

Modeliranje ponašanja pješaka u blizini pješačkih prijelaza temeljeno na Markovljevim lancima

Starčević Ratković, Ivana

Master's thesis / Diplomski rad

2022

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Mechanical Engineering and Naval Architecture / Sveučilište u Zagrebu, Fakultet strojarstva i brodogradnje**

Permanent link / Trajna poveznica: <https://urn.nsk.hr/urn:nbn:hr:235:204334>

Rights / Prava: [In copyright](#)/[Zaštićeno autorskim pravom.](#)

Download date / Datum preuzimanja: **2024-07-15**

Repository / Repozitorij:

[Repository of Faculty of Mechanical Engineering and Naval Architecture University of Zagreb](#)



SVEUČILIŠTE U ZAGREBU
FAKULTET STROJARSTVA I BRODOGRADNJE

DIPLOMSKI RAD

Ivana Starčević Ratković

Zagreb, 2022.

SVEUČILIŠTE U ZAGREBU
FAKULTET STROJARSTVA I BRODOGRADNJE

DIPLOMSKI RAD

Mentor:

Doc. dr. sc. Branimir Škugor

Studentica:

Ivana Starčević Ratković

Zagreb, 2022.

Izjavljujem da sam ovaj rad izradila samostalno koristeći znanja stečena tijekom studija i navedenu literaturu.

Ovaj rad nastao je zahvaljujući istraživačima s *Institut für Kraftfahrzeuge* u Aachenu. Ovim putem im zahvaljujem za omogućavanje pristupa skupu podataka za obradu.

Zahvaljujem mentoru doc.dr.sc. Branimiru Škugoru na ideji za rad, pruženoj pomoći, savjetima i susretljivosti te asistentici Lei Pavelko, mag. ing. el. tech. inf., na pruženim informacijama o geografskim lokacijama pješačkog prijelaza.

Zahvaljujem i svim svojim kolegicama, kolegama i prijateljima na podršci tijekom cijelog studija. Hvala vam za svaki savjet, svako druženje te za sva zajednička učenja, ali i za susretljivost i razumijevanje za izbjivanje s ponekih socijalnih situacija u zadnjih par tjedana. Vi ste jedan od razloga zbog kojih mi je žao što ovo studiranje nije bar malo dulje trajalo.

Najveće zahvale idu mojoj obitelji na neprestanoj podršci tijekom svih godina studiranja, a posebno mojoj sestri na lektoriranju svih mojih radova, uz pokoji suvišan komentar.

Na kraju, posebno zahvaljujem Jakovu. Hvala ti za svaku skuhanu kavu, *brainstorming* sesiju, pregled svake verzije rada; sveukupno, za nezamislivu podršku. Ovaj rad posvećen je tebi.

Ivana Starčević Ratković



SVEUČILIŠTE U ZAGREBU
FAKULTET STROJARSTVA I BRODOGRADNJE



Središnje povjerenstvo za završne i diplomske ispite
Povjerenstvo za diplomske radove studija strojarstva za smjerove:
proizvodno inženjerstvo, računalno inženjerstvo, industrijsko inženjerstvo i menadžment,
inženjerstvo materijala te mehatronika i robotika

Sveučilište u Zagrebu Fakultet strojarstva i brodogradnje	
Datum:	Prilog:
Klasa:	602-14/22-6/1
Ur. broj:	15-1703-22-

DIPLOMSKI ZADATAK

Student: **IVANA STARČEVIĆ RATKOVIĆ** Mat. br.: 0108073383

Naslov rada na hrvatskom jeziku: **Modeliranje ponašanja pješaka u blizini pješačkih prijelaza temeljeno na Markovljevim lancima**

Naslov rada na engleskom jeziku: **Markov chain-based modelling of pedestrian crossing behavior at unsignalized crosswalks**

Opis zadatka:

Iako su brojni zadaci autonomne vožnje niže razine u dobroj mjeri riješeni, autonomna vožnja u prisutnosti ostalih sudionika u prometu kao što su primjerice pješaci i biciklisti još uvijek predstavlja značajan izazov. Stoga je precizno modeliranje njihovog ponašanja i integracija pripadnih modela u simulacijska okruženja za autonomnu vožnju od presudne važnosti za razvoj autonomnih vozila. U sklopu diplomskog rada stoga je potrebno:

- provesti pregled literature i navesti glavne pristupe u modeliranju ponašanja pješaka u blizini pješačkih prijelaza, te opisati najčešće pokazatelje koji se koriste u pripadnim analizama (TTA, TTAa, PET, i sl.)
- za odabrani skup podataka vezanih uz interakcije vozila i pješaka provesti pred-obradu i analizu s ciljem pripreme podataka za modeliranje ponašanja pješaka te dobivanje osnovnih uvida (kao skup podataka primjerice koristiti javno dostupne inD podatke, <https://www.ind-dataset.com/>)
- modelirati ponašanje pješaka s obzirom na odabrane snimljene podatke korištenjem metode Markovljevih lanaca (pripadnu matricu prijelaznih vjerojatnosti implementirati u obliku rijetke matrice temeljene na rječniku ključeva)
- provesti validaciju izvedenog modela korištenjem Monte Carlo simulacija s obzirom na snimljene podatke
- razmotriti mogućnost primjene dobivenog modela za optimalno upravljanje brzinom vozila koje prilazi pješačkom prijelazu s pješacima.

U radu je potrebno navesti korištenu literaturu i eventualno dobivenu pomoć.

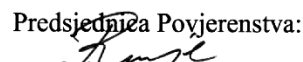
Zadatak zadan:
29. rujna 2022.

Rok predaje rada:
1. prosinca 2022.

Predviđeni datum obrane:
12. prosinca do 16. prosinca 2022.

Zadatak zadao:


doc. dr. sc. Branimir Skugor

Predsjednica Povjerenstva:

prof. dr. sc. Biserka Runje

SADRŽAJ

SADRŽAJ	I
POPIS SLIKA	III
POPIS TABLICA.....	VI
POPIS OZNAKA	VII
POPIS KRATICA	VIII
SAŽETAK.....	IX
SUMMARY	X
1. UVOD.....	1
2. ODABIR SKUPA PODATAKA	5
2.1. Zahtjevi na skup podataka.....	5
2.2. inD skup podataka.....	5
2.2.1. O skupu podataka.....	5
2.2.2. Prednosti skupa	7
2.2.3. Metoda prikupljanja podataka.....	7
2.2.4. Usporedba s postojećim skupovima podataka	8
3. PREDOBRAĐA I ANALIZA SKUPA PODATAKA	10
3.1. Odabir podataka	10
3.1.1. Program za vizualizaciju podataka	10
3.1.2. Odabir promatranih objekata	16
3.2. Reduciranje informacija o objektima	20
3.2.1. Reduciranje iznosa udaljenosti na jednu vrijednost.....	21
3.2.2. Haversinusna formula	23
3.2.3. Udaljenost od pješačkog prijelaza.....	26
3.2.4. Brzina i akceleracija pješaka i vozila.....	30
4. MODELIRANJE PONAŠANJA PJEŠAKA	31
4.1. Markovljevi lanci	31
4.1.1. Definicija Markovljevih lanaca.....	31
4.1.2. Markovljev lanac u obliku rječnika ključeva.....	32

4.1.2.1.	Rječnik ključeva.....	32
4.1.2.2.	Primjena rječnika ključeva za pohranu elemenata matrice.....	34
4.2.	Princip modeliranja ponašanja pješaka u blizini pješačkih prijelaza.....	35
4.2.1.	Odabir stanja za kreiranje modela.....	35
4.2.1.1.	Odabir rezolucija za podatke o pješacima	36
4.2.1.2.	Odabir rezolucija za podatke o vozilima	43
4.2.1.3.	Kreiranje rječnika i krajnji odabir rezolucija varijabli.....	50
4.2.2.	Algoritam za kreiranje modela.....	52
5.	VALIDACIJA IZVEDENOG MODELA	54
5.1.	Monte Carlo simulacija.....	54
5.1.1.	Definicija i povijest.....	54
5.1.2.	Primjena	54
5.1.3.	Python algoritam slučajnog odabira.....	55
5.1.4.	Nasumičnost u simulacijama	56
5.2.	Opis korištenih metoda validacije.....	56
5.3.	Validacija pomoću 50 simulacija nasumičnog početnog stanja.....	58
5.3.1.	Raspodjela vrijednosti dobivenih simulacijama	59
5.3.2.	Grafička usporedba simuliranih i testnih vrijednosti	65
5.4.	Validacija pomoću 100 simulacija predefiniranog početnog stanja	83
5.4.1.	Korijen srednje kvadratne vrijednosti pogreške rezultata.....	83
5.4.2.	Točnost procjene očekivanog ishoda	87
5.4.3.	Vrijeme do dolaska i vrijeme do sudara	89
6.	PRIMJENA MODELA.....	94
7.	ZAKLJUČAK.....	96
	LITERATURA.....	99

POPIS SLIKA

Slika 1.	Raskrižja prikazana u inD datasetu, na lokacijama, redom: Bendplatz, Frankenburg, Heckstrasse, Neuköllner Strasse [5]	6
Slika 2.	Vizualizacija podataka pomoću danog programskog koda	15
Slika 3.	Primjer prikaza dijagrama za objekt rednog broja 49 na snimci 19	16
Slika 4.	Interakcija pješaka i vozila gdje vozilo u bližem traku ne nailazi pravocrtno	19
Slika 5.	Interakcija pješaka i vozila gdje nije moguće sa sigurnošću utvrditi razlog usporavanja vozila	19
Slika 6.	Položaj lokalnog koordinatnog sustava, [https://www.ind-dataset.com/format] ..	22
Slika 7.	Primjer trajektorija u inD skupu podataka, [5]	23
Slika 8.	Udaljenost dvije točke na sferi, [9]	25
Slika 9.	Ishodište novokreiranih koordinatnih sustava uz prikaz pješaka, vozila, pločnika, kolnika i pješačkog prijelaza te smjerova kretanja pješaka i vozila, [11]	26
Slika 10.	Rubovi pješačkog prijelaza	27
Slika 11.	Princip izračuna apsolutne udaljenosti pješaka od pješačkog prijelaza: geografski položaji predstavljaju točke u kojima su poznate geografske koordinate, [11]	28
Slika 12.	Shematski prikaz Markovljevih lanaca: a) prvog reda, b) drugog reda, [12]	32
Slika 13.	Prikaz jednog od mogućih prelazaka unutar Markovljevog lanca sastavljenog od dva stanja	34
Slika 14.	Usporedni prikaz klasičnog prikaza Markovljevog lanca, matrice prijelaznih vrijednosti i pripadajućeg rječnika ključeva	35
Slika 15.	Raspodjela neobrađenih vrijednosti udaljenosti pješaka	37
Slika 16.	Raspodjela udaljenosti pješaka u slučaju rezolucije od 0,5 m	37
Slika 17.	Raspodjela brzina pješaka u slučaju udaljenosti od 0,75 m	38
Slika 18.	Raspodjela udaljenosti pješaka u slučaju rezolucije od 1 m	38
Slika 19.	Raspodjela brzina pješaka	39
Slika 20.	Raspodjela brzina pješaka u slučaju rezolucije od 0,25 m/s	40
Slika 21.	Raspodjela brzina pješaka u slučaju rezolucije od 0,5 m/s	40
Slika 22.	Raspodjela ubrzanja pješaka	41

Slika 23.	Raspodjela ubrzanja pješaka u slučaju rezolucije od $0,25 \text{ m/s}^2$	42
Slika 24.	Raspodjela ubrzanja pješaka u slučaju rezolucije od $0,5 \text{ m/s}^2$	42
Slika 25.	Raspodjela ubrzanja pješaka u slučaju rezolucije od 1 m/s^2	43
Slika 26.	Raspodjela snimljenih udaljenosti vozila	44
Slika 27.	Raspodjela iznosa udaljenosti vozila u slučaju rezolucije od $7,5 \text{ m}$	45
Slika 28.	Raspodjela iznosa udaljenosti vozila u slučaju rezolucije od 10 m	45
Slika 29.	Raspodjela snimljenih brzina vozila.....	46
Slika 30.	Raspodjela iznosa brzine vozila u slučaju rezolucije od 1 m/s	47
Slika 31.	Raspodjela iznosa brzine vozila u slučaju rezolucije od 3 m/s	47
Slika 32.	Raspodjela snimljenih ubrzanja vozila.....	48
Slika 33.	Raspodjela iznosa ubrzanja vozila u slučaju rezolucije od 1 m/s^2	49
Slika 34.	Raspodjela iznosa ubrzanja vozila u slučaju rezolucije od $1,5 \text{ m/s}^2$	49
Slika 35.	Broj vrijednosti po ključu unutar rječnika stanja	52
Slika 36.	Broj vremenskih okvira po interakciji.....	57
Slika 37.	Broj vremenskih okvira po kreiranoj simulaciji.....	59
Slika 38.	Raspodjela vrijednosti udaljenosti pješaka za sve podatke unutar 50 simulacija .	60
Slika 39.	Raspodjela vrijednosti udaljenosti vozila za sve podatke unutar 50 simulacija....	61
Slika 40.	Raspodjela vrijednosti brzine pješaka za sve podatke unutar 50 simulacija	62
Slika 41.	Raspodjela vrijednosti brzine vozila za sve podatke unutar 50 simulacija	63
Slika 42.	Raspodjela vrijednosti ubrzanja pješaka za sve podatke unutar 50 simulacija	64
Slika 43.	Raspodjela vrijednosti ubrzanja vozila za sve podatke unutar 50 simulacija	65
Slika 44.	Prikaz testne putanje pješaka u odnosu na 9 simuliranih putanja	66
Slika 45.	Prikaz testne putanje vozila u odnosu na 9 simuliranih putanja.....	67
Slika 46.	Prikaz testne brzine pješaka u odnosu na 9 simuliranih brzina.....	68
Slika 47.	Prikaz testne brzine vozila u odnosu na 9 simuliranih brzina	70
Slika 48.	Prikaz testnog ubrzanja pješaka u odnosu na 9 simuliranih ubrzanja	71
Slika 49.	Prikaz testnog ubrzanja vozila u odnosu na 9 simuliranih ubrzanja	72
Slika 50.	Prikaz udaljenosti pješaka sa snimke <i>Train 4</i> u odnosu na 4 simulirane udaljenosti	74

Slika 51.	Prikaz udaljenosti vozila sa snimke <i>Train 4</i> u odnosu na 4 simulirane udaljenosti	74
Slika 52.	Prikaz brzine pješaka sa snimke <i>Train 4</i> u odnosu na 4 simulirane brzine.....	75
Slika 53.	Prikaz brzine vozila sa snimke <i>Train 4</i> u odnosu na 4 simulirane brzine	75
Slika 54.	Prikaz ubrzanja pješaka sa snimke <i>Train 4</i> u odnosu na 4 simulirana ubrzanja ...	76
Slika 55.	Prikaz ubrzanja vozila sa snimke <i>Train 4</i> u odnosu na 4 simulirana ubrzanja.....	76
Slika 56.	Usporedbeni prikaz udaljenost pješaka i vozila snimke <i>Test 4</i>	77
Slika 57.	Usporedbeni prikaz brzine pješaka i vozila snimke <i>Test 4</i>	78
Slika 58.	Usporedbeni prikaz ubrzanja pješaka i vozila snimke <i>Test 4</i>	78
Slika 59.	Usporedbeni prikaz udaljenosti pješaka i vozila <i>Simulacije 16</i>	79
Slika 60.	Usporedbeni prikaz brzine pješaka i vozila <i>Simulacije 16</i>	80
Slika 61.	Usporedbeni prikaz ubrzanja pješaka i vozila <i>Simulacije 16</i>	80
Slika 62.	Usporedbeni prikaz udaljenosti pješaka i vozila <i>Simulacije 29</i>	81
Slika 63.	Usporedbeni prikaz brzine pješaka i vozila <i>Simulacije 29</i>	82
Slika 64.	Usporedbeni prikaz ubrzanja pješaka i vozila <i>Simulacije 29</i>	82
Slika 65.	RMSE za svih 6 stanja.....	85
Slika 66.	Snimljene i snimulirane vrijednosti TTA-a	92
Slika 67.	Snimljene i simulirane vrijednosti TTA (bez iznosa jednakih ∞).....	92
Slika 68.	Apsolutna greška TTA po snimci.....	93

POPIS TABLICA

Tablica 1. Statističke informacije o lokacijama snimanja (kratica VRU proizlazi iz <i>Vulnerable Road Users</i> i odnosi se na pješake i bicikliste).	9
Tablica 2. Sadržaj datoteka <i>Recording Meta Information</i>	11
Tablica 3. Sadržaj datoteka <i>Track Meta Information</i>	12
Tablica 4. Sadržaj datoteka <i>Tracks</i>	12
Tablica 5. Analiza broja interakcija pješaka i vozila.	17
Tablica 6. Identifikacijske oznake odabranih interakcija pješaka i vozila.	20
Tablica 7. Identifikacijske oznake odabranih interakcija pješaka i vozila.	27
Tablica 8. Rezolucije podataka unutar različitih setova.	50
Tablica 9. Oznake simulacija po rednom broju.	56
Tablica 10. Numeracija interakcija po rednom broju.	57
Tablica 11. Srednji RMSE po svakoj varijabli.	86
Tablica 12. Numeracija interakcija po rednom broju.	88
Tablica 13. Usporedba snimljenih i simuliranih TTA.	91

POPIS OZNAKA

Oznaka	Mjerna jedinica	Opis oznake
a_{pd}	m/s^2	ubrzanje pješaka
a_v	m/s^2	ubrzanje vozila
d	m	udaljenost neke dvije točke po duljini glavne kružnice sfere
d_{pd}	m	udaljenost pješaka od pješačkog prijelaza
d_v	m	udaljenost vozila od pješačkog prijelaza
h	-	haversinus
n	-	broj stanja, podataka
p_{ij}	-	vjerojatnost prelaska iz stanja i u stanje j
P	-	vjerojatnost
\mathbf{P}	-	matrica prijelaznih vjerojatnosti
r	m	polumjer sfere
s	m	udaljenost
s_i	-	moгуće stanje i
S	-	domena mogućih stanja
TTA	s	vrijeme do dolaska
v	m/s	brzina
v_{pd}	m/s	brzina pješaka
v_v	m/s	brzina vozila
x	m	vrijednost na x osi kartezijevog koordinatnog sustava
x_i	-	stvarni iznos i -te varijable
x_m	-	prosječni iznos varijable
$x_{p,i}$	-	predviđeni iznos varijable u i -tom trenutku
X_i	-	stvarno stanje i
X	-	domena, tj. lanac stvarnih stanja
y	-	vrijednost na y osi kartezijevog koordinatnog sustava
Δv_{pd}	m/s	razlika u brzini pješaka
Δv_v	m/s	razlika u brzini vozila
ϑ	$^\circ, \text{rad}$	središnji kut sfere
φ, λ	$^\circ$	geografska širina i duljina
σ	-	standardna devijacija
σ^2	-	varijanca

POPIS KRATICA

Kratika	Opis
CITR	- Odjel za istraživanje upravljanja i inteligentnog transporta (eng. <i>Control and Intelligent Transportation Reserach</i>)
CSV	- format datoteke – zarezom odvojene vrijednost (eng. <i>Comma Separated Values</i>)
DUT	- Sveučilište Dalian (eng. <i>Dalian University of Technology</i>)
GPS	- globalni sutav pozicioniranja (eng. <i>Global Positioning System</i>)
ID	- identifikacijska oznaka (eng. <i>Identification</i>)
ika	- Institut za automobilsko inženjerstvo (njem. <i>Institut für Kraftfahrzeuge</i>)
IMU	- uređaj za mjerenje ubrzanja (eng. <i>Inertial Measurement Unit</i>)
inD	- skup podataka snimljenih dronom na raskrižju (eng. <i>Intersection Drone Dataset</i>)
ISO	- Međunarodna organizacija za standardizaciju (eng. <i>International Organization for Standardization</i>)
JAAD	- baza podataka za promatranje raskrižja u kontekstu autonombne vožnje (eng. <i>Joint Attention in Autonomous Driving</i>)
LiDAR	- optički mjerni instrument (eng. <i>Light Detection and Ranging</i>)
MCMC	- Monte Carlo simulacija na temelju Markovljevih lanaca (eng. <i>Markov Chain Monte Carlo simulation</i>)
NLP	- neurolingvističko programiranje (eng. <i>Natural Language Processing</i>)
PD	- pješak (eng. <i>Pedestrian</i>)
RMSE	- korijen srednje kvadratne pogreške (eng. <i>Root Mean Square Error</i>)
SI	- internacionalni sustav mjernih jedinica (fr. <i>Système International d'Unités</i>)
TTA	- vrijeme do dolaska (eng. <i>Time To Arrive</i>)
TTC	- vrijeme do sudara (eng. <i>Time To Collision</i>)
UTM	- poprečna Mercatorova projekcija (eng. <i>Universal Transverse Mercator</i>)
VRU	- oznaka (ranjivih) sudionika u prometu (eng. <i>Vulnerable Road Users</i>)

SAŽETAK

Iako su brojni zadaci autonomne vožnje niže razine u dobroj mjeri riješeni, autonomna vožnja u prisutnosti ostalih sudionika u prometu, kao što su primjerice pješaci i biciklisti, još uvijek predstavlja značajan izazov. Stoga je precizno modeliranje njihovog ponašanja i integracija pripadnih modela u simulacijska okruženja za autonomnu vožnju od presudne važnosti za razvoj autonomnih vozila. U sklopu diplomskog rada obrađena je problematika modeliranja ponašanja pješaka u blizini pješačkih prijelaza temeljena na snimljenim podacima. Pritom je proveden pregled literature na tu temu te su izdvojeni najčešći pokazatelji koji se koriste u pripadnim analizama. Za potrebe modeliranja i dobivanje osnovnih uvida odabran je javno dostupan skup snimljenih podataka koji sadrži značajan broj interakcija vozila i pješaka. Modeliranje ponašanja pješaka s obzirom na odabrane snimljene podatke provedeno je korištenjem metode Markovljevih lanaca, gdje je pripadna matrica prijelaznih vjerojatnosti implementirana u obliku rijetke matrice temeljene na rječniku ključeva u svrhu smanjenja zahtjeva na memoriju i poboljšanje računalne učinkovitosti. Provedena je i validacija izvedenog modela korištenjem Monte Carlo simulacija s obzirom na snimljene podatke. Naposljetku je razmotrena mogućnost primjene dobivenog modela za optimalno upravljanje brzinom vozila koje prilazi pješačkom prijelazu s pješacima te je iznesen zaključak i mogućnost dodatnog usavršavanja modela.

Ključne riječi: Markovljevi lanci, modeliranje, Monte Carlo simulacije, pješaci

SUMMARY

Although many tasks of lower-level autonomous driving have been solved to a good extent, autonomous driving in the presence of other road users such as pedestrians and cyclists still represents a significant challenge. Therefore, precise modeling of their behavior and integration of related models into simulation environments for autonomous driving is of crucial importance for the development of autonomous vehicles. As part of the thesis, the issue of modeling pedestrian behavior near pedestrian crossings based on recorded data was addressed. At the same time, a review of the literature on the subject was carried out, and the most common indicators used in related analyzes were singled out. For the purposes of modeling and obtaining basic insights, a publicly available set of recorded data containing a significant number of vehicle-pedestrian interactions was selected. Pedestrian behavior modeling with regard to selected recorded data was carried out using the Markov chain method, where the associated matrix of transition probabilities was implemented in the form of a sparse matrix based on a dictionary of keys in order to reduce memory requirements and improve computing efficiency. Validation of the derived model was also carried out using Monte Carlo simulations with respect to recorded data. Finally, the possibility of applying the obtained model for optimal control of the speed of vehicles approaching a crosswalk with pedestrians was considered, and the conclusion as well as the possibility of additional refinement of the model were presented.

Key words: Markov chains, modeling, Monte Carlo simulations, pedestrians

1. UVOD

Kao jedan od pravaca u kojem se kreće razvoj tehnologije automobila, svakako se smatra razvoj djelomično ili potpuno autonomnih vozila. Sama ideja autonomije u vožnji potječe čak i iz 1920-ih, dok je prvi konkretan rezultat prikazan u Japanu 1977. [1] Japansko poluautonomno vozilo sadržavalo je analogno računalo i dvije kamere za definiranje pozicije vozila. Takvo vozilo moglo se kretati samo unutar označenih cestovnih linija. Današnja tehnologija koja se primjenjuje u autonomnoj vožnji je puno naprednija i puno šira – vozila koriste razne senzore, poput kamera, LiDAR-a, radara, GPS-a, IMU-a, a informacije sa svih se sjedinjuju za potpun prikaz informacije o položaju vozila i njegovoj okolini. Ograničenje u vidu korištenja senzora na vozilu svakako su zahtjevi na računalo, u vidu brzine kojom se temeljem informacija sa senzora dobiva povratna informacija o upravljanju vozilom. No, skupa tehnologija priječi komercijalizaciju autonomnih vozila.

Kako bi se procijenila točnost modela za korištenje pri bilo kojem aspektu autonomne vožnje, neophodno je i zakonski obavezno obaviti velik broj testiranja. Testiranja je potrebno obaviti za velik broj jedinstvenih uvjeta na točno određen način. Ona mogu biti dugotrajna, a razlozi za isto su razni. Jedan od njih svakako je čekanje na željene uvjete. Također, pri testiranju u stvarnim uvjetima pojavljuje se popriličan broj zakonskih i etičkih pitanja – smije li se snimanje obavljati na određenoj lokaciji u određeno vrijeme, je li za isto potrebna jedna ili više dozvola, od koga je potrebno zatražiti dozvolu, postoji li ikakva mogućnost ugrožavanja ostalih sudionika u prometu? Nadalje, sva snimanja u stvarnom okruženju kao prioritet moraju osigurati sigurnost sudionika u prometu, čime je nemoguće dobiti snimke situacija kada, na primjer, vozilo nailazi na grupu pješaka velikom brzinom ili kada pješak izlijeće pred vozilo koje se kreće određenom brzinom. U potrebi za takvim testiranjima opasnim za pješaka, koriste se lutke koje reprezentiraju ljude (eng. *dummy*). One su ljudima slične samo po izgledu te za neke naprednije sustave prepoznavanja nisu korisne – korištenje lutaka ne daje klasične ljudske pokrete niti mogućnost neverbalne komunikacije.

Zakonske regulative poprilično su oštre po pitanju dopuštenja za testiranje autonomnih ili sličnih vozila na javnim površinama. Testiranja autonomnih vozila na cestama su skupa i zahtjevnija za izvedbu zbog količine zakonskih propisa koje je potrebno ispoštovati te dobivanja dozvole za puštanje vozila u pogon kada nije sa sto-postotnom sigurnošću moguće reći što će

biti idući korak koji će vozilo poduzeti. Samim time nepovjerljivost javnosti je prema toj novoj tehnologiji velika i svaka vijest o neuspjehu stvarnog testiranja ima velik odjek u javnosti. Osim prethodno navedenih ograničenja, značajan problem predstavlja činjenica da vozilo može djelovati adekvatno unutar nekog testnog okruženja, no ne i unutar svih ostalih, točnije, nije moguće sa sigurnošću reći da će vozilo imati željeno ponašanje izvan kontroliranih uvjeta. Primjerice, vozni uvjeti u Europi, gdje je vozilo testirano, neće se nužno poklapati s voznim uvjetima, na primjer, u Aziji.

Zbog navedenih razloga, ali i mnogih drugih, logičan izbor je virtualno testiranje vozila.

Virtualno testiranje autonomnih vozila pomoću simulatora naširoko je usvojeno. Najčešće podrazumijeva kreiranje modela ponašanja objekta te promatranje i korigiranje njegovog ponašanja na skupu podataka. Skup podataka može biti stvaran ili umjetno kreiran. Tehnologija se uglavnom razvija i testira u virtualnom simulacijskom okruženju, što postavlja velike zahtjeve na sama simulacijska ograničenja koja moraju čim vjernije reproducirati uvjete iz stvarne vožnje.

Kao jednim od najnepredvidljivijih faktora u okolini vozila smatraju se drugi sudionici u prometu kao što su pješaci i biciklisti. Njihovo kretanje, iako je po prometnim propisima ograničeno na točno određene površine, često podrazumijeva kretanja na neočekivanim i cestovnim površinama te prijelaze preko ceste preko neoznačenih dijelova. Kretanje pješaka je često nepredvidivo te se ne može sa potpunom sigurnošću predvidjeti čak ni temeljem skupa informacija dobivenih o pješaku unutar većeg broja prethodnih vremenskih trenutaka. Na kretanje mogu utjecati razni faktori kao što su dob i spol osobe, osobna percepcija prostora, veličina grupe u kojoj se nalazi pješak, razina pažnje, učestalost imitacije ponašanja ostalih sudionika u prometu i tako dalje. [2] Pješaci su, također, najranjivija skupina sudionika u prometu te se njihovoj sigurnost pridaje najveći značaj.

Postoje brojna istraživanja vezana uz ponašanje pješaka pješaka u interakciji s nadolazećim vozilom u blizini pješačkog prijelaza. Primjerice, u slučaju autonomnih vozila, očekuje se dulje čekanje pješaka na prelazak preko pješačkog prijelaza radi veće razine nepovjerenja. U slučaju klasičnih vozila s vozačem, često se odluka o prelasku temelji na neverbalnoj komunikaciji, kao što je očni kontakt ili mahanje ruke (također ovisno o kulturi).

Najtočniji skup podataka za virtualno testiranje, to jest, skup na temelju kojeg bi se unificiralo ponašanje pješaka, svakako bi bio onaj u kojem pješak ne bi znao razlikovati je li vozilo klasično ili autonomno. Isto tako, s obzirom na navedeno, za učenje autonomnih vozila korisno

je iskorištavati skupove podataka koji uključuju interakciju pješaka i konvencionalnog vozila jer su u tim slučajevima pješaci neoprezniji i ponašanje im je sličnije svakodnevnom (kontrast dodatnom oprezu uslijed nepovjerenja prema autonomnim vozilima).

Nažalost, broj javno dostupnih skupova podataka temeljem kojih bi se mogle analizirati značajke prijelaza pješaka preko pješačkog prijelaza u ovisnosti o nadolazećem vozilu nije velik. Jedan od takvih skupova naziva se JAAD skup podataka (eng. *Joint Attention in Autonomous Driving (JAAD) Dataset*). [3] On sadrži podatke o interakcijama pješaka i vozila u urbanim i ruralnim sredinama. U njemu pješaci prelaze pješačke prijelaze individualno ili u grupama. Iz razloga što je snimanje obavljeno iz vozila, pješaci su često zaklonjeni raznim objektima. Skup podataka analiziran je u više istraživanja, a većina njih odnosi se na psihologiju koja pogoni odluke pješaka, u vidu gestikulacija, duljeg ili kraćeg gledanja prije prelaska ceste, usporavanja, kimanja glavom i sl. Ovaj skup podataka kreiran je primarno za svrhe psihološke analize ponašanja pješaka.

Zasada najveći skup podataka koji se odnosi na ponašanje pješaka u prometu je PIE skup podataka (eng. *Pedestrian Intention Estimation Dataset*). Sadrži više od 6 sati snimaka snimljenih kamerom montiranom na vozilo. Uz to, pruža informacije o vozilu kao što su njegova brzina, smjer kretanja i geografske koordinate te informacije o prometnoj infrastrukturi. Specifičan je po tome što daje opširne informacije o ponašanju pješaka. Sadrži informacije o 1842 pješaka. [4]

Kako bi se riješilo pitanje ponašanja pješaka pri nailasku vozila, cilj ovog rada je kreiranje modela koji prikazuje ponašanje pješaka u blizini pješačkog prijelaza kada na njega pravocrtno nailazi vozilo u pješaku bližem voznom traku, pomoću Markoljevih lanaca. Rad se bazira na javno dostupnom skupu podataka pod nazivom *inD Dataset (Intersection Drone Dataset)*, koji sadrži velik broj snimljenih podataka. [5] Položaji pješaka i vozila te njihove brzine i ubrzanja su jasno dani za svaki vremenski okvir. Uz javno dostupne informacije, dan je i programski alat za vizualizaciju podataka. Svi podaci snimljeni su bespilotnom letjelicom (dronom) iz zraka te se na taj način nikako nije utjecalo na ponašanje sudionika u prometu. Dron je cijelo snimanje odradio na istom mjestu te je time omogućio da se podaci o pješacima i vozilima jednostavno uspoređuju jer se nalaze u istom okružju i u istim okolišnim uvjetima. Sljedeći korak u radu je analiza skupa podataka te uklanjanje suvišnih podataka koji se ne odnose na interakciju pješak - vozilo. Skup podataka snimljen u nekom vremenskom trenutku te prilagođen uspoređivanju s preostalim podacima unutar drugih vremenskih trenutaka smatra se stanjem iz kojeg se prelazi

u neko drugo ili neka druga stanja te stanje u koje se dolazi. Ti skupovi stanja temelj su za kreiranje Markovljevih lanaca. Važno je napomenuti da je vremenski razmak između dva stanja uvijek jednak. Markovljevi lanci međusobno povezuju sva stanja nekog događaja na način da povezuju trenutno i buduće stanje (ili više mogućih budućih stanja) uz dan iznos vjerojatnosti za prelazak u to stanje. Također, moguće je ne prijeći u drugo stanje, nego ostati u istom. Specifičnost Markovljevih lanaca je u tome što se odnos između dva stanja gleda samo na temelju utjecaja prethodnog stanja na trenutno ili trenutnog na buduće.

Pri kreiranju Markovljevih stanja za prikaz vjerojatnosti koristi se matrica prijelaznih vrijednosti (*eng. Transition Probability Matrix, TPM*). Implementacija takve matrice ovdje je zamijenjena rječnikom (struktura podataka dostupna u programskom jeziku *Python*), kako bi se smanjila računalna zahtjevnost te ubrzala simulacija jedne šetnje pješaka od početnog do krajnjeg stanja. Svi postupci za analizu podataka, kreiranje rječnika i simulacija kreirani su pomoću programskog jezika *Python*. Opisano je više vrsta metoda validacije dobivenih simulacija te je zaključno dana ocjena kvalitete modela uz mogućnost korištenja dobivenog modela za optimalno upravljanje brzinom vozila te su dane ideje o mogućnostima dodatnog razvoja modela.

Neke od validacija određenih primjena modela rezultirale su značajnom točnošću te je time pokazano da se razmatrana metoda modeliranja može koristiti u slične svrhe za druge skupove podataka. U obzir se uzima samo apsolutna udaljenost od pješačkog prijelaza, te pravocrtnost voznog traka u kojem se nalazi nailazeće vozilo. Također, za primjenu ovog modela s istim rezultatima, potrebno je promatrati interakciju samo jednog pješaka sa samo jednim vozilom.

2. ODABIR SKUPA PODATAKA

2.1. Zahtjevi na skup podataka

Kao što je navedeno u pregledu literature, najčešći tip prikupljanja podataka, onaj sa snimanjem iz vozila, često negativno utječe na rezultate jer su sudionici u prometu uglavnom svjesni da ih se snima i promatra, što uvelike utječe na njihovo ponašanje u prometu te se takvo ponašanje tada ne može kategorizirati kao uobičajeno ili standardno. Na drugu stranu, takav način snimanja omogućava pogodnosti kao što su primjećivanje izraza lica ili točno definiranje pokreta pješaka. Velik problem je i manjak preglednosti kod snimanja iz vozila, potencijalnoj zaklonjenosti subjekata, manjku informacija o duljini vozila, prethodnom kretanju promatranog sudionika u prometu te mnogi drugi.

Kako bi se izbjegle prethodno navedene i još poneke pogreške, za ovaj rad je odabran *inD* skup podataka (eng. *inD dataset*, gdje kratica *inD* predstavlja *Intersection Drone Dataset*). U odnosu na prethodne skupove podataka, u ovom su skupu podataka snimke napravljene iz zraka, dronom, iz ptičje perspektive s visine veće od 60 m. Time je dobivena preglednost cijelog raskrižja (željene lokacije snimanja), a samo prikupljanje podataka nije utjecalo na ponašanje sudionika u prometu. Snimanje dronom podacima daje dodatnu vrijednost iz razloga što je lakše staviti fokus promatranja samo na objekte, jer je okolina u svakom trenutku jednaka te ju je jednostavno eliminirati. [5]

2.2. inD skup podataka

2.2.1. O skupu podataka

inD skup podataka rezultat je rada istraživača s Aachenskog Instituta za automobilsko inženjerstvo (*ika*, njem. *Institut für Kraftfahrzeuge*). Sadrži 10 sati snimki s četiri različita raskrižja njemačkog grada Aachena, kako je prikazano na Slika 1. Skup podataka daje putanje za 13 599 objekata, to jest, sudionika u prometu, uključujući vozila, pješake i bicikliste (pod kategoriju biciklista spadaju i motociklisti).

Unutar skupa podataka postoje informacije o 8233 vozila i 5366 preostalih sudionika u prometu. Sva snimanja su obavljena između 2017. i 2019. godine, a podrazumijevala su prethodno navedena četiri raskrižja, s time da im je zajednički faktor svakako činjenica da ne sadrže prometnu signalizaciju (semafore), da je ograničenje brzine 50 km/h te da sadrže pločnik za nesmetano kretanje pješaka. [5]



Slika 1. Raskrižja prikazana u inD datasetu, na lokacijama, redom: Bendplatz, Frankenburg, Heckstrasse, Neuköllner Strasse [5]

Ovaj skup podataka kreiran je s ciljem poticanja istraživanja na temu sigurnosti u autonomnoj vožnji, kreiranja modela prometa, analize prometa, analize ponašanja vozača i ostalih sudionika u prometu. Namijenjen je istraživanjima te je time otvorenog tipa, što znači da mu je moguće pristupiti uz dopuštenje. *inD* skup podataka je odabran za ovaj rad jer je najopsežniji javno dostupan skup tog tipa, uz iznimnu točnost određivanja položaja sudionika u prometu. Najvažnije, sadrži i snimku na kojoj se nalazi pješački prijelaz bez signalizacije, što je temelj istraživanja ovog rada. Samim time, raskrižje odabrano za razmatranje u ovom radu je raskrižje u Frankenburgu, iz razloga što je jedino koje sadrži pješački prijelaz (gornja desna pod-slika na Slici 1). Nalazi se blizu središta grada, a nedaleko je i park te velika količina stambenih zgrada zbog čega se na snimci pojavljuje velik broj pješaka i biciklista. Interakcija pješaka i vozila događa se samo na pješačkim prijelazima, dok se interakcija vozila i biciklista odvija i na pješačkim prijelazima i na raskrižju.

2.2.2. Prednosti skupa

inD skup podataka opisan je u radu iz 2020. godine. [5] U istom radu se spominje kako je on napredak u odnosu na prethodne javno dostupne skupove. Kao primjer se navodi *BIWI Walking Pedestrians* skup s dva podskupa, oba nastala snimanjem iz zraka sa zgrada u Zürichu, koji su malenog opsega i sadrže samo pješake. Nadalje, *Stanford Drone Dataset* za prikupljanje podataka koristi dronove – evidentirano je više od 10 000 putanja različitih sudionika u prometu na području sveučilišnog kampusa, no, s obzirom na lokaciju snimanja, manje od 6 % putanja se odnosi na vozila. 2019. godine objavljeni su skupovi podataka pod nazivima *CITR* i *DUT*. *DUT* je također kreiran snimanjem kretnji pješaka na kampusu, uz prethodno navedene nedostatke snimanja u tim uvjetima, a *CITR* sadrži kontrolirana ponašanja vozila na parkiralištu.

Uzimajući prethodno navedeno u obzir, istraživači s Aachena kreirali su skup podataka sa sljedećim prednostima:

- Očuvano je prirodno ponašanje sudionika u prometu time što je uređaj za snimanje bio izvan njihovog vidnog polja
- Skup podataka sadrži više tisuća putanja različitih sudionika u prometu kako bi se osigurala raznovrsnost
- Višestrukost lokacija i vremena snimanja (prikupljanja podataka), snimanje na javnim površinama
- Praćenje svih vrsta sudionika u prometu
- Točnost geografskog pozicioniranja sudionika u prometu (greška < 0,1 m)
- Uključenost i prikaz cestovne infrastrukture i prometne signalizacije

2.2.3. Metoda prikupljanja podataka

Prvi korak u prikupljanju podataka bio je odabir lokacije za snimanje. Četiri navedena raskrižja odabrana su temeljem informacija o gustoći prometa te o sastavu prometa (izgled raskrižja, postojanje prometne signalizacije itd.). Nadalje, snimanje je obavljeno u uvjetima vedrog vremena, prikladnog prirodnog osvjetljenja te uz slabe vjetrove. Dron je snimao sa visine do 100 m (ovisno o površini raskrižja koju je trebalo uhvatiti kamerom). U naknadnoj obradi je izvedena korekcija distorzije leće i svi vremenski okviri su stabilizirani kako bi odgovarali prvom okviru.

Iz razloga što je ručno označavanje svakog od objekata na svakom vremenskom okviru bilo previše zahtjevno, za obradu podataka o položaju objekata korišteni su algoritmi računalnog vida temeljeni na neuronskim mrežama. Za učenje neuronske mreže, prvo je ručno označeno 400 objekata na način da je njihova površina označena pomoću kvadratnih okvira (eng. *bounding box*), a skup podataka za treniranje mreže uvećan je standardnim postupcima poput rotacije i deformacije slike. Za dodjeljivanje klase svakom pikselu korištena je semantička segmentacija. Primjerice, definiralo se predstavlja li neki piksel vozilo ili pozadinu. Na taj način se skupu piksela iste klase dodijelio kvadratni okvir.

Za definiranje vrste praćenog objekta korištene su dvije neuronske mreže. Jedna je bila odgovorna za objekte malih dimenzija, poput biciklista i pješaka, a druga za ostale praćene objekte. Klasifikacija objekata u kategorije temeljila se na očekivanim vrijednostima veličine, brzine i položaja objekta.

S obzirom na to da su se sve detekcije obavljale pojedinačno za svaki vremenski okvir, bilo ih je potrebno spojiti kako bi se kreirale trajektorije i kako bi se sve vrijednosti povezale uz određeni objekt. Ovdje je korišten algoritam za praćenje, koji je, uz kreiranje trajektorija, uklanjao lažno pozitivne detekcije te interpolirao informacije o objektu u slučajevima kada se objekt ne bi pojavio u nekoliko vremenskih okvira. Konačno, sve su trajektorije pretvorene iz koordinata slike u metričke koordinate, ovisne o lokalnom koordinatnom sustavu.

2.2.4. Usporedba s postojećim skupovima podataka

U odnosu na preostale, već navedene, javno dostupne skupove podataka, *inD* skup podataka prednjači u broju potpunih informacija i trajektorija pješaka (v. Tablicu 1). Iako veću količinu informacija o sudionicima u prometu sadrže i skupovi *Stanford Drone Dataset* i *Interaction dataset*, njihove mane očituju se u malim brojevima interakcija pješaka i vozila (u *Stanford Drone Datasetu* većina vozila je parkirana) te manjku informacija o ranjivim korisnicima u prometu (pješacima, biciklistima) čime je nemoguće promatrati interakcije između istih (*Interaction dataset*).

Tablica 1. Statističke informacije o lokacijama snimanja (kratica VRU proizlazi iz *Vulnerable Road Users* i odnosi se na pješake i bicikliste).

	<i>Bendplatz</i>	<i>Frankenburg</i>	<i>Heckstrasse</i>	<i>Neukoellner Strasse</i>
Broj snimanja	11	12	3	7
Ukupno trajanje snimki, u minutama	185	243	51	110
Broj vozila	2503	2436	1196	2098
Broj VRU-a	1235	3799	83	249

3. PREDOBRAĐA I ANALIZA SKUPA PODATAKA

Kako bi se kreirao model ponašanja pješaka, potrebno je odabrati, izdvojiti te prilagoditi podatke ideji rješenja predstavljenoj u uvodu.

3.1. Odabir podataka

3.1.1. Program za vizualizaciju podataka

Prvi korak u kreiranju modela pješaka svakako je odabir podataka za obradu. *inD* skup podataka sastoji se od podataka izdvojenih iz 33 snimke. Za svaku snimku dane su četiri datoteke. Prva je slikovna, prikazuje raskrižje na koje se snimka odnosi. Preostale tri datoteke su u *.csv* formatu (eng. *Comma Separated Values*, vrijednosti odvojene zarezom). Kao što i sam naziv govori, ovaj format predstavlja tekstualnu datoteku u kojoj su informacije odvojene zarezima. Zapisi u *.csv* formatu vrlo su zastupljeni u raznim bazama podataka i općenito tabličnim skupovima informacija, a velika većina programskih paketa svih namjena ima ugrađenu mogućnost otvaranja dokumenata ovog formata. Jedan od čestih načina primjene upravo je i korištenje ovog formata samo kao posredničkog, kako bi se informacije mogle prenijeti iz jednog u drugi programski paket, kada to direktnim putem nije moguće obaviti između dotičnih programa.

Nazivi triju prethodno spomenutih datoteka, u *.csv* formatu, s podacima korištenim u ovom radu glase: *XX_recordingMeta.csv*, *XX_tracksMeta.csv* te *XX_tracks.csv*, gdje *XX* predstavlja jedinstvenu identifikacijsku oznaku svake snimke, odnosno, *ID* (eng. kratica za *identification*). U ovom slučaju, *ID* je broj vrijednosti između 0 i 32.

Prva datoteka od navedene tri daje generalni pregled informacija u snimci, druga sadrži kratki pregled svih objekata unutar snimke, sa svrhom izdvajanja objekata, na primjer, prema klasi ili kategoriji (automobil, autobus/kamion, pješak, bicikl, motocikl). Treća datoteka predstavlja iznose svih vremenski ovisnih vrijednosti za svaki objekt. Ona je najopsežnija od prethodno navedenih i najviše će se koristiti u ovom radu. Detaljniji popis informacija sadržanih u datotekama dan je u Tablici 2, Tablici 3 i Tablici 4..

Tablica 2. Sadržaj datoteka *Recording Meta Information*.

<i>Naziv</i>	<i>Opis</i>	<i>Mjerna jedinica</i>
recordingId	identifikacijska oznaka snimke, svaka snimka sadrži vlastitu oznaku	-
locationId	identifikacijska oznaka lokacije snimanja	-
frameRate	frekvencija izrade vremenskih okvira videa	Hz
speedLimit	ograničenje brzine	m/s
weekday	dan u tjednu tokom kojeg se snimalo	-
startTime	sat unutar koje je započelo snimanje	hh
duration	trajanje snimanja	s
numTracks	broj praćenih objekata	-
numVehicles	broj praćenih vozila	-
numVRUs	broj praćenih ranjivih sudionika u prometu	-
latLocation	grubo određene koordinate zemljopisne širine na lokaciji snimanja	deg
lonLocation	grubo određene koordinate zemljopisne dužine na lokaciji snimanja	deg
xUtmOrigin	x vrijednost UTM koordinate ishodišta lokalnog koordinatnog sustava za lokaciju snimanja	m
yUtmOrigin	y vrijednost UTM koordinate ishodišta lokalnog koordinatnog sustava za lokaciju snimanja	m
orthoPxToMeter	faktor skaliranja od orto piksela slike do UTM metara (vrijednost potrebna za vizualizaciju)	m/px

Tablica 3. Sadržaj datoteka *Track Meta Information*.

<i>Naziv</i>	<i>Opis</i>	<i>Mjerna jedinica</i>
recordingId	identifikacijska oznaka snimke, svaka snimka sadrži vlastitu oznaku	-
trackId	identifikacijska oznaka objekta	-
initialFrame	početni vremenski okvir objekta	-
finalFrame	zadnji vremenski okvir objekta	-
numFrames	ukupan broj vremenskih okvira	-
width	širina praćenog objekta (postavljena na nulu kod ranjivih sudionika u prometu)	m
length	duljina praćenog objekta (postavljena na nulu kod ranjivih sudionika u prometu)	m
class	kategorija praćenog objekta	-

Tablica 4. Sadržaj datoteka *Tracks*.

<i>Naziv</i>	<i>Opis</i>	<i>Mjerna jedinica</i>
recordingId	identifikacijska oznaka snimke, svaka snimka sadrži vlastitu oznaku	-
trackId	identifikacijska oznaka objekta	-
frame	vremenski okvir za koji su dane informacije	-
trackLifetime	trenutna „starost“ objekta u određenom vremenskom okviru	-
xCenter	x koordinata težišta objekta u lokalnom koordinatnom sustavu	m
yCenter	y koordinata težišta objekta u lokalnom koordinatnom sustavu	m
heading	pravac kretanja u lokalnom koordinatnom sustavu	deg
width	širina praćenog objekta (postavljena na nulu kod ranjivih sudionika u prometu)	m

length	duljina praćenog objekta (postavljena na nulu kod ranjivih sudionika u prometu)	m
xVelocity	brzina u smjeru x-osi lokalnog koordinatnog sustava	m/s
yVelocity	brzina u smjeru y-osi lokalnog koordinatnog sustava	m/s
xAcceleration	ubrzanje u smjeru x-osi lokalnog koordinatnog sustava	m/s ²
yAcceleration	ubrzanje u smjeru y-osi lokalnog koordinatnog sustava	m/s ²
lonVelocity	longitudinalna (uzdužna) brzina	m/s
latVelocity	lateralna (bočna) brzina	m/s
lonAcceleration	longitudinalno (uzdužno) ubrzanje	m/s ²
latAcceleration	lateralno (bočno) ubrzanje	m/s ²

Odabir podataka započinje pregledom snimki, radi upoznavanja sa skupom podataka i vizualnog određivanja interakcija pješaka s vozilima. Za detaljniju analizu odabrane su snimke pod oznakama 18-29, jer se jedine odnose na raskrižje s pješačkim prijelazom (sva raskrižja prikazana su na Slici 1).

Uz prethodno navedene datoteke, službenim odobravanjem pristupa skupu podataka omogućen je i pristup repozitoriju koji sadrži izvorne programske datoteke za učitavanje, obradu i vizualizaciju podataka pomoću programskih jezika *MATLAB* i *Python*.

Alat za vizualizaciju učitava podatke i prikazuje ih na fotografiji raskrižja. Moguće je vizualizirati određene vremenske okvire ili pokrenuti cijelu snimku. Također, moguće je prikazati dodatne podatke kao što su brzine objekata i njihova identifikacijska oznaka.

Za uspješno pokretanje programskog koda u programskom jeziku *Python* potrebno je, redom, odabrati postojeće ili kreirati novo okruženje (eng. *environment*) pomoću naredbi: `conda create -n drone-dataset-tools38 python=3.8` (za verziju *Pythona* 3.8) te `conda activate drone-dataset-tools38`. Idući korak je instalacija potrebnih paketa naredbom `pip3 install -r requirements.txt`. Za pokretanje vizualizacijskog alata potrebno je prebaciti se u direktorij pod nazivom *src/*, uz naredbu `cd src`.

Na Slici 2 prikazan je radni okvir programa za vizualizaciju podataka, koji se pokreće pokretanjem programskog koda pod nazivom `run_track_visualization.py`. Kako bi se pokrenula vizualizacija sa željenim parametrima, u terminal je neophodno upisati vrstu skupa podataka

(*ind*, *round*, *exid* ili *unid*) nakon izraza - *-dataset*. Nadalje, potrebno je eksplicitno definirati koju snimku je potrebno pokrenuti. Isto se postiže naredbom - *-recording* nakon koje slijedi redni broj snimke. Primjerice, za pokretanje snimke 19 uz zadane postavke prikaza koristi se naredba `python run_track_visualization.py --dataset ind --recording 19`. Ovakva naredba dovoljna je za vizualizaciju uz standardne postavke gdje je većina dodatnih prikaza isključena. U slučaju potrebe za suprotnim, dodatni prikazi se uključuju upisivanjem naredbe uz napomenu *True* pri pokretanju vizualizacije – zadana postavka za svaki od dodatnih prikaza je *False*. Sve dodatne naredbe su opcionalne i ovise o potrebama korisnika koji pregledava ili analizira podatke.

Također, prije pokretanja vizualizacijskog alata pomoću linije koda, potrebno je razmotriti lokaciju datoteka korištenih za analizu na računalu (već spominjane *.csv* datoteke). U slučaju da se datoteke ne nalaze u zadanom direktoriju, potrebno je navesti put do istih.

Za potrebe ovog rada, korištena je naredba `python run_track_visualization.py --dataset_dir D:\Dokumenti\FSB\Diplomski\inD-dataset-v1.0\data\ --dataset ind --recording 19 --show_bounding_box True --show_trajectory True --annotate_track_id True`. Prethodni primjer pokazuje veću količinu dodatnih naredbi. Ove dodatne naredbe, redom, postavljaju put do direktorija u kojem se nalaze *.csv* datoteke s podacima, navode vrstu skupa podataka, redni broj snimke, aktiviraju prikaz kvadratnih okvira koji okružuju površinu objekta te ih time predstavljaju (prethodno spomenuti *bounding box*). Naredba `--show_trajectory True` omogućava prikaz prethodne trajektorije vozila, a naredba `--annotate_track_id True` uz prikaz okvira objekta dodaje i brojanu identifikacijsku oznaku.

Slika 2 predstavlja vizualizaciju podataka u vremenskom okviru rednog broja 2500. Željeni vremenski okvir moguće je postaviti upisivanjem njegovog rednog broja pod *Set frame*, ili je moguće doći do njega pritiskom na srednju oznaku (trokut usmjeren desno, poznat kao oznaka za *Play*) koja pokreće snimku za po određeni broj vremenskih okvira. Preostali simboli na dnu prikaza predstavljaju mogućnost prelaska na sljedeći ili prethodni ili preskakanje za nekoliko vremenskih okvira unaprijed ili unatrag. Tijekom pokretanja vizualizacije, uz naredbu - *-playback speed* moguće je dodati broj *n* koji pokazuje razmak narednih vremenskih okvira koji će se prikazati prilikom pritiska na gumb *Play* (primjerice, upisivanjem broja 5, prikazuje se svaki peti vremenski okvir). Pritiskom na *Reset* prikaz se vraća u početni, nulti, vremenski okvir. U gornjem lijevom kutu radnog okvira vizualizacijskog alata nalaze se tri opcije reprezentirane s tri simbola. Opcija na trećem mjestu, prikazana simbolom povećala, daje

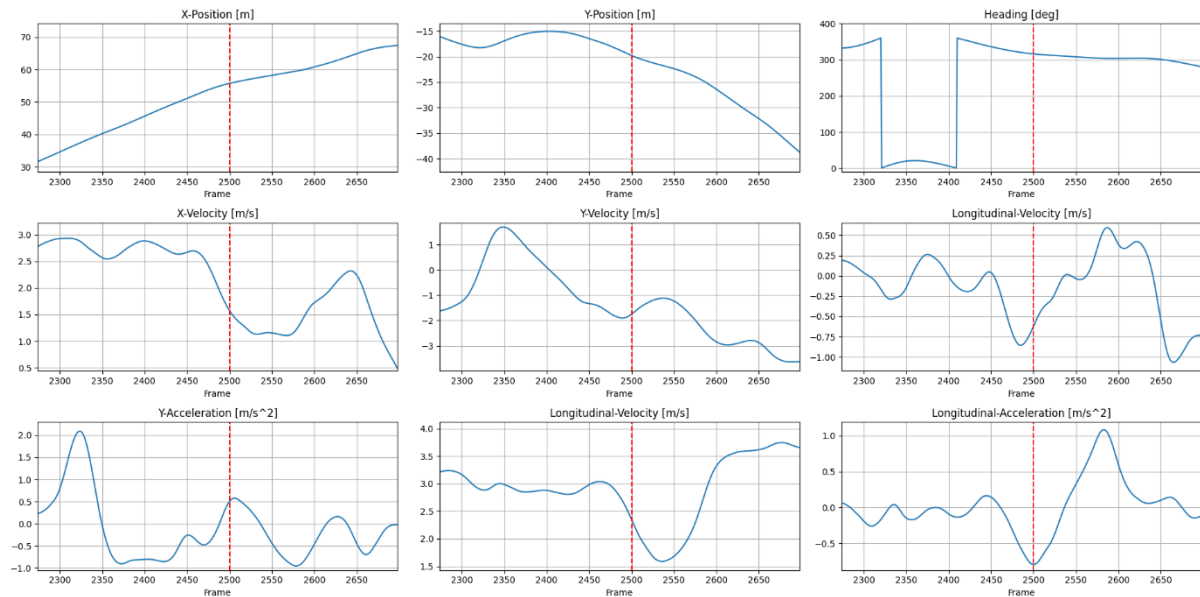
možnost uvećanja odabranog dijela slike, dok opcija na drugom mjestu omogućuje pomicanje po slici. Prva opcija prikazana simbolom kuće vraća vizualni prikaz u početno stanje.



Slika 2. Vizualizacija podataka pomoću danog programskog koda

Neke od ovdje neuključenih opcija omogućuju prikazivanje preostale putanje objekta, orijentacije vozila simbolom trokuta ili oznake temeljene na vrsti objekta, brzini, orijentaciji ili duljini pojavljivanja objekta u snimci. Svaka kategorija objekta definirana je svojom bojom (ovdje neprikazano) – na primjer, automobili su svijetloplavi, pješaci crveni, bicikli i motocikli žute boje. Inače, promatrani objekti na slici ni u kojem trenutku nisu eksplicitno prikazani kako se cijela vizualizacija ne bi zagušila informacijama.

Odabirom nekog objekta i pritiskom na njegovu identifikacijsku oznaku, otvara se novi prozor s dijagramima smjera, položaja, brzine i akceleracije istog objekta. Primjerice, Slika 3 prikazuje devet dijagrama unutar određenog vremenskog intervala koji istovremeno prikazuju prethodno navedene informacije o objektu. Moguće je uvećati željeni dio (odsjek) dijagrama za detaljniji uvid u vrijednosti na koordinatnim osima. Koordinatne osi dijagrama su osi lokalnog koordinatnog sustava, o čemu će više biti dano u narednim potpoglavljima.



Slika 3. Primjer prikaza dijagrama za objekt rednog broja 49 na snimci 19

3.1.2. Odabir promatranih objekata

Sljedeći korak pri pregledu snimki odnosio se na evidentiranje svih interakcija pješaka i vozila u blizini pješačkog prijelaza. Za svaki skup pješaka i vozila zabilježeno je prilazi li pješak pješačkom sa sjeverne ili južne strane, je li vozilo u bližem ili daljem voznom traku, kreće li se vozilo pravocrtno ili možda već stoji, nalaze li se na pješačkom prijelazu ili u blizini jedan ili više pješaka i prelaze li istovremeno ili iz suprotnih smjerova.

Iz prethodnih podataka izdvojen je broj pješaka (prikazanih kraticom PD - eng. *pedestrian*) koji prelaze preko pješačkog prijelaza, za svaku snimku, kako je vidljivo u Tablici 5. Prvi stupac prikazuje identifikacijsku oznaku snimke, drugi sveukupan broj pješaka, dok se treći odnosi na broj grupa ili pojedinačnih pješaka koji prelaze preko ceste u slučaju kada u njima bližem traku vozilo nailazi pravocrtno (u trak nije došlo skretanjem). Četvrti stupac prikazuje broj prijelaza u kojima sudjeluje jedan pješak i jedno vozilo.

Tablica 5. Analiza broja interakcija pješaka i vozila.

<i>ID snimke</i>	<i>Ukupan broj PD koji prelaze preko pješačkog</i>	<i>Broj grupa PD ili samostalnih PD u slučaju da vozilo nailazi pravocrtno u bližem traku</i>	<i>Broj samostalnih PD u slučaju da vozilo nailazi pravocrtno u bližem traku</i>
18	40	11	6
19	35	7	4
20	61	10	7
21	61	6	4
22	52	11	8
23	59	5	3
24	42	6	6
25	50	9	6
26	54	6	4
27	66	13	6
28	55	7	6
29	54	17	11
Σ	629	108	71

Iz Tablice 5 je vidljivo kako nijedna snimka ne prevladava značajno po broju interakcija u odnosu na ukupan broj od 71 interakcije. Temeljem toga, odlučeno je da će se za daljnju obradu koristiti podatci iz svih dvanaest snimki koje se unutar skupa podataka odnose na snimke s pješačkim prijelazom.

Idući dio obrade odnosi se na detaljnije informacije o pješacima i vozilima te karakteristikama njihove interakcije.

Obrada se temeljila na prikupljanju i uspoređivanju informacija o identifikacijskom broju pješaka i vozila, informaciji dolazi li pješak sa sjeverne ili južne strane prijelaza te informacije o tome tko prvi prelazi pješački prijelaz (pješak ili vozilo). Sve te informacije promatrane su unutar vremenskih okvira koji su, proizvoljno odabrano, ograničeni trenutkom kada je pješak na otprilike sedam metara od pješačkog prijelaza do trenutka kada pješak u potpunosti prijeđe preko pješačkog prijelaza. Naknadno je primjećeno da se ta ograničenja ne mogu uvijek

primijeniti, jer se, na primjer, vozilo nekad ne pojavljuje na snimci kada je pješak na sedam metara od pješačkog prijelaza, o čemu je više rečeno kasnije.

Uzevši u obzir činjenicu da su neke od 71 interakcije trajale po 20-ak vremenskih jedinica po 40 ms, a neke se odnosile na situaciju da vozilo već stoji radi prolaska prethodnog pješaka, odbačeno je 11 od prethodno navedenih 71 interakcija. Novih 60 interakcija podijeljeno je temeljem smjera prilaska pješaka pješačkom prijelazu – 35 slučajeva odnosilo se na prilazak pješaka sa sjeverne strane, dok se preostalih 25 odnosi na prilazak pješaka s južne strane.

Kako bi se zadržala temeljna pretpostavka da na ponašanje jednog vozila utječe ponašanje samo jednog pješaka i obrnuto, što je pokazano i eliminiranjem interakcija koje su potencijalno u obzir uzimale ostale faktore, odlučeno je da se ne razmatraju situacije gdje pješak prilazi pješačkom prijelazu s južne strane, a vozilo sa zapadne. Odluka je donesena na temelju vizualnog prikaza raskrižja. Iako su prethodno eliminirane situacije gdje vozilo u promtarani vozni trak dolazi skretanjem, nije moguće odrediti usporava li vozilo u slučaju prilaska pješačkom prijelazu sa zapadne strane samo radi pješaka ili i radi činjenice da je nešto prije pješačkog prijelaza prošao kroz neobilježeno raskrižje. Slika 5 prikazuje jednu od prethodno opisanih situacija. S obzirom na oznake na cesti, vidljivo je da vozilo dolaskom na raskrižje mora usporiti radi mogućnosti nailaska drugih sudionika u prometu s njemu lijeve ili desne strane.

Na Slici 4 prikazana je interakcija između pješaka i vozila u 2310. vremenskom okviru snimke 19. Vidljivo je kako vozilo koje se nalazi u bližem traku u odnosu na pješaka u taj trak ne nailazi pravocrtno te time njegovo usporavanje nema nužno veze s pojavom pješaka u blizini pješačkog prijelaza.

Frame: 2310/32542



Slika 4. Interakcija pješaka i vozila gdje vozilo u bližem traku ne nailazi pravocrtno

Frame: 22356/32542



Slika 5. Interakcija pješaka i vozila gdje nije moguće sa sigurnošću utvrditi razlog usporavanja vozila

Konačno odabran skup interakcija za razmatranje se, na kraju, odnosi na 35 slučajeva gdje pješak nailazi sa sjeverne strane prijelaza, a vozilo se kreće u smjeru istok → zapad (v. Slika 2). Odabrani skup snimki prikazan je u Tablica 6.

Tablica 6. Identifikacijske oznake odabranih interakcija pješaka i vozila.

<i>ID snimke</i>	18	18	19	19	19	19	20	20	20	20	21	22
<i>ID pješaka</i>	90	311	49	133	199	320	189	263	281	361	63	151
<i>ID vozila</i>	94	318	50	138	201	318	194	266	285	367	60	166
<i>ID snimke</i>	22	22	22	23	23	24	24	24	25	25	25	25
<i>ID pješaka</i>	451	458	541	53	188	1	39	122	63	183	205	245
<i>ID vozila</i>	462	464	551	58	192	34	40	130	68	180	211	248
<i>ID snimke</i>	26	27	27	28	28	28	28	28	29	29	29	
<i>ID pješaka</i>	38	60	158	23	99	134	320	467	235	404	498	
<i>ID vozila</i>	4	66	163	25	108	149	326	477	241	419	505	

3.2. Reduciranje informacija o objektima

U ovom potpoglavlju opisan je izbor značajnih informacija za daljnju obradu i kreiranje simulacije.

U prethodnom potpoglavlju, točnije, u Tablici 4, navedene su sve informacije o objektu koje su dostupne za određeni vremenski trenutak u datotečnom *.csv* formatu. Kako neke od njih nemaju značaj za simuliranje ponašanja pješaka, korisno ih je eliminirati ili reducirati.

Za manipulaciju *.csv* datotekama i daljnji nastavak rada, odabran je programski jezik *Python*. Ta odluka temelji se na već postojećem modulu *csv* (u radu korištena verzija 0.0.13) koji se instalira te učita u programski jezik pomoću naredbe *import csv*. Prednosti modula su jednostavno učitavanje informacija i pisanje u datoteke pomoću funkcija *csv.reader* i *csv.writer*. U ovom radu korištene su i klase *csv.DictReader* i *csv.DictWriter*, za čitanje i pisanje podataka u obliku rječnika ključeva.

Pomoću prethodno navedenih funkcija, kreiran je program za učitavanje datoteka *XX_tracks.csv*, za svih 12 identifikacijskih brojeva, te su iz istih izdvojeni podatci samo o

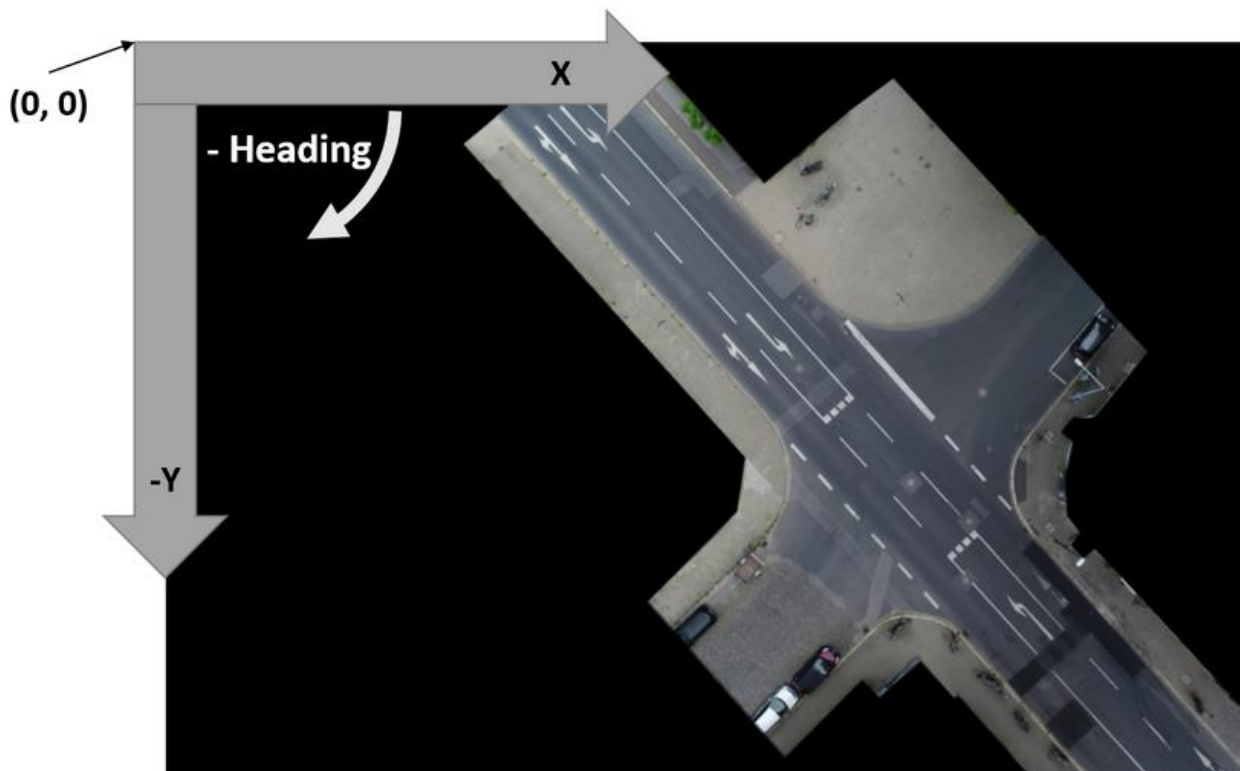
željenim pješacima i vozilima, navedenim u Tablici 6. Ti podatci spremljeni su u novu datoteku pod nazivom *potrebni.csv*.

Daljnja obrada odnosi se na datoteku *potrebni.csv*. Od 17 danih informacija po vremenskom koraku svakog od promatranih objekata, prvotno su eliminirane informacije o trenutnoj starosti objekta (*trackLifetime*), širini (*width*), pravcu kretanja (*heading*), bočnoj brzini (*lonVelocity*) i bočnom ubrzanju (*lonAcceleration*). Kao relevantni za kretanje pješaka i vozila, odabrani su njihovi položaji (položaji njihovih težišta, *xCenter* i *yCenter*), brzine (*xVelocity*, *yVelocity*, *lonVelocity*) i ubrzanja (*xAcceleration*, *yAcceleration*, *lonAcceleration*). Od preostalih podataka, zadržane su i informacije o identifikacijskim brojevima snimke i objekta te o duljini objekata, kako bi se lakše diferenciralo između pješaka i vozila.

3.2.1. Reduciranje iznosa udaljenosti na jednu vrijednost

Da bi se smanjila računalna zahtjevnost i sveukupno trajanje simulacija, preostale podatke potrebno je reducirati na još manji broj podataka. Primjerice, za udaljenost je odabrano da će se razmatrati apsolutna udaljenost objekta od bilo koje točke u kojoj se nalazi do točke u kojoj dolazi do pješačkog prijelaza te ga prelazi – udaljenost će se računati pomoću duljine putanje. Kako bi se udaljenost mogla prikazati u SI sustavu i time uspoređivati s drugim varijablama, položaje iz lokalnog koordinatnog sustava najprije je bilo potrebno pretvoriti u geografske koordinate.

U datotekama *XX_recordingMeta.csv* sadržane su informacije o UTM koordinatama ishodišta lokalnog koordinatnog sustava. Ishodište lokalnog koordinatnog sustava prikazano je na Slici 6. Iako slika prikazuje drugo raskrižje u odnosu na odabrano za ovaj rad, koordinatni sustav je za svako raskrižje jednako položen i orijentiran – nalazi se u gornjem lijevom vrhu slike.

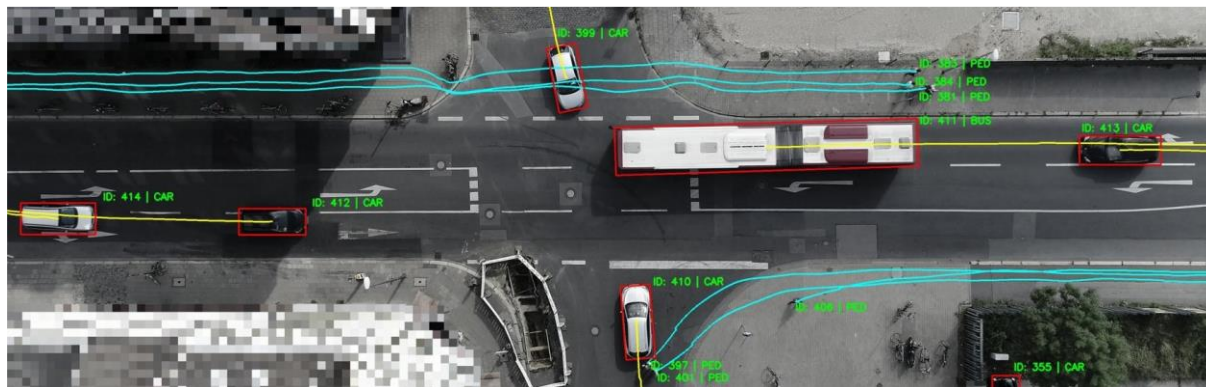


Slika 6. Položaj lokalnog koordinatnog sustava, [<https://www.ind-dataset.com/format>]

Za dobivanje položaja središta mase pješaka u UTM koordinatama, temeljem uputa sa službenih stranica skupa podataka, vrijednostima $xCenter$ i $yCenter$ dodane su vrijednosti $xUtmOrigin$ i $yUtmOrigin$.

Za izračun udaljenosti vozila korišten je ponešto drugačiji pristup – kako su širina i duljina pješaka jednake nuli, njegov centar mase precizno predstavlja točku u kojoj se on nalazi kada stupa na pješački prijelaz. S obzirom na to da je kod vozila centar mase pomaknut za pola duljine vozila u odnosu na dio vozila koji prvi pristupa pješačkom prijelazu, kod pretvorbe u UTM koordinate, bilo je potrebno oduzeti i polovicu iznosa duljine vozila kako bi se razmatrala prva točka pristupa pješačkom prijelazu, ona najbliža pješaku.

Slika 7 prikazuje primjer trajektorija dobivenih kao rezultat unutar *inD* skupa podataka. Vidljivo je da trajektorije započinju u sjecištu dijagonala pravokutnika koji obrubljuje vozilo. Na slici su vidljive trajektorije pješaka i njihovo ishodište.



Slika 7. Primjer trajektorija u inD skupu podataka, [5]

Idući korak pri izračunu udaljenosti pješaka i vozila predstavlja prebacivanje iz UTM koordinata u geografsku širinu i dužinu. Za točan izračun udaljenosti od pješačkog prijelaza, neophodno je računanje udaljenosti preko geografskih koordinata – zbog zakrivljenosti Zemljine plohe.

I u ovom slučaju je također korišten *Python*. Kreiran je jednostavan program u kojem je učitani modul *utm* (korištena verzija 0.7.0). [6] On u sebi sadrži ugrađenu funkciju *to_latlon*, koja pretvara UTM koordinate u geografsku širinu i dužinu. Odabrane vrijednosti koordinata zaokružene su na 6 decimalnih mjesta radi povećane preciznosti pri računanju udaljenosti.

Za izračun udaljenosti dvije točke koristi se haversinusna formula.

3.2.2. Haversinusna formula

Haversinusna formula precizno određuje udaljenosti između dvije točke na sferi temeljem njihovih geografskih širina i dužina. Proizlazi iz sferne trigonometrije, a koristila se za precizno računanje udaljenosti u astronomiji i u navigaciji. Pri sljedećem razmatranju poslužit će prikaz na Slici 8.

Temelj haversinusne formule je haversinus, poznat kao i haverzirani sinus ili semiversus. Zapisuje se kao $\text{haversin}(\vartheta)$, $\text{semiversin}(\vartheta)$, $\text{semiversinus}(\vartheta)$, $\text{havers}(\vartheta)$, $\text{hav}(\vartheta)$, $\text{sem}(\vartheta)$ ili $\text{hv}(\vartheta)$. Predstavlja polovicu vrijednosti versinusa (obnutog sinusa).

Versinus se može izraziti preko sinusne ili kosinusne funkcije:

$$\text{versin}(\vartheta) = 2 \sin\left(\frac{\vartheta}{2}\right) = 1 - \cos(\vartheta) \quad (1)$$

Analogno tome, a prema samoj definiciji haversinusa, on se definira kao:

$$\text{havrsin}(\vartheta) = \frac{1 - \cos(\vartheta)}{2} \quad (2)$$

Gdje su:

<i>versin</i>	-	versinus
<i>sin</i>	-	sinus
<i>cos</i>	-	kosinus
<i>havrsin</i>	-	havrsinus
ϑ	-	kut [°/rad]

Kako bi se precizno definirala havrsinusna formula, može se pretpostaviti da je središnji kut između bilo koje dvije točke na sferi jednak:

$$\vartheta = \frac{d}{r} \quad (3)$$

Gdje su:

d	[m]	udaljenost neke dvije točke po duljini glavne kružnice sfere
r	[m]	polumjer sfere

Havrsinusna formula također omogućava da se havrsinus od ϑ izračuna direktno iz geografske širine i geografske dužine. One su ovdje prikazane simbolima φ za geografsku širinu i λ za geografsku dužinu. Time proizlazi formula:

$$\text{hav}(\vartheta) = \text{hav}(\varphi_2 - \varphi_1) + \cos(\varphi_1) \cdot \cos(\varphi_2) \cdot \text{hav}(\lambda_2 - \lambda_1) \quad (4)$$

Iz razloga što kosinusi mogu uzrokovati degradaciju rezolucije pri malim kutovima, češće se primjenjuje jednadžba:

$$\text{hav}(\vartheta) = \text{hav}(\varphi_2 - \varphi_1) + (1 - \text{hav}(\varphi_1 - \varphi_2) - \text{hav}(\varphi_1 + \varphi_2)) \cdot \text{hav}(\lambda_2 - \lambda_1) \quad (5)$$

Ovdje su:

φ_1, φ_2	[°]	geografska širina točke 1 i točke 2
λ_1, λ_2	[°]	geografska dužina točke 1 i točke 2

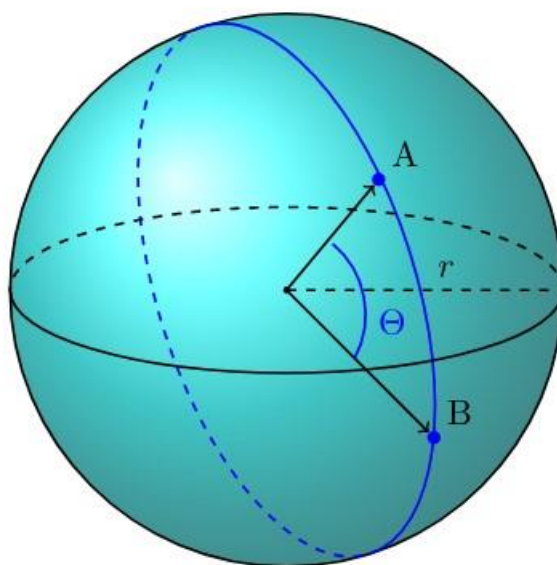
Kako bi se izračunala udaljenosti d , koristi se inverz havrsinusa ili inverz sinusa (arkus sinus):

$$d = r \cdot \text{archav}(h) = 2r \cdot \arcsin(\sqrt{h}) \quad (6)$$

Što se prošireno može zapisati kao:

$$\begin{aligned}
 d &= 2r \cdot \arcsin \left(\sqrt{\text{hav}(\varphi_2 - \varphi_1) + (1 - \text{hav}(\varphi_1 - \varphi_2) - \text{hav}(\varphi_1 + \varphi_2)) \cdot \text{hav}(\lambda_2 - \lambda_1)} \right) = \\
 &= 2r \cdot \arcsin \left(\sqrt{\sin^2 \left(\frac{\varphi_2 - \varphi_1}{2} \right) + \left(1 - \sin^2 \left(\frac{\varphi_2 - \varphi_1}{2} \right) - \sin^2 \left(\frac{\varphi_2 + \varphi_1}{2} \right) \right) \cdot \sin^2 \left(\frac{\lambda_2 - \lambda_1}{2} \right)} \right) = \\
 &= 2r \cdot \arcsin \left(\sqrt{\sin^2 \left(\frac{\varphi_2 - \varphi_1}{2} \right) + \cos(\varphi_1) \cdot \cos(\varphi_2) \cdot \sin^2 \left(\frac{\lambda_2 - \lambda_1}{2} \right)} \right) \quad (7)
 \end{aligned}$$

Slična jednadžba može se izvesti i preko kosinusa, no rezultat koji daje nije točan kada su u pitanju točke čija je udaljenost manja od približno jednog kilometra. U primjerima vezanim uz ovaj rad, apsolutni iznosi udaljenosti se kreću u rasponu od 0 m do 50-ak m (kasnije prikazano), za što nikako nije moguće koristiti jednadžbu zapisanu preko kosinusa. U oba slučaja, jednadžbe su samo aproksimacije pri primjeni na geografske koordinate Zemlje iz razloga što Zemlja nije idealna sfera. Njen polumjer se razlikuje na polovima i na ekvatorima. Uzevši sve to u obzir, točnost izračuna dobivenog haversinusnom ili kosinusnom jednadžbom sadrži grešku do 0,5 %. [8]



Slika 8. Udaljenost dvije točke na sferi, [9]

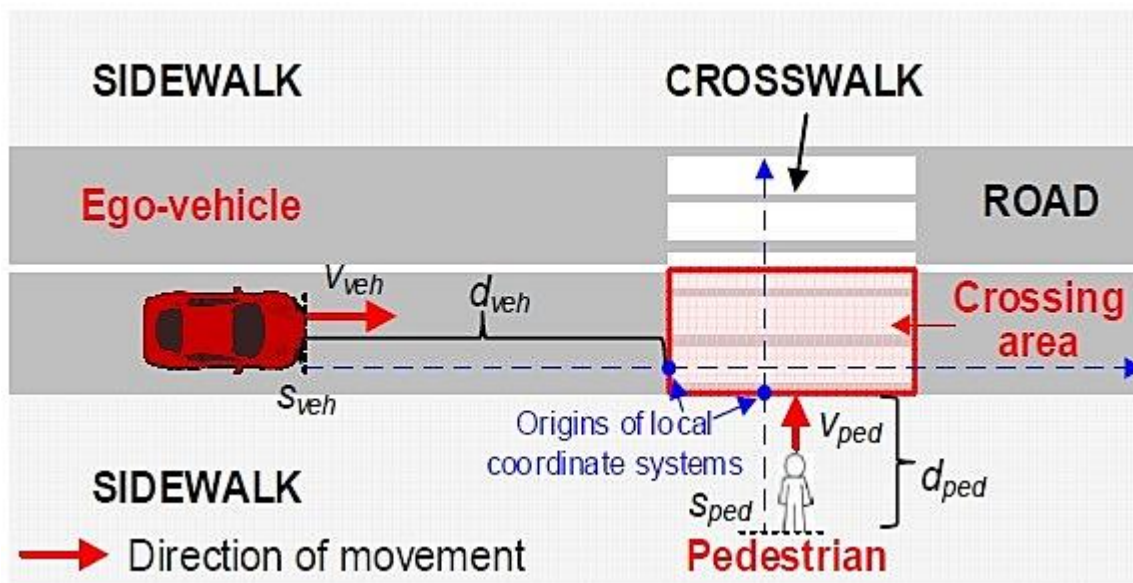
Iako se jednadžba (7) može jednostavno zapisati u programskom kodu i direktno koristiti za izračun, i ovdje je odabrana primjena već gotovog modula za *Python*, pod nazivom *haversine* (verzija 2.7.0). [10] Modul daje mogućnost izračuna udaljenosti bilo koje dvije točke uzevši u obzir njihove geografske koordinate te rješenje daje u kilometrima, miljama ili čak i nautičkim

miljama. Također, omogućuje i izračun geografske lokacije (točke) iz dane informacije o vektoru (udaljenosti i smjeru) i startnoj točki.

3.2.3. Udaljenost od pješačkog prijelaza

Idući korak u radu je primjena haversinusa da bi se položaj pješaka reducirao na jednu vrijednost umjesto trenutne dvije, kako bi se uklonili vremenski okviri u kojima su pješak ili vozilo pretjerano daleko ili se u njima ne odvija ništa značajno za cijelu interakciju prelaska preko pješačkog prijelaza.

Umjesto danog lokalnog koordinatnog sustava, postavljaju se novi, jednodimenzionalni koordinatni sustavi: jedan sadrži koordinatnu os paralelnu s uzdužnom osi kretanja pješaka, dok je os drugog koordinatnog sustava okomita na prvu te samim time paralelna sa uzdužnom osi kretanja vozila. Ishodište koordinatnog sustava pješaka nalazi se na početku pješačkog prijelaza sa sjeverne strane, dok je ishodište koordinatnog sustava pješaka na istočnom rubu pješačkog prijelaza. Novi koordinatni sustavi, iz perspektive pješaka, prikazani su na Slici 9.



Slika 9. Ishodište novokreiranih koordinatnih sustava uz prikaz pješaka, vozila, pločnika, kolnika i pješačkog prijelaza te smjerova kretanja pješaka i vozila, [11]

Slika 10 prikazuje geografski položaj pješačkog prijelaza. Ovdje je korišten i programski paket *MATLAB*; on omogućuje iscrtavanja točaka, linija ili objekata u obliku poligona na podržanim geografskim kartama ili projiciranim koordinatnim sustavima. Sve to omogućuje funkcija *geoplot*, kao dio skupa alata pod nazivom *Mapping (Mapping Toolbox)*. Funkcija se

koristi na način da se specificiraju rubne točke prijelaza. Te točke su definirane geografskim koordinatama. Pješački prijelaz je prvotno pronađen preko *Googleove* usluge *Maps*, a rubne točke, njihove geografske dužine i širine, određene su vizualnim putem, aproksimativno, zbog smanjene rezolucije slike. Preostale točke, koje se nalaze između rubnih geografski određenih, određene su interpolacijom.



Slika 10. Rubovi pješačkog prijelaza

Svih osam rubnih točaka pješačkog prijelaza sa pripadajućim koordinatama prikazano je u Tablici 7. Sve koordinate prikazane su na šest decimala radi preciznosti – šesta decimala je prva decimala u kojoj se neke od njih kreću razlikovati.

Tablica 7. Identifikacijske oznake odabranih interakcija pješaka i vozila

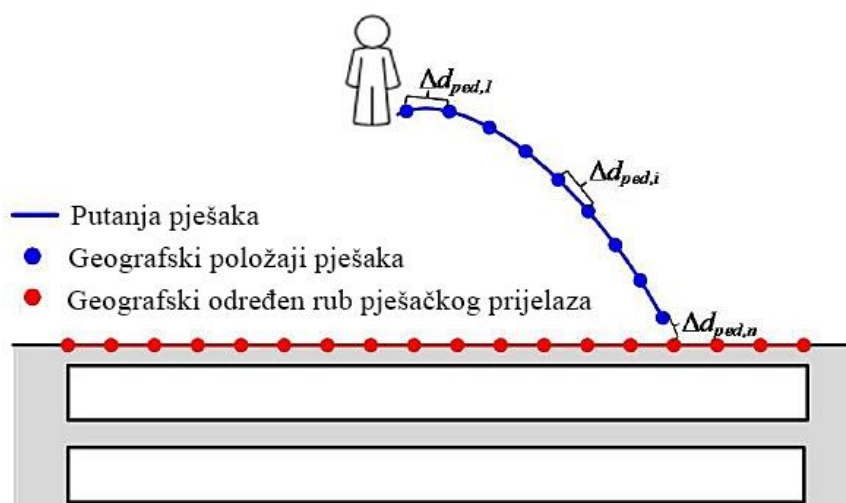
<i>Redni broj rubne točke</i>	<i>Položaj na slici</i>	<i>Geografska širina</i>	<i>Geografska dužina</i>
1	Sjeverni rub, lijevo	50,768417	6,102355
2	Sjeverni rub, desno	50,768425	6,102403
3	Istočni rub, gore	50,768423	6,102403
4	Istočni rub, dolje	50,768381	6,102419
5	Južni rub, lijevo	50,768378	6,102419
6	Južni rub, desno	50,768371	6,102372
7	Zapadni rub, gore	50,768415	6.102354
8	Zapadni rub, dolje	50.768374	6.102371

Izračunata širina pješačkog prijelaza je 3,49 m, dok duljina iznosi 4,76 m. Izračun je obavljen korištenjem haversinusne formule (7) za udaljenost rubnih koordinata sjevernog ruba (širina prijelaza) te za udaljenost rubnih koordinata zapadnog ruba (duljina prijelaza).

Iz rubnih točaka prijelaza pomoću jednadžbe pravca određeni su pravci koji predstavljaju sjeverni, južni, istočni i zapadni rub prijelaza, prema općepoznatom izrazu kako slijedi:

$$y - y_1 = \frac{y_2 - y_1}{x_2 - x_1} (x - x_1). \quad (8)$$

Kako bi se uklonile suvišne informacije o pješaku ili vozilu, kreirana je skripta u *Pythonu*. Princip po kojem funkcionira je taj da se prvo prolazi kroz cijelu snimku (sva stanja, sve vremenske okvire) promatranog objekta. U trenutku kad taj objekt stupi na pješački prijelaz, točnije, prijeđe predefimiran sjeverni rub kao pješak ili istočni rub kao vozilo, smatra se da je donesena odluka o tome prelazi li prijelaz prvo pješak ili vozilo. Informaciju o tome je li pješak prešao sjeverni rub pješačkog daje program temeljem toga nalaze li se njegove koordinate, uvrštene u jednadžbu pravca sjevernog ruba, izvan ili unutar skupa koji definira taj pravac. Analogan izračun, samo s pravcem koji definira istočni rub prijelaza, obavlja se za vozilo.



Slika 11. Princip izračuna apsolutne udaljenosti pješaka od pješačkog prijelaza: geografski položaji predstavljaju točke u kojima su poznate geografske koordinate, [11]

Pregledom podataka i same preglednosti pješačkog prijelaza i nogostupa koji do njega vodi, određeno je da će se početnom točkom smatrati točka na 7 m od prve točke unutar pješačkog prijelaza, po putanji pješaka. Za vozilo početna točka nije definirana na uniforman način nego

je usko vezana uz položaj pješaka, budući da je promatranje pješakovog ponašanja prioriternije pri kreiranju modela.

Potprogram, određivši prvu točku stupanja na prijelaz kao točku promjene po pitanju nalaženja unutar/izvan skupa koji definira taj pravac, od prve točke, to jest, od tog trenutka unatrag računa udaljenost za svaki prethodni korak te ih zbraja dok ne dođe do udaljenosti od 7 m za pješaka. Prvi vremenski okvir u kojem je pješak na udaljenosti većoj od 7 m po svojoj putanji uzima se kao početni te se svi vremenski okviri nižeg rednog broja vremenskog okvira zanemaruju te uklanjaju u nastavku obrade. Također, bitno je napomenuti kako ne idu svi pješaci pravocrtno niti slijede jedan smjer, već kod nekih dolazi do šetanja u drugim smjerovima te naknadnog vraćanja prema smjeru prelaska pješačkog prijelaza. Isto tako, u nekim vremenskim okvirima udaljenost pješaka od prijelaza ostaje nepromijenjena u odnosu na prethodni korak ili čak i više njih (primjer je stajanje i čekanje na prelazak na pješačkom prijelazu).

Kao kraj putanje određen je trenutak kada pješak u potpunosti prijeđe pješački prijelaz, ili kada vozilo izađe iz geografskog prostora prikazanog na snimci. Matematički, odnosno programski definirano, taj trenutak se određuje kao trenutak kada se pješakove koordinate nalaze izvan pravokutnika koji predstavlja površinu pješačkog prijelaza. Kraj putanje predstavlja i zadnji vremenski okvir za koji se promatraju informacije o pješaku i vozilu. Nužno je da se informacije o pješaku i vozilu promatraju za isti broj vremenskih okvira te da su početna i krajnja vrijednost tih okvira jednake, kako bi se kretanje moglo promatrati i uspoređivati u točno određenim trenucima.

U nekoliko od 35 razmatranih slučajeva, kada je pješak na 7 m od stupanja na prijelaz, identifikacijski broj vozila u tom okviru još uvijek ne postoji. Kako bi se ispunio prethodni zahtjev za jednakošću svih vremenskih okvira, odbačeni su vremenski okviri pješaka bez vozila te je u tim slučajevima udaljenost pješaka u početnom razmatranom okviru manja od predviđenih 7 m.

Nakon obavljenih izračuna, varijable $xCenter$ i $yCenter$ zamjenjuju se varijablom $distance$, koja predstavlja udaljenost od ishodišta pješačkog prijelaza po putanji objekta. Varijabla $distance$ poprima negativne vrijednosti za slučajeve kada pješak ili vozilo nisu stupili na pješački prijelaz, a pozitivne nakon prelaska pravca koji određuje njima bliži rub prijelaza.

3.2.4. Brzina i akceleracija pješaka i vozila

Analogno udaljenosti, i brzina i akceleracija su odabrane samo kao varijable u uzdužnom smjeru, to jest, smjeru kretanja. Time su potpuno odbačene vrijednosti *xVelocity*, *yVelocity* te *xAcceleration*, *yAcceleration*.

Dobivene vrijednosti se sada spajaju u jednu vrijednost za jedan vremenski okvir (jednu interakciju) – vrijednosti za pješaka se spajaju s informacijama o vrijednostima vozila u istim vremenskim okvirima za iste snimke. Ovaj postupak također je automatiziran pomoću *Python* koda.

Zaključno, potpuna obrada podataka dala je 35 snimki koje sadrže sveukupno 7221 interakciju između pješaka i vozila.

Svaki od obrađenih vremenskih okvira spremljen je u datoteku naziva *reducirani.csv*, po rednom broju pojavljivanja u snimkama.

4. MODELIRANJE PONAŠANJA PJEŠAKA

4.1. Markovljevi lanci

4.1.1. Definicija Markovljevih lanaca

Kako bi se ponašanje pješaka moglo modelirati te time i prikazati unutar simulacije, odabrana stanja povezana su u Markovljeve lance.

Markovljevi lanci predstavljaju nizove stanja sustava, a proučavaju nizove slučajnih varijabli kod kojih su skupovi stanja iz vremenske i domene ostalih varijabli diskretni. [12] To su stohastički procesi koji imaju tzv. „svojstvo zaboravljivosti“ – na buduće, odnosno sadašnje stanje, utječe samo konačan, odabran broj prethodnih stanja sustava. [12] Na taj se način razlikuju Markovljevi lanci prvog, drugog, i n -tog reda, ovisno o broju prethodnih stanja koja utječu na trenutno, odnosno buduće stanje sustava. U svakom trenutku sustav može prijeći u neko novo stanje, ili može ostati u istom stanju, a ove promjene stanja nazivaju se tranzicijama. [12]

Ako je lanac stanja X_0, X_1, \dots, X_n Markovljev lanac prvog reda, za sve izbore stanja s_0, s_1, \dots, s_n iz domene stanja S vrijedi:

$$P(X_n = s_n | X_{n-1} = s_{n-1}, \dots, X_0 = s_0) = P(X_n = s_n | X_{n-1} = s_{n-1}), \quad (9)$$

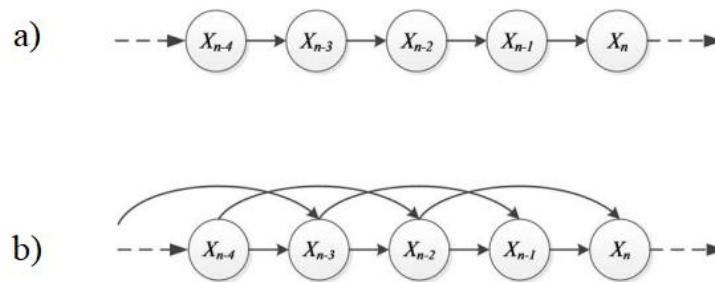
gdje n predstavlja trenutno stanje, a $0, \dots, n-1$ prošlost. Dakle, sadašnje stanje s_n ovisi samo o prethodnom s_{n-1} , ali ne i o načinu na koji je proces dospio u prethodno stanje, tj. vrijednostima procesa u ranijim trenucima.

Slično, u slučaju Markovljevog lanca drugog reda, za lanac stanja X_0, X_1, \dots, X_n mora vrijediti:

$$P(X_n = s_n | X_{n-1} = s_{n-1}, \dots, X_0 = s_0) = P(X_n = s_n | X_{n-1} = s_{n-1}, X_{n-2} = s_{n-2}), \quad (10)$$

što znači da je uvjetna vjerojatnost dolaska u stanje X_n jednakog iznosa uzmu li se u obzir prethodna dva, ili sva prethodna stanja u postojećem lancu stanja. Odnosno, vjerojatnost dolaska u trenutno stanje n ovisi samo o prethodna dva stanja $n-1$ i $n-2$.

Shematski prikazi Markovljevih lanaca prvog i drugog reda dani su na Slici 12.



Slika 12. Shematski prikaz Markovljevih lanaca: a) prvog reda, b) drugog reda, [12]

Analogne tvrdnje i izrazi mogu se primijeniti za proizvoljno odabran red Markovljevog lanca. Markovljevi lanci mogu se podijeliti na stacionarne i nestacionarne, ovisno o tome je li prijelazna vjerojatnost ovisna o koraku, tj. trenutku. U razmatranom slučaju radi se o Markovljevom lancu sa svojstvom stacionarnosti, za koji je matrica prijelaznih vjerojatnosti (prvog reda) s k stanja:

$$P = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1k} \\ p_{21} & p_{22} & \cdots & p_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ p_{k1} & p_{k2} & \cdots & p_{kk} \end{pmatrix}. \quad (11)$$

Elementi ove matrice predstavljaju vjerojatnost prijelaza dvaju uzastopnih slučajnih varijabli X_n i X_{n-1} iz stanja s_i u stanje s_j kako slijedi:

$$p_{ij} = P(X_n = s_i | X_{n-1} = s_j), \quad (12)$$

a moraju biti nenegativni, tj. $p_{ij} \geq 0$, te njihov zbroj u svakom retku matrice prijelaznih vjerojatnosti mora biti jednak jedinici.

4.1.2. Markovljev lanac u obliku rječnika ključeva

4.1.2.1. Rječnik ključeva

Nedostatak bilo kojeg Markovljevog lanca svakako je eksponencijalni porast zahtjeva na memoriju s brojem diskretnih stanja. Pod time se prvenstveno misli na nule, to jest, na prikaze nepostojanja korelacije ili sljedivosti između dvaju stanja (nepostojanje vjerojatnosti). U ovakvim primjerima gdje su parametri realni, kao što su udaljenosti, brzine i ubrzanja, pojavljuje se veći broj kombinacija stanja između kojih je vjerojatnost prijelaza jednaka nuli

(tj. nijedan prijelaz među tim stanjima nije zabilježen u snimljenim podacima). Sve to je fizikalno jednostavno shvatljivo – pješak, na primjer, ne može iz stanja u kojem ima brzinu 0,5 m/s u idućem trenutku (nakon 40 ms) prijeći u stanje s brzinom od 3 m/s.

Pregledom podataka uočeno je da broj pojavljivanja svih stanja nije ravnomjerno raspoređen. Neka od stanja prelaze samo u jedno, dok se kod nekih pojavljuje situacija da prelaze u nekoliko desetaka drugih stanja. Samim time korištenje matrice prijelaznih vrijednosti dodatno gubi smisao – pojavljivost velikog broja nula u matrici je neizbježna.

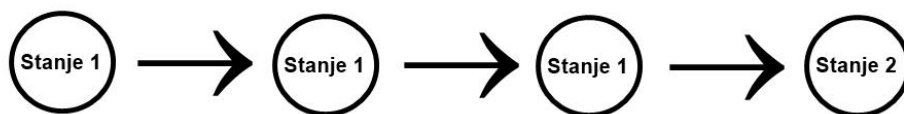
Matrica gdje kao elementi većinski prevladavaju nule naziva se rijetka matrica (eng. *sparse matrix*). Takve matrice su česta pojava u strojnom učenju, primjerice, u obradi prirodnog jezika (NLP - eng. *Natural Language Processing*).

Dodatni problem matrice prijelaznih vrijednosti je činjenica da su, s obzirom na broj promatranih varijabli, sva stanja opisana sa šest varijabli, što vodi na matricu s 12 dimenzija. Porast broja varijabli stanja nužno uključuje i porast memorijskih zahtjeva za obradu podataka, što je u ovom slučaju neopravdano s obzirom na to da je većina elemenata matrice jednaka nuli te su te informacije manje korisne. Mogućnosti su ili smanjivanje dimenzionalnosti matrice ili promjena pristupa. S obzirom na to da bi smanjenje dimenzionalnosti matrice utjecalo na točnost modela, odlučeno je kako će se umjesto matrice prijelaznih vrijednosti koristiti struktura podataka temeljena na rječniku ključeva unutar *Pythona*. Karakteristika rječnika je što sadrži podatke u obliku povezanih parova nazvanih ključ (eng. *key*) i vrijednosti (eng. *value*). Svaki ključ može biti povezan s više vrijednosti, no sam ključ se u rječniku pojavljuje samo jednom. Elementi u rječniku su promjenjivi, a ovisno o verziji *Python* programa koja se koristi, mogu imati definiran nepromjenjiv red. Primjerice, sve prethodne verzije *Pythona*, zaključno s verzijom 3.6, nisu podrazumijevale da rječnici imaju definiran poredak te se u tim verzijama nije moglo pristupiti elementima rječnika koristeći brojevu oznaku za određivanje njihovog položaja (eng. *index*). Vrijednosti unutar rječnika mogu biti bilo kojeg tipa (*string*, *integer*, *boolean* ili *list*). Neke od često korištenih naredbi vezanih uz rječnik uključuju *len* (eng. duljina) za vraćanje informacije o broju elemenata u rječniku, vraćanje informacije o vrijednosti pod ključem pomoću metode *get(key)*, jednostavno dodavanje novih ključeva i vrijednosti pomoću naredbe *dictionary[key]=value*, uklanjanje određenog elementa ili brisanje svih elemenata uz naredbe *pop* ili *clear*, vraćanje kopije rječnika uz *copy*.

4.1.2.2. Primjena rječnika ključeva za pohranu elemenata matrice

Kao što je već navedeno, umjesto svih elemenata rijetke matrice, u rječnik se spremaju elementi među kojima je korelacija različita od nule. Iako bi rijetka matrica sadržavala vjerojatnosti o prelasku iz prethodnog u sljedeće stanje, te se iste vrijednosti ne prebacuju direktno u vrijednosti nekog od ključeva. Za potrebe ovog rada predložen je modificirani rječnik, gdje su vrijednosti nekog ključa sva stanja u koja se može doći iz stanja određenog ključem.

U nekim slučajevima prethodno stanje je jednako sljedećem – tada se kao ključ odabire prethodno stanje, a isto to, sljedeće stanje, postavlja kao vrijednost. U slučaju da u više vremenskih okvira slijede jednaka stanja, sva se dodjeljuju istom ključu. Svako od tih stanja zapisuje se u obliku liste čiji su elementi udaljenost, brzina i akceleracija pješaka te udaljenost, brzina i akceleracija vozila. Na Slici 13 dan je primjer situacije gdje se iz *Stanja 1* nakon nekoliko trenutaka ostajanja u istom stanju prelazi u *Stanje 2*. Slika 14 prikazuje korelaciju između klasičnog prikaza Markovljevog lanca, matrice prijelaznih vrijednosti i pripadajućeg rječnika ključeva.



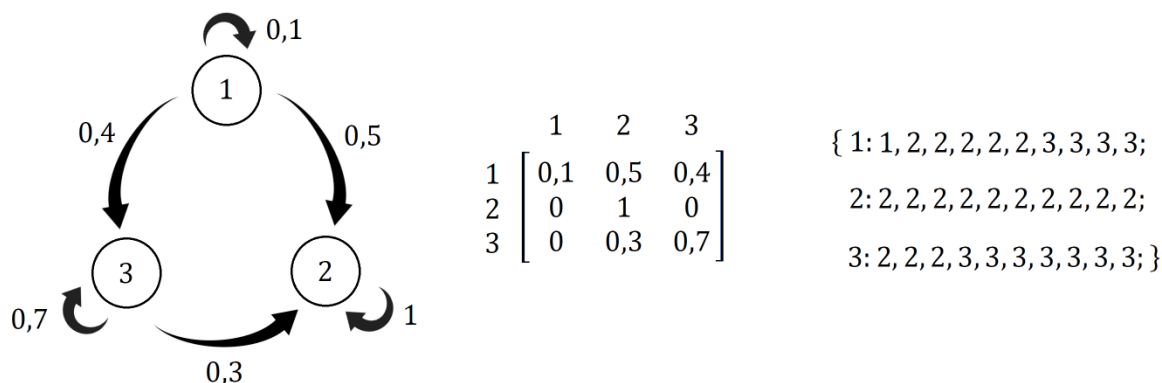
Slika 13. Prikaz jednog od mogućih prelazaka unutar Markovljevog lanca sastavljenog od dva stanja

Algoritam za kreiranje rječnika glasi kako slijedi:

1. učita se datoteka sa svim stanjima
2. kreiraju se prazne liste za skup svih ključeva i vrijednosti rječnika
3. prolazi se kroz sva stanja u dokumentu
4. ako se neko od tih stanja ne nalazi među ključevima rječnika, dodaje se u taj popis
5. svako stanje koje slijedi nakon stanja koje je jednako ključu rječnika dodaje se u listu vrijednosti pod tim ključem
6. u slučaju da je ključ zadnji element neke snimke, njegova vrijednost postaje prazna lista.

Algoritam se završava prolaskom kroz popis svih stanja među snimljenim podacima u datoteci. Krajnji rezultat algoritma je rječnik koji međusobno povezuje sva dana stanja. Tako dobiveni

algoritam temelj je za kreiranje simulacije, o čemu će više riječi biti u narednom poglavlju. Prednost rječnika je što se informacije o korelaciji između dva stanja ne gube te je uz par programskih naredbi rječnik lako ponovno pretvoriti u rijetku matricu. Krajnji oblik rječnika s informacijama o broju ključeva, vrijednosti i uporabi prikazan je u narednim potpoglavljima.



Slika 14. Usporedni prikaz klasičnog prikaza Markovljevog lanca, matrice prijelaznih vrijednosti i pripadajućeg rječnika ključeva

4.2. Princip modeliranja ponašanja pješaka u blizini pješačkih prijelaza

4.2.1. Odabir stanja za kreiranje modela

Datoteka *reducirani.csv*, koja sadrži sva stanja, sastoji se od, redom, podataka o identifikacijskoj oznaci snimke, identifikacijskoj oznaci pješaka te vozila, udaljenosti pješaka, udaljenosti vozila, brzini pješaka i vozila, ubrzanju pješaka te ubrzanju vozila. S obzirom na činjenicu da je svaki element stanja prikazan s informacijama na pet decimalnih mjesta, lako je zaključiti da je mogućnost pronalaženja dva ista skupa svih stanja, ili čak i dva pojedina ista elementa, vrlo malena ili nepostojeća. Ako bi se ti iznosi bez ikakvih promjena uzimali kao elementi rječnika ili očekivana sljedeća stanja, kreirane simulacije bile bi identične snimljenim – vjerojatnost prelaska u neko drugo stanje, osim snimljenog, bila bi nepostojeća.

Za osiguranje raznovrsnosti u vidu stanja u koje je moguće prijeći iz nekog odabranog stanja, odlučeno je sve vrijednosti reducirati na neke koje neće uvelike utjecati na točnost informacija, a opet će omogućiti međusobno povezivanje stanja iz različitih snimki.

Kako bi se kreirao rječnik, prvo je potrebno odabrati interakcije koja bi se u njemu nalazile.

Od postojećih 7221 opisanih interakcija u 7221 vremenskom okviru, 6464 je odabrano za skup učenja (kreiranje rječnika), a preostalih 757 za skup testiranja. Ta podjela temeljena je na odabiru 30 snimki za kreiranje rječnika te 5 za testne snimke (30 odabranih snimki zajedno

sadrži 6464 interakcije). Sveukupan broj prijelaza među svim snimljenim interakcijama je 7186. Svrstavanjem iznosa udaljenosti, brzine i ubrzanja vozila unutar određenih rezolucija, taj broj se znatno povećava.

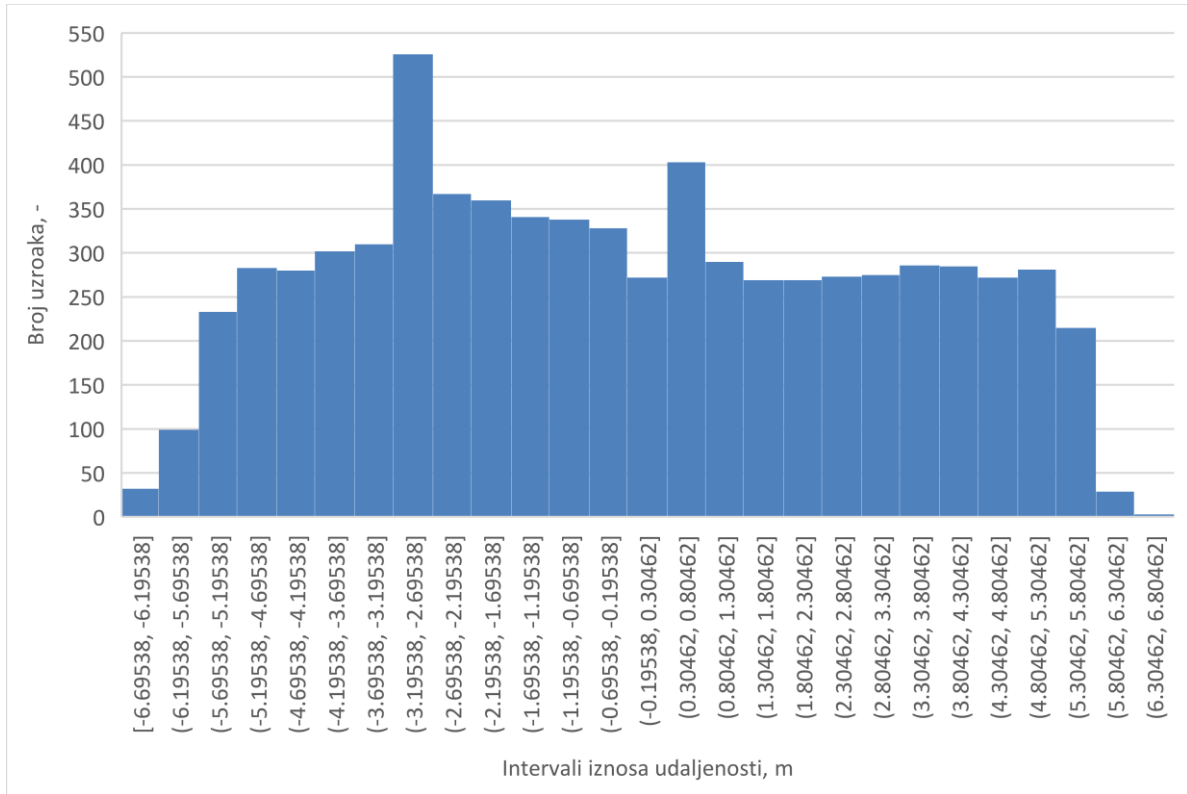
Za definiranje rezolucija bilo je potrebno pregledati iznose podataka o udaljenosti, brzini i ubrzanju te prikazati njihovu raspodjelu u sveukupnim podacima. Cilj smještanja vrijednosti koje opisuju gibanje objekta unutar rezolucija je kategorizirati ih u neku od situacija koje su prethodno analizirane i za koje je određeno željeno djelovanje vozila. Na taj način bi se kreirani model morao koristiti i na podacima koji nisu nužno ekvivalentni snimljenim skupovima podataka.

4.2.1.1. Odabir rezolucija za podatke o pješacima

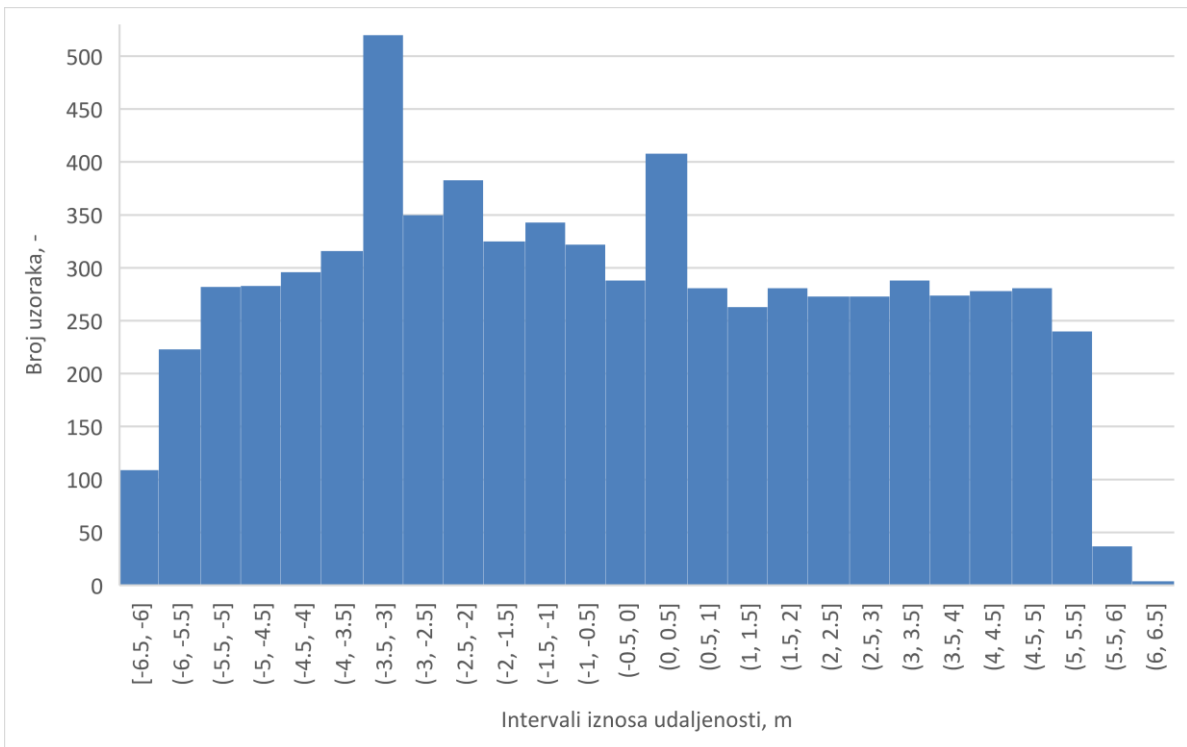
Budući da se ovaj rad primarno bavi ponašanjem pješaka u blizini pješačkog prijelaza, a tek onda vozilima, veći je fokus stavljen na rezolucije podataka vezanih uz pješake. To znači da su rezolucije vozila grublje kada je isto išlo u korist većim rezolucijama podataka o pješaku.

Kako bi se odredila optimalna rezolucija za predstavljanje udaljenosti pješaka, primarno su prikazani iznosi svih udaljenosti na histogramu. Jedan stupac sadržava vrijednosti unutar 0,5 m. Najmanja promatrana udaljenost je -6,69538 m, dok je najveća 6,80462 m. Vidljivo je da je raspodjela udaljenosti od stupanja na pješački prijelaz poprilično slična (broj uzoraka je poprilično sličan za sve intervale), što je i više nego očekivano s obzirom na to da se pri definiranju stanja za kreiranje modela referenciralo na udaljenost pješaka do i od stupanja na pješački prijelaz.

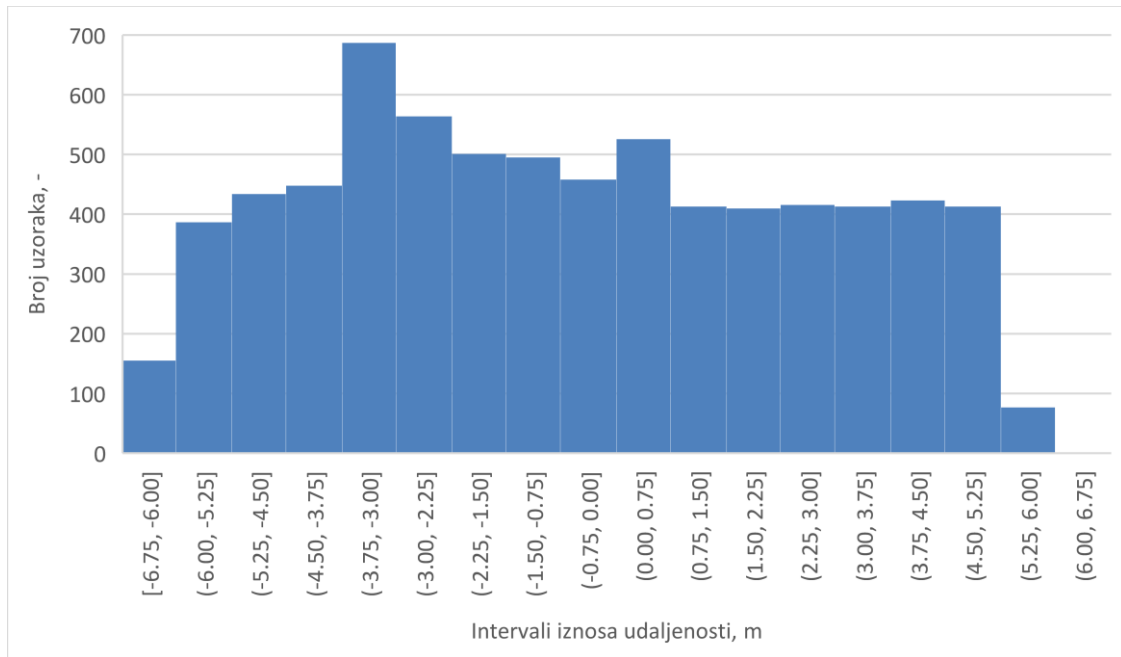
Razmatrane su rezolucije od 0,5 m, 0,75 m i 1 m. Raspodjela svih snimljenih udaljenosti te raspodjela za svaku rezoluciju prikazane su na pripadnim dijagramima. [Slike 15 – 18]



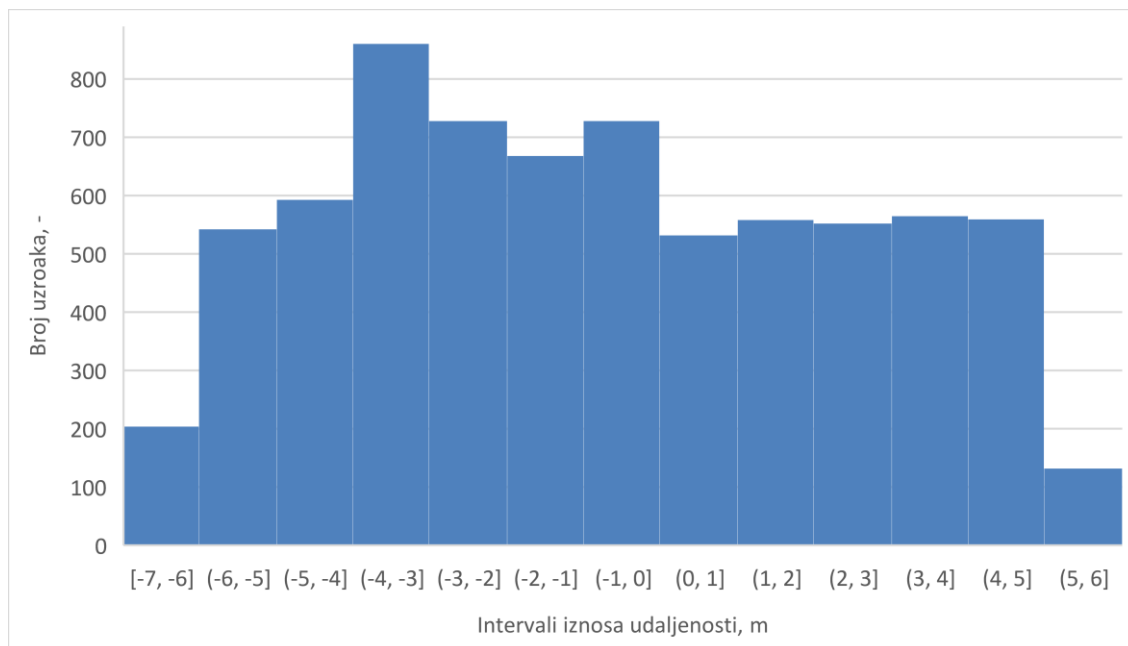
Slika 15. Raspodjela neobrađenih vrijednosti udaljenosti pješaka



Slika 16. Raspodjela udaljenosti pješaka u slučaju rezolucije od 0,5 m



Slika 17. Raspodjela brzina pješaka u slučaju udaljenosti od 0,75 m



Slika 18. Raspodjela udaljenosti pješaka u slučaju rezolucije od 1 m

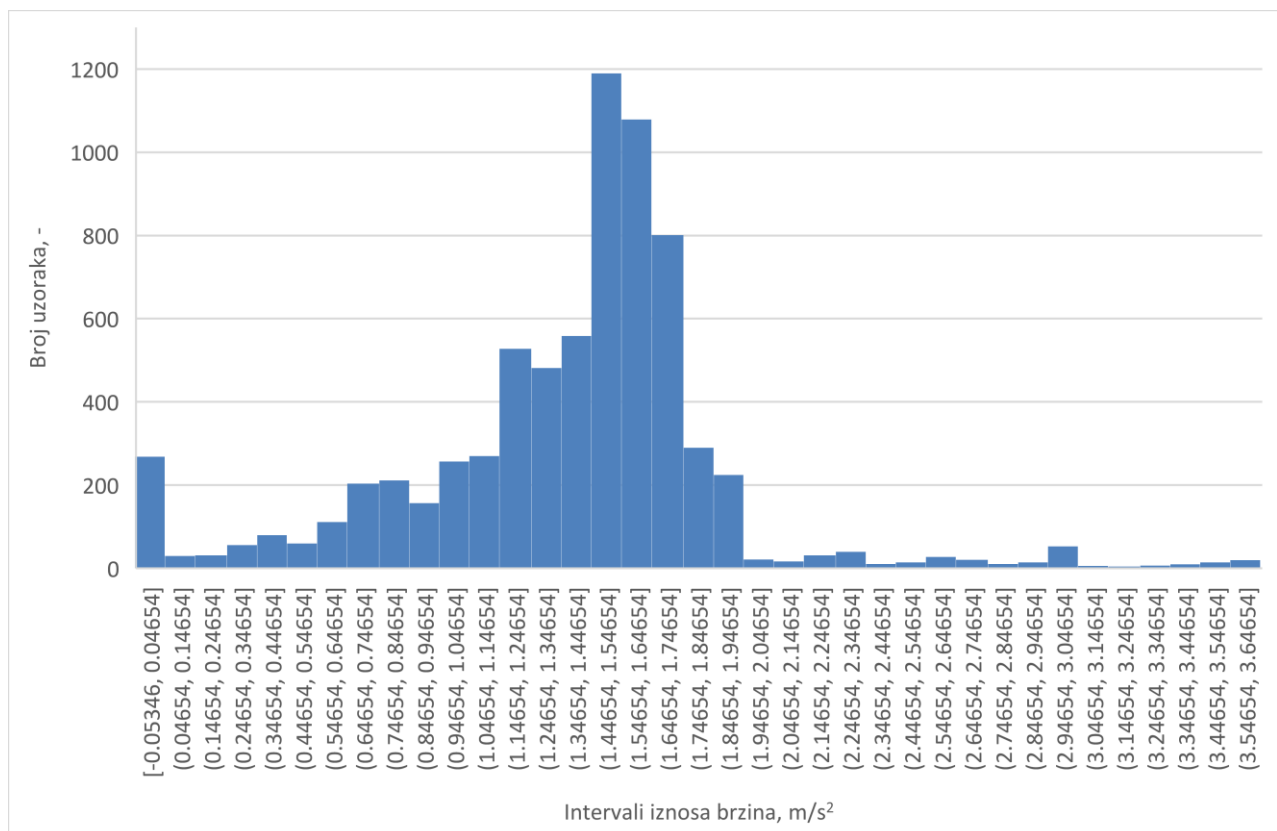
Uvidom u dijagrame vidljivo je da su, neovisno o rezoluciji, dijagrami i dalje jednake raspodjele kao i set početnih podataka.

S obzirom na velik broj uzoraka kod rezolucije udaljenosti od 0,5 m, ista je odbačena te su za kasnije kreiranje setova rezolucija odabrane rezolucije od 0,75 m i 1 m.

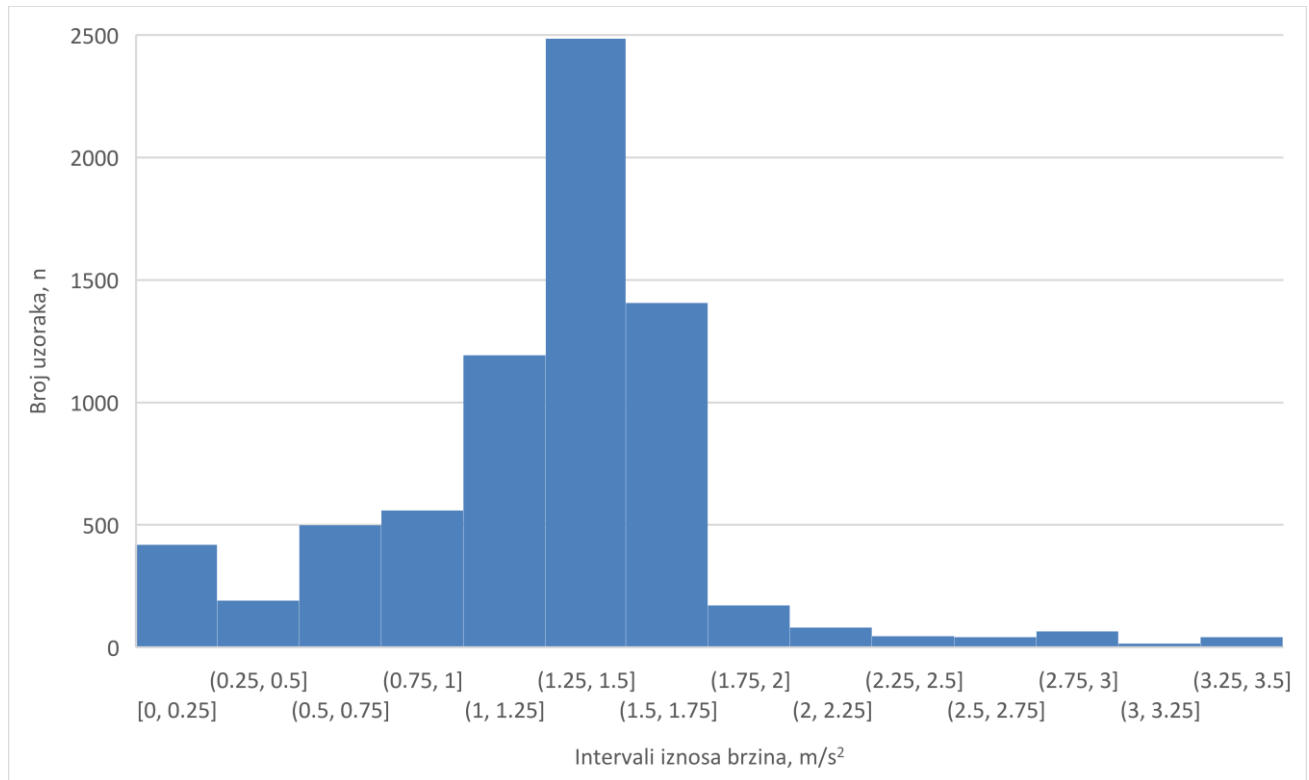
Nastavno na udaljenost pješaka, kao sljedeća varijabla po značaju smatra se varijabla iznosa brzine pješaka. Dijagram na Slici 19 prikazuje raspodjelu brzina pješaka unutar svih 7221 stanja.

Pri odabiru rezolucije brzine pješaka razmatrane su rezolucije od 0,5 m/s i 0,25 m/s. Za rezoluciju od 0,25 m/s, dobiva se 14 uzoraka, dok za rezoluciju od 0,5 m/s proizlazi 7 uzoraka. [Slika 20, Slika 21] Rezolucija od 1 m nije razmatrana iz razloga što se većina podataka nalazi između 0 m/s i 2 m/s, što bi generalno dalo premalo stanja da bi simulacijski model bio sličan stvarnosti.

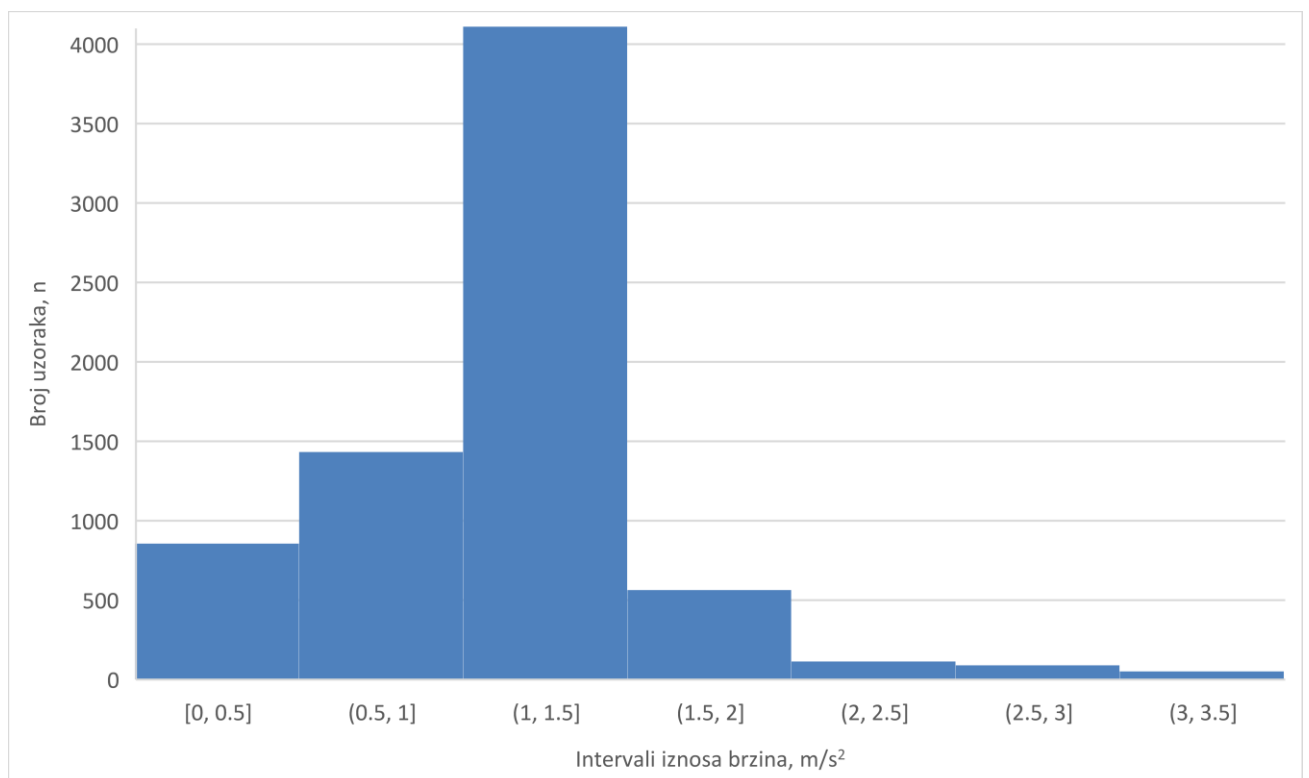
Samim time, za brzinu pješaka odabrana je rezolucija od 0,5 m/s jer istovremeno pokazuje veću preciznost pri definiranju brzine, a ne uzrokuje pojavu prevelikog broja uzoraka. Opcija rezolucije brzine od 0,25 m/s najbližnja je originalnoj raspodjeli vrijednosti, no daje prevelik opseg vrijednosti.



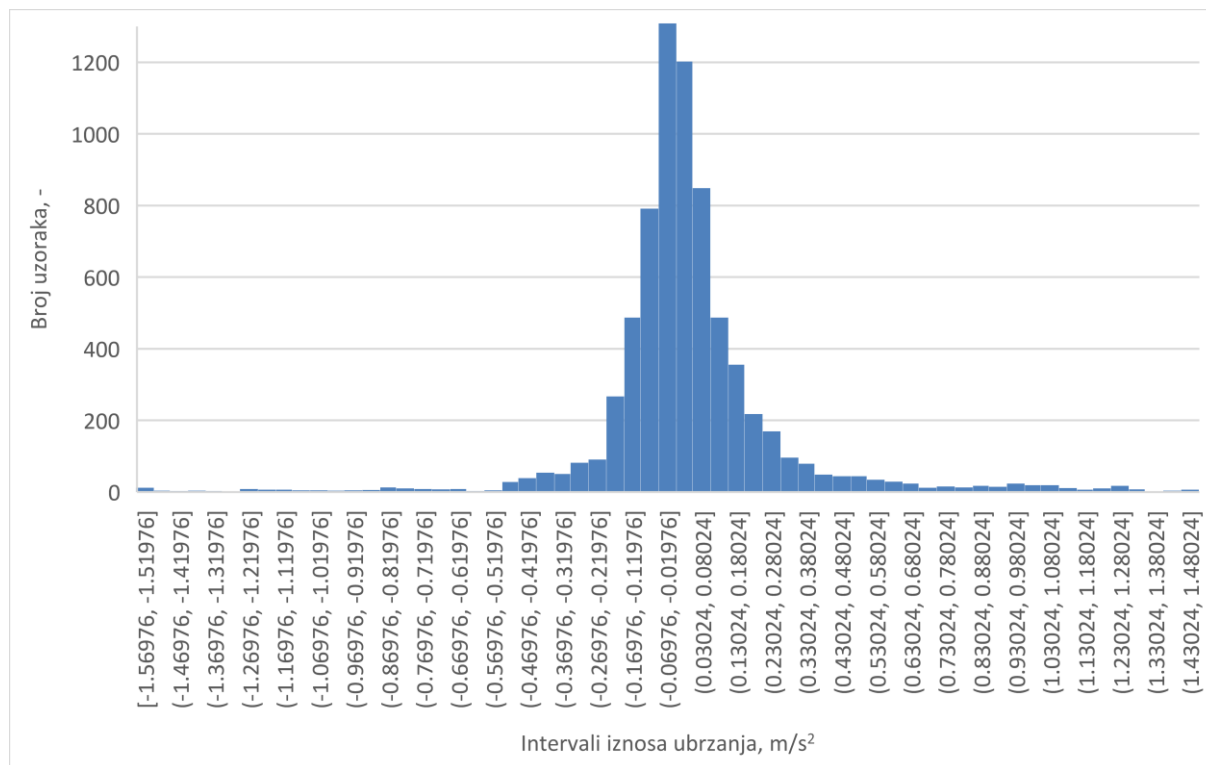
Slika 19. Raspodjela brzina pješaka



Slika 20. Raspodjela brzina pješaka u slučaju rezolucije od 0,25 m/s



Slika 21. Raspodjela brzina pješaka u slučaju rezolucije od 0,5 m/s



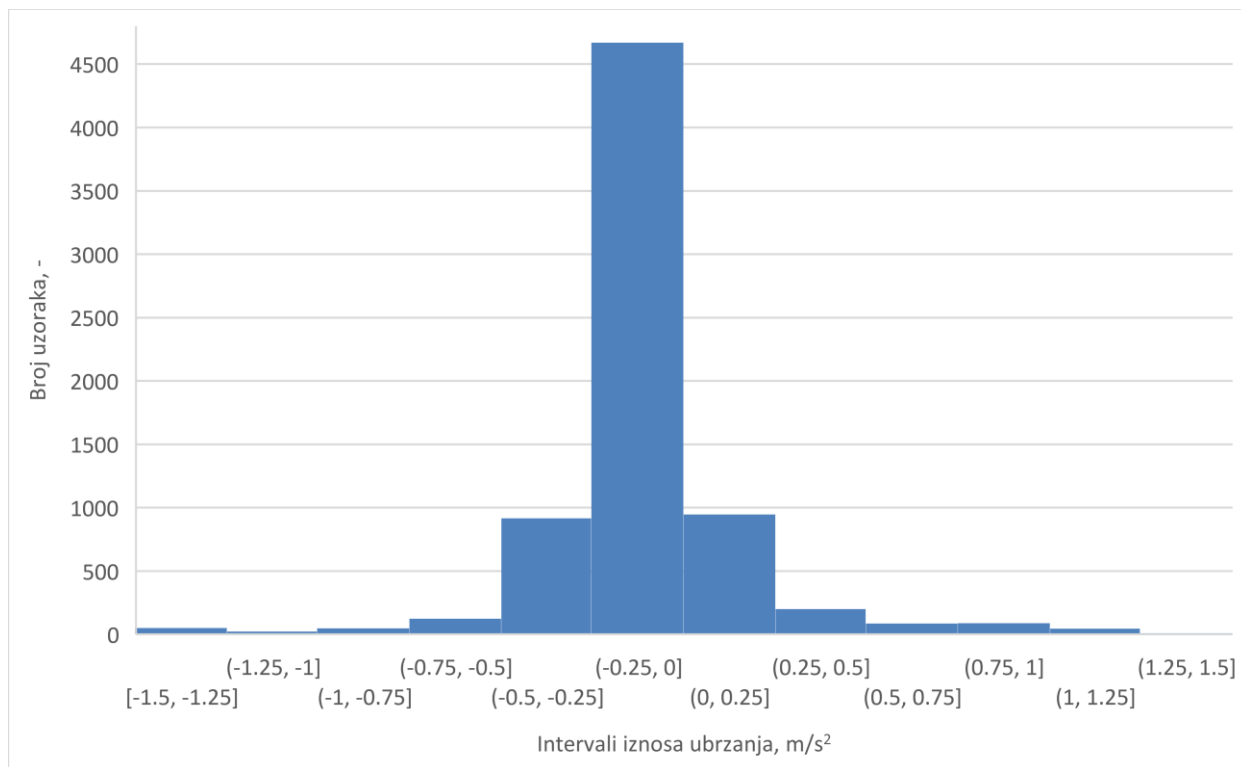
Slika 22. Raspodjela ubrzanja pješaka

Analogno prikazima za prethodne vrijednosti, prikazana su i sva ubrzanja pješaka dobivena u snimljenim podacima.

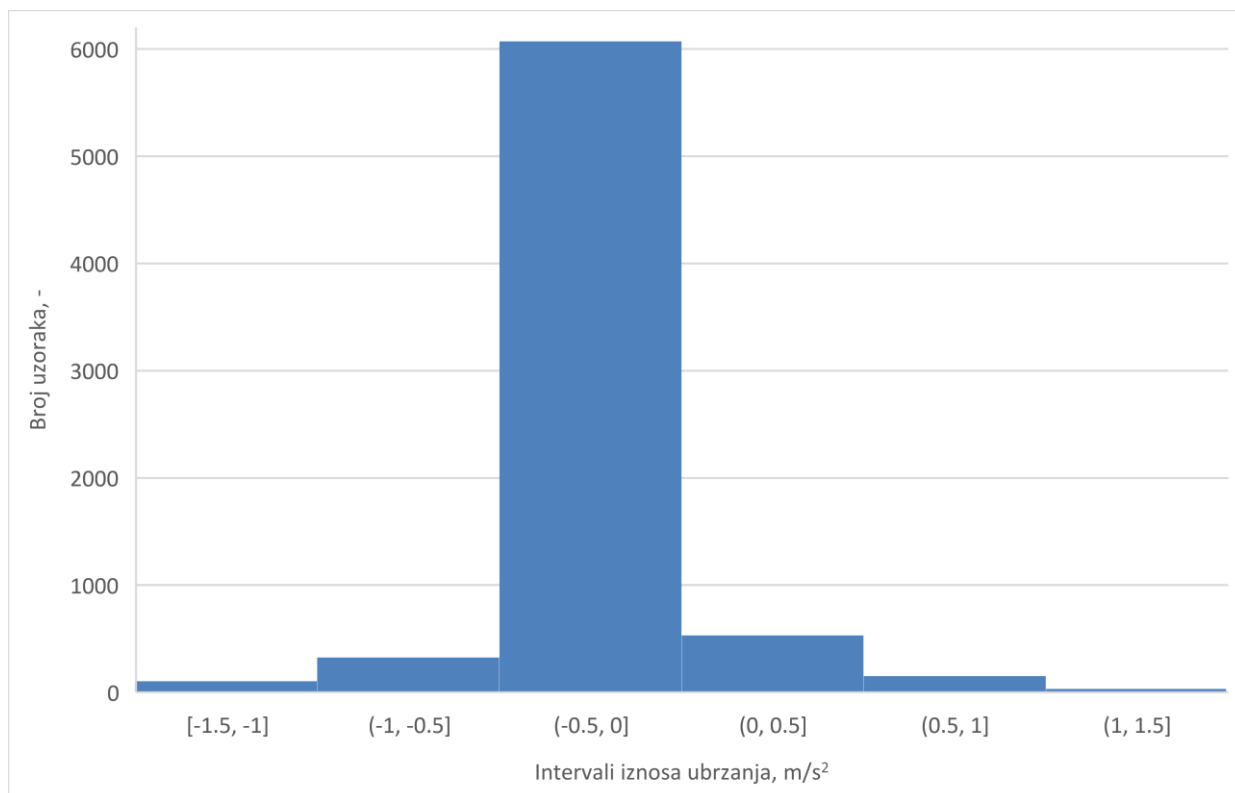
Na Slici 22 je vidljivo da se iznosi ubrzanja kreću između iznosa od $-1,56976 \text{ m/s}^2$ i $1,48024 \text{ m/s}^2$. Pošto taj raspon vrijednosti nije prevelik, zaokruživanje iznosa ubrzanja na cijele vrijednosti dalo bi iznose od -2 m/s^2 , -1 m/s^2 , 0 m/s^2 i 1 m/s^2 , što je poprilično pogrubljanje uzevši u obzir raspon različitosti početnih vrijednosti.

Za rezoluciju ubrzanja pješaka razmatrane su rezolucije od $0,25 \text{ m/s}^2$, $0,5 \text{ m/s}^2$ te od 1 m/s^2 .

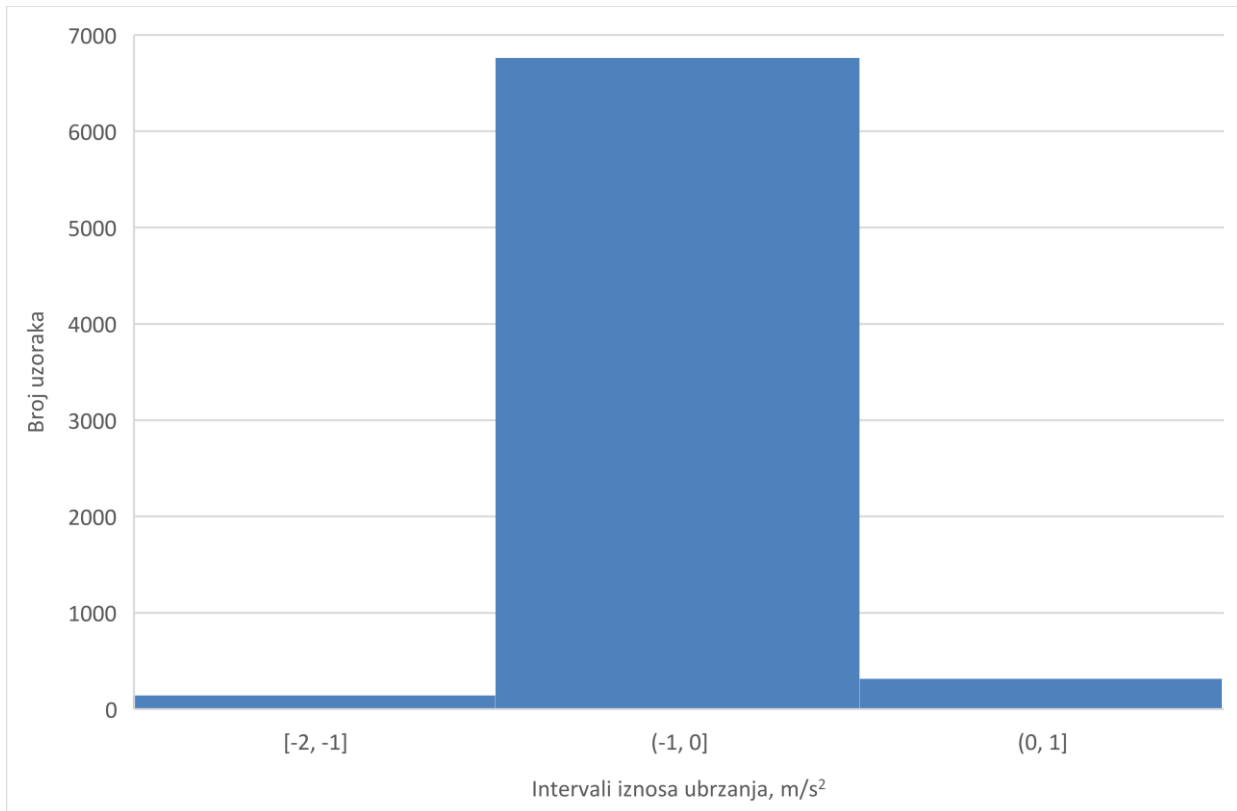
Raspodjela vrijednosti prikazana je dijagramima na Slici 23, Slici 24 i Slici 25.



Slika 23. Raspodjela ubrzanja pješaka u slučaju rezolucije od 0,25 m/s²



Slika 24. Raspodjela ubrzanja pješaka u slučaju rezolucije od 0,5 m/s²



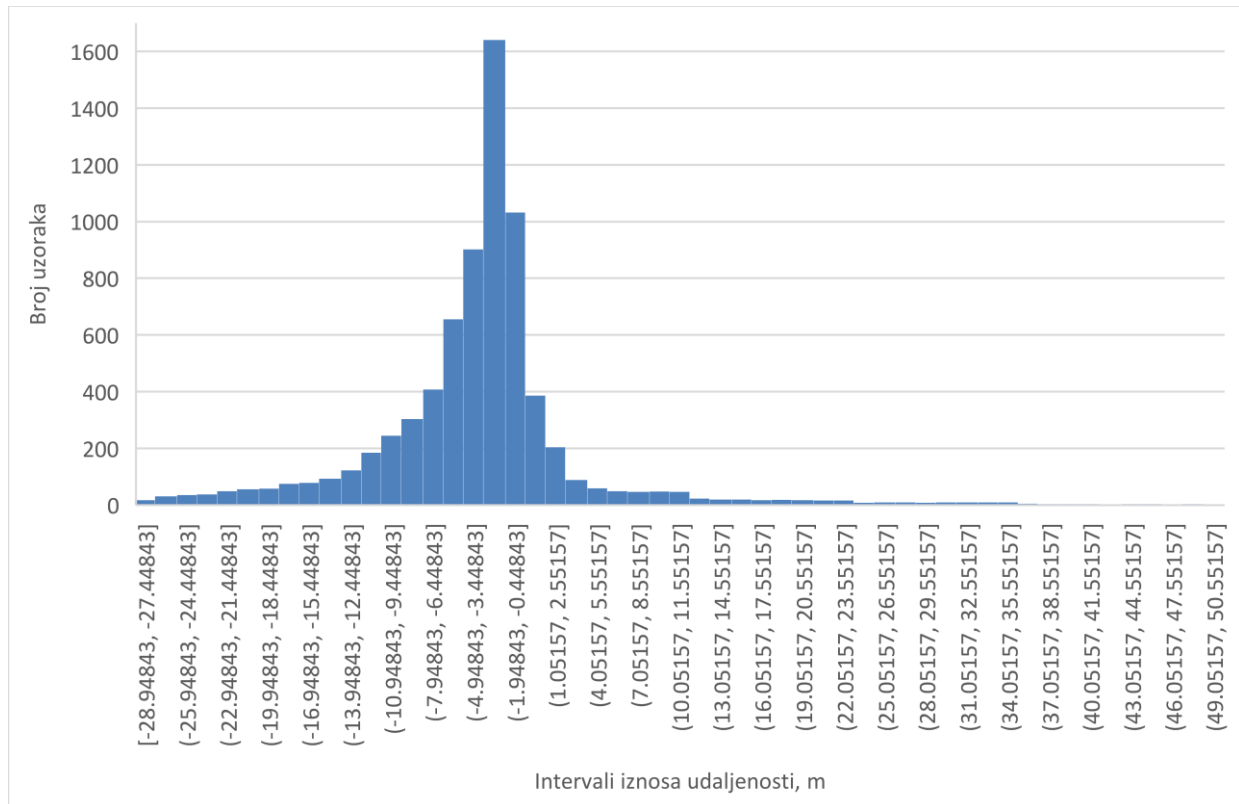
Slika 25. Raspodjela ubrzanja pješaka u slučaju rezolucije od 1 m/s²

Uvidom u prethodne tri slike vidljivo je da je većina vrijednosti koncentrirana između -1 m/s² i 0 m/s² te je kao rezolucija odabrana ona od 1 m/s².

4.2.1.2. Odabir rezolucija za podatke o vozilima

Analogno odabiru rezolucija za varijable vezane uz pješake, odabiru se rezolucije za vrijednosti varijabli vozila. Kao što je već napomenuto, veći fokus je stavljen na analizu ponašanja pješaka te su time varijable vezane uz vozila grublje raspodijeljene.

Slika 26 prikazuje raspodjelu neobrađenih udaljenosti vozila.

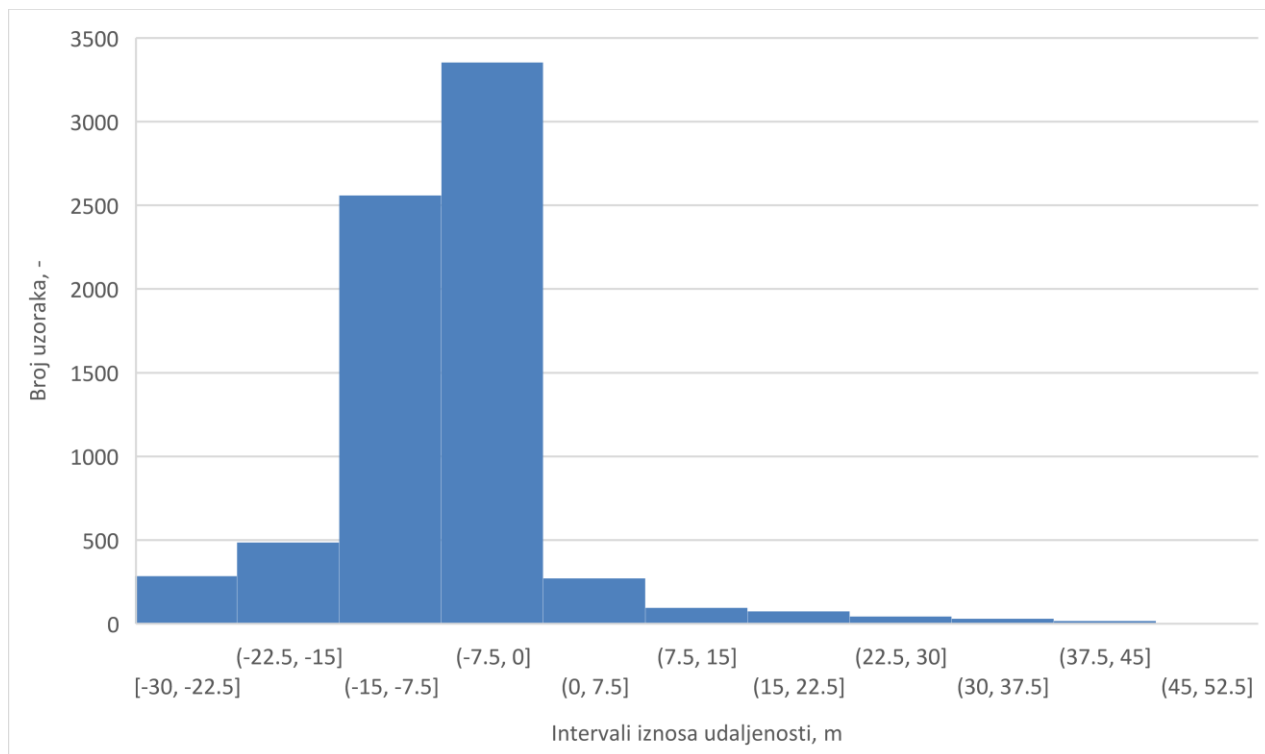


Slika 26. Raspodjela snimljenih udaljenosti vozila

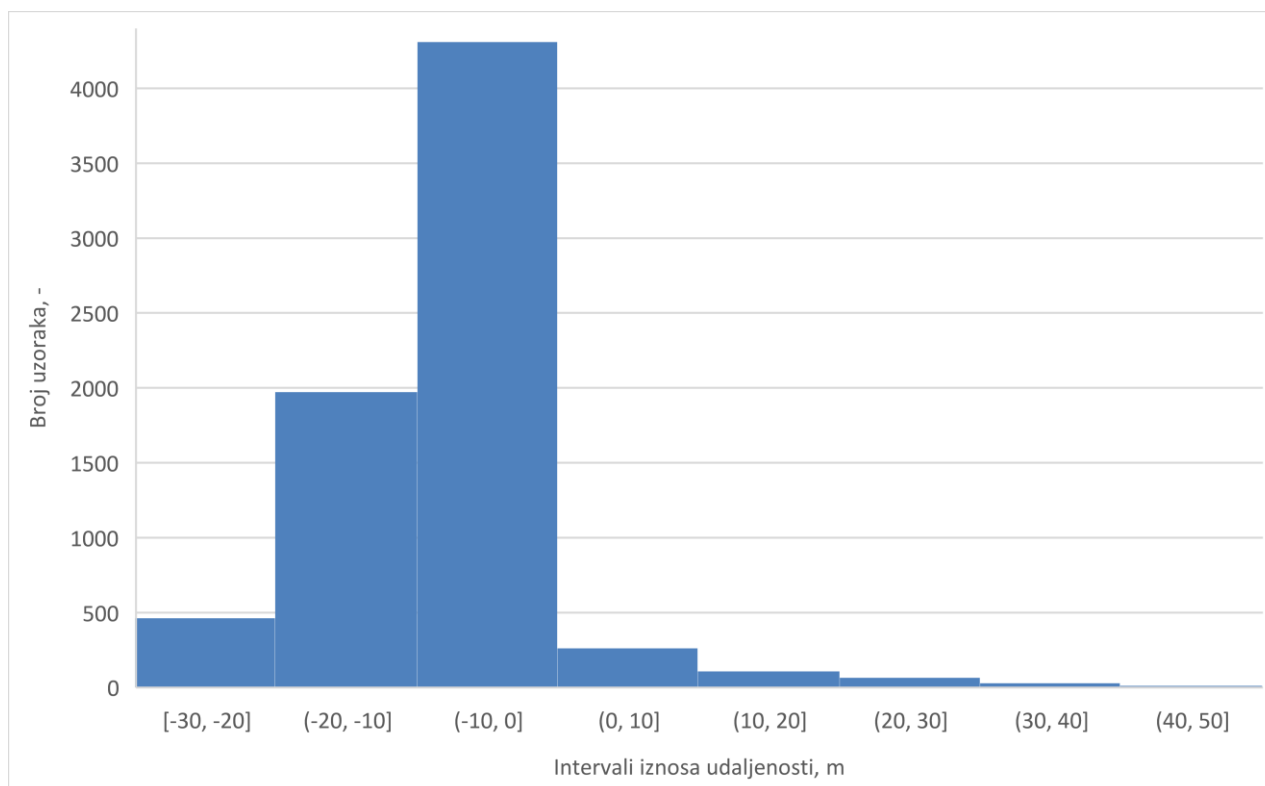
Uzevši u obzir dijagram na slici gore, dolazi se do informacije kako se udaljenost vozila unutar snimljenih vrijednosti kreće u rasponu od -28,94843 m do 50,55157 m. Tako širok raspon rezolucije rezultat je razlika u položaju vozila u odnosu na pješaka u rasponu vremenskih okvira uzetih za obradu i zapravo je i poželjan s obzirom na to da je svakako cilj dobiti model koji je treniran u čim raznovrsnijim uvjetima.

Što se tiče odabira rezolucije, u slučaju zaokruživanja iznosa udaljenosti vozila na cijele brojeve, i dalje bi se dobilo 80 unikatnih vrijednosti, što je svakako previše. Za rezoluciju od 5 m ta vrijednost iznosi 16, što je puno manje od 80, ali i dalje pretjerano s obzirom na činjenicu da je se većina stanja kreće u rasponu od -10 do 2 m.

Naredni grafovi prikazuju raspodjelu vrijednosti kada su uzete rezolucije od 7,5 m [Slika 27] i 10 m [Slika 28].



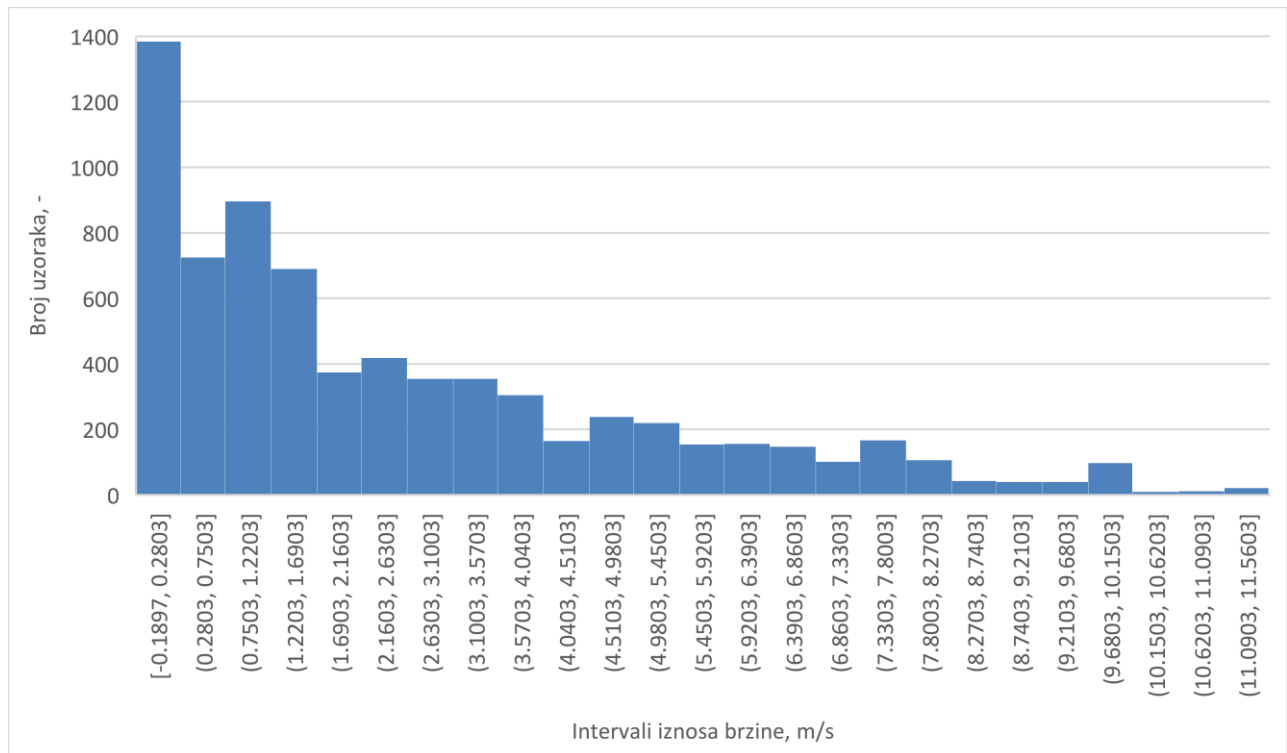
Slika 27. Raspodjela iznosa udaljenosti vozila u slučaju rezolucije od 7,5 m



Slika 28. Raspodjela iznosa udaljenosti vozila u slučaju rezolucije od 10 m

S obzirom na približno jednak broj vrijednosti kod obje rezolucije, obje su uzete za kasnije razmatranje u setovima rezolucija, prikazanim u sljedećem potpoglavlju.

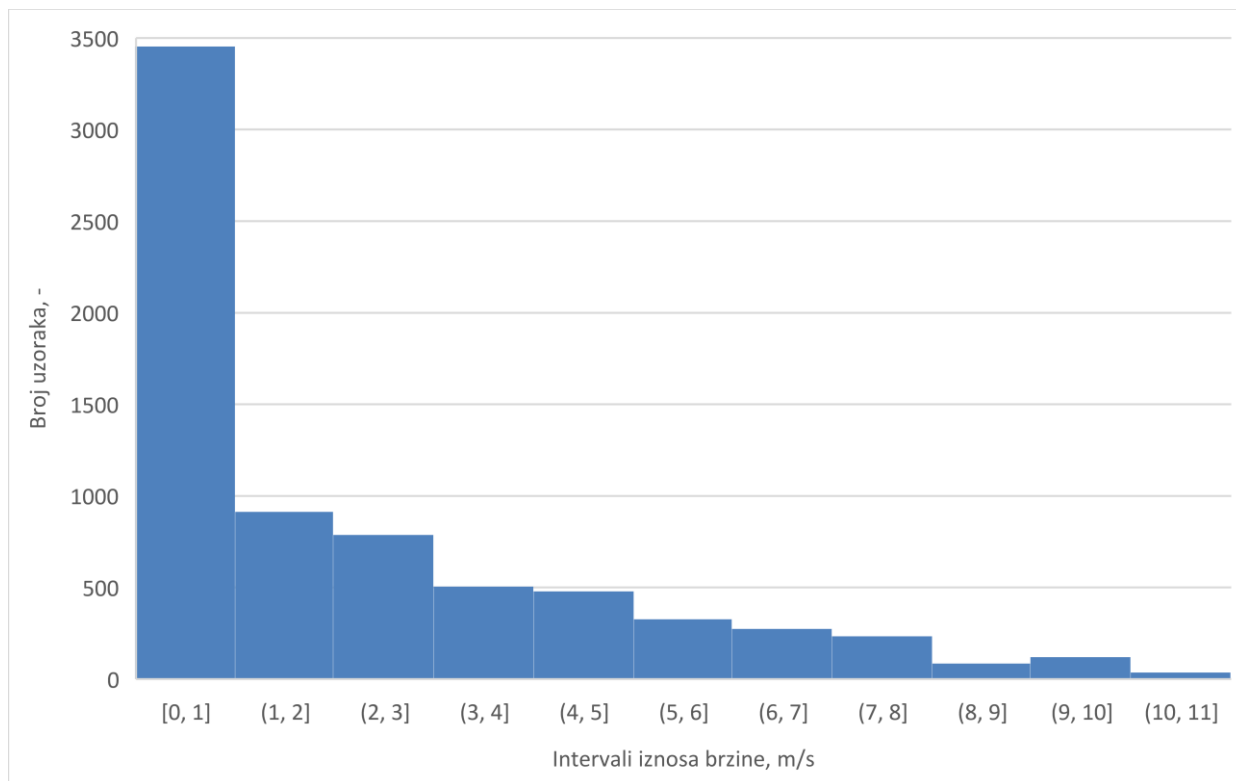
Nadalje, iz dijagrama na Slici 29 vidljivo je da se raspon brzina vozila kreće od -0,1897 m/s do 11,5603 m/s.



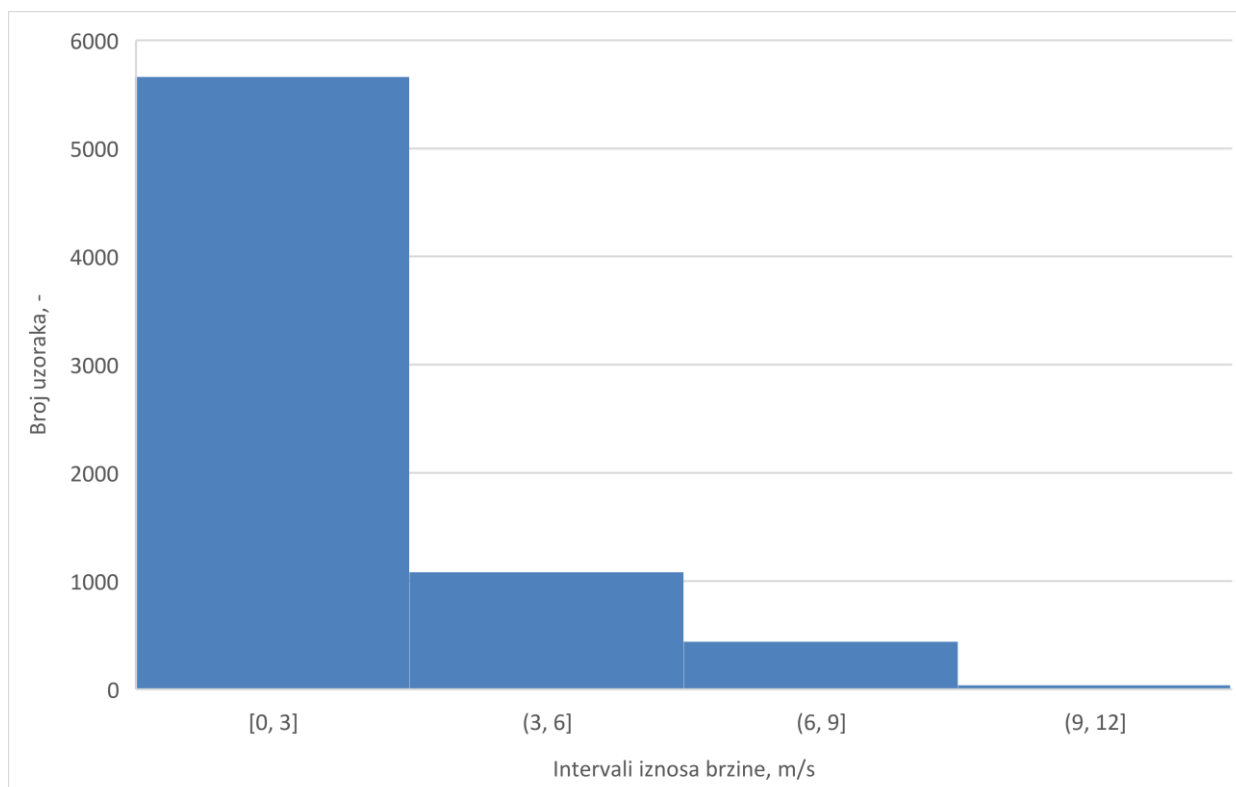
Slika 29. Raspodjela snimljenih brzina vozila

Ako bi se za brzinu vozila uzimale cjelobrojne vrijednosti, broj različitih mogućnosti bio bi 12. Svako dodatno pogrubljivanje rezolucije pokazuje znatne razlike s fizikalne strane jer je razlika pri ponašanju pješaka u situaciji kada nailazi vozilo brzinom od 12 m/s u odnosu na brzinu od 1 m/s puno veća nego kada nailazi vozilo na udaljenosti 12 m u odnosu na 1 m.

Budući da se većina podataka nalazi na lijevom dijelu dijagrama, to jest, brzine su većinom manjih iznosa, kompromis u vidu rezolucije brzine i smanjivanja točnosti je rezolucija od 3 m/s prikazana na Slici 31.



Slika 30. Raspodjela iznosa brzine vozila u slučaju rezolucije od 1 m/s



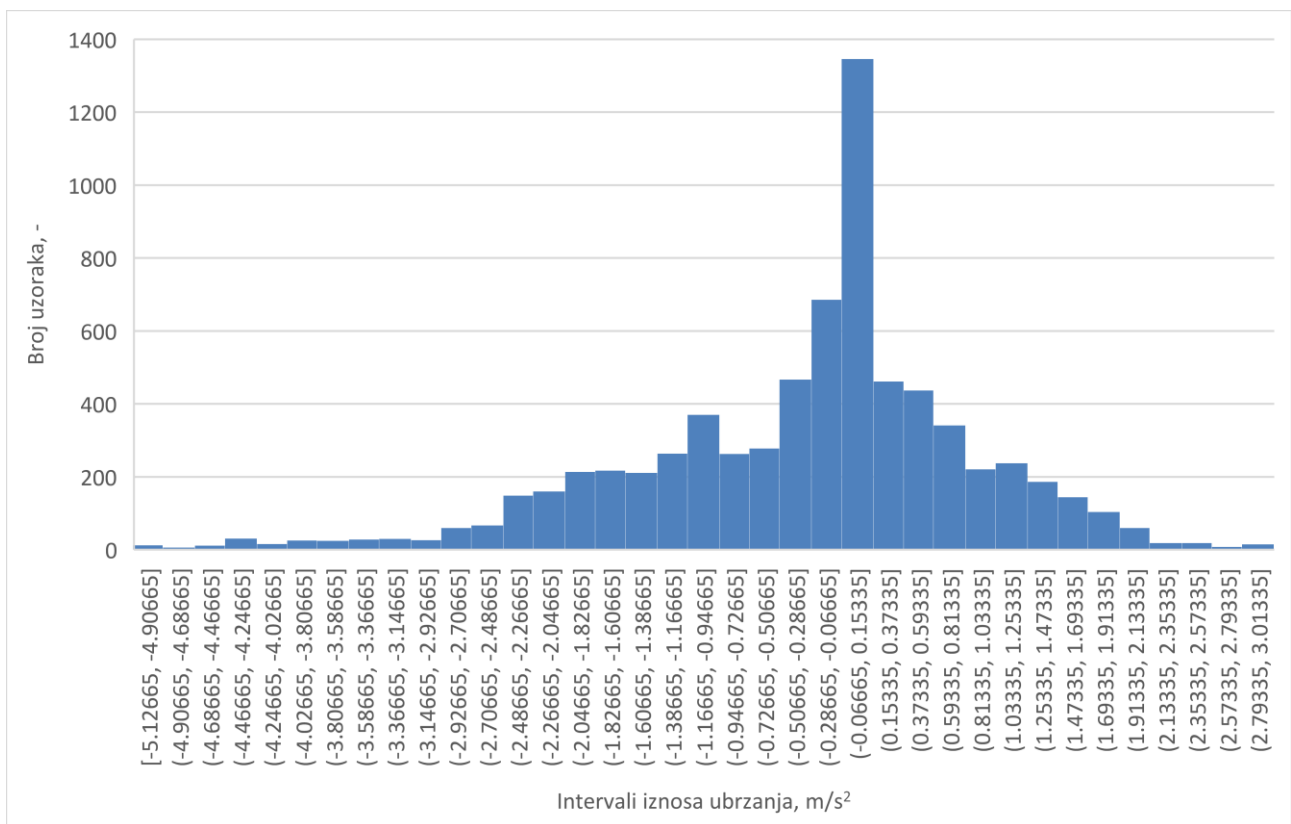
Slika 31. Raspodjela iznosa brzine vozila u slučaju rezolucije od 3 m/s

Slika 32 prikazuje raspodjelu ubrzanja vozila unutar snimljenih podataka.

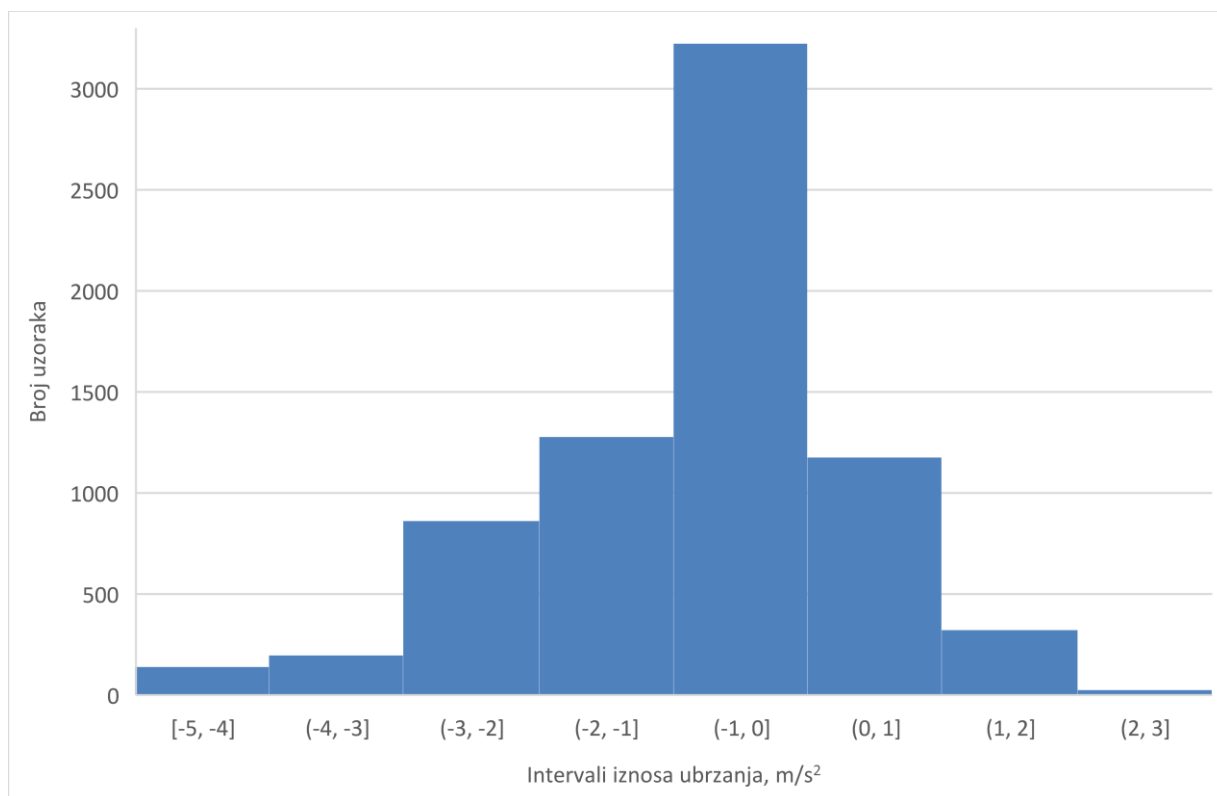
Najmanje zabilježeno ubrzanje je $-5,12665 \text{ m/s}^2$, dok najveće iznosi $3,01335 \text{ m/s}^2$.

Pri odabiru rezolucije za iznos ubrzanja vozila, razmatrane su rezolucije od 1 m/s^2 i $1,5 \text{ m/s}^2$. Prva rezolucija dala je pojavnost od 9 unikatnih vrijednosti, dok druga rezolucija definira 6 različitih vrijednosti. Ponovno, s obzirom na količinu vrijednosti u blizini iznosa od 0 m/s^2 , kao dovoljnom rezolucijom i bez uzrokovanja pretjerane greške u modelu odabrana je rezolucija od $1,5 \text{ m/s}^2$.

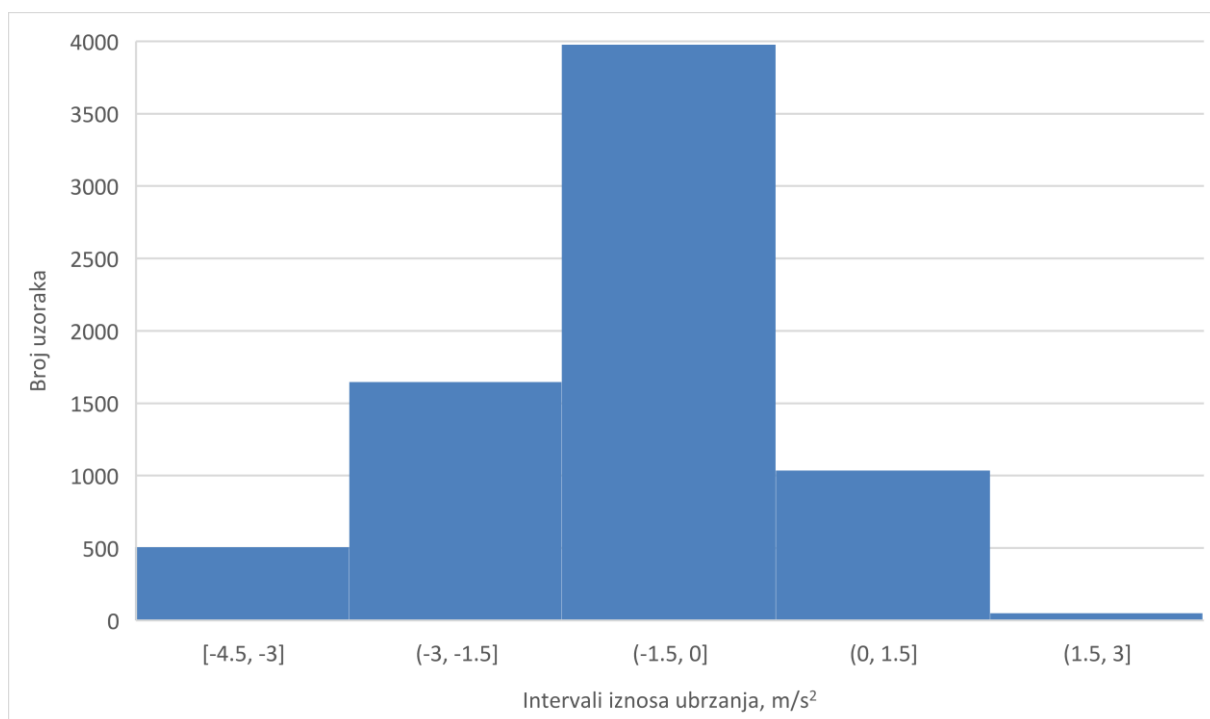
Slika 33 i Slika 34 prikazuju raspodjelu iznosa brzine vozila u slučaju rezolucija od 1 m/s^2 te slijedno $1,5 \text{ m/s}^2$.



Slika 32. Raspodjela snimljenih ubrzanja vozila



Slika 33. Raspodjela iznosa ubrzanja vozila u slučaju rezolucije od 1 m/s²



Slika 34. Raspodjela iznosa ubrzanja vozila u slučaju rezolucije od 1,5 m/s²

Kako su prve tri informacije unutar svakog vremenskog okvira suviše u kreiranju rječnika, nakon odabira rezolucija i pretvorbe u iste se odbacuju. Za isto se koristi potprogram *uklanjanje.py* u *Pythonu* koji iz datoteke *reducirani.csv* učitava pojedinu vrijednost te preskače njene prve tri informacije, a ostale sprema u dokument *uklonjeno.csv*. Svako stanje prikazano je u obliku liste s pojedinim elementima odvojenim zarezima. Novonastala stanja sastoje se od 6 varijabli relevantnih za međusobno uspoređivanje.

4.2.1.3. Kreiranje rječnika i krajnji odabir rezolucija varijabli

Temeljem zaključaka o rezolucijama udaljenosti, brzine i ubrzanja, izvedenim u prošlom potpoglavlju, u užu izbor za odabir rezolucija stavljena su tri seta rezolucija. Sva tri usporedno prikazuje Tablica 8.

Tablica 8. Rezolucije podataka unutar različitih setova.

	<i>Set 1</i>	<i>Set 2</i>	<i>Set 3</i>
Udaljenost pješaka	1 m	0,75 m	0,75 m
Udaljenost vozila	7,5 m	7,5 m	10 m
Brzina pješaka	0,5 m/s	0,5 m/s	0,5 m/s
Brzina vozila	3 m/s	3 m/s	3 m/s
Ubrzanje pješaka	1 m/s ²	1 m/s ²	1 m/s ²
Ubrzanje vozila	1,5 m/s ²	1,5 m/s ²	1,5 m/s ²
Broj ključeva (diskretnih stanja Markovljevog lanca)	504	599	572
Broj ključeva bez testnog seta	466	540	514

Za izračun rezolucija, točnije, pretvaranje podataka u podatke određenog raspona, kreiran je potprogram u *Pythonu* naziva *rezolucije.py*. Princip funkcioniranja mu je takav da učitava listu pojedinog stanja iz dokumenta *uklonjeno.csv* te za odabrani element liste stanja preračunava snimljeni iznos u iznos željene rezolucije. S obzirom na to da je položaj elementa u listi stanja unificiran (jednak za svako stanje), lociranje željenog elementa vrši se pozivanjem položaja istog unutar stanja (stanje je definirao kao lista vrijednosti informacija o stanju, primjerice udaljenosti, brzine i ubrzanja, u *Pythonu*). Promjena rezolucije vrši se za svaki element, naredbom za naredbom.

Naknadno, svako potpuno izmijenjeno stanje sprema se u datoteku koja ponovno sadrži 7221 listu (skup podataka) stanja.

S obzirom da je potrebno odabrati između 3 različita seta rezolucija, kreirane su tri datoteke s rezolucijama, naziva analognih nazivima setova – *rezolucija1.csv*, *rezolucija2.csv* te *rezolucija3.csv*. Svaka od tih datoteka koristi se kao ulaz u potprogram koji kreira rječnik sadržan od stanja te vraća informaciju o broju ključeva (unikatnih vrijednosti).

Potprogram *rjecnik.py*, kao što je već navedeno, kao ulaz uzima jednu od datoteka sa stanjima određene rezolucije. Prije toga, u potprogramu je kreirana lista naziva *kljucevi*, kojoj je cilj sadržavati sve ključeve. Isto tako, kreirana je i lista naziva *vrijednosti*. Iz jedne od datoteka *rezolucijaN.csv* (generalnim zapisom za sve tri datoteke smatra se *rezolucijaN.csv*) učitavaju se pojedine liste stanja, te se provjerava postoje li vrijednosti te liste koje podrazumijevaju udaljenosti, brzine i akceleracije u listi ključeva. Ako ne postoje, kao podatak se iz tipa liste (eng. *list*) pretvaraju u niz (eng. *tuple*), jer ključ rječnika ne može biti lista, te u tom formatu spremaju u listu *kljucevi*.

U slučaju da se stanje već nalazi u listi *kljucevi*, smatra se da je došlo do situacije u snimci u kojoj se stanje između dva vremenska okvira nije promijenilo. Samim time, trenutno stanje se sprema kao vrijednost prethodnog ključa jer ono slijedi stanje ključa (stanje u koje se prelazi iz stanja ključa), pozicionirajući se na prvo mjesto prve novokreirane liste *vrijednostipokljucu* unutar liste *vrijednosti*. Isto se ponavlja za svako stanje koje slijedi vrijednost prethodno postavljenu kao vrijednost ključa. Svakom ključu je kao vrijednost pridružena lista svih stanja u koje on prelazi (lista *vrijednostipokljucu*).

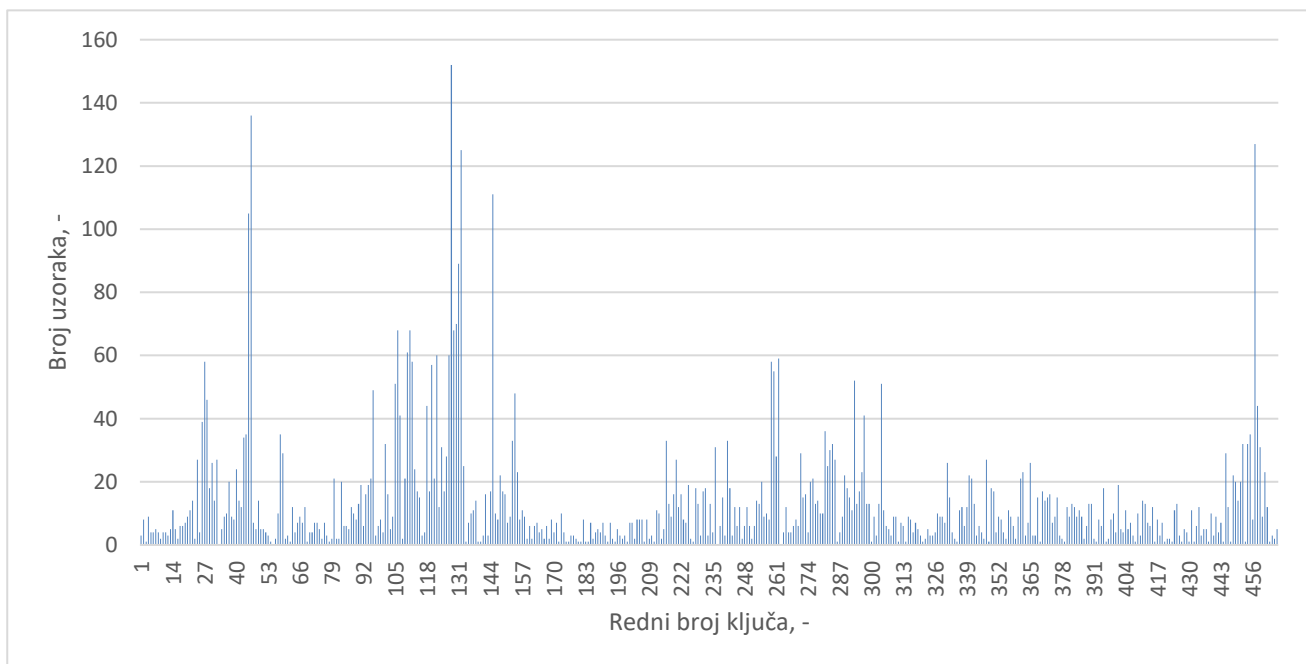
Ovdje postoje i iznimke. S obzirom na to da se datoteka *rezolucijaN.csv* sastoji od svih stanja svih snimki, čitanjem svakog reda u datoteci prelazi se iz jedne snimke u drugu. Zadnje stanje neke snimke se nužno mora postaviti kao ključ s vrijednosti prazne liste u slučaju da naknadno ne postoji istovrijedno stanje sa slijednim stanjem. Nikako ne smije doći do situacije gdje bi se iz zadnjeg stanja snimke prešlo u prvo stanje iduće snimke, ako je fizikalna korelacija među njima nemoguća.

Kreirane liste *kljucevi* i *vrijednosti* postavljaju se kao elementi rječnika ekvivalentno svojim nazivima. Gotov rječnik sprema se u svojstveni dokument iz kojeg se kasnije učitava radi kreiranja simulacija.

I prije samog stvaranja rječnika, nakon obrade svih stanja iz dokumenata *rezolucijaN.csv*, moguće je jednostavnom naredbom dobiti broj elemenata u listi *kljucevi*. Taj broj predstavlja

broj unikatnih stanja iz kojih se može prijeći u neko drugo stanje. Broj ključeva u svakom rječniku daje Tablica 8 u zadnjim redovima.

S obzirom na sličnost rezolucija, od predloženih setova je odabran *Set 1* jer sadrži najmanje različitih ključeva, što znači da postoji više vrijednosti za svaki ključ u odnosu na preostala dva seta. Broj vrijednosti po setu je izuzetno značajan za održavanje slučajnosti pri odabiru sljedećeg stanja. Histogram na Slici 35 prikazuje broj vrijednosti po ključu novokreiranog rječnika stanja.



Slika 35. Broj vrijednosti po ključu unutar rječnika stanja

4.2.2. Algoritam za kreiranje modela

Nakon definiranja rječnika, naredni korak je kreiranje modela. S obzirom na postojanje rječnika koji sam po sebi predstavlja sva stanja i sve vjerojatnosti prelaska iz jednog stanja u neko sljedeće, potrebno je kreirati algoritam za automatizirano generiranje skupa stanja. Budući da je cilj da simulirana stanja u čim većem broju budu slična stvarnima, potrebno je odabrati skup početnih stanja. Za ta stanja odabrana su sva u kojima se pješak nalazi na udaljenosti od 5 ili više metara od početne točke sjevernog dijela pješačkog prijelaza. Da nije postavljeno to ograničenje, kao početno stanje bi bilo moguće odabrati i stanje koje se nalazi na sredini pješačkog prijelaza ili pak pri kraju, nakon čega slijedi možda samo jedno stanje. Takva simulacija, koja bi se sastojala od sveukupno 2 stanja, nije referentna za usporedbu sa stvarnim snimkama.

Lista sa svim početnim stanjima nazvana je *pocetni* i kreira se učitavanjem svakog pojedinog stanja te uspoređivanjem prve vrijednosti u listi s brojem -5. U slučaju da je broj veći ili jednak -5, stanje se upisuje u listu *pocetni*.

Princip funkcioniranja algoritma za kreiranje modela je sljedeći: iz liste *pocetni* nasumično se odabire stanje i postavlja kao prvi izabrani ključ. Kao sljedeće stanje odabire se vrijednost pod tim ključem, također nasumično, te se njegova vrijednost postavlja kao idući ključ. Cijeli proces se ponavlja sve dok odabrano stanje, u ulozi ključa, za vrijednosti nema praznu listu ili dok ne sadrži samo jedno stanje kao vrijednost, a to stanje, to jest vrijednost, je jednaka ključu. Ovdje se kreiranje simulacije zaustavlja, sva stanja se redom spremaju u dokument *simulacije.csv* te se ponovno kreće u novi odabir prvog ključa iz liste *pocetni*. Više o principu nasumičnog odabira i korištenju dobivenih simulacija rečeno je u narednom poglavlju.

5. VALIDACIJA IZVEDENOG MODELA

5.1. Monte Carlo simulacija

5.1.1. Definicija i povijest

Monte Carlo simulacija, još poznata kao Monte Carlo metoda, odnosno simulacija vjerojatnosti s većim brojem mogućnosti, matematička je tehnika koja se koristi kako bi se procijenili mogući ishodi nesigurnog, tj. neizvjesnog događaja. Metoda je razvijena još za vrijeme Drugog svjetskog rata kako bi se poboljšalo donošenje odluka u neizvjesnim uvjetima, a ime je dobila prema poznatom gradu kasina, upravo iz razloga što se sama središnja ideja metode temelji na elementu slučajnosti, slično kao i rulet. [13]

Od njezina uvođenja, ova metoda koristi se za procjenu rizika u brojnim područjima i djelatnostima – od umjetne inteligencije i burzovnih vrijednosti preko predviđanja prodaje do projekt menadžmenta. Također, nudi brojne prednosti nad prediktivnim modelima s fiksnim ulaznim podacima, na primjer, omogućava provođenje analize osjetljivosti ili određivanje korelacije ulaznih podataka. Analiza osjetljivosti omogućava donositeljima odluka uvid u utjecaj pojedinih ulaznih podataka na izlaz, tj. ishod, dok korelacija omogućava razumijevanje odnosa među bilo kojim ulaznim podacima. [13]

5.1.2. Primjena

Za razliku od klasičnih predviđajućih modela koji koriste skup konstantnih, odnosno fiksnih ulaza, Monte Carlo simulacija predviđa set ishoda temeljen na procijenjenom rasponu vrijednosti ulaznih podataka. Drugim riječima, Monte Carlo simulacija gradi model mogućih rezultata prema raspodjeli vjerojatnosti, npr. jednolikoj ili normalnoj, za bilo koju varijablu čija je vrijednost nesigurna, odnosno neizvjesna. Zatim ponovno računa rezultate u više iteracija, koristeći uvijek različit skup nasumičnih brojeva između gornje i donje definirane granice. U realnim primjenama, ovakav iterativan postupak provodi se više tisuća puta kako bi se dobio velik broj mogućih ishoda.

Neovisno o konkretnom alatu koji se koristi, provođenje Monte Carlo simulacije odvija se u sljedeća tri koraka:

1. postavljanje predviđajućeg modela, identificirajući zavisne varijable – ishode, te i sve nezavisne varijable, tj. ulazne podatke

2. određivanje raspodjele vjerojatnosti nezavisnih varijabli. Korištenje zapisanih podataka, osobne procjene ili bilo koje druge tehnike ili metode moguće je i potrebno kako bi se definirao raspon mogućih vrijednosti i dodijelili težinski faktori svakoj varijabli
3. ponavljajuće provođenje samih simulacija, generirajući time nasumične vrijednosti nezavisnih varijabli (rezultata, tj. ishoda). Ovaj korak provodi se dok se ne stvori reprezentativan uzorak približno beskonačnog broja mogućih kombinacija.

Prilikom obrade podataka, korisno je odrediti varijancu σ^2 i standardnu devijaciju σ (odstupanje), kao glavne pokazatelje raspodjele. Varijanca varijable je kvadrirana razlika njezinog stvarnog iznosa x_i i očekivanog iznosa, odnosno prosjeka x_m kako slijedi: (13)

$$\sigma^2 = \frac{1}{n} \sum_{i=1}^n (x_i - x_m)^2 \quad (13)$$

a standardno odstupanje je drugi korijen varijance. [13]

Prilikom korištenja Monte Carlo simulacija, drugi korak opisanog postupka otvara mogućnost korištenju brojnih alata, postupaka i metoda za definiranje i preračunavanje raspodjele vjerojatnosti nezavisnih varijabli. Upravo na taj način moguće je napraviti podjelu, tj. razlikovati podvrste Monte Carlo simulacija. U ovom radu, korištena je raspodjela vjerojatnosti na temelju kreiranog Markovljevog lanca prvog reda, čime je korištenu metodu moguće nazivati i Monte Carlo simulacija na temelju Markovljevih lanaca (MCMC – eng. *Markov Chain Monte Carlo simulation*). Kao sam algoritam generiranja nasumičnih brojeva odabran je *Pythonov* algoritam slučajnog odabira.

5.1.3. *Python algoritam slučajnog odabira*

Korištenje *Pythonovog* algoritma slučajnog odabira (eng. *random*) ne daje potpuno nasumične podatke; podatke dobivene ovom metodom prikladnije je nazivati pseudo-nasumičnim podacima (eng. *pseudorandom*). Navedeni modul za stvaranje „nasumičnih“ podataka koristi algoritam *Mersenne Twister*. [14] Ovaj algoritam generira 53-bitne realne brojeve, a njegov period, u smislu duljine perioda prije ponavljanja, iznosi $2^{19937}-1$, što ga čini vrlo moćnim. Ovako dobivene podatke, iako se čine nasumičnima, moguće je reproducirati. Međutim, za statističko modeliranje i provođenje simulacija, korištenje ovakvog algoritma, zbog njegove brzine, u potpunosti je prikladno. [15] Dodatno, s obzirom na opseg, odnosno količinu podataka

s kojima je u ovom radu rađeno, problemi s periodičkim ponavljanjem niza pseudo-nasumičnih podataka neće se pojaviti.

5.1.4. Nasumičnost u simulacijama

Nasumičnost u kreiranim simulacijama osigurana je korištenjem prethodno navedene naredbe *random()* unutar programskog jezika *Python*. Iako vjerojatnosti izbora neke vrijednosti kao sljedećeg stanja nisu eksplicitno navedene, daju se iščitati te izraziti temeljem broja jednakih vrijednosti u nekom ključu. Primjerice, ako stanje 1 u ulozi ključa sadrži, redom, stanja 2, 2, 3, 4, vjerojatnost da će se iz stanja 1 prijeći u stanje 2, to jest, da će se odabrati stanje 2, duplo je veća nego vjerojatnost da će se odabrati stanje 3 ili 4. Za dani primjer, vjerojatnost prelaska iz stanja 1 u stanje 2 je 0,5, dok vjerojatnost prelaska u stanje 3 ili stanje 4 iznosi po 0,25.

5.2. Opis korištenih metoda validacije

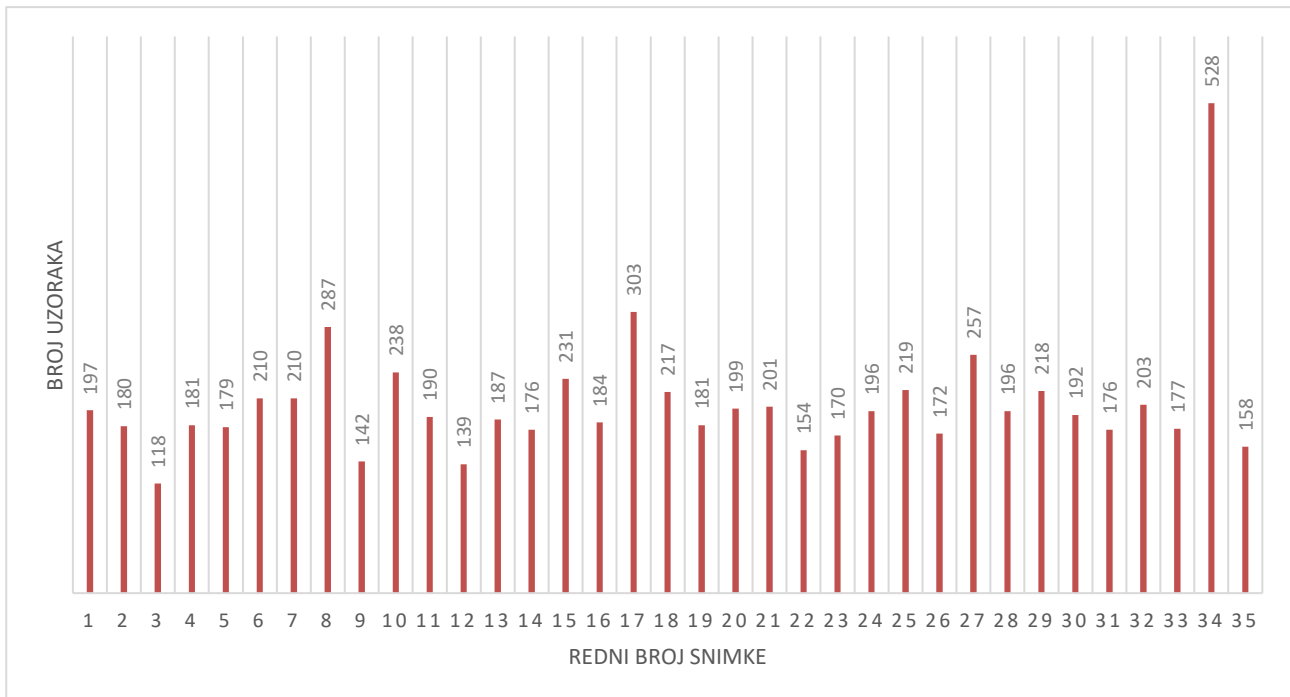
U ovom radu prikazano je više oblika validacije dobivenih rezultata (modela prelaska preko pješačkog prijelaza). Za provjeru točnosti simulacija, od 35 početnih setova, 5 je uzeto za naknadnu validaciju modela, to jest, odabrano je 5 interakcija koje neće sudjelovati u kreiranju rječnika. Krajnji cilj je tih 5 interakcija usporediti s rezultatima dobivenim simulacijom. Dalje se navode pod nazivom „testne simulacije“.

Tablica 9. Oznake simulacija po rednom broju.

	<i>Svi podaci</i>	<i>Treniranje</i>	<i>Testni set</i>
<i>Ukupni broj setova</i>	35	30	5
<i>Pješak prelazi prvi</i>	31	27	4
<i>Vozilo prelazi prvo</i>	4	3	1
<i>Pješak prelazi prvi</i>	88,57 %	90 %	80 %
<i>Vozilo prelazi prvo</i>	11,43 %	10 %	20 %

Tablica 9 prikazuje temelje odluke o broju od 5 testnih simulacija te broj i udio interakcija gdje pješak ili vozilo prvi prelaze preko pješačkog prijelaza, što je značajno za jednu od naknadnih validacija.

Testne simulacije su odabrane kako bi bile najbližije prosjeku svih ostalih simulacija. To je značilo da se težilo otprilike jednakoj raspodjeli situacija gdje pješak prelazi prvi preko pješačkog prijelaza u odnosu na situacije gdje isto radi vozilo prije pješaka. Također, kriterij je bio da testne simulacije imaju čim standardnija stanja po pitanju početnih stanja te da je trajanje originalne snimke prosječne duljine. Sva trajanja simulacija prikazana su histogramom na Slika 36.



Slika 36. Broj vremenskih okvira po interakciji

Redni broj interakcija na Slici 36 analogan je redosljedu kojim su interakcije prvi put zapisane u Tablici 6. Za testne interakcije odabrane su one pod rednim brojevima 11, 13, 16, 21 i 30 te su preimenovane u testne interakcije rednih brojeva 1-5. Redni brojevi prilagođeni su novonastalom skupu podataka za kreiranje rječnika. Tablica 10 prikazuje numeraciju svih interakcija koja se naknadno koristi kako bi se reducirao broj informacija. U istoj tablici, pod nazivima *Test 1* do *Test 5*, moguće je vidjeti kojih pet interakcija je odabrano za testni set.

Tablica 10. Numeracija interakcija po rednom broju.

Redni broj	1	2	3	4	5	6	7	8	9	10	11	12
ID snimke	18	18	19	19	19	19	20	20	20	20	22	22
ID pješaka	90	311	49	133	199	320	189	263	281	361	151	458

<i>ID vozila</i>	94	318	50	138	201	318	194	266	285	367	166	464
<i>Redni broj</i>	13	14	15	16	17	18	19	20	21	22	23	24
<i>ID snimke</i>	22	23	24	24	24	25	25	25	26	27	27	28
<i>ID pješaka</i>	541	188	1	39	122	183	205	245	38	60	158	23
<i>ID vozila</i>	551	192	34	40	130	180	211	248	4	66	163	25
<i>Redni broj</i>	25	26	27	28	29	30	Test 1	Test 2	Test 3	Test 4	Test 5	
<i>ID snimke</i>	28	28	28	29	29	29	21	22	23	25	28	
<i>ID pješaka</i>	99	320	467	235	404	498	63	451	53	63	134	
<i>ID vozila</i>	108	326	477	241	419	505	60	462	58	68	149	

U prvoj vrsti validacije, odabran je skup početnih stanja svake simulacije iz liste *pocetni* te se simulacija pokrenula. Dobiveno je 50 rezultata koji su prikazani grafički, zajedno s najslabijom interakcijom iz testnog skupa. Uz testni skup, iz skupa od 30 simulacija čija se stanja nalaze u rječniku, nekoliko skupova je naknadno odabrano za skupove pod nazivima i uspoređeno s rezultatima prethodnih 50 simulacija.

Druga vrsta validacije podrazumijevala je pokretanje 100 simulacija iz identičnog početnog stanja svake od 30 interakcija, te za pet testnih slučajeva. Kod testnih slučajeva, početno stanje se ne nalazi u rječniku te je odabrano stanje iz rječnika najslabije istom. Za svaku od simuliranih varijabli izračunat je korijen srednje kvadratne vrijednosti pogreške (RMSE, eng. *root mean squared error*) te je ista podijeljena sa brojem simulacija.

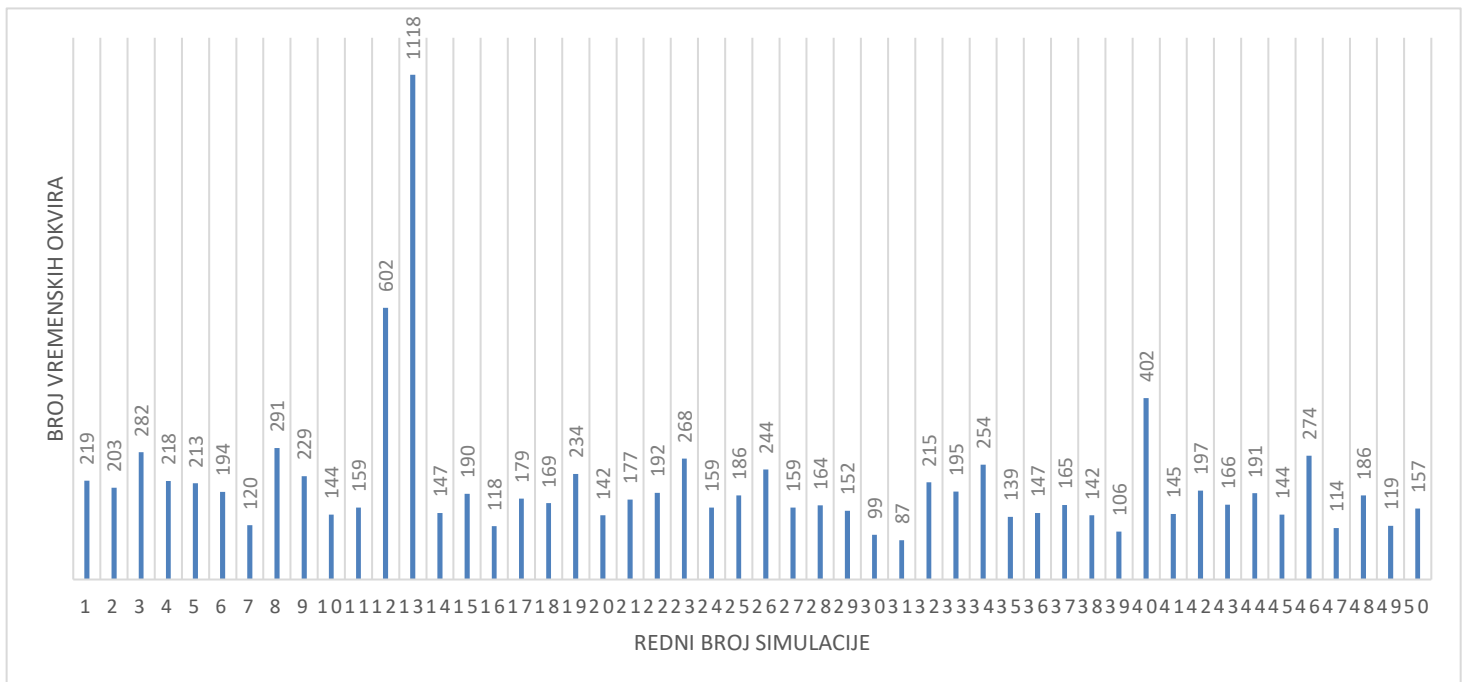
Kao treća vrsta validacije, za proteklih 35 x 100 simulacija, izračunat je srednji TTA (vrijeme do dolaska, eng. *Time to Arrival*) te uspoređen s vrijednostima iz snimljenog seta podataka.

5.3. Validacija pomoću 50 simulacija nasumičnog početnog stanja

Prva metoda validacije podrazumijeva kreiranje 50 simulacija s početnim stanjima iz liste *pocetni*. Početno stanje je pri kreiranju svake simulacije nasumično izabrano iz te liste.

Slika 37 prikazuje trajanje simulacija u broju vremenskih okvira. U prosjeku se većina simulacija zadržava na oko 190 vremenskih okvira, dok su jedine značajne iznimke simulacije pod rednim brojevima 12, 13 i 40. U sva tri slučaja, došlo je do „zapinjanja“ u istom stanju.

Navedeno zapinjanje se može objasniti na primjeru *Simulacije 13*. Uvidom u slijed svih stanja te simulacije, vidljivo je da se više puta ponavlja stanje ('0.0', '0.0', '0.0', '0.0', '0.0', '0.0'). U ovom stanju, redom, nalaze se udaljenost pješaka, udaljenost vozila, brzina pješaka, brzina vozila te ubrzanje pješaka nakon kojeg slijedi ubrzanje vozila. Unutar rječnika, ključ identičnog formata sadrži 127 vrijednosti, od kojih 126 iznosi ['0.0', '0.0', '0.0', '0.0', '0.0', '0.0'], a samo jedna ['0.0', '0.0', '0.5', '0.0', '0.0', '0.0']. To znači da je vjerojatnost da se među vrijednostima tog ključa odabere stanje sa svim nulama 99,21% (dobivena dijeljenjem broja stanja s nulama s ukupnim brojem vrijednosti, 126/127). Čak i kada bi se odabralo preostalo stanje, ono, kao ključ, među vrijednostima sadrži vrijednost jednaku stanju sa svim nulama. U slučaju *Simulacije 13* to se i dogodilo te je simulacija ponovno zapela u vrijednostima sa svim nulama. Iako su vjerojatnosti za takve događaje rijetke, i dalje su moguće, no poznavanje očekivanog rezultata simulacije lako omogućuje eliminiranje takvih ili sličnih dobivenih simulacija.



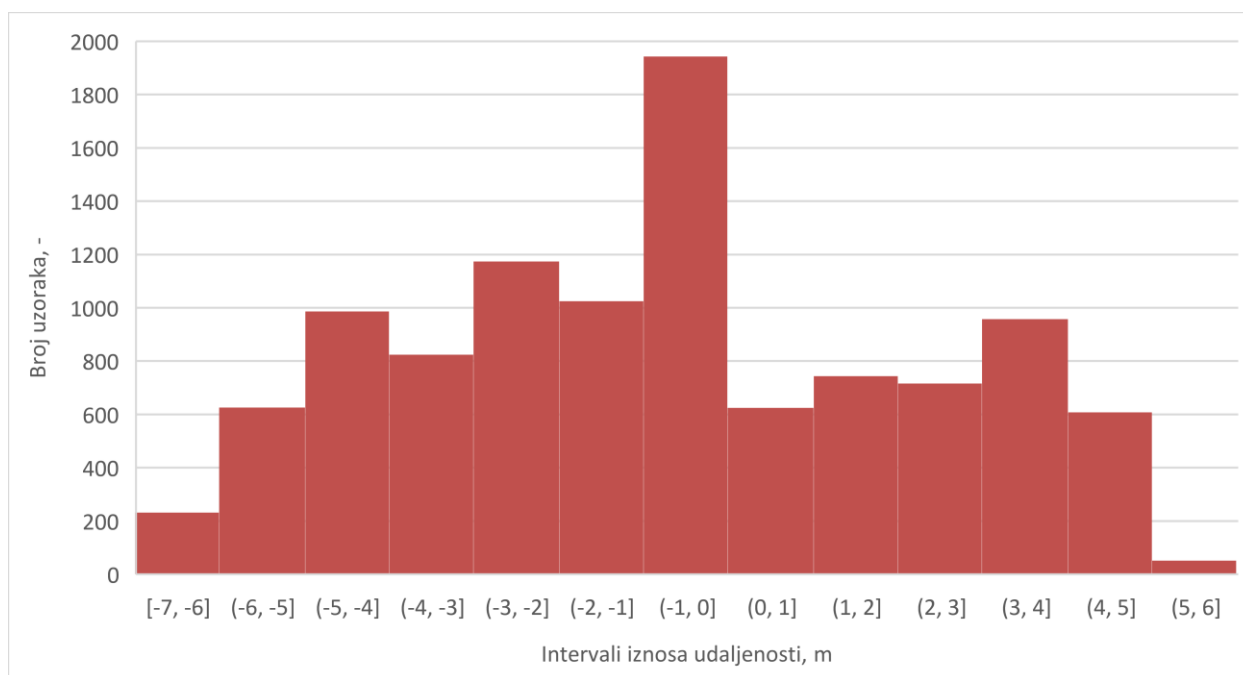
Slika 37. Broj vremenskih okvira po kreiranoj simulaciji

5.3.1. Raspodjela vrijednosti dobivenih simulacijama

Nadalje, analizirana je raspodjela vrijednosti za udaljenost, brzinu i ubrzanje pješaka i vozila dobivenih simulacijama. Svi prikazi vidljivi su na dijagramima na Slikama 38 - 43.

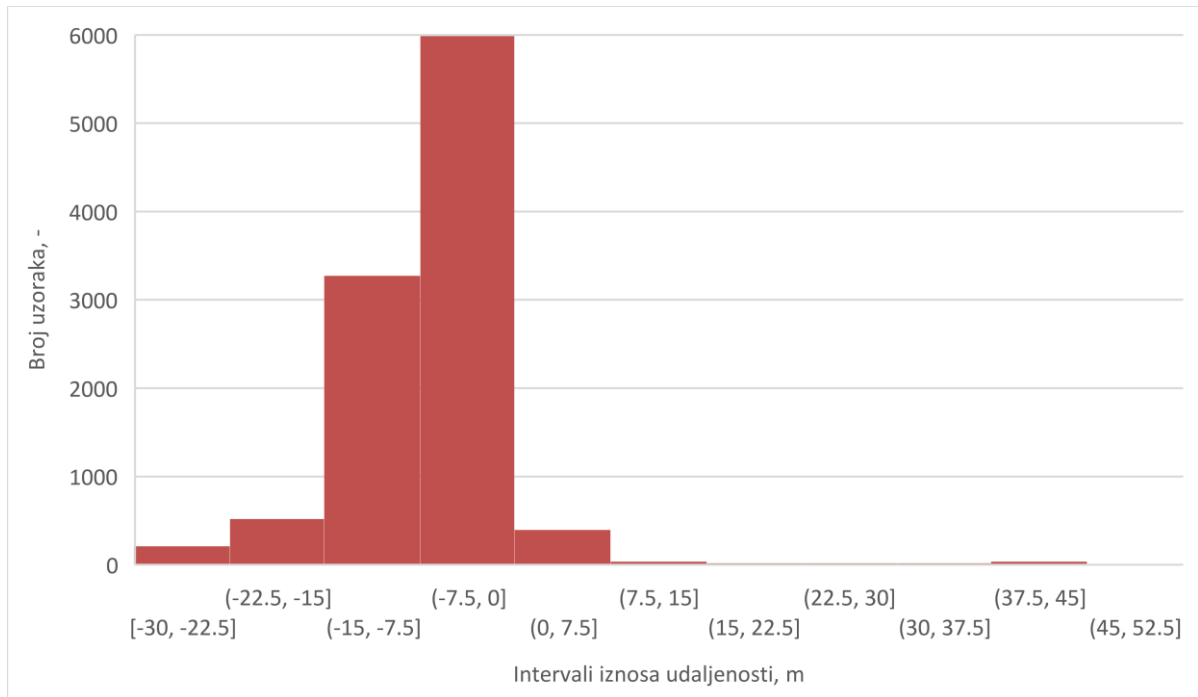
Prva od navedenih slika prikazuje raspodjelu vrijednosti udaljenosti pješaka unutar svih dobivenih simulacija. Od svih 6 prikazanih raspodjela, očekivano je kako će ova rezultirati

najsličnijim brojem uzoraka po vrijednosti. Isto se nije dogodilo s obzirom na prethodno spomenutu *Simulaciju 13* koja sadrži velik broj stanja na udaljenosti od 0 m. S obzirom na to da je ciljani prikaz raspodjele onaj na Slici 18, usporedbom oba dijagrama vidljivo je da raspodjela udaljenosti pješaka za ostale intervale vrijednosti zadovoljava. Jedina značajna razlika je manji broj udaljenosti u četvrtom stupcu koji prikazuje broj stanja na udaljenosti od (- 4, -3] m, što je vjerojatno posljedica nedovoljnog broja simulacija. Također, nije isključena ni mogućnost da je do toga došlo zbog činjenice da je jedno od početnih stanja snimke na -4 m, a ovdje se stanja takve udaljenosti pješaka nisu smatrala početnima za simulacije.



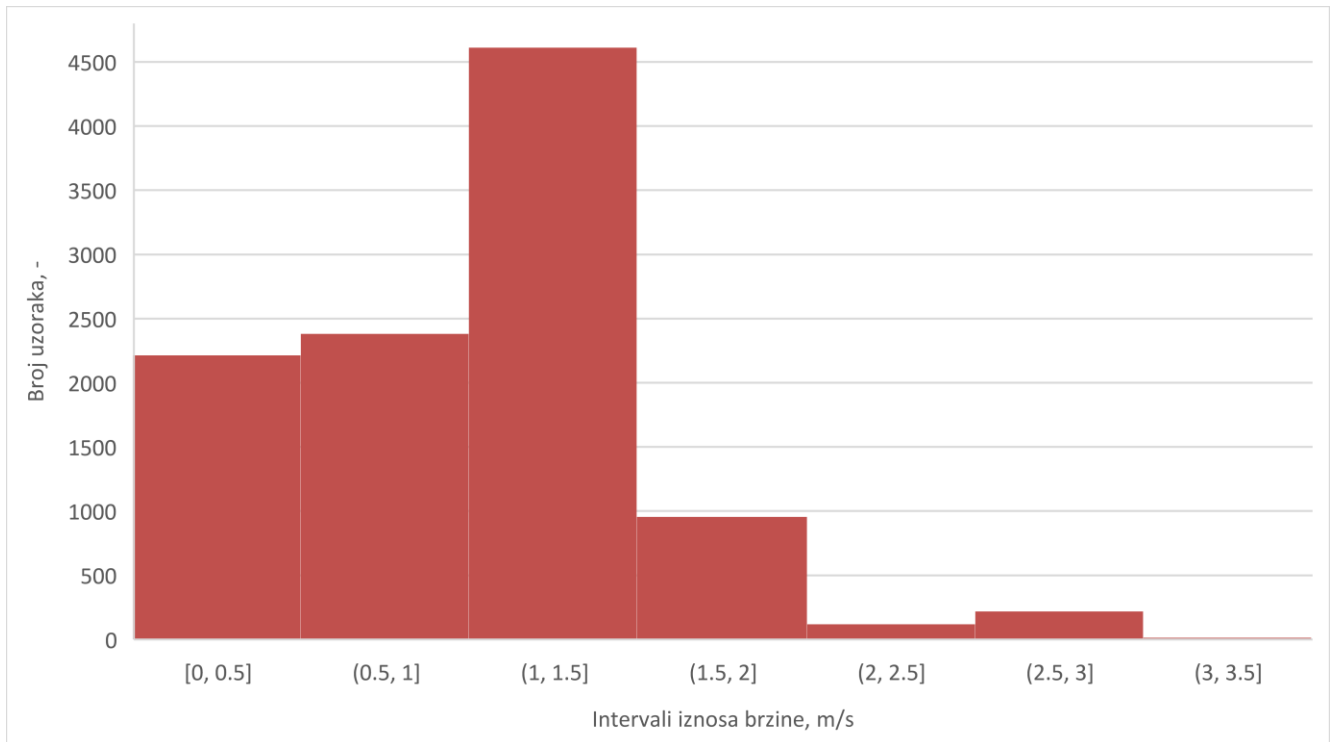
Slika 38. Raspodjela vrijednosti udaljenosti pješaka za sve podatke unutar 50 simulacija

Slično raspodjeli udaljenosti pješaka, i raspodjela udaljenosti vozila slijedi obrazac prikazan na Slici 27. I ovdje je vidljiv povećan broj uzoraka na udaljenosti od 0 m u odnosu na snimljene vrijednosti. Jedan od uzroka tome je svakako simulacija rednog broja 13. Od ostalih razlika, moguće je primjetiti smanjen broj uzoraka s udaljenostima većim od 7,5 m.



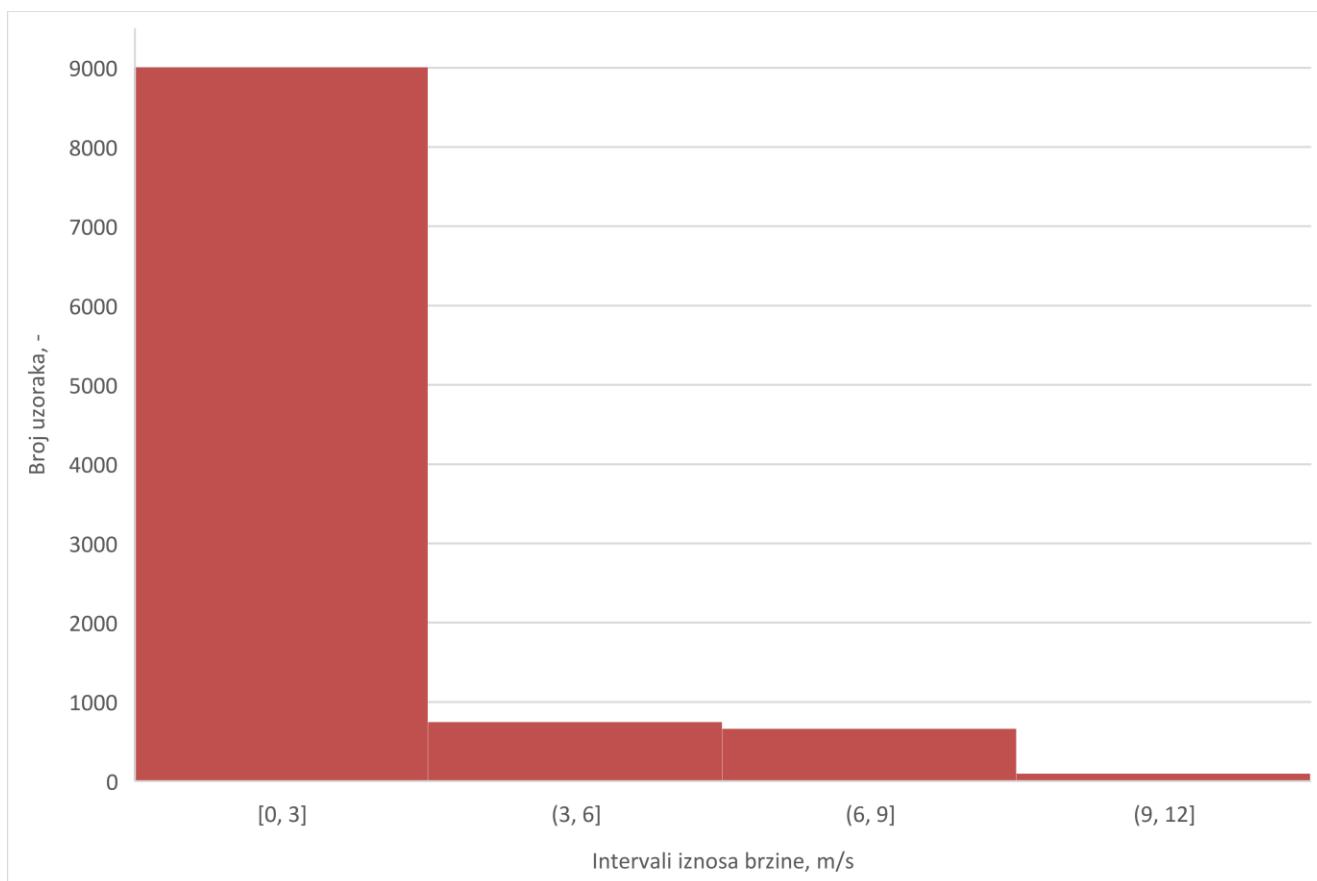
Slika 39. Raspodjela vrijednosti udaljenosti vozila za sve podatke unutar 50 simulacija

Raspodjela brzina pješaka za svih 50 simulacija po svom obrascu potpuno slijedi raspodjelu na Slici 21. Značajnija razlika može se, ponovno, primjetiti kao posljedica ponavljanja stanja svim nulama *Simulacije 13*. U odnosu na histogram na Slici 21, vidljiva je i povećana učestalost dobivene vrijednosti brzine od 3 m/s.



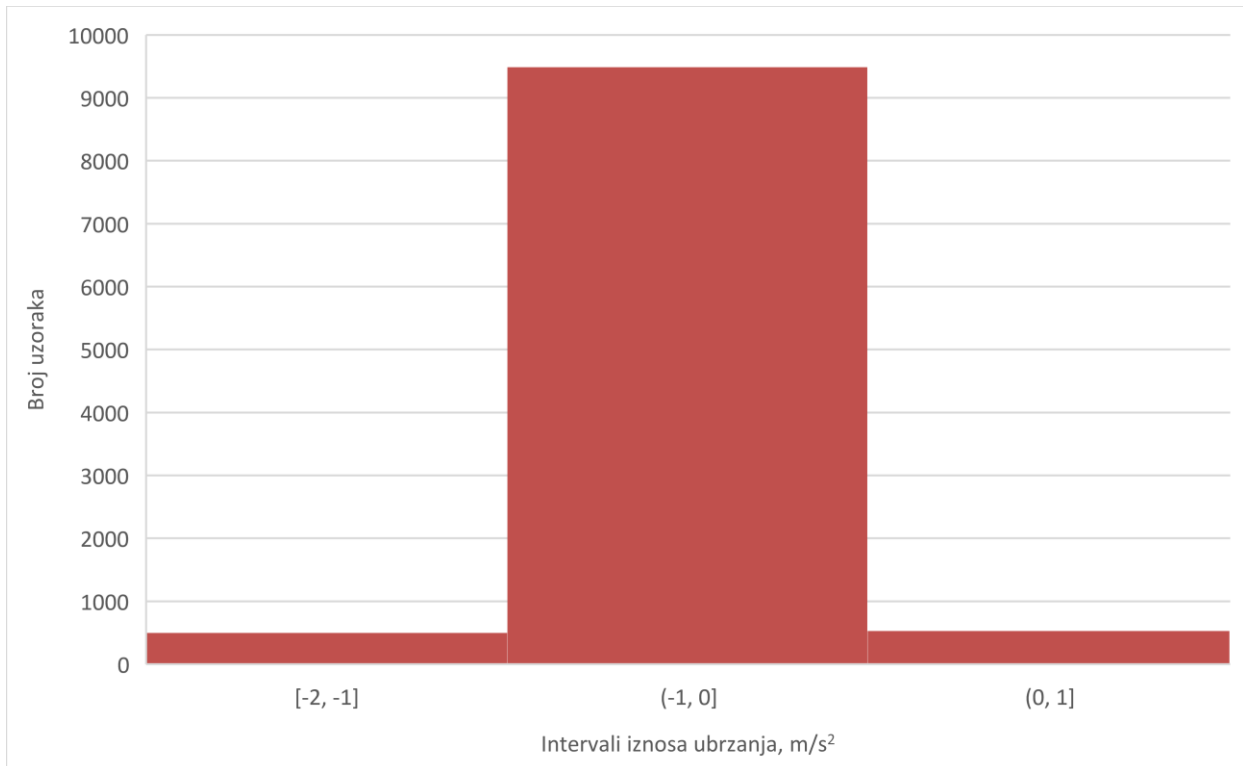
Slika 40. Raspodjela vrijednosti brzine pješaka za sve podatke unutar 50 simulacija

Slika 41 prikazuje raspodjelu brzina vozila koja poprilično nalikuje željenoj i očekivanoj raspodjeli dobivenih rezultata. U odnosu na referentnu Sliku 31, primjetljiva je razlika u broju pojava sadržanih u drugom stupcu (brzine od 3 m/s do 6 m/s) u odnosu na pojave trećeg stupca (brzine od 6 m/s do 9 m/s). Kao i u svim prethodnim primjerima za iznose i raspodjelu varijabli udaljenosti i brzine, i ovdje je povećan udio koji sadrži vrijednosti brzine od 0 m/s do 3 m/s.



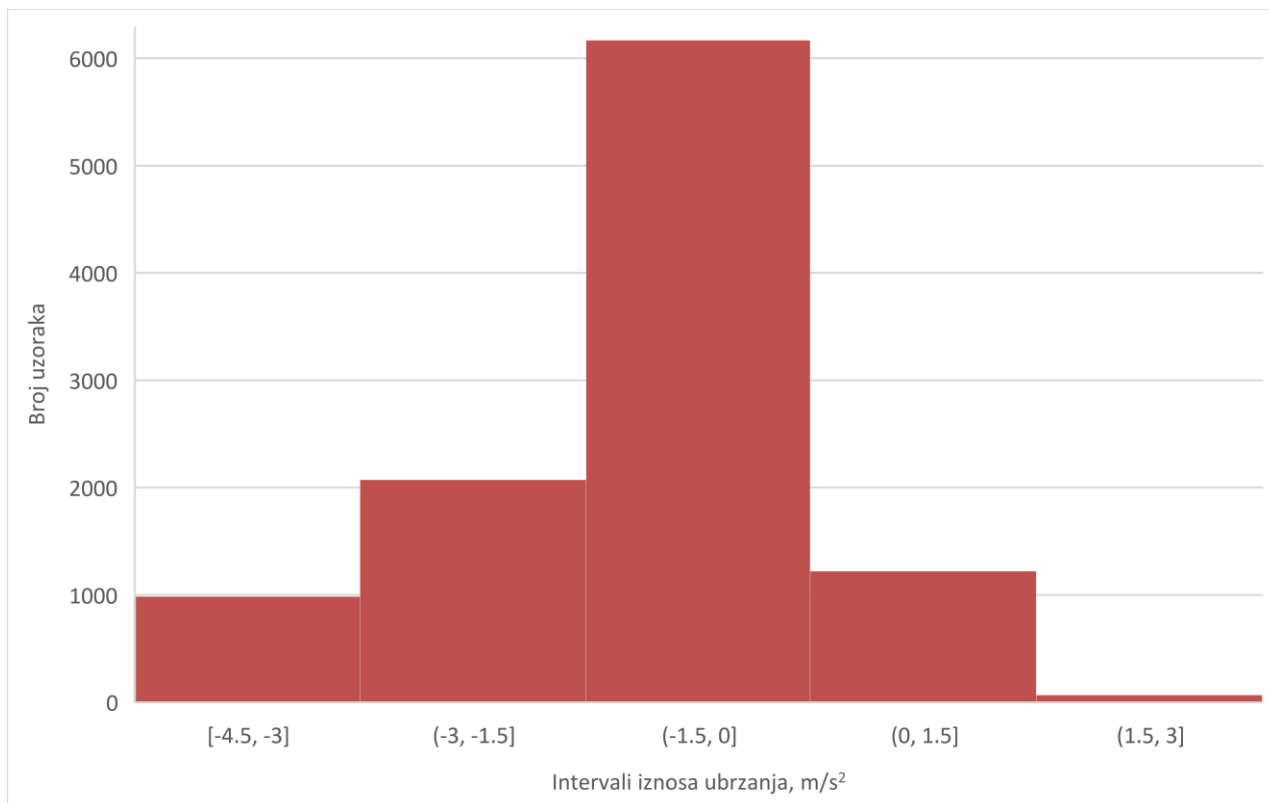
Slika 41. Raspodjela vrijednosti brzine vozila za sve podatke unutar 50 simulacija

Raspodjela ubrzanja pješaka također slijedi onu na referentnoj Slici 25. Manje razlike mogu se primjetiti u povećanom broju uzoraka unutar prvog stupca (ovdje je broj pojava u prvom stupcu skoro pa jednak broj pojava u trećem stupcu), za razliku od referentne Slike 21 gdje je skup trećeg stupca s vrijednostima ubrzanja pokazivao duplo veću učestalost u odnosu na prvi stupac. Isto tako, očekivano je povećana učestalost vrijednosti unutar drugog stupca kao posljedica velikog broja stanja *Simulacije 13*.



Slika 42. Raspodjela vrijednosti ubrzanja pješaka za sve podatke unutar 50 simulacija

Kao zadnja raspodjela prikazuje se raspodjela vrijednosti ubrzanja vozila za sve podatke unutar 50 simulacija. [Slika 43] I ovdje je zaključak više-manje jednak prethodnima – stupac s vrijednostima od $-1,5 \text{ m/s}^2$ do 0 m/s^2 sadrži veći udio uzoraka od referentnog na Slici 34. Od preostalih značajnih informacija, vidljiv je porast broja vrijednosti unutar prvog stupca iznosa ubrzanja.



Slika 43. Raspodjela vrijednosti ubrzanja vozila za sve podatke unutar 50 simulacija

5.3.2. Grafička usporedba simuliranih i testnih vrijednosti

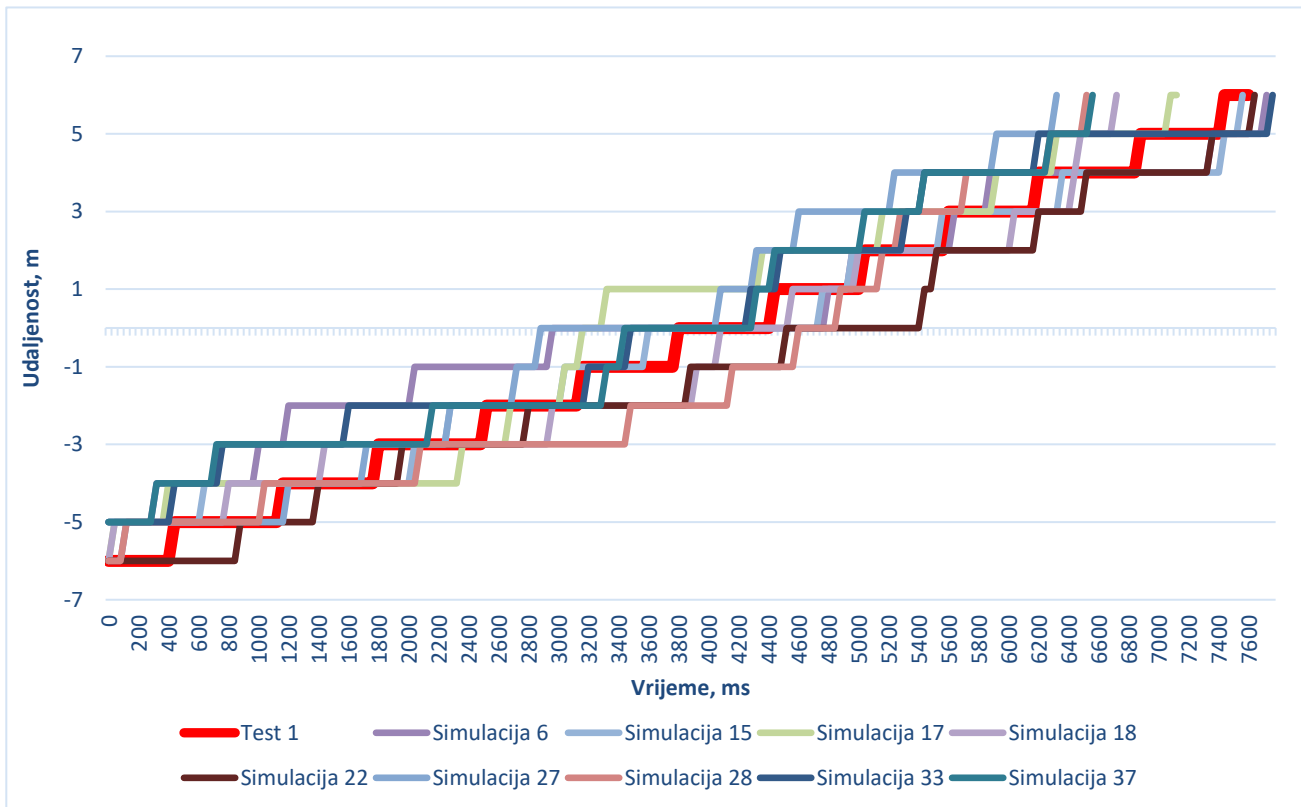
U ovom potpoglavlju na Slikama 44 – 49 prikazana je usporedba dobivenih vrijednosti s jednom od testnih vrijednosti. S obzirom na to da je simuliranih vrijednosti 50, a testnih 5, odabrano je da će se grafički uz jednu testnu vrijednost istovremeno prikazati samo devet simuliranih vrijednosti kako bi se izbjegla velika količina podataka te time i nepreglednost.

Od testnih vrijednosti odabrana je snimka naziva *Test 1*. Na svim dijagramima označena je jarkom crvenom bojom te je, u odnosu na druge linije koje iscertavaju rezultate, značajnije podebljana.

Slika 44 prikazuje usporedbu udaljenosti od pješačkog prijelaza, točnije, putanju pješaka. Iako je izračunato da je duljina pješačkog prijelaza 4,76 m, rijetko koji pješak je taj put prošao točno uzdužno te time završio kretanje nakon točno 5 m (rezolucijom je 4,76 m pretvoreno u 5 m).

Iz dijagrama je vidljivo da nijedna od simulacijom dobivenih putanja ne slijedi testnu putanju u svakom koraku. Vremenska raspodjela po udaljenostima je poprilično slična, uz činjenicu da sve simulacije u prosjeku imaju slično trajanje (navedeno na ordinati u milisekundama).

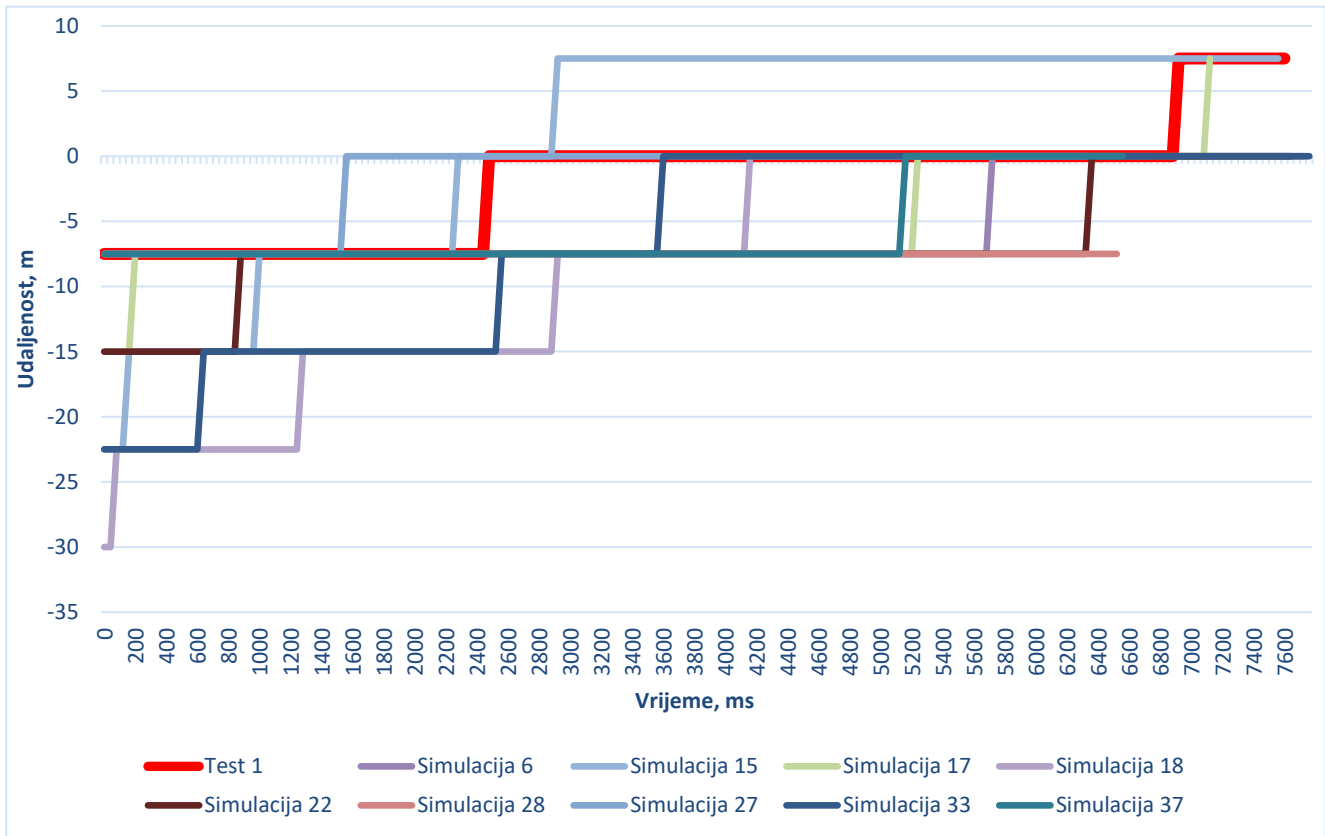
Uzevši u obzir da stanja snimke *Test 1* nisu sudjelovala u kreiranju rječnika, ovaj grafički prikaz pokazuje da prikazane simulacije potpuno zadovoljavaju jer poprilično realistično pokazuju kretanje pješaka.



Slika 44. Prikaz testne putanje pješaka u odnosu na 9 simuliranih putanja

Slika 45 prikazuje simulacijom dobivene vrijednosti za putanju, to jest, udaljenost vozila unutar trajanja simulacije. I ovdje je vidljivo da simulacije donekle slijede obrazac prikazan testnim setom, iako su razlike poprilično očekivane. Kao što je već prije spomenuto, za kreiranje skupa podataka za simulacije primarni fokus bio je na ponašanju pješaka te na njihovoj udaljenosti te sa samim time ni početna ni krajnja udaljenost vozila nikako nije ujednačila. Iz tog razloga, ovdje je dovoljan pokazatelj uspješnosti simulacije to što svaka od prikazanih simulacija kao početak ima vrijednost manju od nule te da se zadržava na nuli, što automatski znači da je vozilo u procesu propuštanja pješaka. Neke od putanja vozila se prije kraja simulacija ni ne približe pješačkom prijelazu, primjerice, putanja dobivena *Simulacijom 28*. U vidu procjene točnosti i smislenosti takvog rezultata (ovdje se za vrijeme trajanja cijele simulacije vozilo nalazi na jednakoj udaljenosti od prijelaza), potrebno je promotriti i dijagrame brzine i akceleracije vozila, koji su prikazani na Slici 47 i Slici 49. Iako je na prvu moguće pomisliti da vozilo u

Simulaciji 28 stoji na 7,5 m od pješačkog prijelaza, postoji mogućnost da se ono izuzetno sporo kretalo unutar raspona vrijednosti koji su definiranjem rezolucije spali pod 7,5 m. Ovakav slučaj je samo jedan od primjera koji pokazuje koliko je važno razumijevanje podataka pri analizi dobivenih simulacija.



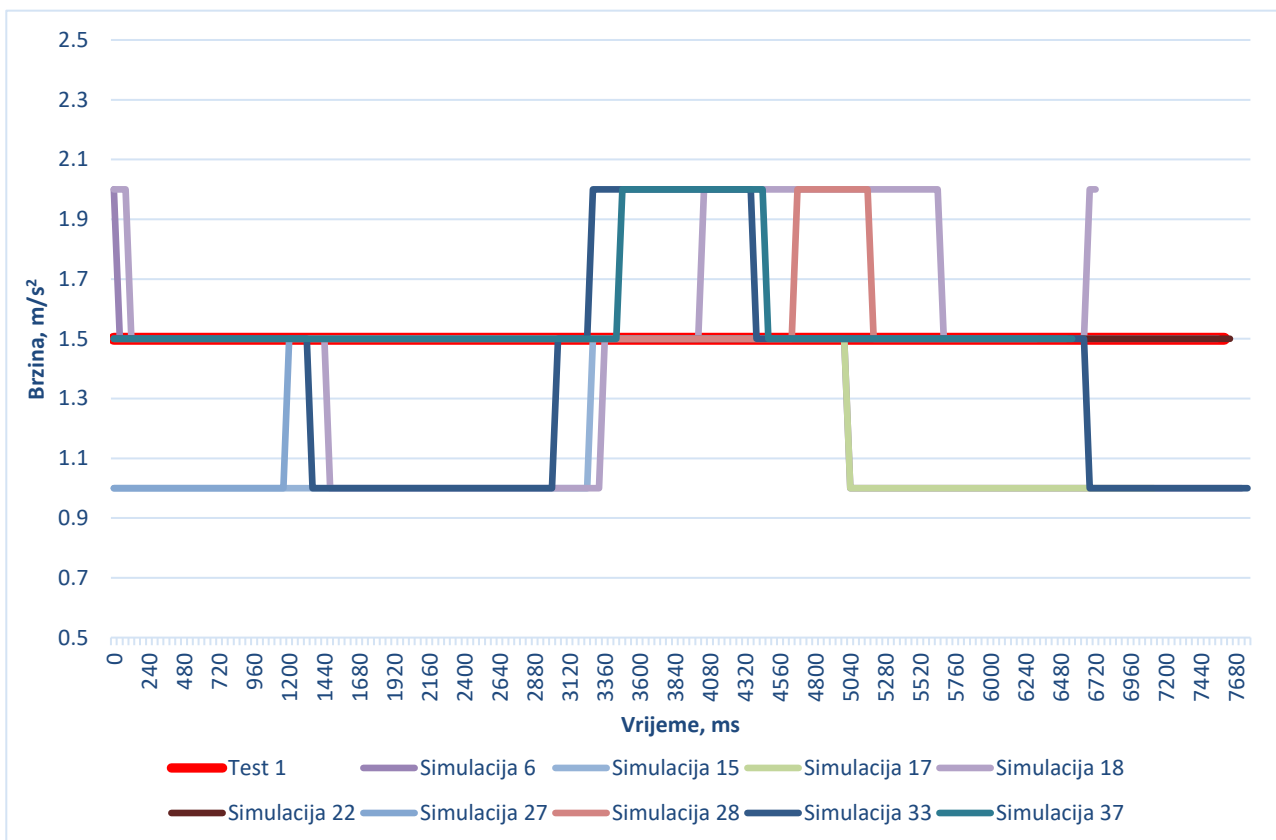
Slika 45. Prikaz testne putanje vozila u odnosu na 9 simuliranih putanja

Na Slici 46 prikazana je raspodjela brzina pješaka po simulaciji. Iako testni set kao brzinu prikazuje konstantnu vrijednost, dobivene simulacije variraju u raspodjeli brzina.

I samu snimljenu vrijednost *Testa 1* moguće je jednostavno objasniti – iz dijagrama na Slici 44 i Slici 45 vidljivo je da je do pješačkog prijelaza prije pješaka stiglo vozilo. Vozilo je ostalo na istoj udaljenosti, što može značiti da se ili nalazilo unutar raspona od 0 m ili da je izuzetno usporilo. Oboje pješaku daje informaciju o tome da mu vozilo daje prednost prelaska te nije bilo potrebe da pješak uspori kako bi se uvjerio u mogućnost sigurnog prolaska. Naravno, dane rezolucije za brzinu pješaka značajno su reducirale primjetljivost razlika u njegovoj brzini te je svakako moguće da je pješak u nekom trenutku prije stupanja na pješački prijelaz usporio ili stupanjem na pješački prijelaz ubrzao.

Od preostalih simulacija potrebno je istaknuti i *Simulaciju 6*. Ovdje je vidljivo kako je pješak u nekim trenucima povećavao i smanjivao svoju brzinu. Isto može biti uzrokovano malom udaljenosti vozila od pješačkog prijelaza od početka simulacije, a uzrok može biti i u prirodi ponašanja pješaka. Vidljivo je da povećanje brzine kretanja pješaka korelira s vremenskim okvirom u kojem dolazi do stupanja na pješački prijelaz te je moguće zaključiti da je u ovom slučaju pješak stupanjem na prijelaz ubrzao svoje kretanje. Kao i kod vozila, isto je moguće potvrditi uvidom u dijagram raspodjele ubrzanja [Slika 48].

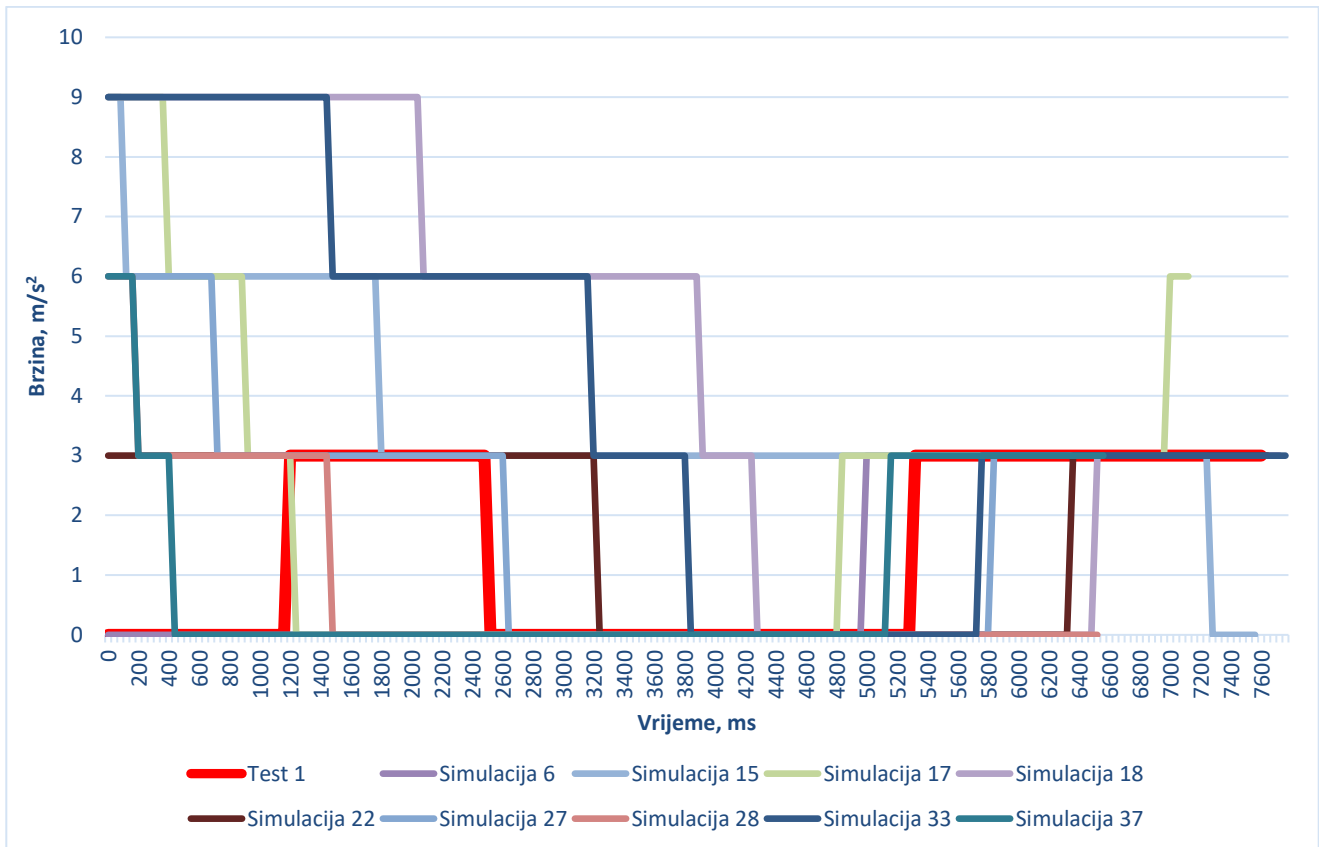
Slične analize mogu se izvesti za bilo koju od prikazanih simulacija. Ovdje će se još kratko razmotriti brzina kretanja pješaka u *Simulaciji 33* u odnosu na brzinu kretanja pješaka u simulaciji istog rednog broja. Vidljivo je da brzina kretanja vozila [Slika 47] opada s vremenom, a tek u vremenskom trenutku kada brzina pada na 0 m/s, dolazi i do povećanja brzine kretanja pješaka. Lako je zaključiti da je u ovom slučaju pješak čekao da se vozilo potpuno ili skoro pa potpuno zaustavi kako bi sigurno prešao preko pješačkog prijelaza.



Slika 46. Prikaz testne brzine pješaka u odnosu na 9 simuliranih brzina

Za brzine vozila, prikazane na Slici 47, slično kao i za udaljenosti vozila, zbog raznovrsnosti podataka o vozilima ne postoji očekivana raspodjela vrijednosti kao što je to stvar za informacije o pješacima. Ono što je svakako moguće zamijetiti je česta raspodjela brzina u vidu smanjivanja brzine prilaskom pješačkom prijelazu (za vozila koja se u početku simulacije nisu nalazila u blizini prijelaza). Vozila koja se početkom simulacije nalaze u blizini pješačkog prijelaza ili do kraja simulacije ostaju unutar jednakog raspona brzine ili, nakon prelaska pješaka preko pješačkog prijelaza, ubrzavaju.

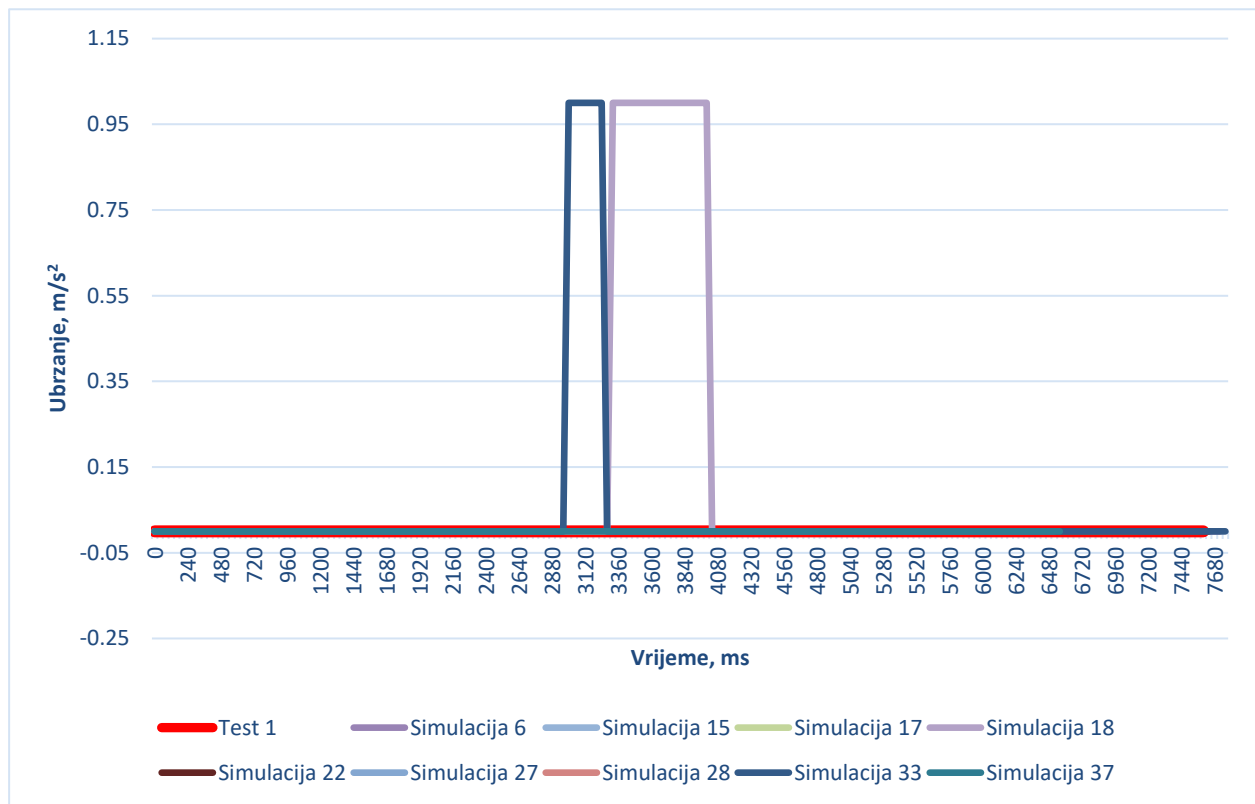
Uz to, ovdje je svakako potrebno istaknuti brzinu vozila u testnom skupu podataka. Vozilo je u jednom trenutku pokazalo povećanu brzinu, no približavanjem pješačkom prijelazu brzina je pala na nulu te ponovno narasla u vremenskim okvirima nakon prelaska pješaka preko prijelaza. Objašnjenje za takvu raspodjelu brzina moguće je pronaći u više faktora – jedan od njih svakako može biti činjenica da vozač nije na vrijeme primjetio pješaka te da mu je daljnji plan bio naknadno ubrzavati, što se promijenilo primjećivanjem pješaka, dok je druga mogućnost da se vrijednost brzine vozila cijelo vrijeme kretala oko 1,5 m/s što ju je ponekad svrstalo u vrijednost od 3 m/s, a ponekad u 0 m/s. Treća mogućnost je da je na brzinu vozila utjecalo postojanje parkirališta koje se nalazi desno od vozila u smjeru njegovog kretanja, uz cestovni trak. Moguće da je time brzina vozila bila pod utjecajem vozila koje je, primjerice, dalo naznaku izlaženja s parkirnog mjesta. Samo promatranje dijagrama nije dovoljno za određivanje uzroka promjenama bilo koje varijable te je sve navode potrebno promatrati u kontekstu okoline u kojoj se vozilo nalazi.



Slika 47. Prikaz testne brzine vozila u odnosu na 9 simuliranih brzina

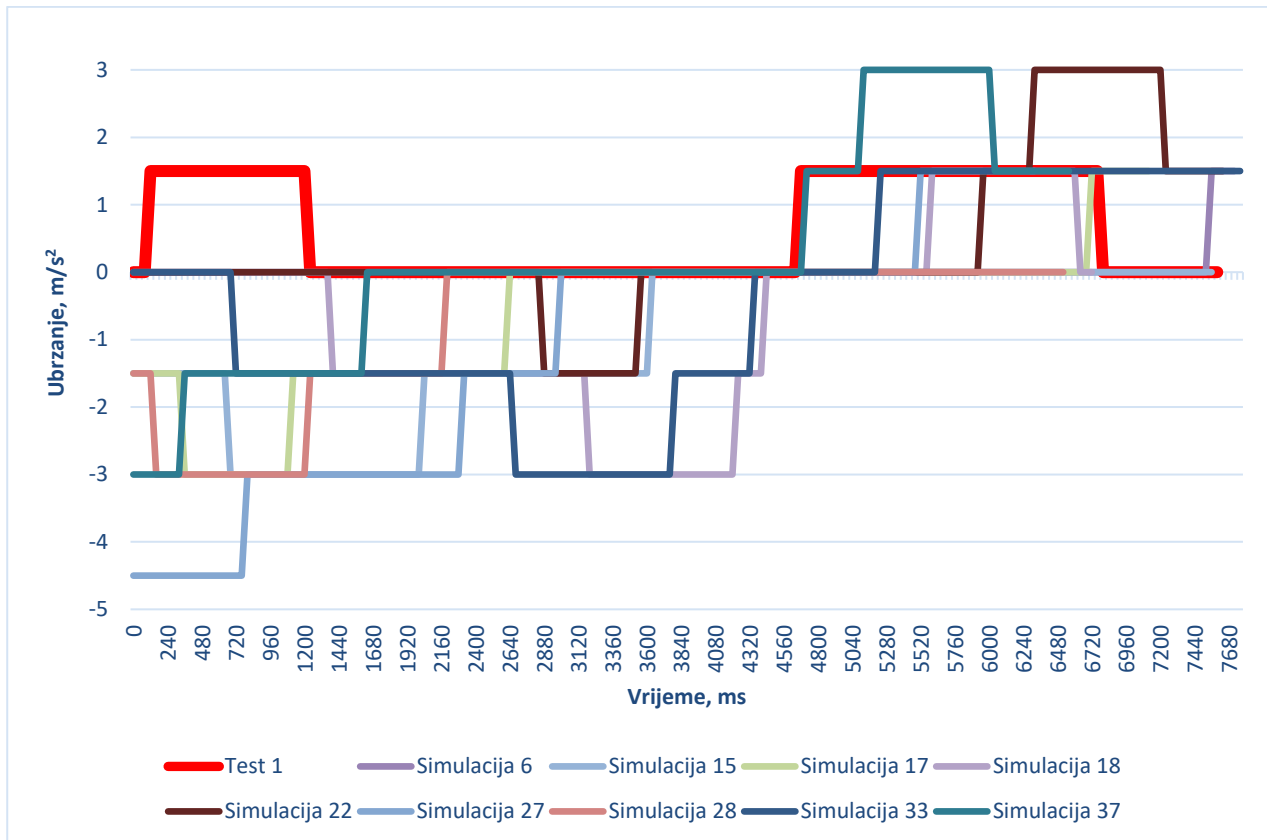
Slika 48 prikazuje raspodjelu ubrzanja pješaka simuliranih vrijednosti i testne vrijednosti. S obzirom na nepostojanje nagle promjene pri odluci o prijelasku pješačkog prijelaza, većina ubrzanja je u ovim situacijama jednaka nuli. Iznimke su *Simulacija 33* i *Simulacija 18*. Kod *Simulacije 33* ubrzanje se događa prije stupanja na pješački prijelaz, vjerojatno kako bi se vozilu dodatno istaknula pješakova namjera prelaska ceste. Kod *Simulacije 18* ubrzanje se događa također prije stupanja na pješački prijelaz. Ovdje je razlog potencijalno isti, pogotovo kada se uzme u obzir da se ubrzanje pojavilo kada se vozilo nije nalazilo na udaljenosti iznosa 0 m od

pješačkog prijelaza – pješaku nije bilo jasno hoće li vozilo stati pa je potencijalno ubrzao kako bi čim prije prešao pješački prijelaz.



Slika 48. Prikaz testnog ubrzanja pješaka u odnosu na 9 simuliranih ubrzanja

Dijagrami ubrzanja vozila pretežito prikazuju usporenja vezana uz dolazak do pješačkog prijelaza. [Slika 49] Iako je u prikazanoj *Simulaciji 15* prvi objekt koji prelazi pješački prijelaz vozilo, ono ne ubrzava ni prilaskom pješačkom prijelazu ni nakon što se od njega udalji. Naime, moguće je vidjeti da u trenucima prije prelaska preko pješačkog prijelaza usporava, što je vjerojatno mjera preodstrožnosti u slučaju da na pješački prijelaz neočekivano naleti pješak ili biciklist s jasno izraženom namjerom prelaska.



Slika 49. Prikaz testnog ubrzanja vozila u odnosu na 9 simuliranih ubrzanja

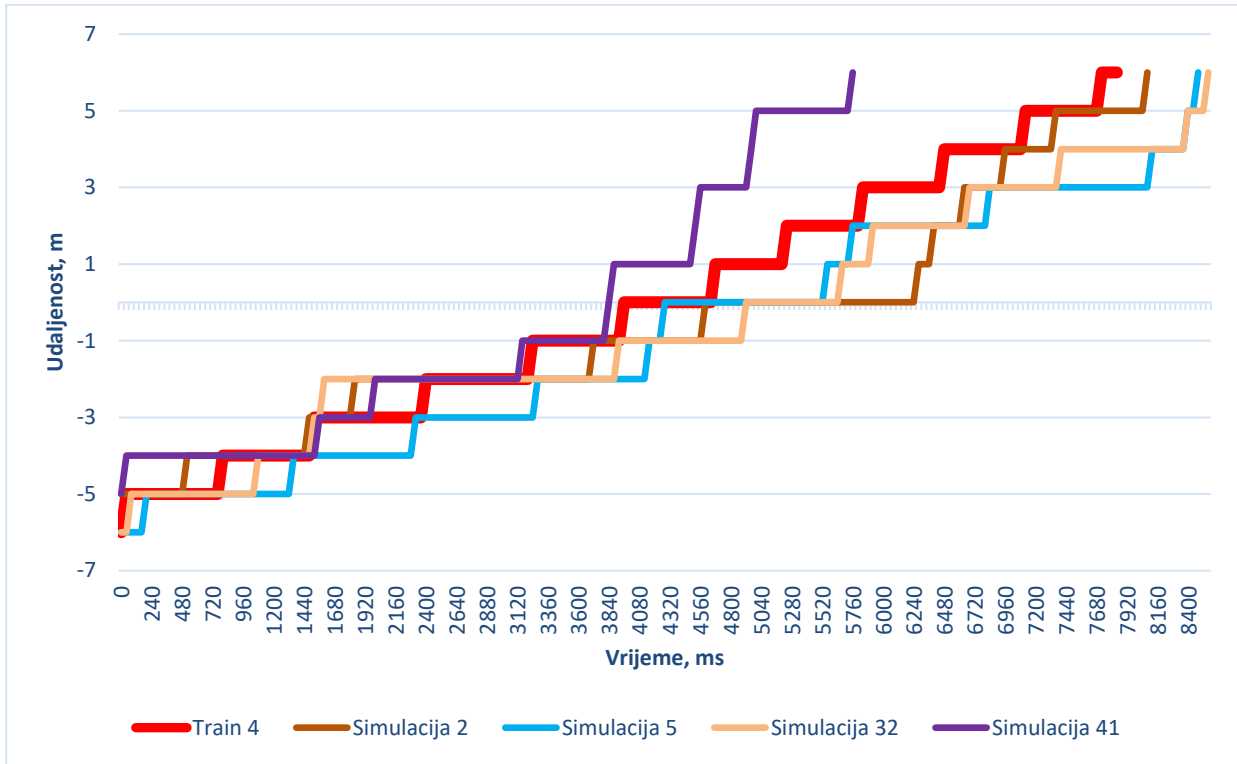
Uz prikaz testne snimke usporedno sa simulacijama, napravljen je i usporedni prikaz u odnosu na jednu od snimki čija se stanja nalaze u rječniku. Radi se o snimci pod rednim brojem 24 navedenoj u Tablici 10. U ovom slučaju usporedno se razmatraju samo 4 simulacije, *Simulacije* 2, 5, 32 i 41, kako bi prikaz rezultata bio vizualno jasniji (pregledniji). Za razliku od prethodne analize rezultata, ova analiza je pretežito usmjerena na usporedbu sveukupnih informacija o stanjima simulacije u odnosu na snimku.

Snimka 24, dalje navođena kao *Train 4*, prikazuje situaciju u kojoj vozilo prelazi preko pješačkog prijelaza prije pješaka, kao što je situacija i za *Simulaciju 2* te *Simulaciju 32*. Isto je vidljivo usporedbom trenutaka kada pješak i vozilo prijeđu iz udaljenosti od 0 m na udaljenost od 1 m (za pješaka) ili 7,5 m (za vozilo). U slučaju snimke *Train 4* isto se događa s vremenskim razmakom od otprilike 800 ms, dok je u *Simulaciji 32* potrebno približno 1800 ms. Po vremenskom razmaku prednjači *Simulacija 2* gdje pješak stupa na prijelaz otprilike 2500 ms nakon vozila. Tijekom trajanja preostale dvije simulacije vozilo nije stiglo prijeći preko pješačkog prijelaza.

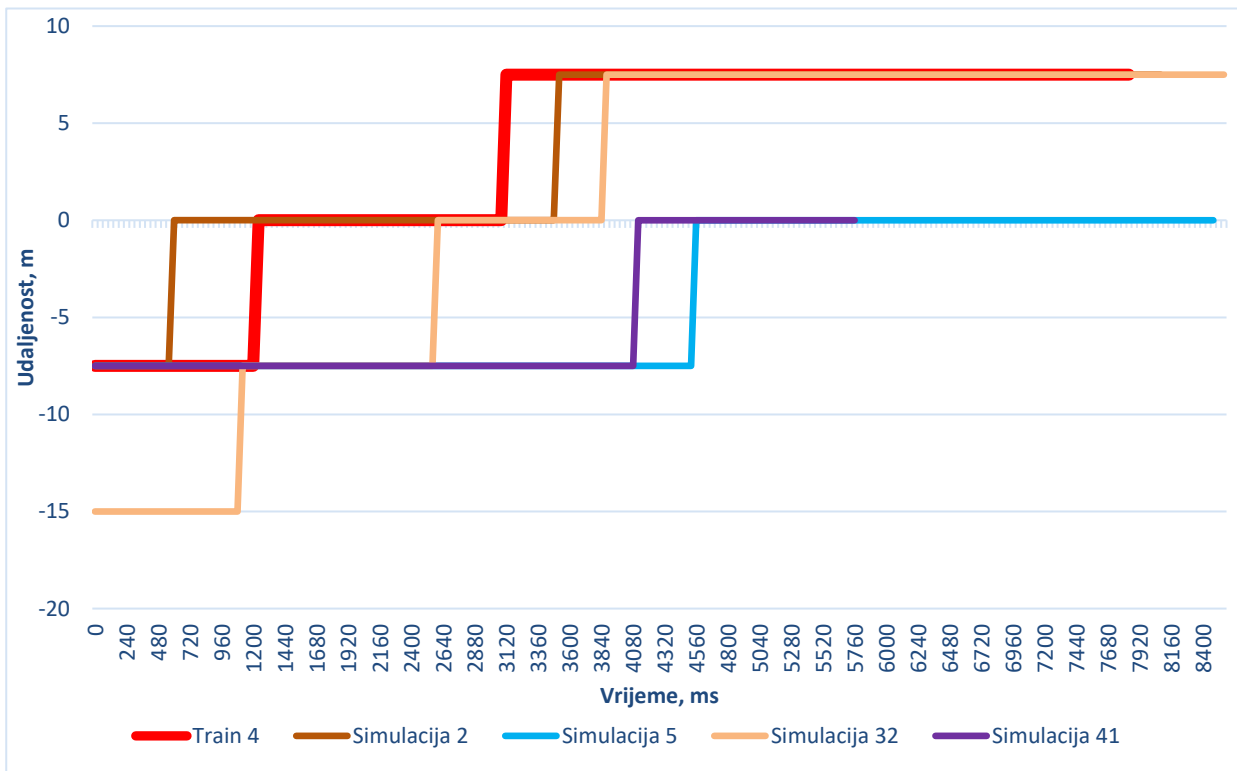
S obzirom na brzinu pješaka skupa *Train 4*, vidljivo je kako ona opada u trenutku kada je pješak na 2 m od stupanja na pješački prijelaz, a vozilo pokazuje tendenciju nastavljanja putanje. Brzina vozila je unutar istog perioda konstantna. Brzina pješaka povećava se u trenutku u kojem se vozilo već odmaknulo 7,5 m od pješačkog prijelaza. U tom vremenskom periodu, zanimljivo je promotriti kako ubrzanja ni usporenja kod pješaka nema, dok je kod vozila vidljivo povećanje ubrzanja tek nakon udaljavanja od prijelaza. Pri približavanju pješačkom prijelazu, vozilo smanjuje brzinu. Dodatno o analizi *Snimke 24* može se pronaći u poglavlju 5.4.2.

Isto se može primijeniti na *Simulacije 2 i 32* – dijagrami iznosa njihovih udaljenosti, brzina i ubrzanja slijede isti obrazac, ovisan o trenutcima u kojima dolazi do prijelaza preko pješačkog prijelaza. Jedinu značajnu razliku prikazuje dijagram ubrzanja vozila, gdje vozilo sa *Simulacije 32* usporava približivši se pješačkom prijelazu na 7,5 m te ubrzava nakon što dođe na jednaku udaljenost od točke stupanja na pješački prijelaz.

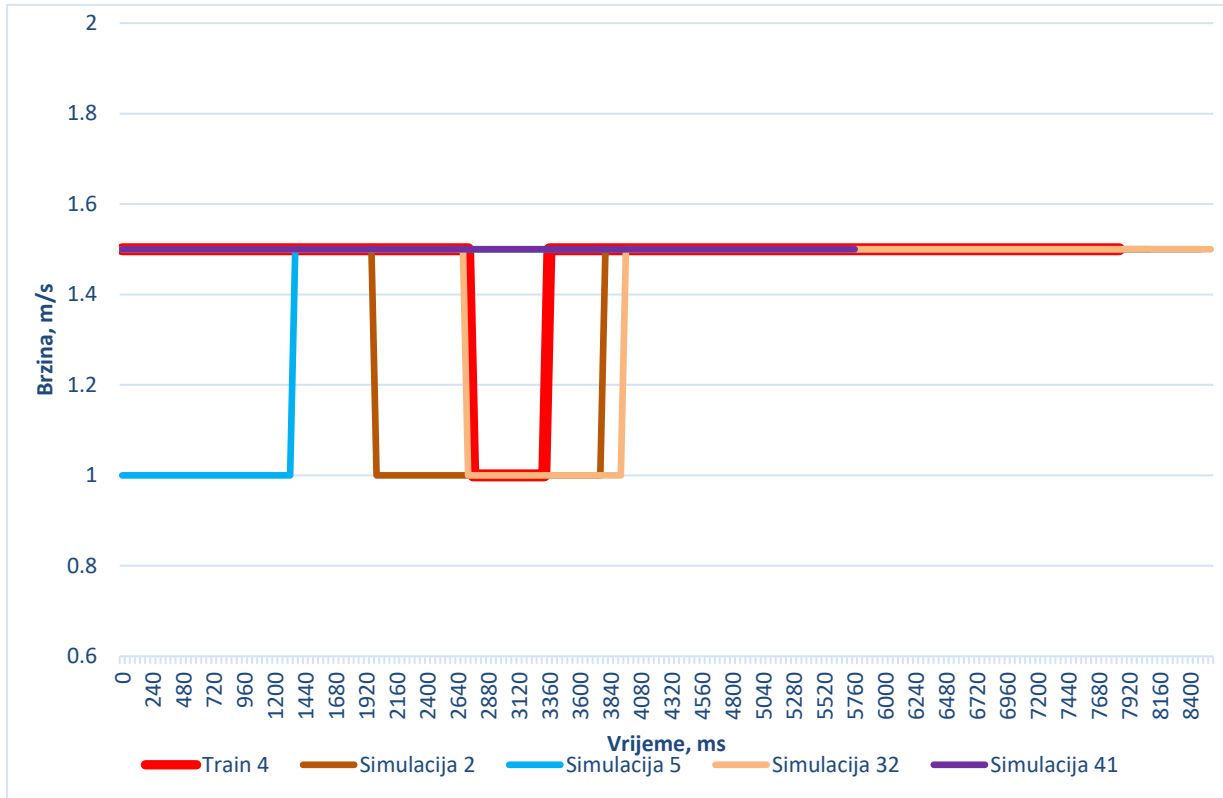
Za preostale dvije simulacije vidljivo je da im brzine padaju na nulu sukladno približavanju pješačkom prijelazu. U *Simulaciji 5* se može primjetiti da je prisutna značajna deceleracija te da je do smanjenja brzine na 0 m/s došlo tek u trenutku kada je i udaljenost vozila i pješaka od pješačkog prijelaza bila 0 m, što može značiti da je vozač u vozilu naglo donio odluku o potrebi zaustavljanja ispred pješačkog prijelaza. Prethodno objašnjenje odnosi se na prikaze na Slika 50- Slika 55.



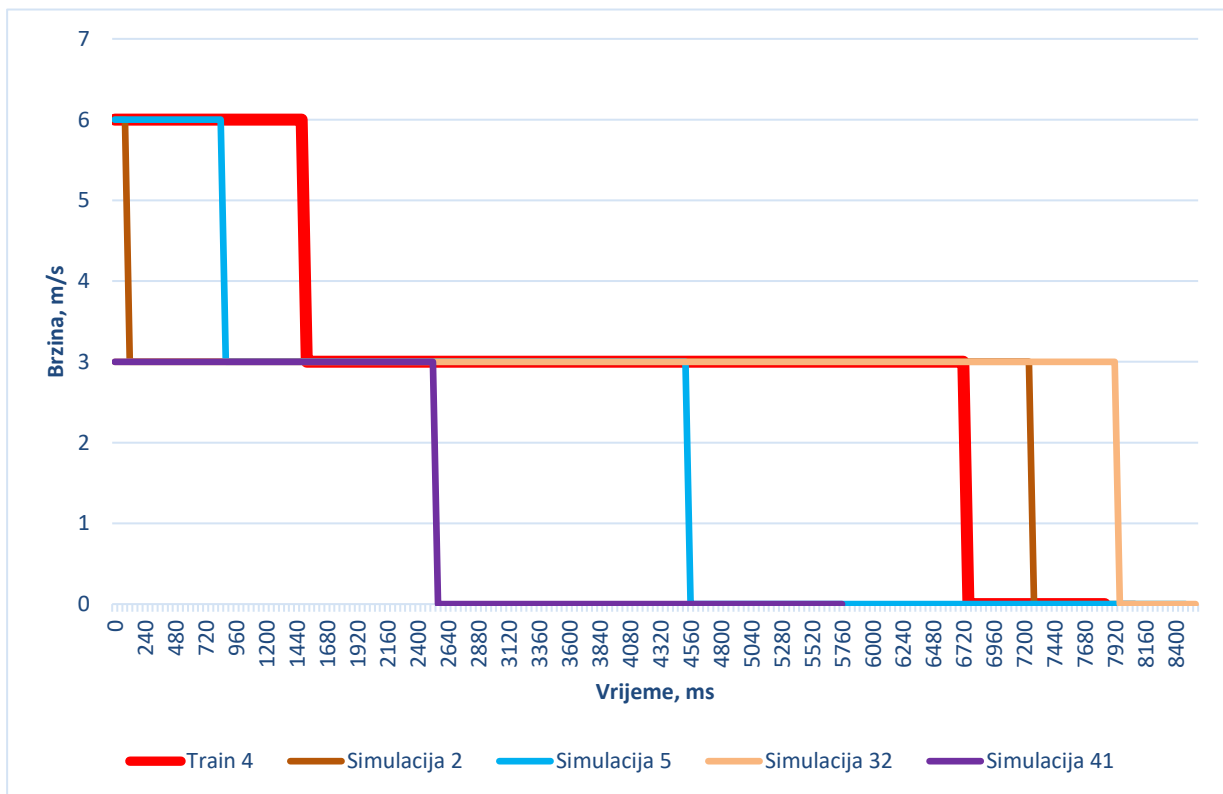
Slika 50. Prikaz udaljenosti pješaka sa snimke *Train 4* u odnosu na 4 simulirane udaljenosti



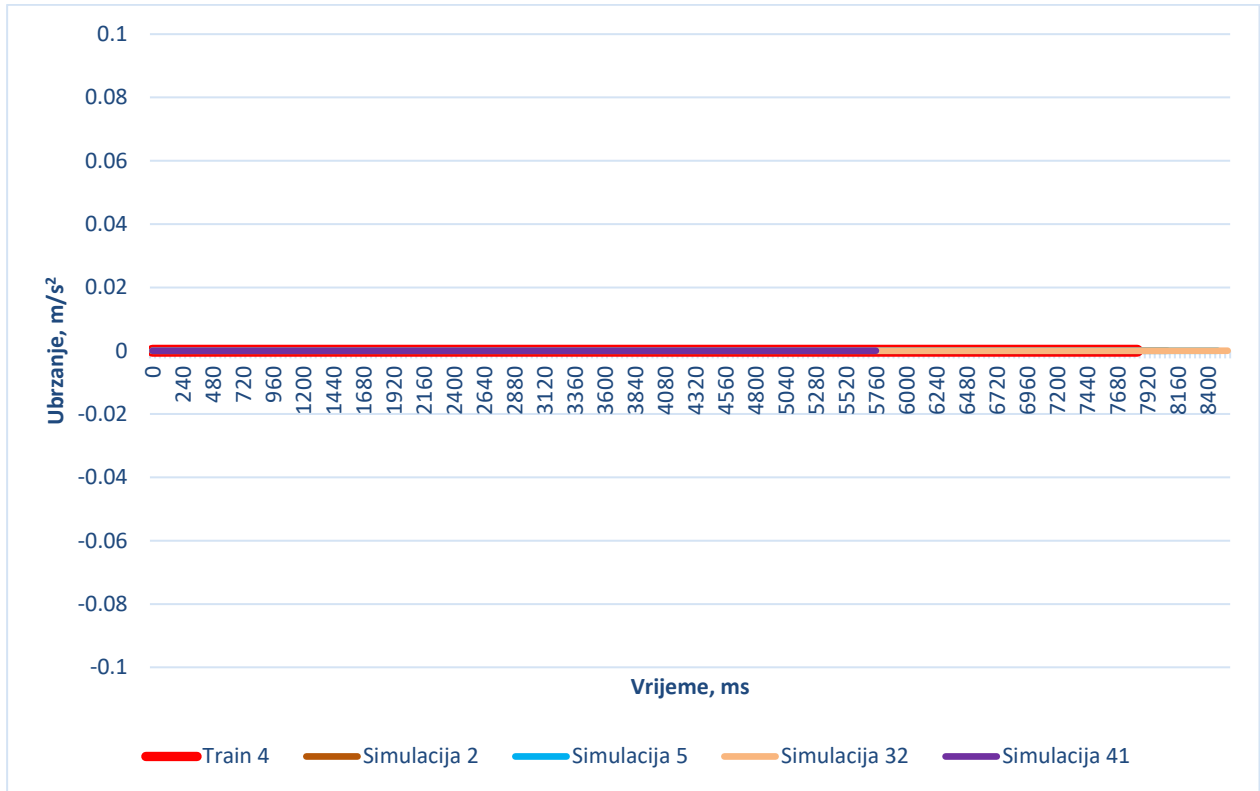
Slika 51. Prikaz udaljenosti vozila sa snimke *Train 4* u odnosu na 4 simulirane udaljenosti



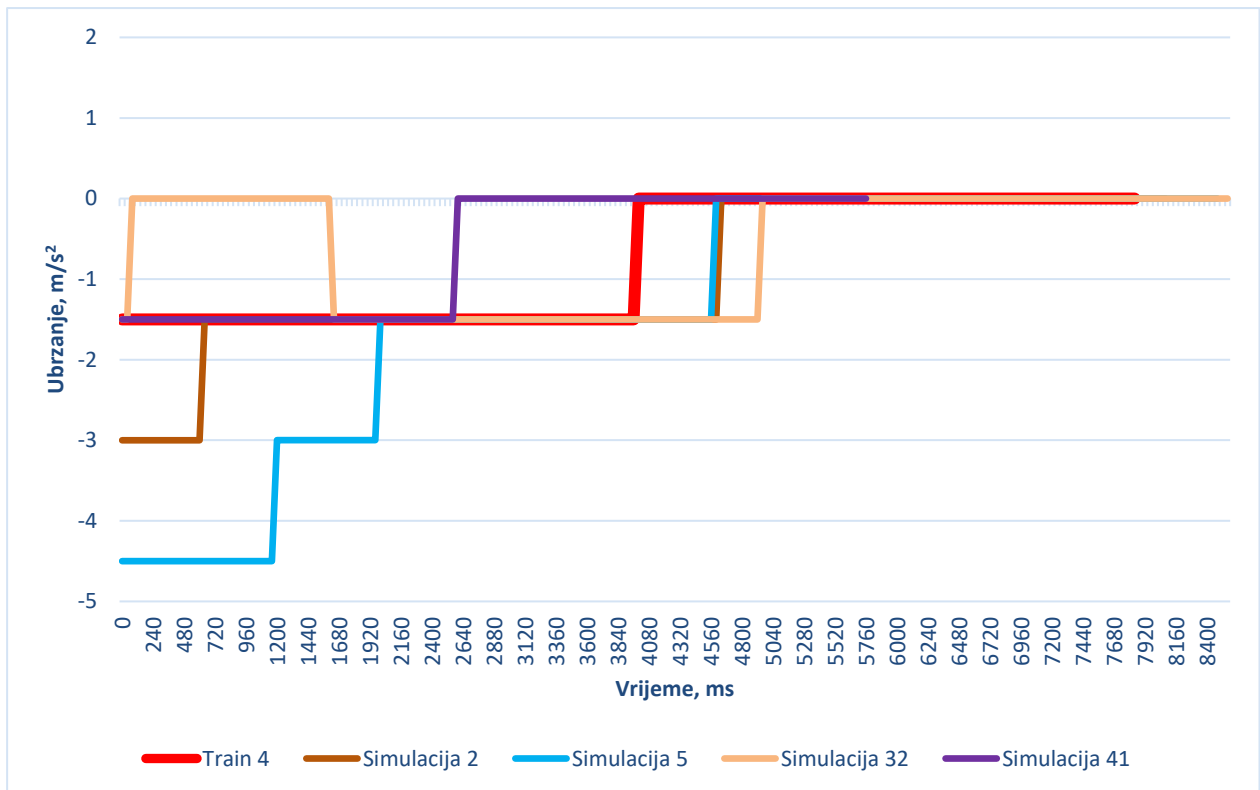
Slika 52. Prikaz brzine pješaka sa snimke *Train 4* u odnosu na 4 simulirane brzine



Slika 53. Prikaz brzine vozila sa snimke *Train 4* u odnosu na 4 simulirane brzine



Slika 54. Prikaz ubrzanja pješaka sa snimke *Train 4* u odnosu na 4 simulirana ubrzanja



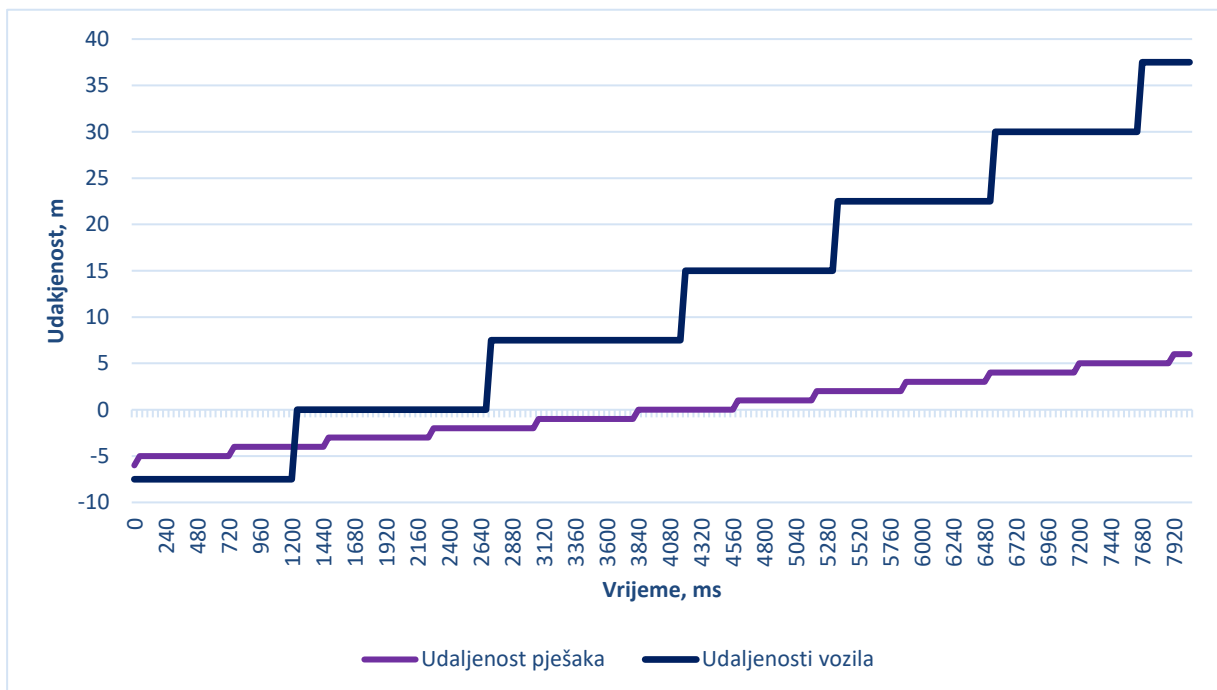
Slika 55. Prikaz ubrzanja vozila sa snimke *Train 4* u odnosu na 4 simulirana ubrzanja

Uz prethodne analize sveukupnog ponašanja prikazano je po tri dijagrama za šest stanja za tri snimke ili simulacije gdje su istovremeno prikazane vrijednosti udaljenosti pješaka i vozila te, analogno, brzina i ubrzanje pješaka i vozila. Umjesto milisekundi, trajanje simulacije prikazano je u broju vremenskih okvira. Razmak između dva vremenska okvira je 40 ms.

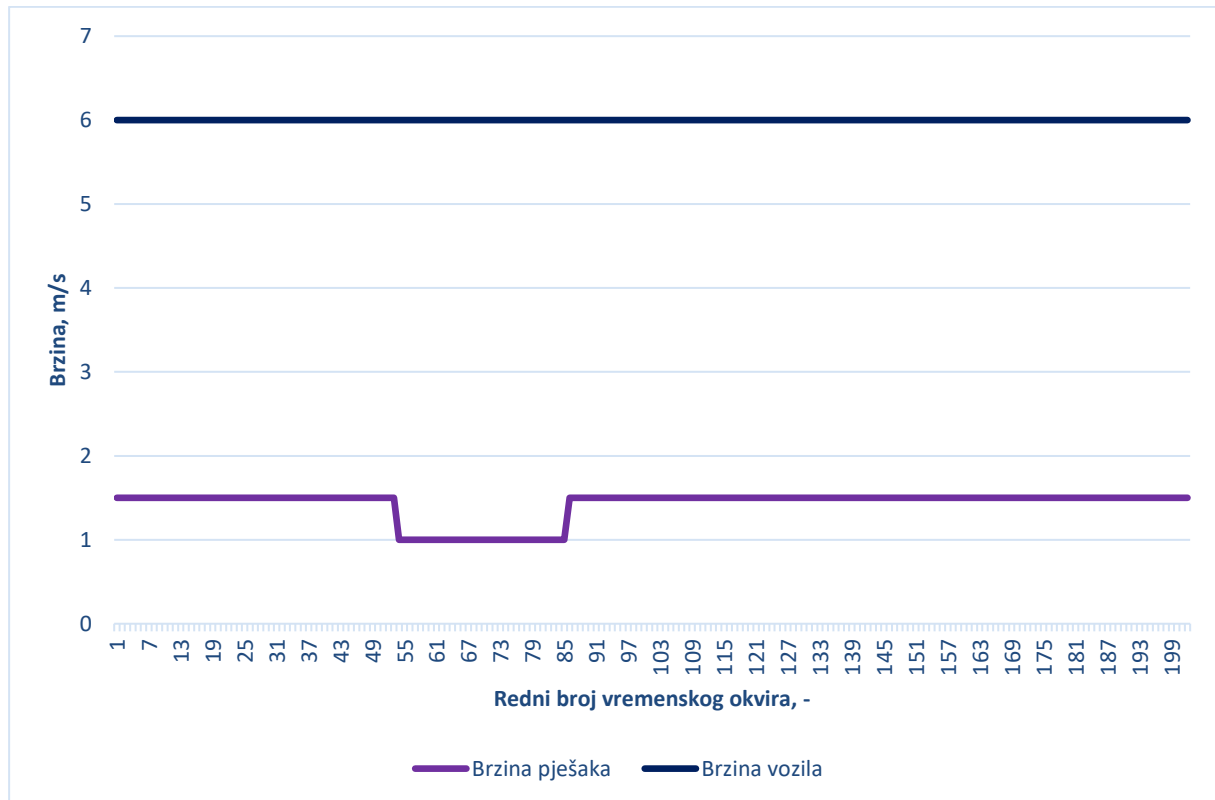
Snimka Test 4

Prvi skup podataka odnosi se na snimku *Test 4*. Ova snimka je odabrana jer je jedna od rijetkih unutar cijelokupnog skupa podataka u kojima dolazi do situacije da vozilo prelazi pješački prijelaz prije pješaka.

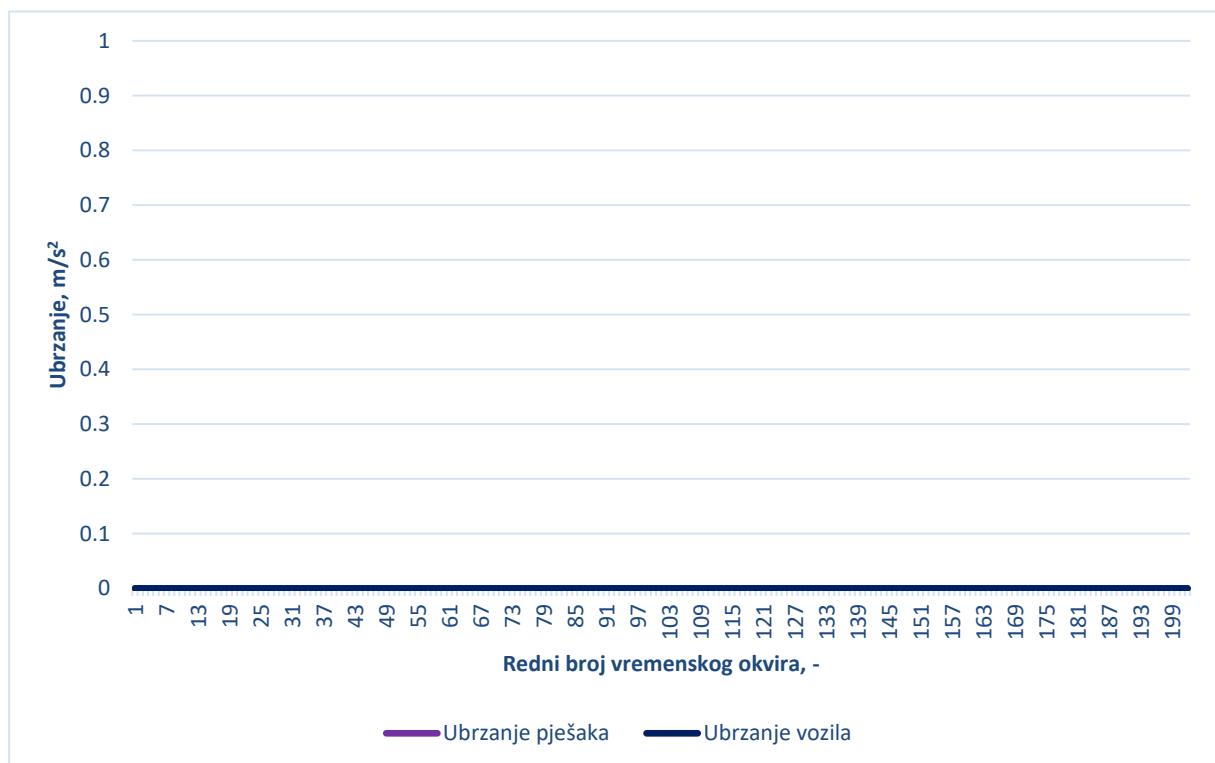
S obzirom na neveliku razliku između početnih udaljenosti pješaka i vozila (razlika je samo $\Delta d = 1,5$ m) [Slika 56], značajno je razmotriti mogući uzrok nepropuštanja pješaka. Iako se udaljenost pješaka od pješačkog prijelaza smanjivala te ga je vozilo zasigurno moralo primjetiti, barem dok je pješak bio na 3 m ili 4 m od prijelaza, vozilo ni u kojem trenutku nije usporilo svoje kretanje niti smanjilo svoju brzinu. Iako kontekst iza ovakvog ponašanja ne može biti jasan samo iz promatranog dijagrama, može se pretpostaviti da je ovakvo ponašanje suprotno od očekivanog po pitanju kulturo i zakonski definiranog propuštanja pješaka. Samim time, ovakva snimka ne bi smjela biti referentna za kreiranje stanja te za prenošenje ovog iskustva na vozilo kao poželjnog. [Slika 57, Slika 58]



Slika 56. Usporedbeni prikaz udaljenost pješaka i vozila snimke *Test 4*



Slika 57. Usporedbeni prikaz brzine pješaka i vozila snimke *Test 4*



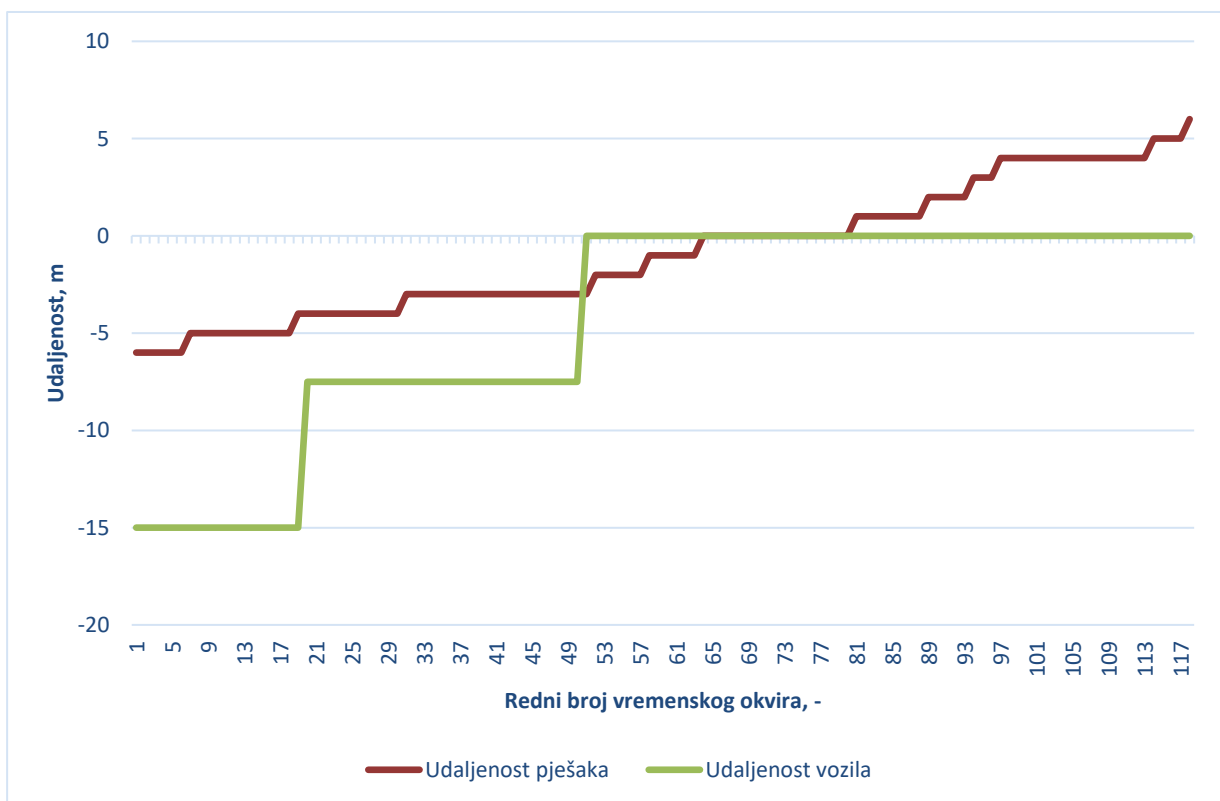
Slika 58. Usporedbeni prikaz ubrzanja pješaka i vozila snimke *Test 4*

Simulacija 16

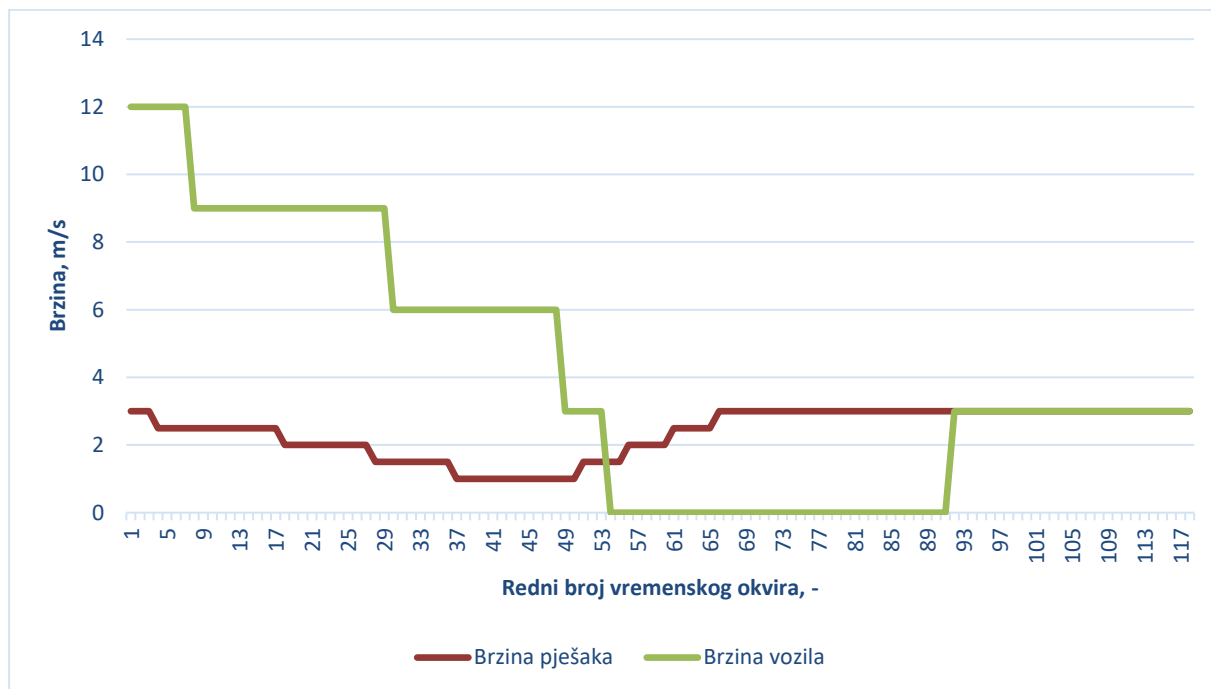
Za sljedeći prikaz odabrana je *Simulacija 16*. Odabir je temeljen na varijaciji u brzinama i akceleracijama i pješaka i vozila. Te vrijednosti u većini ostalih simulacija iznose nula.

U ovom slučaju vidljivo je kako je vozilo smanjilo svoju brzinu približavajući se pješačkom prijelazu. Isto tako, primjećujući smanjenje brzine, pješak je svoju brzinu povećao. Brzina pješaka se tijekom trajanja cijele simulacije smanjivala i povećavala. Pretpostavka je da je razlog tome što pješak na većoj udaljenosti od vozila nije bio siguran što će vozilo poduzeti. [Slika 59] Pokazano je da je teže procijeniti brzinu daljeg vozila, sa stajališta pješaka. [2]

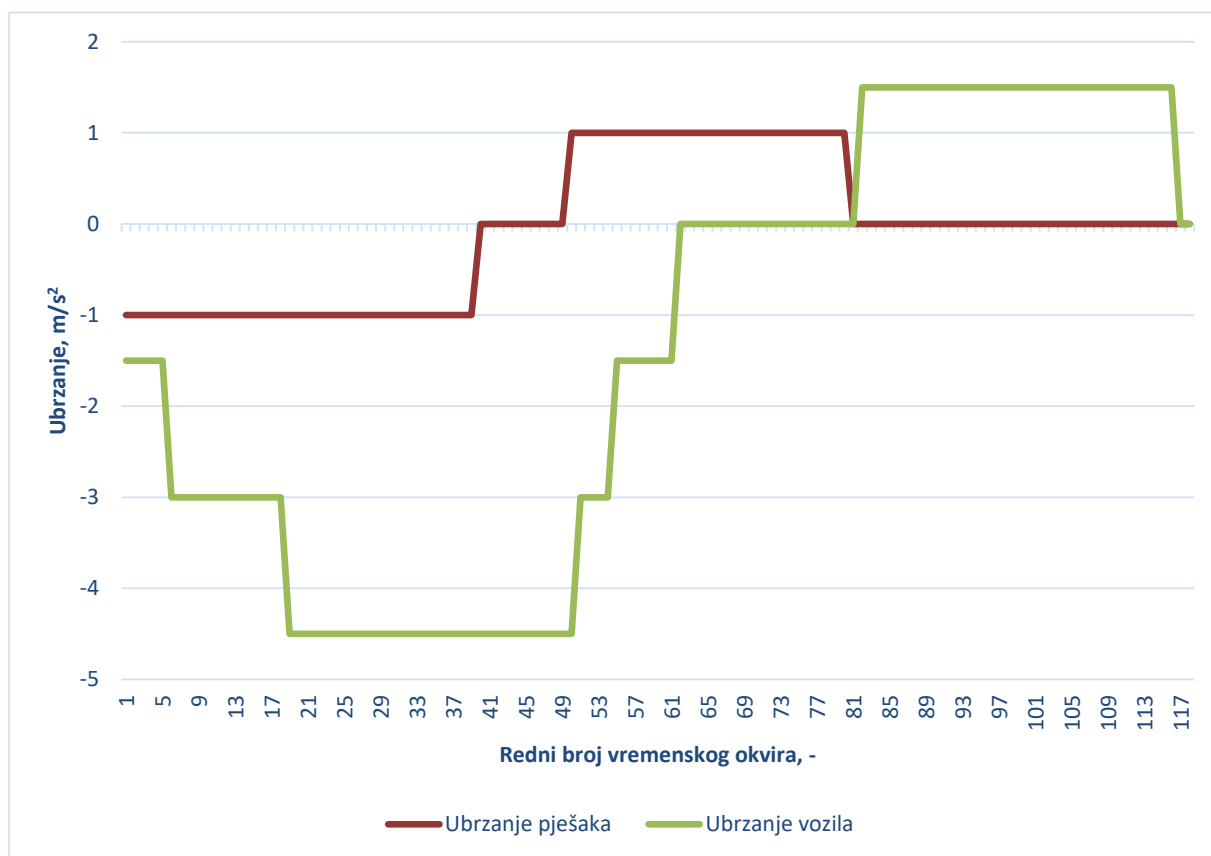
Akceleracija i deceleracija pješaka i vozila koreliraju s njihovim prethodno opisanim ponašanjem u danim trenucima. [Slika 60, Slika 61]



Slika 59. Usporedbeni prikaz udaljenosti pješaka i vozila *Simulacije 16*



Slika 60. Usporedbeni prikaz brzine pješaka i vozila *Simulacije 16*

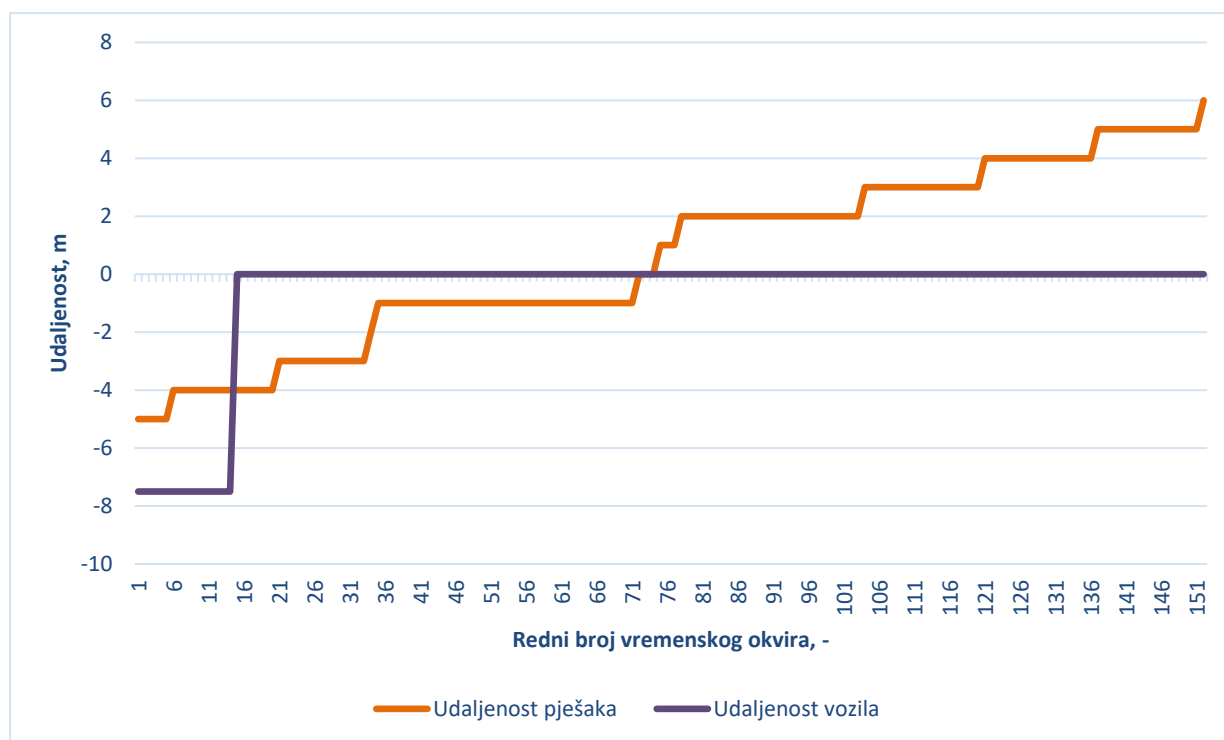


Slika 61. Usporedbeni prikaz ubrzanja pješaka i vozila *Simulacije 16*

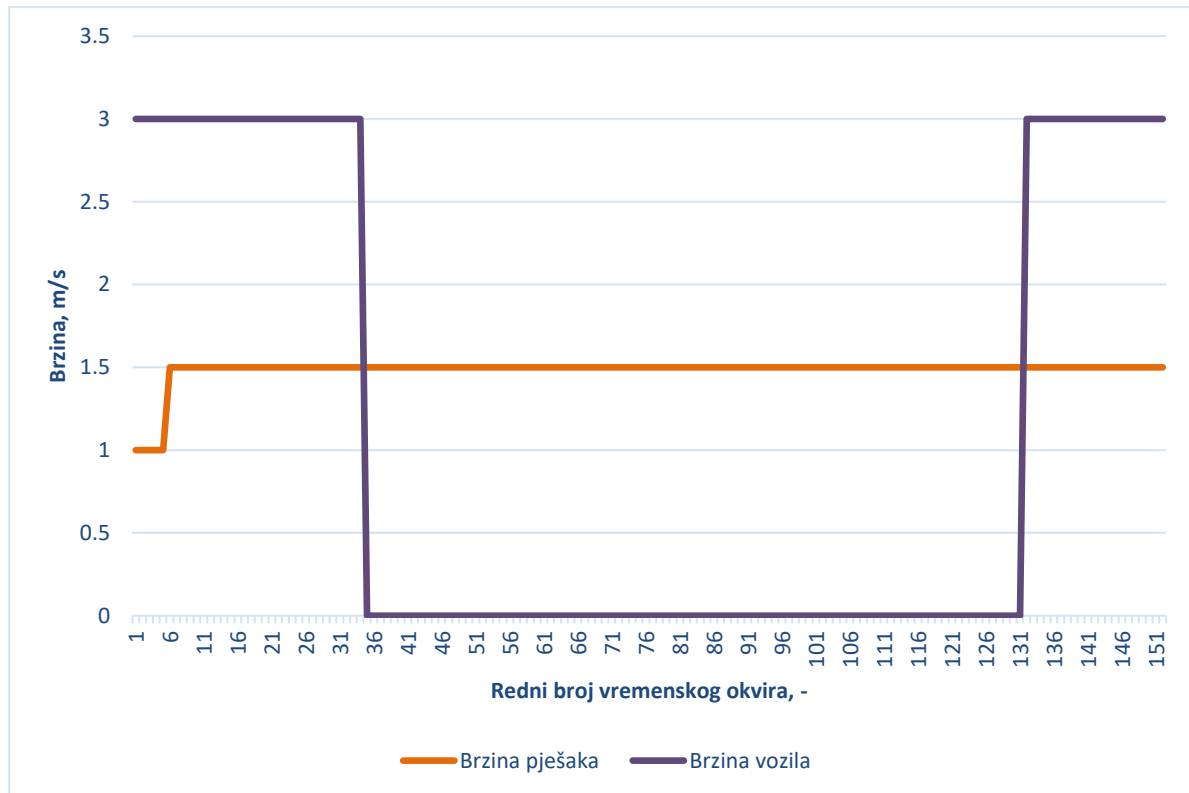
Simulacija 29

Sljedeći dijagrami prikazuju grafičke prikaze vrijednosti udaljenosti, brzina i ubrzanja za *Simulaciju 29*. Značajno je promotriti kako, iako je vozilo stalo na udaljenosti od 0 m od pješačkog prijelaza, postoji značajan i trajan porast njegove brzine i ubrzanja u trenucima kada je ono i dalje "stajalo". [Slika 62] Brzina vozila je porasla na 3 m/s i zadržala se na toj vrijednosti unutar 20 vremenski okvira sveukupnog trajanja 80 ms. [Slika 63] Iz tih vrijednosti lako je izračunati prijeđeni put množenjem brzine i trajanja simulacije. Kao rezultat se dobije 0,24 m, što i dalje naveliko spada pod rezoluciju udaljenosti vozila koja sadrži iznose unutar 0 m. Također, vidljivo je da se vozilo zaustavilo puno prije nego što je pješak došao do ruba pješačkog prijelaza, te time nije bilo potrebe da pješak mijenja brzinu svog kretanja ili ubrzava/usporava. [Slika 64]

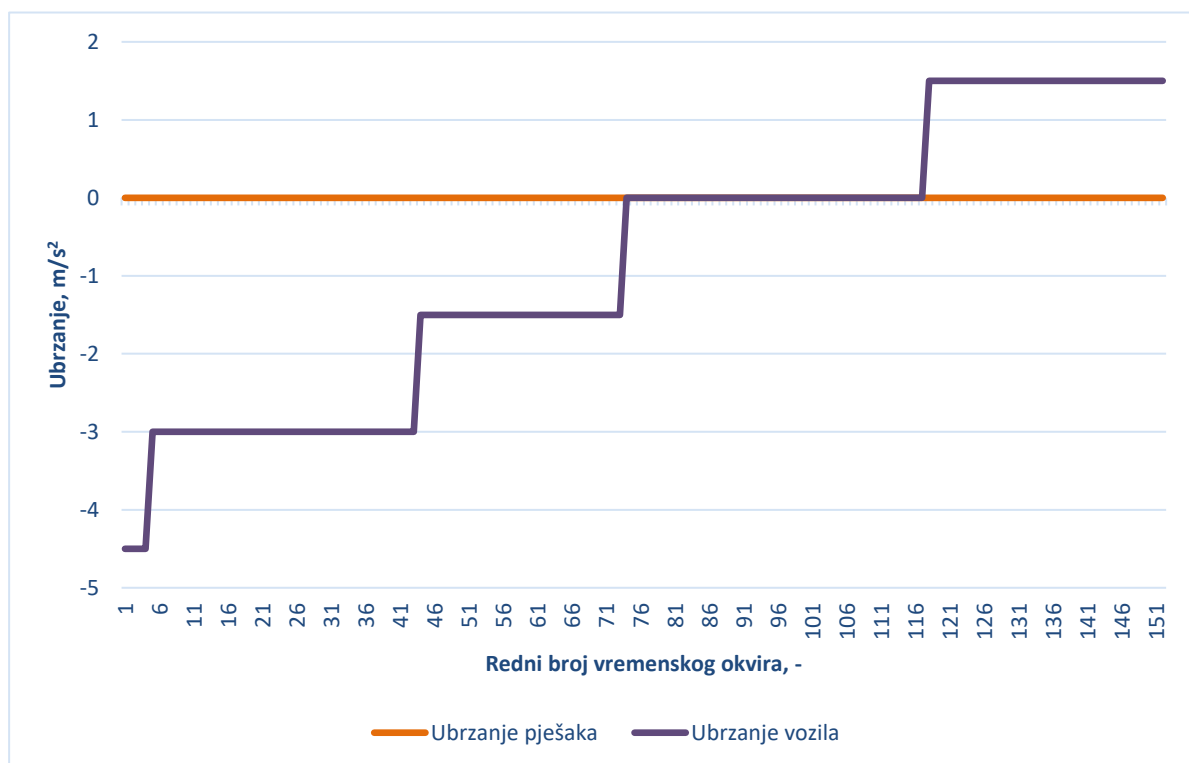
Potrebno je napomenuti da je ubrzavanje ili usporavanje pri prelaženju pješačkog prijelaza često rezultat individualne odluke – neki od pješaka osjećaju nelagodu i potencijalnu opasnost pri zadržavanju na cesti, dok drugi ne izražavaju značajnu razliku u hodaњу po pločniku ili po pješačkom prijelazu.



Slika 62. Usporedbeni prikaz udaljenosti pješaka i vozila *Simulacije 29*



Slika 63. Usporedbeni prikaz brzine pješaka i vozila *Simulacije 29*



Slika 64. Usporedbeni prikaz ubrzanja pješaka i vozila *Simulacije 29*

5.4. Validacija pomoću 100 simulacija predefiniranog početnog stanja

Sljedeća metoda validacije podrazumijeva usporedbu snimljenog seta podataka sa 100 simulacija, u slučaju kada im je početno stanje jednako. I ovdje je model kreiran pomoću istog rječnika koji se prvi put spominje u četvrtom poglavlju. S obzirom na činjenicu da se početna stanja za pet testnih stanja eksplicitno ne nalaze u rječniku jer nisu ni sudjelovala u njegovom kreiranju, odabrana su stanja najsličnija njima.

Pregledom podataka ustanovljeno je da je prvo stanje za snimke *Test 1* i *Test 3* identično prvom stanju za snimku rednog broja 3, dok je prvo stanje snimke *Test 5* identično prvom stanju pod rednim brojem 20.

Za početno stanje za svih 100 simulacija koje se uspoređuju sa snimkom pod nazivom *Test 2*, odabrano je stanje pod rednim brojem 28. Razlika između vrijednosti varijabli u *Testu 2* i snimci 28 je u brzini pješaka koja je za $\Delta v_{pd} = 0.5$ m/s veća u *Testu 2*.

Najsličnije početno stanje početnom stanju snimke *Test 4* je stanje snimke 27. U ovom slučaju razlika u iznosima varijabli iznosi $\Delta v_v = 3$ m/s. Ovakva razlika je značajnija od razlike za *Test 2*, što se naknadno može vidjeti u rezultatima.

Za kreiranje 100 simulacija po svakoj snimci (sveukupno je kreirano 3500 simulacija), modificiran je prethodno kreirani program u *Pythonu*. Kao početna stanja nisu se više uzimala stanja iz liste svih stanja s većom udaljenosti od 5 m, već je kreirana nova lista koja je u sebi sadržavala svih 35 početnih stanja (uz prethodnu napomenu o odabranim najsličnijim stanjima za snimke *Test 2* i *Test 4*). Program je kreiran tako da prolazi kroz listu početnih stanja te za svaku kreira 100 slučajnih slijedova stanja. Podaci o svim novokreiranim simulacijama spremljeni su u obliku *.csv* datoteke za pojedinu snimku.

5.4.1. Korijen srednje kvadratne vrijednosti pogreške rezultata

Nadalje, pri statističkoj obradi podataka, kako bi se odredila sličnost simulacija stvarnim snimkama, odabrana je metoda validacije pomoću srednje kvadratne vrijednosti pogreške (veličine slične varijanci), odnosno njezinog drugog korijena – RMSE (eng. *Root Mean Squared Error*), tj. veličine slične standardnoj devijaciji. [16]

Ove se veličine uobičajeno koriste kao pokazatelji kvalitete rezultata simulacija, odnosno jednostavni pokazatelji odstupanja simulacijom predviđenih vrijednosti od zabilježenih, tj. stvarnih ili eksperimentalno dobivenih podataka. [16]

Korijen srednje kvadratne vrijednosti pogreške računa se prema izrazu:

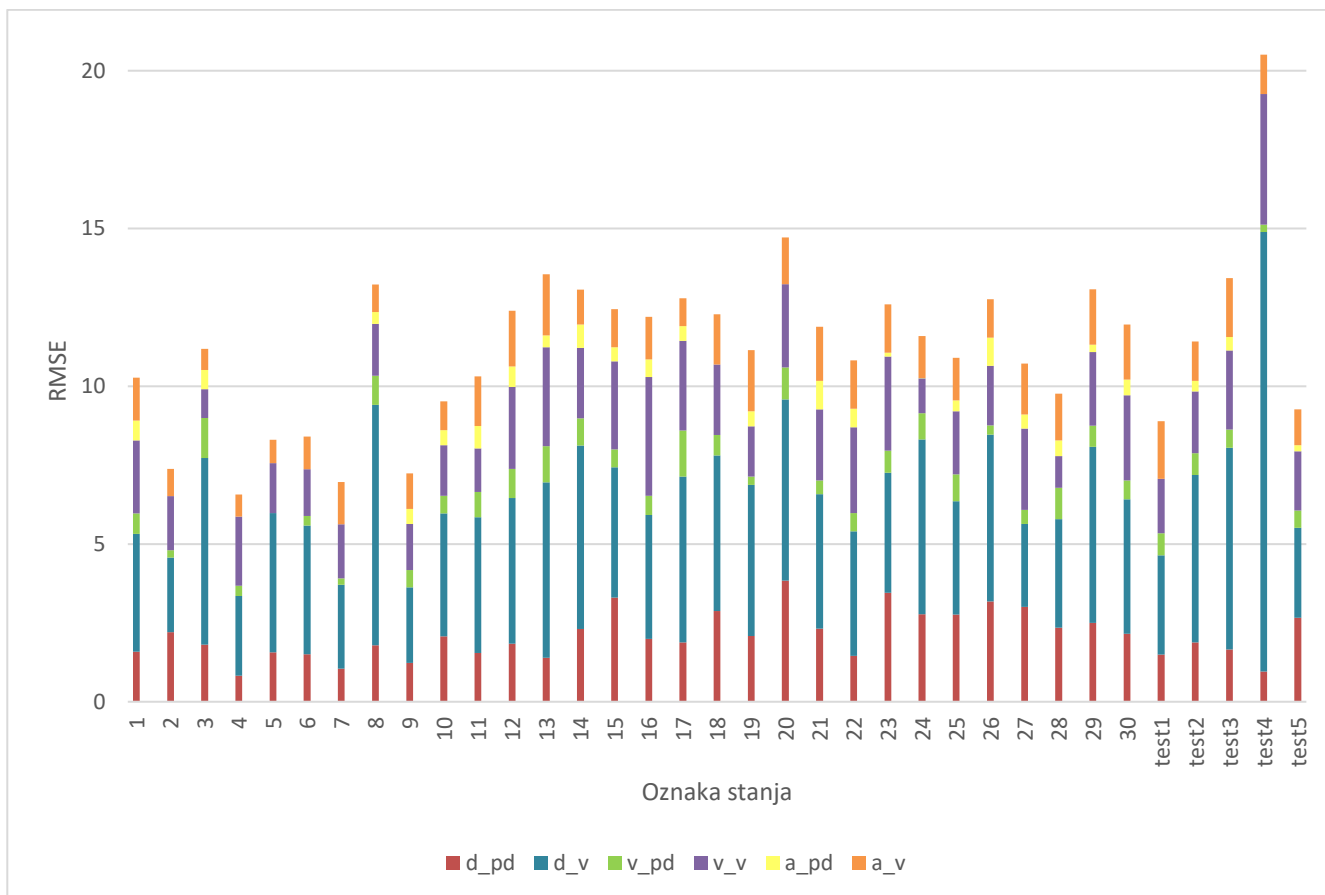
$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - x_{p,i})^2}, \quad (14)$$

koji slijedi iz prethodno definiranog izraza (13), gdje je jedina razlika što u izrazu (13) x_m označava prosjek, dok je u izrazu (14) $x_{p,i}$ predviđena, odnosno simulacijom dobivena vrijednost razmatrane varijable.

Slika 65 prikazuje raspodjelu srednje kvadratne vrijednosti po varijabli stanja za svako stanje.

Oznake predstavljaju sljedeće:

1. d_{pd} – položaj pješaka ili udaljenost pješaka od pješačkog prijelaza
1. d_v – položaj vozila ili udaljenost vozila od pješačkog prijelaza
2. v_{pd} – brzina pješaka
3. v_v – brzina vozila
4. a_{pd} – ubrzanje pješaka
5. a_v – ubrzanje vozila



Slika 65. RMSE za svih 6 stanja

Iz samog grafičkog prikaza je vidljivo da suma svih grešaka u prosjeku iznosi 11,24, a varira između iznosa 6,56 za simulacije pod brojem 4 i 14,71 za simulacije pod rednim brojem 20. Iznimka je suma grešaka za simulaciju testnog skupa podataka broja 4. Ondje ukupna greška iznosi 20,51. Najveći udio u toj grešci ima iznos greške položaja pješaka, čak 13,93. Slijedi ga iznos greške brzine vozila iznosa 4,14.

Obrazloženje za tako velike greške može se pronaći u činjenici da je za simulaciju skupa za *Test 4* korištena snimka koja tom testu nije potpuno identična, već im je slično samo početno stanje. U vidu početnog stanja, napomenuta je greška u razlici brzine vozila od 3 m/s, što pojašnjava iznos greške iste. Obrazloženje za grešku u položaju vozila je činjenica da se po pitanju udaljenosti vozila gledaju velike rezolucije, od 7,5 m. Ako postoji, svaka greška u iznosu stvarnog i simuliranog stanja je najmanjeg iznosa od 7,5 m, što sumarno može dati veliku pogrešku. Za razliku od udaljenosti vozila, preostale rezolucije se odnose na razliku od 0,5, 1 ili najviše 3 jedinice unutar vremenskih okvira te je rijetkost da će takve rezolucije uzrokovati veću grešku.

Iako je i za simuliranje *Testa 2* korištena snimka različita od originalne, njihove razlike nisu toliko značajne po pitanju svih stanja te je samim time i suma svih pogrešaka nešto malo veća od prosjeka – iznosi 11,42.

Egazaktne vrijednosti RMSE-a navedene su u Tablici 11.

Tablica 11. Srednji RMSE po svakoj varijabli.

<i>Br. simulacije / varijabla</i>	<i>Udaljenost pješaka, d_pd</i>	<i>Udaljenost vozila, d_v</i>	<i>Brzina pješaka, v_pd</i>	<i>Brzina vozila, v_v</i>	<i>Ubrzanje pješaka, a_pd</i>	<i>Ubrzanje vozila, a_v</i>
1	1.584199	3.735718	0.651231	2.308331	0.633145	1.359039
2	2.208142	2.359857	0.23486	1.709985	0	0.670675
3	1.805962	5.916443	1.272206	0.908295	0.612372	0.697183
4	0.823846	2.523193	0.33243	2.186046	0	0.697183
5	1.563043	4.416267	0	1.581453	0	0.73928
6	1.503215	4.078103	0.30884	1.477856	0	1.039567
7	1.046835	2.658883	0.199318	1.714635	0	1.341769
8	1.786238	7.625601	0.918551	1.647685	0.376003	0.872483
9	1.231493	2.396679	0.54632	1.461928	0.479005	1.122018
10	2.070891	3.901222	0.559435	1.593779	0.479613	0.919781
11	1.546053	4.303064	0.795358	1.382194	0.716489	1.57021
12	1.844134	4.613853	0.924528	2.591046	0.655551	1.759505
13	1.394467	5.560899	1.149121	3.126458	0.375612	1.938324
14	2.303271	5.813775	0.869052	2.233675	0.7379	1.099367
15	3.298553	4.136254	0.566598	2.782626	0.449287	1.211882
16	1.987278	3.930974	0.605186	3.772205	0.558598	1.346347
17	1.880274	5.253519	1.466158	2.838613	0.471016	0.876598
18	2.872297	4.930706	0.653788	2.231954	0	1.594057
19	2.082023	4.789575	0.260504	1.596232	0.471016	1.936496
20	3.837639	5.744471	1.017632	2.632486	0	1.481476
21	2.318321	4.257759	0.435287	2.25351	0.903174	1.713389

22	1.457147	3.946315	0.576449	2.713517	0.594994	1.532541
23	3.449844	3.80564	0.702456	2.984902	0.116727	1.53406
24	2.768383	5.543008	0.837614	1.0982	0.014397	1.332796
25	2.763615	3.591769	0.85511	1.992167	0.352829	1.345498
26	3.178427	5.287316	0.298117	1.879575	0.898543	1.215036
27	3.003514	2.632843	0.443311	2.573504	0.453983	1.610749
28	2.349181	3.438968	0.989549	1.007906	0.496649	1.482545
29	2.500768	5.575612	0.675981	2.331784	0.23092	1.757268
30	2.15512	4.264136	0.592757	2.699997	0.501173	1.742046
Test 1	1.49916	3.1417	0.697492	1.724748	0	1.832286
Test 2	1.882045	5.302034	0.693041	1.956984	0.335606	1.247254
Test 3	1.659012	6.392854	0.580219	2.498606	0.425739	1.86704
Test 4	0.95449	13.92998	0.232261	4.139002	0	1.253295
Test 5	2.661003	2.852682	0.546558	1.881707	0.184073	1.140036

5.4.2. Točnost procjene očekivanog ishoda

Jedna od mogućnosti korištenja ovog simuliranog seta podataka svakako je predviđanje ponašanja pješaka u vidu odluke o stupanju na pješački prijelaz prije ili nakon prolaska vozila. Iako dani set ne sadrži dovoljan broj situacija u kojima vozilo ne propušta pješaka, simulacije su dale razne rezultate. Zbrajanjem broja vrijednosti u kojima prvo prolazi pješak (označene oznakom p) u svih 100 simulacija po snimci te dijeljenje tog broja sa 100 daje udio situacija s prvotnim prolaskom pješaka. Analogno se dobivaju rezultati za udio prvog prolaska vozila u odnosu na cjelokupan broj simulacija. Zbroj željenih vrijednosti za simulaciju uvijek iznosi 1, a kao potpun ishod simulacije smatra se varijabla čije je ponavljanje učestalije u 100 simulacija (na primjer, ako je to varijabla p , pješak prelazi pješački prijelaz prije vozila).

Dobivena vrijednosti uspoređena je s vrijednošću u snimkama, koja uvijek iznosi 1, za pješaka ili za vozilo. Smatra se da je simulacija legitiman prikaz stvarne snimke ako je udio očekivanih prelazaka pješačkog prijelaza sličan ili donekle identičan snimljenom. Prikaz rezultata dan je u Tablici 12.

Tablica 12. Numeracija interakcija po rednom broju.

<i>Redni broj snimke</i>	1	2	3	4	5	6	7	8	9	10
<i>Snimljeni ishod</i>	p	p	p	p	p	p	p	v	p	p
<i>Vjerojatnost snimljenog ishoda</i>	1	1	1	1	1	1	1	1	1	1
<i>Vjerojatnost simuliranog ishoda</i>	0,97	0,99	0,98	0,99	0,88	0,95	0,99	0,97	0,96	0,94
<i>Redni broj</i>	11	12	13	14	15	16	17	18	19	20
<i>Snimljeni ishod</i>	p	p	p	p	p	p	p	p	p	p
<i>Vjerojatnost snimljenog ishoda</i>	1	1	1	1	1	1	1	1	1	1
<i>Vjerojatnost simuliranog ishoda</i>	0,99	0,84	0,95	0,97	0,86	0,97	0,86	0,98	0,9	0,97
<i>Redni broj</i>	21	22	23	24	25	26	27	28	29	30
<i>Snimljeni ishod</i>	p	p	p	v	p	p	v	p	v	p
<i>Vjerojatnost snimljenog ishoda</i>	1	1	1	1	1	1	1	1	1	1
<i>Vjerojatnost simuliranog ishoda</i>	0,95	0,84	0,07	0,08	0,97	0,32	0,83	0,97	0,99	0,20
<i>Redni broj</i>	Test 1		Test 2		Test 3		Test 4		Test 5	
<i>Snimljeni ishod</i>	p		p		p		v		p	
<i>Vjerojatnost snimljenog ishoda</i>	1		1		1		1		1	
<i>Vjerojatnost simuliranog ishoda</i>	0,94		0,97		0,97		0,02		0,96	

S obzirom na dobivene rezultate, lako je zaključiti kako simulacije u većini slučajeva zadovoljavaju (dobivaju se očekivane vrijednosti za 89 % snimki). Od četiri snimke koje ne zadovoljavaju, u dvije je prioritetno vozilo, dok se u druge dvije zamijećuje prvi prelazak pješaka. Najmanja preciznost dobivena je za simulaciju snimke *Test 4*. Takva simulacija ni nije bazirana na stanjima iz snimke te je rezultat očekivan – osim snimke *Test 4*, ne postoji ni jedna druga snimka sličnih vrijednosti da bi krajnji rezultat bio zadovoljavajuć. Kod snimki 24, 26 i 30, krivac za netočne rezultate vjerojatno je činjenica da su im stanja poprilično unikatna su time manje šanse da se takvo stanje odabere iz vrijednosti rječnika ključeva. Naravno, uvijek postoji mogućnost da bi se na većem uzorku podataka neka od tih vrijednosti približila željenoj. *Snimka 24*, navedena pod nazivom *Train 4*, prikazana je i grafički na Slici 50 – 55. Iz te analize je vidljivo da među 50 simulacija nasumično odabranog početnog stanja nije pronađena simulacija koja bi imalo odgovarala originalnoj *Snimci 24*. Budući da ni pomoću 100 nasumičnih simulacija početnog stanja jednakog početnom stanju *Snimke 24* nije dobiven dovoljan broj adekvatnih simulacija koje bi rezultirale prelaskom vozila preko pješačkog prijelaza prije pješaka, zaključak je da je i prethodni skup simulacija, njih 50, reprezentativan za ono što se naknadno potvrđuje unutar 100 simulacija.

Pregledom podataka o *Snimci 24* zaključeno je kako njena stanja nisu neuobičajena koliko su rijetka u odnosu na preostale snimke sa sličnim uvjetima (veći je broj snimki u kojima pješak prelazi prvi te je time veći broj takvih stanja).

5.4.3. Vrijeme do dolaska i vrijeme do sudara

Brojni psihološki faktori utječu na prelaske pješaka preko ceste. Pješaci, kao za autonomna vozila jedna od najteže predvidivih skupina sudionika u prometu, s vozačima komuniciraju na neverbalne načine i njihovo je ponašanje ovisno o brojnim unutarnjim i vanjskim faktorima. Među najvažnijim faktorima koji definiraju ponašanje pješaka u prometu i njihove prelaske ceste identificirani su brzina nadolazećeg vozila i njegova udaljenost. Ova dva faktora, uzimajući pritom ili ne uzimajući u obzir i brzinu pješaka, moguće je povezati u jedan, kombinirani parametar u vidu vremena. Na taj način pojednostavljuje se, zadržavajući pritom i fizikalni smisao, analiza utjecaja ovih veličina u razmatranoj problematici.

Vrijeme do dolaska (eng. *time to arrival, TTA*), metoda je validacije dobivenog rješenja. Ova veličina predstavlja vrijeme do kontakta uzimajući u obzir pokretnog promatrača i nepokretnu,

tj. mirujuću metu. S druge strane, vrijeme do sudara – TTC (eng. *Time To Collision*) veličina je opisana pokretnim promatračem i pokretnom metom. [17]

U ovom radu detaljnije je razmatrana samo veličina TTA – vrijeme do dolaska. Uzimajući prikladnu brzinu vozila v i udaljenost vozila od pješaka s , vrijeme do dolaska računa se prema izrazu:

$$TTA = \frac{s}{v}. \quad (15)$$

U literaturi se značenja TTA i TTC često preklapaju.

Za izračun vremena dolaska, kreiran je novi programski kod. On je podrazumijevao učitavanje datoteka sa simulacijama te pronalaženje trenutka prije stupanja pješaka na pješački prijelaz (zadnji vremenski okvir u kojem je udaljenost pješaka iznosila -1). Za taj vremenski okvir, zapamćeni su iznosi varijabli udaljenosti i brzine vozila, te je izračunat njihov omjer, to jest, vrijeme dolaska, sukladno jednadžbi (15). Sve vrijednosti za izračunata vremena dolaska spremile su se u novokreiranu datoteku naziva *tta.csv*.

Slijedno tome je učitana *.csv* datoteka sa svim originalnim snimkama, iz koje je za pojedinačnu snimku na identičan način izračunato vrijeme dolaska.

Idejno, naredni korak pretpostavljao je izračun prosječne vrijednosti svih 100 dobivenih vremena po snimci. No, u više navrata kao rezultat vremena dolaska dobiveno je beskonačno vrijeme, što je i fizikalno smisleno s obzirom na to da se radilo o trenutcima u kojima je vozilo stajalo (brzina vozila bila je 0 m/s). U takvim slučajevima je, ovisno o broju simulacija u kojima se kao iznos TTA pojavljuje beskonačnost, za prosječno vrijeme odabrana ili beskonačnost ili neki prosječno dobiven broj. Na primjer, kada bi se unutar 100 simulacija za neku snimku u par simulacija pojavio iznos od beskonačno sekundi, simulacije takvog rezultata bi se eliminirale i krajnja suma svih TTA po simulaciji bi se podijelila s brojem simulacija u kojem je kao rezultat dobiven realan broj. U tim slučajevima, simulacije s rezultatom od $TTA = \infty$ nisu smatrane validnima za daljnji izračun.

Nasuprot tome, kada bi se u većini slučajeva kao vrijednost TTA dobila beskonačnost, ona je odabrana kao prosjek svih dobivenih vrijednosti.

Beskonačnost kao rezultat nije moguće izbjeći. Vjerojatna je mogućnost da je u ovom radu velik broj brzina automobila svrstan pod 0 m/s, radi rezolucije, iako se stvarna vrijednost može razlikovati od 0 m/s. Rezolucija je poprilično gruba za računanje TTA-a, no i dalje je ostavljena u ovom formatu radi preostalih izračuna te dobivanja kvalitetnog modela. No, i u slučaju vrlo

malih brzina različitih od nule, TTA bi ispao poprilično velikog iznosa, što znači da ni pogreška u vidu $TTA = \infty$ nije nužno daleka od stvarne situacije.

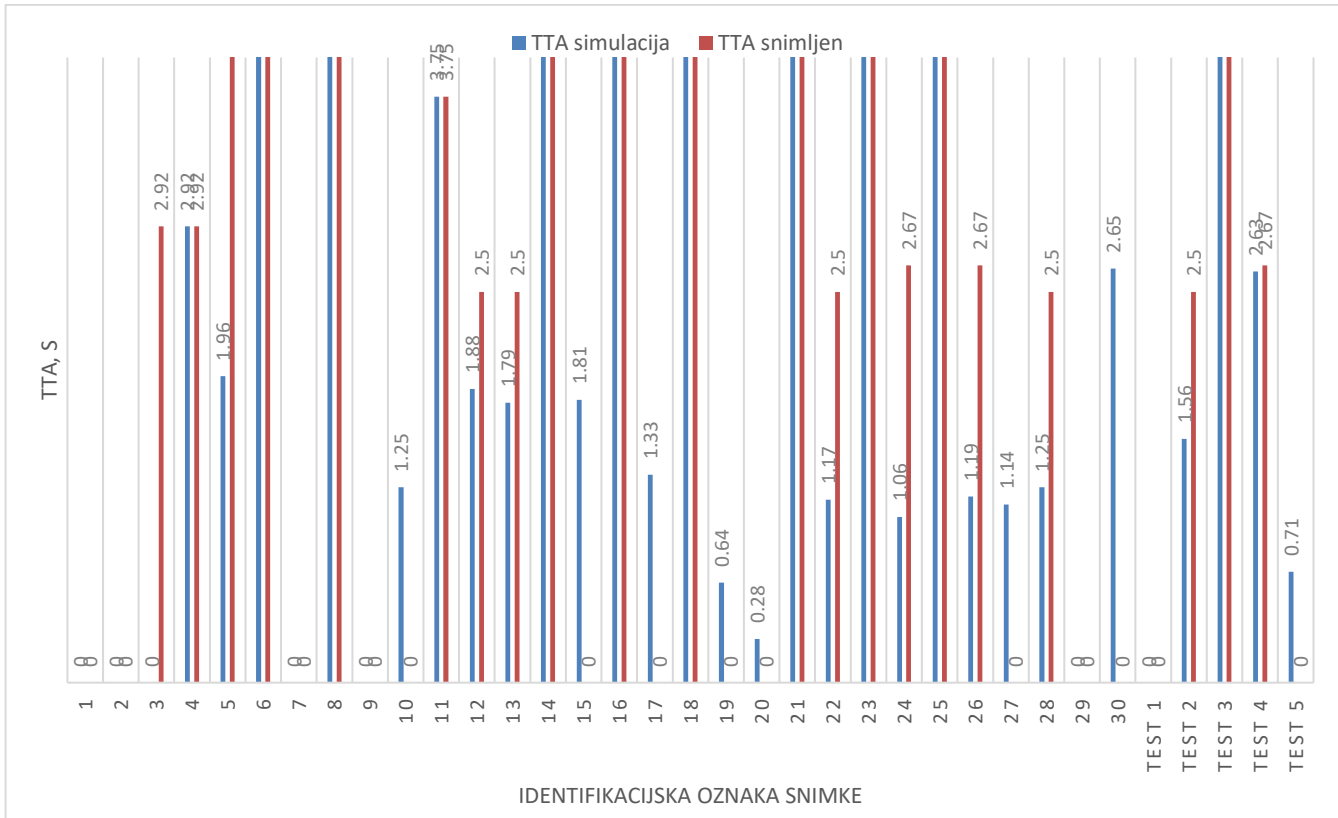
Tablica 13 i Slika 66 prikazuju tekstualni i grafički prikaz odnosa snimljenih i simulacijski dobivenih vrijednosti za TTA. Prva slika prikazuje dijagram s odnosom svih vrijednosti, dok su na drugoj slici uklonjeni svi redni brojevi snimki gdje je snimljena ili simulirana vrijednost TTA jednaka beskonačnosti.

Neovisno o slici, sami podaci iz tablice daju informaciju o poprilično velikoj grešci između snimljene i simulirane vrijednosti te se samim time ovaj model ne može koristiti kao adekvatan za kvalitetnu procjenu TTA. Jedan od razloga je svakako i onaj već spomenut – rezolucije brzine, ali i udaljenosti, pregrube su za izračun TTA u kojem, u stvarnoj primjeni, i druga decimala vrijednosti sekunde čini razliku pri odabiru aktivnosti koje računalo vozila treba poduzeti.

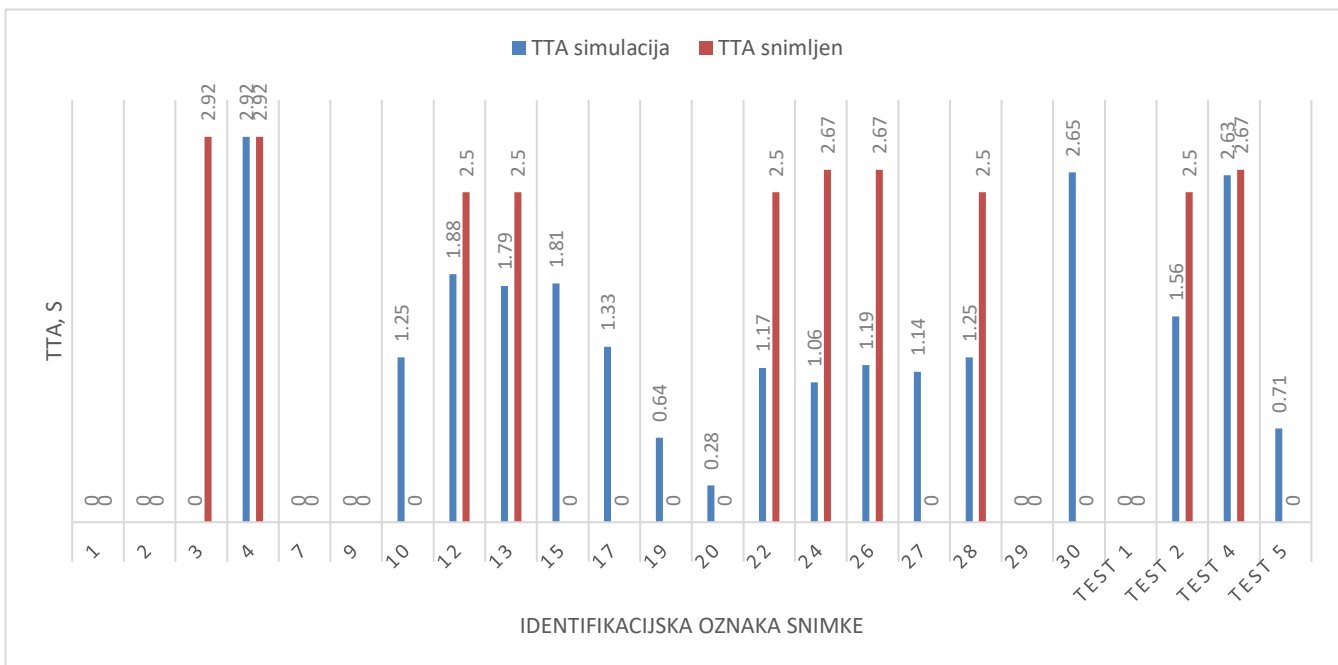
Tablica 13. Usporedba snimljenih i simuliranih TTA

<i>Redni broj</i>	1	2	3	4	5	6	7	8	9	10	11
<i>Snimljeni ishod</i>	0	0	2,92	2,92	∞	∞	0	∞	0	0	3,75
<i>Simulirani ishod</i>	0	0	0	2,92	1,96	∞	0	∞	0	1,25	3,75
<i>Apsolutna razlika</i>	0	0	2,92	0	∞	0	0	0	0	1,25	0
<i>Redni broj</i>	12	13	14	15	16	17	18	19	20	21	22
<i>Snimljeni ishod</i>	2,5	2,5	∞	0	∞	0	∞	0	0	∞	2,5
<i>Simulirani ishod</i>	1,88	1,79	∞	1,81	∞	1,33	∞	0,64	0,28	∞	1,17
<i>Apsolutna razlika</i>	0,62	0,71	0	1,81	0	1,33	0	0,64	0,28	0	∞
<i>Redni broj</i>	23	24	25	26	27	28	29	30			
<i>Snimljeni ishod</i>	∞	2,67	∞	2,67	0	2,5	0	0			
<i>Simulirani ishod</i>	∞	1,06	∞	1,19	1,14	1,25	0	2,65			
<i>Apsolutna razlika</i>	0	1,61	0	1,48	1,14	1,25	0	2,65			
<i>Redni broj</i>	Test 1		Test 2		Test 3		Test 4		Test 5		
<i>Snimljeni ishod</i>	0		2,5		∞		2,67		0		
<i>Simulirani ishod</i>	0		1,56		∞		2,63		0,71		
<i>Apsolutna razlika</i>	0		0,94		0		0,04		0,71		

Slika 66 prikazuje odnos snimljenog i simuliranog TTA. Na Slici 67 prikazan je uspoređan z

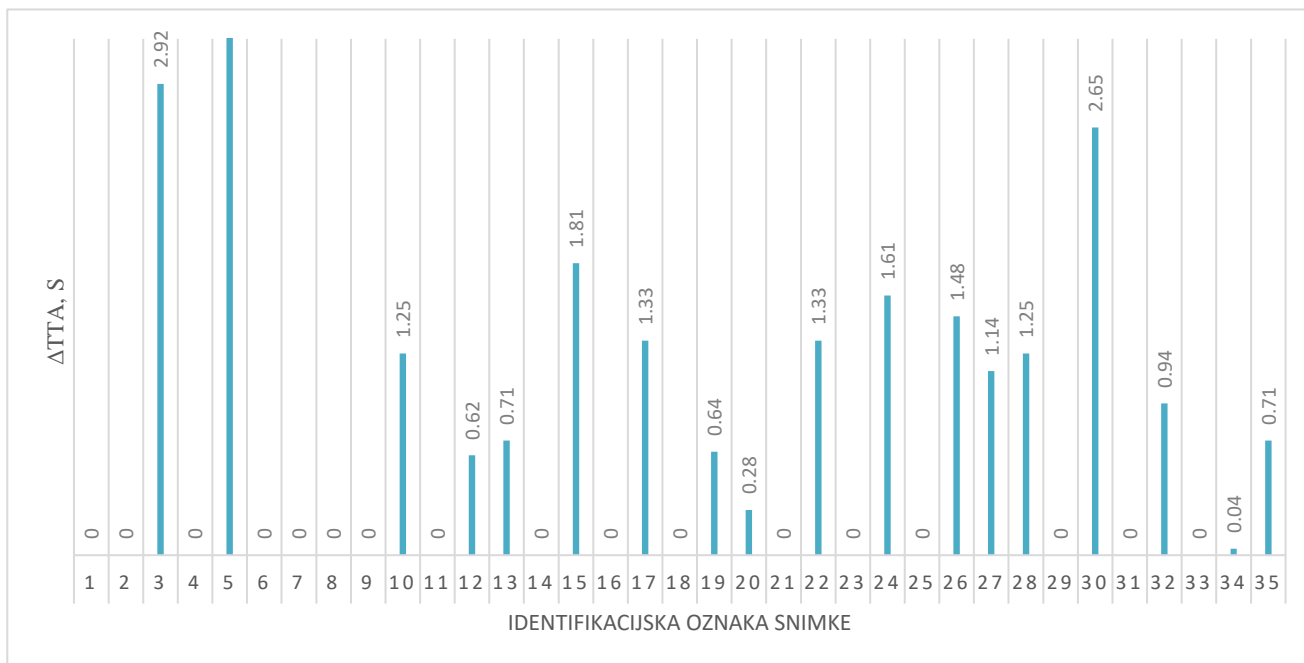


Slika 66. Snimljene i snimulirane vrijednosti TTA-a



Slika 67. Snimljene i simulirane vrijednosti TTA (bez iznosa jednakih ∞)

Slika 68 prikazuje iznos apsolutne greške po snimci. U 18 od 35 snimaka vidljiva je greška različita od nule, dok je u 11 od tih 18 simulacija greška veća od 1 sekunde, što prikazuje izrazitu nepreciznost modela.



Slika 68. Apsolutna greška TTA po snimci

6. PRIMJENA MODELA

U ovom poglavlju razmatra se mogućnost primjene dobivenog modela za optimalno upravljanje brzinom vozila koje prilazi pješačkom prijelazu s pješacima.

Kako bi se ovaj model mogao primjenjivati u istu svrhu, prvo je potrebno definirati je li njegovo ponašanje u slučaju prilaska pješačkom prijelazu adekvatno, točnije, hoće li se vozilo zaustaviti i propustiti pješaka kada pješak pokaže jasnu namjeru stupanja na pješački prijelaz. Isto tako, postavlja se pitanje može li vozilo i sa kojom točnošću procijeniti nakane pješaka. Za takve slučajeve nema prostora za greške.

U cilju utvrđivanja mogućnosti takve primjene, potrebno je obaviti dodatna testiranja uz rezultate prikazane u potpoglavlju 5.4.2. točnije, potrebno je da uz dodatna testiranja i rezultati simuliranih ishoda prikazanih u Tablici 12 budu jednaki rezultatima snimljenih ishoda.

Prvo će biti raspravljeno kako je moguće poboljšati rezultate već dobivene simulacijom prikazanom u ovom radu (točnije, kako poboljšati model), a naknadno će biti napomenuto nekoliko dodatnih vrsta testiranja za provjeru točnosti.

Kako bi se poboljšali rezultati modela stvorenog u ovom radu, neophodno je promijeniti rezoluciju. Naime, pri svrstavanju udaljenosti vozila u rezolucije, udaljenost od 3,8 m smatrat će se udaljenošću od 7,5 m, dok će se udaljenost od 3,5 m smatrati udaljenošću od 0 m. Pri potrebi za sto-postotnom točnošću procjene ponašanja, rezolucije je potrebno smanjiti na čim manje vrijednosti, ovisno o tome koliko skup podataka dozvoljava. Rezolucija se ne smije smanjiti na štetu smanjenja broja vrijednosti unutar ključa rječnika.

U slučaju da se smanjenjem rezolucije smanji broj vrijednosti pod ključem nekog stanja, neophodno je proširiti skup podataka s većim brojem vrijednosti. Proširivanje skupa podataka je svakako poželjno za kreiranje većeg spektra jedinstvenih simuliranih događaja.

Uz sve navedeno, svakako je potrebno provjeriti je li se svaki od sudionika u prometu na snimkama ponašao u skladu s prometnim propisima. Primjerice, uvidom u snimke *inD* skupa podataka vidljivo je da se vozilo nije zaustavilo svaki put kada bi se na udaljenosti unutar 1 m od pješačkog prijelaza pojavio pješak. Uzimanje fragmenata takvih situacija kao stanja za kreiranje simulacije nikako ne može rezultirati željenim (ispravnim) ishodom.

Na kraju, neophodno je razumjeti što se od simulacije očekuje te samim time eliminirati podatke simulacije čiji su rezultati nesuvisli te simulacije koje nikako ne odgovaraju očekivanoj vrsti

simulacija. Još bolji pristup bio bi traženje uzroka za postojanje nesuvislih simulacija te eliminiranje istih.

Preostale mogućnosti testiranja odnose se i na testiranja s podacima iz ovog skupa. Primjerice, moguće je modificirati vrijednosti u rječniku na način da se uklone sva stanja osim onih koja za dani set vrijednosti pokazuju maksimum brzine i akceleracije vozila, te opcionalno minimum udaljenosti vozila. Reduciranjem vrijednosti rječnika na takve skupove, omogućuje se kreiranje simulacija u kojima vozilo temeljem svoje brzine i ubrzanja ili ne pokazuje namjeru propuštanja pješaka ili će je pokazati tek u krajnjem slučaju, na manjoj udaljenosti od one na kojoj bi se očekivalo smanjenje akceleracije vozila. Iako je takva simulacija umjetno stvorena (uklanjanjem stanja iz vrijednosti ključa smanjuje se nasumičnost) te je broj jedinstvenih simulacija u takvim uvjetima ograničen, i takav skup podataka može poslužiti za neki oblik validacije ponašanja vozila.

Manipulacijom vrijednostima iz rječnika moguće je kreirati i situacije gdje vozilo, nakon usporavanja pri približavanju pješačkom prijelazu, donese drugačiju odluku te krene ubrzavati. Isto se može primijeniti na ponašanje pješaka, gdje bi se pješak mogao približiti pješačkom prijelazu, ili čak stupiti na njega, no u međuvremenu promijeniti odluku i odmaknuti se od ceste. Naknadno, rezultati ovih testiranja mogu se primijeniti na određivanje brzine vozila pri upravljanju vozilom. Uz očekivano usporavanje u blizini pješačkog prijelaza, moguće je to usporavanje svesti na potreban minimum, ako se u blizini pješačkog prijelaza ne nalazi pješak ili ako taj pješak zasigurno ne pokazuje namjeru prelaska preko prijelaza. Naravno, za određivanje optimalne brzine vozila, moguće je promatrati i putanju pješaka – pješak se može nalaziti u blizini pješačkog prijelaza bez da mu je smjer kretanja orijentiran prema prijelazu te bez da planira kretanje u narednim trenucima koliko je potrebno vozilu da prijeđe tu dionicu ceste.

7. ZAKLJUČAK

U ovom radu prikazan je model ponašanja pješaka u blizini pješačkih prijelaza temeljen na Markovljevim lancima. Model je kreiran na temelju postojećih podataka o prijelazu ceste, sakupljenim u poprilično idealnim uvjetima – snimanje se odvijalo tijekom dana, vrijeme je bilo vedro, a vremenske neprilike nisu smanjivale vidljivost ni pješacima ni vozačima. Snimanje je obavljeno dronom iz zraka, što je uklonilo utjecaj na ponašanje sudionika u slučaju svjesnosti činjenice da ih se promatra. Sve to kao rezultat dalo je model koji može u kratkom vremenu generirati velik broj simulacija, gdje sva stanja potječu iz fragmenata stvarnih događanja.

Model koji je nastao kao rezultat ovog rada pokazao se vjerodostojnim za daljnju primjenu. Provedene su tri metode validacije. U prvom skupu validacije, odabran je predefiniран skup početnih stanja svake simulacije te se simulacija pokrenula. Dobiveno je 50 rezultata koji su prikazani grafički, zajedno s najbližijom interakcijom iz testnog skupa. Uz testni skup, iz skupa od 30 simulacija čija se stanja nalaze u rječniku, pet je naknadno odabrano i uspoređeno s rezultatima prethodnih 50 simulacija. Rezultati simulacija prikazani su histogramima, to jest, prikazana je raspodjela iznosa udaljenosti, brzina i ubrzanja na svojstvenim dijagramima za pješake i za vozila. Količina nasumično dobivenih iznosa odgovarala je onoj u skupu podataka za učenje. Raspodjele na histogramima su se podudarale s raspodjelama na histogramima kreiranim sa snimljenog skupa podataka, to jest, skup podataka validiran je pomoću dijagrama koji su prikazivali sve snimljene iznose kojima se opisuju stanja. Vidljiva je korelacija između ponašanja pješaka opisanog prijednom udaljenošću u odnosu na njegovu brzinu i ubrzanje u istim vremenskim okvirima. Isto je potvrđeno i za vozila. Nadalje, usporedbenim grafičkim prikazom, uz prethodne, vidljivo je kako je odnos prelaska pješaka i vozila preko prijelaza realističan te da nije došlo do situacija gdje bi prelazak oba subjekta interferirao.

Druga vrsta validacije podrazumijevala je pokretanje 100 simulacija iz identičnog početnog stanja svake od 30 interakcija te za pet testnih slučajeva. Kod testnih slučajeva, početno stanje se ne nalazi u rječniku te je odabrano stanje iz rječnika najbližije istom. Za svaku od simuliranih varijabli izračunat je korijen srednje kvadratne vrijednosti pogreške (RMSE, eng. *root mean squared error*) te je isti podijeljen sa brojem simulacija. Rezultati su pokazali da RMSE ne odstupa za testni set podataka u odnosu na set podataka za treniranje. Naime, RMSE je na nekim testnim skupovima ispao manji u odnosu na RMSE izračunat za neke od snimki korištenih za treniranje. RMSE međusobno ne odstupa značajno ni za jedan od mjerenih

elemenata stanja interakcija, a u slučajevima gdje je izuzetno veći od RMSE-a ostalih snimaka, dano je jednostavno objašnjenje. Na istom skupu od 100 simulacija po snimci dan je i izračun točnosti procjene prvog prelaska preko pješačkog prijelaza. Dok je za neke snimke točnost iznosila i do 99 %, u par simulacija, točnost je bila samo 2 %. Sveukupna točnost procjene prednosti prelaska je 89 %, što je s matematičke strane zadovoljavajuće, a sa strane realne primjene i implementacije u vozilo, potpuno nezadovoljavajuće.

Kao treća vrsta validacije, za proteklih 35 x 100 simulacija, izračunat je srednji TTA (eng. *Time to Arrival*) te uspoređen s vrijednostima iz snimljenog seta podataka. Točno procijenjen TTA pojavljuje se u samo 49 % snimki. Od preostalog skupa gdje TTA nije dobro procijenjen, 61 % njih sadržava grešku veću od 1 sekunde. Ne postoji razlika u redu veličine greške na skupu za učenje i skupu za testiranje. Štoviše, greška u skupu za testiranje iznosi 40 %, što je manje od prosječne vrijednosti sveukupne greške. Prethodno rečeno potvrđuje da su simulacije realan prikaz i testnog skupa podataka i skupa za učenje, no zbog raznih razloga, među kojima je svakako i rezolucija vrijednosti korištena u modelu, model nije dovoljno dobar da bi se koristio za procjenu TTA.

Iako je ovaj model dokazano vjerodostojan za korištenje u istim ili donekle sličnim uvjetima u kojima je i snimljen, naveliko u obzir ne uzima dodatne ljudske faktore. Primjeri su dob ili spol pješaka ili vozača, njihovo trenutno raspoloženje, razina njihove pažnje usmjerene na promet, iskustvo sudjelovanja u prometu... Cilj ovog rada bio je i objektivizirati i normirati ponašanje pješaka i vozača. Samim time zanemarila se ona subjektivna strana ljudskog ponašanja koje ne mora nužno biti logična. Računala funkcioniraju na način da će često donijeti najpovoljniju odluku za sve subjekte; nasuprot tome, ljudsko ponašanje se često pokazalo kao netemeljeno na logici i najmanjem utrošku resursa, vremena ili energije.

Kao proširenje ovog rada, moguće je kreirati simulacije vezane uz prelazak pješaka preko pješačkog prijelaza kada vozilo nailazi u daljnjem traku, kada vozilo u bližem ili daljem traku već stoji ili simulacije prelaska većeg broja pješaka preko prijelaza. Ondje je moguće uzimati u obzir veliku varijaciju u situacijama, na primjer, kada pješaci kreću s iste početne točke, kada se kreću u suprotnim smjerovima te kada idu u istom smjeru, ali drugi pješak dolazi do pješačkog prijelaza sa zakašnjenjem u odnosu na prvog pješaka. Skup podataka sadrži i podatke vezane uz bicikliste te bi se isti također mogli modelirati, u ulozi subjekata koji prelaze pješački ili u ulozi subjekata na kolniku. Baš zbog te varijacije u vidu redovitog prelaska iz stanja sličnog

ponašanju pješaka u stanje slično ponašanju vozila, ponašanje biciklista je zahtjevnije za modeliranje.

Idealne simulacije cijelog prometa nema, no kada se traži model gdje se na temelju stvarnih ponašanja kreira umjetan skup ponašanja, Markovljevi lanci se pokazuju kao odličan alat. Jednostavno ih je kreirati čak i na temelju već postojećeg rješenja (unijeti svoje vrijednosti u već postojeći programski kod), a nisu računalno zahtjevni. Trajanje jedne simulacije od 200-tinjak stanja, koje uključuje čitanje iz rječnika i slučajan odabir sljedećeg stanja, u prosjeku traje 2-3 sekunde. Također, pri treniranju autonomnih vozila, velik problem je skup podataka za obradu ili učenje vozila, a ovakvim pristupima moguće je generirati skupove po značaju ekvivalentne stvarnim snimkama.

LITERATURA

- [1] Prvo poluatonomno vozilo: <https://computerhistory.org/blog/where-to-a-history-of-autonomous-vehicles/?key=where-to-a-history-of-autonomous-vehicles> (pristupljeno: 29.10.2022.)
- [2] Rasouli, A.; Kotseruba, I.; Tsotsos, J.K.: Agreeing to cross: How drivers and pedestrians communicate. *2017 IEEE Intelligent Vehicles Symposium (IV)*, 2017., pp. 264-269.
- [3] Rasouli, A.; Kotseruba, I.; Tsotsos, J.K.: Are They Going to Cross? A Benchmark Dataset and Baseline for Pedestrian Crosswalk Behavior, *2017 IEEE International Conference on Computer Vision Workshops (ICCVW)*, 2017., pp. 206-213
- [4] Rasouli, A.; Kotseruba, I.; Kunic, T.; Tsotsos, J.: PIE: A Large-Scale Dataset and Models for Pedestrian Intention Estimation and Trajectory Prediction, *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, 2019., pp. 6261-6270
- [5] Bock, J.; Krajewski, R.; Moers, T.; Runde, S.; Vater, L.; Eckstein, L.: The inD Dataset: A Drone Dataset of Naturalistic Road User Trajectories at German Intersections" *2020 IEEE Intelligent Vehicles Symposium (IV)*, 2020., pp. 1929-1934
- [6] Modul UTM: <https://pypi.org/project/utm/#description> (pristupljeno: 20.11.2022.)
- [7] Haversinusna formula: <https://community.esri.com/groups/coordinate-reference-systems/blog/2017/10/05/haversineformula> (pristupljeno: 24.11.2022.)
- [8] Točnost izračuna udaljenosti na Zmeljnoj plohi: <https://blog.mapbox.com/fast-geodesic-approximations-with-cheap-ruler-106f229ad016> (pristupljeno: 29.11.2022.)
- [9] Haversinusna formula: <https://www.baeldung.com/cs/haversine-formula> (pristupljeno: 9.11.2022.)
- [10] Modul haversine: <https://pypi.org/project/haversine/> (pristupljeno: 29.10.2022.)
- [11] Pavelko, L.; Škugor, B.; Deur, J.; Ivanović, V.; Tseng, H.E.: Modification and Experimental Validation of a Logistic Regression Vehicle-Pedestrian Model
- [12] Šekoronja, B. Markovljevi lanci, podloge za vježbe iz kolegija Umjetna inteligencija, Sveučilište u Zagrebu: Fakultet strojarstva i brodogradnje;
- [13] Monte Carlo simulacije: <https://www.ibm.com/cloud/learn/monte-carlo-simulation> (pristupljeno: 29.11.2022.)
- [14] Funkcija Random: <https://docs.python.org/3/library/random.htm> (pristupljeno: 27.11.2022.)
- [15] Funkcija Random: <https://realpython.com/python-random/> (pristupljeno: 11.10.2022.)

- [16] RMSE: <https://c3.ai/glossary/data-science/root-mean-square-error-rmse/> (pristupljeno: 24.11.2022.)
- [17] Solini, H.; Andre, J.: Time-to-arrival estimations to simulated pedestrians, *Accident Analysis & Prevention*, Vol. 146, 2020.