

Određivanje pozicije mobilnih GSM korisnika korišćenjem *Support Vector Regression* metode

Majda Petrić, Nataša Nešković, Aleksandar Nešković
Katedra za telekomunikacije, Laboratorija za radio
komunikacije
Elektrotehnički fakultet Univezriteta u Beogradu
Beograd, Republika Srbija
majdapetric@hotmail.com, natasha@etf.rs, neshko@etf.rs

Miloš Borenović
Vlatacom
Beograd, Republika Srbija
milos.borenovic@gmail.com

Sadržaj—Potreba mobilnih operatora da omoguće servis hitnog poziva, kao i veliko interesovanje za razvojem komercijalnih LBS (*Location Based Service*) servisa, doveli su u proteklih nekoliko godina do intenzivnog razvoja tehnika za određivanje pozicije mobilnih korisnika. Osnovni cilj datih istraživanja jeste postizanje što veće preciznosti u pozicioniranju, ali uz što je moguće manje izmena na strani mreže i mobilnog terminala. U okviru ovog rada predstavljen je model za pozicioniranje mobilnih GSM (*Global System for Mobile Communications*) korisnika korišćenjem SVR (*Support Vector Regression*) metode. Realizovana su četiri tipa predloženog modela, korišćenjem različitih *kernel* funkcija. Ispitivanje preciznosti predloženih modela vršeno je kroz testiranje na realnim podacima prikupljenim u urbanom okruženju.

Ključne riječi—pozicioniranje; GSM; SVR.

I. UVOD

Ideja za razvojem metoda za pozicioniranje korisnika u javnim mobilnim sistemima prvenstveno je nastala iz potrebe za određivanjem lokacije mobilnog korisnika u slučaju iniciranja E112 (Evropa), odnosno E911 (USA), servisa za hitne pozive. Prema odluci FCC (*Federal Communication Commission*), do 11. septembra 2012. godine, mobilni operatori su u obavezi da omoguće lociranje korisnika, odnosno obezbede određivanje geografske dužine i širine njegove trenutne pozicije, sa preciznošću od minimalno 300m, u roku od 6 minuta. S druge strane, razvoj tehnika za određivanje pozicije mobilnih korisnika u ćelijskim sistemima doveo je i do mogućnosti razvoja niza komercijalnih LBS (*Location Based Service*) servisa poput: servisa lokalne i regionalne pretplate, tarifiranja zavisno od lokacije korisnika, pružanja informacija o najbližim objektima od interesa (npr. bankomatima, pumpama, bolnicama, restoranima, itd.), pružanja servisnih informacija (npr. stanje u saobraćaju, vremenska prognoza, itd.), praćenja vozila, praćenja pošiljki, navigacije do zadate adrese, oglašavanja, itd.

Minimalna zahtevana preciznost tehnika za pozicioniranje mobilnih korisnika u javnim mobilnim sistemima definisana je od strane FCC, a zavisi od toga da li se proračun prostornih koordinata lokacije vrši na strani mreže (*network-based* tehnike) ili na strani mobilnog terminala (*mobile-based* tehnike) [1].

Postoji već nekoliko rešenja za određivanje pozicije mobilnih korisnika u GSM (*Global System for Mobile Communications*), DCS (*Digital Communication System*) i UMTS (*Universal Mobile Telecommunication System*) mrežama. Predložene tehnike međusobno se razlikuju po parametrima GSM/UMTS signala na osnovu čijih vrednosti se vrši procena pozicije mobilnog korisnika, kao i na osnovu metode koja se koristi za određivanje nelinearne relacije koja postoji između vrednosti datih parametara i prostornih koordinata trenutne lokacije mobilnog korisnika. Postojeće tehnike pozicioniranja zasnivaju se na: korišćenju informacija o identitetu servisne bazne stanice (*Cell-ID*), merenju snage signala baznih stanica na mestu prijema (*Received Signal Strength*, RSS), korišćenju vrednosti TA (*Timing Advance*) i RTT (*Round Trip Time*) parametara, merenju vremena propagacije GSM/UMTS signala (*Time Of Arrival*, TOA) ili razlike u vremenu propagacije GSM/UMTS signala više različitih baznih stanica (*Time Difference Of Arrival*, TDOA), određivanju pravca nailaska signala (*Angle Of Arrival*, AOA), merenju amplituda i faza direktne i *multipath* komponenta UMTS signala (*Power Delay Profile*, PDP) ili kombinovanju nekih od prethodno navedenih pristupa. U zavisnosti od korišćenog pristupa, postojeće tehnike za pozicioniranje mobilnih korisnika zahtevaju određene izmene na strani mreže, mobilnog terminala ili oba.

