

<https://doi.org/10.37528/FTTE/9788673954165/POSTEL.2022.039>

## ANALIZA PROTOKOLA RUTIRANJA BAZIRANIH NA UČENJU POTKREPLJIVANJEM ZA VANET MREŽE

Nenad Jevtić, Pavle Bugarčić  
Univerzitet u Beogradu - Saobraćajni fakultet  
n.jevtic@sf.bg.ac.rs, p.bugarcic@sf.bg.ac.rs

**Rezime:** *Bežične ad hoc mreže za vozila (Vehicular Ad hoc Networks, VANETs) karakteriše velika mobilnost čvorova i česte promene u mrežnoj topologiji, što značajno otežava proces rutiranja paketa. Pokazalo se da tradicionalni protokoli rutiranja nisu u stanju da blagovremeno isprate ove promene. Iz tog razloga su razvijeni protokoli bazirani na učenju potkrepljivanjem, koji omogućavaju stalno praćenje promena u mrežnom okruženju i prilagođavanje procesa rutiranja tim promenama. U ovom radu je izvršena analiza i poređenje primene ove dve vrste protokola, sa ciljem da se dokaže svrsishodnost pristupa baziranog na učenju potkrepljivanjem kod VANET mreža. Za predstavnika tradicionalnih protokola izabran je AODV (Ad hoc On-Demand Distance Vector) protokol, dok je za predstavnika protokola baziranih na učenju potkrepljivanjem izabran ARPRL (Adaptive Routing Protocol based on Reinforcement Learning). Rezultati simulacija su pokazali da ARPRL protokol pokazuje značajno bolje mrežne performanse u pogledu procenta izgubljenih paketa i kašnjenja paketa s kraja na kraj.*

**Ključne reči:** *protokoli rutiranja, učenje potkrepljivanjem, VANET*

### 1. Uvod

Sa razvojem pametnih gradova i inteligentnih transportnih sistema, bežične *ad hoc* mreže za vozila (*Vehicular Ad hoc Networks*, VANETs) postaju sve značajnije. VANET mreže predstavljaju posebnu kategoriju bežičnih *ad hoc* mreža (*Wireless Ad hoc Networks*, WANETs), kod kojih mrežu čvorova čini skup vozila koja međusobno komuniciraju putem bežičnih *ad hoc* linkova. Karakterišu ih česte promene u mrežnoj topologiji i velika brzina kretanja čvorova, koji iz tog razloga mogu veoma brzo izaći iz dometa čvorova koji šalju podatke. Promene u mrežnoj topologiji prouzrokuju česte prekide linkova između čvorova u mreži, što dovodi do degradacije različitih mrežnih performansi, kao što su kašnjenje s kraja na kraj (*end-to-end delay*, E2ED), protok (*throughput*), stepen uspešno isporučenih paketa (*packet delivery ratio*, PDR), itd.

Kako bi se prevazišli ovi problemi i unapredio proces rutiranja u VANET mrežama, neki autori predlažu korišćenje veštačke inteligencije prilikom izbora optimalne putanje za slanje podataka [1-3]. Najčešće korišćena oblast veštačke

inteligencije u protokolima rutiranja za VANET mreže je mašinsko učenje (*Machine Learning*, ML), a tip ML koji daje najperspektivnije rezultate je učenje potkrepljivanjem (*Reinforcement Learning*, RL) koje je detaljnije opisano u [4]. Glavna karakteristika ovog tipa učenja je stalna interakcija agenta učenja sa okruženjem, što omogućava praćenje i prilagođavanje promenama u okruženju. Zbog toga je učenje potkrepljivanjem posebno pogodno za primenu u mrežama čija se topologija često menja.

U ovom radu će biti izvršena analiza i poređenje rezultata primene tradicionalnih protokola rutiranja i protokola rutiranja baziranih na učenju potkrepljivanjem u VANET mrežama. Za predstavnika tradicionalnih protokola izabran je AODV (*Ad hoc On-Demand Distance Vector*) protokol [5], kao jedan od najpopularnijih reaktivnih protokola rutiranja za WANET mreže. Sa druge strane, za predstavnika protokola rutiranja baziranih na učenju potkrepljivanjem izabran je ARPRL (*Adaptive Routing Protocol based on Reinforcement Learning*) protokol [3]. Indikatori mrežnih performansi na osnovu kojih je izvršeno poređenje primene ovih protokola su procenat izgubljenih korisničkih paketa i prosečno kašnjenje paketa sa kraja na kraj mreže.

