

Република Србија
Криминалистичко-полицијски универзитет
Департман информатике и рачунарства



Докторска дисертација

**Један приступ кластеровању саобраћајних
незгода у урбаним срединама**

МЕНТОР:

Проф. др Милан Гњатовић

КАНДИДАТ:

Иван М. Кошанин

Београд, 2023.

Republic of Serbia

University of Criminal Investigation and Police Studies

Department of Information Technology



Doctoral Dissertation

**An Approach to Traffic Accident Clustering
in Urban Areas**

SUPERVISOR:

Prof. Dr. Milan Gnjatović

STUDENT:

Ivan M. Košanin

Belgrade, 2023

Наслов докторске дисертације: „Један приступ кластеровању саобраћајних незгода у урбаним срединама“.

Сажетак: Ова теза предлаже нов приступ кластеровању саобраћајних незгода у урбаним срединама намењеном детектовању и селектовању критичних путних сегмената, тј., путних сегмената високог и просторно пролонгираног ризика за настанак саобраћајних незгода. Овај приступ полази од претпоставке да је свака саобраћајна незгода описана геопросторним координатама и временским тренутком у којем се десила, а намењен је за примену у условима ограничених људских и техничких ресурса. На нивоу спецификације, предложени алгоритам промовише стабилност резултата у времену и просторну колокацију незгода у различитим периодима, а санкционише величину селектованих кластера. На методолошком нивоу, овај алгоритам се разликује од других алгоритама који се примењују за кластеровање саобраћајних незгода по томе што се не заснива на тзв. густини опсервација, али испољава главне предности ових приступа: дозвољава кластере произвољног облика, не захтева априорно спецификовање броја кластера и искључује одступајуће опсервације. Коначно, приступ ће бити оцењен на корпусу који садржи јавно доступне податке о локацијама саобраћајних незгода са повређеним или погинулим лицима које су се десиле у трогодишњем периоду од јануара 2019. до децембра 2021. године у три највећа града у Републици Србији: Београду, Новом Саду и Нишу.

Кључне речи: саобраћајна незгода, просторно пролонгирани ризик, кластеровање, Гешталт-теорија, одређивање колена криве.

Научна област: Рачунарске науке.

Ужа научна област: Рачунарство.

The title of the doctoral dissertation: “An Approach to Traffic Accident Clustering in Urban Areas”.

Abstract: This thesis introduces a novel approach to traffic accident clustering in urban areas aimed at detecting and selecting critical road segments, i.e., road segments of spatially prolonged and high traffic accident risk. This approach assumes that a traffic accident is described by its positional coordinates and time of occurrence and is intended for application in circumstances of limited human or technical resources for traffic monitoring and management. At the specification level, the proposed algorithm promotes the stability of clustering results through time and inter-period accident spatial collocation, and penalize the size of selected clusters. At the methodological level, the proposed algorithm is not density-based, as are most other state-of-the-art clustering algorithms applied in the context of traffic accident analysis, but still keeps their main advantages: it allows for clusters of arbitrary shapes, does not require an a-priori given number of clusters, and excludes noisy observations. Finally, the proposed approach is validated with respect to publicly available data on traffic accidents involving injury or death that occurred in three of the largest cities of Serbia (Beograd, Novi Sad and Niš) over a three-year period that runs from January 2019 to December 2021.

Keywords: traffic accident, spatially prolonged risk, clustering, Gestalt theory, knee detection.

Scientific field: Computer Sciences.

Narrow scientific field: Scientific Computing.

САДРЖАЈ

Списак слика.....	viii
Списак табела	xi
1. Увод.....	1
2. Мотивација и преглед релевантних истраживања	3
2.1 Мотивација.....	3
2.2 Преглед релевантних истраживања.....	5
3. Основни алгоритам за кластерованье саобраћајних незгода	11
3.1 Основни појмови	11
3.2 Алгоритам за кластерованье.....	12
3.3 Селекција кластера.....	13
3.4 Илустрација алгоритма.....	15
3.4.1 Алати	15
3.4.2 Вредност просторног прага.....	15
3.4.3 Подаци.....	16
3.4.4 Извршавање алгоритма.....	17
3.4.5 Стабилност резултата у времену	21
3.5 Закључак.....	24
4. Доменски–специфичан критеријум за процену резултата кластерованья	25
4.1 Мера за процену резултата кластерованья саобраћајних незгода	25
4.1.1 Стабилност резултата кластерованья у времену.....	25
4.1.2 Просторна колокација незгода у различитим периодима	29
4.1.3 Релативна величина кластера.....	29
4.1.4 Интегрисана мера за процену резултата кластерованья.....	30
4.2 Аутоматско израчунавање вредности просторног прага	31

4.3	Резултати	35
4.3.1	Детаљан приказ резултата за Београд	40
4.3.2	Детаљан приказ резултата за Нови Сад	57
4.3.3	Детаљан приказ резултата за Ниш.....	68
4.4	Закључак.....	75
5.	Дискусија	77
5.1	Екстерна евалуација	77
5.2	Принцип близине у Гешталт-теорији.....	81
6.	Закључак.....	83
	Литература	85
	Стручна биографија	91
	Изјава о ауторству.....	92
	Изјава о истоветности штампане и електронске верзије докторске дисертације	93
	Изјава о коришћењу	94

Списак слика

Слика 2.1 Илустрација принципа близине у Гешталт-теорији. Слика је преузета из (Gnjatović et al. 2022).....	4
Слика 3.1 Приказ саобраћајних незгода са повређеним или погинулим лицима у централним општинама Београда у периоду од јануара 2021. до децембра 2021. Слика је генерисана применом компоненте ArcMap софтверског пакета ESRI ArcGis (https://www.esri.com) и преузета из (Gnjatović et al. 2022).....	17
Слика 3.2 Приказ селектованих кластера у Београду током 2021. године. Слика је генерисана применом компоненте ArcMap софтверског пакета ESRI ArcGis (https://www.esri.com) и преузета из (Gnjatović et al. 2022).....	18
Слика 3.3 Приказ селектованих кластера у Београду током 2021. године и локација камерних стубова у Београду у марту 2022. Слика је генерисана применом компоненте ArcMap софтверског пакета ESRI ArcGis (https://www.esri.com) и преузета из (Gnjatović et al. 2022).....	19
Слика 3.4 На левој страни: Приказ саобраћајних незгода са повређеним или погинулим лицима које су се догодиле на општини Звездара током (а) 2021., (в) 2020. и (д) 2019. године, респективно. На десној страни: Приказ селектованих кластера на основу анализе података из (б) 2021., (г) 2020. и (ђ) 2019. године, респективно. Локације камерних места (у складу са стањем из марта 2022. године) дате су због комплетности приказа. Слика је генерисана применом компоненте ArcMap софтверског пакета ESRI ArcGis (https://www.esri.com) и преузета из (Gnjatović et al. 2022).....	23
Слика 4.1 (а) Све саобраћајне незгоде са повређеним и погинулим лицима које су се догодиле на општини Звездара током периода P_1 . (б) Четири селектована кластера. (в) Све саобраћајне незгоде са погинулим или повређеним лицима које су се догодиле на општини Звездара током периода P_2 . Додатно, сваки од кластера приказаних на сл. 4.1(б) представљен је на сл. 4.1(в) минималним конвексним обухватајућим многоуглом. Слика је генерисана применом компоненте ArcMap софтверског пакета ESRI ArcGis (https://www.esri.com) и преузета из (Кошанин et al. 2023).....	27
Слика 4.2. (а) Наглашена конкавност (тј., тачка локалног максимума) и (б) ненаглашена конкавност у скуп D . Слика је преузета из (Кошанин et al. 2023).....	34
Слика 4.3 (а) График нормализоване интегрисане мере η у односу на примењену нормализовану вредност прага τ за Београд. (б–г) Графички приказ разлике између тачака	

нормализованог скупа података и главне дијагонале јединичног квадрата за Београд, Нови Сад и Ниш, респективно. Слика је преузета из (Кошанин et al. 2023).....	37
Слика 4.4 Селектовани кластери на општини Барајево.....	52
Слика 4.5 Селектовани кластери на општини Чукарица.....	52
Слика 4.6 Селектовани кластери на општини Гроцка.....	52
Слика 4.7 Селектовани кластери на општини Лазаревац.....	52
Слика 4.8 Селектовани кластери на општини Младеновац.....	53
Слика 4.9 Селектовани кластери на општини Нови Београд.....	53
Слика 4.10 Селектовани кластери на општини Обреновац.....	53
Слика 4.11 Селектовани кластери на општини Палилула.....	53
Слика 4.12 Селектовани кластери на општини Раковица.....	54
Слика 4.13 Селектовани кластери на општини Савски венац.....	54
Слика 4.14 Селектовани кластери на општини Сопот.....	54
Слика 4.15 Селектовани кластери на општини Стари град.....	54
Слика 4.16 Селектовани кластери на општини Сурчин.....	55
Слика 4.17 Селектовани кластери на општини Вождовац.....	55
Слика 4.18 Селектовани кластери на општини Врачар.....	55
Слика 4.19 Селектовани кластери на општини Земун.....	55
Слика 4.20 Селектовани кластери на општини Звездара.....	56
Слика 4.21 Селектовани кластери на општини Бач.....	65
Слика 4.22 Селектовани кластери на општини Бачка Паланка.....	65
Слика 4.23 Селектовани кластери на општини Бачки Петровац.....	65
Слика 4.24 Селектовани кластери на општини Бечеј.....	65
Слика 4.25 Селектовани кластери на општини Беочин.....	66
Слика 4.26 Селектовани кластери на општини Нови Сад.....	66
Слика 4.27 Селектовани кластери на општини Србобран.....	66
Слика 4.28 Селектовани кластери на општини Сремски Карловци.....	66
Слика 4.29 Селектовани кластери на општини Темерин.....	67
Слика 4.30 Селектовани кластери на општини Тител.....	67
Слика 4.31 Селектовани кластери на општини Врбас.....	67
Слика 4.32 Селектовани кластери на општини Жабалъ.....	67
Слика 4.33 Селектовани кластери на општини Алексинац.....	73
Слика 4.34 Селектовани кластери на општини Дољевац.....	73
Слика 4.35 Селектовани кластери на општини Гаџин Хан.....	73

Слика 4.36 Селектовани кластери на општини Мeroшина.....	73
Слика 4.37 Селектовани кластери на општини Ниш.....	74
Слика 4.38 Селектовани кластери на општини Нишка Бања.....	74
Слика 4.39 Селектовани кластери на општини Ражањ.....	74
Слика 4.40 Селектовани кластери на општини Сврљиг.....	74
Слика 5.1 Шест (од девет, в. сл. 5.2) централних општина Београда са локацијама камерних стубова у августу 2020. године. За сваку општину приказани су кластери добијени применом вредности просторног прага $\tau=150\text{м}$. Сваки кластер саобраћајних незгода представљен је минималним конвексним обухватајућим многоуглом. Слика је генерисана применом компоненте ArcMap софтверског пакета ESRI ArcGis (https://www.esri.com) и преузета из (Košanin et al. 2023).....	79
Слика 5.2 Три (од девет, в. сл. 5.1) централних општина Београда са локацијама камерних стубова у августу 2020. године. За сваку општину приказани су кластери добијени применом вредности просторног прага $\tau=150\text{м}$. Сваки кластер саобраћајних незгода представљен је минималним конвексним обухватајућим многоуглом. Слика је генерисана применом компоненте ArcMap софтверског пакета ESRI ArcGis (https://www.esri.com) и преузета из (Košanin et al. 2023).....	80
Слика 5.3 Разлика у концептуализацији принципа близине у Гешталт-теорије у (i) Osbourn & Martinez (1995) и (ii) приступ у овој тези. Слика је преузета из (Gnjatović et al. 2023).....	81

Списак табела

Табела 2.1 Преглед скорашњих истраживања у којима се примењује кластеровање у контексту анализе саобраћаја.....	6
Табела 3.1: Резиме резултата кластеровања и селектовања доминантних кластера добијених применом основног алгоритма на јавно доступне податке о саобраћајним незгодама са повређеним или погинулим лицима које су се догодиле током 2021. године у Београду.....	18
Табела 3.2: Опис резултата кластеровања и селектовања доминантних кластера добијених применом основног алгоритма на јавно доступне податке о саобраћајним незгодама са повређеним или погинулим лицима које су се догодиле током 2021. године у Београду. Приказани идентификациони бројеви кластера одговарају идентификационим бројевима датим у легендама на сл. 3.2 и 3.3.....	19
Табела 3.3 Резиме добијених резултата кластеровања након примене алгоритма на јавно доступне податке о саобраћајним незгодама са погинулим и повређеним лицима које су се догодиле на општини Звездара током 2019., 2020., 2021., респективно.....	22
Табела 3.4 Опис селектованих кластера добијених применом алгоритма на јавно доступне податке о саобраћајним незгодама са погинулим и повређеним лицима које су се догодиле на општини Звездара током 2019., 2020. и 2021., респективно.....	22
Табела 4.1 Процењивање стабилности резултата кластеровања у времену. Алгоритам се примењује на податке о саобраћајним незгодама са повређеним или погинулим лицима у Београду током периода P_1 и P_2 , за вредност прага $\tau=170m$. Сви децимални бројеви у табели заокружени су на три децимална места	28
Табела 4.2 Релативна величина свих селектованих кластера у Београду током периода P_1 за вредност прага $\tau=170m$. Сви децимални бројеви у табели заокружени су на три децимална места.....	30
Табела 4.3. Бројеви саобраћајних незгода са погинулим и повређеним лицима које су се десиле у трогодишњем периоду од јануара 2019. до децембра 2021. године у три највећа града у Републици Србији: Београду, Новом Саду и Нишу.....	35
Табела 4.4 Вредности мера представљених у секцији 4.1, добијених кад се представљени алгоритам примени на податке о саобраћајним незгодама које су се догодиле у Београду у периоду 2019-2021. и вредности просторног прага из секвенце (4.25). Сви децимални бројеви у табели заокружени су на три децимална места.....	36

Табела 4.5 Резултати кластеровања за Београд ($\tau^*=150\text{м}$). Сви децимални бројеви у табели заокружени су на три децимална места.....	38
Табела 4.6 Резултати кластеровања за Нови Сад ($\tau^*=170\text{м}$). Сви децимални бројеви у табели заокружени су, кад је прикладно, на три децимална места.....	39
Табела 4.7 Резултати кластеровања за Ниш ($\tau^*=160\text{м}$). Сви децимални бројеви у табели заокружени су, кад је прикладно, на три децимална места.....	39
Табела 4.8 Сумарни подаци резултата кластеровања саобраћајних незгода у Београду ($\tau^*=150\text{м}$).....	41
Табела 4.9 Збир података о апсолутним и релевантним површинама селектованих кластера по Општинама у Београду.....	42
Табела 4.10 Површине појединачних кластера на општини Звездара.....	42
Табела 4.11 Површине појединачних кластера на општини Раковица.....	43
Табела 4.12 Површине појединачних кластера на општини Врачар.....	43
Табела 4.13 Површине појединачних кластера на општини Вождовац.....	43
Табела 4.14 Површине појединачних кластера на општини Палилула.....	44
Табела 4.15 Површина кластера на општини Стари Град.....	44
Табела 4.16 Површине појединачних кластера на општини Савски Венац.....	44
Табела 4.17 Површина кластера на општини Нови Београд.....	44
Табела 4.18 Површине појединачних кластера на општини Земун.....	45
Табела 4.19 Површине појединачних кластера на општини Сурчин.....	45
Табела 4.20 Површине појединачних кластера на општини Чукарица.....	46
Табела 4.21 Површине појединачних кластера на општини Младеновац.....	47
Табела 4.22 Површине појединачних кластера на општини Гроцка.....	47
Табела 4.23 Површине појединачних кластера на општини Лазаревац.....	49
Табела 4.24 Површине појединачних кластера на општини Сопот.....	50
Табела 4.25 Површине појединачних кластера на општини Барајево.....	51
Табела 4.26 Површине појединачних кластера на општини Обреновац.....	51
Табела 4.27 Сумарни подаци резултата кластеровања саобраћајних незгода у Новом Саду ($\tau^*=170\text{м}$).....	58
Табела 4.28 Збир података о апсолутним и релевантним површинама селектованих кластера по Општинама у Новом Саду.....	59
Табела 4.29 Површине појединачних кластера на општини Бач.....	59
Табела 4.30 Површине појединачних кластера на општини Бачка Паланка.....	60
Табела 4.31 Површине појединачних кластера на општини Бачки Петровац.....	60

Табела 4.32 Површине појединачних кластера на општини Бечеј.....	61
Табела 4.33 Површине појединачних кластера на општини Беочин.....	61
Табела 4.34 Површине појединачних кластера на општини Нови Сад.....	62
Табела 4.35 Површине појединачних кластера на општини Србобран.....	62
Табела 4.36 Површине појединачних кластера на општини Сремски Карловци.....	62
Табела 4.37 Површине појединачних кластера на општини Темерин.....	63
Табела 4.38 Површина кластера на општини Тител.....	63
Табела 4.39 Површине појединачних кластера на општини Врбас.....	64
Табела 4.40 Површине појединачних кластера на општини Жабаљ.....	64
Табела 4.41 Сумарни подаци резултата кластеровача саобраћајних незгода у Нишу ($\tau^*=160\text{m}$).....	69
Табела 4.42 Збир података о апсолутним и релевантним површинама селектованих кластера по Општинама у Нишу.....	70
Табела 4.43 Површине појединачних кластера на општини Алексинац.....	70
Табела 4.44 Површине појединачних кластера на општини Дољевац.....	71
Табела 4.45 Површина кластера на општини Гацин Хан.....	71
Табела 4.46 Површине појединачних кластера на општини Меровина.....	71
Табела 4.47 Површине појединачних кластера на општини Ниш.....	71
Табела 4.48 Површине појединачних кластера на општини Нишка Бања.....	72
Табела 4.49 Површине појединачних кластера на општини Ражањ.....	72
Табела 4.50 Површине појединачних кластера на општини Сврљиг.....	72
Табела 5.1 Резултати екстерне евалуације.....	80

1. Увод

Саобраћајне незгоде представљају глобални друштвени и здравствени проблем. Процењено је да у саобраћајним незгодама годишње погине 1,35 милиона људи, док до педесет милиона људи бива повређено (Aroga, 2021), а трошкови на нивоу држава износе приближно три процента годишњег бруто домаћег производа (European Parliament, 2022). Надзор и управљање саобраћајем представљају један од начина за повећање нивоа безбедности саобраћаја. У складу с тим, извесна истраживачка пажња већ је посвећена проблему аутоматског детектовања места са повећаним ризиком од саобраћајних незгода. Ова дисертација посматра специфични аспект овог проблема, који се односи на аутоматско детектовање и селектовање критичних путних сегмената који су кандидати за примарни предмет надзора и управљања у условима ограничених људских и техничких ресурса.

Предмет научног истраживања у овој дисертацији односи се на развој новог приступа аутоматском кластеровачу саобраћајних незгода у урбаним срединама намењеном за контекстно-зависно детектовање и селектовање путних сегмената високог и просторно пролонгираног ризика за настанак саобраћајних незгода.

Циљ научног истраживања у овој дисертацији односи се на развој и валидирање новог приступа аутоматском кластеровачу саобраћајних незгода у урбаним срединама, у складу с постављеним предметом истраживања. Основни алгоритам за кластероваче извршава се у две фазе. Прва фаза представља адаптацију графовски заснованог алгоритма за сегментацију дигиталних слика инспирисану изабраним принципима Гешталт-теорије, са једним хиперпараметром који се односи на праг растојања. Друга фаза представља селектовање кластера који се доминантни у односу на број саобраћајних незгода, засновано на адаптацији алгоритма за бинаризовање дигиталних слика без улазних параметара. Поред тога, уводи се нов, доменски-специфичан критеријум за процену резултата кластеровача, који промовише стабилност резултата у времену и просторну колокацију незгода у различитим периодима, а санкционише величину селектованих кластера. На основу овог критеријума и адаптације алгоритма за одређивање колена криве врши се аутоматско одређивање вредности улазног хиперпараметра.

Општа хипотеза научног истраживања у оквиру предложене докторске дисертације може се формулисати на следећи начин: На основу података о геопросторним координатама и временским тренуцима саобраћајних незгода које су се десиле у датој

урбаној средини и у датом временском периоду могу се детектовати путни сегменти високог и просторно пролонгираног ризика за настанак саобраћајних незгода.

Посебна хипотеза научног истраживања у оквиру предложене докторске дисертације може се формулисати на следећи начин: Аутоматско детектовање путних сегмената високог и просторно пролонгираног ризика за настанак саобраћајних незгода може се реализовати кластеровањем података о геопросторним координатама и временским тренуцима саобраћајних незгода на начин који дозвољава кластере произвољног облика, не захтева априорно спецификовање броја кластера и искључује одступајуће опсервације.

Резултати научног истраживања спроведеног у оквиру ове дисертације укључују дизајн и валидирање новог рачунарског алгоритма за аутоматско кластеровање саобраћајних незгода у урбаним срединама намењеном за контекстно-зависно детектовање и селектовање путних сегмената високог и просторно пролонгираног ризика за настанак саобраћајних незгода. Предложени алгоритам ће бити оцењен на корпусу који садржи јавно доступне податке о локацијама и временима саобраћајних незгода које су се десиле у периоду 2019-2021. у три највећа града Републике Србије (в. Gnjatović et al. 2022; Košanin et al. 2023).

Остатак дисертације структурисан је на следећи начин: Друго поглавље представља преглед релевантних истраживања и алгоритама у области аутоматског кластеровања саобраћајних незгода. Треће поглавље уводи основни алгоритам за кластеровање саобраћајних незгода, са једним хиперпараметром који се односи на праг растојања, а потом га илуструје на реалним подацима. Четврто поглавље представља доменски-специфичан критеријум за процену резултата кластеровања саобраћајних незгода, на основу којег се потом врши аутоматско израчунавање хиперпараметра алгоритма за кластеровање, што је демонстрирано на реалним подацима. Пето и шесто поглавље дискутују представљене доприносе и закључују дисертацију.

2. Мотивација и преглед релевантних истраживања

Кластерованье има важну улогу у анализирању саобраћајних незгода. Две истраживачке линије добијају највећу пажњу у овој области. Прва линија се односи на саобраћајну логистику, на пример: анализа саобраћајног оптерећења и рутирање возила итд. Друга линија се односи на сигурност саобраћаја, на пример: откривање образаца саобраћајних незгода, откривање црних тачака и критичних сегмената пута итд.

Истраживање представљено у овој тези тематски припада другој истраживачкој линији. У секцији 2.1 мотивисан је приступ кластерованью који је изложен у овој тези. У секцији 2.2 дат је преглед скорашњих истраживања у овој области.

2.1 Мотивација

Истраживачком питању кластерованья саобраћајних незгода посвећена је значајна пажња (Jeong et al. 2022; в. и Manap et al. 2021; в. и Santos et al. 2021; в. и Gutierrez-Osorio & Pedraza 2020; в. и Sun et al. 2020). За свеобухватни преглед, читаоци могу погледати (Jeong et al. 2022; Wokaba et al. 2022; Islam et al. 2021). Неки од често примењиваних алгоритама за кластерованье (на пример, алгоритам k-means, Lloyd 1982; Selim & Ismail 1984) узима број кластера као улазни параметар (в. и Dukan & Kovari, 1984). У пракси, посматрани подаци се вишеструко кластерују за различите бројеве кластера, а потом се одређује оптималан број кластера по неком задатом критеријуму. Један од критеријума заснива се на дисперзији унутар кластера (Tibshirani et al. 2001):

$$WCSS(t) = \sum_{i=1}^t \left(\frac{1}{2n_i} \sum_{j,k \in C_i} d_{jk} \right), \quad (2.1)$$

где су:

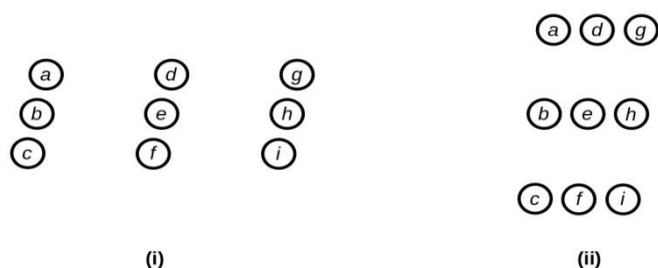
- t – број кластера,
- C_i – i -ти кластер,
- n_i – број опсервација додељен кластеру C_i ,
- d_{jk} – растојање између опсервација j и k .

Граф дисперзије унутар кластера у односу на примењени број кластера обично садржи колена које указује на оптималан број кластера (Tibshirani et al. 2001). Алтернативни метод да се одреди оптималан број кластера описан је у (Rousseeuw 1987).

Захтев да се унапред одреди број кластера представља ограничење. Додатно, алгоритам k-means узима у обзир цео скуп података и генерише кластере сферичног облика, што није нужно прикладно за представљање кластера саобраћајних незгода (Islam et al. 2021). Да би се ово ограничење превазишло, алгоритам базиран на густини, DBSCAN (Ester et al. 1996), искључује одступајуће опсервације и дозвољава генерисање кластера произвољних облика. Уместо унапред задатог броја кластера, овај алгоритам прихвата два различита параметра: максимални радијус околине и минимални број тачака за формирање густог региона. Алгоритам OPTICS (Ankerst et al. 1999) представља проширење алгоритма DBSCAN, које генерише структуру кластера датог скупа података засновану на густини.

У (Islam et al. 2021) приказано је да алгоритми кластеровача засновани на густини имају боље перформансе него алгоритам k-means у контексту анализе саобраћајних незгода. Слично њима, алгоритам представљен у овој тези дозвољава генерисање кластера произвољног облика и не захтева да се унапред одреди број кластера. Предложен приступ кластеровачу није базиран на густини, али је инспирисан принципом близине у Гешталт-теорији (Wertheimer 1938). Према овом принципу, када су људи суочени са великим бројем истих визуелних стимуланса (на пример: тачка на дводимензионалној равни), најприроднија форма груписања заснива се на најмањим интервалима.