Dalji razvoj tehnika za određivanje pozicije mobilnih korisnika u ćelijskim sistemima usmeren je ka postizanju što veće preciznosti uz što manje izmene na strani mreže ili mobilnog terminala. Iako tehnike za pozicioniranje mobilnih korisnika koje se zasnivaju na korišćenju infrastrukture javnih mobilnih sistema imaju manju preciznost od postojećih tehnika baziranih na upotrebi satelitskih sistema (GPS, *Global Positioning System*), njihova osnovna prednost jeste u mogućnosti njihove primene, osim u *outdoor*, i u *indoor* okruženjima.

U okviru ovog rada biće predstavljen *network-based* model za određivanje pozicije mobilnih GSM korisnika korišćenjem SVR (*Support Vector Regression*) algoritma za estimaciju prostornih koordinata nepoznate lokacije na osnovu izmerenih vrednosti snaga signala baznih stanica na datoj lokaciji. Predložen model za pozicioniranje opisan je u okviru drugog poglavlja. Način prikupljanja podataka za treniranje i testiranje

modela, kao i njegovo formiranje, predstavljeni su u okviru trećeg poglavlja. Rezultati testiranja i verifikacije predloženog modela prikazani su u okviru četvrtog poglavlja, dok su zaključak i smernice za dalje istraživanje dati u petom poglavlju.

II. MODEL ZA POZICIONIRANJE

A. RSS (Received Signal Strength) pristup

Predložena tehnika za određivanje pozicije mobilnog korisnika korišćenjem infrastrukture GSM sistema zasnovana je na RSS pristupu. Odnosno, estimacija nepoznatih prostornih koordinata mobilnog korisnika vrši se na osnovu vrednosti snage signala servisne i šest susednih GSM baznih stanica (BS) koje mobilni terminal (MT) izmeri na datoj lokaciji. Prednost korišćenja ovog pristupa je prvenstveno u već standardizovanom postupku merenja nivoa signala servisne i susednih BS od strane mobilnog terminala, koji se vrši periodično radi omogućavanja *handover-a*. Takođe, slanje izveštaja servisnoj BS o izmerenim RSS vrednostima (*RxLev* parametrima) podržano je od strane postojeće signalizacije kroz slanje standardizovanog *Measurement Report-a*. Time su korišćenjem ovakvog pristupa izbegnute bilo kakve hardverske i softverske izmene na strani mobilnog terminala. S druge strane, korišćenje TOA i AOA pristupa zahteva odgovarajuće izmene na strani mobilnog terminala [2]. Druga prednost korišćenja RSS pristupa jeste u dostupnosti informacije o RSS vrednostima za veći broj baznih stanica. Naime, prethodno razvijene metode za rešavanje problema lokalizacije mobilnog korisnika u ćelijskim sistemima pokazale su da je tačnost pozicioniranja utoliko veća, ukoliko su odgovarajući parametri GSM/UMTS signala, na osnovu kojih se vrši estimacija pozicije, dobijeni od većeg broja baznih stanica [2]-[4]. Sa druge strane, vrednosti parametara poput TA, RTT i PDP su raspoložive samo za servisnu baznu stanicu (RTT parametar može se dobiti i za BS koje čine aktivni set, ali sam parametar nije podržan u svim UMTS mrežama, dok merenje PDP nije standardizovano, kao i slanje informacije o njemu servisnoj BS). Konačno, prednost korišćenja RSS pristupa jeste i u činjenici da je manje osetljiv na NLOS (*Non Line of Sight*) uslove propagacije u urbanim i suburbanim zonama, koji metode koje se zasnivaju na merenju vremena propagacije (TA, RTT, TDOA) ili pravca nailaska signala (AOA) u datim sredinama čini značajno manje preciznim [2].

