



РЕЦЕНЗИЯ

От: проф. д-р Красен Стефанов Стефанов
СУ „Св. Климент Охридски“ - ФМИ
Информатика и компютърни науки

Относно: дисертационен труд за присъждане на образователна и научна степен **„доктор“** по научна специалност в УНСС.

Основание за представяне на рецензията: участие в състава на научното жури по защита на дисертационния труд съгласно Заповед №1088/14.04.2021 на Ректора на УНСС.

Автор на дисертационния труд: *Николай Ангелов Къскатийски*
Тема на дисертационния труд: *Методи за пренос и обработка на данни от Интернет на обектите в платформа за големи данни*

1. Информация за дисертанта

Николай Ангелов Къскатийски е роден през 1977 г. в София. Завършва Образцовия техникум по финна механика и оптика М. В. Ломоносов в София. Придобива магистърска степен по Автоматика и механика в Техническия университет – София и магистърска степен по специалност „Бизнес администрация“ в УНСС. Получава допълнителна квалификация в областта на управлението на проекти Project Management Professional (PMP) сертификат от Project Management Institute.

Николай Къскатийски има богат опит от работа по софтуерни проекти в областта на застрахователния и финансов сектор. Някои от заеманите от него длъжности са:

- Разработчик на софтуер в OBS – Bulgaria (2002-2009)
- Младши ръководител проект в CSC – Bulgaria (2009-2016)
- Ръководител проект в DXC – Bulgaria (2016-2020)
- ИТ директор в DXC – Bulgaria (2020 – до сега)

Николай Ангелов Къскатийски е зачислен в задочна докторантура със заповед на ректора на УНСС №1149/09.05.2017. Обучението е субсидирано от държавата по докторска програма „Приложение на изчислителната техника в икономиката“, професионално направление 3.8.

Икономика към катедра „Информационни технологии и комуникации“ със срок на обучение 4 години - от 28.03.2017г. до 28.03.2021г. и научен ръководител: доц. д-р Любен Боянов.

2. Обща характеристика на представения дисертационен труд

Дисертационният труд е разработен в четири глави, към които има Въведение, Заключение, Литература, Приложение, Индекс на фигурите и Индекс на Таблиците. Работата е в рамките на 165 страници. Изложението се простира на 139 страници, списъкът на информационните източници в рамките на 7 страници. Авторът използва 102 литературни източника, от които 3 на български и 99 на английски език. Приложението е с обем от 14 страници. Дисертационният труд включва 22 таблици и 81 фигури.

Дисертационният труд изследва протоколите за пренос на данни при Интернет на Обектите (ИНО) и връзката им със системи за големи данни. За потвърждение на актуалността на темата може да се съди от факта, че днес повече от всякога хората и бизнеса се стремят да извлекат предимство събирайки информация от разнородни източници и ИНО се явяват, като удобен инструмент, който може да предостави това. Съществуват множество протоколи по които може да се осъществи пренос на данни между ИНО, обоснования подбор на протокол, който да се използва при изграждането на конкретна система може да подобри значително ефективността и.

Въведението на дисертационния труд обхваща съдържанието на разработката и представя целта, обектът, предметът, задачите за реализиране на целта и работните хипотези.

Основната цел на дисертационния труд е да изследва приложимостта и особеностите на комуникационните протоколи в ИНО от гледна точка системи за големи данни.

Обект на представения труд са методите за комуникация от ИНО към системите за обработка и съхранение на данни в частност големи данни. Предмет на дисертационния труд са подходите за предаване и приемане на данни от ИНО в системи за обработка на големи данни, тяхната обработка и интегриране, както и методологията за избор на комуникационни протоколи за данни от ИНО.

Задачите на работата включват изследване на съвременни протоколи за пренос на информация, както и дефиниране на критерии за оценка на приложимост на отделните протоколи.

Основните хипотези над които работи Николай Къскатийски са добре формулирани и се отнасят за възможността да се посочат и класифицират комуникационни протоколи, които са приложими за работа

с Ино, както и че е възможно да се дефинира методология по която да се подобри ефективността на преноса на данни от Ино към системи за работа с големи данни.

В първа глава авторът прави анализ на основните характеристики на големите данни. Дадени са процесите при работа с този вид данни. Представени са и сравнителни характеристики на редица от известните платформи за работа с големи данни – основно за положителните им страни и какво още се желае от тях. Отбелязват се два от най-широко разпространени подхода за приемане и обработка на данни от различни източници, включително и от Ино – данни в реално време и пакетни данни. В тази връзка се обръща внимание и на Ламбда и Капа архитектурите. Авторът дефинира основни понятия в Ино с цел по-задълбочено запознаване с предметната област, представя най-общата архитектура на Ино и съществуващите референтни архитектури.

