

НАСТАВНО-НАУЧНОМ ВЕЋУ

Предмет: Реферат о урађеној докторској дисертацији кандидата Саше Стојановића, дипломираног инжењера електротехнике

Одлуком Наставно-научног већа Електротехничког факултета у Београду бр. 5011/07-3 од 23.09.2015. године, именовани смо за чланове Комисије за преглед и оцену докторске дисертације кандидата Саше Стојановића под насловом

Процена сличности процедура у бинарном коду

После прегледа достављене Дисертације и других пратећих материјала и разговора са Кандидатом, Комисија је сачинила следећи

РЕФЕРАТ

1. УВОД

1.1. Хронологија одобравања и израде дисертације

Кандидат Саша Стојановић, дипломирани инжењер електротехнике, уписао је докторске студије на Електротехничком факултету у Београду у школској 2007/2008. години. На основу одлуке Наставно–научног већа бр. 2944/2 од 11.10.2007. године, Студијски програм је започео у пролећном семестру школске 2007/2008, па се рок за завршетак докторских академских студија рачуна од почетка тог семестра, сагласно Статуту Универзитета у Београду и Статуту Електротехничког факултета. По истеку законског рока за завршетак докторских академских студија, на захтев студента, одобрено је продужење рока за завршетак ових студија за још два семестра, сагласно члану 92. став 4 Статута Универзитета у Београду, као и додатно продужење за годину дана, на основу Одлуке бр. 24-06/10-2007/5011 од 4.3.2015. године.

Кандидат је пријавио тему за израду докторске дисертације на Електротехничком факултету у Београду 29.01.2015. године. Комисија за студије трећег степена Електротехничког факултета у Београду разматрала је предлог теме за израду докторске дисертације и предлог Комисије о оцени подобности теме и кандидата упутила Наставно – научно већу на усвајање 03.02.2015. године. Наставно-научно веће именovalo је Комисију за оцену услова и прихватање теме докторске дисертације на седници одржаној 04.03.2015. године, у саставу др Вељко Милутиновић, редовни професор (ментор, Универзитет у Београду – Електротехнички факултет), др Јелица Протић, ванредни професор (Универзитет у Београду – Електротехнички факултет), др Горан Квашчев, доцент (Универзитет у Београду – Електротехнички факултет), др Ненад Митић, ванредни професор (Универзитет у

Београду – Математички факултет). Извештај Комисије за оцену услова и прихватање теме докторске дисертације Наставно-научно веће је прихватило на седници одржаној 17.03.2015. године, а Веће научних области техничких наука Универзитета у Београду дало је сагласност на тему 06.04.2015. године (број одлуке 61206-1404/2-15).

Кандидат је предао докторску дисертацију 01.09.2015. године. На седници одржаној 08.09.2015. године Комисија за студије трећег степена дала је сагласност за образовање Комисије за преглед и оцену докторске дисертације. На седници Наставно-научног већа одржаној 15.09.2015. године именована је Комисија за преглед и оцену докторске дисертације Саше Стојановића, дипломираног инжењера електротехнике, под насловом „Процена сличности процедура у бинарном коду“. Чланови комисије су: др Вељко Милутиновић, редовни професор (ментор, Универзитет у Београду – Електротехнички факултет), др Захарије Радивојевић, доцент (Универзитет у Београду – Електротехнички факултет), др Ненад Митић, ванредни професор (Универзитет у Београду – Математички факултет), др Драган Бојић, доцент (Универзитет у Београду – Електротехнички факултет), др Милош Цветановић, доцент (Универзитет у Београду – Електротехнички факултет).

1.2. Научна област дисертације

Научна област дисертације је Електротехника и рачунарство, а ужа научна област је Рачунарска техника и информатика. За ову ужу научну област матичан је Електротехнички факултет у Београду.

Дисертација је рађена под менторством редовног професора др Вељка Милутиновића. Ментор испуњава законске услове за ментора, бави се научним радом у ужој области Рачунарска техника и информатика.