U slučaju predložene tehnike, za svaku ćeliju GSM sistema definiše se poseban SVR model za pozicioniranje, na osnovu podataka prikupljenih merenjem u njenoj zoni pokrivanja.

Pojedine faze u postupku određivanja lokacije mobilnog korisnika prikazane su na Sl. 1. Nepoznata pozicija mobilnog korisnika određuje se tako što MT prosleđuje servisnoj BS podatke o izmerenim RSS vrednostima na datoj lokaciji kroz slanje *Measurement Report-a*. Servisna BS prosleđuje date podatke dalje ka određenom SMLC (*Serving Mobile Location Center*) centru, gde se odabir konkretnog SVR modela, koji će se primeniti za estimaciju pozicije datog korisnika, vrši na osnovu *Cell-ID* bazne stanice sa najvišim nivoom signala (tj. najvećom vrednošću RSS parametra na posmatranoj lokaciji). Razlog zbog koga se ne koristi SVR model koji odgovara ćeliji servisne bazne stanice, jeste u činjenici da usled određenih

ograničenja sistema (izbegavanja *ping-pong* efekta kod *handover-a*, korišćenja *load balancing-a*, itd.) servisna bazna stanica ne mora uvek biti i ona koja je najbliža mobilnom korisniku.

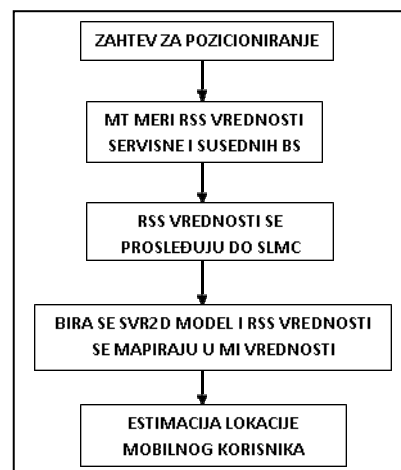
Nakon selekcije odgovarajućeg SVR modela za pozicioniranje, vrši se mapiranje RSS vrednosti koje je izmerio mobilni terminal u odgovarajuće ulaze modela (MI, *Model Inputs*). MI vrednosti predstavljaju RSS vrednosti 7 referentnih GSM baznih stanica koje imaju najveću „vidljivost“ u oblasti koju pokriva dati SVR model za pozicioniranje. Razlog za uvođenje datog mapiranja jeste u činjenici da na različitim lokacijama unutar oblasti koju pokriva posmatrani model za pozicioniranje postoji i različita „vidljivost“ baznih stanica. Takođe, na pojedinim lokacijama nisu „vidljive“ svih šest susednih BS. Dat problem je poznat i kao „*missing value*“ problem, odnosno problem nepostojeće vrednosti [5]. Prilikom mapiranja, izmerene RSS vrednosti, koje pripadaju baznim stanicama koje nisu referentne za odabrani model, odbacuju se. S druge strane, za MI vrednosti onih referentnih BS, koje nisu vidljive sa trenutne lokacije mobilnog korisnika, uzima se vrednost praga prijema, odnosno -110dBm.

