebe koncepta 'uskog grla teorije informacija'



Slika 2. Problemi procene lokalne gustine (lokalne jačine zavisnosti) za $D=2$ signale. Kvadrati prikazuju problem preklapanja – iste tačke pripadaju različitim lokalnim okruženjima. Kružnice prikazuju problem parametara – procenjena gustina zavisi od praga, tj. od prečnika kružnice.

[11]. D -dimenzionalne tačke u prostoru gustine kopule predstavljaju stanja konačnog, diskretnog i potpuno povezanog aperiodičnog i nesvodljivog (tj. regularnog) Markovljevog lanca. Što su tačke u prostoru bliže, to je verovatnoća prelaza između odgovarajućih stanja veća. Ako je d_{ij} rastojanje između tačke i i tačke j u kopula domenu, tada je tranziciona verovatnoća između odgovarajućih stanja

$$p_{ij} = \frac{e^{-\gamma d_{ij}}}{\sum_{j=1}^N e^{-\gamma d_{ij}}} \quad (2)$$

U jednačini (2) γ predstavlja parametar. Traziciona matrica $\mathbf{\Pi}=[p_{ij}]$ je stohastička po vrstama (desno stohastička) sa vektorom stabilnih verovatnoća stanja $\boldsymbol{\pi} = [\pi_1 \dots \pi_N]$. Verovatnoća posmatranog stanja π_i , $i=1, \dots, N$, veća je ako su verovatnoće da sistem iz ostalih stanja prelazi u to stanje velike. Na osnovu jednačine (2), verovatnoće rastu ako rastojanje između ostalih stanja i posmatranog stanja opada, tj. ako su ostale tačke 'gusto' raspoređene oko posmatrane tačke. Iz toga sledi da verovatnoće stanja π_i predstavljaju dobru procenu gustine.

U [11] proračuna verovatnoća stanja zasniva se na rešavanju problema:

$$\lim_{n \rightarrow \infty} \mathbf{\Pi}^n = \begin{bmatrix} \pi_1 & \dots & \pi_N \\ \vdots & \dots & \vdots \\ \pi_1 & \dots & \pi_N \end{bmatrix} \quad (3)$$

putem sopstvenih vrednosti i sopstvenih vektora tranzicione matrice. Najveća sopstvena vrednost stohastičke matrice je $\lambda=1$ [12], a njen sopstveni vektor $\boldsymbol{\pi}$, poznat i kao Peron-Frobenius sopstveni vektor [12], predstavlja stabilnu raspodelu verovatnoća stanja $\boldsymbol{\pi} = [\pi_1 \dots \pi_N]$. Izračunavanje sopstvenog vektora $\boldsymbol{\pi}$ za $\lambda=1$ svodi se na klasičan način rešavanja jednačina stanja [13]:

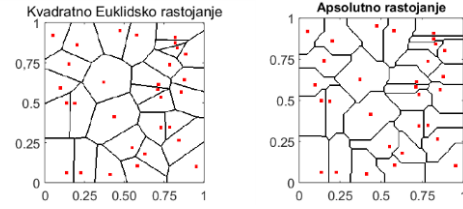
$$\boldsymbol{\pi} \cdot (\mathbf{\Pi} - \lambda \cdot \mathbf{I}) = \mathbf{0}_N, \text{ uz uslov } \boldsymbol{\pi} \cdot \boldsymbol{\xi}_N^T = 1, \lambda = 1 \quad (4)$$

gde \mathbf{I} označava jediničnu matricu, $\mathbf{0}_N$ i $\boldsymbol{\xi}_N$ označavaju vektore dimenzija $1 \times N$ ispunjene svim nulama i svim jedinicama, a T predstavlja transponovanje.

