Prošireni sažetak / Extended abstract

Otkrivanje znanja i prediktivni modeli u kardiologiji

Knowledge discovery and predictive models in cardiology

Guy Carrault

Campus de Beaulieu, Université de Rennes I, Rennes, Francuska

Campus de Beaulieu, Université de Rennes I, Rennes, France

Ovim priopćenjem želimo ukazati kako razmatranje vremenske dinamike srčanih funkcija poboljšava interpretaciju kardiološke multivariatne vremenske serije. Opisat će se dva različita pristupa koja će se predstaviti na dva različita primjera.

Prva metodologija kombinira prostorno-vremensko neizrazito (fuzzy) kodiranje i multikorespodencijsku analiza (MCA). MCA se primjenjuje da bi se: 1) smanjio broj dimenzija podataka i osigurali novi sintetski indeksi temeljeni na "faktorskim osima" dobivenim od MCA; 2) tumačile dobivene faktorske osi u fiziološkim uvjetima i 3) analizirala promjena stanja bolesnika projekcijom prikupljenih podataka u ravninu nazvanu "faktorska ravnina" formiranu od prve dvije faktorske osi. Kao ilustracija, podaci dobiveni implatibilnim uređajima s dva senzora (senzor transtorakalnog otpora i akcelerometar) analizirani su da bi se otkrila njihova potencijalna primjena za praćenje pacijenata liječenih kardijalnom resinkronizacijskom terapijom (CRT). Da bi se klasificirali različiti uzorci nastanka, nakon smanjenja broja dimenzija podataka dobivenih od MCA, uvedena je mjera sličnosti za grupiranje promatranih podataka dobivenih od 41 pacijenta s CRT. Istraživanje cijele baze podataka omogućilo je otkriće da su dobiveni klasteri u usporedbi s napomenama iz zdravstvenog kartona pojedinog bolesnika te je zauzelo dva područja u faktorskoj ravnini, jedno vezano uz pogoršanje zdravlja pacijenta i drugo vezano uz stabilno kliničko stanje.

U drugoj metodologiji dinamiku pojedinih vremenskih serija odlikuje niz skrivenih polu-Markovljevih modela (HSMM). U tom slučaju metoda dijagnostike koristi ne samo trenutne vrijednosti nego je također predložena i uporaba unutarnje dinamike vremenske serije. Predložen je skriveni polu-Markovljev model da bi se otkrila vremenska evolucija promatranih serija te su ispitane različite predprocesne metode ovih serija. Temeljem usporedbi vjerojatnosti stvoren je detektor koji za opservirano promatranje generira referentni ili patološki HSMM. Za ilustraciju analizirane su unutarnje dinamike RR serije od 18 nedonoščadi iz neonatalnih jedinica intenzivnog liječenja (NICU) da bi se otkrili prediktivni uzorci bradikardije. Neobična i ponavljajuća bradikardija kod nedonoščadi često otkriva važne bolesti kao sto je to kasnonastala sepsa. Obzirom na visok morbiditet i mortalitet povezan s infekcijom, neophodno je brza detekcija. Pristup je bio kvantitativno ocijenjen. Usporedbom dva konvencionalna detektora korištena u NICU, ovaj pokazuje poboljšanje od oko 13% u osjetljivosti i 7% u specifičnosti te također reducira kašnjenje dijagnostike od oko tri sekunde u usporedbi s konvencionalnim detektorima.

Ključne riječi: prediktivni modeli, otkrivanje znanja, kardiologija.

In this communication, we would like to show how the consideration of the temporal dynamics of cardiac features improve interpretation of cardiac multivariate time series. Two different approaches will be described and demonstration will be illustrated over two examples.

The first methodology combines spatiotemporal fuzzy coding and multiple correspondence analysis (MCA). MCA is applied in order to: 1) reduce the dimensionality of the data and provide new synthetic indexes based on the "factorial axes" obtained from MCA; 2) interpret these factorial axes in physiological terms; and 3) analyze the evolution of the patient's status by projecting the acquired data into the plane formed by the first two factorial axes named "factorial plane." As an illustration, data obtained from an implantable device with two sensors (a transthoracic impedance sensor and an accelerometer) are analyzed in order to discover their potential application for the follow-up of patients treated with a cardiac resynchronization therapy (CRT). In order to classify the different evolution patterns, after reducing the dimensionality of the data by MCA, a similarity measure is introduced to cluster the observed data set from 41 CRT patients. Exploration of all the data base allows to discover that the obtained clusters, compared with the annotations on each patient's medical record, occupy two areas on the factorial plane, one being correlated with a health degradation of patients and the other with a stable clinical state.

In the second methodology, the dynamics of individual time series are characterized by a set of Hidden Semi-Markovian Models (HSMM). In this case, a detection method that exploits not only the instantaneous values, but also the intrinsic dynamics of the temporal series, is proposed. The hidden semi-Markov model is proposed to discover the temporal evolution of observed series and different pre-processing methods of these series are investigated. The detector is based on the comparison of the likelihood that a given observation being generated by a reference HSMM or a pathologic HSMM. As an illustration, the intrinsic dynamics of the RR series of 18 preterm new-borns acquired in neonatal intensive care units (NICU) are analysed in order to discover predictive patterns of the bradycardia. Unusual and recurrent bradycardias in preterm babies often revealed important disorders such as late-onset sepsis. In view of the high morbidity and mortality associated with infection, rapid detection is requested. The approach was quantitatively evaluated. Compared to two conventional detectors used in NICU, our detector shows an improvement of around 13% in sensitivity and 7% in specificity and also reduces the detection delay of approximately 3 seconds with respect to conventional detectors.

Keywords: predictive models, knowledge discovery, cardioogy.

*Corresponding author — E-mail: guy.carrault@univ-rennes1.fr

Literature

1. Guéguin M, Roux E, Hernández AI, Porée F, Mabo P, Graindorge L, Carrault G. Exploring time series retrieved from cardiac implantable devices for optimizing patient follow-up, IEEE Trans Biomed Eng. 2008;55(10):2343-52.

2. Altuve M, Carrault G, Beuchee A, Pladys P, Hernández AI. On-line apnea-bradycardia detection using hidden semi-Markov models. Conf Proc IEEE Eng Med Biol Soc. 2011;2011:4374-7.

3. Beuchée A, Carrault G, Bansard JY, Boutaric E, Bétrémieux P, Pladys P. Uncorrelated randomness of the heart rate is associated with sepsis in sick premature infants, Neonatology. 2009;96(2):109-14.