

Procesi, podaci, ljudi i stvari: četiri IoE osnove prikupljanja medicinskih podataka

Konstantinos Katzis
Department of Computer Science
and Engineering, European University Cyprus,
Nicosia, Cyprus
k.katzis@euc.ac.cy

Tamara Škorić
Fakultet tehničkih nauka
Univerzitet u Novom Sadu
Novi Sad, Srbija
ceranic@uns.ac.rs

Sladana Jovanović
Preduzeće za telekomunikacije
“Telekom Srbija”, A.D.
Beograd, Srbija
sladjanajo@telekom.rs

Dragana Bajić
Fakultet tehničkih nauka
Univerzitet u Novom Sadu
Novi Sad, Srbija
dragana.bajic@gmail.com

Milan Jovanović
“Endava”.
Beograd, Srbija
milan.jovanovic@endava.com

Branislav Milovanović
Medicinski fakultet
Univerzitet u Beogradu
Beograd, Srbija
branislav_milovanovic@vektor.net

Sažetak — U radu se analiziraju fundamentalne osnove Interneta svega (IoE – Internet of everything) i njihove realizacije u kontekstu medicinskih aplikacija. Poseban akcenat je stavljen na formiranje velikih baza medicinskih podataka sačinjenih od signala koje korisnici prikupljaju tokom svakodnevnih aktivnosti a zatim odbace da im ne opterećuju memorijske resurse. Na prikupljenim podacima prikazane su definicione postavke volonterskog prikupljanja podataka (crowd sensing) i to kvalitet informacija, motivacija volontera, prenos signala nasuprot prenosu parametara uz mogućnosti uštede. Ukazano je na značaj razvoja robusnih metoda za analizu volonterski prikupljenih podataka.

Ključne riječi – IoE u zdravstvu, prikupljanje podataka, crowd sensing.

I. UVOD

Rad je posvećen opisu medicinskih aplikacija u kontekstu Interneta svega, realizovanog putem “crowd sensinga”, to jest putem volonterskog prikupljanja podataka. Zasnovan je na činjenici da preko 75% svetske populacije poseduje takozvane “pametne” mobilne uređaje – telefone, satove, tablete, različite delove fitnes opreme. Svaki takav uređaj poseduje veliki broj senzora, procesor zavidnih mogućnosti i primopredajnik, stoga je u mogućnosti da snima, obrađuje i prenosi najrazličitije podatke. Na žalost, mnoštvo podataka se odmah nakon pregleda odbaci, a ponekad se čak ni ne pregleda. Čovečanstvo se time lišava velikog izvora medicinskih informacija čija bi analiza mogla da pomogne daljem razvoju dijagnostike i prognostike.

Jedna od platformi koja se bori protiv odbacivanja podataka je crowd sensing, koji omogućava da podaci nađu korisnu primenu za dobrobit i pojedinca i zajednice.

Crowd sensing je strogo definisan po pitanju korišćenja resursa. Njegova inicijalna tipologija se odnosila na zaštitu životne sredine, na društvo i na urbane probleme [1] gde je količina podataka za prenos relativno skromna. Medicinski signali su dodati kasnije, sa svim ograničenjima koje unose privatnost i poverljivost podataka [2-6]. Kao naknadno dodati, morali su da se uklape u već postojeća ograničenja. Jedan od najznačajnijih je da se u udaljeni centar za obradu ne prenose snimljeni signali, već samo parametri koji se iz signala izdvajaju na licu mesta [1], a sami signali se odbacuju. Time postavlja pitanje pouzdanosti: volonterski-amaterski snimljeni signali mogu da sadrže artefakte pa parametri izvučeni iz njih mogu da budu pogrešni. Pouzdanost je važna jer je jedna od primena crowd sensinga formiranje baza medicinskih podataka o volonterima koji svoje signale snimaju u različitim svakodnevnim situacijama [7]. Te baze su rudnik podataka za medicinska istraživanja pa je potrebno da zapisi budu tačni. Baza o hipertenzivnim pacijentima je već u procesu formiranja [7].

