

НАЗИВ ФАКУЛТЕТА Факултет техничких наука

ИЗВЕШТАЈ О ОЦЕНИ ДОКТОРСКЕ ДИСЕРТАЦИЈЕ

-обавезна садржина- свака рубрика мора бити попуњена

(сви подаци уписују се у одговарајућу рубрику, а назив и место рубрике не могу се мењати или изоставити)

I ПОДАЦИ О КОМИСИЈИ
<p>1. Датум и орган који је именовao комисију На основу предлога матичне катедре и одлуке Наставно-научног већа Факултета техничких наука, Декан Факултета техничких наука донео је решење број 012-199/09-2017 од 31. 08. 2017.</p> <p>2. Састав комисије са знаком имена и презимена сваког члана, звања, назива уже научне области за коју је изабран у звање, датума избора у звање и назив факултета, установе у којој је члан комисије запослен:</p> <p>Др Александар Купусинац, ванредни професор, члан Комисије, председник Примењене рачунарске науке, 19.05.2016., Факултет техничких наука, Нови Сад</p> <p>Др Дарко Стефановић, доцент, члан Комисије, Информационо-комуникациони системи, 15.10.2012., Факултет техничких наука, Нови Сад</p> <p>Др Дејан Вукобратовић, ванредни професор, члан Комисије Телекомуникације и обрада сигнала, 01.04.2014., Факултет техничких наука, Нови Сад</p> <p>Др Владимир Божовић, ванредни професор, члан Комисије Математика, 12.06.2015., Природно-математички факултет у Подгорици</p> <p>Др Дубравко Ћулибрк, ванредни професор, члан Комисије, ментор Информационо-комуникациони системи, 17.01.2013., Факултет техничких наука, Нови Сад</p>
II ПОДАЦИ О КАНДИДАТУ
<p>1. Име, име једног родитеља, презиме: Катарина, Душан Гаврић</p> <p>2. Датум рођења, општина, држава: 14.11.1985., Сента, Република Србија</p> <p>3. Назив факултета, назив студијског програма дипломских академских студија – мастер и стечени стручни назив Факултет техничких наука, Индустрijско инжењерство и менаџмент, Дипломирани инжењер менаџмента - мастер</p> <p>4. Година уписа на докторске студије и назив студијског програма докторских студија 2010. / 2011., Индустрijско инжењерство / Инжењерски менаџмент</p> <p>5. Назив факултета, назив магистарске тезе, научна област и датум одбране: /</p>
<p>6. Научна област из које је стечено академско звање магистра наука: /</p>

III НАСЛОВ ДОКТОРСКЕ ДИСЕРТАЦИЈЕ:

Истраживање великих количина података о покретним објектима

IV ПРЕГЛЕД ДОКТОРСКЕ ДИСЕРТАЦИЈЕ:

Навести кратак садржај са назнаком броја страна, поглавља, слика, шема, графикона и сл.

Докторска дисертација кандидаткиње Катарине Гаврић, под насловом **Истраживање великих количина података о покретним објектима**, изложена је у 5 поглавља на 121 страни. Попис коришћене литературе са 150 наслова наведен је на 12 страна, а садржај дисертације на 2 стране. Докторска дисертација садржи 12 табела, 28 слика и 2 прилога. Садржај докторске дисертације је следећи:

1. Увод
2. Приказ релевантног истраживања у области
3. Методологија (имплементација)
4. Резултати истраживања
5. Закључци
6. Прилози
7. Литература

VI ВРЕДНОВАЊЕ ПОЈЕДИНИХ ДЕЛОВА ДОКТОРСКЕ ДИСЕРТАЦИЈЕ:

Поглавље 1 – Увод

У уводном поглављу кандидаткиња истиче значај утицаја савремених информационих технологија на генерисање огромне количине података о локацији људи и покретних објеката. Ови подаци представљају ресурс који се може искористити како би се постигао много виши степен разумевања навика, кретања и образаца понашања људи у циљу развијања модела од значаја за различите научне области и унапређење техничких решења у гранама попут урбанизма, маркетинга, епидемиологије итд. Примарни извор података о локацији и кретању људи данас представљају мобилни телефони, било да се ради о записима о активностима корисника мобилних телефона (енгл. Caller Data Records (CDRs)) које прикупљају телефонске компаније за потребе наплате услуга, или о директним записима о локацији које генеришу паметни телефони коришћењем позиционих система (енгл. Global Positioning System (GPS)). Као последица тога, последњих година је дошло до наглог скока количине јавно доступног кориснички-генерисаног садржаја, што је мотивисало истраживаче да развију начине за његову анализу како би открили шаблоне и трендове скривене у том садржају и на основу њих дошли до нових сазнања. Ти напори често подразумевају употребу техника и метода истраживања података (енгл. Data Mining) како би се решили различити проблеми: од детекције атрактивних локација, откривања карактеристичних путања којима се људи крећу, до естимације будућих вредности ризика ширења инфективних болести.