Koncept Voronoi regiona [14] prvi put je primenjen za procenu gustine verovatnoće u [3]. Definiše se na sledeći način: neka je A skup svih koordinata a u $[0, 1]^D$ prostoru gustine kopule. Neka je $W_k = [u_{1k}, \dots, u_{nk}]$, $k=1, \dots, N$, D -dimenzionalna tačka PIT signala. Tada Voronoi region R_k^D oko tačke W_k sadrži sve koordinate a koje su bliže posmatranoj tački W_k nego bilo kojoj drugoj tački W_j , $j=1, \dots, N$, $j \neq k$ u smislu izabrane definicije rastojanja $d(a, W_k)$. Formalnije,

$$R_k^D = \{a \in A \mid d(a, W_k) \leq d(a, W_j), \forall j \neq k\}. \quad (5)$$

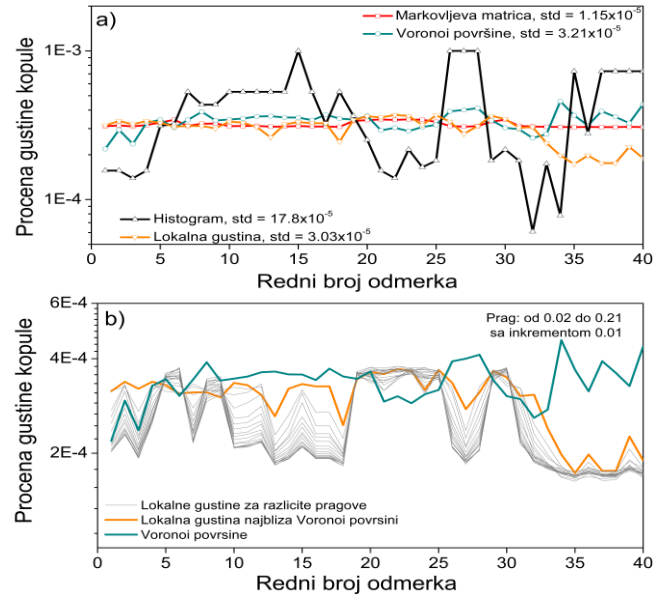
D -dimenzionalne 'površine' Voronoi regiona su inverzno proporcionalne gustini zavisnosti (kopula gustini). Povećanje gustine implicira smanjenje prostora između tačaka i, posledično, smanjenje Voronoi površina. Pored toga, region R_k^D razdvaja tačku W_k od ostalih tačaka po kriterijumu najvećeg rastojanja, tako da je površina oko tačke W_k jednoznačno određena, jedinstvena i bez preklapanja a sve površine se dopunjavaju do celokupne površine kopula domena. Međutim, iako nije potrebno definisati prag, postoji stepen slobode: moguće je birati način proračuna rastojanja između koordinate a i tačke W_k . Po pravilu se koristi kvadratno Euklidsko rastojanje ali mogu i druga, što dovodi do drugačijih Voronoi dekompozicija. Na Slici 3 prikazana je Voronoi dekompozicija istog segmenta signala za slučaj kvadratnog Euklidskog i apsolutnog rastojanja.



Slika 3. Dvodimenzionalni Voronoi regioni istog signala proračunati na osnovu različitih definicija rastojanja; levi panel: kvadratno Euklidsko rastojanje; desni panel: apsolutno rastojanje. Crvene tačke predstavljaju dvodimenzionalne PIT SBP i PIT PI signale.

IV. REZULTATI

Slika 4 prikazuje fluktuacije jačine zavisnosti SBP i PI signala kontrolnog pacova. Signali su dužine $N = 3000$ a prikazano je četrdeset odmeraka. Slika 4a prikazuje fluktuacije kada je procena rađena Markovljevom matricom (parametar γ ne utiče na procenu), Voronoi dekompozicijom sa kvadratom Euklidskog rastojanja, lokalnom dekompozicijom sa pragom 0.09 i histogramom sa 10×10 ćelija histograma. U sliku je upisana i standardna devijacija std kopula gustine za sva četiri postupka. Slika 4b prikazuje fluktuacije za različite pragove pri primeni metode lokalne gustine.

