

Indoor model predikcije nivoa električnog polja baziran na SVM algoritmu

Vladimir Slavković, Aleksandar Nešković, Nataša Nešković
Laboratorija za radiokomunikacije, Katedra za telekomunikacije
Elektrotehnički fakultet Univerziteta u Beogradu
Beograd, Republika Srbija
vslavkovic@etf.rs, neshko@etf.rs, natasha@etf.rs

Sadržaj—U ovom radu predstavljen je novi indoor model predikcije nivoa električnog polja. Pri razvoju modela korišćen je Support Vector Machine (SVM) algoritam, čime su prevaziđeni neki od osnovnih nedostataka postojećih determinističkih i statističkih modela. Razvoj i analiza performansi predloženog SVM modela zahteva generisanje trening, validacionog i test skupa podataka. U tu svrhu sprovedena su merenja nivoa električnog polja u okviru prizemlja zgrade Elektrotehničkog fakulteta u Beogradu. U okviru rada analiziran je uticaj veličine trening skupa na performanse predloženog modela predikcije. Na osnovu ove analize dat je predlog o potrebnoj veličini trening skupa, nakon čega su prikazane performanse razvijenog modela. Analiza dobijenih rezultata je pokazala da je model brz i pouzdan.

Ključne reči—indoor; model predikcije; radio propagacija; SVM algoritam.

I. UVOD

Predikcija nivoa električnog polja je veoma kompleksan i težak zadatak. Razvijen je veliki broj različitih modela predikcije [1]-[4]. Postojeći modeli predikcije se grubo mogu podeliti na statističke (bazirane na empirijskim podacima) i determinističke (koji slede neki mehanizam propagacije) modele.

Statistički modeli implicitno uzimaju u razmatranje uticaje kompletnog okruženja. U cilju razvoja statističkih modela neophodno je sprovoditi merenja nivoa električnog polja. Problem koji ispoljavaju statistički modeli odnosi se na njihovu tačnost. Tačnost statističkih modela zavisi ne samo od tačnosti merenja, već i od sličnosti između okruženja koje se analizira i okruženja gde su sprovedena merenja. Sa druge strane, numerička efikasnost ovih modela je zadovoljavajuća.

Deterministički modeli se baziraju na principima fizičke propagacije signala, pa se stoga mogu primeniti na različita okruženja bez uticaja na tačnost predikcije. Osnovni problem determinističkih modela predstavlja njihova proračunska efikasnost, naročito u slučaju kada ih je neophodno primeniti na celokupan sistem, odnosno na veliki broj baznih stanica. Pored toga, implementacija ovih modela često zahteva velike baze podataka, koje je ponekad veoma teško ili čak nemoguće generisati. Međutim, ukoliko se deterministički modeli ispravno implementiraju moguće je postići veću tačnost predikcije nego u slučaju statističkih modela [3].

Dakle, osnovni problem statističkih modela predstavlja njihova tačnost, dok kod determinističkih modela proračunska efikasnost nije na zadovoljavajućem nivou. Osnovne nedostatke navedenih modela moguće je prevazići razvojem modela predikcije korišćenjem SVM (*Support Vector Machine*) algoritma. SVM algoritam je već uspešno implementiran za predikciju nivoa električnog polja u mikročelijskom okruženju [5].

SVM algoritam spada u grupu novih *machine learning* algoritama, koji se zasnivaju na teoriji statističkog učenja [6, 7]. Osnovna ideja SVM algoritma jeste da se u skupu konačnog broja podataka uoče modeli (relacije) koji za taj skup važe. Učenjem ovih relacija SVM algoritam stiče osobinu generalizacije, koja zapravo znači da će algoritam biti u stanju da izvrši predikciju nad novim podacima generisanim od strane istog izvora.