B. SVR_{2D} model za pozicioniranje

Pronalaženje nelinearne relacije koja postoji između RSS vrednosti izmerenih na nekoj lokaciji (ulaznih podataka) i prostornih koordinata date lokacije (izlaznih podataka) vrši se korišćenjem SVR metode. SVR predstavlja klasu SVM (*Support Vector Machine*) algoritama za analizu modela (*pattern analysis* ili *pattern recognition*), odnosno algoritama za pronalaženje opštih relacija u skupu podataka [6], a namenjen je rešavanju problema regresije.

SVR algoritam sastoji se iz dve faze:

- trening faze (*off-line*) – u okviru koje se trenira model za već poznate parove ulazno-izlaznih podataka snimljenih na određenom broju lokacija i otkriva nelinearna relacija koja postoji između njih,
- verifikaciona faza (*online*) – u okviru koje se koristi prethodno naučena relacija za estimaciju prostornih koordinata nepoznate lokacije na osnovu poznatih RSS vrednosti.



Slika 1. Postupak određivanja lokacije mobilnog GSM korisnika.

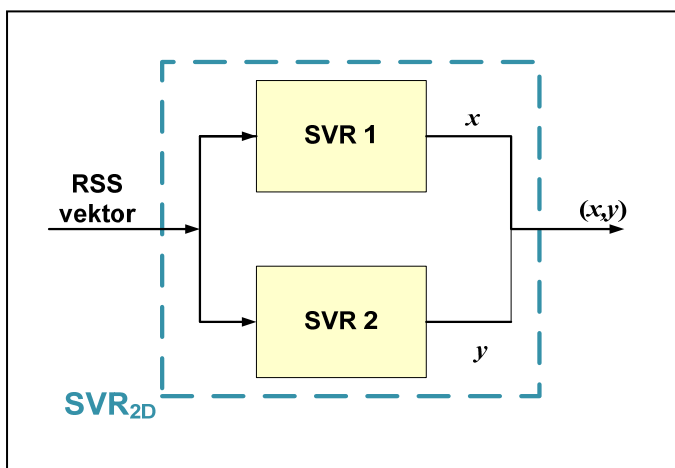
Osnovni koncept na kome se zasnivaju svi SVM algoritmi jeste preslikavanje podataka iz originalnog u karakteristični prostor veće dimenzionalnosti (*feature domen*), gde se sada opšte relacije koje postoje između seta ulaznih i izlaznih podataka mogu predstaviti linearnim funkcijama [6]. Problem povećanja kompleksnosti algoritma, usled povećanja dimenzionalnosti, rešen je korišćenjem specijalnih *kernel* funkcija. *Kernel* funkcije omogućavaju implicitno uvođenje karakterističnog prostora korišćenjem tkz. „*kernel trika*“ [7], koji eliminiše potrebu za poznavanjem same funkcije preslikavanja podataka iz originalnog u karakterističan prostor. Na taj način, SVM metode omogućavaju korišćenje karakterističnog prostora, čija dimenzionalnost sada može biti i beskonačna, a da kompleksnost proračuna ostane ista kao u originalnom prostoru.

Kako određivanje lokacije mobilnog korisnika predstavlja dvodimenzionalni problem, trenirana su dva SVR modela. SVR1 model treniran je sa ciljem pronalaženja veze između RSS vrednosti i prostorne koordinate x , dok je SVR2 model treniran u cilju pronalaženja relacije između RSS vrednosti i prostorne koordinate y . Zajedno, ta dva modela čine SVR_{2D} model za određivanje pozicije mobilnog korisnika u dvodimenzionalnom prostoru, koji je prikazan na Sl. 2.

III. GENERISANJE MODELA ZA POZICIONIRANJE

U okviru ovog rada realizovan je SVR_{2D} model koji se odnosi na makro ćeliju u urbanoj zoni grada Beograda. Posmatrani SVR_{2D} model za pozicioniranje realizovan je korišćenjem MATLAB programa. Za datu makro ćeliju simulirana su četiri različita tipa SVR_{2D} modela, koji se zasnivaju na korišćenju ERBF (*Exponential Radial Basis Function*), *Laplacian*, *Generalized T-Student* i *Cauchy kernel* funkcije. Kako RSS vrednosti u okviru neke makro ćelije imaju raspodelu blisku Gausovoj, korišćena je kvadratna funkcija greške kao kriterijum za optimalno treniranje modela [3].