Ovaj rad povezuje problem volonterskog prikupljanja medicinskih podataka u svima dostupnu medicinsku bazu sa 4 konstitutivna elementa Interneta svega (IoE) [8]. Rad se suprotstavlja crowd sensing standardima [1] i pokazuje da se ušteda u prenosu može ostvariti selekcijom signala koji se prenose. Istraživanje je urađeno na 460 subjekata koji su

uključivali i pacijenate sa različitim kardiovaskularnim dijagnozama i zdrave volontere.

II. ČETIRI IOE OSNOVE PRIKUPLJANJA MEDICINSKIH PODATAKA

Internet svega počiva na četiri stožera koje čine međusobno povezani uređaji i objekti (stvari), povezani ljudi, podaci koji se koriste za inteligentne odluke i procesi koji omogućavaju da se prave informacije isporuče na pravo mesto u pravo vreme.

A. Ljudi

Osnova IoE sa primenom u zdravstvu su ljudi. Prikupljanje podataka koje pojedinci snimaju u slobodno vreme iz sopstvene radoznalosti i odbacuju nakon pregleda predstavlja volontersku aktivnost. Pojedinci nemaju ličnu korist od prikupljanja podataka pa se ne može očekivati da troše vreme na prenos zapisanog signala na uređaj sa fiksnim napajanjem i eventualnu dodatnu obradu. Stoga, umesto zapisa, ove aplikacije na zahtev korisnika automatski, u realnom vremenu, prenose snimane podatke u udaljeni centar. Čak i tada, kako su eksperimenti pokazali, volonteri su zainteresovani samo na početku. Tabela 1 prikazuje rečit gubitak interesovanja mladih zdravih dobrovoljaca za eksperiment: tokom desetodnevnog ispitivanja njihov broj se smanjio sa 10 na 3, a trajanje (kratkog) eksperimenta se prepolovilo [7].

Tabela 1: Gubitak interesovanja volontera

	Broj aktivnih volonera	Prosečno trajanje merenja [s]
Prvi dan:	10	260 ± 52
Deseti dan:	3	120 ± 14

U radovima [9-12] opisano je nekoliko izrazito teorijskih algoritama za povećanje motivacije volontera za prikupljanje podataka. Radovi su multivarijabilni i u opštem slučaju pretpostavljaju više istovremenih volonterskih aktivnosti i nekoliko načina nagrađivanja (uključujući i ekonomski momenat). Neprimenljivi su za slučaj samo jedne aktivnosti i nepostojanje novčanih nagradnih resursa.

U radu [7] prikazan je jedan način motivacije: hipertenzivni pacijenti su, tokom korišćenja aplikacije, bivali upozoreni da smanje trenutne aktivnosti kada njihovi parametri izađu iz osega koji se toleriše. Pokazalo se da su dve komponente – svest o bolesti i lični benefit – dovoljne da motivišu pacijente da aktivno doprinesu formiranju baze.

B. Podaci

Određnica za podatke u kontekstu IoE u zdravstvu je kvalitet informacija (QoI) [13]. Kvalitet informacija se, pre svega, odnosi na odsustvo artefakata (združeno ime za sve smetnje, izobličenja i patologije koje pogađaju medicinske podatke).

Medicinska istraživanja zahtevaju vizuelnu korekciju artefakata. Najtvrdi stanovišta čak ne dozvoljavaju ni primenu filtera razvijenih isključivo za medicinske signale određene vrste i naknadnu vizuelnu korekciju zaostalih grešaka. Vizuelna korekcija artefakata je, međutim, kontradiktorna principu minimalnog napora koji predstavlja uslov za rad volontera.