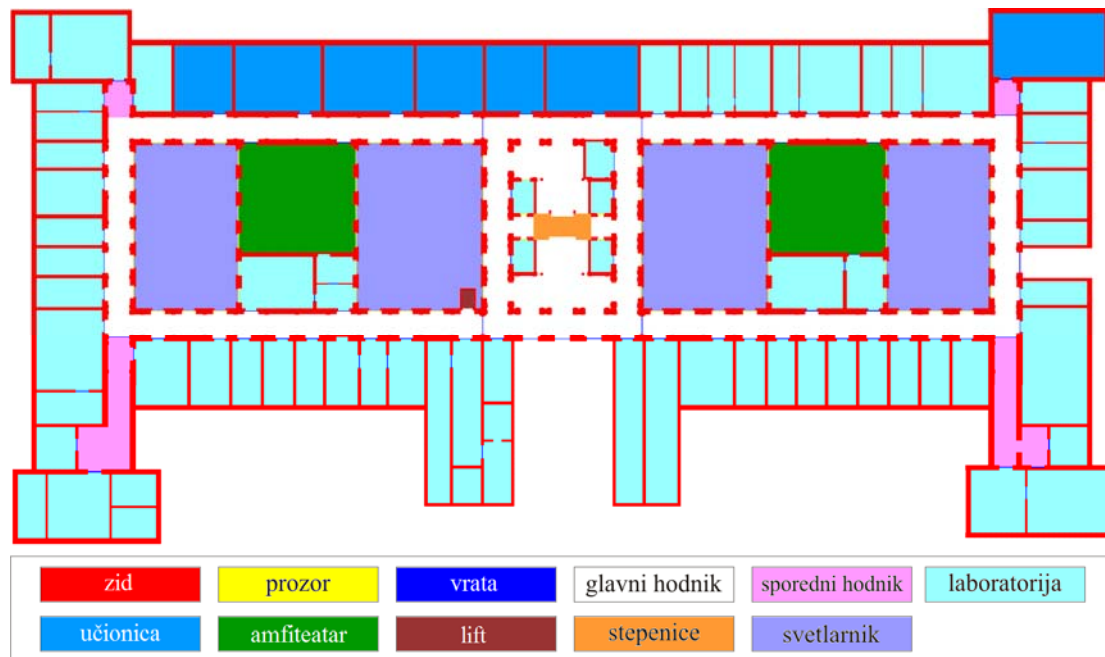
Novi model je realizovan za frekvencijski opseg 900MHz. Međutim, treba napomenuti da se generalna ideja predloženog modela može jednostavno primeniti na bilo koji frekvencijski opseg.

II. BAZA PODATAKA

Tačnost modela predikcije u velikoj meri zavisi od tačnosti korišćene baze podataka. Implementacija SVM modela zahteva bazu podataka dobijenu na osnovu 2D plana indoor okruženja, pri čemu je posmatrano okruženje klasifikovano u nekoliko kategorija. Klasifikaciju okruženja je najjednostavnije sprovoditi tako što se na osnovu skeniranog 2D plana okruženja napravi kolor slika, na kojoj su različitim bojama predstavljene različite definisane kategorije okruženja. U ovom radu definisano je 11 različitih kategorija okruženja: zid, vrata, prozor, glavni hodnik, sporedni hodnik, laboratorija/kancelarija, učionica, amfiteatar, lift, svetlarnik (spoljašnji prostor) i stepenice. Na prethodno opisani način kreirana je baza podataka za prizemlje zgrade Elektrotehničkog fakulteta u Beogradu (Sl. 1). Rezolucija kreirane baze podataka iznosi 5×5 cm.

III. SVM MODEL

SVM model predstavljen u radu bazira se na primeni kernel tehnike [6, 7]. Primenom kernel tehnike podaci se implicitno preslikavaju u pogodno izabran karakterističan prostor, gde



Slika 1. Klasifikovano okruženje prizemlja zgrade Elektrotehničkog fakulteta u Beogradu (150.2 × 72.9 m).

relacije koje se mogu uočiti u takvom skupu podataka postaju linearne. Kernel tehnika se zapravo zasniva na primeni kernel funkcije. Nakon preslikavanja podataka u karakterističan prostor koriste se algoritmi bazirani na linearnoj algebri, geometriji i statistici kako bi se u skupu preslikanih podataka otkrio odgovarajući model [6, 7].