Treniranje, validacija i verifikacija predloženog SVR_{2D} modela izvršeni su na osnovu realnih podataka prikupljenih tokom merne kampanje u vidu *drive test*-a sprovedenog na teritoriji grada Beograda.



Slika 2. Grafički prikaz SVR_{2D} modela za pozicioniranje.

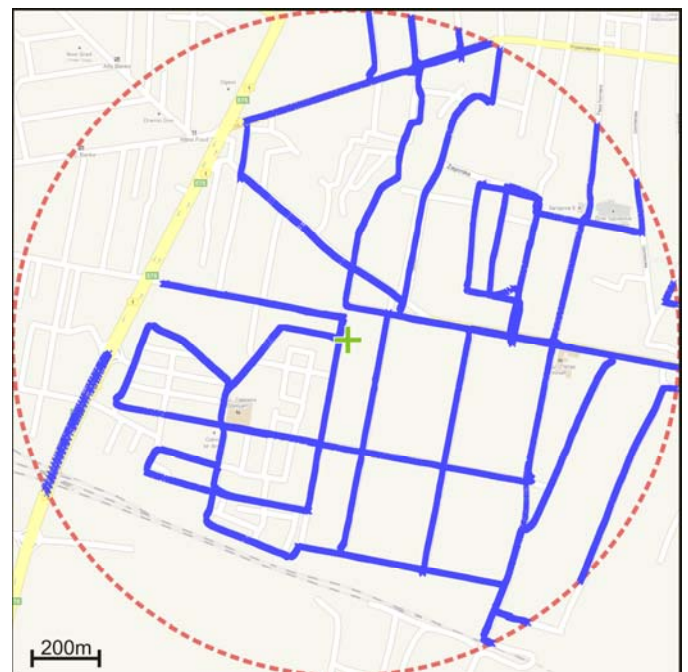
Merenje RSS vrednosti baznih stanica na pojedinim lokacijama (mernim tačkama) vršeno je pomoću *Rohde&Schwarz TSMQ* mrežnog skenera. Geografske koordinate datih lokacija dobijene su korišćenjem diferencijalnog GPS prijemnika, čija medijanska greška pozicioniranja iznosi manje od 5m. Za preuzimanje izmerenih RSS i GPS podataka korišćen je laptop sa „*R&S Romes v4*“ softverom.

Geografska oblast koju pokriva realizovani SVR_{2D} model je oblika kruga sa centrom u poziciji GSM bazne stanice na koju se odnosi dati model. Usvojeno je da se poluprečnik date oblasti definiše tako da se 99.9% svih mernih tačaka, u kojima najveća izmerena RSS vrednost pripada datoj BS, nađe unutar posmatrane zone. Tako dobijena zona pokrivanja posmatranog SVR_{2D} modela je poluprečnika 1km i obuhvata 31391 mernu tačku. Geografska oblast na koju se odnosi realizovani SVR_{2D} model prikazana je na Sl. 3. Zelenom bojom označena je pozicija GSM bazne stanice na koju se odnosi model, a plavom bojom ulice duž kojih je vršeno merenje. Pozicija BS je uzeta za referentnu, odnosno vrednosti x i y koordinata definisane su relativno u odnosu na nju kao koordinatni početak.

Sva četiri SVR_{2D} modela trenirana su sa podacima prikupljenim na 3139 slučajno odabranih lokacija, koje predstavljaju 10% od ukupnog broja mernih tačaka iz geografske oblasti koju pokriva model.

IV. REZULTATI TESTIRANJA I VERIFIKACIJE MODELA ZA POZICIONIRANJE

Preciznost predloženih SVR_{2D} modela ispitivana je kroz određivanje greške pozicioniranja (*Distance Error*, DE), koja predstavlja rastojanje između tačne i procenjene geografske lokacije mobilnog korisnika.