1.3. Биографски подаци о кандидату

Саша Стојановић, дипломирани инжењер електротехнике и рачунарства, рођен је 27. септембра 1982. године у Фочи, Социјалистичка Федеративна Република Југославија, од мајке Руже и оца Доброслава. Основну школу "Свети Сава" и први разред гимназије завршио је у Фочи. Даље школовање наставио је у Математичкој гимназији у Београду. За време основног и средњег школовања учествовао и освајао награде на многобројним такмичењима из математике, физике и информатике, међу којима су и два међународна такмичења светског ранга.

Електротехнички факултет Универзитета у Београду уписао је 2001. године као 35. на ранг листи. Основне петогодишње студије на Електротехничком факултету завршио је за непуних 5 година, при чему је сваки испит положио у првом испитном року. Дипломирао је 15. маја 2006. године на Одсеку за рачунарску технику и информатику, са просечном оценом 9,64. Дипломски рад на тему "Оперативни систем реалног времена за MSP430x44x" одбранио је са оценом 10 код проф. др Вељка Милутиновића. Докторске студије уписао је 2008. године на Смеру рачунарска техника и информатика, где је до сада положио све испите са оценом 10 и обавио два студијска истраживачка рада и један научно-стручни рад.

Од септембра 2006. године ангажован је при Катедри за рачунарску технику и информатику, где је 22. јануара 2007. изабран у звање Сарадника у настави, а 06. априла 2009. изабран у звање Асистента. У претходном периоду био је ангажован за потребе држања вежби на табли и у лабораторији на предметима: Програмирање 1, Програмирање 2, Оперативни системи 1, Оперативни системи 2, Системско програмирање, Системски софтвер, Микропроцесорски системи, Рачунарски ВЛСИ системи и Програмирање мобилних уређаја. Повремено је био ангажован и за потребе држања вежби на табли из предмета: Експертски системи, Инфраструктура за електронско пословање, Практикума из оперативних система и Претраживање и истраживање података на интернету.

Коаутор је на три рада у међународним часописима са фактором утицаја (енг. *impact factor*) са SCI листе, од чега су два у директној вези са дисертацијом, два рада у домаћим часописима, као и неколико радова на домаћим и међународним конференцијама. Учествовао је у припреми и извођењу неколико туторијала, од чега су два била на познатим међународним конференцијама HiPeas 2013 и ISCA 2013.

2. ОПИС ДИСЕРТАЦИЈЕ

2.1. Садржај дисертације

Докторска дисертација садржи насловну страну и кратак резиме рада на српском и енглеском језику, садржај, седам поглавља, преглед коришћене литературе, кратку биографију аутора, изјаву о ауторству, изјаву о истоветности штампане и електронске верзије докторске дисертације и изјаву о коришћењу. Поглавља су насловљена: 1. Увод, 2. Преглед литературе, 3. Дефиниција проблема, 4. Окружење за евалуацију, 5. Евалуација постојећих алата и приступа, 6. Предложени приступ и евалуација и 7. Закључак. Дисертација садржи 156 стране, 38 слика и 8 табела.

2.2. Кратак приказ појединачних поглавља

Прво поглавље представља увод у дисертацију. У овом поглављу је уведен проблем који је решаван у дисертацији. Описан је проблем комерцијалне употребе софтверске библиотеке са доступним изворним кодом унутар бинарног кода неког производа. Како би се открила употреба неопходно је експертско знање како би се свака од процедура из софтверске библиотеке упоредила са сваком од процедура у бинарном коду, што је временски изузетно захтевно. Описан је и значај проблема, као и фактори који уносе комплексност у посматрани проблем. На крају је описан предлог приступа чији је основни циљ да убрза потрагу за коришћеном софтверском библиотеком тако што ће смањити број поређења процедура која експерт треба да уради.