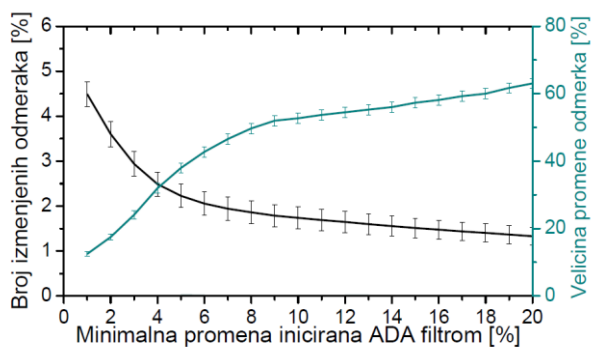
Budući da su artefakti prisutni i kada se signali snimaju u kontrolisanim profesionalnim okruženjima kao što su medicinske ustanove, problem artefakata u slobodnom prikupljanju podataka mobilnim uređajima je još izraženiji. Senzore postavljaju sami volonteri koji za to nisu obučeni. Usled labavo nameštenih senzora dolazi do njihovog pomeranja koje izaziva artefakte. Nadalje, volonteri se kreću dok rade merenja što još više pospešuje trenje senzora (i čini signal nestacionarnim, što je dodatni problem). Konačno, pod artefaktom se podrazumeva i signal sa patologijom.

Iz signala sa artefaktima ne mogu da se pouzdano procene parametri i bilo bi dobro da se prenese celokupni zapisan signal kako bi se artefakti kvalitetno otklonili u udaljenom centru. Međutim, velika količina podataka će opterećivati prenosni medijum i povećati potrošnju baterije - radi se o mobilnim uređajima a predajnik predstavlja najveći potrošač.

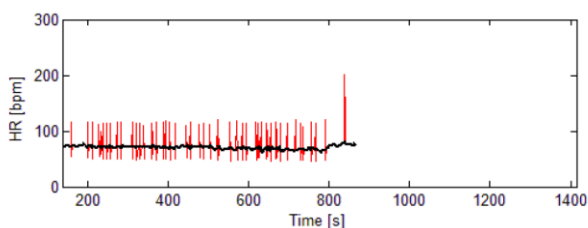
Mogući način uštede je procena da li je informacija dovoljno kvalitetna da bi zaslužila da se prenese. U tom smislu je analizirano 460 nizova sa R-R intervalima (R-R interval je rastojanje dva uzastopna R vrha na elektrokardiografskom – EKG - zapisu). Najčešći artefakti u ovom signalu su posledica lošeg EKG signala, grešaka u izdvajanja R-R intervala iz dobrog EKG signala i srčane patologije. Kriterijumi za odbacivanje su dužina signala i broj korigovanih odmeraka.

Za dužinu signala postoje standardi: mora da traje bar 5 minuta [14]). Za broj korigovanih odmeraka ne postoji propis i, sem preliminarnih istraživanja u [15], ni odgovarajuće preporuke. Ako je broj korigovanih odmeraka suviše velik, analiza takvog signala neće imati smisla.

Po kriterijumu dužine odbačeno je 60 signala, a zatim su signali filtrirani putem modifikovanog ADA filtera, dizajniranog upravo za R-R signal i njegov inverzni parnjak, signal srčanog ritma (heart rate, HR) [16]. Modifikacija originalnog ADA filtera je jednostavna: ako se filtrirani odmerak razlikuje od polaznog odmerka za manje od P%, vraća mu se polazna vrednost.



Slika 1. Broj promenjenih odmeraka signala [%] i veličina odgovarajuće promene, u zavisnosti od tolerisanog minimalnog praga ADA filtra. Grafici prikazuju vrednost usrednjenu nad 400 pacijenata \pm s.e.m.



Slika 2. Signal srčanog ritma (HR) izražen u otkucajima u minuti (beats per minute, bpm). Originalni signal sa tipičnim artefaktima prikazan je crveno, signal korigovan ADA filtrom prikazan je crno.