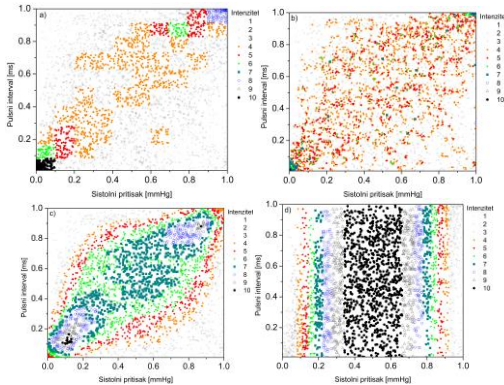


Slika 4. Procena jačine povezanosti SBP i PI signala (gustina kopule) za kontrolnog pacova. a) Četiri metode procene i standardna devijacija svakog od dobijenih signala; b) Prikaz lokalne gustine za različite pragove.

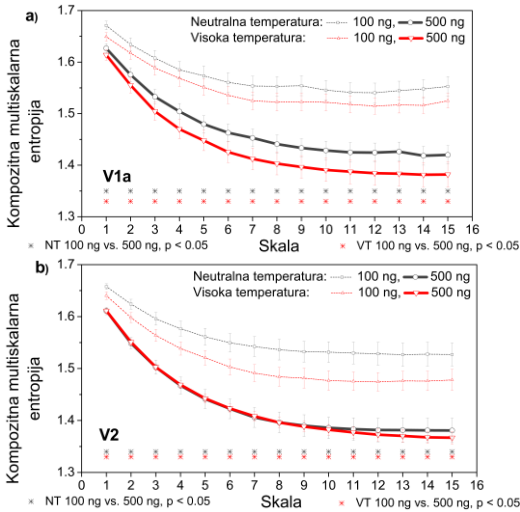
Slika 5 prikazuje kopula gustinu (strukturu zavisnosti) jednog od kontrolnih pacova procenjenju različitim metodama. I dok histogram, Voronoi dekompozicija i kvadratna procena lokalne gustine prikazuju očekivanu sliku, Markovljeva procena stabilnih verovatnoća stanja se ne uklapa: zasniva se na simetričnoj matrici rastojanja $d(i, j)$ sa stohastičkim vrstama. To znači da se svaka vrsta skalira sopstvenim faktorom, odnosno vrste se obrađuju pod međusobno različitim uslovima što dovodi do pojave linija. Sličan efekat se uočava i kada se

lokalna gustina procenjuje na osnovu pragova iz matrice rastojanja i cirkularne površine za lokalnu procenu.

Konačno, slika 6 prikazuje kompozitnu multiskalarnu entropiju procenjenju analizom SBP i PI signala transformisanih u niz fluktuacija jačine povezanosti ta dva signala. Uočava se da metoda diferencira efekat različitih doza primenjenig antagonista.



Slika 5. Procena gustine kopule; a) histogram; b) Voronoi površine; c) lokalna gustina (kvadratne granice); d) Markovljeve matrice.



Slika 6. Kompozitna multiskalarna entropija procenjena nad transformisanim nizovima jačine međusobne zavisnosti SBP i PI signala laboratorijskih životinja pod različitim dozama V1a i V2 antagonista

V. ZAKLJUČAK

Rad analizira metode procene gustine verovatnoće primenjene na proceduru združavanja više zavisnih signala u jedan vremenski signal koji prikazuje fluktuacije jačine njihove međusobne zavisnosti. Kao važan rezultat pokazalo se da metode procene koje se zasnivaju na matricama rastojanja između pojedinih tačaka nisu pogodne jer su matrice simetrična a metode funkcionišu liniju – po – liniju i unose tipične artefakte koji su posledica različitih skaliranja po liniji (u slučaju Markovljevih matrica) a, vrlo verovatno, i različitih koordinatih sistema, slično artefaktima koji se javljaju pri Radonovoj transformaciji. To je značajno napomenuti jer se Markovljeve matrice koriste za procenu gustine verovatnoće u novom konceptu Teorije

informacija, ‘uskom grlu’ gde je, za razliku od SBP-PI signala, mnogo teže proveriti tačne vrednopsti.