У другом поглављу дат је преглед релевантне литературе. На почетку је дат општи осврт на поређење кода и прве области у којима је поређење кода коришћено. Потом су побројане области у којима се поређење кода користи и описане околности у којима се поређење кода ради у свакој од наведених области, а затим је дат преглед приступа за откривање софтверских клонова и плагијаризма у којима је поређење кода централни део скоро свих приступа. Приступу су класификовани на основу репрезентације кода која се користи приликом поређења. На крају сваке класе приступа извучени су закључци о предностима и манама у погледу примене одговарајућих алата на проблем посматран у раду. На крају су наведене и области у којима се поређење кода појављује само у неким приступима. У оквиру описа ових области описани су сценарији у којима се појављује потреба за поређењем кода као и приступи чија примена на проблем посматран у дисертацији би потенцијално била интересантна и који су имплементирани за потребе поређења са предложеним приступом.

У трећем поглављу дефинисан је проблем који је решаван у оквиру дисертације. Наведене су све претпоставке и ограничења под којима ће проблем бити решаван у дисертацији и постављени су јасни циљеви којима предложено решење треба да тежи. На крају поглавља проблем је исказан и кроз шест истраживачких питања на која дисертација треба да одговори.

У четвртном поглављу описано је окружење за евалуацију предложеног и постојећих приступа. У првом делу описани су подаци који се користе при тестирању. Описан је поступак припреме података за употребу, као и начин формирања узорака за тестирање. У другом делу описане су битне карактеристике приступа за поређење кода из угла експерта

kojem су намењени, а затим су у складу са битним карактеристикама уведене мере успеха које ће се користити у процесу евалуације. Мере које су предложене за употребу у процесу евалуације се уобичајено користе и у другим областима за оцену квалитета претраживања и то су одзив, прецизност, F и F_2 мера.

У петом поглављу приказана је евалуација постојећих приступа. У евалуацији су коришћена два скупа алата. Први скуп представљају постојећи алати који на улазу очекују изворни код, док другу групу представљају алати, који су на основу описа одговарајућих приступа у доступној литератури, имплементирани за потребе поређења са предложеним приступом и који на улазу прихватају код у бинарном облику. Применом на проблему посматраном у дисертацији, један део евалуираних приступа је постигао високу прецизност, док је одзив имао малу вредност. Остали евалуирани приступи су постигли већи одзив, с тим да је постигнута прецизност знатно смањена. Описани резултати су указали на потребу за увођењем новог приступа.

У шестом поглављу описан је предложени приступ заједно са евалуацијом. На почетку је уведен приступ заснован на софтверским метрикама који се састоји од две фазе. У првој фази се из процедура издвајају вредности метрика, након чега се пореде метрике тражене процедуре и сваке од процедура која потенцијално потиче од изворног кода тражене процедуре, чиме се добијају парцијалне мере сличности. Парцијалне мере сличности се даље трансформишу на основу претходног знања, након чега се за сваке две процедуре неком од предложених формула комбинују вредности већег броја парцијалних мера сличности како би се израчунала једна, тотална мера сличности. Приступ предлаже 19 софтверских метрика, од којих су четири инспирисане постојећим софтверским метрикама, док су остале предложене наменски за примену на бинарни код. Предложено је и пет начина за трансформацију парцијалних мера сличности, као и 7 различитих формула за рачунање тоталне мере сличности. Израчуната мера тоталне сличности се користи у другој фази приступа за сортирање сумњивих процедура у складу са нерастућом сличношћу према траженој процедури. Описани приступ је потом евалуиран и упоређен са постојећим алатима коришћеним и у поглављу 5. У евалуацији је поред 7 предложених формула за рачунање тоталне мере сличности евалуирана и могућност примене три технике машинског учења. Евалуација је показала да предложени приступ постиже већи одзив од свих алата са којима је упоређен, а да при томе задржи и високу прецизност. Потом су анализирани недостаци предложеног приступа у погледу количине информација које се задрже за потребе поређења процедура, што је резултовало увођењем додатних техника за побољшање резултата предложеног приступа. Предложено је пет техника, од којих неке задржавају додатне информације о поређеним процедурама, док неке додатно елиминишу информације које негативно утичу на поређење процедура. Предложене технике су потом евалуиране и упоређене са постојећим приступима, укључујући и алате који су имплементирани на основу описа у одговарајућим радовима. Евалуација је показала да увођење техника може да побољша резултате предложеног приступа, као и да се употребом техника постиже скоро двоструко већи одзив у поређењу са најбољим од постојећих приступа. На крају шестог поглавља издвојени су одговори на истраживачка питања постављена у трећем поглављу.