На пример, за скуп тачака који је приказан на слици 2.1(i), најприродније уређење било би abc/def/ghi, али за скуп приказан на слици 2.1(ii), природно кластероваче било би adg/beh/cfi. Важно је напоменути да на овај принцип природног груписања не утиче повећање броја тачака (Wertheimer 1938).



Слика 2.1 Илустрација принципа близине у Гешталт-теорији. Слика је преузета из (Gnjatović et al. 2022).

У овој тези представљен је нови приступ аутоматском просторном кластеровању саобраћајних незгода у урбаним срединама инспирисан принципом близине у Гешталт-теорији. На нивоу спецификације, алгоритам има за циљ проналажење путних сегмената високог и просторно пролонгираног ризика за настанак саобраћајних незгода. Сегмент пута је посматран као пут просторно пролонгираног ризика ако је повезан са непразним скупом N локација саобраћајних незгода, које се могу сматрати просторно блиским у смислу транзитивног затварања. Прецизније, нека је R релација дефинисана на скупу N на следећи начин:

$$R = \{(n_i, n_j) \mid d(n_i, n_j) \leq \tau\}, \quad (2.2)$$

где су:

- n_i и n_j – саобраћајне незгоде,
- τ – просторни праг,
- $d(n_i, n_j)$ – просторно растојање између саобраћајних незгода n_i и n_j .

Кластер се формира као транзитивно затварање релације R , а детектовање сегмента пута просторно пролонгираног ризика постиже се кластеризовањем. Додатно, сегмент пута се сматра путем високог саобраћајног ризика, ако одговара кластеру који је доминантан у односу на број незгода.

2.2 Преглед релевантних истраживања

Преглед скорашњих истраживања у којима се примењује кластеровање (Aggarwal & Reddy 2014) у контексту анализе саобраћаја дат је у табели 2.1 (в. Košanin et al. 2023). У наставку рада дат је кратак преглед ових радова.

У (Zhao et al. 2023) приказана је анализа саобраћајног оптерећења која је спроведена применом унапређеног алгоритма k -means. Анализу дугорочних ефеката саобраћајног оптерећења на мостовима тешко је спровести због различитости саобраћајних токова у погледу интензитета саобраћаја и врсте возила. У овом раду посматра се симулација саобраћајног оптерећења на мостовима. Подаци су прикупљени применом система за мерење тежине возила у покрету, а потом су из њих генерисан стохастички ток саобраћаја у више трака. Добијени резултати показали су да се

саобраћајно оптерећење може поделити на кластере са идентичним карактеристикама расподеле.

Табела 2.1 Преглед скорошњих истраживања у којима се примењује кластероване у контексту анализе саобраћаја.

Истраживачки проблем	Методолошки приступ	Референца
анализа саобраћајног оптерећења	унапређени алгоритам k-means	(Zhao et al. 2023)
анализа саобраћајних загушења	самоорганизоване мапе (неуралне мреже)	(Zang et al. 2023)
класификација стања у саобраћају	алгоритам k-medoids	(Shang et al. 2022)
идентификација нивоа путне мреже	алгоритам k-means	(Hernández et al. 2021)
анализа саобраћајних загушења	сиво релационо кластероване	(Zhang et al. 2016)
издвајање образаца саобраћајних незгода	алгоритам ROCK	(Esenturk et al. 2022a)
издвајање образаца саобраћајних незгода	алгоритам COOLCAT	(Esenturk et al. 2022b)
анализа фактора саобраћајних незгода	алгоритам k-means	(Niu, et al. 2022)
моделовање саобраћајних незгода	компаративна анализа класификатора заснованих на машинском учењу	(Bokaba, et al. 2022)
идентификовање црних тачака	алгоритам HDBSCAN	(Wang, et al. 2023)
анализа саобраћајних загушења	алгоритам k-means	(Li & Huang 2023)
анализа ризичног понашања у возњи	алгоритам k-means	(Chen, et al. 2023)
оптимално рутирање путање	модификовани алгоритам k-medoids	(Shah, et al. 2022)
анализа саобраћајних незгода са погинулим и тешко повређеним пешацима	метода KDE	(Rampinelli, et al. 2022)
систем за управљање саобраћајем	алгоритам DBSCAN	(Lilhore, et al. 2022)
анализа тежине саобраћајне незгоде	алгоритам DBSCAN	(Jeong et al. 2022)
процена безбедности на аутопуту	алгоритам k-means	(Baek 2022)
анализа озбиљности саобраћајних незгода са пешацима	метода KDE	(Bajada & Attard 2021)

У (Zang et al. 2023) приказана је анализа саобраћајног загушења која је добијена коришћењем самоорганизованих мапа (тј., неуралних мрежа). Анализа саобраћајног загушења има за циљ да се формулише ефикасна стратегија за управљање саобраћајним загушењима. Истраживања саобраћајних загушења често су посвећена утврђивању образаца на основу саобраћајних токова, густине и брзине саобраћаја на одређеном подручју. Циљ посматраног истраживања односи се на препознавање образаца саобраћајних загушења на подручју града анализом индекса саобраћајних перформанси у временском периоду од 699 дана у 2018., 2019., и 2021. години у Пекингу. Метода самоорганизованих мапа у колаборацији са побољшаним алгоритмом за одређивање броја кластера, коришћена је за добијање образаца на основу индекса саобраћајних перформанси (TPI) кроз време. Саобраћајна загушења су кластерована у односу на дане у недељи (понедељак, петак, субота итд.) и према догађајима (фестивалима итд.).

У (Shang et al. 2022) приказано је класификовање стања у саобраћају засновано на алгоритму k-medoids. Класификовање стања у саобраћају примењује се као подршка управљању саобраћајним загушењима. У раду је представљена нова хибридна метода за класификовање стања у саобраћају заснована на ненадгледаном кластеровању. Алгоритам k-medoids користи се за кластеровање података о брзини саобраћаја са више тачака у посматраном подручју, а затим се бирају тачке које представљају центре кластера за даљу анализу. Након тога се примењује алгоритам за спектрално кластеровања на податке о брзини, протоку и заузетости циљних тачака.

У (Hernández et al. 2021) приказана је идентификација нивоа путне мреже која је добијена коришћењем алгоритма k-means. Анализирани су подаци о саобраћају из 2019. године у Баскији (Шпанија). Главна променљива је просечан дневни саобраћај, који представља главни индикатор у многим областима идентификације путне мреже. У првој фази примењена је анализа из које су добијене дескриптивне статистичке информације, а затим је извршено кластеровање над различитим добијеним променљивима. Након тога, променљива од интереса је процењена за целу путну мрежу применом радијаних функција. Показано је да је кластеровање погодно за идентификацију значаја појединачних сегмената пута, а интерполација заснована на радијалним функцијама за процену просечног годишњег дневног саобраћаја.

У (Zhang et al. 2016) приказана је анализа саобраћајног загушења која је добијена коришћењем модела за сиво релационо кластеровање. У раду су прикупљене информације о брзини, густини и обиму саобраћаја на путевима у урбаним срединама, који су потом коришћене за анализу саобраћајног загушења. Резултати показују да је

просечна тачност предложеног модела знатно већа од тачности добијених применом алгоритама k-means и fuzzy c-means.

У (Esenturk et al. 2022a) приказано је издвајање образаца саобраћајних незгода применом алгорита ROCK. Анализе овог типа могу се користити за (а) проналажења веза између услова пре настанка саобраћајних незгода и (б) утврђивање стратегија за спречавање саобраћајних незгода. Доприноси посматраног рада односе се на оба аспекта анализе: (а) идентификовање образаца у саобраћајним незгодама и (б) разматрање могућности примене повезаних и аутоматизованих возила. Увођење оваквих возила у саобраћај на путевима није могуће уколико се не изврши тестирање на великом броју саобраћајних сценарија, како би се доказала безбедност. У раду први циљ је идентификовање трендова у хетерогеним подацима о саобраћајним незгодама из базе података UK's STATS19, применом алгорита ROCK. На добијеним кластерима спроведене су фреквенцијска и геометријска анализа. На основу примене методе за рударење података и геометријске интерпретације кластера, предложен је приступ генерисању сценарија за повезана и аутоматизована возила.

У (Esenturk et al. 2022b) приказано је издвајање образаца саобраћајних незгода које је добијено коришћењем алгорита COOLCAT.

У (Niu, et al. 2022) приказана је анализа фактора саобраћајних незгода које је добијена коришћењем алгорита k-means. На јавно доступним подацима о саобраћајним незгодама примењени су алгоритми за рударење података (тј., априорни алгоритам). Из података су издвојени главни атрибути, над којим је потом примењен алгоритам k-means, како би се класификовао ниво саобраћајних незгода према броју лица.

У (Vokaba, et al. 2022) приказано је моделовање саобраћајних незгода које је добијено коришћењем упоредне анализе класификатора заснованих на машинском учењу. У раду је извршена анализа перформанси различитих класификатора (наивни бајесовски класификатор, логистичка регресија, k-nearest neighbors, AdaBoost, метода потпорних вектора, ансамбл стабала одлучивања) који су примењени на подацима о саобраћајним незгодама које су се догодиле у Гаутенгу у Јужноафричкој Републици. Класификатори су процењени на основу пет мера: тачност, средња квадратна грешка, прецизност, одзив и ROC-крива.

У (Wang, et al. 2023) приказано је идентификовање црних тачака које је добијено коришћењем алгорита HDBSCAN. Представљен је модел за идентификацију црних тачака који интегрише географски информациони система и хијерархијско кластеровање засновано на густини. Модел је валидиран на подацима о 3536 саобраћајних незгода које

су се догодиле у Хангџоу, Кини у периоду од првог августа до тридесет и првог октобра 2020. године.

У (Li & Huang 2023) приказана је анализа саобраћајних загушења које је добијена коришћењем алгоритам *k-means*. У раду се врши фузија динамичких и статичких карактеристика саобраћаја и предлаже графовски заснована репрезентација мреже путева. Потом се критични делови путева одређују применом алгоритма за кластеровање.

У (Chen, et al. 2023) приказана је анализа ризичног понашања у вожњи која је добијена коришћењем алгоритам *k-means*. У раду су посматрана четири индикатора за процену ризика понашања у вожњи који се односе на бочну стабилност возила, уздужну стабилност возила, праћење возила и промену саобраћајне траке. Корелација између ових индикатора одређена је применом Пирсоновог коефицијента корелације, при чему је показано да су слабо корелисани. Кластеровањем су одређене четири категорије ризичне вожње.

У (Shah, et al. 2022) приказан је приступ рутирању возних путања која је добијена коришћењем модификованог алгоритма *k-medoids*. Предложени метод генерише ад-хок мреже возила у циљу одређивања субоптималних путања. Кластеровање се спроводи у односу на два главна критеријума: избегавање саобраћајних незгода и остваривање ефикасног и поузданог преноса порука у условима високе мобилности.

У (Rampinelli, et al. 2022) приказана је анализа саобраћајних незгода са погинулим и тешко повређеним пешацима која је добијена коришћењем методе KDE. У раду се анализирају фактори ризика, из просторне и статистичке перспективе, који се односе на карактеристике пешака, судара и окружења и доприносе различитим нивоима повреде пешака у Сантијагу, Чилеу. Применом методе KDE идентификују се просторни кластери који садрже високу концентрацију погинулих и тешко повређених пешака. Након тога, коришћењем скупа података о саобраћајним незгодама за цео град, одређују се фактори ризика који утичу на тежину повреде и судара пешака на просторном кластеру.

У (Lilhore, et al. 2022) приказан је систем за управљање саобраћајем који је добијен коришћењем алгоритма DBSCAN. Адаптивни систем за управљање саобраћајем заснован је на Интернету ствари и методама машинског учења. Адаптивни систем садржи велики број сценарија за управљање саобраћајем, а његов дизајн се заснива на три основа ентитета: возилу, инфраструктури и догађају. Циљ примењеног кластеровања јесте откривање случајних аномалија.

У (Jeong et al. 2022) приказана је анализа тежине саобраћајних незгода која је добијена коришћењем алгоритма DBSCAN. У раду, главни циљ је идентификација фактора и типова који утичу на саобраћајне незгоде у Сеулу. Над скупом података о саобраћајним незгодама у Сеулу спроведене су анализе различитим техникама надгледаног и ненадгледаног машинског учења. Критични фактори који утичу на озбиљност саобраћајне незгоде (алгоритам заснован на регресији) и доминантни типови саобраћајних незгода применом алгоритама надгледаног машинског учења (ансамбли и регресиони модели), а доминантни типови саобраћајних незгода применом алгоритама ненадгледаног машинског учења (кластероване). Важно је напоменути да фактори средине (време, годишње доба, дан у недељи) не утичу на тежину саобраћајних незгода у Сеулу.

У (Baek 2022) приказана је процена безбедности на аутопуту која је добијена коришћењем алгоритма k-means. У раду се предлаже метод класификације који је примењен на административна подручја у Јужној Кореји. Метод је примењен на податке добијене од стране 536 станице за процену обима саобраћаја на националним путевима у Јужној Кореји у 2019. години. Главну објашњавајућу променљиву у раду представља однос просечног годишњег обима саобраћаја по сату и дану, респективно.

У (Bajada & Attard 2021) приказана је анализа озбиљности саобраћајних незгода са пешацима која је добијена коришћењем метода KDE. Овај тип саобраћајних незгода карактеристичан је за урбана подручја, а у посматраном раду посматране су саобраћајне незгоде које су се догодиле у периоду од 2004. до 2018. године на Малти. Прво су посматрани односи између променљивих (старост, пол, месец, дан, време, округ, повреда и категорија) користећи методу анализе вишеструких подударарања (Multiple Correspondence Analysis), а затим се добијени резултати кластеровани применом методе KDE.

Коначно, треба истаћи да се различити приступи кластерованја саобраћајних незгода комбинују са применом географских информационих система (Wang, et al. 2023; Satria & Castro 2016), што се такође налази на линији истраживања представљеног у овој дисертацији.

3. Основни алгоритам за кластеровање саобраћајних незгода

У овом поглављу биће представљен двостепени приступ кластеровању и селектовању саобраћајних незгода (в. Gnjatović et al., 2022). Поглавље је структуисано на следећи начин. У секцији 3.1 уведени су основни појмови који ће бити коришћени у представљању посматраног алгоритма. Секција 3.2 описује графовски заснован приступ просторном кластеровању саобраћајних незгода, а секција 3.3 приступ адаптивном селектовању кластера који су доминантни у односу на број саобраћајних незгода које обухватају.

3.1 Основни појмови

Саобраћајна незгода n_i представљена је следећом уређеном тројком:

$$n_i = (id_i, \varphi_i, \lambda_i), \quad (3.1)$$

где су:

- id_i – јединствени идентификациони број саобраћајне незгоде n_i ,
- φ_i и λ_i – геопросторне координате саобраћајне незгоде n_i , изражене у радијусима.

Просторно растојање између саобраћајних незгода n_i и n_j израчунава се на основу следеће формуле (Sinnott, 1984):

$$d(n_i, n_j) = 2 \cdot R \cdot \text{atan2}(\sqrt{a(n_i, n_j)} \cdot \sqrt{1 - a(n_i, n_j)}), \quad (3.2)$$

где су:

$$a(n_i, n_j) = \sin^2 \frac{\varphi_j - \varphi_i}{2} + \cos \varphi_i \cos \varphi_j \sin^2 \frac{\lambda_j - \lambda_i}{2}, \quad (3.3)$$

функција atan2 означава адаптацију функције аркус тангенс спецификоване за једнозначно израчунавање вредности угла, а константа R представља средњи полупречник Земље, чија је вредност једнака $6371 \cdot 10^3$ м. Додатно, нека је τ просторни праг који представља улазни параметар алгоритма за кластеровање и нека је $N = \{n_1, n_2, \dots, n_k\}$ скуп саобраћајних незгода које су се догодиле у датом периоду.

3.2 Алгоритам за кластеровање

Предложени приступ кластеровању представља прилагођење графовски заснованог алгоритма за сегментовање дигиталних слика (Felzenszwalb & Huttenlocher 2004; в. и Gnjatović, Mašek & Adamović 2019) и извршава се кроз следеће кораке:

1. У сваком тренутку извршавања алгоритма, тренутни резултат кластеровања представљен је низом целих бројева:

$$C = (c(n_1), c(n_2), \dots, c(n_k)), \quad (3.4)$$

где важи $(\forall 1 \leq i \leq k)(c(n_i) \in \{1, 2, \dots, k\})$, а $c(n_i)$ представља идентификациони број кластера којем је саобраћајна незгода n_i додељена. У првом кораку алгоритма, свака саобраћајна незгода додељена је засебном кластеру:

$$(\forall 1 \leq i \leq k)(c(n_i) = i). \quad (3.5)$$

2. Нека је $D(N, \tau)$ скуп свих комбинација без понављања друге класе над скупом N , тј., скуп свих неуређених парова саобраћајних незгода чије је међусобно растојање мање од прага τ или му је једнако. Другим речима, скуп $D(N, \tau)$ садржи парове просторно блиских саобраћајних незгода које су кандидати да буду у истом кластеру. Без губитка општости, скуп $D(N, \tau)$ може се дефинисати на следећи начин:

$$D(N, \tau) = \{(n_i, n_j) | \{n_i, n_j\} \subset N \wedge i < j \wedge d(n_i, n_j) \leq \tau\}. \quad (3.6)$$

3. Генерише се секвенца која садржи све елементе из скупа $D(N, \tau)$ уређене неоппадајуће по удаљености између саобраћајних незгода.

$$\hat{D}(N, \tau) = (n_{i_1}, n_{j_1}), (n_{i_2}, n_{j_2}), \dots, (n_{i_m}, n_{j_m}). \quad (3.7)$$

4. Пролази се кроз секвенцу $\hat{D}(N, \tau)$, почевши од прве и закључно с последњом позицијом. Током проласка кроз секвенцу, за сваки уређени пар $\delta_p = (n_i, n_j)$ у секвенци

$\widehat{D}(N, \tau)$ врши се провера да ли саобраћајне незгоде n_i и n_j припадају различитим кластерима, тј., да ли важи $c(n_i) \neq c(n_j)$. Ако је овај услов испуњен, врши се спајање кластера којим посматране саобраћајне незгоде припадају:

$$\begin{aligned}
 & \text{for } (1 \leq p \leq |\widehat{D}(N, \tau)|) \{ \\
 & \quad \text{let } \delta_p = (n_i, n_j) \\
 & \quad \text{if } (c(n_i) \neq c(n_j)) \text{ then} \\
 & \quad \quad \text{for } (1 \leq q \leq |C|) \\
 & \quad \quad \quad \text{if } (c(n_q) = c(n_j)) \text{ then } (c(n_q) \leftarrow c(n_i)) \\
 & \quad \} \\
 & \}
 \end{aligned} \tag{3.8}$$

На овај начин, кластери су одређени на основу транзитивног затварања бинарне релације дефинисане у другом кораку горњег алгоритма. Резултати кластерована у овој фази алгоритма репрезентовани су низом C након завршетка четвртог корака у горњем алгоритму. У општем случају, број формираних кластера једнак је броју различитих вредности у низу C .

3.3 Селекција кластера

У другој фази алгоритма издваја се подскуп кластера који су у датом контексту доминантни у односу на број саобраћајних незгода које обухватају. Нека је $\chi(C)$ хистограм низа C :

$$\chi(C) = \{(c_1, p_1), (c_2, p_2), \dots, (c_t, p_t)\}, \tag{3.9}$$

где су:

- c_i – идентификациони број кластера,
- p_i – број саобраћајних незгода које припадају c_i ,
- t – број кластера у низу C ,
- и $1 \leq i \leq t$.

Алгоритам за адаптивно селектовање кластера представља прилагођење методе за адаптивно одређивање вредности глобалног прага у поступку бинаризовања дигиталне слике (Shih 2010; в. и Gnjatović et al., 2018) и извршава се кроз следеће кораке:

1. Почетна вредност прага μ_0 једнака је просечном броју саобраћајних незгода по кластеру у низу C :

$$\mu_0 = \frac{1}{|\chi(C)|} \sum_{i=1}^{|\chi(C)|} p_i. \quad (3.10)$$

2. За текућу вредност прага μ_i где је $i \geq 0$, скуп $\chi(C)$ се дели на два дисјунктна подскупа:

$$\begin{aligned} \chi_1 &= \{(c, p) \mid (c, p) \in \chi(C) \wedge p \leq \mu_i\}, \\ \chi_2 &= \{(c, p) \mid (c, p) \in \chi(C) \wedge p > \mu_i\}, \end{aligned} \quad (3.11)$$

а следећа вредност прага μ_{i+1} израчунава се на следећи начин:

$$\mu_{i+1} = \frac{1}{2} \left(\frac{1}{|\chi_1|} \sum_{i=1}^{|\chi_1|} p_i + \frac{1}{|\chi_2|} \sum_{i=1}^{|\chi_2|} p_i \right). \quad (3.12)$$

3. Ако промена вредности прага није значајна, тј., ако важи:

$$|\mu_i - \mu_{i+1}| \leq \frac{1}{2}, \quad (3.13)$$

алгоритам се завршава, а прагу μ додељује се вредност μ_{i+1} . У супротном, процес се враћа на корак 2.

Коначно, подскуп кластера који су у датом контексту доминантни у односу на број саобраћајних незгода које обухватају одређује се на следећи начин:

$$C = \{c \mid (c, p) \in \chi(C) \wedge p > \mu\}, \quad (3.14)$$

при чему треба имати у виду да μ представља адаптивно израчунати праг за потребе издвајања доминантних кластера из низа C , док τ (в. секцију 3.2) представља просторни праг и улазни параметар алгоритма.

3.4 Илустрација алгоритма

У овој секцији илустрована је функционалност основног алгоритма за кластеровање саобраћајних незгода на реалним подацима.

3.4.1 Алати

Прототип система базиран је на приступу представљеном у секцији 3.3. Графички прикази просторних података и процена површина кластера изведени су применом компоненте ArcMap софтверског пакета ESRI ArcGis (<https://www.esri.com>).

3.4.2 Вредност просторног прага

Просторни праг τ који је уведен у секцији 3.2. представља улазни параметар алгоритма за кластеровање. За потребе илустровања функционалности алгоритма, вредност прага τ износи 200 метара из следећег разлога. Дозвољено ограничење на путевима у урбаним срединама према Закону о безбедности саобраћаја Републике Србије износи 50 километара на час (National Assembly of the Republic of Serbia, 2022). Да бисмо у обзир узели однос између предвиђене и стварне брзине у урбаним срединама, посматраћемо један од минималних прекршаја, тј., прекорачење брзине до 20 километара на час (National Assembly of the Republic of Serbia, чл. 333, 2022). Стога, полазна претпоставка је да возач управља својим возилом брзином од 70 километара на час, а вредност просторног прага у овом примеру дефинишемо као раздаљину које возило пређе за временски интервал од 10 секунди (тј., приближно 200 метара).

Иако је просторном прагу τ додељена одређена вредност за потребе овог примера, битно је напоменути да је он дефинисан као улазни параметар. Генерално, вредност просторног прага одређује се у односу на спољне критеријуме, који варирају у зависности од контекста примене алгоритма. Истраживачко питање аутоматског одређивања прага τ размотрено је у наредном поглављу.

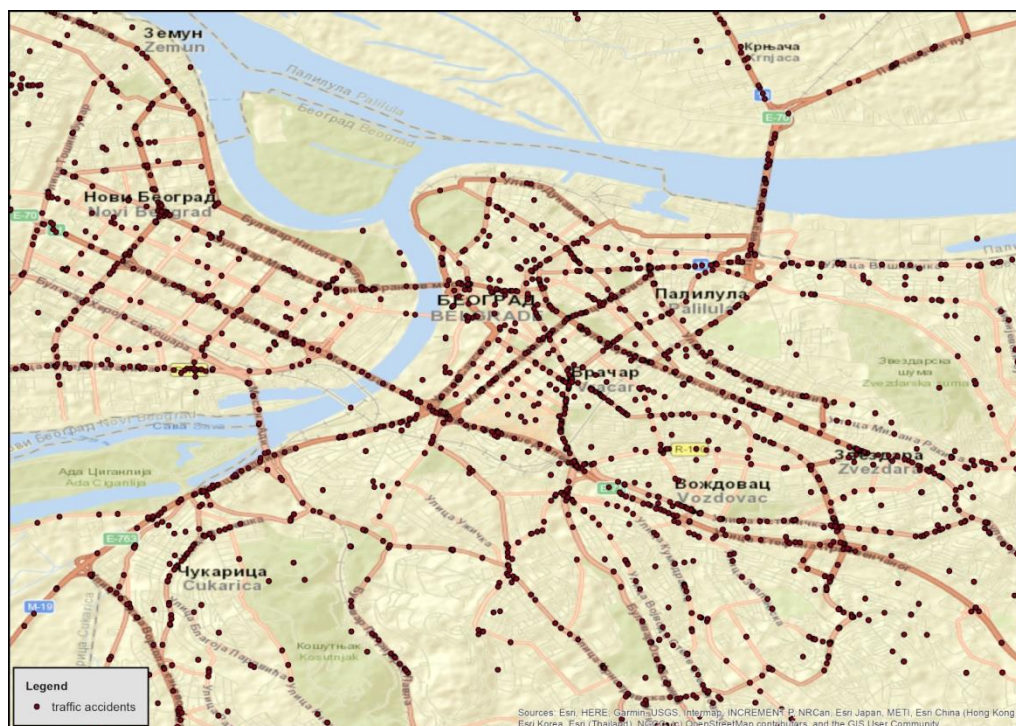
3.4.3 Подаци

За потребе илустровања функционалности основног алгоритма за кластеровање примењени су јавно доступни подаци о саобраћајним незгодама Министарства унутрашњих послова Републике Србије. Посматрани подскуп података садржи информације о 15366 саобраћајних незгода које су се десиле на територији града Београда у периоду од јануара 2021. године до децембра 2021. године (Republic of Serbia, 2021). Саобраћајне незгоде из овог подскупа података могу се поделити у три групе:

- 11294 саобраћајних незгода са материјалном штетом,
- 3996 саобраћајних незгода са повређеним лицима,
- 76 саобраћајних незгода са погинулим лицима.