Във втора глава Николай Къскатийски прави обзор на четири модела на комуникация и посочва предимствата им при работа с Ино. Докторантът синтезира най-важните особености на редица комуникационни протоколи и посочва ключови е характеристики, които е възможно да окажат влияние на работата с Ино. Отбелязва тенденциите в развитието на платформите за работа с Ино. В работата е засегнат въпросът с мащабируемостта на изчислителните системи, което може да е важно качество при избор на дадена реализация в света на Ино. Добавен е обзор на основни похвати използвани при количествено оценяване на описателни характеристики.

В трета глава дисертантът дефинира критерии, отнасящи се за системи с участието на Ино, като критериите са подбрани на базата на допитване, както и на сравнения в научни публикации и разговори с експерти. Тези критерии позволяват да се оцени приложимостта на даден протокол като се съставя таблица с числово измерение на приложимостта на даден протокол към съответен критерий. Авторът предлага методология, чрез която при избор на значими критерии за дадена реализация може да се даде оценка за приложимост на даден протокол.

В четвърта глава авторът описва проведените експерименти, които верифицират методологията, предложена в глава трета. Даден е пример с платформа от Ино, която е лесна за прилагане, поради което се разглежда като подходяща за потвърждение на изложението. Платформата, използвана за опитите, както и елементите са подбрани от такива, които са широко използвани на пазара. Експериментите са разделени на две основни групи. Първата е как един обект предава данни по различни протоколи, като се измерва количеството пренесени данни и енергията,

използвани за единица време. Авторът тества пет различни протокола (MQTT, CoAP, AMQP, DDS и HTTP) със структурирани, полуструктурирани и неструктурирани данни. Втората група експерименти проследява влиянието на броя съобщения получени от Ино източници за времето, необходимо за тяхното съхранение в система за големи данни. С цел верификация са използвани две широко разпространи платформи за пренос на съобщения - Apache NiFi и Apache Kafka. Резултатите и при двете са много сходни. Посочено е че големия брой съобщения води до затруднение при запис им в разпределената файлова система (HDFS), която принципно се използва в системи за работа с големи данни, работещи с Хадуп. Авторът предлага решение, чрез използване на буфер за натрупване на съобщенията преди техния запис във файл. Изследва се влиянието на буфера над системата за приемане и съхранение на съобщения.

В Заключение, авторът обобщава постиженията си, като посочва своите научни и научно-приложни приноси.

Много добро и силно впечатление прави голямото количество използвана литература, която е основно от престижните източници в областта на Информатиката като IEEE и ACM, Springer и др. Всички те сочат за много висока осведоменост на дисертанта по проблематиката, разглеждана в дисертационния труд.

3. Оценка на получените научни и научно-приложни резултати

Докторантът прави анализ на различни архитектури за обработка на големи обеми от данни на базата на реални практически експерименти, включващи инсталиране и конфигуриране на множество системи със свободен достъп и решаване на практически задачи върху тях. На базата на тези експерименти той проучва голямо количество от комуникационни протоколи, което води до създаване на аналитичен обзор.

Николай Къскатийски създава методология за избор на комуникационен протокол за пренос на данни от Интернет на обектите (Ино) в системи на големи данни. Същата се прилага в проекта към Центъра за компетентност "Дигитализация на икономиката в среда на Големи данни" и дава много добри резултати. В рамките на същия проект той предлага нов подход, водещ до подобряване на цялостната архитектура за записване на данни от Ино в клъстер на Hadoop. За целта той предлага използване на допълнителен буфер в системата за управление на потоци от данни Apache NiFi.

Тук трябва да изтъкнем иновативния характер на използваната

методология за анализ на архитектурите и комуникационните протоколи използвани за Ино, тъй като подобни практически изследвания са нещо съвсем ново в България.

Получените научни и научно-приложни резултати позволяват категорично да се направи извода, че докторантът е изпълнил всички поставени цели и задачи в дисертационния труд.

4. Оценка на научните и научно-приложни приноси

Научните и научно-приложните приноси на докторанта са тясно свързани с постигнатите от него научни и приложни резултати. Приносите, посочени от дисертанта, могат да бъдат обобщени по следния начин.

Научни приноси:

- Направен е обзор и са класифицирани основните характеристики на комуникационните протоколи за Ино - MQTT, CoAP, AMQP, DDS и HTTP;
- Въз основа на анализ на тези протоколи е изведена характеристика с количествени показатели;
- Създадена е методология за избор на комуникационен протокол.