У седмом поглављу дат је закључак. У закључку су укратко приказани циљеви и преглед предложеног приступа заједно са постигнућима. На крају су дати неки могући даљи правци истраживања. У осмом поглављу дат је преглед коришћене литературе.

3. ОЦЕНА ДИСЕРТАЦИЈЕ

3.1. Савременост и оригиналност

Предмет истраживања у дисертацији је поређење процедура са циљем да се процени сличност између њих. Поређење процедура представља један подскуп приступа за поређење кода. Прва истраживања у области поређења кода урађена су пре неколико деценија и од

тада је ова област у сталном фокусу истраживача, о чему сведоче многобројни радови који се и даље објављују у признатим међународним часописима и на познатим међународним конференцијама. Због сталног раста софтверске индустрије и притиска који раст врши на софтверске компаније, проблем поређења кода све више добија на значају, посебно у области провере нарушености ауторских права, чиме се бави и ова дисертација. Стога се може закључити да се дисертација бави савременим проблемом који све више добија на значају.

У дисертацији је предложен нов и оригиналан приступ за процену сличности процедура у бинарном коду. Приступ је заснован на употреби софтверских метрика које су део предложеног приступа и које су осмишљене за примену на бинарном коду. Поред метрика, приступ предлаже механизме за поређење вредности предложених софтверских метрика, трансформацију резултата поређења вредности метрика и формуле за рачунање сличности између процедура. Приступ предлаже и 5 нових и оригиналних техника које имају за циљ повећање успешности основног приступа на који се примене. Предложени приступ и технике су у дисертацији упоређени са постојећим приступима чија примена на посматрани проблем је била могућа. Резултати показују да је на посматраном проблему предложени приступ постигао знатно боље резултате у односу на остале поређене алате, што додатно потврђује оригиналност приступа предложеног у дисертацији.

3.2. Осврт на референтну и коришћену литературу

Кандидат је детаљно претражио и упознао се са одговарајућом литературом. У дисертацији је прецизно наведено 70 библиографских референци на литературу која је у вези са темом дисертације. Литература садржи и најновије радове релевантне за тему дисертације, као и одговарајуће радове чији је кандидат коаутор. Једно поглавље дисертације посвећено је анализи приступа потенцијално применљивих на проблем посматран у дисертацији, при чему је обухваћен већи број области у којима се ради поређење кода, како оних у којима је поређење кода централни део већине решења, тако и оних у којима се поређење кода појављује у ретким приступима. Искazi у том поглављу, као и у осталим деловима дисертације, добро су поткрепљени цитатима одговарајућих радова.

3.3. Опис и адекватност примењених научних метода

Истраживање у оквиру докторске дисертације обухватило је следеће фазе:

1. систематично проучавање литературе из области у којима се појављују алати за поређење кода са посебним освртом на применљивост на проблем посматран у дисертацији,
2. развој новог приступа за поређење процедура у бинарном коду који се базира на процени сличности међу процедурама, при чему су примењена и сазнања до којих се дошло у току проучавања литературе,
3. имплементација експерименталног алата који имплементира предложени приступ, као и постојеће приступе за које не постоје алати или постојећи алати нису погодни за директну примену на посматрани проблем,
4. дефинисање окружења за евалуацију предложеног и постојећих приступа,
5. евалуација могућности примене постојећих приступа на посматрани проблем,
6. евалуација предложеног приступа и поређење са постојећим приступима,

7. критичка анализа предложеног приступа у циљу проналажења могућих побољшања,

8. евалуације предложених побољшања и поређење са постојећим приступима.