Rad je organizovan na sledeći način. U drugom poglavlju su opisani principi funkcionisanja AODV i ARPRL protokola rutiranja. U trećem poglavlju je izvršeno poređenje i analiza mrežnih performansi dobijenih prilikom korišćenja ova dva protokola. U četvrtom poglavlju su data zaključna razmatranja.

## 2. Protokoli rutiranja za VANET mreže

Protokoli rutiranja odgovorni su za određivanje i održavanje optimalne putanje za prosleđivanje podataka u mreži. Kod VANET mreža svi čvorovi učestvuju u rutiranju podataka. Glavni izazov sa kojim se susreću protokoli rutiranja u ovim mrežama je velika mobilnost mrežnih čvorova, odnosno česte promene mrežne topologije. U ovakvim uslovima može doći do kasnog otkrivanja prekida linka na putanji koja se koristi za slanje podataka, što za posledicu ima gubitak paketa i smanjenje protoka u mreži. Iz tog razloga su neophodni novi pristupi u rutiranju, koji podrazumevaju brže otkrivanje promena u mreži i blagovremeni izbor nove putanje u slučaju prekida linkova na staroj putanji. U ovom poglavlju će biti prikazan, sa jedne strane, tradicionalni pristup rutiranju podataka, a sa druge strane noviji pristup koji podrazumeva uključivanje veštačke inteligencije u proces rutiranja. Najpre će biti opisan princip funkcionisanja AODV protokola, kao jednog od najpoznatijih tradicionalnih reaktivnih protokola rutiranja, a zatim i princip funkcionisanja ARPRL protokola koji spada u protokole rutiranja baziranje na učenju potkrepljivanjem.

### 2.1 AODV protokol

AODV protokol omogućava dinamično, samopokretno, *multihop* rutiranje između mobilnih čvorova koji žele da uspostave i održavaju komunikaciju u *ad hoc* mreži. Protokol se sastoji iz dva glavna mehanizma: „Otkrivanje putanje” (*Route Discovery*) i „Održavanje putanje” (*Route Maintenance*), koji zajedno rade kako bi čvorovi mogli da otkrivaju i održavaju putanje za proizvoljna odredišta u mreži.

Kada se javi potreba za slanjem podataka, AODV protokol aktivira mehanizam otkrivanja putanje uz korišćenje RREQ (*Route Request*) i RREP (*Route Replay*) kontrolnih paketa. Čvor koji želi da pošalje pakete podataka najpre šalje difuzne RREQ

pakete sa ciljem da pronade najbolju putanju do odredišta. RREQ paket može primiti neki međučvor ili odredišni čvor. Ukoliko je paket primio međučvor, najpre će ažurirati putanju do izvornog čvora u svojoj tabeli rutiranja, a zatim proveriti da li ima odgovarajuću putanju do željenog odredišta i ako ima poslati je izvornom čvoru putem RREP paketa. U suprotnom difuzno će reemitovati RREQ paket dalje u mrežu. Kada odredišni čvor primi RREQ paket, on prvo kreira (ili ažurira) svoju tabelu rutiranja podatkom o putanji do izvornog čvora, pa šalje RREP paket ka izvornom čvoru istom putanjom kojom je primio RREQ paket. Kada primi RREP paket, izvorni čvor počinje sa slanjem paketa podataka ka odredištu istom putanjom kojom je primio RREP paket. Ukoliko primi više RREP paketa, izvorni čvor bira putanju sa najmanjim brojem hopova do odredišta.

Mehanizam održavanja putanje se odvija pomoću *Hello* kontrolnih paketa. Čvorovi periodično šalju *Hello* pakete ka svojim susedima kako bi obavestili ostale čvorove da li su aktivni. Ukoliko čvor ne primi *Hello* paket u unapred definisanom vremenskom intervalu od nekog od ranijih suseda, smatraće da je link ka tom susedu u prekidu. U slučaju prekida linka čvor slanjem RERR (*Route Error*) paketa obaveštava sve svoje susede koji su koristili taj link da više nije dostupan. Svaki čvor ima svoju tabelu rutiranja u koju su smešteni podaci o rutama memorisanim u skladu sa unapred definisanim vremenom (*Delete Period*). Nakon isteka ovog vremena, podaci iz tabela rutiranja u čvorovima se brišu.