У складу са тим, кандидаткиња је навела циљеве дисертације као и мотивацију за њену израду. Као основни циљ се наводи евалуација могућности коришћења све веће количине јавно доступних података о локацији и кретању људи како би се дошло до нових сазнања, развили модели понашања и кретања људи и применили за решавање практичних проблема. Иако у данашње време постоји значајан број студија које се баве овим или сличним истраживачким питањима, још увек не постоји адекватан универзални модел за разумевање високо информативне димензије географског простора на основу података које корисници генеришу. Појавом и брзим растом "нетрадиционалних" података, отворене су могућности за проналажење занимљивих, корисних образаца у кретању, навикама и понашању корисника, што је свакако представљао изазов за истраживање ових и сличних питања у оквиру ове дисертације.

Поглавље 2 – Приказ истраживања у области

У овом поглављу кандидаткиња је направила свеобухватан преглед релевантне литературе и приказала тренутно стање у области. То обухвата: истраживање података генерисаних од стране мобилних телефона, као и кориснички-генерисаног мултимедијалног садржаја, људску динамику, активности и мобилност, као и истраживање података у области дигиталне епидемиологије.

Такође, описани су процеси обраде података у области истраживања података и откривања знања, као и кратак историјат развоја кључних модела који се данас користе за истраживање у предметној области. Поред тога, приказани су и примери метаподатака, као извора података који су касније коришћени приликом анализа, као и кратак опис друштвених мрежа за дељење мултимедијалног садржаја.

Поглавље 3 – Методологија (имплементација)

У трећем поглављу је приказана методологија која је коришћена приликом истраживања у оквиру ове тезе. У првом делу поглавља су приказани различити скупови података који су коришћени током истраживачког рада: скуп података са друштвених платформи Flickr и YouTube, демографски подаци са странице Africor, два скупа података преваленције вируса сиде на Обали Слоноваче (ДХС подаци), као и Д4Д скуп података који је садржао записе мобилних телефона компаније "Оранж", такође за територију Обале Слоноваче. Свеобухватно истраживање је подељено у четири фазе: откривање локација које корисници посећују (детекцију атрактивних локација применом техника кластеризације и визуализације података), откривање карактеристичних трајекторија корисника (путања кретања у посматраном подручју) и анализу њихових навика везаних за путовања, анализу метаподатака придружених корисничким сервисима (Flickr) као алтернативу записима мобилних телефона и естимацију будућих вредности ризика ширења инфективних болести.

За анализу података у оквиру прве три фазе коришћене су методе за истраживање података: селекција обележја и алгоритми за нумеричку предикцију, кластеровање, технике визуализације велике количине гео-референцираних података (топлотне мапе, OPTICS алгоритам за кластеровање), као и технике статистичке дистрибуције: закон скраћене снаге (енгл. Truncated Power law), Lévy flight, као и модели зановани на функцијама густине вероватноће (енгл. Probability Density Function). Подаци записа мобилних телефона који су коришћени у последњој фази су анализирани применом статистичких метода: регресија гребена (енгл. Ridge Regression), регресија вектора (енгл. Support Vector Regression), те рекурзивна елиминација атрибута (енгл. Recursive Feature Elimination). Коришћене методе су обрађене и приказане у овом поглављу.

Поглавље 4 – Резултати истраживања

У овом поглављу, кандидаткиња је евалуирала низ метода истраживања података на различитим скуповима података, а резултате упоредила са реалним статистичким подацима. У оквиру истраживања о карактеристичним трајекторијама и навикама везаних за путовања, показано је да YouTube видео снимци могу да се користе за идентификацију основних образаца у људском понашању, као и за анализу временске динамике и природе њихових активности. У оквиру наредног истраживања где су Flickr подаци коришћени као алтернатива записима мобилних телефона, кандидаткиња је показала да се законитости које се могу издвојити посматрањем података мобилних телефона, могу пронаћи и у Flickr подацима, као и да јавно доступни подаци имају потенцијал да омогуће истраживачима да спроведу анализе на већим просторним размерама (континент/широм света).