Наведени поступци у основи припадају и теоријским и експерименталним истраживањима, и у потпуности одговарају проблему и постављеном циљу дисертације. Примењене експерименталне методе су адекватне и валидне.

3.4. Применљивост остварених резултата

У оквиру дисертације представљен је оригиналан приступ за убрзавање потраге за библиотеком која је доступна у облику изворног кода а која је коришћена унутар бинарног кода. Имплементација предложеног приступа је верификована у случају откривања употребе библиотеке MatrixSSL која је издата под двоструком лиценцом и која се користи у оквиру BusyBox софтверског пакета. Верификација је показала да је употребом алата могуће пронаћи већи део софтверске библиотеке уз вишеструко скраћивање уложеног експертског времена, чиме је откривање употребе софтверске библиотеке постало практично изводљиво. Уз веома једноставно претпроцесирање бинарног кода у виду дисасемблирања и развој новог лексера, експериментални алат је могуће врло лако применити и на многе друге случајеве у којима је потребно открити неовлашћену употребу софтверске библиотеке у бинарном коду.

3.5. Оцена достигнутих способности кандидата за самостални научни рад

Кандидат је у изради дисертације показао способност за самостални научни рад. Израдио је систематичну и критичку анализу постојећих решења, уз анализу могућности њихове примене на посматрани проблем. Развио је оригиналан приступ за процену сличности процедура у бинарном коду који је искоришћен за убрзавање откривања употребе софтверске библиотеке. Показао је да је предложени приступ изузетно користан и применљив у пракси. Резултате својих истраживања објавио је у два часописа од међународног значаја са признатим фактором утицаја.

4. ОСТВАРЕНИ НАУЧНИ ДОПРИНОС

4.1. Приказ остварених научних доприноса

Допринос изложене докторске дисертације је у домену поређења процедура у бинарном коду. Као саставни делови дисертације садржани су следећи научни доприноси:

- Идентификација области у којима постоје приступи применљиви на проблем откривања неовлашћене употребе софтверске библиотеке.
- Систематизација и класификација постојећих приступа у контексту применљивости на проблем откривања неовлашћене употребе софтверске библиотеке.
- Формирање новог приступа за процену сличности процедура у бинарном коду и убрзавање потраге за употребом софтверске библиотеке.
- Критичка анализа недостатака предложеног приступа и увођење техника за побољшање резултата.
- Развој експерименталног алата који имплементира предложени приступ.
- Развој експерименталних алата за постојеће приступе за које не постоји адекватан алат.

- Евалуације постојећих приступа.
- Евалуација предложеног приступа и предложених техника за побољшање резултата.

4.2. Критичка анализа резултата истраживања

Увидом у дисертацију и циљеве истраживања, Комисија констатује да је кандидат успешно одговорио на постављене изазове, и да резултати оправдавају почетна очекивања. Предложен је нов и оригиналан приступ за убрзање откривања употребе софтверске библиотеке проценом сличности процедура заснованом на софтверским метрикама. Предложене софтверске метрике, заједно са предложеним формулама за поређење и рачунање сличности процедура, чине приступ јединственим и отпорнијим на измене које у поређене кодове унесу преводиоци. У контексту посматраног проблема дисертација представља значајан искорак у односу на постојеће приступе обзиром да представља један од веома ретких алата намењених посматраном проблему и да у односу на остале алате постиже значајно боље резултате. Дисертација отвара и многе нове правце за даље истраживање и усавршавање приступа у циљу побољшања резултата које приступ постиже у посматраним околностима. Резултати истраживања су од интереса свима који на тржишту нуде софтверску библиотеку са јавно доступним изворним кодом, за чију комерцијалну употребу је потребно прихватити услове дефинисане у посебној лиценци. Приступ је посебно интересантан компанијама које услужно раде проверу нарушености ауторских права.