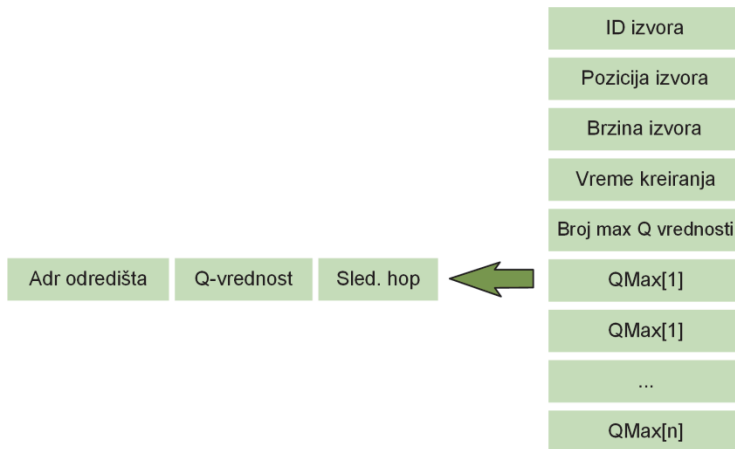
## 2.2 ARPRL protokol

Sa ciljem unapređenja performansi mreže koje degradiraju zbog brzih promena u mrežnoj topologiji koje AODV protokol ne može uspešno da prati, u ovom poglavlju opisan je ARPRL protokol koji se bazira na učenju potkrepljivanjem. Konkretno, u ovom protokolu je korišćen algoritam Q-učenja (*Q-learning*) [4] koji predstavlja jedan od najčešće korišćenih tipova učenja potkrepljivanjem.

Za rutiranje podataka svaki čvor koristi svoju Q-tabelu, koja se sastoji od Q-vrednosti koje se ažuriraju razmenom *Hello* paketa između susednih čvorova, prijemom paketa podataka i pomoću povratnih informacija sa MAC (*Medium Access Control*) sloja. Pored Q-tabele svaki čvor održava tabelu suseda, kako bi u svakom trenutku znao skup dostupnih susednih čvorova. Ova tabela se ažurira samo razmenom *Hello* paketa. Kada čvor želi da pošalje pakete podataka ka odredištu, on prvo proverava u svojoj Q-tabeli da li ima sledeći hop ka tom odredištu. Ukoliko nema, inicira proces pronalaženja putanje korišćenjem LPREQ (*Learning Probe Request*) i LPREP (*Learning Probe Reply*) kontrolnih paketa, koji je veoma sličan procesu otkrivanja putanje kod AODV protokola.

Svako vozilo u mreži se ponaša kao agent učenja i neprekidno prikuplja informacije o stanju linkova u mreži kroz interakciju sa svojim susedima periodičnim slanjem *Hello* paketa. Struktura *Hello* paketa prikazana je na slici 1. Svaki *Hello* paket sadrži sledeća polja: ID, poziciju i brzinu izvornog čvora, vreme kreiranja *Hello* paketa, broj maksimalnih Q-vrednosti sadržanih u *Hello* paketu, na šta se nadovezuje niz maksimalnih Q-vrednosti. Ukoliko čvor ima više potencijalnih putanja do određenog odredišta, izabraće onu sa najvećom Q-vrednošću koja predstavlja maksimalnu Q-vrednost za to odredište. Ponavljanjem ovog procesa za sve odredišne čvorove u mreži kreira se niz maksimalnih Q-vrednosti. Svaka maksimalna Q-vrednost sadrži adresu

odredišnog čvora, odgovarajuću Q-vrednost i adresu sledećeg hopa na putanji ka odredištu.