Алгоритам за кластеровање је демонстриран на тзв. критичним саобраћајним незгодама, у које убрајамо последње две групе, тј., 4072 (3996 + 76) незгода са повређеним или погинулим лицима. За сваку саобраћајну незгоду, прототип система узима у обзир само идентификатор незгоде и њене геопросторне координате. Мапа на којој је приказан подскуп саобраћајних незгода са повређеним или погинулим лицима које су се догодиле у централним општинама Београда током 2021. године приказана је на сл. 3.1.

Поред тога, да би се илустровала стабилност резултата у времену, алгоритам се додатно примењује на податке о критичним саобраћајним незгодама које су се догодиле на територији општине Звездара током трогодишњег периода од јануара 2019. до децембра 2021. године (Republic of Serbia 2021, 2020, 2019).



Слика 3.1 Приказ саобраћајних незгода са повређеним или погинулим лицима у централним општинама Београда у периоду од јануара 2021. до децембра 2021. Слика је генерисана применом компоненте ArcMap софтверског пакета ESRI ArcGis (<https://www.esri.com>) и преузета из (Gnjatović et al. 2022).

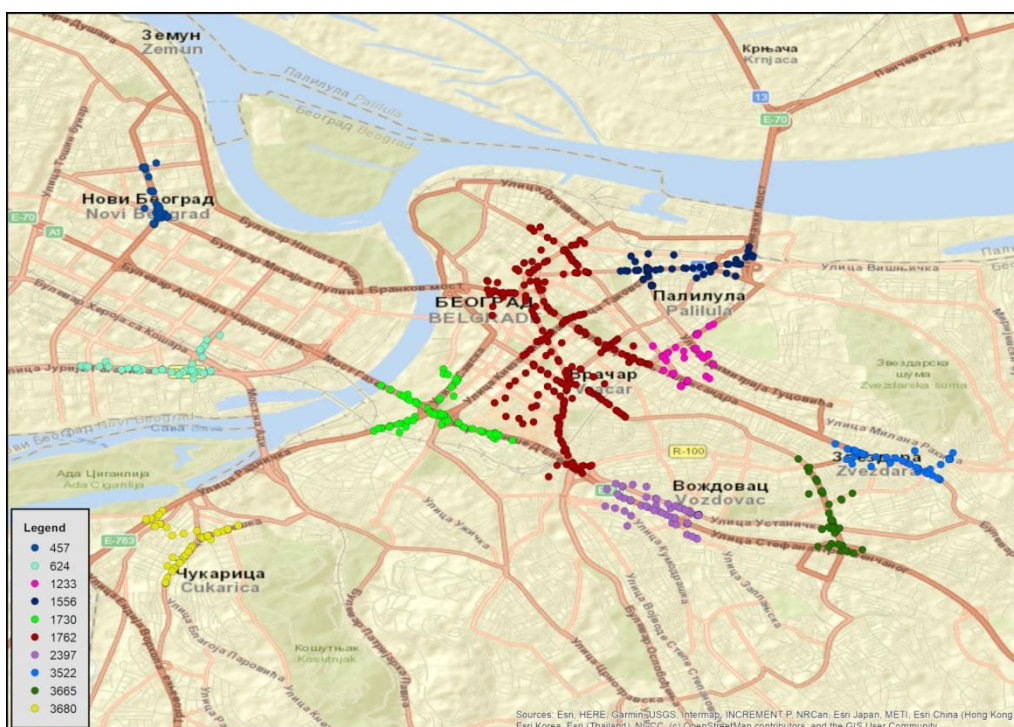
3.4.4 Извршавање алгоритма

У првој фази предложеног алгоритма, 4072 саобраћајних незгода подељено је у 1439 кластера. Просечан број саобраћајних незгода по кластеру износи 2,796, са стандардном девијацијом 8,909. У другој фази алгоритма селектовано је само десет доминантних кластера. Просечан број саобраћајних незгода по кластеру износи 73,3, са стандардном девијацијом 69,103 (в. табелу 3.1).

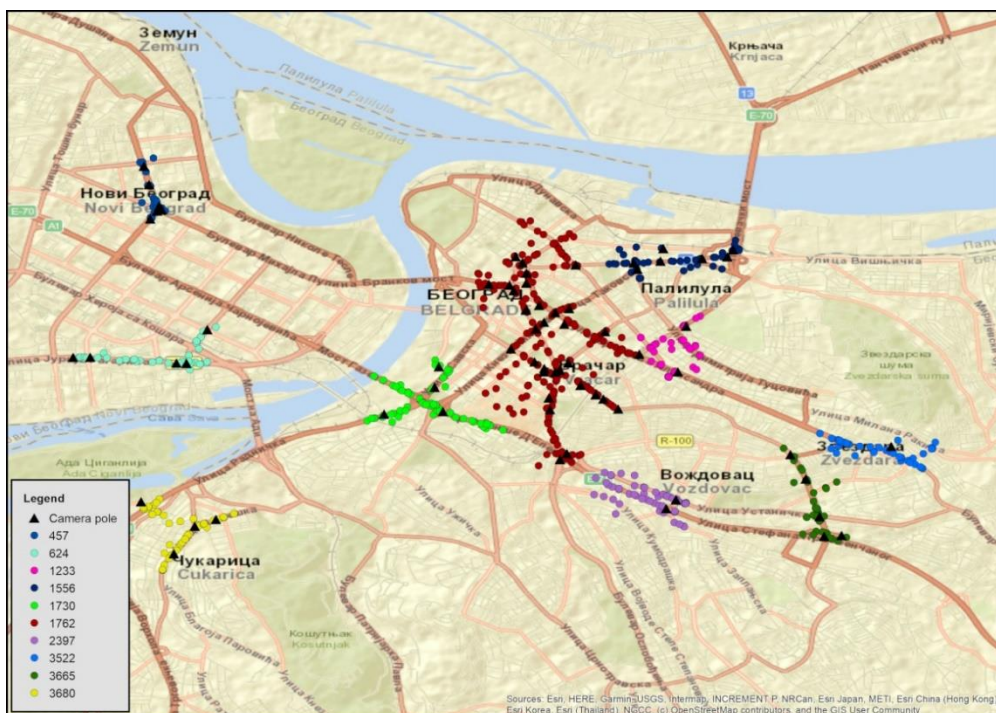
Мапа са селектованим кластерима приказана је на сл. 3.2. Иако приказује само централне општине Београда, ова мапа садржи свих десет издвојених кластера. Бројеви критичних саобраћајних незгода по кластерима приказани су у другом реду табеле 3.2. Идентификациони бројеви кластера приказани у табели одговарају идентификационим бројевима датим у легендама на сл. 3.2 и 3.3.

Табела 3.1: Резиме резултата кластеровања и селектовања доминантних кластера добијених применом основног алгоритма на јавно доступне податке о саобраћајним незгодама са повређеним или погинулим лицима које су се догодиле током 2021. године у Београду.

	Прва фаза (кластеровање)	Друга фаза (селектовање)
# саобраћајних незгода:	4072	733
# кластера:	1439	10
Просечан # саобраћајних незгода по кластеру:	2,796	73,3
Стандардна девијација:	8,909	69,103



Слика 3.2 Приказ селектованих кластера у Београду током 2021. године. Слика је генерисана применом компоненте ArcMap софтверског пакета ESRI ArcGis (<https://www.esri.com>) и преузета из (Gnjatović et al. 2022).



Слика 3.3 Приказ селектованих кластера у Београду током 2021. године и локација камерних стубова у Београду у марту 2022. Слика је генерисана применом компоненте ArcMap софтверског пакета ESRI ArcGis (<https://www.esri.com>) и преузета из (Gnjatović et al. 2022).

Табела 3.2: Опис резултата кластеровања и селектовања доминантних кластера добијених применом основног алгоритма на јавно доступне податке о саобраћајним незгодама са повређеним или погинулим лицима које су се догодиле током 2021. године у Београду. Приказани идентификациони бројеви кластера одговарају идентификационим бројевима датим у легендама на сл. 3.2 и 3.3.

Идентификатор кластера:	1762	1730	2397	1556	3680	457	1233	3665	3522	624
# саобраћајних незгода:	279	66	62	58	53	45	44	44	43	39
Површина кластера (км²):	1,75	0,30	0,33	0,22	0,20	0,06	0,23	0,15	0,13	0,29
# камерних стубова:	33	4	2	7	4	6	3	5	1	5

Скуп мера које се често примењују за процењивање резултата кластеровања саобраћајних незгода укључује коефицијент силуете (Rousseeuw 1987), индекс Калински-Харабаш (Calinski & Harabasz 1974), индекс Дејвис-Боулдин (Davies & Bouldin 1979) итд. Међутим, ове мере су опште и независне од контекста примене алгоритма за кластеровање. За разлику од њих, у овом поглављу се примењује квалитативни процес

који је заснован на критеријуму везаном за саобраћај, уз напомену да је доменски-специфичан критеријум за процену резултата кластеровања детаљније размотрен у наредном поглављу.

Резултати прелиминарне анализе изложени у овом поглављу могу се сматрати задовољавајућим: десет селектованих кластера покрива приближно само 0,11 процената укупне површине града (тј., 3,65 км² од приближно 3233 км², в. табелу 3.2) обухватајући притом чак 18 процента свих критичних саобраћајних незгода (тј., 733 од 4072, в. табелу 3.1). У циљу даље илустрације, упоредићемо резултате кластеровања са локацијама стубова за саобраћајне и друге камере који су изведени из јавно доступних информација Министарства унутрашњих послова (Ministry of Interior 2022). Притом, важно је напоменути следеће:

- Локације камерних стубова одређене су од стране трећег лица, независно од овог истраживања.
- Представљени алгоритам је агностичан у односу на локације камерних стубова тј., ове локације нису узимане у обзир приликом кластеровања.
- Подаци о саобраћајним незгодама који се користе за генерисање кластера прикупљени су током 2021. године. У том периоду, камере на посматраним локацијама још увек нису биле званично активирани, тј., нису имале утицај на понашање возача у посматраном периоду.

У складу с овим, скуп локација камерних стубова може се посматрати као индиректна објашњавајућа променљива у датом контексту.

Од 464 камерна места која се налазе у Београду, седамдесет се налази у оквиру селектованих кластера. Бројеви камерних стубова по кластерима приказани су у табели 3.2, а сл. 3.3 приказује селектоване кластере у Београду током 2021. године и локације камерних стубова у Београду у марту 2022. Десет селектованих кластера покрива 0,11 процената површине града и обухвата 15 процената камерних стубова.

3.4.5 Стабилност резултата у времену

Да би се илустровала стабилност резултата кластеровања у времену, представљени алгоритам је примењен на подацима о критичним саобраћајним незгодама које су се десиле у једној географској области у различитим периодима. Претходна секција показује резултате анализе спроведене на подацима који се односе на површину целог града Београда (тј., 3233 км²) и период од јануара 2021. до децембра 2021. У овој секцији, алгоритам је примењен само на податке о критичним саобраћајним незгодама које су се десиле на територији општине Звездара. Површина ове општине износи 31,11км², што чини 0,96 процената укупне површине Београда.

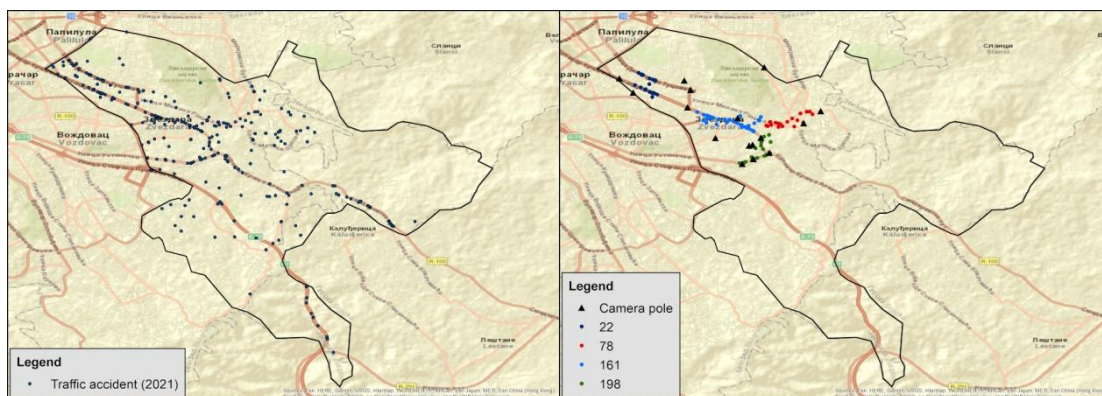
Представљени алгоритам, са вредношћу просторног прага $\tau=200\text{м}$, примењен је на јавно доступне податке о саобраћајним незгодама са погинулим или повређеним лицима које су се догодиле на општини Звездара у временском периоду од три године, тј., од јануара 2019. године до децембра 2021. године. У складу са циљем ове дисертације да се представи приступ намењен за примену у условима ограничених људских и техничких ресурса, општина Звездара је изабрана зато што представља једну од централних општина са најмање камерних места, тј., 16 од 464 камерних места у Београду. Приказ саобраћајних незгода које су се догодиле на овој општини током 2021., 2020. и 2019. године дат је на сл. 3.4 (а, в, д), респективно. Одговарајући прикази селектованих кластера по посматраним годинама дати су на сл. 3.4 (б, г, ђ), респективно. Локације камерних места (у складу са стањем из марта 2022. године) дате су због комплетности приказа. Резиме резултата кластеровања приказан је у табели 3.3. Селектовани кластери су описани у табели 3.4.

Табела 3.3 Резиме добијених резултата кластеровања након примене алгоритма на јавно доступне податке о саобраћајним незгодама са погинулим и повређеним лицима које су се догодиле на општини Звездара током 2019., 2020., 2021., респективно.

Година	Сумарни подаци	Прва фаза (кластеровање)	Друга фаза (селектовање)
2021	# саобраћајних незгода:	317	116
	# кластера:	97	4
	Просечан # незгода по кластеру:	3,186	29
	Стандардна девијација:	5,930	8,573
2020	# саобраћајних незгода:	282	101
	# кластера:	95	5
	Просечан # незгода по кластеру:	2,905	20,2
	Стандардна девијација:	5,020	11,214
2019	# саобраћајних незгода:	349	136
	# кластера:	93	5
	Просечан # незгода по кластеру:	3,699	27,2
	Стандардна девијација:	6,561	11,444

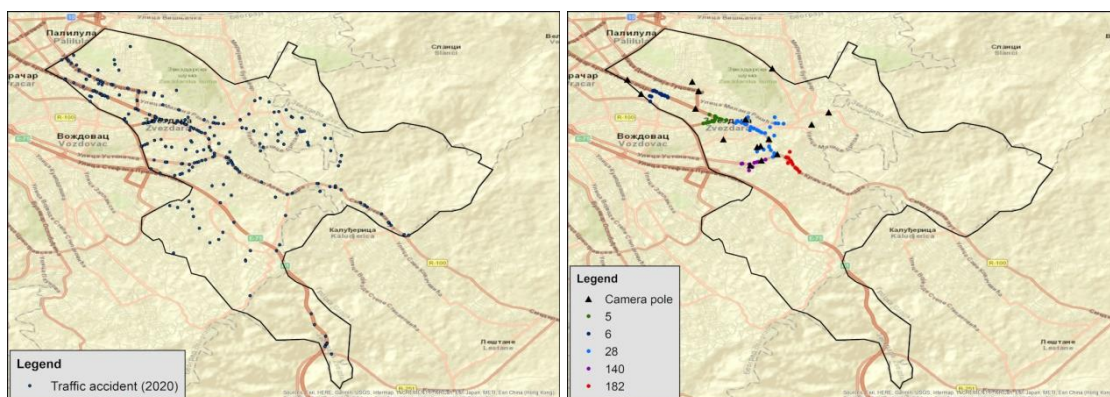
Табела 3.4 Опис селектованих кластера добијених применом алгоритма на јавно доступне податке о саобраћајним незгодама са погинулим и повређеним лицима које су се догодиле на општини Звездара током 2019., 2020. и 2021., респективно.

Година	Сумарни подаци	Селектовани кластери				
2021	Идентификатор кластера:	161	22	198	78	
	# саобраћајних незгода:	43	28	25	20	
	Површина кластера [км ²]:	0,132	0,193	0,171	0,189	
2020	Идентификатор кластера:	28	5	6	140	182
	# саобраћајних незгода:	42	19	16	12	12
	Површина кластера [км ²]:	0,272	0,073	0,065	0,121	0,109
2019	Идентификатор кластера:	194	72	192	116	80
	# саобраћајних незгода:	50	23	22	21	20
	Површина кластера [км ²]:	0,486	0,235	0,317	0,163	0,145



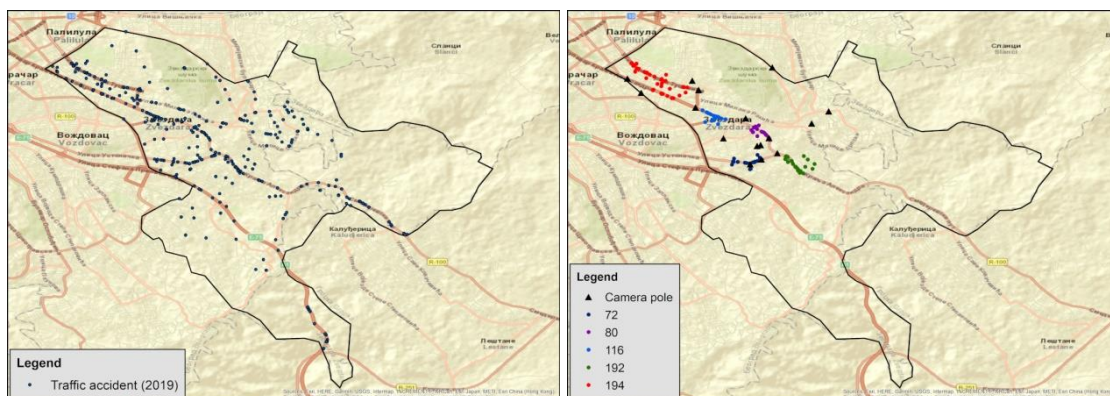
(a)

(б)



(в)

(г)



(д)

(ђ)

Слика 3.4 На левој страни: Приказ саобраћајних незгода са повређеним или погинулим лицима које су се догодиле на општини Звездара током (а) 2021., (в) 2020. и (д) 2019. године, респективно. На десној страни: Приказ селектованих кластера на основу анализе података из (б) 2021., (г) 2020. и (ђ) 2019. године, респективно. Локације камерних места (у складу са стањем из марта 2022. године) дате су због комплетности приказа. Слика је генерисана применом компоненте ArcMap софтверског пакета ESRI ArcGis (<https://www.esri.com>) и преузета из (Gnjatović et al. 2022).

Стабилност резултата у времену може се илустровати у односу на два аспекта:

- удео укупног броја критичних саобраћајних незгода на општини Звездара које припадају селектованим кластерима,
- површину преклапања кластера засебно селектованих у посматраним годинама.

У односу на први аспект, може се уочити следеће:

- У 2021., четири селектована кластера покривају приближно 2,2 процента површине општине (тј., $0,685\text{km}^2$ од $31,11\text{km}^2$) и обухватају 36,59 процента свих критичних саобраћајних незгода на општини (тј., 116 од 317).
- У 2020., пет селектованих кластера покривају приближно 2,06 процента површине општине (тј., $0,64\text{km}^2$ од $31,11\text{km}^2$) и обухватају 35,82 процента свих критичних саобраћајних незгода на општини (тј., 101 од 282).
- У 2019., пет селектованих кластера покрива приближно 4,33 процента површине општине (тј., $1,346\text{km}^2$ од $31,11\text{km}^2$) и обухватају 38,97 процента свих критичних саобраћајних незгода на општини (тј., 136 од 349).

Може се приметити да је удео критичних саобраћајних незгода које припадају селектованим кластерима стабилан у посматраном трогодишњем периоду (тј., 36,59, 35,82 и 38,97 процената, респективно).

У односу на други аспект, може се уочити значајно преклапање између селектованих кластера у све три године. Површина преклапања је $0,353\text{ km}^2$, што чини 51,53 процената селектоване површине у 2021. години, 55,15 процената селектоване површине у 2020. години и 26,23 процената селектоване површине у 2019. години.

3.5 Закључак

У овом поглављу представљен је двостепени приступ кластеровану саобраћајних незгода и илустрован на реалним подацима. За потребе илустровања алгоритма, просторном прагу τ , који представља улазни параметар алгоритма за кластероване, додељена је одређена вредност. У наредном поглављу биће размотрено истраживачко питање аутоматског одређивања вредности овог прага и уведен доменски-специфичан критеријум за процену резултата кластероване.

4. Доменски–специфичан критеријум за процену резултата кластеровања

У овом поглављу биће представљен доменски–специфичан критеријум за процену резултата кластеровања саобраћајних незгода, који ће потом бити примењен за аутоматско израчунавање вредности овог прага τ , тј., улазног параметра алгоритма за кластеровање представљеног у претходном поглављу (Кошанин et al, 2023).

Поглавље је структурисано на следећи начин: Секција 4.1 уводи меру за процену резултата кластеровања саобраћајних незгода. Секција 4.2 представља адаптацију алгоритма за одређивање колена криве за аутоматско одређивање вредности улазног параметра τ . Секција 4.3 представља резултате примене предложеног приступа на реалним подацима.

4.1 Мера за процену резултата кластеровања саобраћајних незгода

У овој секцији биће уведена мера за процену резултата кластеровања саобраћајних незгода заснована на три одвојене, али повезане мере:

- стабилност резултата кластеровања у времену,
- просторна колокација незгода у различитим периодима,
- релативна површина селектованих кластера.

4.1.1 Стабилност резултата кластеровања у времену

Процена стабилности резултата кластеровања у времену врши се након примене представљеног алгоритма за кластеровање на податке о саобраћајним незгодама које су прикупљене на истом географском подручју током два различита и непреклапајућа периода, које означавамо P_1 и P_2 , где P_1 претходи P_2 .

Полазна претпоставка је да град има n општина, представљених следећим вектором:

$$M = m_1, m_2, \dots, m_n. \quad (4.1)$$

Алгоритам кластеровања се примењује засебно за сваку општину. За дату вредност прага τ_j и општину m_i , извршавају се следећи кораци:

1. Основни алгоритам кластеровања примењује се на податке о саобраћајним незгодама које су се догодиле на општини m_i током периода P_1 .
2. Израчунају се удели (у процентима) свих саобраћајних незгода на општини m_i током периода P_1 и P_2 , респективно, које припадају кластерима селектованим у кораку 1. Ови удели означавају се са $s_1(m_i, \tau_j, P_1)$ и $s_2(m_i, \tau_j, P_2)$, респективно.

Пример 4.1. Усвојимо следећу поставку улазних параметара:

- општина (m): Звездара,
- вредност прага $\tau=170m$,
- период P_1 : од јануара 2019. године до децембра 2020. године,
- период P_2 : од јануара 2021. године до децембра 2021. године.

Резултати извршавања алгоритма представљеног у претходном поглављу за усвојену поставку улазних параметара могу се сажети на следећи начин:

1. Током периода P_1 , 631 саобраћајних незгода са повређеним или погинулим лицима догодила се на општини Звездара. Сл. 4.1(а) представља мапу ових саобраћајних незгода. Применом алгоритма за кластеровање на овом скупу података добијена су четири кластера, представљена на сл. 4.1(б).
2. Четири селектована кластера садрже 257 саобраћајних незгода. Удео критичних саобраћајних незгода на посматраној општини током периода P_1 које припадају селектованим кластерима износи:

$$s_1(m, \tau, P_1) = \frac{257}{631} = 40,729\%. \quad (4.2)$$

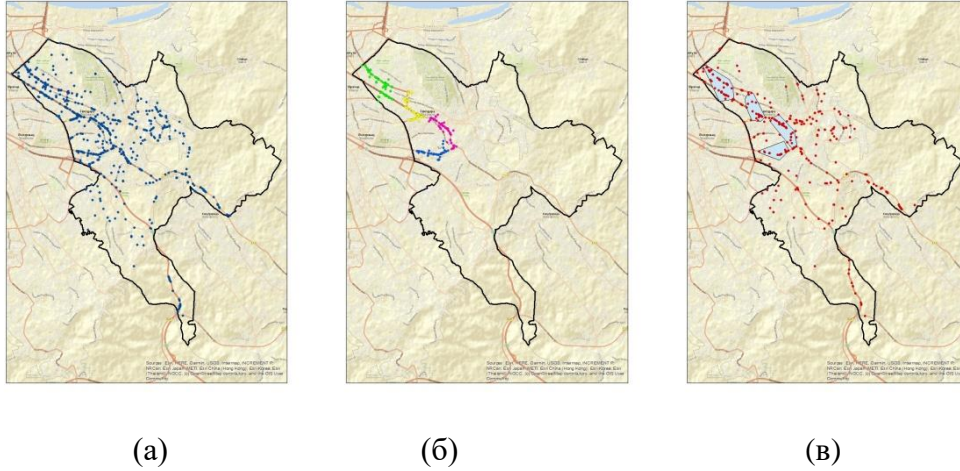
3. Током периода P_2 , 317 саобраћајних незгода са повређеним или погинулим лицима догодило се на општини Звездара. Слика 4.1(в) представља мапу ових саобраћајних незгода. Додатно, сваки од кластера приказаних на сл. 4.1(б) представљен је на сл. 4.1(в) минималним конвексним обухватајућим многоуглом. Број саобраћајних незгода које су се догодиле током овог периода, а припадају кластерима селектованим у кораку 1 јесте 131. Удео саобраћајних незгода износи:

$$s_2(m, \tau, P_2) = \frac{131}{317} = 41,325\%. \quad (4.3)$$

Када се претходно описани низ корака примени за све општине у скупу M , резултат се може представити са два вектора:

$$S_1(M, \tau_j, P_1) = s_1(m_1, \tau_j, P_1), s_1(m_2, \tau_j, P_1), \dots, s_1(m_n, \tau_j, P_1), \quad (4.4)$$

$$S_2(M, \tau_j, P_2) = s_2(m_1, \tau_j, P_2), s_2(m_2, \tau_j, P_2), \dots, s_2(m_n, \tau_j, P_2).$$



Слика 4.1 (а) Све саобраћајне незгоде са повређеним и погинулим лицима које су се догодиле на општини Звездара током периода P_1 . (б) Четири селектована кластера. (в) Све саобраћајне незгоде са погинулим или повређеним лицима које су се догодиле на општини Звездара током периода P_2 . Додатно, сваки од кластера приказаних на сл. 4.1(б) представљен је на сл. 4.1(в) минималним конвексним обухватајућим многоуглом. Слика је генерисана применом компоненте ArcMap софтверског пакета ESRI ArcGis (<https://www.esri.com>) и преузета из (Košanin et al. 2023).