SVM model ima 13 ulaznih parametara i jedan izlaz. Uzimajući u obzir da svi ulazi SVM modela treba da imaju vrednosti u intervalu [0,1], definisani su sledeći ulazni parametri:

- normalizovano rastojanje (stvarno rastojanje podeljeno sa 200 m) između predajnika (Tx) i prijemnika (Rx). Treba napomenuti da u slučaju razmatranog okruženja, maksimalno moguće rastojanje Tx-Rx iznosi približno 200 m, pa je u skladu sa tim izvršena normalizacija ovog parametra;
- procenat zastupljenosti zidova, glavnog hodnika, sporednog hodnika, laboratorija/kancelarija, učionica, amfiteatara, liftova, svetlarnika i stepenica na direktnoj putanji Tx-Rx (devet ulaza);
- broj prozora na direktnoj putanji Tx-Rx;
- broj vrata na direktnoj putanji Tx-Rx;
- relativni težinski faktor koji opisuje tip okruženja u blizini prijemnika (primenom ad hoc pravila usvojeni su sledeći različiti težinski faktori za definisane kategorije okruženja: 0.2 – glavni hodnik, 0.3 – sporedni hodnik, 0.4 – laboratorija/kancelarija, 0.5 - učionica).

Svi prethodno navedeni ulazni parametri određuju se automatski na osnovu baze podataka 2D plana indoor okruženja. Kao što se može videti, većina ulaznih parametara definisana je duž prave linije između predajnika i prijemnika.

Slična strategija definisanja ulaznih parametara, doduše u slučaju mikročelijskog modela predikcije, može se pronaći u [4, 5].

Predloženi SVM model ima samo jedan izlazni parametar, a to je normalizovani nivo električnog polja (stvarni nivo podeljen sa 120 dB μ V/m). Treba napomenuti da maksimalni očekivani nivo električnog polja u slučaju razmatranog *indoor* okruženju iznosi 120 dB μ V/m, pa je u skladu sa tim izvršena normalizacija izlaznog parametra.

IV. REALIZACIJA MODELA

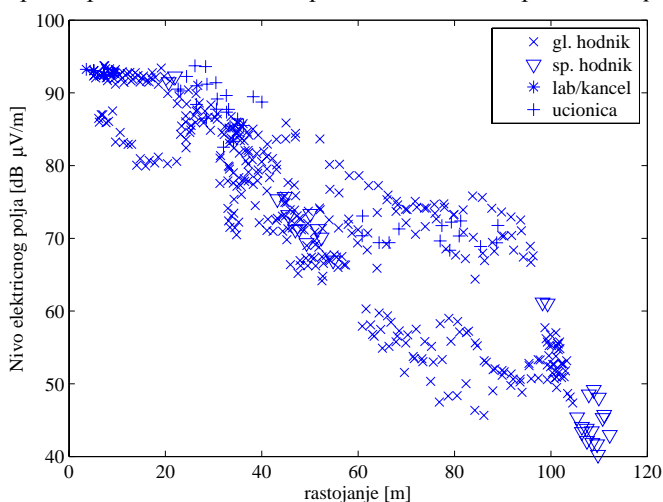
U cilju razvoja i verifikacije predloženog modela, sprovedena su merenja nivoa električnog polja u okviru prizemlja zgrade Elektrotehničkog fakulteta u Beogradu. Zgradu karakterišu debeli spoljašnji i unutrašnji zidovi, velika drvena vrata i prozori sa drvenim okvirima srednje veličine. Tokom merenja, nemodulisani (CW - *continuous wave*) RF nosilac snage 1W emitovan je korišćenjem vertikalno polarizovane 2 dBd omnidirekcionih antena. Visina predajne antene bila je 2 m. U cilju automatskog prikupljanja podataka prijemnik (*Rohde&Schwarz ESPI Test Receiver*) povezan je sa PC laptopom i postavljen na pokretna kolica. Prijemna antena *Anritsu MP524C* (merna kalibrisana antena za opseg 520-1000 MHz) bila je montirana na visini od 1.75m. Period odabiranja sistema iznosio je 10ms. Tokom merenja pozicija pokretnih kolica snimana je ručno. Prikupljeno je ukupno preko 600 000 vrednosti nivoa električnog polja. Rezultati merenja ukazuju na veliki dinamički opseg nivoa električnog polja (oko 76 dB).

Brzi feding je eliminisan usrednjanjem izmerenih nivoa električnog polja na $\pm 5\lambda$ (± 1.65 m) intervalu. Interval usrednjanja izabran je u skladu sa tipom razmatranog okruženja, a prema [8]. Kako bi se ispitala merna ponovljivost, odnosno utvrdila maksimalna dostižna tačnost modela predikcije nivoa električnog polja, merenja nivoa električnog

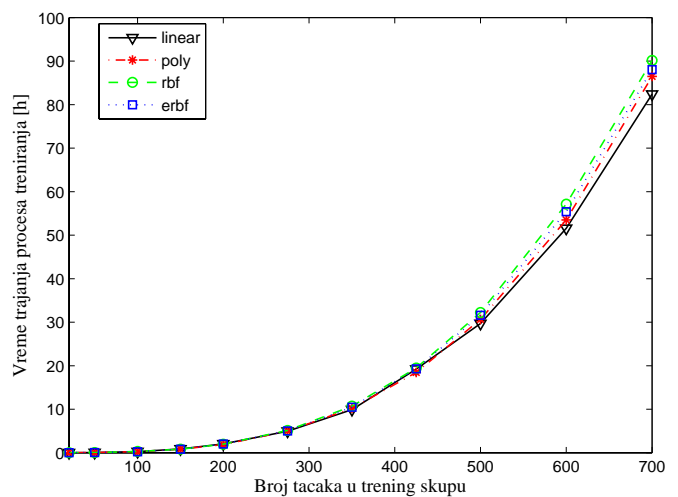
polja su ponovljena za nekoliko stotina tačaka. Na osnovu prvobitnih i ponovljenih merenja utvrđeno je da srednja greška iznosi -0.36 dB, a standardna devijacija 1.36 dB. Navedene vrednosti definišu maksimalnu dostižnu tačnost modela predikcije nivoa električnog polja. U cilju postizanja što realnijih uslova propagacije radio signala, merenja su sprovedena tokom aktivnog dela dana, čime je uzet u obzir uticaj ljudi.

Na osnovu merenja formirano je ukupno 4625 setova podataka, pri čemu se jedan set podataka sastoji od 13 ulaznih parametara i jednog izlaza. Ovi podaci su podeljeni u tri skupa: trening, validacioni i test skup. Trening skup je korišćen za treniranje algoritma. Validacioni skup je korišćen za izbor optimalnih parametara optimizacije (p_1 – parametar kernel funkcije, i C – *box constraint*). Više o optimizacionim parametrima moguće je pronaći u [6, 7]. Konačno, test skup je korišćen za procenu performansi obučenog algoritma na novim podacima, odnosno na podacima koji nisu bili dostupni algoritmu u toku trening faze. Veličina trening skupa je varirana kako bi se ispitao uticaj veličine trening skupa na performanse razvijenog modela. Razmatrane su sledeće veličine trening skupa: 20, 50, 100, 150, 200, 275, 350, 425, 500, 600 i 700. U okviru validacionog skupa izdvojeno je 150 setova podataka, dok se u okviru test skupa nalaze svi preostali podaci. Na Sl. 2 prikazani su izmereni nivoi električnog polja, pri čemu je radi preglednosti prikazan svaki deseti izmereni nivo električnog polja. Usvojena je pretpostavka da se prijemnik može naći na glavnom hodniku, sporednom hodniku, u učionici ili u laboratoriji/kancelariji.

U ovom radu izvršena je analiza primene četiri različite kernel funkcije na problem predikcije nivoa električnog polja u *indoor* okruženju: linearna, polinomska, rbf (*radial basis function*) i erbf (*exponential rbf*) kernel funkcija. U cilju razvoja modela korišćen je *svm toolbox* MATLAB softverskog paketa [9, 10]. Ovaj toolbox između ostalih uključuje i sledeće funkcije: *svdatanorm*, *svr* i *svroutput*. Funkcija *svdatanorm* normalizuje podatke na opseg [-1, 1], funkcija *svr* se koristi za treniranje modela, odnosno učenje relacija koje u dostupnom treninng skupu podataka postoje, dok se funkcija *svroutput* koristi za predikciju nivoa električnog polja. Treniranje modela zapravo podrazumeva izbor optimalnih vrednosti parametara p_1



Slika 2. Izmereni nivoi električnog polja.