Slika 3. Geografska oblast pokrivanja realizovanog SVR_{2D} modela.

Krosvalidacija prethodno istreniranih modela izvršena je korišćenjem podataka sa 40% slučajno odabranih lokacija (12556 mernih tačaka), dok je preostalih 50% (15696 mernih tačaka) korišćeno za verifikaciju modela.

U okviru tabele 1 prikazani su rezultati verifikacije za četiri različita tipa SVR_{2D} modela. Prilikom treniranja, odabir optimalnih vrednosti za određene parametre SVR_{2D} modela (npr. vrednosti parametara *kernel* funkcija) vršen je tako da se prilikom krosvalidacije datih modela dobije minimalna srednja vrednost i standardna devijacija greške pozicioniranja. Potom je vršena verifikacija tako istreniranog modela.

Performanse predloženih SVR_{2D} modela analizirane su sa stanovišta srednje vrednosti greške pozicioniranja (Srednja DE), standardne devijacije greške pozicioniranja (σ DE), medijanske vrednosti greške pozicioniranja (Medijanska DE), kao i vrednosti greške pozicioniranja koja nije premašena u 67% slučajeva (67% DE), odnosno u 95% slučajeva (95% DE).

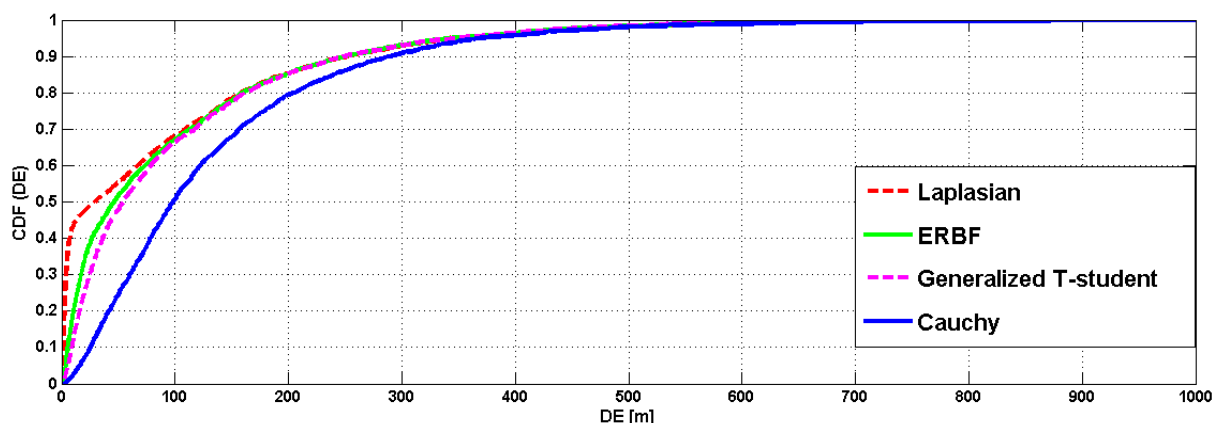
Na osnovu tabele 1 može se zaključiti da su najbolji rezultati po pitanju vrednosti greške pozicioniranja dobijeni za SVR_{2D} model sa *Laplacian kernel* funkcijom, kod koga je u 50% slučajeva greška pozicioniranja manja od 29.3m.