LITERATURA

- [1] A. Sklar, “Fonctions de répartition à n dimensions et leurs marges”. *Publications de l’Institut de Statistique de l’Université de Paris*, vol. 8, pp. 229–231, 1959.
- [2] H. Joe, **Multivariate Models and Dependence Concepts**, Chapman and Hall/CRC, London, 2017.
- [3] D. Bajić, T. Škorić, S. Milutinović-Smiljanić, and N. Japundžić-Žigon “Voronoi decomposition of cardiovascular dependency structures in different ambient conditions: an entropy study”, *Entropy*, vol. 21, no. 11, paper no. 1103-24 pages, 2019.
- [4] B. W. Silverman: **Density Estimation for Statistics and Data Analysis**, CRC Press, Boca Raton, FL, USA, 1986
- [5] N. Wessel, A. Voss, H. Malberg, Ch. Ziehmann, H..U. Voss, A. Schirdewan, U. Meyerfeldt, J. Kurths, “Nonlinear analysis of complex phenomena in cardiological data”, *Herzsch. Elektrophys*, vol. 11(3), pp. 159-173, 2000.
- [6] S. Milutinović, D. Murphy, N. Japundžić-Žigon: “The role of central vasopressin receptors in the modulation of autonomic cardiovascular controls: a spectral analysis study”, *American Journal Of Physiology-Regulatory Integrat*, vol. 291 no. 6, pp r1579-r1591, 2006.
- [7] M. Merkle, **Verovatnoća i statistika za inženjere i studente tehnike**, četvrto izmenjeno i dopunjeno izdanje, Akademska misao, Beograd, 2016.
- [8] S. Jovanović, T. Škorić, O. Šarenac, S. Milutinović-Smiljanić, N. Japundžić-Žigon, D. Bajić, “Copula as a dynamic measure of cardiovascular signal interactions”, *Biomedical Signal Processing and Control*, vol. 43, pp 250-264, 2018,
- [9] T. Tasić, S. Jovanović, Omer Mohamoud, T. Škorić, N. Japundžić-Žigon, D. Bajić, “Dependency structures in differentially coded cardiovascular time series”, *Computational and Mathematical Methods in Medicine*, Hindawi Publisher, vol. 17, no. 1, pp. 1 - 17, 2017,
- [10] M. C. Jeruchim, “Techniques for Estimating the Bit Error Rate in the Simulation of Digital Communication Systems,” *IEEE Journal on Selected Areas in Communications*, vol. 2, no. 1, pp. 153–170, 1984.
- [11] N. Tishby, F. C. Pereira, W. Bialek, “The Information Bottleneck Method”, *Proceedings of the 37th annual Allerton Conference on Communication, Control, and Computing*, pp. 368–377, 1999.
- [12] E. T. Beasley, Jr, P. M. Gibson, “A relationship between characteristic values and vectors”, *Proc. Aw. Math. Soc.* vol. 43 pp 71-78, 1974.
- [13] G. Lukatela, **Statistička teorija telekomunikacija i teorija informacija**, Građevinska knjiga, Beograd, 1981.
- [14] G. Voronoi, “Nouvelles applications des paramètres continous à la théorie des formes quadratiques”, *Journal für die Reine und Angewandte Mathematik*. vol. 133, pp. 97–178, 1908.

ABSTRACT

This paper analyses an original transform of dependency structures of multivariable signals from the probabilistic to the temporal domain, Dependency structures are multidimensional probability densities that show the coupling strength of the observed signals. These structures are described by the static model known as ‘copula density’. The dynamic model of coupling strength is a new and original time series. It enables the analysis of temporal fluctuations of coupling strength. Such an analysis is of importance for medical applications, so cardiovascular signals of laboratory animals were analyzed. A key transformation parameter is copula density, so several estimation methods are compared within this paper.

Temporal transformations of multivariable dependency structures

Dragana Bajić, Tamara Škorić, Marko Mozetić and Nina Japundžić-Žigo