У последњем делу поглавља приказани су резултати студије која је показала да се необрађени ("сирови") подаци о коришћењу мобилних телефона, који се типично прикупљају са циљем наплате услуга оператера мобилне телефоније, могу користити за издвајање знања значајног за различите намене. Анализом записа мобилних телефона и издвајањем комплексних обележја мобилности, конективности и кретања корисника, кандидаткиња је успела да идентификује кључне елементе који утучу на стопу инфекције ХИВ вируса, док је визуализацијом честих путања и међурегионалних миграција настојала да објасни просторну структуру епидемије. Резултати спроведене анализе указују на то да су јаке везе и жаришта у комуникацији и мобилности корисника мобилних телефона повезана са жариштима ХИВ вируса, те да се оваквом анализом може предвидети будући тренд ширења болести.

Поглавље 6 – Закључци

У последњем поглављу су изнета закључна разматрања у вези са за применом техника и метода истраживања података у домену јавно доступног, кориснички-генерисаног мултимедијалног садржаја, као и у домену коришћења записа мобилних телефона за различите намене. Поред тога, отворена су многа питања која су проистекла из резултата истраживања приказаног у дисертацији. Такође су приказани и могући даљи правци истраживања, као и могућност практичне примене предложених решења како би се побољшало разумевање феномена из реалног живота који се рефлектују у прикупљеним подацима.

VI СПИСАК НАУЧНИХ И СТРУЧНИХ РАДОВА КОЈИ СУ ОБЈАВЉЕНИ ИЛИ ПРИХВАЋЕНИ ЗА ОБЈАВЉИВАЊЕ НА ОСНОВУ РЕЗУЛТАТА ИСТРАЖИВАЊА У ОКВИРУ РАДА НА ДОКТОРСКОЈ ДИСЕРТАЦИЈИ

- Sanja Brdar, **Katarina Gavrić**, Dubravko Čulibrk, Vladimir Crnojević, **Unveiling spatial epidemiology**

of HIV with mobile phone data, Scientific Reports 6, Article number: 19342, doi:10.1038/srep19342, ISSN 2045-2322. (2016)

- **Katarina Gavrić**, Dubravko Ćulibrk, Vladimir Crnojević, **Deriving basic law of human mobility using community – contributed multimedia data**, The International Conference on Pattern Recognition Applications and Methods (ICPRAM), pp. 543-246, ISBN 978-989-8565-41-9 (2013)

- **Katarina Gavrić**, Milan Mirković, Dubravko Ćulibrk, Vladimir Crnojević, **Using YouTube data to analyze human continent-level mobility**, The International Conference on Computational Aspects of Social Networks (CASON 2011), MirLabs, pp. 207-210, ISBN 978-1-4577-1131-2 (2011)

- **Katarina Gavrić**, Dubravko Ćulibrk, Predrag Lugonja, Milan Mirković, Vladimir Crnojević, **Detecting Attractive Location and Tourists' Dynamics Using Geo-referenced Images**, The 10th International Conference on Telecommunication in Modern Satellite Cable Broadcasting Services (Telsiks 2011), pp. 208-211, ISBN 978-1-4577-2017-8 (2011)

VII ЗАКЉУЧЦИ ОДНОСНО РЕЗУЛТАТИ ИСТРАЖИВАЊА

Истраживање приказано у дисертацији је довело до следећих конкретних закључака:

- У оквиру фазе откривања локација које корисници посећују (детекција атрактивних локација применом техника кластеризације и визуелизације података), кандидаткиња је показала да метаподаци кориснички-генерисаног мултимедијалног садржаја могу да се користе за идентификацију основних образаца у људском понашању, као и за анализу временске динамике кретања. У сврху овог истраживања, коришћени су метаподаци придружени видеима корисничког сервиса YouTube, а добијени резултати су упоређени са реалним подацима из тог подручја.
- Анализом метаподатака корисничког сервиса Flickr као алтернативе записима мобилних телефона, кандидаткиња је показала да се законитости које се могу издвојити посматрањем једног могу пронаћи и у другом скупу података, те да јавно доступни подаци имају потенцијал да омогуће истраживачима да спроводе анализе на већим географским размерама (имајући у виду малу доступност записа мобилних телефона).
- У оквиру фазе анализе записа мобилних телефона и издвајања значајних обележја на основу мобилности, конективности и кретања корисника, кандидаткиња је приказала кључне елементе који утичу на стопу ширења инфекције ХИВ вируса, док је визуелизацијом честих путања и међурегионалних миграција настојала да објасни просторну структуру епидемије.