У општем случају, општине у граду могу се разликовати по површини, броју становника, густини саобраћаја и другим факторима. Међутим, у предложеном приступу процени стабилности резултата кластеровања у времену, све општине се третирају као подједнако важне. Стога, за дату вредност прага τ_j стабилност резултата кластеровања у времену процењује се као косинусна сличност између вектора датих у (4.4):

$$s(M, \tau_j, P_1, P_2) = \frac{\sum_{k=1}^n (s_1(m_k, \tau_j, P_1) \cdot s_2(m_k, \tau_j, P_2))}{\sqrt{\sum_{k=1}^n s_1^2(m_k, \tau_j, P_1) \cdot \sum_{k=1}^n s_2^2(m_k, \tau_j, P_2)}}. \quad (4.5)$$

Пошто су сви елементи вектора у (4.4) позитивни, вредност $s(M, \tau_j, P_1, P_2)$ увек припада опсегу $[0,1]$, где вредност 1 представља максималну стабилност (τ_j , сличност између вектора), а вредност 0 представља минималну стабилност.

Табела 4.1 Процењивање стабилности резултата кластеровања у времену. Алгоритам се примењује на податке о саобраћајним незгодама са повређеним или погинулим лицима у Београду током периода P_1 и P_2 , за вредност прага $\tau=170$ м. Сви децимални бројеви у табели заокружени су на три децимална места.

Период P_1 [2019 ~ 2020]				Период P_2 [2021]		
Општина (m)	# незгода	# селектованих незгода	$s_1(m, \tau, P_1)$	# незгода	# селектованих незгода	$s_2(m, \tau, P_2)$
Барајево	123	54	43,902	51	21	41,176
Гроцка	295	105	35,593	150	40	26,667
Лазаревац	281	98	34,875	123	31	25,203
Младеновац	206	59	28,641	113	30	26,549
Нови Београд	1126	401	35,613	537	186	34,637
Обреновац	376	129	34,309	220	59	26,818
Палилула	962	308	32,017	384	132	34,375
Раковица	297	105	35,354	141	44	31,206
Савски венац	572	311	54,371	270	155	57,407
Сопот	100	28	28,000	49	10	20,408
Стари град	316	263	83,228	176	145	82,386
Сурчин	231	86	37,229	125	30	24,000
Вождовац	879	324	36,860	447	177	39,597
Врачар	412	272	66,019	182	139	76,374
Земун	800	257	32,125	393	126	32,061
Звездара	631	257	40,729	317	131	41,325
Чукарица	797	258	32,371	394	132	33,503
Укупно	8404	3315	39,446	4072	1588	38,998

Пример 4.2. Поставку улазних параметара усвојену у Примеру 4.1, можемо да применимо да бисмо проценили стабилност резултата кластеровања за све општине у Београду. Добијени резултати представљени су у табели 4.1. Елементи вектора $S_1(M, \tau_j, P_1)$ и $S_2(M, \tau_j, P_2)$ дефинисаних у (4.4) приказани су у четвртој и седмој колони табеле. У складу са (4.5), стабилност резултата кластеровања за усвојене параметре износи:

$$s(M, \tau, P_1, P_2) = 0,990. \quad (4.6)$$

4.1.2 Просторна колокација незгода у различитим периодима

Просторна колокација незгода у периодима P_1 и P_2 на нивоу града M дефинише се као удео (тј., проценат) свих саобраћајних незгода које су се догодиле у граду M током периода P_2 , а припадају подручјима покривеним кластерима добијеним кад се алгоритам за кластеровање примени на скуп свих саобраћајних незгода које су се догодиле у M током периода P_1 . Ову меру означавамо: $c(M, \tau, P_1, P_2)$.

Пример 4.3. На основу последњег реда табеле 4.1, може се увидети следеће:

- Укупан број саобраћајних незгода са повређеним или погинулим лицима током периода P_2 износи 4072.
- Број саобраћајних незгода које су се догодиле током периода P_2 , а припадају подручјима покривеним кластерима добијеним кад се алгоритам за кластеровање примени на скуп свих саобраћајних незгода са повређеним или погинулим лицима које су се догодиле током периода P_1 износи 1588.

Просторна колокација незгода у периодима P_1 и P_2 на нивоу Београда износи:

$$c(M, \tau, P_1, P_2) = \frac{1588}{4072} = 38,998\%. \quad (4.7)$$

4.1.3 Релативна величина кластера

У представљеном приступу, површина покривена кластером саобраћајних незгода концептуализована је као површина минималног конвексног обухватајућег многоугла (в. слику 4.1(в)). У складу са овом концептуализацијом, релативна величина селектованих кластера дефинише се као удео површине града M покривене кластерима који се добију када се основни алгоритам за кластеровање примени за сваку општину града засебно на скупове саобраћајних незгода које су се догодиле током периода P_1 . Релативну величину селектованих кластера означавамо $r(M, \tau, P_1)$.

Пример 4.4 Усвајајући исту поставку улазних параметара као у Примеру 4.2, табела 4.2 приказује следеће податке за сваку општину: број селектованих кластера, површину коју заузимају селектовани кластери, површину општине и релативну величину селектованих кластера на нивоу општине. Релативна величина селектованих кластера на нивоу града може се израчунати на основу последњег реда у табели 4.2:

$$r(M, \tau, P_1) = \frac{26,502}{3231,469} = 0,820\%. \quad (4.8)$$

Табела 4.2 Релативна величина свих селектованих кластера у Београду током периода P_1 за вредност прага $\tau=170m$. Сви децимални бројеви у табели заокружени су на три децимална места.

Општина	# селектованих кластера	Површина селектованих кластера [км ²]	Површина општине [км ²]	Релативна величина кластера [%]
Барајево	19	0,011	212,831	0,005
Гроцка	22	0,072	299,349	0,024
Лазаревац	21	0,123	382,540	0,032
Младеновац	9	0,151	338,764	0,045
Нови Београд	1	7,057	40,756	17,316
Обреновац	3	0,644	409,588	0,157
Палилула	2	4,563	450,351	1,013
Раковица	7	0,222	30,025	0,739
Савски венац	2	2,321	14,082	16,484
Сопот	11	0,003	270,506	0,001
Стари град	1	2,232	5,376	41,527
Сурчин	20	0,053	288,303	0,018
Вождовац	3	2,233	148,409	1,505
Врачар	1	2,424	2,911	83,256
Земун	8	1,500	149,682	1,002
Звездара	4	1,612	31,087	5,186
Чукарица	6	1,280	156,909	0,815
Укупно	140	26,502	3231,469	0,820

4.1.4 Интегрисана мера за процену резултата кластеровања

Основни алгоритам за кластеровање представљен у претходном поглављу намењен је за примену у условима ограничених људских и техничких ресурса. У складу са тим, интегрисану меру за процену резултата кластеровања дефинишемо тако да промовише стабилност резултата у времену и просторну колокацију незгода у различитим периодима,

а санкционише величину селектованих кластера. За дати град M , вредност прага τ_j и периоде P_1 и P_2 , интегрисана мера дефинише се на следећи начин:

$$\eta(M, \tau_j, P_1, P_2) = \frac{s(M, \tau_j, P_1, P_2) \cdot c(M, \tau_j, P_1, P_2)}{r(M, \tau_j, P_1, P_2)}, \quad (4.9)$$

где су:

- $s(M, \tau_j, P_1, P_2)$ - стабилност резултата кластерованја у времену,
- $c(M, \tau_j, P_1, P_2)$ - просторна коалиција незгода у различитим периодима,
- $r(M, \tau_j, P_1, P_2)$ - релативна величина селектованих кластера.

Пример 4.5 Узимајући у обзир (4.6)-(4.8), вредност интегрисане мере за процену резултата кластерованја добијене за усвојену поставку параметара износи:

$$\eta(M, \tau, P_1, P_2) = 47,082. \quad (4.10)$$

4.2 Аутоматско израчунавање вредности просторног прага

За дате периоде P_1 и P_2 и град M , интегрисана мера за процену резултата кластерованја (4.9) може се представити као функција са једним улазним параметром који представља вредност просторног прага, тј., $\eta(\tau)$. Ова редукција, омогућава вишеструку примену алгоритма за кластерованје, представљеног у претходном поглављу, над саобраћајним незгодама које су се догодиле у граду M , током периода P_1 и P_2 , варирајући вредност прага. У овој секцији представљен је алгоритам за аутоматско израчунавање оперативне вредности просторног прага заснован на интегрисаној мери (4.9).

Оперативна вредност прага индикована је коленом криве која представља зависност вредности интегрисане мере $\eta(\tau)$ од примењене вредности прага τ . У наставку је представљен алгоритам за детектовање колена криве. Нека је \hat{D} скуп података који садржи n опсервација за које треба открити колено:

$$\hat{D} = \{(\hat{\tau}_i, \hat{\eta}_i) \mid 1 \leq i \leq n \wedge \hat{\tau}_i \geq 0 \wedge \hat{\eta}_i \geq 0\}, \quad (4.11)$$

где $\hat{\tau}_i$ представља вредност просторног прага, $\hat{\eta}_i$ представља вредност интегрисане мере за процену rezultata кластеровања добијених за вредност прага $\hat{\tau}_i$, при чему су вредности прага равномерно распоређене, тј.:

$$(\exists t \in \mathbf{R}, t > 0)(\forall 1 \leq i < n)(\hat{\tau}_{i+1} - \hat{\tau}_i = t). \quad (4.12)$$

Вредност $\hat{\tau}_i$ и $\hat{\eta}_i$ се прво нормализују на опсег $[0, 1]$, без промене њихових расподела (Satorää et al., 2011), тј.:

$$\bar{D} = \{(\bar{\tau}_i, \bar{\eta}_i) \mid \bar{\tau}_i = \frac{\hat{\tau}_i - \hat{\tau}_{\min}}{\hat{\tau}_{\max} - \hat{\tau}_{\min}} \wedge \bar{\eta}_i = \frac{\hat{\eta}_i - \hat{\eta}_{\min}}{\hat{\eta}_{\max} - \hat{\eta}_{\min}} \wedge (\hat{\tau}_i, \hat{\eta}_i) \in \hat{D}\}, \quad (4.13)$$

где су:

$$\hat{\tau}_{\min} = \min_{1 \leq i \leq n} \hat{\tau}_i, \quad \hat{\eta}_{\min} = \min_{1 \leq i \leq n} \hat{\eta}_i, \quad (4.14)$$

$$\hat{\tau}_{\max} = \max_{1 \leq i \leq n} \hat{\tau}_i, \quad \hat{\eta}_{\max} = \max_{1 \leq i \leq n} \hat{\eta}_i. \quad (4.15)$$

Да би се изабрали кандидати за колено криве, посматрају се разлике између нормализованих тачака скупа (4.13) и линеарне функције $f(\tau) = 1 - \tau$, која представља главну дијагоналу јединичног квадрата на који је оригинални скуп (4.11) нормализован. Нови скуп података који обухвата посматране разлике изводи се на следећи начин:

$$D = \{(\tau_i, \eta_i) \mid \tau_i = \bar{\tau}_i \wedge \eta_i = 1 - \bar{\tau}_i - \bar{\eta}_i \wedge (\bar{\tau}_i, \bar{\eta}_i) \in \bar{D}\}. \quad (4.16)$$

Да би се одредило колено криве, идентификује се најконкавнија тачка (τ, η) на кривој која представља скуп D . Слично (Satorää et. all 2011), скуп кандидата за колено криве дефинише се као скуп који садржи тачке наглашене конкавности, тј., садржи локалне максимуме скупа D :

$$K_1 = \{(\tau_i, \eta_i) \mid 1 < i < n \wedge \eta_i > \eta_{i-1} \wedge \eta_i > \eta_{i+1} \wedge (\tau_i, \eta_i) \in D\}. \quad (4.17)$$

Ако скуп K_1 није празан, конкавност у датој тачки (τ_i, η_i) овог скупа процењује се као угао криве у тачки:

$$\gamma_1(\tau_i, \eta_i) = \arctan\left(\frac{\tau_i - \tau_{i-1}}{|\eta_i - \eta_{i-1}|}\right) + \arctan\left(\frac{\tau_{i+1} - \tau_i}{|\eta_{i+1} - \eta_i|}\right)$$

$$= \begin{cases} \arctan\left(\frac{\frac{\tau_i - \tau_{i-1}}{|\eta_i - \eta_{i-1}|} + \frac{\tau_{i+1} - \tau_i}{|\eta_{i+1} - \eta_i|}}{1 - \frac{\tau_i - \tau_{i-1}}{|\eta_i - \eta_{i-1}|} \cdot \frac{\tau_{i+1} - \tau_i}{|\eta_{i+1} - \eta_i|}}\right), & \text{ако } \frac{\tau_i - \tau_{i-1}}{|\eta_i - \eta_{i-1}|} \cdot \frac{\tau_{i+1} - \tau_i}{|\eta_{i+1} - \eta_i|} < 1, \\ \arctan\left(\frac{\frac{\tau_i - \tau_{i-1}}{|\eta_i - \eta_{i-1}|} + \frac{\tau_{i+1} - \tau_i}{|\eta_{i+1} - \eta_i|}}{1 - \frac{\tau_i - \tau_{i-1}}{|\eta_i - \eta_{i-1}|} \cdot \frac{\tau_{i+1} - \tau_i}{|\eta_{i+1} - \eta_i|}}\right) + \pi, & \text{ако } \frac{\tau_i - \tau_{i-1}}{|\eta_i - \eta_{i-1}|} \cdot \frac{\tau_{i+1} - \tau_i}{|\eta_{i+1} - \eta_i|} > 1, \\ \frac{\pi}{2}, & \text{у супротном,} \end{cases} \quad (4.18)$$

што је представљено на сл. 4.2(a). Најконкавнија тачка у скупу K_1 бира се минимизовањем процењеног угла:

$$(\tau^*, \eta^*) = \operatorname{argmin}_{(\tau_i, \eta_i) \in K_1} \gamma_1(\tau_i, \eta_i). \quad (4.19)$$

У супротном, ако је скуп K_1 празан (тј., ако је секвенца тачака у скупу D монотono нерастућа) услов конкавности (4.17) се релаксира. У овом случају, скуп кандидата за колена криве чине све конкавне тачке (тј., не само тачке наглашене конкавности) на кривој представљеној скупом D .

Имајући у виду да је функција $\eta(\tau)$ дискретна, при чему су вредности τ_i равномерно распоређене (в. (4.12)), њен други извод може се представити на следећи начин:

$$\begin{aligned} \frac{\partial^2 \eta(\tau = \tau_i)}{\partial \tau^2} &= \frac{\partial\left(\frac{\partial \eta(\tau = \tau_i)}{\partial \tau}\right)}{\partial \tau} \\ &= \frac{\eta(\tau = \tau_{i+1}) - 2 \cdot \eta(\tau = \tau_i) + \eta(\tau = \tau_{i-1})}{t^2}. \end{aligned} \quad (4.20)$$

Скуп свих конкавних тачака може се формално представити као скуп тачака у којима је други извод мањи од нуле:

$$K_2 = \{(\tau_i, \eta_i) \mid 1 < i < n \wedge 2\eta_i > \eta_{i-1} + \eta_{i+1} \wedge (\tau_i, \eta_i) \in D\}, \quad (4.21)$$

што је у складу са условом примењеним у (Zhao et al. 2008). Конкавност у било којој тачки (τ_i, η_i) у скупу K_2 процењује се као угао у тој тачки (τ_i, η_i) :

$$\gamma_2(\tau_i, \eta_i) = \begin{cases} \arctan\left(\frac{\tau_i - \tau_{i-1}}{|\eta_i - \eta_{i-1}|}\right) - \arctan\left(\frac{\tau_{i+1} - \tau_i}{|\eta_{i+1} - \eta_i|}\right) + \pi, & \text{ако } \eta_{i-1} < \eta_i < \eta_{i+1}, \\ \arctan\left(\frac{\tau_{i+1} - \tau_i}{|\eta_{i+1} - \eta_i|}\right) - \arctan\left(\frac{\tau_i - \tau_{i-1}}{|\eta_i - \eta_{i-1}|}\right) + \pi, & \text{ако } \eta_{i-1} > \eta_i > \eta_{i+1}, \\ \arctan\left(\frac{\tau_i - \tau_{i-1}}{|\eta_i - \eta_{i-1}|}\right) + \frac{\pi}{2}, & \text{ако } \eta_{i-1} < \eta_i = \eta_{i+1}, \\ \arctan\left(\frac{\tau_{i+1} - \tau_i}{|\eta_{i+1} - \eta_i|}\right) + \frac{\pi}{2}, & \text{ако } \eta_{i-1} = \eta_i > \eta_{i+1}. \end{cases} \quad (4.22)$$

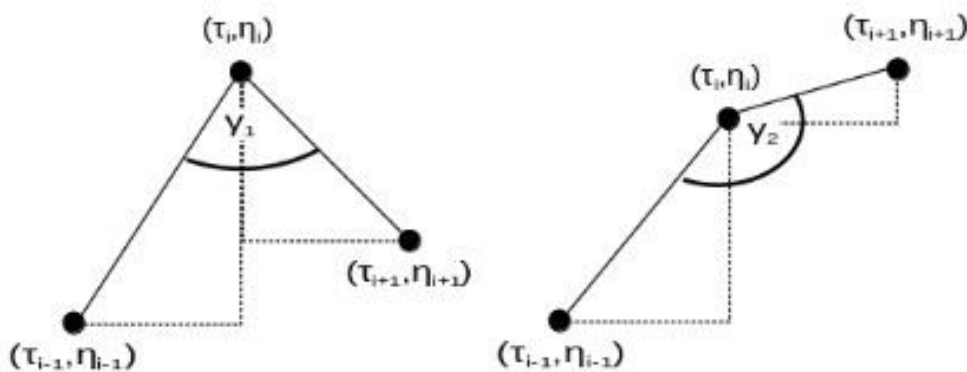
Први случај у (4.22) представљен је на сл. 4.2(б). Најконкавнија тачка у скупу K_2 је одређује се минимизовањем процењеног угла:

$$(\tau^*, \eta^*) = \operatorname{argmin}_{(\tau_i, \eta_i) \in K_2} \gamma_2(\tau_i, \eta_i). \quad (4.23)$$

Коначно, пошто је вредност τ^* - изведена из (4.19) или (4.23) – нормализована (в.(4.13)), оперативна вредност прага добија се на следећи начин:

$$\hat{\tau}^* = \hat{\tau}_{\min} + \tau^*(\hat{\tau}_{\max} - \hat{\tau}_{\min}). \quad (4.24)$$

Лако се показује да је скуп (4.21) увек надскуп скупа (4.17). Међутим, ова два услова разматрају се одвојено и у наведеном редоследу због разлога ефикасности. Алгоритам за аутоматско израчунавање вредности прага илустрован је у следећој секцији.



Слика 4.2. (а) Наглашена конкавност (тј., тачка локалног максимума) и (б) ненаглашена конкавност у скуп D . Слика је преузета из (Кошанин et al. 2023).

4.3 Резултати

Представљени приступ је примењен на скуп који садржи реалне податке о 18880 саобраћајних незгода са повређеним или погинулим лицима које су се десиле у трогодишњем периоду од јануара 2019. до децембра 2021. године у три највећа града у Републици Србији: Београду, Новом Саду и Нишу (в. табелу 4.3). У складу са (4.12), разматра се следећи низ вредности прага:

$$T \equiv 100m, 110m, 120m, \dots, 400m, \quad (4.25)$$

међу којима се бира оперативна вредност прага.

Табела 4.3. Бројеви саобраћајних незгода са погинулим и повређеним лицима које су се десиле у трогодишњем периоду од јануара 2019. до децембра 2021. године у три највећа града у Републици Србији: Београду, Новом Саду и Нишу.

Град	2019	2020	2021	Укупно
Београд	4684	3720	4072	12476
Нови Сад	1710	1464	1574	4748
Ниш	607	521	528	1656
Укупно	7001	5705	6174	18880

Вредности свих мера представљених у секцији 4.1, добијених кад се представљени алгоритам примени за град Београд у посматраном периоду и вредности просторног прага из секвенце (4.25) приказане су у табели 4.4. График нормализоване интегрисане мере η у односу на примењену нормализовану вредност прага τ приказан је на сл. 4.3(а). Додатно, изведене разлике између тачака нормализованог скупа података и главне дијагонале јединичног квадрата назначене су на сл. 4.3(б). Селектована вредност оперативног прага износи:

$$\hat{\tau}^*(M = \text{Београд}, P_1 = [2019 \sim 2020], P_2 = [2021]) = 150m. \quad (4.26)$$

Табела 4.4 Вредности мера представљених у секцији 4.1, добијених кад се представљени алгоритам примени на податке о саобраћајним незгодама које су се догодиле у Београду у периоду 2019-2021. и вредности просторног прага из секвенце (4.25). Сви децимални бројеви у табели заокружени су на три децимална места.

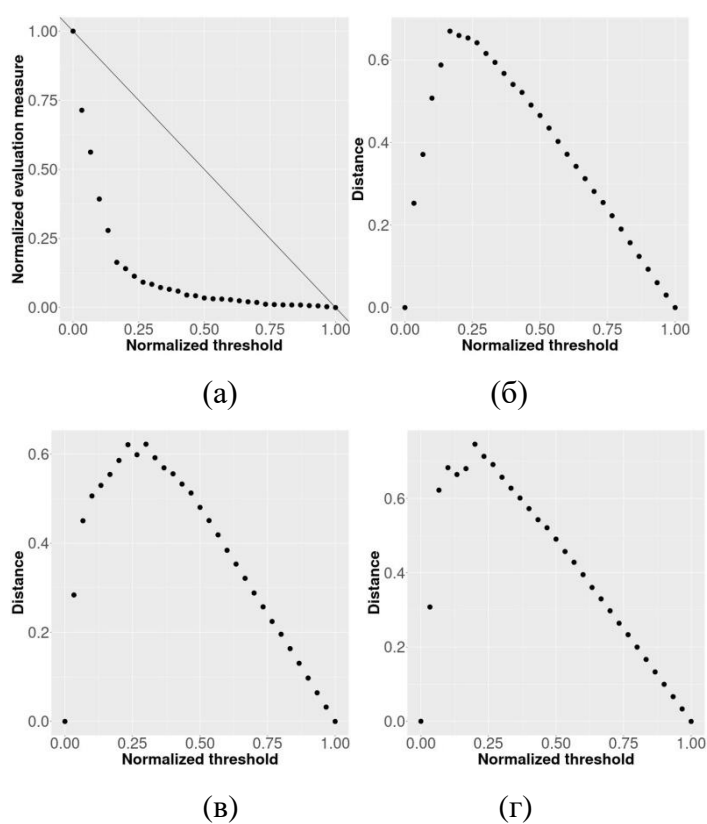
Вредност прага $\hat{\tau}$	Стабилност резултата кластеровања $s(\hat{\tau})$	Релативна величина селектованих кластера $r(\hat{\tau})$	Просторна колокација незгода $c(\hat{\tau})$	Интегрисана мера за кластеровање саобраћајних незгода $\hat{\eta}(\hat{\tau})$
100	0,980	0,001	0,297	302,465
110	0,982	0,001	0,286	220,035
120	0,975	0,002	0,327	176,388
130	0,980	0,002	0,319	127,509
140	0,981	0,004	0,368	94,753
150	0,978	0,005	0,329	61,536
160	0,987	0,007	0,363	54,953
170	0,990	0,008	0,390	47,082
180	0,991	0,010	0,416	40,827
190	0,992	0,011	0,440	38,685
200	0,993	0,013	0,454	35,308
210	0,994	0,014	0,466	33,419
220	0,993	0,015	0,461	31,504
230	0,994	0,017	0,477	27,453
240	0,995	0,019	0,510	26,685
250	0,996	0,022	0,532	24,345
260	0,996	0,024	0,569	23,539
270	0,996	0,025	0,591	23,275
280	0,997	0,026	0,587	22,600
290	0,997	0,028	0,594	21,453
300	0,997	0,029	0,589	20,412
310	0,997	0,030	0,601	19,732
320	0,998	0,035	0,629	17,872
330	0,998	0,036	0,637	17,489
340	0,998	0,037	0,641	17,186
350	0,998	0,037	0,642	17,118
360	0,999	0,038	0,647	17,047
370	0,999	0,040	0,661	16,406
380	0,999	0,041	0,667	16,178

390	0,999	0,045	0,682	15,263
400	0,999	0,048	0,690	14,371

Разлике између нормализованог сета тачака и дијагонале јединичног квадрата добијене кад се представљени алгоритам примени на податке о саобраћајним незгодама које су се догодиле током периода 2019-2021. у Новом Саду и Нишу представљене су на сл. 4.3(в-г), респективно. Селектоване вредности оперативних прага износе:

$$\hat{\tau}^*(M = \text{Нови Сад}, P_1 = [2019 \sim 2020], P_2 = [2021]) = 170\text{м}, \quad (4.27)$$

$$\hat{\tau}^*(M = \text{Ниш}, P_1 = [2019 \sim 2020], P_2 = [2021]) = 160\text{м}.$$



Слика 4.3 (а) График нормализоване интегрисане мере η у односу на примењену нормализовану вредност прага τ за Београд. (б–г) Графички приказ разлике између тачака нормализованог скупа података и главне дијагонале јединичног квадрата за Београд, Нови Сад и Ниш, респективно. Слика је преузета из (Кошанин et al. 2023).

Резултати добијени кад се основни алгоритам кластеровања примени на податке о саобраћајним незгодама у Београду, Новом Саду и Нишу, респективно, узимајући израчунате оперативне вредности прагова за сваки од ових градова, приказане су у табелама 4.5-4.7. Свака од ових табела садржи број селектованих кластера за сваку

општину, површину селектованих кластера, површину општине и релативне величине селектованих кластера у односу на површину општине, односно града.

Табела 4.5 Резултати кластеровања за Београд ($\tau^*=150\text{m}$). Сви децимални бројеви у табели заокружени су на три децимална места.

Општина	# селектованих кластера	Површина селектованих кластера [км ²]	Површина општине [км ²]	Релативна величина кластера [%]
Барајево	19	0,008	212,831	0,004
Гроцка	57	0,046	299,349	0,015
Лазаревац	21	0,085	382,540	0,022
Младеновац	14	0,050	338,764	0,015
Нови Београд	1	3,575	40,756	8,772
Обреновац	3	0,516	409,588	0,126
Палилула	2	4,258	450,351	0,945
Раковица	6	0,132	30,025	0,440
Савски венац	2	1,855	14,082	13,176
Сопот	11	0,003	270,506	0,001
Стари град	1	1,197	5,376	22,260
Сурчин	41	0,035	288,303	0,012
Вождовац	4	1,324	148,409	0,892
Врачар	2	1,171	2,911	40,240
Земун	8	0,653	149,682	0,436
Звездара	6	1,064	31,087	3,423
Чукарица	5	0,899	156,909	0,573
Укупно	203	16,873	3231,469	0,522

Табела 4.6 Резултати кластеровања за Нови Сад ($\tau^*=170m$). Сви децимални бројеви у табели заокружени су, кад је прикладно, на три децимална места.

Општина	# селектованих кластера	Површина селектованих кластера [км ²]	Површина општине [км ²]	Релативна величина кластера [%]
Бач	6	0,010	367,268	0,003
Бачка Паланка	19	0,052	589,496	0,009
Бачки Петровац	7	0,001	158,257	0,00003
Беочин	8	0,001	184,105	0,001
Бечеј	21	0,036	486,196	0,007
Нови Сад	2	2,523	698,816	0,361
Србобран	13	0,005	283,939	0,002
Сремски Карловац	8	0,001	50,538	0,002
Темерин	14	0,031	169,525	0,019
Тител	1	0,0001	260,600	0,00002
Врбас	10	0,131	375,326	0,035
Жабалъ	18	0,003	399,566	0,001
Укупно	127	2,794	4023,633	0,069

Табела 4.7 Резултати кластеровања за Ниш ($\tau^*=160m$). Сви децимални бројеви у табели заокружени су, кад је прикладно, на три децимална места.

Општина	# селектованих кластера	Површина селектованих кластера [км ²]	Површина општине [км ²]	Релативна величина кластера [%]
Алексинач	6	0,116	706,335	0,016
Дољевац	10	0,001	121,275	0,001
Гаџин Хан	1	0,00015	324,931	0,00005
Мерошина	2	0,00005	193,089	0,00002
Ниш	2	1,594	449,929	0,354
Нишка Бања	5	0,002	146,185	0,001
Ражањ	2	0,00003	288,512	0,00001
Сврљиг	6	0,001	496,894	0,00014
Укупно	34	1,714	2727,151	0,063

4.3.1 Детаљан приказ резултата за Београд

У овој секцији је дат детаљан приказ релевантних резултата добијених приликом примене предложеног приступа на податке о саобраћајним незгодама које су се десиле у посматраним периодима на територији Београда, примењујући претходно израчунати просторни праг ($\tau^*=150\text{m}$).

Табела 4.8 приказује расподелу критичних саобраћајних незгода по општинама, расподелу кластера по општинама, расподелу селектованих кластера и критичних саобраћајних незгода по општинама, вредности вектора дефинисаних у (4.4) и процену стабилности резултата кластерованја.

Табела 4.9 садржи збирне податке о апсолутним и релативним површинама селектованих кластера по општинама, док табеле 4.10-4.26 садрже површине појединачних кластера по општинама.

На сл. 4.4-4.20 приказани су селектовани кластери по општинама. Границе општина назначене су црвеном бојом, а границе кластера црном бојом. За сваки кластер је назначен и једнозначни идентификатор, који служи као референца у односу на табеле 4.10-4.26. Сл. 4.4-4.20 су генерисане применом компоненте ArcMap софтверског пакета ESRI ArcGis (<https://www.esri.com>).

Табела 4.8 Сумарни подаци резултата кластеровања саобраћајних незгода у Београду ($\tau^*=150m$).

Општина (m)	Период P ₁ [2019 ~ 2020]								Период P ₂ [2021]		Стабилност резултата	
	# нез-года	# кла-стера	Просечан # незгода по кластеру	Ст. дев.	# селектованих незгода	# селектованих кластера	Просечан # незгода по селектованом кластеру	Ст. дев.	# нез-года	# селектованих незгода	s ₁ (m, τ , P ₁)	s ₂ (m, τ , P ₂)
Барајево	123	86	1,395	0,944	53	19	2,789	1,239	51	20	43,089	39,216
Гроцка	295	182	1,604	1,540	167	57	2,930	2,239	150	49	56,610	32,667
Лазаревац	281	179	1,553	1,295	94	21	4,476	1,893	123	27	33,452	21,951
Младеновац	206	131	1,534	1,315	65	14	4,643	2,057	113	32	31,553	28,319
Нови Београд	1126	156	7,147	19,065	215	1	215,000	0,000	537	86	19,094	16,015
Обреновац	376	181	2,072	5,038	117	3	39,000	9,899	220	56	31,117	25,455
Палилула	962	245	3,857	14,746	293	2	146,500	60,500	384	130	30,457	33,854
Раковица	297	99	2,889	3,806	92	6	15,333	5,558	141	39	30,976	27,660
Савски венац	572	81	6,938	21,171	265	2	132,500	23,500	270	134	46,329	49,630
Сопот	100	83	1,205	0,576	28	11	2,545	0,656	49	10	28,000	20,408
Стари град	316	25	12,520	38,352	198	1	198,000	0,000	176	104	62,658	59,091
Сурчин	231	149	1,537	1,207	121	41	2,951	1,592	125	44	52,381	35,200
Вождовац	879	213	4,075	10,500	299	4	74,750	6,495	447	153	34,016	34,228
Врачар	412	26	15,500	31,559	216	2	108,000	48,000	182	114	52,427	62,637
Земун	800	256	3,070	4,857	200	8	25,000	10,124	393	111	25,000	28,244
Звездара	631	151	4,106	8,679	250	6	41,667	15,829	317	121	39,620	38,170
Чукарица	797	257	3,066	7,022	216	5	43,200	23,928	394	108	27,102	27,411
s(M, τ, P₁, P₂)											0,978	

Табела 4.9 Збир података о апсолутним и релевантним површинама селектованих кластера по Општинама у Београду.

Општина	Површина општине [км ²]	Површина кластера [км ²]	Релативна површина кластера [%]
Звездара	31,087	1,064	3,423
Раковица	30,025	0,132	0,440
Врачар	2,911	1,171	40,240
Вождовац	148,409	1,324	0,892
Палилула	450,351	4,258	0,945
Стари град	5,376	1,197	22,260
Савски венац	14,082	1,855	13,176
Нови Београд	40,756	3,575	8,772
Земун	149,682	0,653	0,436
Сурчин	288,303	0,035	0,012
Чукарица	156,909	0,899	0,573
Младеновац	338,764	0,050	0,015
Гроцка	299,349	0,046	0,015
Лазаревац	382,540	0,085	0,022
Сопот	270,506	0,003	0,001
Барајево	212,831	0,008	0,004
Обреновац	409,588	0,516	0,126
Укупно	3231,469	16,873	0,522

Табела 4.10 Површине појединачних кластера на општини Звездара.

Идентификатор кластера	Површина [км ²]
36	0,035793
78	0,062735
142	0,213148
288	0,344626
306	0,071178
387	0,336556
Укупно	1,064036

Табела 4.11 Површине појединачних кластера на општини Раковица.

Идентификатор кластера	Површина [км ²]
33	0,016736
53	0,034237
56	0,018399
60	0,009400
197	0,019373
253	0,034050
Укупно	0,132195

Табела 4.12 Површине појединачних кластера на општини Врачар.

Идентификатор кластера	Површина [км ²]
13	0,736576
108	0,434816
Укупно	1,171392

Табела 4.13 Површине појединачних кластера на општини Вождовац.

Идентификатор кластера	Површина [км ²]
67	0,455328
71	0,314715
250	0,397648
733	0,156752
Укупно	1,324443

Табела 4.14 Површине појединачних кластера на општини Палилула.

Идентификатор кластера	Површина [км ²]
367	3,797220
422	0,460432
Укупно	4,257652

Табела 4.15 Површина кластера на општини Стари град.

Идентификатор кластера	Површина [км ²]
148	1,196700
Укупно	1,196700

Табела 4.16 Површине појединачних кластера на општини Савски венац.

Идентификатор кластера	Површина [км ²]
18	1,265473
203	0,589960
Укупно	1,855433

Табела 4.17 Површина кластера на општини Нови Београд.

Идентификатор кластера	Површина [км ²]
638	3,575020
Укупно	3,575020

Табела 4.18 Површине појединачних кластера на општини Земун.

Идентификатор кластера	Површина [км ²]
161	0,045133
213	0,115715
250	0,038849
300	0,025057
316	0,161160
424	0,093228
458	0,052082
616	0,122030
Укупно	0,653254

Табела 4.19 Површине појединачних кластера на општини Сурчин.

Идентификатор кластера	Површина [км ²]
0	0,000034
1	0,000166
2	0,000049
4	0,000618
5	0,000110
6	0,000023
7	0,000033
8	0,000026
11	0,000138
12	0,000967
14	0,002441
15	0,000171
18	0,000186
24	0,000193
25	0,000721
34	0,000052
36	0,000046
40	0,000011
41	0,000065

42	0,018334
43	0,002847
44	0,000045
51	0,000711
63	0,000006
69	0,000184
72	0,003268
75	0,000131
78	0,001429
79	0,000052
95	0,000162
120	0,000468
125	0,000325
131	0,000002
136	0,000017
138	0,000075
159	0,000223
160	0,000069
161	0,000644
183	0,000008
188	0,000101
209	0,000167
Укупно	0,035323

Табела 4.20 Површине појединачних кластера на општини Чукарица.

Идентификатор кластера	Површина [км²]
42	0,089806
121	0,045507
143	0,676676
317	0,060198
344	0,026804
Укупно	0,898991

Табела 4.21 Површине појединачних кластера на општини Младеновац.

Идентификатор кластера	Површина [км ²]
2	0,000206
6	0,008771
10	0,000625
15	0,000389
26	0,002218
33	0,000081
56	0,000044
61	0,000316
75	0,002182
78	0,026083
87	0,000193
89	0,006387
118	0,001602
140	0,000424
Укупно	0,049522

Табела 4.22 Површине појединачних кластера на општини Гроцка.

Идентификатор кластера	Површина [км ²]
0	0,000008
1	0,000104
5	0,000119
7	0,000173
9	0,003410
10	0,000043
12	0,000085
17	0,013111
19	0,000102
24	0,000024
25	0,000985
27	0,000631
33	0,001206
35	0,000104
39	0,000078
41	0,000074

44	0,000210
45	0,015719
53	0,000038
59	0,000009
65	0,000151
66	0,000009
67	0,000162
73	0,000062
74	0,002584
78	0,000217
83	0,000080
84	0,000059
86	0,000074
93	0,000183
94	0,000136
96	0,000009
99	0,000638
108	0,000173
114	0,000042
115	0,000104
122	0,000046
131	0,000151
139	0,000895
141	0,000010
142	0,000540
143	0,000033
146	0,000123
150	0,000014
152	0,000032
156	0,000011
160	0,000174
162	0,000134
166	0,000155
172	0,001143
217	0,000145
225	0,000495
243	0,000155
245	0,000198

258	0,000193
271	0,000008
272	0,000160
Укупно	0,045736

Табела 4.23 Површине појединачних кластера на општини Лазаревац.

Идентификатор кластера	Површина [км²]
1	0,000056
3	0,000097
4	0,000038
12	0,000985
26	0,000425
41	0,000209
46	0,000110
56	0,013250
79	0,000008
80	0,000052
81	0,000045
84	0,000606
89	0,000250
96	0,002186
103	0,003377
108	0,005883
110	0,006393
121	0,030796
122	0,000273
152	0,019475
155	0,000866
Укупно	0,085380

Табела 4.24 Површине појединачних кластера на општини Сопот.

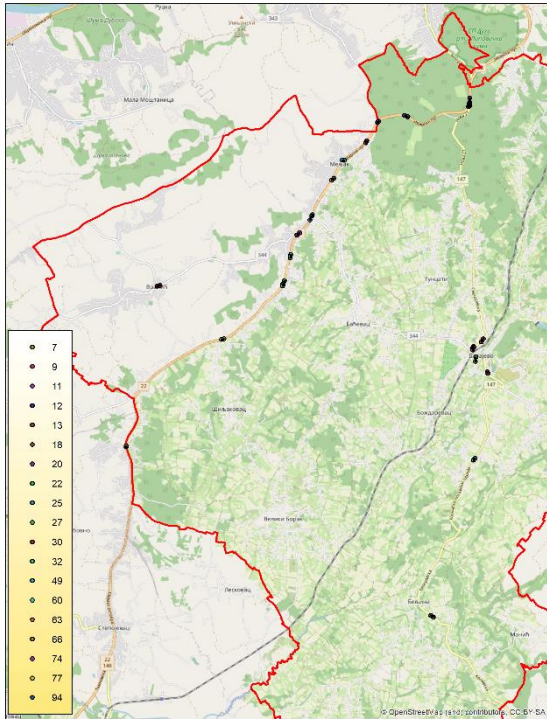
Идентификатор кластера	Површина [км²]
0	0,000189
4	0,000643
8	0,001057
15	0,000724
21	0,000158
26	0,000179
39	0,000031
40	0,000059
43	0,000340
47	0,000003
57	0,000039
Укупно	0,003422

Табела 4.25 Површине појединачних кластера на општини Барајево.

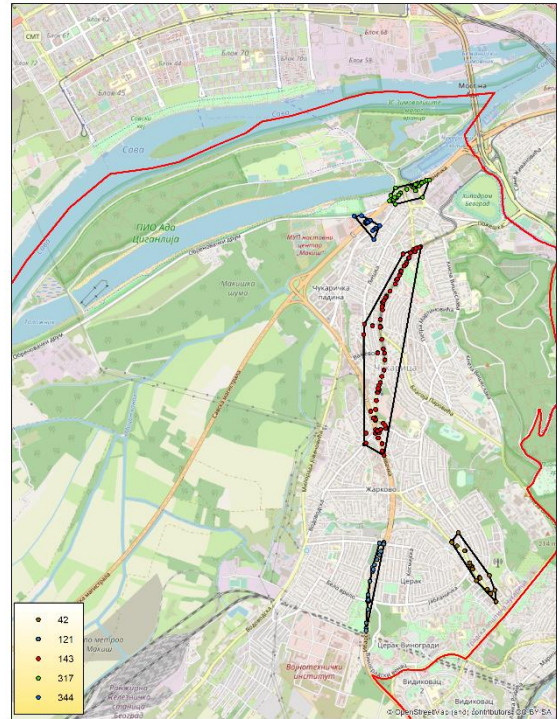
Идентификатор кластера	Површина [км ²]
7	0,000149
9	0,001548
11	0,000155
12	0,000072
13	0,000062
18	0,000031
20	0,004190
22	0,000249
25	0,000462
27	0,000168
30	0,000168
32	0,000190
49	0,000133
60	0,000111
63	0,000010
66	0,000140
74	0,000059
77	0,000129
94	0,000170
Укупно	0,008197

Табела 4.26 Површине појединачних кластера на општини Обреновац.

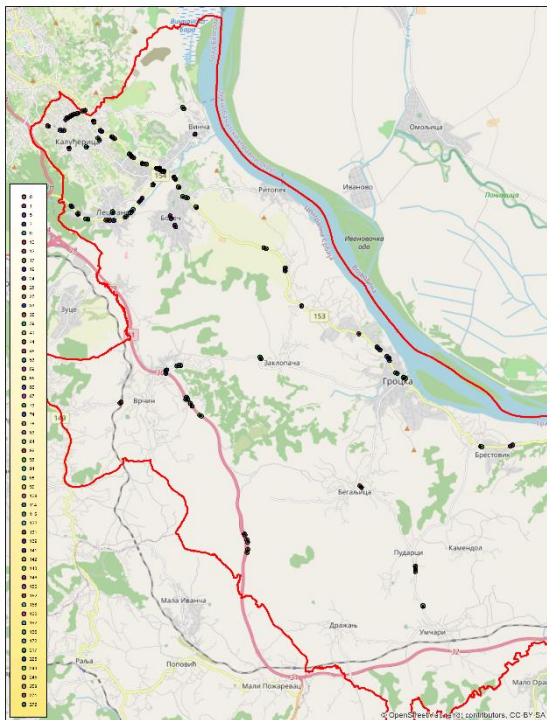
Идентификатор кластера	Површина [км ²]
34	0,181743
78	0,156931
211	0,177678
Укупно	0,516352



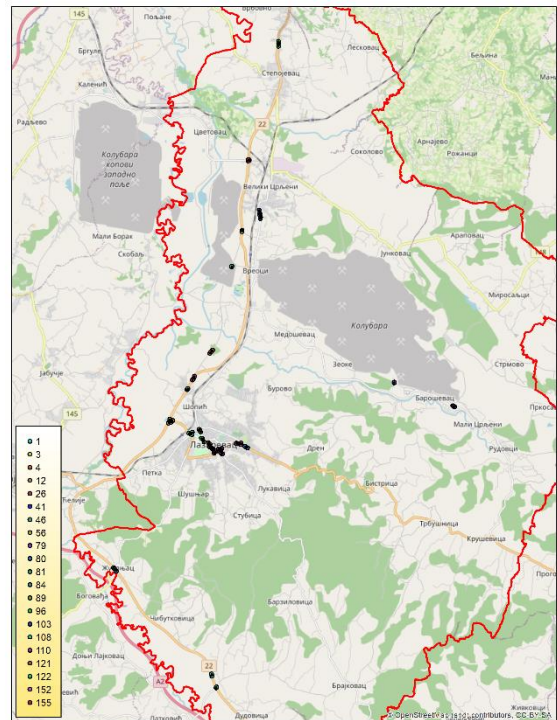
Слика 4.4 Селектовани кластери на општини Барајево.



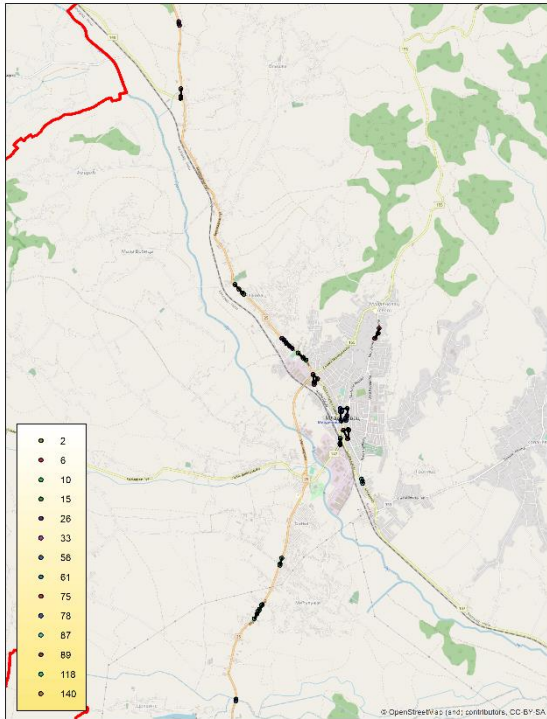
Слика 4.5 Селектовани кластери на општини Чукарица.



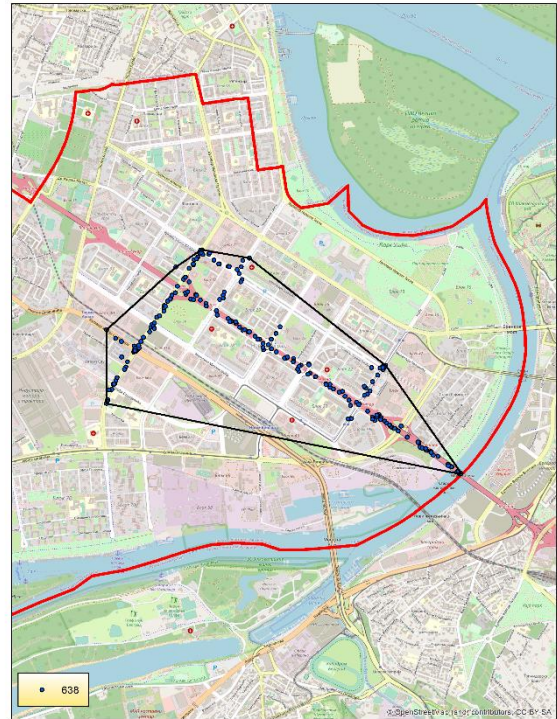
Слика 4.6 Селектовани кластери на општини Гроцка.



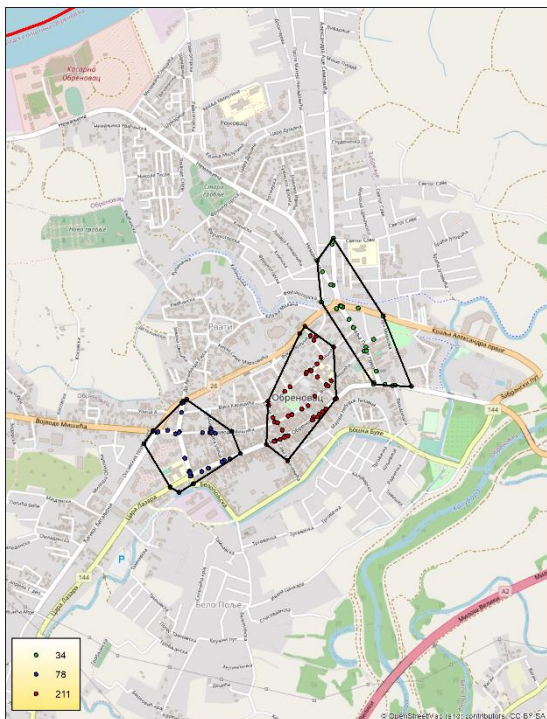
Слика 4.7 Селектовани кластери на општини Лазаревац.



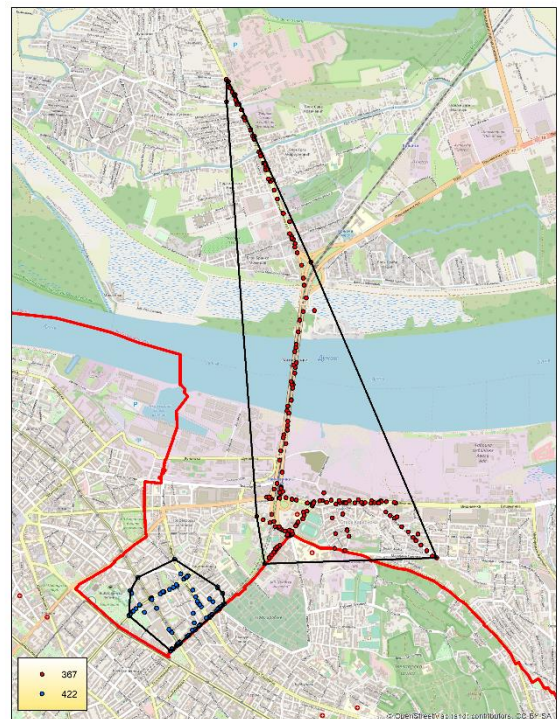
Слика 4.8 Селектовани кластери на општини Младеновац.



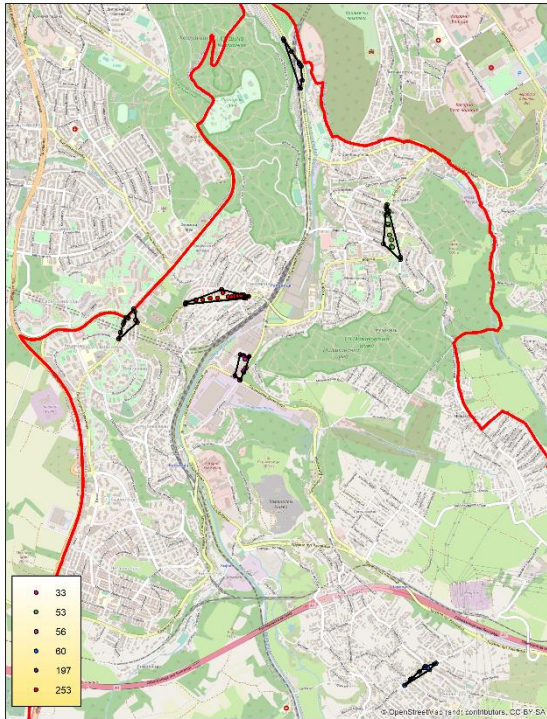
Слика 4.9 Селектовани кластери на општини Нови Београд.



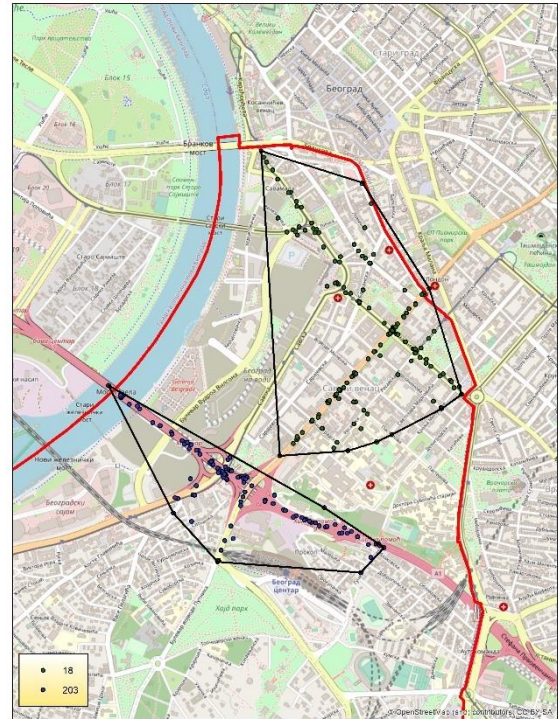
Слика 4.10 Селектовани кластери на општини Обреновац.



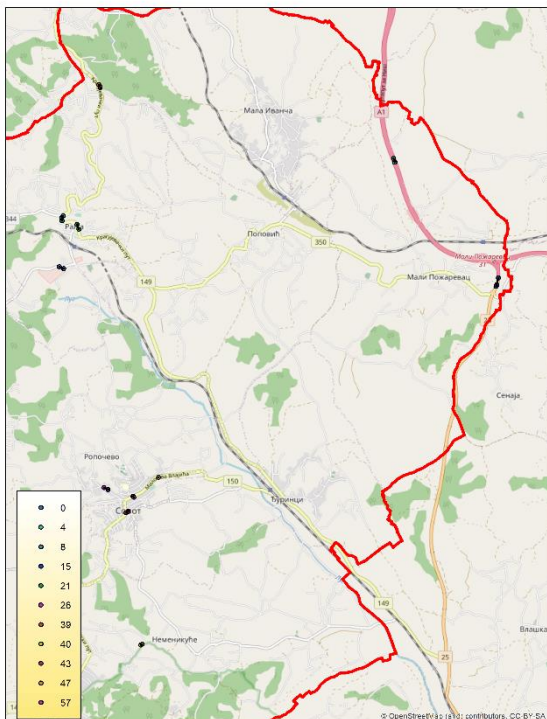
Слика 4.11 Селектовани кластери на општини Палилула.



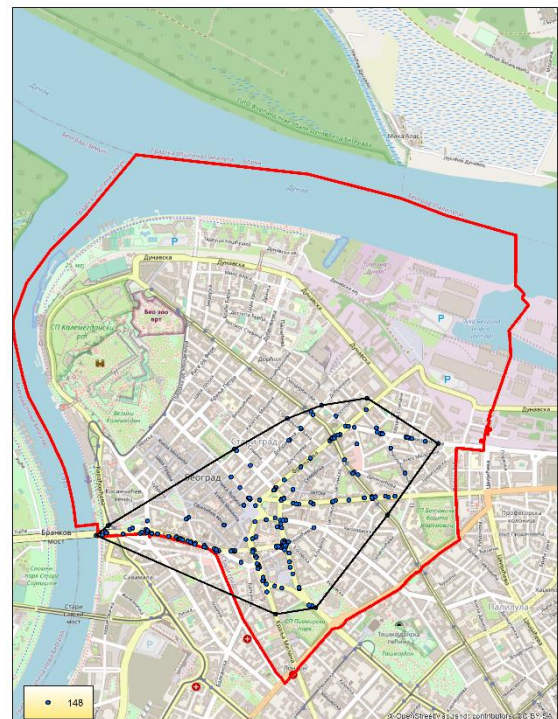
Слика 4.12 Селектовани кластери на општини Раковица.



Слика 4.13 Селектовани кластери на општини Савски венац.



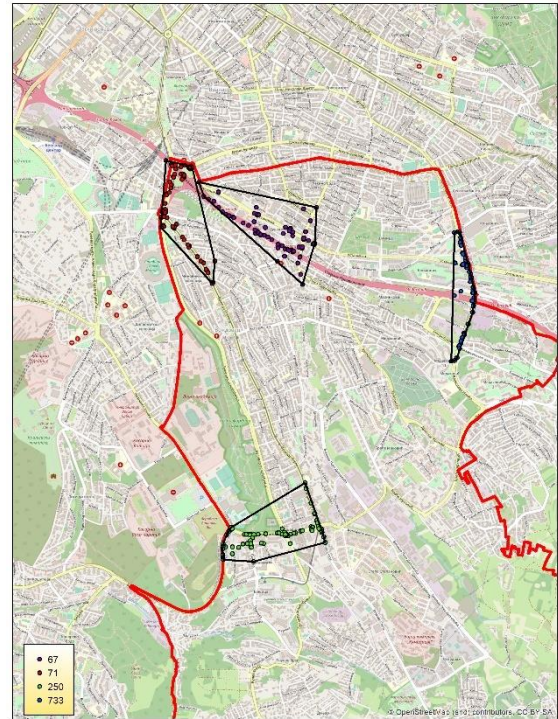
Слика 4.14 Селектовани кластери на општини Сопот.



Слика 4.15 Селектовани кластери на општини Стари град.



Слика 4.16 Селектовани кластери на општини Сурчин.



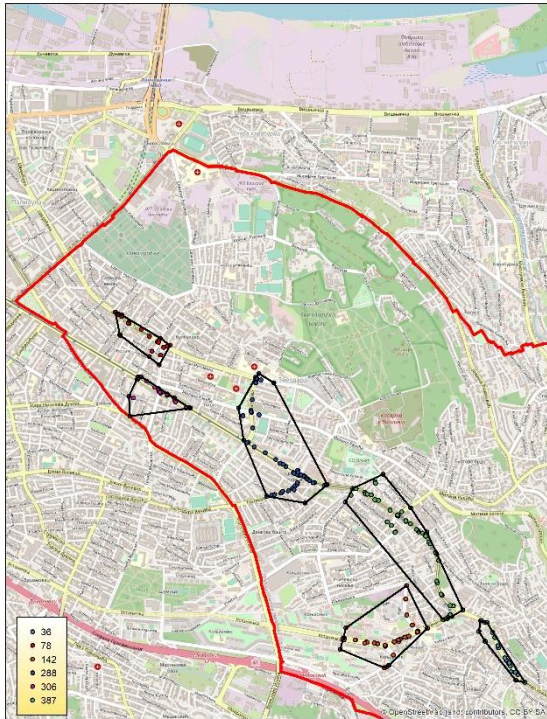
Слика 4.17 Селектовани кластери на општини Вождовац.



Слика 4.18 Селектовани кластери на општини Врачар.



Слика 4.19 Селектовани кластери на општини Земун.



Слика 4.20 Селектовани кластери на општини Звездара.

4.3.2 Детаљан приказ резултата за Нови Сад

У овој секцији је дат детаљан приказ релевантних резултата добијених приликом примене предложеног приступа на податке о саобраћајним незгодама које су се десиле у посматраним периодима на територији Новог Сада, примењујући претходно израчунати просторни праг ($\tau^*=170\text{m}$).

Табела 4.27 приказује расподелу критичних саобраћајних незгода по општинама, расподелу кластера по општинама, расподелу селектованих кластера и критичних саобраћајних незгода по општинама, вредности вектора дефинисаних у (4.4) и процену стабилности резултата кластерована.

Табела 4.28 садржи збирне податке о апсолутним и релативним површинама селектованих кластера по општинама, док табеле 4.29-4.40 садрже површине појединачних кластера по општинама.

На сл. 4.21-4.32 приказани су селектовани кластери по општинама. Границе општина назначене су црвеном бојом, а границе кластера црном бојом. За сваки кластер је назначен и једнозначни идентификатор, који служи као референца у односу на табеле 4.29-4.40. Сл. 4.21-4.32 су генерисане применом компоненте ArcMap софтверског пакета ESRI ArcGis (<https://www.esri.com>).

Табела 4.27 Сумарни подаци резултата кластеровања саобраћајних незгода у Новом Саду ($\tau^*=170m$).

Општина (m)	Период P ₁ [2019 ~ 2020]								Период P ₂ [2021]		Стабилност резултата	
	# нез-года	# кла-стера	Просечан # незгода по кластеру	Ст. дев.	# селектованих незгода	# селектованих кластера	Просечан # незгода по селектованом кластеру	Ст. дев.	# нез-года	# селектованих незгода	s ₁ (m, τ, P ₁)	s ₂ (m, τ, P ₂)
Бач	40	32	1,250	0,559	14	6	2,333	0,471	18	4	35,000	22,222
Бачка Паланка	206	123	1,650	1,482	85	19	4,474	2,036	104	42	41,262	40,385
Бачки Петровац	49	38	1,263	0,636	17	7	2,429	0,728	36	6	34,694	16,667
Беочин	52	42	1,238	0,569	18	8	2,250	0,661	28	4	34,615	14,286
Бечеј	128	91	1,407	0,972	58	21	2,762	1,306	85	22	45,313	25,882
Нови Сад	2122	521	4,063	12,637	348	2	174,000	43,000	1015	149	16,400	14,680
Србобран	69	49	1,408	0,780	33	13	2,538	0,746	33	10	47,826	30,303
Сремски Карловци	30	20	1,500	0,671	18	8	2,250	0,433	13	5	60,000	38,462
Темерин	116	65	1,769	1,262	53	14	3,786	1,264	55	17	45,690	30,909
Тител	34	30	1,133	0,562	4	1	4,000	0,000	18	0	11,765	0,000
Врбас	236	149	1,570	1,933	68	10	6,800	4,792	107	24	28,814	22,430
Жабалъ	92	70	1,314	0,622	40	18	2,222	0,629	62	3	43,478	4,839
s(M, τ, P₁, P₂)											0,939	

Табела 4.28 Збир података о апсолутним и релевантним површинама селектованих кластера по Општинама у Новом Саду.

Општина	Површина општине [км ²]	Површина кластера [км ²]	Релативна површина кластера [%]
Бач	367,268	0,00985	0,00268
Бачка Паланка	589,496	0,05225	0,00886
Бачки Петровац	158,257	0,00053	0,00034
Бечеј	486,196	0,03589	0,00738
Беочин	184,105	0,00122	0,00066
Нови Сад	698,816	2,52348	0,36111
Србобран	283,939	0,00509	0,00179
Сремски Карловци	50,538	0,00090	0,00179
Темерин	169,525	0,03149	0,01858
Тител	260,600	0,00006	0,00002
Врбас	375,326	0,13064	0,03481
Жабалъ	399,566	0,00302	0,00076
Укупно	402,633	2,79442	0,06945

Табела 4.29 Површине појединачних кластера на општини Бач.

Идентификатор кластера	Површина [км ²]
1	0,000005
5	0,000050
10	0,009141
12	0,000003
16	0,000116
23	0,000537
Укупно	0,009852

Табела 4.30 Површине појединачних кластера на општини Бачка Паланка.

Идентификатор кластера	Површина [км ²]
17	0,001111
20	0,000081
21	0,000262
25	0,000003
28	0,001785
29	0,004513
37	0,000007
38	0,005019
39	0,000009
41	0,000039
46	0,000020
50	0,023329
53	0,011393
55	0,000124
101	0,000066
115	0,002682
139	0,001398
160	0,000014
162	0,000396
Укупно	0,052252

Табела 4.31 Површине појединачних кластера на општини Бачки Петровац.

Идентификатор кластера	Површина [км ²]
2	0,000103
4	0,000076
6	0,000042
14	0,000082
19	0,000032
23	0,000095
27	0,000103
Укупно	0,000533

Табела 4.32 Површине појединачних кластера на општини Бечеј.

Идентификатор кластера	Површина [км ²]
8	0,000093
10	0,000006
16	0,017337
18	0,000019
19	0,000135
23	0,000017
24	0,000236
25	0,000009
26	0,000037
32	0,000087
36	0,000939
45	0,000027
48	0,000098
57	0,000231
58	0,015688
60	0,000096
66	0,000124
74	0,000520
82	0,000003
89	0,000184
122	0,000004
Укупно	0,035891

Табела 4.33 Површине појединачних кластера на општини Беочин.

Идентификатор кластера	Површина [км ²]
0	0,000093
1	0,000164
8	0,000270
9	0,000235
13	0,000020
24	0,000170
26	0,000127
36	0,000140
Укупно	0,001219

Табела 4.34 Површине појединачних кластера на општини Нови Сад.

Идентификатор кластера	Површина [км ²]
708	1,577829
849	0,945648
Укупно	2,523477

Табела 4.35 Површине појединачних кластера на општини Србобран.

Идентификатор кластера	Површина [км ²]
3	0,000004
7	0,000016
8	0,003663
10	0,000020
11	0,000245
12	0,000112
23	0,000323
24	0,000133
26	0,000205
29	0,000014
32	0,000131
37	0,000073
41	0,000153
Укупно	0,005093

Табела 4.36 Површине појединачних кластера на општини Сремски Карловци.

Идентификатор кластера	Површина [км ²]
0	0,000042
1	0,000170
4	0,000015
7	0,000207
10	0,000094
11	0,000005
15	0,000309
20	0,000061
Укупно	0,000903

Табела 4.37 Површине појединачних кластера на општини Темерин.

Идентификатор кластера	Површина [км ²]
4	0,003475
5	0,000037
7	0,007451
8	0,000230
9	0,000784
10	0,000582
12	0,001470
14	0,006744
21	0,000015
28	0,001430
40	0,001865
50	0,006686
55	0,000005
58	0,000717
Укупно	0,031490

Табела 4.38 Површина кластера на општини Тител.

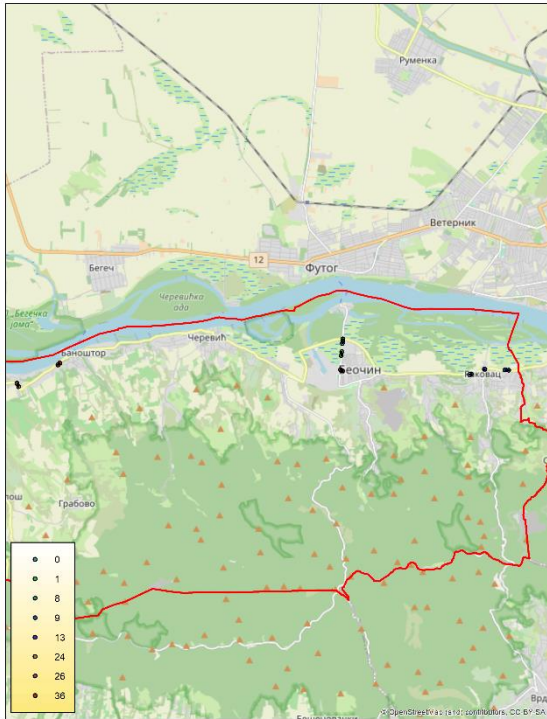
Идентификатор кластера	Површина [км ²]
9	0,000057
Укупно	0,000057

Табела 4.39 Површине појединачних кластера на општини Врбас.

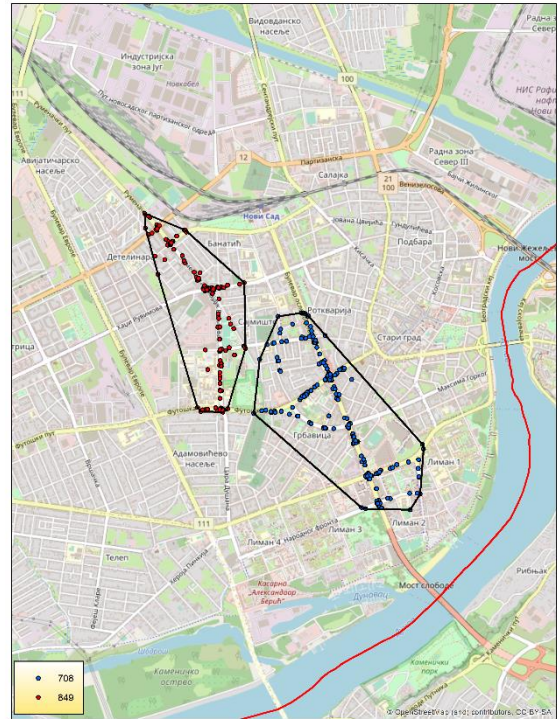
Идентификатор кластера	Површина [км ²]
2	0,000594
26	0,003130
32	0,000754
35	0,101063
48	0,000009
60	0,007903
69	0,001578
92	0,013197
105	0,000708
123	0,001701
Укупно	0,130637

Табела 4.40 Површине појединачних кластера на општини Жабаљ.

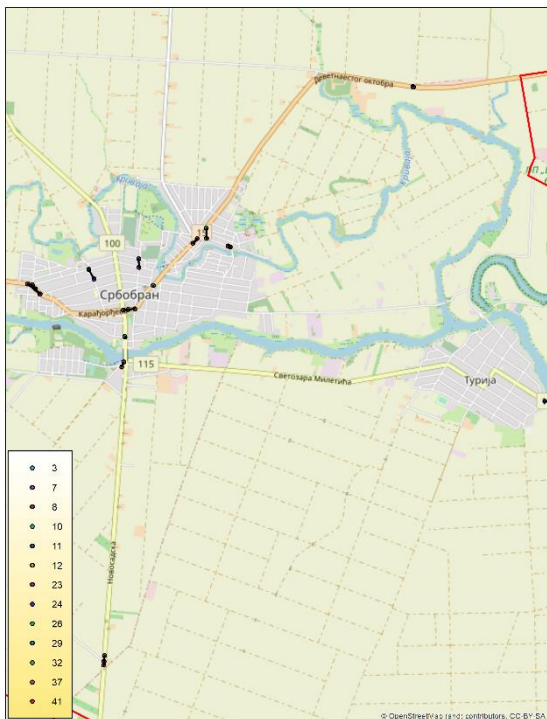
Идентификатор кластера	Површина [км ²]
5	0,000156
9	0,000230
10	0,000153
11	0,000227
13	0,000072
14	0,000014
22	0,000168
23	0,000095
27	0,000003
28	0,000176
30	0,000066
34	0,000008
42	0,000192
48	0,000896
49	0,000265
55	0,000112
63	0,000136
64	0,000051
Укупно	0,003020



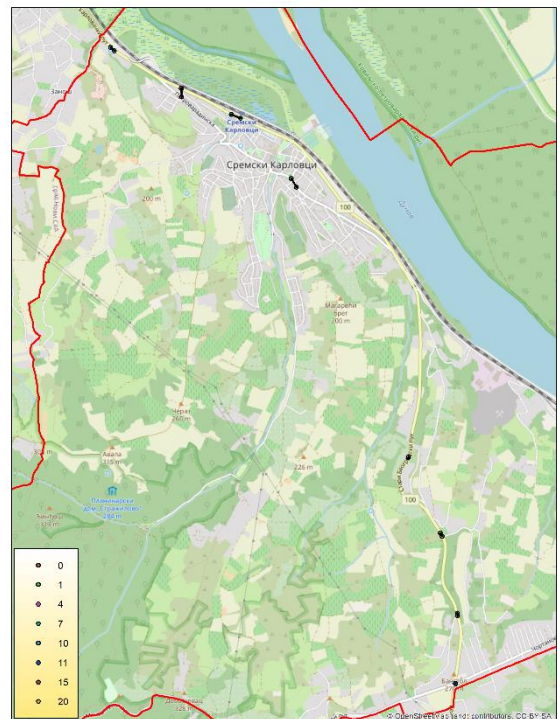
Слика 4.25 Селектовани кластери на општини Беочин.



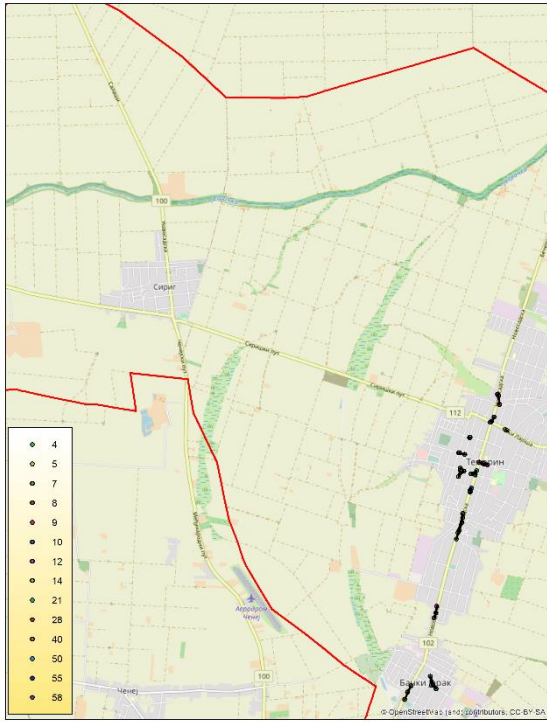
Слика 4.26 Селектовани кластери на општини Нови Сад.



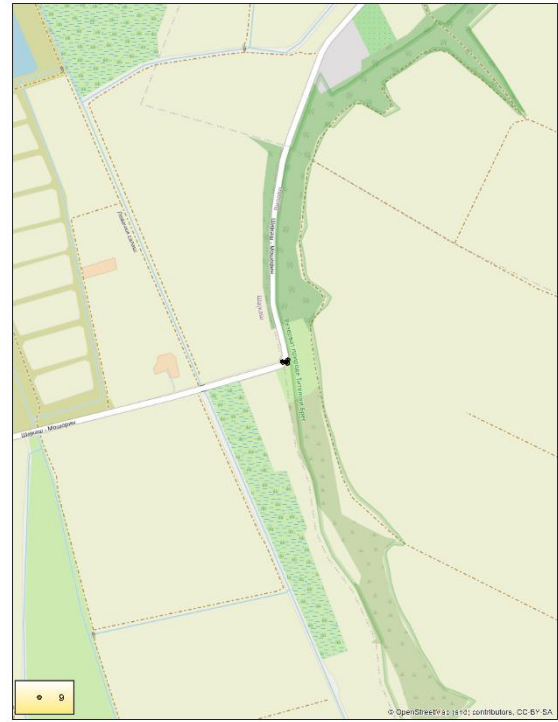
Слика 4.27 Селектовани кластери на општини Србобран.



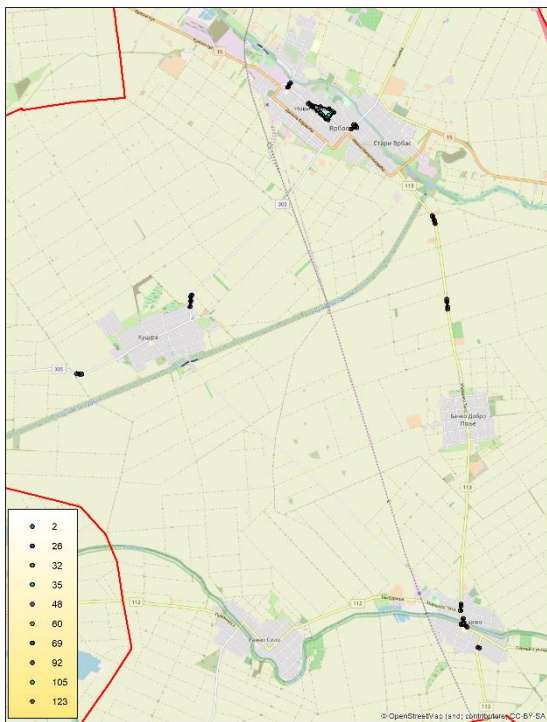
Слика 4.28 Селектовани кластери на општини Сремски Карловци.



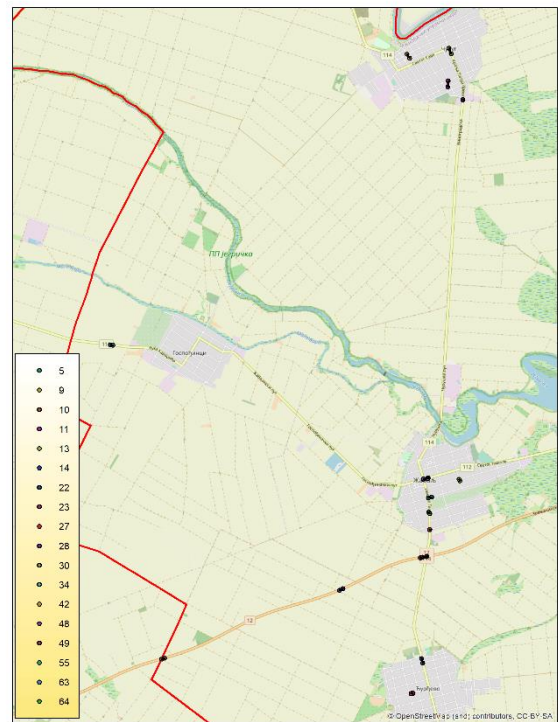
Слика 4.29 Селектовани кластери на општини Темерин.



Слика 4.30 Селектовани кластери на општини Тутел.



Слика 4.31 Селектовани кластери на општини Врбас.



Слика 4.32 Селектовани кластери на општини Жабалј.

4.3.3 Детаљан приказ резултата за Ниш

У овој секцији је дат детаљан приказ релевантних резултата добијених приликом примене предложеног приступа на податке о саобраћајним незгодама које су се десиле у посматраним периодима на територији Ниша, примењујући претходно израчунати просторни праг ($\tau^*=160\text{m}$).

Табела 4.41 приказује расподелу критичних саобраћајних незгода по општинама, расподелу кластера по општинама, расподелу селектованих кластера и критичних саобраћајних незгода по општинама, вредности вектора дефинисаних у (4.4) и процену стабилности резултата кластерована.

Табела 4.42 садржи збирне податке о апсолутним и релативним површинама селектованих кластера по општинама, док табеле 4.43-4.50 садрже површине појединачних кластера по општинама.

На сл. 4.33-4.40 приказани су селектовани кластери по општинама. Границе општина назначене су црвеном бојом, а границе кластера црном бојом. За сваки кластер је назначен и једнозначни идентификатор, који служи као референца у односу на табеле 4.43-4.50. Сл. 4.33-4.40 су генерисане применом компоненте ArcMap софтверског пакета ESRI ArcGis (<https://www.esri.com>).

Табела 4.41 Сумарни подаци резултата кластеровања саобраћајних незгода у Нишу ($\tau^* = 160m$).

Општина (m)	Период P ₁ [2019 ~ 2020]								Период P ₂ [2021]		Стабилност резултата	
	# нез-года	# кла-стера	Просечан # незгода по кластеру	Ст. дев.	# селектованих незгода	# селектованих кластера	Просечан # незгода по селектованом кластеру	Ст. дев.	# нез-года	# селектованих незгода	s ₁ (m, τ , P ₁)	s ₂ (m, τ , P ₂)
Алексинач	159	106	1,472	1,549	37	6	6,167	3,976	82	9	23,270	10,976
Дољевац	73	54	1,352	1,108	29	10	2,900	1,921	36	7	39,726	19,444
Гаџин Хан	23	22	1,045	0,208	2	1	2,000	0,000	4	0	8,696	0,000
Мерошина	37	35	1,057	0,232	4	2	2,000	0,000	27	1	10,811	3,704
Ниш	734	264	2,754	7,387	149	2	74,500	24,500	346	59	20,300	17,052
Нишка Бања	40	34	1,176	0,452	11	5	2,200	0,400	17	4	27,500	23,529
Ражањ	26	22	1,091	0,287	4	2	2,000	0,000	6	0	15,385	0,000
Сврљиг	36	27	1,296	0,597	14	6	2,333	0,471	10	3	38,889	30,000
											s(M, τ , P ₁ , P ₂)	0,940

Табела 4.42 Збир података о апсолутним и релевантним површинама селектованих кластера по Општинама у Нишу.

Општина	Површина општине [км ²]	Површина кластера [км ²]	Релативна површина кластера [%]
Алексинач	706,335	0,11574	0,01639
Дољевац	121,275	0,00093	0,00076
Гаџин Хан	324,931	0,00015	0,00005
Мерошина	193,089	0,00005	0,00002
Ниш	449,929	1,59431	0,35435
Нишка Бања	146,185	0,00173	0,00118
Ражањ	288,512	0,00003	0,00001
Сврљиг	496,894	0,00069	0,00014
Укупно	2727,151	1,71362	0,06284

Табела 4.43 Површине појединачних кластера на општини Алексинач.

Идентификатор кластера	Површина [км ²]
3	0,002572
7	0,008346
13	0,103235
14	0,000368
37	0,001088
104	0,000134
Укупно	0,115743

Табела 4.44 Површине појединачних кластера на општини Доњевац.

Идентификатор кластера	Површина [км ²]
1	0,000014
4	0,000020
6	0,000047
14	0,000032
18	0,000151
21	0,000409
26	0,000047
33	0,000035
36	0,000084
47	0,000087
Укупно	0,000926

Табела 4.45 Површина кластера на општини Гаџин Хан.

Идентификатор кластера	Површина [км ²]
11	0,000151
Укупно	0,000151

Табела 4.46 Површине појединачних кластера на општини Меришина.

Идентификатор кластера	Површина [км ²]
2	0,000019
9	0,000028
Укупно	0,000047

Табела 4.47 Површине појединачних кластера на општини Ниш.

Идентификатор кластера	Површина [км ²]
9	1,263633
11	0,330675
Укупно	1,594308

Табела 4.48 Површине појединачних кластера на општини Нишка Бања.

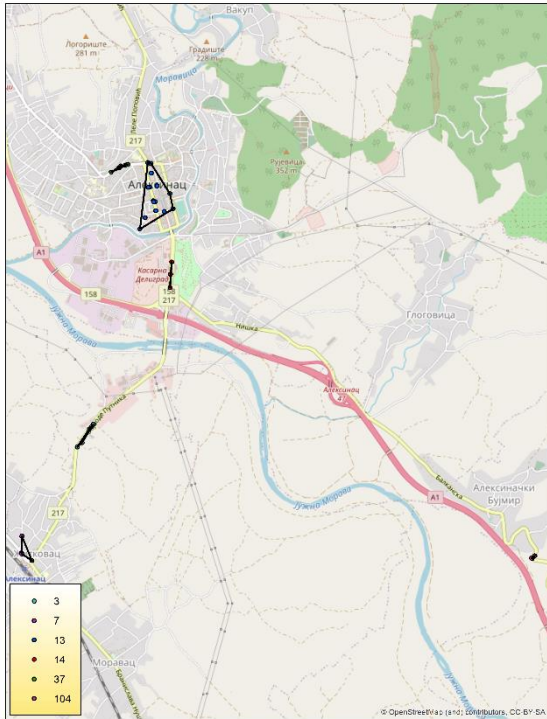
Идентификатор кластера	Површина [км ²]
3	0,000074
7	0,001401
10	0,000185
19	0,000019
35	0,000053
Укупно	0,001732

Табела 4.49 Површине појединачних кластера на општини Ражањ.

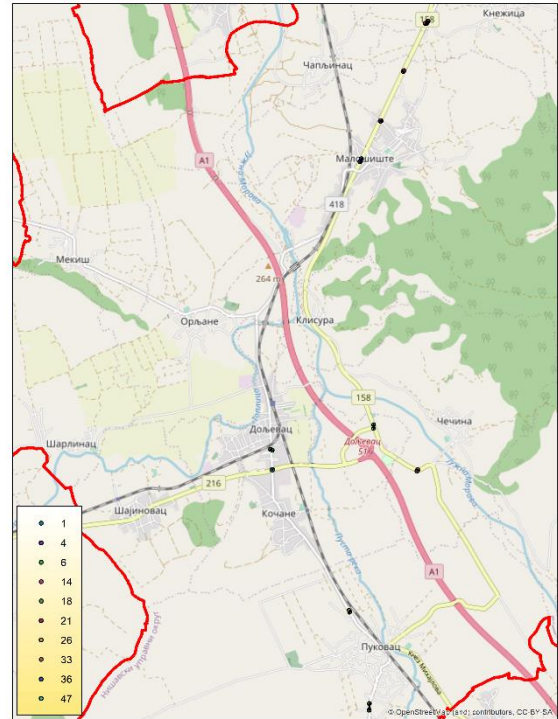
Идентификатор кластера	Површина [км ²]
2	0,000024
21	0,000003
Укупно	0,000027

Табела 4.50 Површине појединачних кластера на општини Сврљиг.

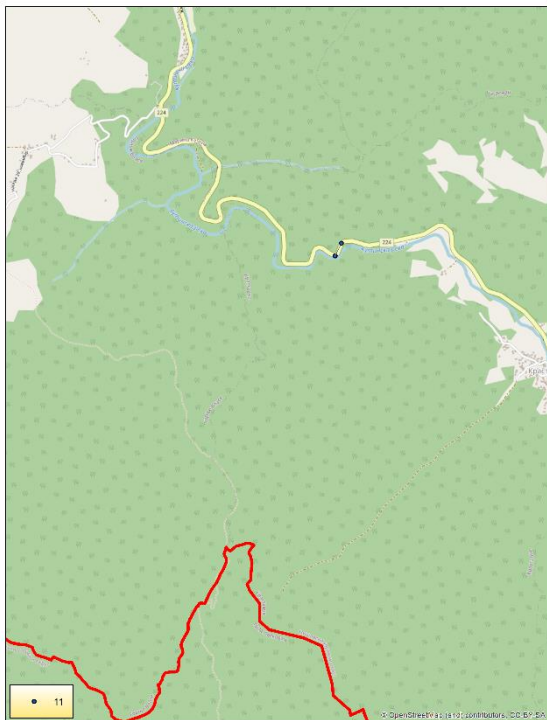
Идентификатор кластера	Површина [км ²]
1	0,000315
3	0,000025
11	0,000061
17	0,000083
18	0,000130
27	0,000074
Укупно	0,000688



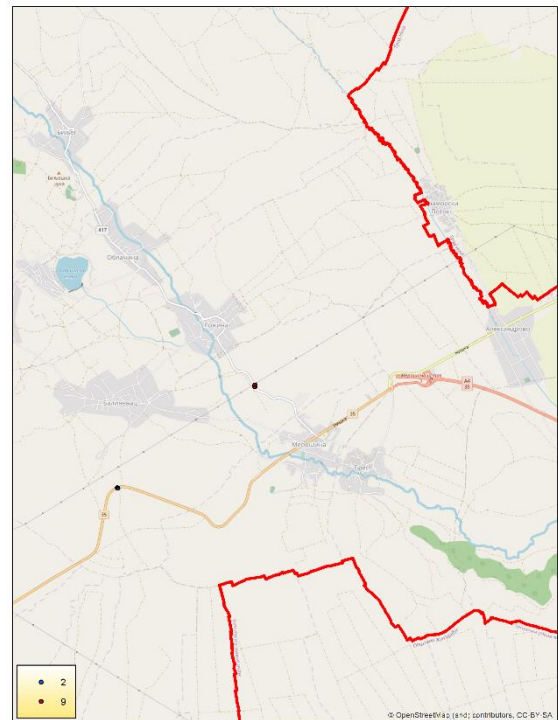
Слика 4.33 Селектовани кластери на општини Алексинац.



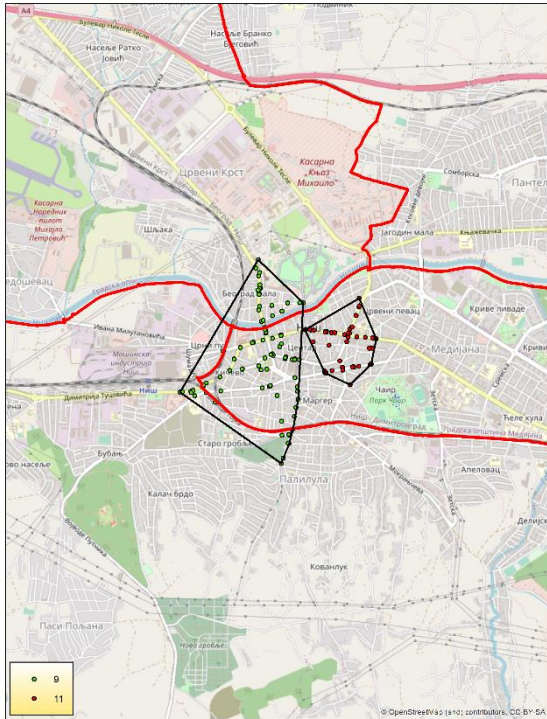
Слика 4.34 Селектовани кластери на општини Дољевац.



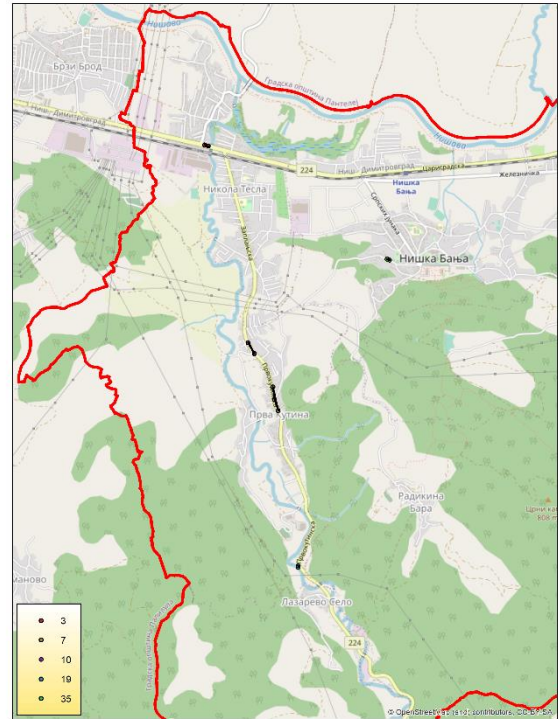
Слика 4.35 Селектовани кластери на општини Гаџин Хан.



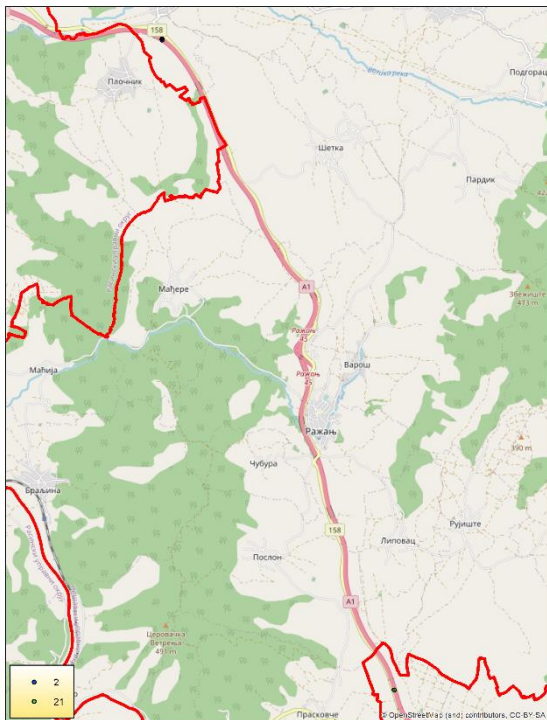
Слика 4.36 Селектовани кластери на општини Мерошина.



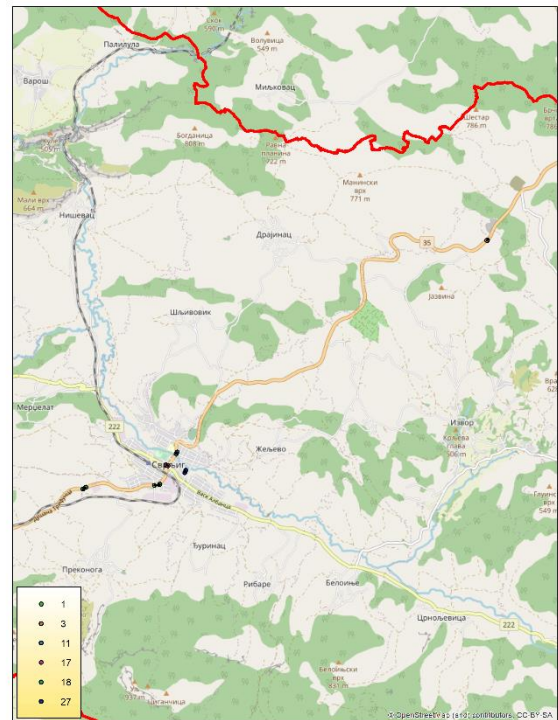
Слика 4.37 Селектовани кластери на општини Ниш.



Слика 4.38 Селектовани кластери на општини Нишка Бања.



Слика 4.39 Селектовани кластери на општини Ражањ.



Слика 4.40 Селектовани кластери на општини Сврџиг.

4.4 Закључак

У овом поглављу представљен је доменски–специфичан критеријум за процену резултата кластерована саобраћајних незгода, који је потом примењен за аутоматско израчунавање вредности прага t , који представља улазни параметар алгоритма за кластероване представљеног у претходном поглављу.

Предложени приступ илустрован је на скупу који садржи реалне податке о 18880 саобраћајних незгода са повређеним или погинулим лицима које су се десиле у трогодишњем периоду од јануара 2019. до децембра 2021. године у три највећа града у Републици Србији: Београду, Новом Саду и Нишу. Добијени резултати кластерована додатно су дискутовани у наредном поглављу.

5. Дискусија

У овом поглављу посматрају се изабрани практични и концептуални аспекти приступа предложеног у овој тези. У секцији 5.1 врши се екстерно процењивање добијених резултата, тј., резултати добијени на основу анализе података о саобраћајним незгодама које су се десиле у Београду процењују се у односу на локације камерних стубова у Београду. У секцији 5.2, предложени приступ кластеровању посматра се из перспективе других релевантних истраживања која су инспирисана принципима Гешталт-теорије.

5.1 Екстерна евалуација

У претходном поглављу демонстрирана је стабилност резултата предложеног алгорита у времену. Из (4.26) и (4.27) може се уочити да алгоритам израчунава сличне вредности просторног прага за све анализиране градове. Додатно, релативне површине селектованих кластера износе само 0,522, 0,069 и 0,063 процената површине града, в. табеле (4.5-4.7).

У овом поглављу, резултати добијени на основу анализе података о саобраћајним незгодама које су се десиле у Београду додатно су процењени екстерно, тј., у односу на локације камерних стубова у Београду. Постављање камерних стубова је био дугорочан процес. За потребе ове анализе, разматрају се локације камерних стубова у одређеном тренутку, односно у августу 2020. године (Ministry of Interior 2023), из следећих разлога:

- Доступност података: Информације о локацијама камерних стубова у августу 2020. године могу се извести из јавно доступних података Министарства унутрашњих послова Републике Србије (Ministry of Interior 2023).
- Екстерни критеријум: Локације камерних места одређене су на основу стручне анализе коју је спровела трећа страна, независно од овог истраживања. Локације камерних стубова су индикативне, између осталог, за критичне саобраћајне тачке и могу се користити као критеријум за екстерну евалуацију.
- Временска прикладност: Посматрају се локације камерних стубова у раним фазама постављана, под претпоставком да је постављање камерних стубова почело од најкритичнијих тачака. Поред тога, кластери добијени применом представљеног алгорита на податке о саобраћајним незгодама које су се десиле у периоду P_1 , тј., од јануара 2019. године до децембра 2020. године.

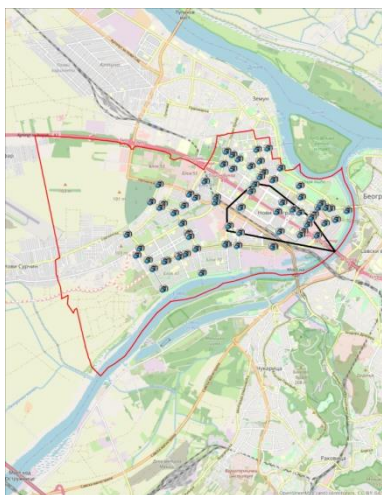
Тренутак у којем се посматрају локације камерних стубова (тј., август 2020. године) временски је близу краја овог периода.

- У посматраним периодима P_1 и P_2 , тј., од јануара 2019. године до децембра 2021. године, камере на посматраним локацијама још увек нису биле званично активирани, тј., нису имала утицај на понашање возача у посматраном периоду.

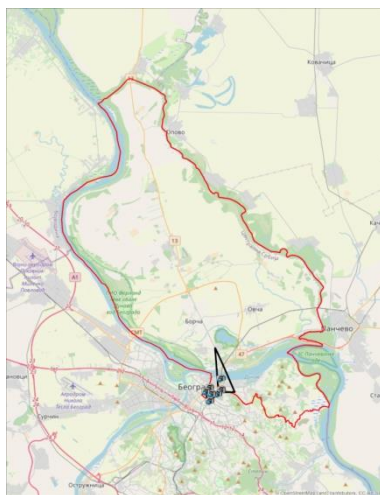
У августу 2020. године, 154 камерних стубова, на којима су се налазиле 392 камере, постављено је у девет од седамнаест општина у Београду. Резултати кластерована за ових девет централних општина и локације камерних стубова представљени су на сл. 5.1 и 5.2.

Расподела камерних стубова и камера на овим општинама представљена је у прве три колоне табеле 5.1. Број и удео камерних стубова и камера који припадају кластерима добијеним применом почетног прага (4.26) представљени су у последње четири колоне ове табеле.

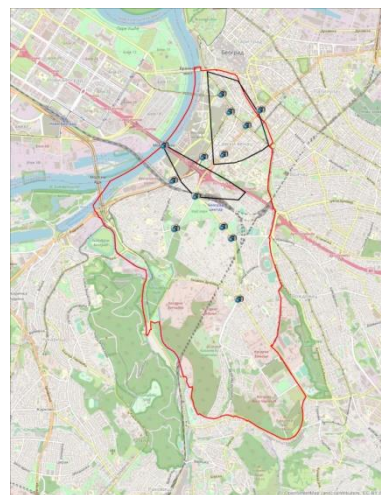
Резултати кластерована могу се сумирати на следећи начин. Укупна површина селектованих кластера представља само 0,522 процента површине града (в. табелу 4.4) и покрива 40,26 процента свих камерних стубова и 35,97 процената камера у граду (в. табелу 5.1). Ови резултати се могу сматрати задовољавајућим, посебно ако се узме у обзир циљ ове дисертације да се представи приступ намењен за примену у условима ограничених људских и техничких ресурса за праћење и управљање саобраћајем. Поред тога, треба приметити да једна општина (тј., Нови Београд, в. сл. 5.1(а)) садржи знатно већи број камера од осталих. Површина ове општине износи само 1,26 процената површине града (тј., 40,756 км² од 3331,469 км², в. табелу 4.5), али се на њој налази 47,40 процената свих камерних стубова (тј., 73 од 154 камерна места) и 55,87 процената свих камера (тј., 219 од 392, в. табелу 5.1). Расподела камера на овој општини одређена је разлозима који нису искључиво везани за саобраћај, због чега није примарно индикативна у односу на критичне саобраћајне тачке. Ако изузмемо ову општину из разматрања, преостали кластери покривају 60,49 процената свих камерних стубова и 63,58 процената камера.



(а) Нови Београд



(б) Палилула



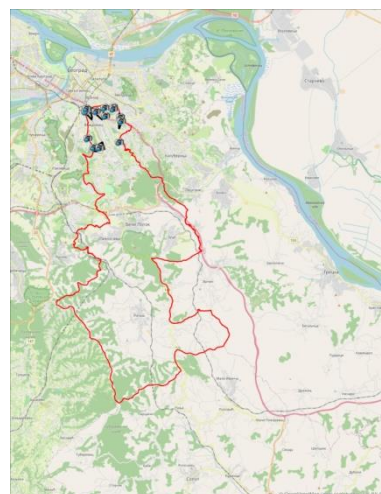
(в) Савски венац



(г) Стари град



(д) Врачар

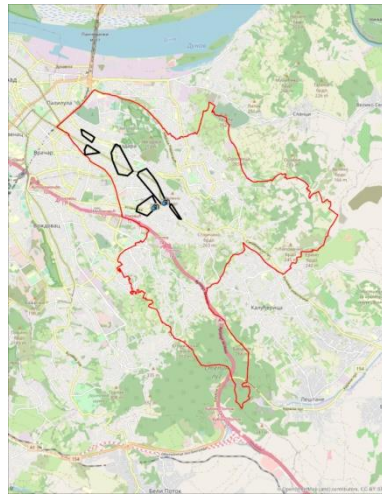


(ђ) Вождовац

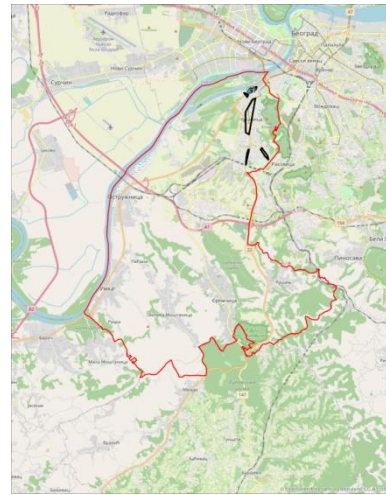
Слика 5.1 Шест (од девет, в. сл. 5.2) централних општина Београда са локацијама камерних стубова у августу 2020. године. За сваку општину приказани су кластери добијени применом вредности просторног прага $\tau=150\text{m}$. Сваки кластер саобраћајних незгода представљен је минималним конвексним обухватајућим многоуглом. Слика је генерисана применом компоненте ArcMap софтверског пакета ESRI ArcGis (<https://www.esri.com>) и преузета из (Кошанин et al. 2023).



(е) Земун



(ж) Звездара



(з) Чукарица

Слика 5.2 Три (од девет, в. сл. 5.1) централних општина Београда са локацијама камерних стубова у августу 2020. године. За сваку општину приказани су кластери добијени применом вредности просторног прага $\tau=150\text{m}$. Сваки кластер саобраћајних незгода представљен је минималним конвексним обухватајућим многоуглом. Слика је генерисана применом компоненте ArcMap софтверског пакета ESRI ArcGis (<https://www.esri.com>) и преузета из (Коџанин et al. 2023).

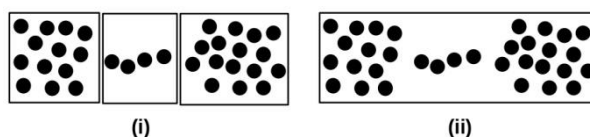
Табела 5.1 Резултати екстерне евалуације.

Општина	# камерних места	# камера	# покривених камерних стубова	Удео покривених камерних стубова [%]	# покривених камера	Удео покривених камера [%]
Нови Београд	73	219	13	17,81	31	14,16
Палилула	7	17	4	57,14	10	58,82
Савски венац	13	20	9	69,23	15	75,00
Стари град	15	42	12	80,00	35	83,33
Врачар	5	10	4	80,00	9	90,00
Вождовац	14	19	8	57,14	11	57,89
Земун	24	57	9	37,50	22	38,60
Звездара	2	6	2	100,00	6	100,00
Чукарица	1	2	1	100,00	2	100,00
Укупно	154	392	62	40,26	141	35,97

5.2 Принцип близине у Гешталт-теорији

Предложени приступ кластеровању може се посматрати из перспективе других релевантних истраживања која су инспирисана принципима Гешталт-теорије.

Идеја о инспирисању рачунарских модела овим принципима није нова. На пример, (Osourn & Martinez 1995) представљају алгоритам за кластеровање базиран на локалном k -димензионалном суседству сваке тачке, дозвољавајући произвољан број кластера и произвољан облик кластера. Међутим, њихова имплицитна концептуализација близине разликује се од концептуализације у овој тези. Према концептуализацији коју усвајају (Osourn & Martinez 1995), образац тачака који је приказан на слици 5.3(i) садржи три кластера: два велика кластера и један мањи кластер који се налази између њих. У нашем приступу, близина тачака (тј., локација саобраћајних незгода) дефинисана је помоћу транзитивног затварања, тако да исти образац садржи само један кластер (в. сл. 5.3(ii)).



Слика 5.3 Разлика у концептуализацији принципа близине у Гешталт-теорије у (i) Osourn & Martinez (1995) и (ii) приступ у овој тези. Слика је преузета из (Gnjatović et al. 2022).

Два различита приступа сегментовању дигиталних слика заснована на принципима Гешталт-теорије предложена су у (Wang & Li 2008; в. и Wu & Zhang 2013). У вези са принципом близине у Гешталт-теорији, ови приступи разматрају разлику интензитета између сегмената слике и имплицитно укључују транзитивно затварање, међутим, оба приступа ограничавају избор сегмената на суседство тренутно наглашеног сегмента. Чак је и алгоритам за сегментовање слике, представљен у (Felzenszwalb & Huttenlocher 2004), на којем је заснован приступ описан у овој тези, укључује процену удаљености два сегмента. У приступу представљеном у овој тези, ово ограничење није присутно, што ће бити размотрено у наставку.

У (Felzenszwalb & Huttenlocher 2004), разлика између сегмената C_i и C_j дефинисана је као минимална тежина гране која их повезује:

$$\text{Diff}(C_i, C_j) = \min_{\substack{v_i \in C_i, \\ v_j \in C_j, \\ (v_i, v_j) \in E}} \omega(v_i, v_j), \quad (5.1)$$

где су v_i и v_j два суседна пиксела (тј., $(v_i, v_j) \in E$) који припадају сегментима C_i и C_j , респективно, а $\omega(v_i, v_j)$ представља апсолутну разлику интензитета пиксела v_i и v_j . У приступу представљеном у овој тези разматра се просторно растојање између локација саобраћајних незгода, а растојање између два кластера дефинише се на исти начин. Са друге стране, да би се утврдило постојање границе између сегмената C_i и C_j , приступ представљен у (Felzenszwalb & Huttenlocher 2004) претпоставља да разлика између ових сегмената мора бити већа од збира њихових интерних разлика $\text{Int}(C_1)$ и $\text{Int}(C_2)$:

$$\text{Diff}(C_i, C_j) > \min\{\text{Int}(C_i) + \tau(C_i), \text{Int}(C_j) + \tau(C_j)\}, \quad (5.2)$$

где су вредности прагова дефинисана као обрнуто пропорционалне величини сегмената, тј., $\tau(C) \sim \frac{1}{|C|}$. За разлику од тога, приступ у овој тези примењује релаксирани услов: да би се утврдило постојање границе између два кластера, њихова међусобна удаљеност мора бити већа од константне вредности прага (в. једначину (3.6) у секцији 3.2).

Образложење за ову одлуку о релаксирању услова везано је за домен ове студије. У складу са циљем да се детектују путни сегменти високог и просторно пролонгираног ризика за настанак саобраћајних незгода, не захтева се јачи услов за утврђивање постојања границе између релативно малих кластера, због чега се у обзир не узимају њихове интерне разлике.

6. Закључак

У овој тези предложен је нов приступ кластеровању саобраћајних незгода у урбаним срединама намењеном за детектовање и селектовање путних сегмената високог и просторно пролонгираног ризика за настанак саобраћајних незгода.

Следеће опште одреднице овог приступа чине га применљивим у различитим саобраћајним контекстима:

- Овај приступ полази од претпоставке да је свака саобраћајна незгода описана само геопросторним координатама и временским тренутком у којем се десила, не узимајући у обзир шири контекст саобраћајне незгоде.
- Предложени приступ се може применити на било коју урбану средину са хијерархијском структуром надлежности контроле саобраћаја.
- Предложени приступ је намењен за примену у условима ограничених људских и техничких ресурса.

Методолошки допринос ове тезе може се сумирати на следећи начин. Основни алгоритам за кластеровање извршава се у две фазе. Прва фаза представља адаптацију графовски заснованог алгоритма за сегментацију дигиталних слика инспирисану изабраним принципима Гешталт-теорије, са једним хиперпараметром који се односи на праг растојања. Друга фаза представља селектовање кластера који су доминантни у односу на број саобраћајних незгода, засновано на адаптацији алгоритма за бинаризовање дигиталних слика без улазних параметара. Поред тога, предложен је нов, доменски-специфичан критеријум за процену резултата кластеровања, који промовише стабилност резултата у времену и просторну колокацију незгода у различитим периодима, а санкционише величину селектованих кластера. На основу овог критеријума и адаптације алгоритма за одређивање колена криве врши се аутоматско одређивање вредности улазног хиперпараметра примењеног у првој фази кластеровања.

Овај приступ се разликује од других алгоритама који се примењују за кластеровање саобраћајних незгода по томе што се не заснива на тзв. густини опсервација, али испољава главне предности ових приступа: дозвољава кластере произвољног облика, не захтева априорно спецификовање броја кластера и искључује одступајуће опсервације. Коначно, приступ је илустрован и позитивно оцењен на корпусу који садржи јавно доступне податке о локацијама саобраћајних незгода са повређеним или погинулим лицима које су се десиле у трогодишњем периоду од јануара 2019. до децембра 2021. године у три највећа

града у Републици Србији: Београду, Новом Саду и Нишу. Из тога се може извести закључак да су хипотезе постављене у уводном поглављу успешно потврђене.

Литература

1. Aggarwal, C. C.; Reddy, C. K. (eds.). *Data Clustering: Algorithms and Applications*. CRC Press. 2014.
2. Ankerst, M.; Breunig, M.M.; Kriegel, H.-P.; Sander, J. OPTICS: Ordering points to identify the clustering structure. In *Proceedings of the ACM SIGMOD International Conference on Management of Data, Philadelphia, PA, USA, 31 May–3 June 1999*; pp. 49–60.
3. Arora, P. Final Report. Independent Evaluation of the United Nations Road Safety Trust Fund (UNRSF) Secretariat. 8 April 2021. DeftEdge Corporation. Available online: https://unece.org/sites/default/files/2021-04/TRANS_FinalReportUNRSF_Apri21_0.pdf (приступљено 10.08).
4. Baek, J. Highway Regional Classification Method Based on Traffic Flow Characteristics for Highway Safety Assessment. *Sensors* 2022, 22, 86.
5. Bajada, T.; Attard, M. A typological and spatial analysis of pedestrian fatalities and injuries in Malta. *Res. Transp. Econ.* 2021, 86, 101023.
6. Bokaba, T.; Doorsamy, W.; Paul, B.S. Comparative Study of Machine Learning Classifiers for Modelling Road Traffic Accidents. *Appl. Sci.* 2022, 12, 828.
7. Calinski, T.; Harabasz, J. A dendrite method for cluster analysis. In *Communications in Statistics*; Taylor & Francis: Oxfordshire, UK, 1974; Volume 3, pp. 1–27.
8. Chen, S.; Cheng, K.; Yang, J.; Zang, X.; Luo, Q.; Li, J. Driving Behavior Risk Measurement and Cluster Analysis Driven by Vehicle Trajectory Data. *Appl. Sci.* 2023, 13, 5675.
9. Davies, D.L.; Bouldin, D.W. A Cluster Separation Measure. *IEEE Trans. Pattern Anal. Mach. Intell.* 1979, PAMI-1, 224–227.
10. Dukan, P.; Kovari, A. Cloud-based smart metering system. In *Proceedings of the 2013 IEEE 14th International Symposium on Computational Intelligence and Informatics (CINTI), Budapest, Hungary, 19–21 November 2013*; pp. 499–502.
11. Esenturk, E.; Wallace, A.G.; Khastgir, S.; Jennings, P. Identification of Traffic Accident Patterns via Cluster Analysis and Test Scenario Development for Autonomous Vehicles. *IEEE Access* 2022a, 10, 6660–6675.
12. Esenturk, E.; Turley, D.; Wallace, A.; Khastgir, S.; Jennings, P. A data mining approach for traffic accidents, pattern extraction and test scenario generation for autonomous vehicles. *Int. J. Transp. Sci. Technol.* 2022b, in press, corrected proof.

13. Ester, M.; Kriegel, H.-P.; Sander, J.; Xu, X. A density-based algorithm for discovering clusters in large spatial databases with noise. In Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD'96), Portland, OR, USA, 2–4 August 1996.
14. European Parliament. EU Road Safety Policy Framework 2021–2030—Recommendations on Next Steps Towards “Vision Zero”, P9_TA(2021)0407, Text Adopted. 6 October 2021. Available online: https://www.europarl.europa.eu/doceo/document/TA-9-20_21-0407_EN.pdf (accessed on 25 March 2022).
15. Felzenszwalb, P.F.; Huttenlocher, D.P. Efficient Graph-Based Image Segmentation. *Int. J. Comput. Vis.* 2004, 59, 167–181.
16. Gnjatović, M.; Košanin, I.; Maček, N.; Joksimović, D. Clustering of Road Traffic Accidents as a Gestalt Problem. *Applied Sciences*, Vol. 12, No. 9, 4543.
17. Gnjatović, M.; Maček, N.; Adamović, S. A non-connectionist two-stage approach to digit recognition in the presence of noise. In Proceedings of the 10th IEEE International Conference on Cognitive Infocommunications (CogInfoCom), Naples, Italy, 23–25 October 2019; pp. 15–20.
18. Gnjatović, M.; Tasevski, J.; Borovac, B.; Maček, N. An entropy-based approach to automatic detection of critical changes in human-machine interaction. In Proceedings of the 9th IEEE International Conference on Cognitive Infocommunications (CogInfoCom), Budapest, Hungary, 22–24 August 2018; pp. 175–178.
19. Gutierrez-Osorio, C.; Pedraza, C.; Modern data sources and techniques for analysis and forecast of road accidents: A review, *Journal of Traffic and Transportation Engineering (English Edition)*, Volume 7, Issue 4, 2020; pp. 432-446.
20. Hernández, H.; Alberdi, E.; Pérez-Acebo, H.; Álvarez, I.; García, M.J.; Eguia, I.; Fernández, K. Managing Traffic Data through Clustering and Radial Basis Functions. *Sustainability* 2021, 13, 2846.
21. Islam, M.R.; Jenny, I.J.; Nayon, M.; Islam, M.R.; Amiruzzaman, M.; Abdullah-Al-Wadud, M. Clustering algorithms to analyze the road traffic crashes. In Proceedings of the 2021 International Conference on Science & Contemporary Technologies (ICSCT), Dhaka, Bangladesh, 5–7 August 2021.
22. Jeong, H.; Kim, I.; Han, K.; Kim, J. Comprehensive Analysis of Traffic Accidents in Seoul: Major Factors and Types Affecting Injury Severity. *Appl. Sci.* 2022, 12, 1790.

23. Košanin, I.; Gnjatović, M.; Maček, N.; Joksimović, D.; 2023. A Clustering-Based Approach to Detecting Critical Traffic Road Segments in Urban Areas. *Axioms*. Vol. 12, No. 6, 509. 2023.
24. Li, Y.; Huang, M. Identification of Critical Road Links Based on Static and Dynamic Features Fusion. *Appl. Sci.* 2023, 13, 5994.
25. Lilhore, U.K.; Imoize, A.L.; Li, C.-T.; Simaiya, S.; Pani, S.K.; Goyal, N.; Kumar, A.; Lee, C.-C. Design and Implementation of an ML and IoT Based Adaptive Traffic-Management System for Smart Cities. *Sensors* 2022, 22, 2908.
26. Lloyd, S. Least squares quantization in PCM. *IEEE Trans. Inf. Theory* 1982, 28, 129–137.
27. Manap, N.; Borhan, M.N.; Yazid, M.R.M.; Hambali, M.K.A.; Rohan, A. Identification of Hotspot Segments with a Risk of Heavy-Vehicle Accidents Based on Spatial Analysis at Controlled-Access Highway. *Sustainability* 2021, 13, 1487.
28. Ministry of Interior, Republic of Serbia. Camera Locations within Belgrade (In Serbian). Available online: <http://www.mup.gov.rs/wps/wcm/connect/b152c15f-16eb-47b3-b9a4-c7f32c2cc1ba/Lokacij+Bg.pdf?MOD=AJPERES&CVID=n-sczZB> (accessed on 25 March 2022).
29. Ministry of Interior, Republic of Serbia. List of Locations of Video Surveillance System Camera Sites in the City of Belgrade. Available online: <http://www.mup.rs/wps/wcm/connect/56a5cf77-df71-440a-a5bd-0d5f92ec8336/lat-Tabela-prelazi.pdf?MOD=AJPERES&CVID=ng1rX1> (accessed on 12 March 2023). (In Serbian).
30. Niu, Z.; Wang, Y.; Sun, S. Correlation Analysis of Traffic Accident Factors based on Mean Clustering. In *ICCSIE '22, Proceedings of the 7th International Conference on Cyber Security and Information Engineering, Brisbane Australia, 23–25 September 2022*; Association for Computing Machinery: New York, NY, USA, 2022; pp. 569–575.
31. National Assembly of the Republic of Serbia. Law on Road Traffic Safety. *Off. Gaz. Repub. Serb.* no. 41/2009-3, 53/2010-12, 101/2011-270, 32/2013-22 (decision of the Constitutional Court), 55/2014-61, 96/2015-106 (other law), 9/2016-178 (decision of the Constitutional Court), 24/2018-70, 41/2018-122, 41/2018-32 (other law), 87/2018-26, 23/2019-3, 128/2020-3 (other law). Available online: <http://www.pravno-informacioni-sistem.rs/SlGlasnikPortal/eli/rep/sgrs/skupstina/zakon/2009/41/1/reg/20201026> (accessed on 1 March 2022).

32. Osbourn, G.C.; Martinez, R.F. Empirically defined regions of influence for clustering analyses. *Pattern Recognit.* 1995, 28, 1793–1806.
33. Rampinelli, A.; Calderón, J.F.; Blazquez, C.A.; Sauer-Brand, K.; Hamann, N.; Nazif-Munoz, J.I. Investigating the Risk Factors Associated with Injury Severity in Pedestrian Crashes in Santiago, Chile. *Int. J. Environ. Res. Public Health* 2022, 19, 11126.
34. Republic of Serbia. Data on Traffic Accidents for 2021 for the Territory of all Police Administrations and Municipalities. Available online: <https://data.gov.rs/s/resources/podatsi-o-saobratshajnim-nezgodama-po-politsijskim-upravama-i-opshtinama/20220125-085458/nez-opendata-2021-20220125.xlsx> (accessed on 1 March 2022).
35. Republic of Serbia. Data on Traffic Accidents for 2020 for the Territory of all Police Administrations and Municipalities. Available online: <https://data.gov.rs/s/resources/podatsi-o-saobratshajnim-nezgodama-po-politsijskim-upravama-i-opshtinama/20210125-210208-095135/nez-opendata-2020-20210125.xlsx> (accessed on 1 March 2022).
36. Republic of Serbia. Data on Traffic Accidents for 2019 for the Territory of all Police Administrations and Municipalities. Available online: <https://data.gov.rs/s/resources/podatsi-o-saobratshajnim-nezgodama-po-politsijskim-upravama-i-opshtinama/20200125-200127-133136/nez-opendata-2019-20200125.xlsx> (accessed on 1 March 2022).
37. Rousseeuw, P.J. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* 1987, 20, 53–65.
38. Santos, D.; Saias, J.; Quaresma, P.; Nogueira, V.B. Machine Learning Approaches to Traffic Accident Analysis and Hotspot Prediction. *Computers* 2021, 10, 157.
39. Satopää, V.; Albrecht, J.; Irwin, D.; Raghavan, B. Finding a “Kneedle” in a Haystack: Detecting Knee Points in System Behavior. In *ICDCSW ’11, Proceedings of the 2011 31st International Conference on Distributed Computing Systems Workshops*, Washington, DC, USA, 20–24 June 2011; Association for Computing Machinery: Minneapolis, MN, USA, 2011; pp. 166–171.
40. Satria, R.; Castro, M.; GIS TOOLS FOR ANALYZING ACCIDENT AND ROAD DESIGN: A REVIEW, XII Conference on Transport Engineering. CIT 2016, Valencia, Spain.

41. Selim, S.Z.; Ismail, M.A. K-means-type algorithms: A generalized convergence theorem and characterization of local optimality. *IEEE Trans. Pattern Anal. Mach. Intell.* 1984, PAMI-6, 81–87.
42. Shah, M.A.; Zeeshan Khan, F.; Abbas, G.; Abbas, Z.H.; Ali, J.; Aljameel, S.S.; Khan, I.U.; Aslam, N. Optimal Path Routing Protocol for Warning Messages Dissemination for Highway VANET. *Sensors* 2022, 22, 6839.
43. Shang, Q.; Yu, Y.; Xie, T. A Hybrid Method for Traffic State Classification Using K-Medoids Clustering and Self-Tuning Spectral Clustering. *Sustainability* 2022, 14, 11068.
44. Shih, F.Y. *Image Processing and Pattern Recognition: Fundamentals and Techniques*; Wiley-IEEE Press: Hoboken, NJ, USA, 2010.
45. Sinnott, R.W. Virtues of the Haversine. *Sky Telesc.* 1984, 68, 158–159.
46. Sun, Y.; Wang, Y.; Yuan, K.; Chan, T.O.; Huang, Y. Discovering Spatio-Temporal Clusters of Road Collisions Using the Method of Fast Bayesian Model-Based Cluster Detection. *Sustainability* 2020, 12, 8681.
47. Tibshirani, R.; Walther, G.; Hastie, T. Estimating the number of clusters in a data set via the gap statistic. *J. R. Stat. Soc. Ser. B* 2001, 63, 411–423.
48. Wang, Z.; Li, B. A two-stage approach to saliency detection in images. In *Proceedings of the 2008 IEEE International Conference on Acoustics, Speech and Signal Processing*, Las Vegas, NV, USA, 31 March–4 April 2008; pp. 965–968.
49. Wang, D.; Huang, Y.; Cai, Z. A two-phase clustering approach for traffic accident black spots identification: Integrated GIS-based processing and HDBSCAN model. *Int. J. Inj. Control. Saf. Promot.* 2023, published online.
50. Wertheimer, M. Laws of organization in perceptual forms. In *A Source Book of Gestalt Psychology*; Ellis, W.D., Kegan, P., Eds.; Trench, Trubner & Company: London, UK, 1938.
51. Wu, J.; Zhang, L. Gestalt saliency: Salient region detection based on Gestalt principles. In *Proceedings of the the 2013 IEEE International Conference on Image Processing*, Melbourne, VIC, Australia, 15–18 September 2013; pp. 181–185.
52. Zang, J.; Jiao, P.; Liu, S.; Zhang, X.; Song, G.; Yu, L. Identifying Traffic Congestion Patterns of Urban Road Network Based on Traffic Performance Index. *Sustainability* 2023, 15, 948.
53. Zhang, Y.; Ye, N.; Wang, R.; Malekian, R. A Method for Traffic Congestion Clustering Judgment Based on Grey Relational Analysis. *ISPRS Int. J. Geo-Inf.* 2016, 5, 71.

54. Zhao, Q.; Hautamaki, V.; Fränti, P. Knee Point Detection in BIC for Detecting the Number of Clusters. In *Advanced Concepts for Intelligent Vision Systems (ACIVS 2008)*; Blanc-Talon, J., Bourennane, S., Philips, W., Popescu, D., Scheunders, P., Eds.; *Lecture Notes in Computer Science*; Springer: Berlin/Heidelberg, Germany, 2008; Volume 5259, pp. 664–673.
55. Zhao, Y.; Guo, X.; Su, B.; Sun, Y.; Zhu, Y. Multi-Lane Traffic Load Clustering Model for Long-Span Bridge Based on Parameter Correlation. *Mathematics* 2023, 11, 274.

Стручна биографија

Иван Кошанин је рођен 5.11.1995. године у Београду.

Основне академске студије информатике и рачунарства уписује 2014. године на Криминалистичко-полицијској академији у Београду. У јуну 2018. године завршава основне студије са просечном оценом 8,76 и стиче звање дипломирани информатичар.

Током основних академских студија, два пута је награђиван од стране министра унутрашњих послова Републике Србије: наградом за најбољег студента треће године информатике и рачунарства и наградом „Сребрна значка“ за најбољег студента прве генерације информатике и рачунарства.

У октобру 2018. године уписује мастер академске студије информатике и рачунарства на Криминалистичко-полицијском универзитету. У октобру 2019. године завршава мастер студије са просечном оценом 9,37 и стиче звање мастер информатичар.

Након завршених основних академских студија информатике и рачунарства на Криминалистичко-полицијској академији, почиње да ради у Министарству унутрашњих послова Републике Србије.

У октобру 2019. године уписује докторске академске студије информатике и рачунарства на Криминалистичко-полицијском универзитету.

Изјава о ауторству

Име и презиме студента докторских студија: Иван Кошанин

Број индекса: 2R1/0001/19

Изјављујем

да је докторска дисертација под насловом:

„Један приступ кластеровану саобраћајних незгода у урбаним срединама“

- резултат сопственог истраживачког рада;
- да дисертација у целини ни у деловима није била предложена за стицање друге дипломе према студијским програмима других високошколских установа;
- да су резултати коректно наведени и
- да нисам кршио/ла ауторска права и користио/ла интелектуалну својину других лица.

У Београду, _____

Потпис студента докторских студија

Изјава о истоветности штампане и електронске верзије докторске дисертације

Име и презиме студента докторских студија: Иван Кошанин

Број индекса: 2R1/0001/19

Студијски програм: Докторске студије информатике

Наслов дисертације: „Један приступ кластеровању саобраћајних незгода у урбаним срединама“

Ментор: проф. др Милан Гњатовић

Изјављујем да је штампана верзија мог докторског рада истоветна електронској верзији коју сам предао/ла ради похрањивања у Дигиталном репозиторијуму Криминалистичко-полицијског универзитета.

Дозвољавам да се објаве моји лични подаци везани за добијање академског назива доктора наука, као што су име и презиме, година и место рођења и датум одбране рада.

Ови лични подаци могу се објавити на мрежним страницама дигиталне библиотеке, у електронском каталогу и у публикацијама Криминалистичко-полицијског универзитета.

У Београду, _____

Потпис студента докторских студија

Изјава о коришћењу

Овлашћујем Универзитетску библиотеку „Светозар Марковић“ да у Дигитални репозиторијум Криминалистичко-полицијског универзитета унесе моју докторску дисертацију под насловом:

„Један приступ кластеровану саобраћајних незгода у урбаним срединама“,

која је моје ауторско дело.

Дисертацију са свим прилозима предао/ла сам у електронском формату погодном за трајно архивирање.

Моју докторску дисертацију похрањену у Дигиталном репозиторијуму Криминалистичко-полицијског универзитета у Београду и доступну у отвореном приступу могу да користе сви који поштују одредбе садржане у одабраном типу лиценце Креативне заједнице (Creative Commons) за коју сам се одлучио/ла.

1. Ауторство (CC BY)

2. Ауторство – некомерцијално (CC BY-NC)

3. Ауторство – некомерцијално – без прерада (CC BY-NC-ND)

4. Ауторство – некомерцијално – делити под истим условима (CC BY-NC-SA)

5. Ауторство – без прерада (CC BY-ND)

6. Ауторство – делити под истим условима (CC BY-SA)

(Молимо да заокружите само једну од шест понуђених лиценци. Кратак опис лиценци је саставни део ове изјаве).

У Београду, _____

Потпис студента докторских студија

1. **Ауторство.** Дозвољаваате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце, чак и у комерцијалне сврхе. Ово је најслободнија од свих лиценци.
2. **Ауторство** – некомерцијално. Дозвољаваате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце. Ова лиценца не дозвољава комерцијалну употребу дела.
3. **Ауторство** – некомерцијално – без прерада. Дозвољаваате умножавање, дистрибуцију и јавно саопштавање дела, без промена, преобликовања или употребе дела у свом делу, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце. Ова лиценца не дозвољава комерцијалну употребу дела. У односу на све остале лиценце, овом лиценцом се ограничава највећи обим права коришћења дела.
4. **Ауторство** – некомерцијално – делити под истим условима. Дозвољаваате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце и ако се прерада дистрибуира под истом или сличном лиценцом. Ова лиценца не дозвољава комерцијалну употребу дела и прерада.
5. **Ауторство** – без прерада. Дозвољаваате умножавање, дистрибуцију и јавно саопштавање дела, без промена, преобликовања или употребе дела у свом делу, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце. Ова лиценца дозвољава комерцијалну употребу дела.
6. **Ауторство** – делити под истим условима. Дозвољаваате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце и ако се прерада дистрибуира под истом или сличном лиценцом. Ова лиценца дозвољава комерцијалну употребу дела и прерада. Слична је софтверским лиценцама, односно лиценцама отвореног кода.