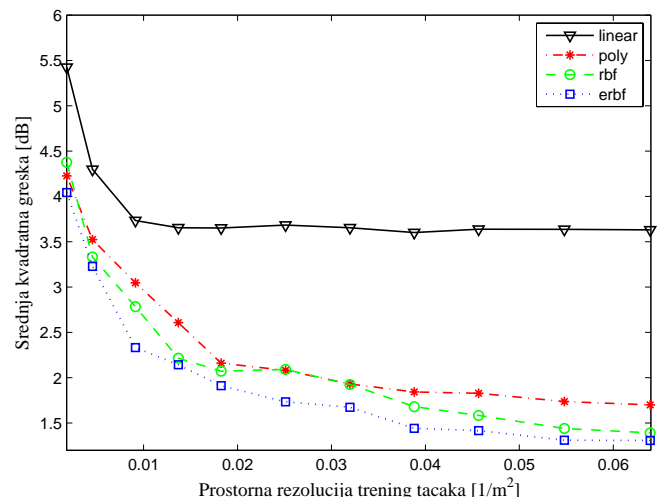


Slika 3. Zavisnost vremena trajanja treniranja i veličine trening skupa.

i C . Vrednosti ovih parametara birane su tako da srednja kvadratna greška (*root mean square* - rms) na validacionom skupu bude minimalna.

Vreme trajanja procesa treniranja SVM modela prevashodno zavisi od veličine trening skupa. Zavisnost vremena trajanja procesa treniranja SVM modela i veličine trening skupa, za različite kernel funkcije, prikazana je na Sl. 3. Očekivano, porast veličine trening skupa dovodi do povećanja vremena trajanja procesa treniranja. Treba primetiti da je vreme trajanja procesa treniranja gotovo identično za sve razmatrane kernel funkcije. Generalno, proces treniranja SVM modela, za adekvatne veličine trening skupa traje nekoliko sati do nekoliko desetina sati na standardnom PC računaru (Intel(R) Core(TM)2 Duo CPU E7200 @ 2.53 GHz, 2.00 GB RAM).

Tačnost SVM modela predikcije zavisi od veličine trening skupa. Zavisnost srednje kvadratne greške predikcije od prostorne rezolucije trening tačaka, za različite kernel funkcije, prikazana je na Sl. 4. Očekivano, kako broj tačaka u trening skupu raste, raste i tačnost SVM modela. Srednja kvadratna greška predikcije naglo opada sve dok prostorna rezolucija trening tačaka ne dostigne vrednost 0.02 1/m^2 , osim u slučaju



Slika 4. Zavisnost srednje kvadratne greške predikcije i prostorne rezolucije trening tačaka.

linearne kernel funkcije gde se ovaj efekat javlja za nešto nižu vrednost prostorne rezolucije. Daljim porastom prostorne rezolucije srednja kvadratna greška opada znatno sporije. Otuda je u ovom radu za vrednost prostorne rezolucije trening tačaka izabrana upravo vrednost 0.02 1/m^2 , što za slučaj razmatranog okruženja približno odgovara trening skupu sa 200 setova podataka. Ovakav trening skup garantuje zadovoljavajuće performanse modela, a ujedno i zadržava vreme rajanja procesa treniranja u realnim granicama. Treba primetiti da linearna kernel funkcija poseduje najlošije performanse, tako da čak i slučaju visoke prostorne rezolucije trening tačaka. srednja kvadratna greška i dalje ostaje iznad 3.5 dB. Sa druge strane, erbf funkcija pokazuje najbolje performanse među analiziranim kernel funkcijama.

V. VERIFIKACIJA MODELA

Kvalitet razvijenog SVM modela verifikovan je nad test skupom podataka, koji nije bio uključen u proces treniranja. Predikcija je dosta brza i traje nekoliko minuta na standardnom PC računaru. Vrednosti statističkih parametara grešaka, koje opisuju performanse predloženog modela za usvojenu veličinu trening skupa (200 setova podataka), za različite kernel funkcije prikazane su u tabeli 1. Vrednosti prikazane u tabeli potvrđuju prethodno iznetu konstataciju da erbf predstavlja najbolji izbor među analiziranim kernel funkcijama. Sve vrednosti statističkih parametara koje su dalje navedene u radu odnose se na erbf kernel funkciju.

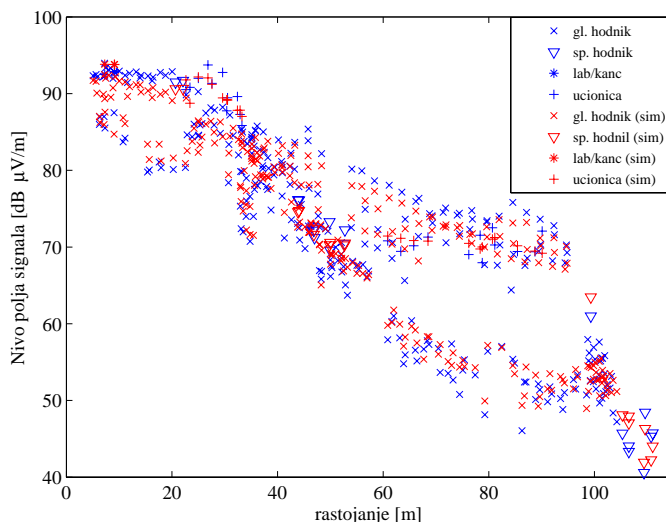
Na Sl. 5 prikazani su rezultati predikcije nivoa električnog polja, zajedno sa odgovarajućim izmerenim vrednostima. Iako je nivo električnog polja multivarijabilna funkcija, prikazana je samo zavisnost nivoa od rastojanja. Kao što se sa slike može zaključiti, rezultati predikcije u velikoj meri prate izmerene nivoove električnog polja.

TABELA I. VREDNOSTI STATISTIČKIH PARAMETARA GREŠKE PREDIKCIJE (Δ - SREDNJA GREŠKA, σ - STANDARDNA DEVIJACIJA).

	Kernel funkcija							
	Linearna		Polinomska		RBF		ERBF	
	Δ (dB)	σ (dB)	Δ (dB)	σ (dB)	Δ (dB)	σ (dB)	Δ (dB)	σ (dB)
Model testiran nad trening podacima	-0.07	3.50	0.06	1.29	0.06	1.28	0.10	1.17
Model testiran nad test podacima	-0.05	3.65	0.08	2.16	0.10	2.07	0.09	1.91

TABELA II. RAZLIKA IZMEĐU IZMERENOG I PROCENJENOG NIVOVA ELEKTRIČNOG POLJA (Δ - SREDNJA GREŠKA, σ - STANDARDNA DEVIJACIJA).

Tip okruženja	ETF	
	Δ (dB)	σ (dB)
Glavni hodnik	0.12	1.92
Sporadni hodnik	-0.22	2.10
Laboratorija/Kancelarija	-1.07	0.18
Učionica	0.07	1.62
Σ	0.09	1.91



Slika 5. Poređenje izmerenog i procenjenog nivoa električnog polja.

Vrednosti statističkih parametara za pojedine kategorije okruženja prikazane su u tabeli 2. Standardna devijacija greške redukcije iznosi 1.91 dB. Imajući u vidu veliku dinamiku merenog signala (oko 76 dB), može se zaključiti da je tačnost razvijenog modela na zavidnom nivou. Treba napomenuti da e u toku ispitivanja merne ponovljivosti uočeno da standardna devijacija prvobitnih i ponovljenih rezultata merenja za tip okruženja laboratorija/kancelarija iznosi 0.17 dB. Razlog dosta niske vrednosti standardne devijacije krije se u činjenici da je laboratorija/kancelarija u kojoj su vršena merenja, neposredno lizu predajnika, tako da nivo električnog polja za ovakav tip okruženja ne ispoljava velike promene u toku vremena. Otuda ne treba da čudi niska vrednost standardne devijacije greške predikcije u slučaju okruženja laboratorija/kancelarija.