Slika 1 prikazuje procenat izmenjenih odmeraka signala $100 \cdot N_{FILT} / N$, ako je minimalna promena odmerka ravna $P \in \{1, \dots, 20\}$ [%] (N je broj odmeraka u signal). Ista slika prikazuje i usrednjenu relativnu apsolutnu razliku Δ_R izmenjenih odmeraka u odnosu na polazne odmerke.

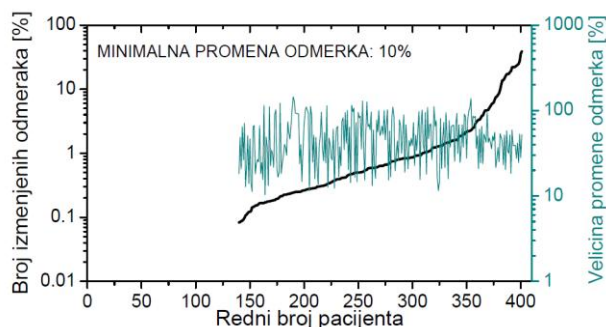
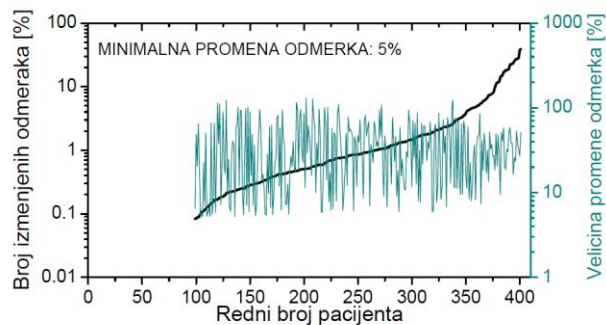
$$\Delta_R = \frac{100}{N_{FILT}} \cdot \sum \frac{abs(x_{FILT} - x_{ORIG})}{x_{ORIG}} [\%].$$

Na slici 2 prikazan je tipičan primer signala srčanog ritma sa artefaktima. Signal je filtriran ADA filtrom ali broj artefakata je suviše veliki i signal ne može da se koristi za dalju analizu, to jest ne mora da se prenosi na udaljeni kraj.

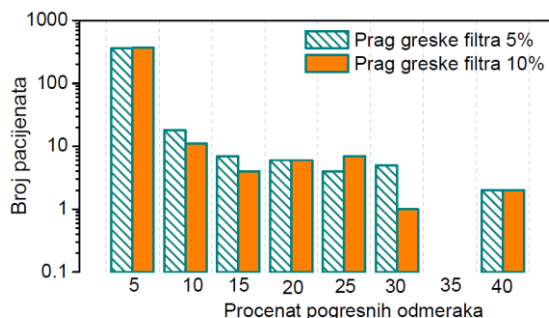
Slike 3 predstavljaju, za 400 pacijenata, procenat promenjenih odmeraka sortiran rastući i odgovarajuću veličinu odstupanja i to kada je prag ADA filtra postavljen na 5% i na 10%.

Slika 4 prikazuje broj signala volontera koji imaju određeni procenat grešaka. Ako se broj pacijenata sa suviše velikim brojem korigovanih odmeraka sabere sa pacijentima sa prekratkim signalima, u proseku treba da se odbaci oko 20% signala. To predstavlja značajnu uštedu jer, kao što je već rečeno, prenos je dominantni potrošač po dva osnova: zauzima propusni opseg i troši bateriju više od bilo čega drugoga. Stoga, testiranjem broja artefakata pomoću ADA filtra ne samo da se rešava problem grešaka, već i ušteda propusnog opsega i

baterije. Ovaj problem je u [7] skraćeno nazvan BBA (battery, bandwidth, artifacts=).



Slika 3. Broj izmenjenih odmeraka za svakog pacijenta u rastućem nizu i odgovarajuća usrednjena relativna vrednost izmene odmeraka za dati signal, kada je minimalna registrovana izmena odmerka filtriranog signala 5% (gornji panel) i 10% (donji panel)



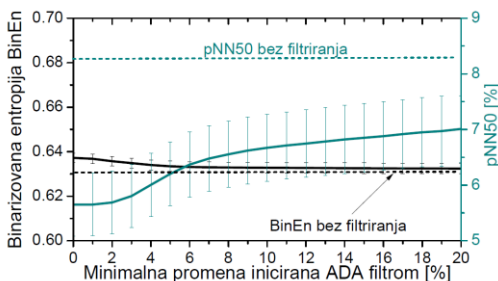
Slika 4. Broj pacijenata sa određenim procentom pogrešnih odmeraka

C. Procesi

Eliminacija artefakata još uvek ne znači da će izdvojena informacija biti pouzdana, tim pre što se standardna metoda – vizuelna korekcija podataka – koristi u laboratorijskim istraživanjima a mnogo manje u kliničkim primenama gde za nju, po sopstvenom priznanju, lekari nemaju vremena. Ovaj aspekt predstavlja temu i za istraživanja u kontekstu ‘ljudi’.

Neophodno je stoga proceniti da li su parametri procenjени iz nepregledanih a automatski obrađenih signala dovoljno pouzdani. Kao ilustrativni primer izabrani su entropija kao parametar nezaobilazan u kliničkim studijama i pNN50 koji svojim imenom kaže da mora da se procenjuje iz NN signala. Entropija je procenjena putem robusne metode dizajnirane za signale sa artefaktima [17]. Signal NN se dobija iz signala R-R tako što se odstrane svi artefakti. Sam parametar je diferencijalan i predstavlja procenat uzastopnih NN intervala koji se međusobno razlikuju za više od 50ms.

Slika 5 prikazuje promenu parametara na svim pacijentima čiji su signali dovoljno dugački za analizu. Binarizovana entropija, razvijena isključivo za primenu u crowd sensing sistemima, radi nezavisno od filtra jer analizira binarno diferencijalno kodovane signale. Diferencijalno kodovanje posmatra razliku uzastopnih odmeraka a binarno kvantovanje tu razliku svodi samo na pozitivnu ili negativnu, tako da artefakti ne mogu mnogo da utiču. Naravno, osetljivost metode je smanjena, ali zadovoljavajuća jer je entropija relativna mera, kao što je u [17] i pokazano. Sa druge strane pNN50 pokazuje značajna odstupanja, naročito kada se uporedi sa odsustvom filtra. Treba voditi računa da su rezultati prikazani kao srednja vrednost \pm s.e.m, što znači da je standardna devijacija na 400 pacijenata dvadeset puta.



Slika 5. Promena parametara BinEn i pNN50 u zavisnosti od parametara filtra; isprekidane linije – parametar procenjen bez filtriranja

D. Objekti i uređaju

Internet ‘stvari’ - objekata i uređaja - (IoT) je najpoznatiji segment Interneta svega. IoT (IoE) u ovom kontekstu omogućava medicinskim uređajima umrežavanje u pametnom okruženju karakterisanom polimornim zahtevima u smislu kašnjenja, truća pouzdanosti, brzine, potrošnje, bezbednosti i slično, istovremeno generišući ogromne količine nestrukturiranih podataka u realnom vremenu. Podaci će izazvati značajni priliv saobraćaja što će povećati i kompleksnost i cenu postojećih mreža. Komprimovanje podataka je jedno, ali ne i jedino rešenje. Bolji pristup bi bio uvođenje jeftinijih mrežnih pristupa, što može da se omoguće primenom nezauzetog dela TV spektra o čemu je detaljno pisano u [7].

Prikupljanje medicinskih podataka za formiranje baze koja se sastoji od ‘amaterskih’ podataka dobijenih od volontera putem njihovih privatnih ‘pametnih’ mobilnih primopredajnih uređaja a pri različitim (ali zabeleženim) aktivnostima generiše vredan izvor informacija za dalje dijagnostičke i prognostičke analize. Već osmišljeni crowd sensing system za hipertenzivne pacijente [7] nije idealan, jer iz signala izvlači samo dvadesetak od stotinu mogućih parametara. Stoga za ozbiljno formiranje baze treba iskoristiti crowd sensing preimućstva, a zaobići ograničenja. Predložena analiza podataka ‘in situ’, umesto izvlačenja parametara, postiže uštede eliminacijom onih signala koji se okarakterišu kao nepogodni za dalju analizu. ADA filter se pokazao kao pogodan za diskriminaciju signala, sa dodatkom praga koji za artefakte proglašava odmerke koji se razlikuju za više od P% od polaznih, a eliminaciju signala radi ako broj pogrešnih odmeraka prelazi zadanu vrednost.

Pokazalo se, takođe, da treba dalje raditi na razvoju robusnih parametara neosetljivih na artefakte. Konačno, za optimizaciju je potrebno voditi računa o svim fundamentalnim stozerima – ljudima, procesima, podacima i uređajima (‘stvarima’) što ovakav koncept volontarističkog kreiranja baze podataka uvodi u vode ‘interneta svega’.

Naredna istraživanja ići će u pravcu daljeg poboljšanja, pre svega hardvera (‘stvari’) iznalaženjem novih resursa za efikasan a po ceni nezahtevan modalitet prenosa podataka što je započeto korišćenjem belog TV opsega [7], analitičkih alata otpornih na karakteristične artefalte (‘procesi’), diskriminaciju podataka koji su neupotrebljivi (‘podaci’) i motivacije pacijenata, povećanjem svesti da ne samo da su analize njihovih podataka bitne za njih same, već da su od velikog značaja za celu zajednicu.

ZAHVALNICA

Ovaj rad je podržan grantom TR32040 Ministarstva prosvete, nauke i tehnološkog razvoja Republike Srbije. Rad je pod pokroviteljstvom EU COST Akcije CA15104 ‘Inclusive Radio Communication Networks for 5G and beyond’ - SEWG-IoT: Internet-of-Things for Health.

LITERATURA

1. Crowdsensing: State of the Art and Privacy Aspects, Available online: <https://resources.infosecinstitute.com/crowdsensing-state-art-privacy-aspects/> (28 November 2018)
2. Chessa, S; Corradi, A.; Foschini L.; Girolami, M. E. Mobile Crowdsensing through Social and Ad Hoc Networking. *IEEE Commun. Mag.* **2016**, *54*, 108-114. Available online: <https://ieeexplore.ieee.org/abstract/document/7509387/authors#authors>
3. Noureen, J.; Asif, M., Crowdsensing: Socio-Technical Challenges and Opportunities. *IJACSA* **2017**, *8*, 363-369. Available online:

- <https://pdfs.semanticscholar.org/716f/f776aa5e0ac25e78fd9002f42f0ce039f8cf.pdf> (28 November 2018)
4. Ganti, R.K.; Ye, F.; Lei, H. Mobile crowdsensing: current state and future challenges, *IEEE Commun. Mag.* **2011**, 32 – 39, Available online: <https://ieeexplore.ieee.org/document/6069707>
 5. Talasila M.; Curtmola R.; Borcea C. Ilr: Improving location reliability in mobile crowd sensing, *IJBDCN* **2013**, 9, 65–85. Available online: <https://web.njit.edu/~mt57/publications/IJBDCN13.pdf>
 6. Mehdi M.; MÄhlmeier G.; Agrawal K.; Pryss R.; Reichert M.; Hauck F. J. Referenceable mobile crowdsensing architecture: A healthcare use case, *Procedia Comput Sci*, **2018**, 134, 445-451. Available online: <https://esit.tinnitusresearch.net/images/Muntazir-Mehdi.pdf>
 7. Jovanović Slađana; Jovanović Milan; Škorić (Ćeranić) Tamara; Jokić Stevan; Milovanović Branislav; Katzis Konstantinos; Bajić Dragana A Mobile Crowd Sensing Application for Hypertensive Patients *Sensors* Vol. 19, No. 2, Str. 400-416, ISBN 1424-8220
 8. The Internet of Everything: How More Relevant and Valuable Connections Will Change the World. Available online: https://www.cisco.com/c/dam/global/en_my/assets/ciscoinnovate/pdfs/loE.pdf (28 November 2018)
 9. Wu, F.-J.; Luo, T. WiFiScout: A Crowdsensing WiFi Advisory System with Gamification-Based Incentive. In Proceedings of the IEEE 11th International Conference on Mobile Ad Hoc and Sensor Systems, Philadelphia, PA, USA, 28–30 October 2014.
 10. Wang, X.; Liu, Z.; Tian, X.; Gan, X.; Guan, Y.; Wang, X. Incentivizing Crowdsensing With Location-Privacy Preserving. *IEEE Trans. Wirel. Commun.* **2017**, 16, 6940–6952.
 11. Jin, H.; Su, L.; Chen, D.; Nahrstedt, K.; Xu, J. Quality of Information Aware Incentive Mechanisms for Mobile Crowd Sensing Systems. In Proceedings of the 16th ACM International Symposium on Mobile Ad Hoc Networking and Computing, Hangzhou, China, 22–25 June 2015.
 12. Jin, H.; Su, L.; Xiao, H.; Nahrstedt, K. Incentive Mechanism for Privacy-Aware Data Aggregation in Mobile Crowd Sensing Systems. *IEEE/ACM Trans. Netw.* **2018**, 26, 2019–2032.
 13. Jaimes L. G; Steele R., Incentivization for Health Crowdsensing, In *IEEE 15th Intl Conf on Dependable, Autonomic and Secure Computing, 15th Intl Conf on Pervasive Intelligence and Computing, 3rd Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress (DASC/PiCom/DataCom/CyberSciTech)*, USA, Nov. 2017;
 14. Task Force of The European Society of Cardiology and The North American Society of Pacing and Electrophysiology "Heart rate variability: Standards of measurement, physiological interpretation, and clinical use", *European Heart Journal* 1996; 17:354–381.
 15. Bajic D, Loncar-Turukalo T, Milovanovic B: Biomedical signals in BANs: pre-channel issues, invited paper, 17th European Wireless, Wien, Austria, 2011, pp 89-92, ISBN 978-3-8007-3343
 16. Wessel N.; Malberg H.; Bauernshmitt R.; Kuths J. Nonlinear methods of cardiovascular physics and their clinical applicability. *IJBC*, 2007, 17, 3325–3371. <https://pdfs.semanticscholar.org/3f43/dc38724e28e2d69b70423c2db4f72474776d.pdf>
 17. Skoric T.; Mohamoud O.; Milovanovic B.; Japundzic-Zigon N.; Bajic D. Binarized cross-approximate entropy in crowd sensing environment. *Comput. Biol. Med.*, 2017, 80, 137-147. Available online: <https://www.sciencedirect.com/science/article/pii/S0010482516303146/>

ABSTRACT

In this paper we analyze fundamental foundations of Internet of Everything (IoE) and their realizations considering the medical applications. A special attention is devoted to the database creation, where the inputs are signals that volunteers collect during their everyday activities, and then discard due to the memory limitations. The collected data served to discuss the basic requirements of the crowd sensing: parameter transmission vs. complete signal transmission. A comparison between robust and classical parameters is performed.

THINGS, PEOPLE, DATA AND PROCESSES: FOUR IoE ESSENTIALS OF THE HEALTH SIGNALS HARVESTING

Konstantinos Katzis, Slađana Jovanović, Milan Jovanović,
Tamara Škorić, Dragana Bajić, Branislav Milovanović