Kumulativna funkcija raspodele (CDF, *Cumulative Distribution Function*) vrednosti greške pozicioniranja za svaki od predloženih SVR_{2D} modela prikazana je na Sl. 4, dok su funkcije gustine verovatnoće (PDF, *Probability Density Function*) grešaka pozicioniranja za simulirane modele prikazane na Sl. 5. Na osnovu Sl. 4. može se videti da se najveći procenat malih DE vrednosti javlja u slučaju primene *Laplacian kernel* funkcije. Međutim, u slučaju *Laplacian* funkcije dobijena je i najveća vrednost standardne devijacije greške pozicioniranja (Sl. 5), tako da su sa stanovišta pojave pojedinih velikih vrednosti grešaka pozicioniranja performance SVR_{2D} modela, baziranih na korišćenju *Laplacian*, ERBF i *Generalized T-student kernel* funkcija, dosta slične. S druge strane, model baziran na upotrebi *Cauchy kernel* funkcije pokazao je lošije performance u odnosu na prethodna tri modela.

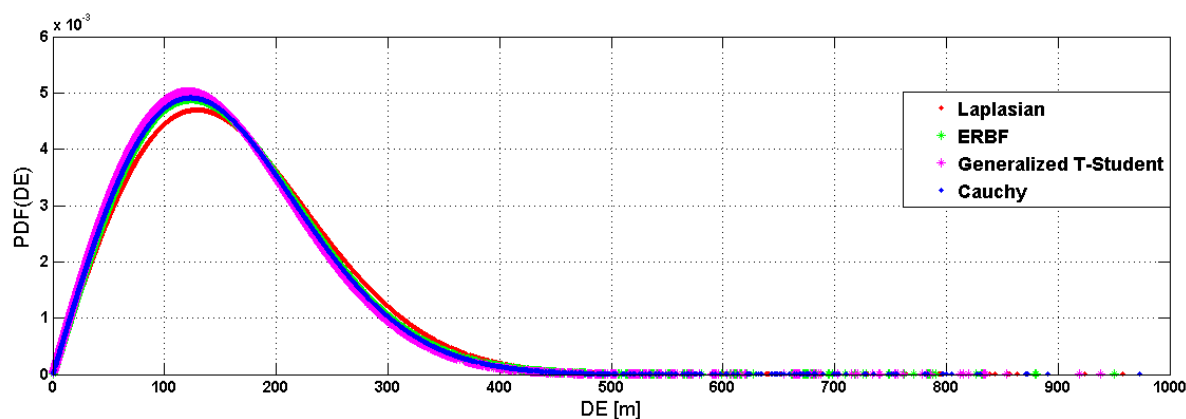
Pojedine velike vrednosti grešaka pozicioniranja mogu se eliminisati uvođenjem tkz. "algoritama za praćenje" ("*tracking algorithm*") [8], [9], gde se na osnovu prethodne pozicije mobilnog korisnika i brzine njegovog kretanja može proceniti da li je nova proračunata pozicija fizički moguća ili ne.

TABELA I. REZULTATI VERIFIKACIJE SVR_{2D} MODELA REALIZOVANIH KORIŠĆENJEM ERBF, LAPLASIAN, GENERALIZED T-STUDENT I CAUCHY KERNEL FUNKCIJA

Tip <i>kernel</i> funkcije	Greška pozicioniranja (<i>Distance Error</i> , DE) SVR _{2D} modela				
	Srednja DE [m]	Medijanska DE [m]	67% DE [m]	95% DE [m]	σ DE [m]
ERBF	96	46.2	99.6	344.8	123.9
<i>Laplacian</i>	88.9	29.3	94.3	342.3	129.1
<i>Generalized T-student</i>	100.2	54.1	102.6	346.9	120.8
<i>Cauchy</i>	134.6	98	146.1	366.1	123.5



Slika 4. Kumulativna funkcija raspodele vrednosti grešaka pozicioniranja (DE).



Slika 5. Funkcije gustine verovatnoće vrednosti greška pozicioniranja (DE).