VIII ОЦЕНА НАЧИНА ПРИКАЗА И ТУМАЧЕЊА РЕЗУЛТАТА ИСТРАЖИВАЊА

У дисертацији је јасно дефинисан циљ истраживања, а резултати су приказани на једноставан и јасан начин. Кандидаткиња је у оквиру истраживања користила адекватне и потврђене научне методе. У објашњењима резултата је користила статистичке показатеље уз помоћ којих је проблем прецизно сагледан и протумачен.

Основну поставку истраживања: да је потребно и могуће применити технике и методе истраживања података у домену анализе јавно доступног, кориснички-генерисаног мултимедијалног садржаја и записа мобилних телефона како би се дошло до нових сазнања односно како би се на ефикаснији начин решили актуелни проблеми у области, кандидаткиња је потврдила истраживањем.

На основу ових показатеља комисија даје **позитивну оцену начина приказа и тумачења резултата истраживања.**

IX КОНАЧНА ОЦЕНА ДОКТОРСKE ДИСЕРТАЦИЈЕ:

VII Да ли је дисертација написана у складу са образложењем наведеним у пријави теме
Дисертација је написана у складу са образложењем наведеним у пријави теме.

VIII Да ли дисертација садржи све битне елементе
Докторска дисертација својим садржајем, резултатима истраживања и начином њиховог тумачења садржи све битне елементе који се захтевају за радове овакве врсте.

1. По чему је дисертација оригиналан допринос науци
Разматрајући целокупну материју поднете докторске дисертације кандидаткиње Катарине Гаврић, **Комисија је закључила да она својим прилазом и тематским одређењем, представља новину у истраживањима у овој области.** Главни допринос дисертације је генерисање нових знања о

кретању, навикама и понашању људи коришћењем јавно доступних података о њиховој локацији и кретању, развијање нових модела којима су боље описане појаве: детекција атрактивних локација, путања кретања људи у посматраном подручју, њихове навике путовања, као и преношење вируса ХИВ на одређеном географском подручју. Истраживањем и извршавањем експеримената добијен је одговор на питање које истраживачке методе предикције дају најбоље резултате на гео-референцираним записима и записима мобилних телефона које људи генеришу, као и у којој мери су погодни за конкретне задатке. Применом различитих скупова података за различита предвиђања уочени су шаблони у понашању људи што свакако даје значајну могућност унапређења квалитета живота људи и њихове околине. Праћењем утицаја различитих фактора на кретање људи и различите обрасце понашања, као што су навике у комуникацији људи (engl. Connectivity Patterns), уочене су потребе за наглашавањем појединих и занемаривањем других фактора. Добијени резултати имају и ширу примену јер се предложени поступци анализе заснивају на математичким и статистичким моделима који се лако могу применити на новим проблемима из области у којима су вршена истраживања, али и на проблемима из других домена као што су управљање вандредним ситуацијама, урбанистичко планирање, управљање ширењем стопе сиромаштва итд.

VIV Недостаци дисертације и њихов утицај на резултат истраживања

Докторска дисертација нема недостатака који би утицали на коначан резултат истраживања.

X ПРЕДЛОГ:

На основу укупне оцене дисертације, комисија предлаже:

- да се докторска дисертација **прихвати**, а кандидату **одобри одбрана**

НАВЕСТИ ИМЕ И ЗВАЊЕ ЧЛАНОВА КОМИСИЈЕ
ПОТПИСИ ЧЛАНОВА КОМИСИЈЕ

2. **Др Александар Купусинац**, ванредни професор, Факултет техничких наука, Нови Сад (у.н.о. Примењене рачунарске науке)

1. **Др Дарко Стефановић**, доцент, Факултет техничких наука, Нови Сад (у.н.о. Информационо-комуникациони системи)

3. **Др Дејан Вукобратовић**, ванредни професор, Факултет техничких наука, Нови Сад (у.н.о. Телекомуникације и обрада сигнала)

4. **Др Владимир Божовић**, ванредни професор, Природно-математички факултет, Подгорица (у.н.о. Математика)

5. Др Дубравко Ђулибрк, ванредни професор, Факултет
техничких наука, Нови Сад (у.н.о. Информационо-
комуникациони системи)

НАПОМЕНА: Члан комисије који не жели да потпише извештај јер се не слаже са мишљењем већине чланова комисије, дужан је да унесе у извештај образложење односно разлоге због којих не жели да потпише извештај.