Научно-приложни приноси:

- Създадена е тестова среда за оценка на реални сценарии за пренос на данни от комуникационни протоколи и е извършена верификация;
- Направено е изследване с пет от най-известните и използвани протоколи за пренос на данни, с използване на структурирани, полуструктурирани и неструктурирани данни, включително за скорост на трафика на пренасяните данни и консумирана енергия от крайните Ино устройства;
- Предложени са блокове за оптимизация на потока приемани данни в модула за приемане и разпределяне на данните - NiFi;
- Намерена е зависимост между броят на получени съобщения и времето за запис в HDFS;
- Направени са изследвания в ефективността на работа на приеманите съобщения при работа с различни пакети за приемане на съобщенията (NiFi и Kafka).
- На базата на направените опити е посочено тясното място при приемането на данни от Ино и е предложена оптимизация с буфер, който значително намалява времето за запис;

5. Оценка на публикациите по дисертацията

Докторантът е направил следните публикации по тематиката на дисертационния труд:

- Nikolai Kaskatiiski, Models and approaches of Internet of Things in education, 7th International Conference on Application of Information and Communication Technology and Statistics in Economy and Education (ICAICTSEE – 2017), November 3-4th, 2017, UNWE, Sofia, Bulgaria, ISSN 2367-7635 (PRINT), ISSN 2367-7643 (ONLINE) достъпна на: <http://icaictsee.unwe.bg/past-conferences/ICAICTSEE-2017.pdf> pp. 554-561.
- Николай Къскатийски, Сравнителна характеристика на данните в застрахователния сектор, Новите информационни технологии и големите данни: Възможности и перспективи при анализите и управленските решения в бизнеса и икономиката и социалната сфера Сборник с доклади от международна научна конференция организирана от катедра „Статистика и иконометрия - със съдействието на факултет „Приложна информатика и статистика - и в партньорство с Националния статистически институт, 6 – 7 юни 2019, издателски комплекс УНСС, ISBN 978-619-232-292-2 стр. 231-242.
- Николай Къскатийски, Size and speed of data generated by Fleet Management Software, VIII Международна научна конференция техника. Технологии. Образование. Сигурност 2020 SCIENTIFIC TECHNICAL UNION OF MECHANICAL ENGINEERING INDUSTRY-4.0, 108, Rakovski Str., 1000 Sofia, Bulgaria, ISSN PRINT 2367-8399, ISSN WEB 2534-8493 достъпна на: <https://stumejournals.com/journals/tm/2020/1/pdf> pp. 13-16
- N Kaskatiiski and L Boyanov, Acquiring and storing Internet of Things data into Hadoop, IOP Conference Series: Materials Science and Engineering, Volume 1031, International Conference on Technics, Technologies and Education 2020 (ICTTE 2020) 4th-6th November 2020, Yambol, Bulgaria, IOP Conf. Ser.: Mater. Sci. Eng. 1031 012060, достъпна на: <https://iopscience.iop.org/article/10.1088/1757-899X/1031/1/012060> .

Тези публикации отговарят на изискванията на УНСС и представят реалистично работата на докторанта.

Поради малкото време след публикуването им не е известно дали има цитирания на тези публикации.

6. Оценка на автореферата

Авторефератът е изготвен в съответствие с изискванията на Правилника за условията и реда за придобиване на научни степени и за заемане на научни длъжности на УНСС. Авторефератът отразява пълно, изчерпателно и точно съдържанието на дисертационния труд, както и на основните научни и приложни приноси.

7. Критични бележки, препоръки и въпроси

Нямам критични бележки към представения труд. Мога да направя следните препоръки за бъдещата научна дейност на докторанта:

- Да се насочи към публикации във водещи научни списания с импакт фактор и импакт ранг
- Да представи резултатите си на водещи световни конференции
- Да се ориентира към съвместни научни проекти с водещи световни научни партньори.

8. Заключение

По обем и съдържание, резултати и приноси, представеният дисертационен труд отговаря напълно на изискванията за присъждане на образователната и научна степен „доктор“. Поради това двама положителна оценка на представения дисертационен труд. Като следствие на всички изложени до тук факти, считам че рецензираният дисертационен труд напълно удовлетворява изискванията на ЗРАСРБ, ПЗРАСРБ и Правилниците за условията и реда за придобиване на научни степени и за заемане на научни длъжности на УНСС. Убедено препоръчвам на уважаемото научно жури да присъди на Николай Ангелов Къскатийски образователната и научна степен „доктор“ в професионално направление 3.8. Икономика, научна специалност „Приложение на изчислителната техника в икономиката“.

16.5.2021 г.

гр. София

Подпис: