

NOVI DINAMIČKI PROTOKOL RUTIRANJA ZA URBANE VANET MREŽE BAZIRAN NA UČENJU POTKREPLJIVANJEM

Pavle Bugarčić, Marija Malnar
Univerzitet u Beogradu - Saobraćajni fakultet,
p.bugarcic@sf.bg.ac.rs, m.malnar@sf.bg.ac.rs

Rezime: *Izbor optimalne putanje za slanje podataka u bežičnim ad hoc mrežama od ključnog je značaja za postizanje zadovoljavajućih mrežnih performansi. Ovaj proces postaje veoma kompleksan ukoliko se topologija mreže brzo menja, kao što je slučaj kod bežičnih ad hoc mreža za vozila (Vehicular Ad-hoc Networks, VANETs). U ovakvim uslovima proces rutiranja može biti značajno unapređen uključivanjem mašinskog učenja u izbor optimalne putanje, a tip učenja koji najviše odgovara visoko dinamičkim mrežama je učenje potkrepljivanjem (Reinforcement Learning, RL). Jedan od najznačajnijih tipova RL za dinamičke bežične ad hoc mreže je Q-učenje (Q-Learning, QL). U ovom radu je predložen novi dinamički protokol rutiranja za urbane VANET mreže baziran na QL (Q-learning based Dynamic Routing Algorithm for urban VANETs, Q-DRAV), koji uključivanjem relevantnih mrežnih parametara u RL proces uspeva značajno da poboljša ukupne mrežne performanse kod VANET mreža. Simulaciona analiza i poređenje sa drugim protokolima rutiranja izvršeni su u NS-3 simulatoru.*

Ključne reči: *učenje potkrepljivanjem, Q-učenje, protokoli rutiranja, VANET mreže, NS-3 simulator*

1. Uvod

Protokoli rutiranja u bežičnim komunikacionim mrežama imaju ulogu da obezbede optimalnu putanju za prenos podataka od izvora do odredišta, preko niza međučvorova. U slučaju visoko dinamičkih mreža, kao što su bežične *ad hoc* mreže za vozila (Vehicular Ad-hoc Networks, VANETs), taj zadatak postaje dosta složeniji. Osnovna karakteristika ovih mreža je relativno velika brzina kretanja čvorova, što prouzrokuje česte promene u mrežnoj topologiji. To dovodi do učestalih prekida linkova koji su deo trenutnih putanja, što značajno otežava održavanje optimalne putanje za slanje podataka. Tradicionalni protokoli rutiranja, poput *Ad hoc On-Demand Distance Vector* (AODV) protokola [1], ne uspevaju pravovremeno da reaguju na brze promene u VANET mrežama. Ovo dovodi do kasne promene putanje koja se koristi za prenos podataka u slučaju prekida nekog linka, što rezultira pogoršanjem ukupnih mrežnih performansi. Brojni autori su pokušali da prevaziđu ovaj problem prilagođavanjem postojećih protokola rutiranja dinamičkoj prirodi ovih mreža, kao i definisanjem odgovarajućih metrika rutiranja sa relativno ograničenim uspehom [2-5].

Rešenje ovog problema može biti u uključivanju mašinskog učenja (*Machine Learning*, ML) u proces izbora optimalne putanje, a tip ML koji najviše odgovara zadacima rutiranja u dinamičkim mrežama je učenje potkrepljivanjem (*Reinforcement Learning*, RL). Ovaj pristup podrazumeva stalnu interakciju agenta učenja sa okruženjem putem preduzimanja određenih akcija, na koje okruženje odgovara dodeljivanjem nagrade (ili kazne) za preduzetu akciju, kao i pružanjem povratnih informacija o novom stanju okruženja nakon preduzete akcije. Najčešće korišćeni RL algoritam u VANET mrežama je Q-učenje (*Q-Learning*, QL). Jedan od prvih, a ujedno i veoma značajan primer primene QL u protokolima rutiranja za VANET mreže, je *Q-Learning AODV* (QLAODV) protokol [6], koji pri izboru optimalne putanje uzima u obzir mobilnost vozila i opterećenje propusnog opsega. Još jedan uspešan protokol koji koristi QL za rutiranje u VANET mrežama je *Adaptive Routing Protocol based on Reinforcement Learning* (ARPRL) [7], koji kao uticajne faktore u procesu rutiranja uzima u obzir stabilnost linka, brzinu kretanja vozila i gubitak paketa. Iako ovi protokoli pokazuju značajno unapređenje mrežnih performansi u poređenju sa tradicionalnim protokolima rutiranja, još uvek postoje značajne mogućnosti za njihovo unapređenje. Iz tog razloga je u ovom radu predložen novi dinamički protokol rutiranja za urbane VANET mreže baziran na QL (*Q-learning based Dynamic Routing Algorithm for urban VANETs*, Q-DRAV). Simulacionom analizom u NS-3 simulatoru pokazano je da ovaj protokol pokazuje značajno bolje performanse u pogledu procenta izgubljenih paketa (*Packet Loss Ratio*, PLR), ostvarenog aplikacionog protoka, prosečnog kašnjenja paketa sa kraja na kraj mreže (*End-To-End Delay*, E2ED) i džitera, u poređenju sa ranijim protokolima.

U nastavku rada su najpre objašnjeni osnovni koncepti učenja potkrepljivanjem. Zatim je detaljno opisan novi Q-DRAV protokol rutiranja. Nakon toga su predstavljeni rezultati simulacione analize u NS-3 simulatoru, u okviru kog je izvršeno testiranje i poređenje novog protokola sa ranijim protokolima rutiranja za VANET mreže. U poslednjem poglavlju su data zaključna razmatranja.

2. Osnovni koncepti učenja potkrepljivanjem

U opštem slučaju, učesnici RL procesa su agent učenja i okruženje. Kroz stalnu interakciju sa okruženjem agent prikuplja povratne informacije, na osnovu kojih prilagođava proces rutiranja trenutnom stanju okruženja. RL proces u jednoj VANET mreži moguće je modelirati na sledeći način. Čvor koji želi da pošalje podatke ka određenom odredištu predstavlja agenta učenja, dok svi ostali čvorovi u mreži predstavljaju njegovo okruženje. Izbor putanje kojom će slati podatke ka odredištu predstavlja akciju koju agent može da preduzme, a ona je ekvivalentna izboru jednog od susednih čvorova da bude sledeći hop na putanji. Nakon slanja paketa, susedni čvorovi obaveštavaju agenta o promeni stanja okruženja i o nagradi za preduzetu akciju.

U protokolima rutiranja za VANET mreže koristi se nekoliko različitih tipova RL algoritama, a najpopularniji je QL. On podrazumeva računanje Q-vrednosti za svaku potencijalnu akciju a_t , koju agent može da preduzme u određenom trenutku t dok je okruženje u stanju s_t . Ove vrednosti agent čuva u svojoj Q-tabeli i na osnovu njih bira narednu akciju koju će preduzeti. U opštem slučaju, za uređeni par (s_t, a_t) agent računa Q-vrednost preko sledeće relacije:

$$Q(s_t, a_t) \leftarrow (1 - \alpha) \cdot Q(s_t, a_t) + \alpha \cdot (R_{t+1} + \gamma \cdot \max_a Q(s_{t+1}, a)) \quad (1)$$

U ovoj relaciji R_{t+1} označava nagradu koju je agent dobio od okruženja za preduzetu akciju a_t , a s_{t+1} predstavlja novo stanje okruženja nakon preduzete akcije a_t . Parametar α označava stepen učenja koji može uzeti vrednost u opsegu $[0,1]$ i definiše kojom brzinom će agent učiti od okruženja. Parametar γ predstavlja diskontni faktor, koji određuje važnost budućih nagrada i takođe može uzeti vrednost u opsegu $[0,1]$. Najzad, vrednost $\max_a Q(s_{t+1}, a)$ predstavlja maksimalnu Q-vrednost do koje agent može doći preduzimanjem jedne od dostupnih akcija a , kada je okruženje u narednom stanju s_{t+1} .

3. Predlog novog RL baziranog protokola rutiranja za VANET mreže

Kako bi se unapredile mrežne performanse VANET mreža, razvijen je novi Q-DRAV protokol rutiranja [8]. Za izbor optimalne putanje za slanje paketa, Q-DRAV koristi QL algoritam koji podrazumeva tri metode za ažuriranje Q-vrednosti. Prva metoda odvija se nakon što vozilo primi *Hello* paket od svog suseda. *Hello* paketi sadrže sve potrebne informacije za ažuriranje Q-vrednosti, na osnovu kojih vozila procenjuju kvalitet putanja koje vode ka odredištu preko nekog od svojih suseda. Ove informacije uključuju koordinate suseda po x i y osi, dostupni propusni opseg suseda (*Neighbor Bandwidth*, NB), broj odredišta za koja sused ima maksimalne Q-vrednosti u svojoj Q-tabeli, kao i sekvencu tih maksimalnih Q-vrednosti zajedno sa informacijama o IP adresama odredišta i sledećeg hopa na putanji do odredišta. Kada čvor c primi *Hello* paket od suseda n , ažurira Q-vrednost za putanju do odredišta d , koja vodi preko tog suseda, na osnovu sledeće jednačine:

$$Q_c(d, n) \leftarrow (1 - \alpha_{c,n}) \cdot Q_c(d, n) + \alpha_{c,n} \cdot (R_{c,n} + \gamma_{c,n} \cdot \max_{y \in \text{Nei}(n)} Q_n(d, y)). \quad (2)$$

Ključni korak kod ažuriranja Q-vrednosti je definisanje vrednosti parametara $\alpha_{c,n}$, $\gamma_{c,n}$ i $R_{c,n}$, preko kojih je potrebno uključiti željene uticajne faktore u QL proces. Za inicijalno ažuriranje Q-vrednosti, stepen učenja ($\alpha_{c,n}$) ima konstantnu vrednost. Eksperimentalnim putem je utvrđeno da se najbolji rezultati mrežnih performansi dobijaju kada se ova vrednost postavi na 0,6. U suprotnom, $\alpha_{c,n}$ se izračunava kao:

$$\alpha_{c,n} = \max(0,6; RDC_{c,n}; LU_{c,n}). \quad (3)$$

Parametar $RDC_{c,n}$ (*relative distance change*) predstavlja relativnu promenu rastojanja između trenutnog čvora c i njegovog suseda n . Izračunava se preko sledeće jednačine:

$$RDC_{c,n} = \sqrt{\left| \frac{\Delta d_{c,n}}{\Delta d_{max}} \right|}, \quad (4)$$

gde $\Delta d_{c,n}$ predstavlja promenu rastojanja između posmatranog čvora c i njegovog suseda n tokom perioda pojedinačnog ažuriranja Q-vrednosti, dok Δd_{max} predstavlja maksimalnu moguću promenu rastojanja tokom tog intervala. Na ovaj način, stepen učenja između čvorova koji brže menjaju udaljenost se ubrzava, jer se pretpostavlja da postoje značajnije promene u kvalitetu veze. Kako bi se brže ažurirale putanje sa manjom pouzdanošću, parametar $\alpha_{c,n}$ uzima u obzir i nepouzdanost linka preko koeficijenta $LU_{c,n}$ (*link unreliability*), koji se računa na osnovu sledeće jednačine:

$$LU_{c,n} = \sqrt{1 - (1 - 2^{-SH_{c,n}}) \cdot \frac{RH_{c,n}}{SH_{c,n}}}. \quad (5)$$

U ovoj relaciji parametar $RH_{c,n}$ (*received hello*) predstavlja broj *Hello* paketa koje je čvor c primio od čvora n , a parametar $SH_{c,n}$ (*sent hello*) broj *Hello* paketa koje je čvor c poslao ka čvoru n . Kod nestabilnih linkova $RH_{c,n}$ je značajno manji od $SH_{c,n}$, tako da je koeficijent $LU_{c,n}$ blizu 1. Ovo utiče na povećanje $\alpha_{c,n}$, odnosno ažuriranje Q-vrednosti se ubrzava. Prema relaciji (3), uvek se bira maksimum između dva pomenuta kriterijuma ($RDC_{c,n}$ i $LU_{c,n}$). Takođe, osigurava se da $\alpha_{c,n}$ nije manje od 0,6, kako ažuriranje Q-vrednosti ne bi bilo previše sporo. Konačni cilj je ubrzati ažuriranje nepouzdatih putanja, dok se istovremeno održava stabilnost Q-vrednosti pouzdatih putanja.

Diskontni faktor ($\gamma_{c,n}$) se koristi za fino podešavanje Q-vrednosti. Za inicijalno ažuriranje Q-vrednosti, ovaj parametar će imati konstantnu vrednost od 0,3. Ova vrednost je takođe izabrana eksperimentalnim putem, nakon što je utvrđeno da daje najbolje mrežne performanse. U suprotnom, $\gamma_{c,n}$ se izračunava preko sledeće jednačine:

$$\gamma_{c,n} = 0,25 + 0,1 \cdot NB_n. \quad (6)$$

Prednost prilikom odabira sledećeg hopa daje se susednim čvorovima sa većim dostupnim propusnim opsegom čvora, putem parametra NB_n . Ovaj parametar se nalazi u *Hello* paketu koji šalje susedni čvor n i izračunava se kao:

$$NB_n = 1 - RUB_n, \quad (7)$$

gde RUB_n (*relative used bandwidth*) predstavlja relativni iskorišćeni propusni opseg susednog čvora n i izračunava se na sledeći način:

$$RUB_n = \frac{UsedBW_n}{MaxBW_n}. \quad (8)$$

U prethodnoj jednačini, $MaxBW_n$ predstavlja maksimalan propusni opseg koji je dostupan čvoru n (definisan karakteristikama bežičnog medijuma za prenos, isti je za sve čvorove), dok $UsedBW_n$ predstavlja iskorišćeni propusni opseg čvora n . Na ovaj način se izbegavaju putanje koje idu preko preopterećenih suseda i smanjuju se zagušenja u mreži.

Parametar $R_{c,n}$ predstavlja nagradu za preduzetu akciju, odnosno za odabir odgovarajućeg suseda za prosleđivanje paketa ka odredištu. Cilj je dati prioritet kraćim putanjama, dok se istovremeno izbegava kreiranje petlji. Ako čvor primi *Hello* paket direktno od odredišta, ovoj putanji svakako treba dati prioritet, tako da se Q-vrednost automatski postavlja na 1 (odnosno na maksimum). U drugim slučajevima, $R_{c,n}$ se definiše kao:

$$R_{c,n} = \begin{cases} 0,6, & 2 \text{ hopa;} \\ -0,5, & \text{petlja;} \\ 0, & \text{inače.} \end{cases} \quad (9)$$

Prvi slučaj se javlja kada je odredišni čvor ujedno i sledeći hop u Q-tabeli susednog čvora n . Takvoj putanji treba dati nagradu, s obzirom na to da ima samo dva

hopa do odredišta. Drugi slučaj se dešava kada je trenutni čvor c ujedno i sledeći hop u tabeli rutiranja susednog čvora n (što znači da će se stvoriti petlja). Takvu putanju treba kazniti, tj. njena Q -vrednost treba da se smanji. U ostalim slučajevima, nema ni nagrade ni kazne.

Poslednji član u jednačini (2), $\max_{y \in \text{Nei}(n)} Q_n(d, y)$, predstavlja maksimalnu Q -vrednost koju u svojoj Q -tabeli ima čvor n za putanju do odredišnog čvora d , preko nekog od svojih suseda y ($\text{Nei}(n)$ predstavlja skup suseda čvora n). Na ovaj način se u izbor optimalne putanje uključuje Q -vrednost drugog hopa od trenutnog čvora c , ka odredišnom čvoru d .

Druga metoda za ažuriranje Q -vrednosti dešava se po prijemu povratnih informacija sa MAC podsloja o gubitku paketa. Cilj je kazniti putanje na kojima dolazi do gubitaka, smanjujući time verovatnoću njihove selekcije pri slanju paketa ka odredištu. Ako čvor c primi informaciju sa MAC podsloja o gubitku paketa poslatih ka susedom čvoru n , ažurira Q -vrednosti za putanje do svih destinacija d_i , koje idu preko tog suseda prateći sledeću jednačinu:

$$Q_c(d_i, n) \leftarrow 0,6 \cdot Q_c(d_i, n). \quad (10)$$

Treća metoda za ažuriranje Q -vrednosti uključuje korišćenje paketa za ispitivanje putanje (*Route Probe Packets*, RPP) i potvrde o prijemu RPP paketa (*RPP Acknowledgment*, RPP-ACK). Iako čvor uvek šalje podatke preko suseda sa najvećom Q -vrednošću, moguće je da celokupna putanja do odredišta preko tog suseda nije optimalna. Stoga, čvor periodično šalje RPP pakete odredištu preko svih ostalih suseda čije Q -vrednosti nisu manje od 95% maksimalne Q -vrednosti. Oba tipa paketa sadrže sledeće informacije: vreme slanja RPP paketa, IP adrese odredišta, izvora i svih međučvorova na putanji od izvora do odredišta (svaki čvor koji primi RPP paket dodaje svoju IP adresu), kao i broj hopova (*Hop Count*, HC) na putanji od izvora do odredišta. Pored ovih zajedničkih polja, RPP-ACK sadrži i informaciju o relativnom propusnom opsegu putanje (*Route Relative Bandwidth*, RRB). Odredišni čvor po prijemu RPP paketa najpre očitava vrednosti svih polja i upisuje ih u RPP-ACK paket. Zatim dodaje RRB polje (inicijalizovano na 0) i šalje paket ka izvornom čvoru istom putanjom kojom je primio RPP. Po prijemu RPP-ACK paketa izvorni čvor računa faktor kvaliteta putanje (*Route Quality Factor*, RQF) na osnovu sledeće jednačine:

$$RQF = RTD \cdot HC \cdot RRB. \quad (11)$$

U ovoj jednačini RTD (*round trip delay*) predstavlja vreme proteklo od slanja RPP paketa do prijema RPP-ACK paketa. Vrednost HC je sadržana u RPP paketu i najpre se postavlja na 0, a onda se pri svakom prijemu ovog paketa (bilo od strane međučvorova ili odredišnog čvora) inkrementira za 1. Na kraju putanje, odredišni čvor očitava ovu vrednost i upisuje je u RPP-ACK paket. Vrednost RRB parametra je sadržana u RPP-ACK paketu, a ažuriraju je čvorovi na putanji od odredišnog do izvornog čvora. Svaki put kada neki međučvor primi RPP-ACK paket, ažurira RRB polje u paketu dodajući primljenoj vrednosti svoj relativni iskorišćeni propusni opseg (koji računa prema jednačini (8)). Nakon ovog ažuriranja, čvor prosleđuje RPP-ACK, a ovaj proces se nastavlja sve dok paket ne stigne do izvornog čvora.

RQF parametar se koristi za odabir putanja sa manjim kašnjenjem paketa, manjim brojem hopova i manje zauzetim propusnim opsegom. To znači da ovaj

parametar takođe treba da bude što niži. Nakon isteka vremenskog ograničenja za prijem RPP-ACK paketa, izvorni čvor određuje koji sused ima najniži RQF i nagrađuje putanju preko tog suseda prema sledećoj relaciji:

$$Q_c(d, n) \leftarrow 1,1 \cdot Q_c(d, n). \quad (12)$$

Putanje preko suseda kojima je poslat RPP, ali RPP-ACK nije primljen unutar definisanog perioda, kažnjavaju se prema relaciji:

$$Q_c(d, n) \leftarrow 0,6 \cdot Q_c(d, n). \quad (13)$$

Još jedan način za poboljšanje performansi VANET mreža je sprečavanje preopterećenja mreže u slučaju velike gustine čvorova. Zbog toga, Q-DRAV protokol osigurava da se *Hello* paketi šalju smanjenim intenzitetom u gustim mrežama. Prvi uslov za uzdržavanje od slanja *Hello* paketa susednom čvoru je da je *Hello* paket poslat tom susedu u prethodnom periodu slanja. Ovo sprečava uzastopno otkazivanje slanja *Hello* paketa i samim tim gubitak informacija o susedima. Drugi uslov je da je ukupan broj maksimalnih Q-vrednosti u paketu veći od definisanog praga (vrednost od 250 je usvojena u ovom protokolu, ali ju je moguće menjati u zavisnosti od veličine mreže). Smanjenje opterećenja mreže vrši se definisanjem verovatnoće slanja *Hello* paketa na sledeći način:

$$p = \begin{cases} (N_{th}/N_{av})^2, & N_{av} > N_{th}; \\ 1, & \text{inače.} \end{cases} \quad (14)$$

U ovoj relaciji N_{av} predstavlja prosečan broj suseda trenutnog čvora u prethodnom periodu (usvojen je period od 50 sekundi), a N_{th} je prag, tj. minimalni broj suseda za sprovođenje smanjenja opterećenja mreže (usvojen prag je 25 suseda). Oba parametra je moguće promeniti u zavisnosti od veličine i gustine mreže. *Hello* paket se šalje određenom susedu sa verovatnoćom p koja se smanjuje sa povećanjem gustine vozila, čime se smanjuje opterećenje mreže.

4. Simulaciona analiza

U ovom poglavlju je predstavljena simulaciona analiza primene novog Q-DRAV protokola rutiranja, koji je upoređen sa AODV, QLAODV i ARPRL protokolima. Za testiranje protokola izabran je NS-3 simulator, kao jedan od najpopularnijih simulatora za testiranje protokola za VANET mreže. U okviru ovog simulatora već je implementiran model AODV protokola. S druge strane, model novog Q-DRAV protokola, kao i modele QLAODV i ARPRL protokola, bilo je neophodno naknadno implementirati kako bi bilo moguće izvršiti njihovo testiranje i poređenje [9]. Implementacioni kôd novog Q-DRAV protokola u NS-3 simulatoru javno je dostupan [10].

4.1. Simulacioni parametri

Protokoli su testirani u dva različita scenarija, koji približno odgovaraju scenarijima u kojima su originalno testirani QLAODV i ARPRL protokoli u [6] i [7],

respektivno. Prvi scenario podrazumeva manju simulacionu oblast sa fiksnim brojem vozila, a cilj je uvideti kako se menjaju mrežne performanse pri variranju maksimalne dozvoljene brzine kretanja vozila. Dimenzije simulacionog područja su 1000 m x 1000 m, sa po 4 horizontalne i vertikalne ulice i po jednom kolovoznom trakom u svakom smeru. Unutar ovog područja kreće se 80 vozila od kojih 30 (nasumično izabranih) generiše *Constant Bit Rate* (CBR) saobraćaj brzine 32 kb/s. Ograničenje maksimalne dozvoljene brzine kretanja vozila varira od 1 m/s do 25 m/s.

Drugi scenario podrazumeva veću simulacionu oblast sa fiksnom maksimalnom dozvoljenom brzinom kretanja vozila, a cilj je uvideti kako se menjaju mrežne performanse pri variranju gustine vozila u mreži. Dimenzije simulacionog područja u ovom scenariju su 2000 m x 2000 m, sa po 5 horizontalnih i vertikalnih ulica. Svaka ulica ima po dve kolovozne trake za kretanje vozila u oba smera, a svaka raskrsnica ima semafor. Maksimalna dozvoljena brzina kretanja vozila je ograničena na 15 m/s. Ukupan broj vozila u mreži varira od 50 do 350 (sa korakom od 50 vozila), pri čemu uvek 20 nasumično izabranih vozila generiše CBR saobraćaj brzine 4 kb/s. Pregled najvažnijih parametara prvog i drugog simulacionog scenarija prikazani su u tabeli 1.

Tabela 1. Parametri prvog i drugog simulacionog scenarija

Parametar	Prvi scenario	Drugi scenario
Dimenzije prostora	1000 m x 1000 m	2000 m x 2000 m
Broj ulica	po 4 horizontalne i vertikalne	po 5 horizontalnih i vertikalnih
Ukupan broj vozila	80	50, 100, 150, 200, 250, 300, 350
Maksimalna dozvoljena brzina kretanja vozila	1 m/s, 5 m/s, 10 m/s, 15 m/s, 20 m/s, 25 m/s	15 m/s
Aplikacioni protok po vozilu	32 kb/s	4 kb/s
Broj vozila koja generišu saobraćaj	30	20

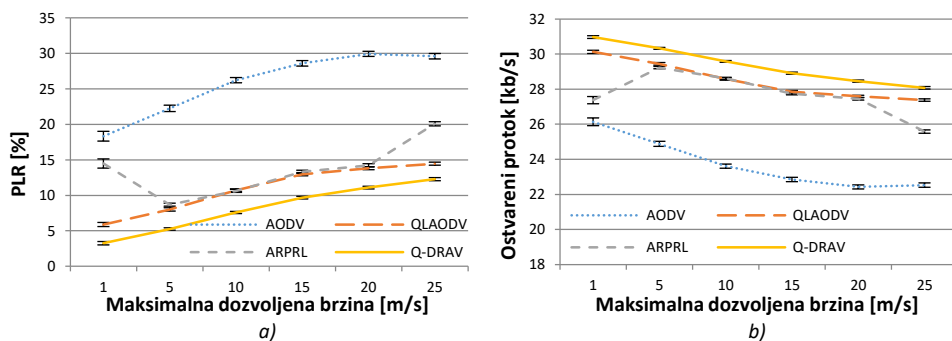
Ostali simulacioni parametri su identični za oba scenarija. Za kreiranje scenarija korišćen je *Manhattan Grid* model mobilnosti i SUMO simulacioni alat, iz kog je model kretanja vozila ekstraktovan i iskorišćen za simulacije u NS-3 simulatoru. Na MAC podsloju se koristi IEEE 802.11p standard za bežične mreže, sa širinom kanala od 10 MHz i protokom od 6 Mb/s. Za model propagacije odabran je *Two Ray Ground*, a na transportnom sloju je korišćen UDP protokol. Veličina paketa na aplikacionom sloju je 512 bajta. Trajanje jedne simulacije je ograničeno na 600 sekundi.

4.2. Rezultati simulacija

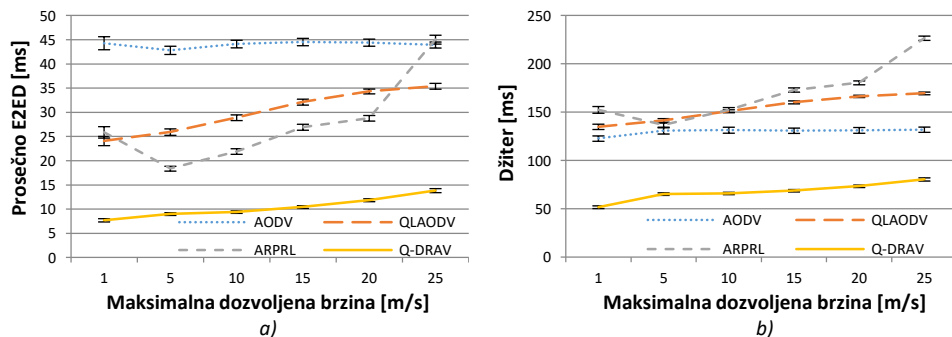
Poređenje protokola izvršeno je posmatrajući PLR, ostvareni aplikacioni protok, E2ED i džiter. Zbog verodostojnosti rezultata, za svaki parametar je izračunata prosečna vrednost na osnovu 100 rezultata simulacija, sprovedenih za različite vrednosti generatora slučajnih brojeva. Pored toga, određen je i 95% interval poverenja za svaku izračunatu vrednost, kako bi se potvrdila pouzdanost rezultata simulacija.

Slike 1a i 1b prikazuju zavisnost PLR i ostvarenog protoka na aplikacionom sloju od maksimalne dozvoljene brzine kretanja vozila u prvom scenariju. Može se zaključiti da QLAODV i ARPRL imaju značajno niže gubitke paketa i veći ostvareni

aplikacioni protokol u poređenju sa AODV protokolom. Primetno je ipak da ARPRL pokazuje priličnu degradaciju ovih parametara pri krajnje malim i velikim maksimalnim dozvoljenim brzinama kretanja vozila. S druge strane, novi Q-DRAV protokol značajno nadmašuje QLAODV i ARPRL po ovim parametrima. Slike 2a i 2b prikazuju zavisnost prosečnog E2ED i džitera od maksimalne dozvoljene brzine kretanja vozila u prvom scenariju. Očigledno je da Q-DRAV pruža ubedljivo najbolje performanse po oba posmatrana parametra. Primetno je da i po ovim parametrima ARPRL pokazuje lošije rezultate pri krajnje malim i velikim maksimalnim dozvoljenim brzinama kretanja vozila.

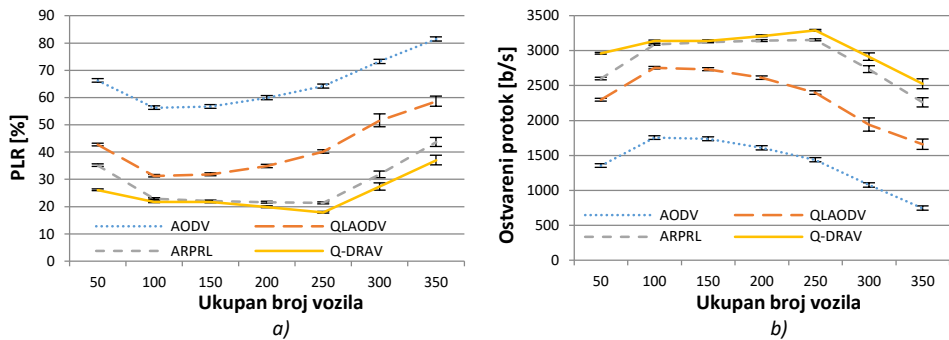


Slika 1. Zavisnost PLR (a) i ostvarenog protoka (b) od maksimalne dozvoljene brzine kretanja vozila

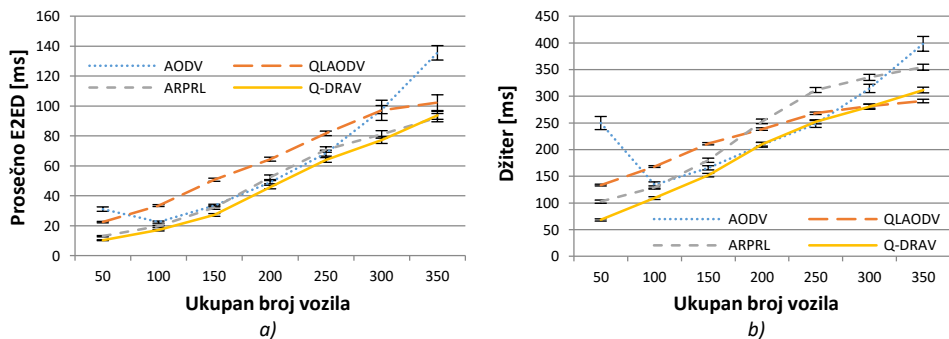


Slika 2. Zavisnost E2ED (a) i džitera (a) od maksimalne dozvoljene brzine kretanja vozila

Slike 3a i 3b prikazuju zavisnost PLR i ostvarenog protoka na aplikacionom sloju od ukupnog broja vozila u drugom scenariju. Primetno je da ARPRL daje značajno bolje rezultate od AODV i QLAODV protokola, ali ga ipak Q-DRAV uvek nadmašuje, posebno u veoma retkim i veoma gustim mrežama. Slike 4a i 4b prikazuju zavisnost prosečnog E2ED i džitera od ukupnog broja vozila u mreži u drugom scenariju. Prema ovim parametrima, Q-DRAV gotovo uvek pokazuje najbolje performanse, osim u slučaju veoma gustih mreža gde ARPRL daje gotovo identično kašnjenje, a QLAODV nešto manji džiter.



Slika 3. Zavisnost PLR (a) i ostvarenog protoka (b) od ukupnog broja vozila u mreži



Slika 4. Zavisnost E2ED (a) i džitera (b) od ukupnog broja vozila u mreži

5. Zaključak

Razvoj inteligentnih transportnih sistema predstavlja jedan od glavnih pravaca unapređenja savremenih saobraćajnih mreža. Bitan segment ovih sistema mogu biti VANET mreže, tako da je od velikog interesa dalje unaprediti njihove performanse. Ove performanse u velikoj meri zavise od protokola rutiranja koji se koristi za izbor optimalne putanje za razmenu podataka između čvorova u mreži, tako da postoji potreba za njihovim stalnim unapređenjem. U ovom radu je pokazano kako se unapređenjem procesa rutiranja pomoću RL mogu značajno unaprediti mrežne performanse VANET mreža, ukoliko se u RL proces uključe adekvatni uticajni faktori. Dobre mrežne performanse novog Q-DRAV protokola pokazuju da može biti adekvatna zamena ranijim protokolima rutiranja u urbanim saobraćajnim uslovima.

Literatura

- [1] C. Perkins, E. Belding-Royer, and Ss Das, "Ad Hoc On demand Distance Vector (AODV) routing", RFC 3561, IETF, October 2003.
- [2] D. R. De Assis, E. C. G. Wille, and J. J. Alves, "New results on the IC_AOMDV protocol for vehicular ad hoc networks in urban areas", Advances in Electrical & Computer Engineering, vol. 23, pp. 21-28, 2023. DOI: 10.4316/AECE.2023.03003

- [3] M. Malnar and N. Jevtic, "An improvement of AODV protocol for the overhead reduction in scalable dynamic wireless ad hoc networks", *Wireless Networks*, vol. 28, pp. 1039-1051, April 2022. DOI: 10.1007/s11276-022-02890-5
- [4] F. Mubarek, S. Aliesawi, K. Ali Alheeti, and N. Alfahad, "Urban-AODV: an improved AODV protocol for vehicular ad-hoc networks in urban environment", *International Journal of Engineering & Technology*, vol. 7, pp. 3030-3036, 2018. DOI: 10.14419/ijet.v7i4.15031
- [5] T. K. Saini and S. C. Sharma, "Recent advancements, review analysis, and extensions of the AODV with the illustration of the applied concept", *Ad Hoc Networks*, vol. 103, pp. 102-148, June 2020. DOI: 10.1016/j.adhoc.2020.102148
- [6] C. Wu, K. Kumekawa, and T. Kato, "Distributed reinforcement learning approach for vehicular ad hoc networks", *IEICE Transactions on Communications*, vol. 93, pp. 1431-1442, June 2010. DOI: 10.1587/transcom.E93.B.1431
- [7] J. Wu, M. Fang, and X. Li, "Reinforcement learning based mobility adaptive routing for vehicular ad-hoc networks", *Wireless Personal Communications*, vol. 101, pp. 2143-2171, May 2018. DOI: 10.1007/s11277-018-5809-z
- [8] P. Bugarčić, N. Jevtić, M. Malnar, and M. Stojanović, "Enhanced QL-based dynamic routing protocol for urban VANETs", *Advances in Electrical & Computer Engineering*, vol. 24, pp. 27-36, November 2024. DOI: 10.4316/AECE.2024.04003
- [9] N. Jevtić and P. Bugarčić, "Comparison of traditional and reinforcement learning based routing protocols in VANET scenario", *Int. Journal for Traffic and Transport Engineering*, vol. 13, pp. 125-137, February 2023. DOI: 10.7708/ijtte2023.13(1).10
- [10] P. Bugarčić, "Code for Q-DRAV routing protocol", 2024, *gitHub*, [Online]. Available at: <https://github.com/pavlebugarcic/qdrav>

Abstract: *The selection of an optimal route for data transmission in wireless ad hoc networks is crucial for achieving satisfactory network performance. This process becomes highly complex when the network topology changes rapidly, as is the case with Vehicular Ad-hoc Networks (VANETs). Under such conditions, the routing process can be significantly improved by incorporating machine learning into the optimal route selection, with the most suitable learning type for highly dynamic networks being reinforcement learning (RL). One of the most prominent RL methods for dynamic wireless ad hoc networks is Q-Learning (QL). In this paper, a new QL-based dynamic routing algorithm for urban VANETs (Q-DRAV) is proposed. By integrating relevant network parameters into the RL process, the proposed approach achieves a significant improvement in overall network performance for VANETs. Simulation analysis and comparison with other routing protocols were conducted in the NS-3 simulator.*

Keywords: *reinforcement learning, Q-learning, routing protocols, VANET, NS-3*

A NEW DYNAMIC ROUTING PROTOCOL FOR URBAN VANETs BASED ON REINFORCEMENT LEARNING

Pavle Bugarčić, Marija Malnar