4.3. Верификација научних доприноса

Кандидат је објавио следеће радове који су у непосредној вези са докторском дисертацијом:

Категорија M22:

1. **Stojanovic, S.**, Radivojevic, Z., Cvetanovic, M.: *Approach for estimating similarity between procedures in differently compiled binaries*, Information and Software Technology, Vol. 58, No. 1, pp.259-271, February 2015. (IF₂₀₁₄:1.046) (ISSN:0950-5849) (doi:10.1016/j.infsof.2014.06.012).

Категорија M23:

1. Radivojevic, Z., Cvetanovic, M., **Stojanovic, S.**: *Comparison of Binary Procedures: A Set of Techniques for Evading Compiler Transformations*, The Computer Journal, Vol. n/a, No. n/a, pp. n/a, September 2015 (online). (IF₂₀₁₄:0.787) (Online ISSN 1460-2067) (doi:10.1093/comjnl/bxv076).

Категорија M51:

1. Berta, K., **Stojanovic, S.**, Radivojevic, Z., Cvetanovic, M.: *Estimation of similarity between functions extracted from x86 executable files*, Serbian Journal of Electrical Engineering, Vol. 12, No. 2, pp. 253-262, 2015 (doi:10.2298/SJEE1502253B).

5. ЗАКЉУЧАК И ПРЕДЛОГ

Дисертација кандидата Саше Стојановића, дипломираног инжењера електротехнике, под насловом „Процена сличности процедура у бинарном коду“ представља оригиналан, савремен и значајан научни допринос. Текст дисертације написан је јасно и разумљиво и добро је организован кроз поглавља и одељке. Циљеви дисертације јасно су формулисани и мотивисани, а резултати истраживања систематски изложени, тако да се научни доприноси могу недвосмислено утврдити. У спроведеним истраживањима предложена је нова метода за процену сличности процедура у бинарном коду. Практична примена предложене методе потврђена је њеном имплементацијом и употребом у оквиру експерименталног алата за убрзавање откривања употребе софтверске библиотеке. Објављивањем резултата својих истраживања у часописима од међународног значаја, кандидат је показао способност за самосталан научни рад, а доприноси истраживања добили су адекватну потврду ваљаности.

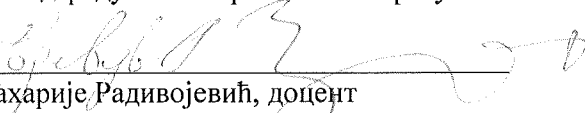
Комисија констатује да дисертација садржи оригиналне научне доприносе, испуњава све законске, формалне и суштинске услове, као и све критеријуме који се уобичајено примењују приликом вредновања докторских дисертација на Електротехничком факултету у Београду. Комисија са задовољством предлаже Наставно-научном већу Електротехничког факултета у Београду да се докторска дисертација под називом „Процена сличности процедура у бинарном коду“ кандидата Саше Стојановића, дипломираног инжењера електротехнике, прихвати, а кандидату одобри усмена одбрана.

У Београду, 07.10.2015. године

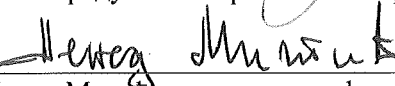
ЧЛАНОВИ КОМИСИЈЕ



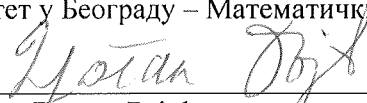
др Вељко Милутиновић, редовни професор
Универзитет у Београду – Електротехнички факултет



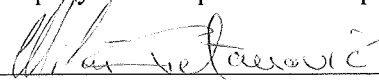
др Захарије Радивојевић, доцент
Универзитет у Београду – Електротехнички факултет



др Ненад Митић, ванредни професор
Универзитет у Београду – Математички факултет



др Драган Бојић, ванредни професор
Универзитет у Београду – Електротехнички факултет



др Милош Цветановић, доцент
Универзитет у Београду – Електротехнички факултет