VI. ZAKLJUČAK

U radu je predstavljen novi SVM model predikcije nivoa električnog polja za indoor okruženje pomoću kojeg se mogu prevazići neki od osnovnih nedostataka statističkih i determinističkih modela predikcije. Predloženi model dosta suptilnije uzima u obzir karakteristike okruženja nego standardni statistički modeli, što rezultuje većom tačnošću modela. Sa druge strane, SVM model nije toliko proračunskih zahtevan kao deterministički modeli. Takođe, implementacija predloženog modela zahteva relativno jednostavne baze podataka.

Analiza performansi SVM modela sprovedena je u okviru prizemlja zgrade Elektrotehničkog fakulteta u Beogradu, koju karakterišu debeli spoljašnji i unutrašnji zidovi, velika drvena vrata i prozori sa drvenim okvirima srednje veličine. Nivo električnog polja u posmatranom okruženju ispoljava veliku dinamiku (oko 76 dB). U okviru rada izvršena je analiza uticaja veličine trening skupa na performanse predloženog modela. Na osnovu ove analize došlo se do zaključka da prostorna rezolucija trening tačaka treba da iznosi 0.02 1/m^2 . Ova vrednost prostorne rezolucije trening tačaka garantuje zadovoljavajuće performanse modela, a ujedno i zadržava vreme trajanja procesa treniranja u realnim granicama. Standardna devijacija

greške predikcije modela iznosi 1.91 dB. Analiza dobijenih rezultata pokazuje da je model brz i pouzdan. U izuzetnim slučajevima moguće je dodatno poboljšati performanse SVM modela dodatnim treniranjem.

LITERATURA

- [1] P. Constatinou, "Properties of wireless RF channels", Wireless LAN systems, Artech House, pp. 129 – 163, 1994.
- [2] K. Pahlavan, A.H. Levesque, "Wireless Information Networks", 2nd ed., Wiley, New Jersey, 2005.
- [3] J. B. Anderson, T. S. Rappaport, S. Yoshida, "Propagation measurements and models for wireless communications channels", IEEE Commun Mag, vol. 33, no. 1, pp. 42-49, 1995.
- [4] W.C.Y. Lee, D.J.Y. Lee, "Microcell prediction in dense urban area", IEEE Trans Veh Tech, vol.47, no. 1, 1998.
- [5] V. Slavkovic, A. Nešković, N. Nešković, "Microcell prediction model based on support vector machine algorithm", Ann. Telecommun., Springer, 2013.
- [6] C. M. Bishop, "Pattern recognition and machine learning", Springer, LLC, New York, USA, 2006.
- [7] J. Shawe-Taylor, N. Cristianini, "Kernel methods for pattern analysis", Cambridge University Press, New York, USA, 2004.
- [8] W.C.Y. Lee, "Estimate of Local Average Power of a Mobile Radio Signal", IEEE Trans Veh Tech, vol. VT-34, no. 1, pp. 22-27, 1985.

[9] S. Gunn, "Support vector machines for classification and regression", Technical Report, School of Electronics and Computer Science, University of Southampton, 1998.

[10] <http://www.isis.ecs.soton.ac.uk/resources/svminfo/>

ABSTRACT

A new indoor prediction model for mobile radio environment is presented in this paper. The Support Vector Machine (SVM) algorithm is used to build the model, making possible to overcome some important disadvantages of both deterministic and statistical models. Development and performance analysis of the proposed model requires training, validation and test data sets. For that purpose, extensive electric field strength measurements were carried out within the ground floor of Faculty of Electrical Engineering in Belgrade. The influence of training set size on model performance is also analyzed. The analysis of the results has shown that proposed model is fast and reliable.

NOVEL INDOOR ELECTRIC FIELD STRENGTH PREDICTION MODEL

Vladimir Slavković, Aleksandar Nešković, Nataša Nešković