V. ZAKLJUČAK

U okviru ovog rada predstavljena su četiri modela za određivanje pozicije mobilnog GSM korisnika korišćenjem RSS pristupa i SVR algoritma. Prednost predloženih modela je u činjenici da ne zahtevaju nikakve izmene na strani mobilnog terminala, kao ni izmene same signalizacije. Kako se poseban model za pozicioniranje implementira na nivou svake ćelije, neophodne promene na strani mreže mogu se uvoditi postepeno, u zavisnosti od potrebe za implementacijom LBS servisa na pojedinim teritorijama. Takođe, predložen model može se bez izmena primeniti i na određivanje lokacije mobilnih DCS i UMTS korisnika. Osnovna prednost korišćenja samog SVR algoritma, kao *pattern analysis* metode, jeste u činjenici da ne zahteva poznavanje i modelovanje specifičnih uslova propagacije signala koji postoje u urbanom okruženju, za razliku od tehnika pozicioniranja koje se zasnivaju na statističkim i propagacionim modelima.

Povećanje preciznosti predložene tehnike moguće je dalje vršiti povećanjem broja RSS vrednosti na osnovu kojih se vrši estimacija pozicije mobilnog korisnika, odnosno korišćenjem istovremeno infrastrukture GSM, DCS i UMTS sistema [3], [8]. Dodatno smanjenje greške pozicioniranja moguće je u slučaju korišćenja infrastrukture više operatora. Na taj način moguće je iz različitih GSM/UMTS signala dobiti više informacija o samom prostoru jer se zbog veće prostorne razdvojenosti baznih stanica smanjuje stepen međusobne korelisanosti datih signala [3]. Takođe, problem pojedinačnih velikih vrednosti greške pozicioniranja može se dodatno rešiti korišćenjem nekog od *tracking* algoritama. Povećanje preciznosti samog SVR algoritma može se realizovati deljenjem oblasti pokrivanja predloženog modela na manje podprostore, čime se ostvaruju bolji početni uslovi za rešavanje problema regresije.

LITERATURA

- [1] http://hraunfoss.fcc.gov/edocs_public/attachmatch/DOC-308377A1.pdf
- [2] L. Zhang, C. Tao, G. Yang, "Wireless positioning: fundamentals, systems and state of the art signal processing techniques," InTech, Rijeka, Croatia, 2011.
- [3] M. Borenović, A. Nešković, D. Budimir, "Multi-system-multy-operator localization in PLMN using neural networks", International Journal of Communication System, Wiley Online Library, 2011.
- [4] H. Laitinen, J. Lahteenmaki, T. Nordstorm, "Database correlation method for GSM localization", IEEE Vehicular Technology Conference, 2001, pp. 2504–2508.
- [5] Z.Wu, C. Li, "Location estimation via Support Vector Regression", IEEE Transactions on mobile computing, March 2007, vol.6, no. 3, pp. 311-321.
- [6] C. M. Bishop, "Pattern recognition and machine learning," Springer Science+Business Media, LLC, New York, USA, 2006.
- [7] J. Shawe-Taylor, N. Cristianini, "Kernel methods for pattern analysis", Cambridge University Press, New York, USA, 2004.
- [8] P. Kemppli, S. Niousiainen "Database Correlation Method for multi-system positioning", IEEE Vehicular Technology Conference, May 2006, pp. 866–870.
- [9] C.M. Takenga, Q. Wen, K. Kyamakya, "On the accuracy improvement issues in GSM location fingerprinting" IEEE Vehicular Technology Conference, September 2006, pp. 1–5.

ABSTRACT

The obligation of mobile operators to implement emergency call services, as well as the big interest in developing commercial LBS (*Location Based Services*), have led to the intensive development of techniques for determination of mobile user's position in the last few years. The main goal of these researches is to achieve the biggest possible accuracy, but with minimal changes in terms of network or mobile handset. In this paper, the model for localization of GSM mobile users using SVR (*Support Vector Regression*) will be proposed. Four types of proposed model have been developed, by using different types of kernel functions. The accuracy of proposed models has been tested using real data, measured in urban environment.

NOVEL METHOD FOR GSM MOBILE USERS' LOCALIZATION BASED UPON SUPPORT VECTOR REGRESSION

M. Petrić, M. Borenović, N. Nešković, A. Nešković