Slika 1. Struktura Hello paketa

Svaki čvor koji ima podatke za slanje ka određenom odredištu bira sledeći hop sa najvećom Q-vrednošću, kako bi izabrao najbolju putanju za slanje podataka. Zato je jako bitno redovno ažuriranje ovih vrednosti, zbog dinamičke prirode VANET mreža. Prvi način ažuriranja Q-vrednosti je preko kontrolnih paketa. Kada primi *Hello* paket, svako vozilo ažurira svoju Q-vrednost na osnovu informacija iz *Hello* paketa. Ako čvor ne primi *Hello* paket od određenog suseda u unapred definisanom intervalu vremena, Q-vrednost za tog suseda biće resetovana na 0. Nakon prijema *Hello* paketa od svog suseda  $n$ , posmatrano vozilo  $c$  će ažurirati svoju Q-tabelu prema sledećoj formuli:

$$Q_c(d, n) = (1 - \alpha_{c,n}) \square Q_c(d, n) + \alpha_{c,n} [R_{c,n} + \gamma_{c,n} \square \max_{y \in Nei(n)} Q_n(d, y)] \quad (1)$$

gde je  $d$  odredišno vozilo,  $Nei(n)$  predstavlja skup suseda čvora  $n$ , a  $\alpha_{c,n}$ ,  $\gamma_{c,n}$  i  $R_{c,n}$  predstavljaju stopu učenja, diskontni faktor i nagradu respektivno, i definišu se na sledeći način:

$$\alpha_{c,n} = \max \left( 0, 2, \frac{|v_c| - |v_n|}{v_{max} - v_{min}} \right) \quad (2)$$

gde  $v_c$  i  $v_n$  predstavljaju brzine kretanja čvorova  $c$  i  $n$ , a  $v_{max}$  i  $v_{min}$  maksimalnu i minimalnu brzinu u mreži;

$$\gamma_{c,n} = \begin{cases} \frac{\sum_{i=1}^N R_{c,n}}{N}, & N \neq 0 \\ 0, & N = 0 \end{cases} \quad (3)$$

gde je  $N$  ukupan broj vozila u mreži;

$$R_{c,n} = C + HMRR_{c,n} + LET_{c,n} \quad (4)$$

gde je  $C$  konstanta i ima vrednost 100, a  $HMRR_{c,n}$  i  $LET_{c,n}$  predstavljaju stepen uspešno primljenih *Hello* poruka i faktor stabilnosti linka respektivno, a računaju se na osnovu sledećih formula:

$$HMRR_{c,n} = \begin{cases} 100 \square \frac{CNT_r(c,n)}{CNT_s(n)}, & CNT_s(n) \geq 15 \\ 100 \square \frac{CNT_r(c,n)}{CNT_s(n)} \square \left(1 - \left(\frac{1}{2}\right)^{CNT_s(n)}\right), & \text{inače} \end{cases} \quad (5)$$

gde je  $CNT_r(c,n)$  broj *Hello* poruka koje je čvor  $c$  primio od čvora  $n$ , a  $CNT_s(n)$  broj *Hello* poruka koje je čvor  $c$  poslao ka čvoru  $n$ ;

$$LET_{c,n} = \begin{cases} 100, & A = 0 \text{ i } B = 0 \\ \min\left(100, \frac{-(AB+CD) + \sqrt{(A^2+C^2) \square R^2 - (AD-BC)^2}}{A^2+B^2}\right), & \text{inače} \end{cases} \quad (6)$$

gde se  $A$ ,  $B$ ,  $C$  i  $D$  računaju preko sledećih formula:

$$A = v_c \cos(\Theta_{v_c}) - v_n \cos(\Theta_{v_n}) \quad (7)$$

gde  $v_c \cos(\Theta_{v_c})$  i  $v_n \cos(\Theta_{v_n})$  predstavljaju projekcije brzina čvorova  $c$  i  $n$  na  $x$  osu respektivno;

$$B = x_c - x_n \quad (8)$$

gde  $x_c$  i  $x_n$  predstavljaju  $x$  koordinate čvorova  $c$  i  $n$  respektivno;

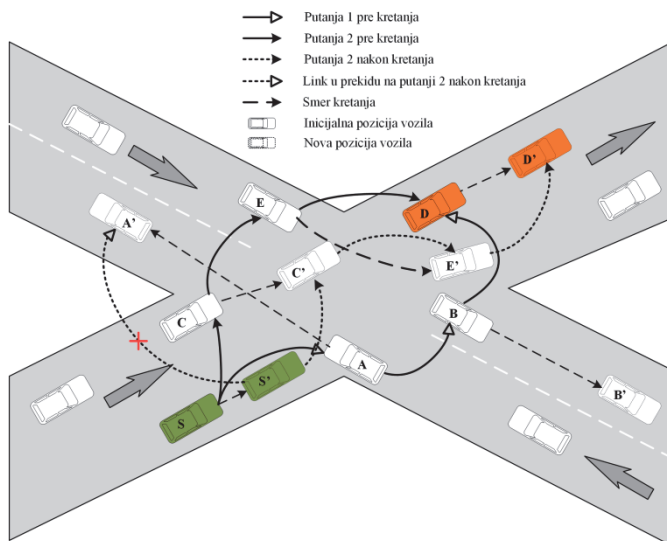
$$C = v_c \sin(\Theta_{v_c}) - v_n \sin(\Theta_{v_n}) \quad (9)$$

gde  $v_c \sin(\Theta_{v_c})$  i  $v_n \sin(\Theta_{v_n})$  predstavljaju projekcije brzina čvorova  $c$  i  $n$  na  $y$  osu respektivno;

$$D = y_c - y_n \quad (10)$$

gde  $y_c$  i  $y_n$  predstavljaju  $y$  koordinate čvorova  $c$  i  $n$  respektivno;

Očigledno je da će čvorovi sa većim koeficijentom  $LET_{c,n}$  imati veće  $Q$ -vrednosti, što znači da se prednost pri odabiru sledećeg hopa daje stabilnijim putanjama. Kako bi se ovo demonstriralo na slici 2 je predstavljen jedan jednostavan scenario u VANET mreži. Komunikacija između vozila S i D je moguća preko dve potencijalne putanje: jedna je preko vozila A ( $S \rightarrow A \rightarrow B \rightarrow D$ ) i druga preko vozila C ( $S \rightarrow C \rightarrow E \rightarrow D$ ). Pošto se vozilo A sve više udaljava od vozila S, dok vozilo C nastavlja pravo kao i vozilo S, prva putanja će izgubiti konekciju nakon nekog vremena zbog prekida linka ( $S' \rightarrow A'$ ), a putanja preko vozila C ostaje validna. Shodno tome, susedno vozilo C je pogodnije da bude izabrano za sledeći hop na putanji od vozila S do vozila D.



Slika 2. VANET scenario koji pokazuje ulogu LET u rutiranju podataka

Drugi način ažuriranja Q-vrednosti je nakon prijema paketa podataka. Kada primi pakete podataka poslate od izvornog čvora  $s$  preko susjednog čvora  $n$ , čvor  $c$  ažurira Q-vrednosti preko formule:

$$Q_c(s, n) = (1 - \alpha_{c,n}) \square Q_c(s, n) + \alpha_{c,n} [R_{c,n} + \gamma_{c,n} \square \max_{y \in Nei(n)} Q_n(s, y)] \quad (11)$$

Treći vid ažuriranja Q-vrednosti je na osnovu povratne informacije sa MAC sloja. Po prijemu obaveštenja sa MAC sloja o gubitku paketa od susjednog čvora  $n$ , čvor  $c$  za svaku destinaciju  $d_i$  ažurira Q-vrednosti preko sledeće formule:

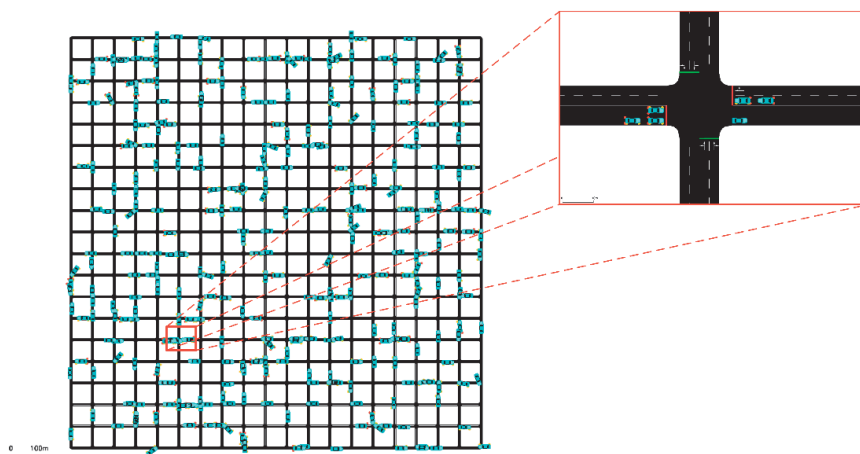
$$Q_c(d_i, n) = 0,5 \square Q_c(d_i, n) \quad (12)$$

što znači da će Q-vrednosti za putanje preko susjednog čvora  $n$  biti sve manje nakon svakog prijema obaveštenja o gubitku paketa na MAC sloju.

### 3. Analiza simulacionih rezultata

Simulacija je izvršena korišćenjem NS-3 (*Network Simulator 3*) mrežnog simulatora [6] i dobro poznatog SUMO (*Simulation of Urban Mobility*) simulatora mobilnosti vozila [7]. Simulacioni scenario koristi uobičajeni *Manhattan Grid* model mobilnosti sa 20 horizontalnih i 20 vertikalnih ulica u polju od 2000 m x 2000 m. Svaka ulica ima po 2 trake za kretanje u oba smera. Kako bi scenario bio što je moguće realniji, svaka peta raskrsnica po horizontalnoj i vertikalnoj osi sadrži semafor. SUMO simulator generiše kretanje vozila u definisanom prostoru, pri čemu se vozila kreću ulicama uz ograničenje brzine od 15 m/s. Protokoli rutiranja su testirani varijacijom gustine saobraćaja, odnosno povećanjem broja vozila od 50 do 250, sa korakom od 50 vozila.

Opisani simulacioni scenario prikazan je na slici 3, gde je kao detalj izdvojena jedna od raskrsnica sa semaforom.



Slika 3. Simulacioni scenario: 2000 m x 2000 m, 20 ulica po svakoj osi, svaka peta raskrsnica sa semaforom

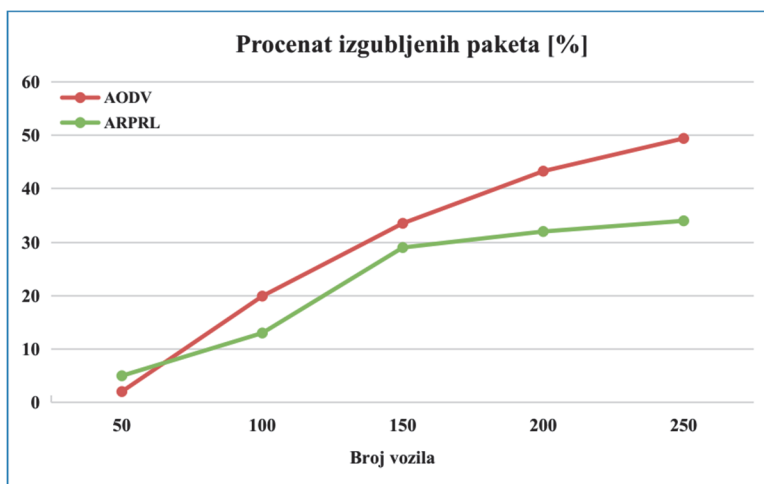
Za komunikacione parametre simulacije pretpostavljene su sledeće vrednosti. Koristi se IEEE 802.11p standard za bežične mreže, sa propusnim opsegom od 10 MHz i protokom od 6 Mb/s. Za model propagacije odabran je *Two Ray Ground*. Testrani su i upoređeni protokoli rutiranja AODV i prethodno opisani ARPRL protokol koji predstavlja modifikaciju AODV protokola baziranu na učenju potkrepljivanjem. Korišćen je UDP (*User Datagram Protocol*) protokol na transportnom sloju. Aplikacioni sloj generiše pakete veličine 512 B koristeći CBR (*Constant Bit Rate*) saobraćaj, pri čemu je ostvareni aplikacioni protok 4 kb/s. Aplikacioni paketski saobraćaj generiše deset nasumično izabranih vozila, dok drugih deset vozila, takođe nasumično odabranih, primaju generisane pakete. Ostali parametri imaju podrazumevana podešavanja za mrežni simulator. Pregled najvažnijih parametara simulacije prikazan je u tabeli 1.

Tabela 1. Pregled parametara simulacije

Parametar	Vrednost
Dimenzije prostora	2000 m x 2000 m
Ulice	Po 20 horizontalnih i vertikalnih ulica
Vreme simulacije	600 s
Broj vozila	50, 100, 150, 200, 250
Maksimalna brzina vozila	15 m/s
Model mobilnosti	<i>Manhattan Grid</i>
Propagacioni model	<i>Two Ray Ground</i>
MAC	IEEE 802.11p, širina kanala 10 MHz, protok 6 Mb/s
Protokol rutiranja	AODV, ARPRL
Transportni protokol	UDP
Aplikacioni protok	4 kb/s (CBR, 10 vozila)
Veličina paketa	512 B

Posmatrana su dva indikatora performansi mreže: procenat izgubljenih korisničkih paketa i prosečno kašnjenje paketa sa kraja na kraj mreže. Urađeno je 200 iteracija simulacija sa različitim podešavanjima za generator slučajnih brojeva i nakon toga izračunata je srednja vrednost za oba indikatora.

Gustina vozila u VANET mrežama ima značajan uticaj na performanse protokola. U ovom radu analiza je izvršena uz fiksiranu maksimalnu brzinu vozila od 15 m/s, a varijacijom broja vozila od 50 do 250 povećava se gustina saobraćaja. Rezultati simulacija prikazani su na slikama 4 i 5.



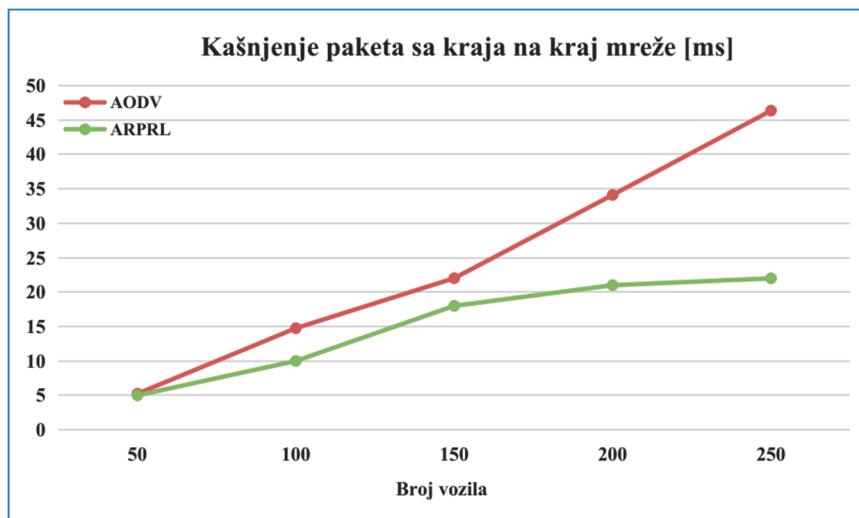
Slika 4. Zavisnost procenta izgubljenih korisničkih paketa od gustine vozila

Slika 4 prikazuje prosečan procenat gubitka korisničkih paketa za oba protokola rutiranja u zavisnosti od broja vozila u simulacionom okruženju. Sa porastom gustine vozila verovatnoća nalaženja kvalitetne putanje raste, te se u izvesnim situacijama može očekivati i poboljšanje rezultata, odnosno smanjenje vrednosti za procenat izgubljenih paketa. Međutim, u razmatranom scenariju, zbog velike dinamike vozila, topologija mreže se brzo menja i dolazi do čestih prekida linkova i slanja RERR paketa, kao i do potrebe za ponovnim slanjem RREQ paketa. Često ponovno otkrivanje putanje dovodi do plavljenja (*flooding*) i degradacije mreže u pogledu gubitka korisničkih paketa zbog povećanog broja kolizija. Ova pojava posebno dolazi do izražaja kada je broj RREQ paketa veliki, odnosno kada u mreži ima puno vozila. Stoga se na slici 4 primećuje jasan trend rasta procenta gubitka paketa za AODV protokol sa povećanjem gustine vozila. Međutim, ARPRL protokol koristi proaktivno Q-učenje koje mu omogućava izbor boljih i stabilnijih putanja, kao i bržu promenu putanje ukoliko dođe do prekida nekog od linkova na toj putanji. Zbog toga ARPRL protokol pokazuje značajno manju degradaciju performansi u pogledu gubitka paketa.

Slika 5 prikazuje prosečno kašnjenje paketa sa kraja na kraj mreže za svaki protokol rutiranja u zavisnosti od broja vozila u simulacionom okruženju. I kod ovog indikatora performanse mreže rezultati pokazuju značajnu degradaciju kod AODV protokola iz gore opisanih razloga. S druge strane, ARPRL protokol ne pokazuje



značajniju degradaciju pri povećanju broja vozila u mreži zadržavajući prosečna kašnjenja paketa u prihvatljivim granicama.



Slika 5. Zavisnost ostvarenog kašnjenja paketa sa kraja na kraj mreže od gustine vozila

#### 4. Zaključak

U ovom radu je izvršena analiza i poređenje rezultata primene AODV i ARPRL protokola rutiranja u VANET mrežama. AODV protokol predstavlja tradicionalni reaktivni protokol rutiranja, što znači da se tek u slučaju javljanja potrebe za slanjem paketa pokrene procedura istraživanja putanje. Očigledno je da AODV protokol nije u stanju da blagovremeno isprati česte promene topologije u VANET mrežama, što rezultuje visokim procentom izgubljenih paketa i velikim kašnjenjem paketa u mreži. Ovo posebno dolazi do izražaja sa povećanjem broja mrežnih čvorova, gde je zbog čestih prekida mrežnih linkova neophodno iznova započinjati proceduru pronalaženja putanje, što dovodi do preopterećenja mreže i degradiranja mrežnih performansi. S druge strane, ARPRL protokol pronalazi optimalnu putanju uz pomoć učenja potkrepljivanjem, koje omogućava stalno praćenje i prilagođavanje promenama u mrežnoj topologiji. Ovo rezultuje značajno boljim mrežnim performansama u pogledu procenta izgubljenih paketa i kašnjenja paketa sa kraja na kraj mreže, što posebno dolazi do izražaja sa povećanjem broja vozila u mreži.

U okviru budućih istraživanja planirana je dalja analiza i testiranje protokola rutiranja baziranih na učenju potkrepljivanjem za VANET mreže, kao i razvoj novog protokola koji bi se još bolje prilagodio dinamičkoj prirodi VANET mreža. Takođe, istraživanje se može proširiti i na FANET (*Flying Ad hoc Networks*) mreže koje postaju sve atraktivnije sa ubrzanim razvojem bespilotnih letelica. Osim toga moguće je, pored opisanog algoritma Q-učenja, testirati primenu unapređenih algoritama učenja potkrepljivanjem, kao što su duboko učenje potkrepljivanjem (*deep reinforcement learning*) i duelno duboko učenje potkrepljivanjem (*dueling deep reinforcement learning*).

## Literatura

- [1] 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
- [2] C. Wu, S. Ohzahata, and T. Kato, "Flexible, Portable, and Practicable Solution for Routing in VANETs A Fuzzy Constraint Q-Learning Approach", *IEEE Transactions on Vehicular Technology*, vol. 62, pp. 4251-4263, July 2013. DOI: 10.1109/TVT.2013.2273945
- [3] 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
- [4] N. Jevtić, M. Malnar, and P. Bugarčić. "Primena učenja potkrepljivanjem u protokolima rutiranja za dinamičke bežične *ad hoc* mreže", in *XXXIX Simpozijum o novim tehnologijama u poštanskom i telekomunikacionom saobraćaju – PosTel 2021*, pp. 249-258, December 2021.
- [5] C. Perkins, E. Belding-Royer, and Ss Das, "Ad Hoc On demand Distance Vector (AODV) routing" RFC 3561, IETF, October 2003.
- [6] NS-3, [Online]. Available at: <https://www.nsnam.org/>
- [7] SUMO, [Online]. Available at: <https://www.eclipse.org/sumo/>

**Abstract:** *Vehicular Ad hoc Networks (VANETs) are characterized by high mobility of nodes and frequent changes in the network topology, which significantly complicates the process of routing data packets. It has been shown that traditional routing protocols are unable to promptly follow these changes. This is the reason why protocols based on reinforcement learning have been developed, as they enable constant monitoring of changes in the network environment and adaptation of the routing process to those changes. In this paper, an analysis and comparison of the application of these two types of protocols are performed, to prove the utility of the approach based on reinforcement learning in VANETs. The AODV (Ad hoc On-Demand Distance Vector) protocol is chosen as the representative of traditional protocols, while the ARPRL (Adaptive Routing Protocol based on Reinforcement Learning) protocol is chosen as the representative of protocols based on reinforcement learning. The simulation results showed that the ARPRL protocol shows significantly better network performance in terms of packet loss ratio and end-to-end delay.*

**Keywords:** *routing protocols, reinforcement learning, VANET*

### ANALYSIS OF REINFORCEMENT LEARNING-BASED ROUTING PROTOCOLS FOR VANETS

Nenad Jevtić, Pavle Bugarčić