

<https://doi.org/10.37528/FTTE/9788673954752/POSTEL.2023.019>

UNAPREĐENJE PROTOKOLA RUTIRANJA ZA VANET MREŽE KORIŠĆENJEM MAŠINSKOG UČENJA

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

Rezime: *Ukoliko se mrežna topologija stalno menja, korišćenje iste putanje za slanje podataka u dužem vremenskom periodu dovodi do degradiranja brojnih mrežnih performansi. Ovaj problem se javlja kod primene tradicionalnih protokola rutiranja u bežičnim ad hoc mrežama za vozila (Vehicular Ad-hoc Networks, VANETs). Zato je korisno pratiti stanje mrežnog okruženja i u skladu sa tim birati najbolju putanju za slanje podataka. Mehanizam kojim se ovo na vrlo efikasan način omogućava naziva se učenje potkrepljivanjem (Reinforcement Learning, RL), koje predstavlja posebnu kategoriju mašinskog učenja. U ovom radu pokazano je kako primena RL unapređuje tradicionalne protokole rutiranja, poređenjem poznatog Ad hoc On-Demand Distance Vector (AODV) protokola, Q-Learning AODV (QLAODV) protokola, koji predstavlja modifikaciju AODV baziranu na RL, i Adaptive Routing Protocol based on Reinforcement Learning (ARPR) protokola koji takođe koristi RL za rutiranje podataka. Rezultati simulacija su pokazali da protokoli bazirani na RL daju značajno bolje performanse u poređenju sa AODV protokolom.*

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

1. Uvod

Bežične *ad hoc* mreže za vozila (*Vehicular ad hoc Networks*, VANETs) predstavljaju posebnu kategoriju bežičnih *ad hoc* mreža (*Wireless ad hoc Networks*, WANETs) kod kojih mrežne čvorove čine vozila koja žele da ostvare bežičnu komunikaciju. *Ad hoc* mreže su specifične po tome što nemaju fiksnu komunikacionu infrastrukturu, već se veza između krajnjih čvorova uspostavlja samo u slučaju potrebe za slanjem podataka. Svi čvorovi u mreži imaju ulogu rutera, a putanja kojom će se prosledivati podaci definiše se pomoću odgovarajućeg protokola rutiranja. Iako u poslednje vreme sve veću pažnju privlači 5G mreža, koja omogućava značajno veće protoke i manji procenat izgubljenih paketa, VANET mreže i dalje imaju određene prednosti zbog kojih su veoma značajan segment inteligentnih transportnih sistema. Najveća prednost VANET mreža leži u njihovoj jednostavnosti i relativno maloj ceni implementacije, s obzirom da nisu potrebna visoka inicijalna ulaganja u mrežnu

infrastrukturu. Takođe, veliki problem 5G mreža je njihova i dalje ne tako široka rasprostranjenost, kako zbog visokih troškova implementacije tako i zbog određenih društvenih neslaganja sa njihovim uvođenjem. Poznato je da u Republici Srbiji i dalje nije uvedena 5G mreža, kao ni u mnogim zemljama u razvoju, tako da je kreiranje VANET mreža najbolji način za ostvarivanje komunikacije između vozila, posebno u urbanim sredinama i na pametnim autoputevima.

Veoma bitan segment za nesmetano funkcionisanje VANET mreža je izbor adekvatnog protokola rutiranja. U ovom radu će biti izvršena simulaciona analiza i poređenje rezultata primene tradicionalnog protokola rutiranja i protokola rutiranja baziranih na učenju potkrepljivanjem. Testiranje je izvršeno u NS-3 (*Network Simulator 3*) [1] simulatoru, koji omogućava implementaciju i testiranja novih protokola rutiranja. Upoređeni su rezultati primene AODV (*Ad hoc On-Demand Distance Vector*) protokola [2] kao predstavnika tradicionalnih protokola rutiranja, čiji je model već implementiran u NS-3 simulatoru, sa rezultatima primene QLAODV (*Q-Learning AODV*) [3] i ARPRL (*Adaptive Routing Protocol based on Reinforcement Learning*) [4] protokola koji se baziraju na učenju potkrepljivanjem. Za testiranje poslednja dva protokola bila je neophodna njihova implementacija u NS-3 simulatoru. Opis i implementacija ARPRL protokola izvršeni su u [5], pa će u ovom radu fokus biti na QLAODV protokolu. Poređenje protokola je izvršeno na osnovu procenta izgubljenih paketa (*Packet Loss Ratio*, PLR), prosečnog kašnjenja s kraja na kraj mreže (*End to End Delay*, E2ED) i procenta korisnog saobraćaja (*Useful Traffic Ratio*, UTR).

Nastavak rada je organizovan na sledeći način. U drugom poglavlju predstavljena je problematika rutiranja u VANET mrežama, zatim je u trećem poglavlju opisan princip funkcionisanja QLAODV protokola. U četvrtom poglavlju su prikazani rezultati simulacione analize gore pomenutih protokola, a u petom poglavlju su data zaključna razmatranja.

2. Rutiranje u VANET mrežama

Protokoli rutiranja u bežičnim komunikacionim mrežama imaju zadatak da obezbede optimalnu putanju za prosleđivanje podataka od izvora do odredišta, preko niza međučvorova. Ovaj zadatak postaje prilično komplikovan u slučaju visoko dinamičnih mreža, kao što su VANET mreže. Osnovna karakteristika VANET mreža je relativno velika brzina kretanja mrežnih čvorova, što prouzrokuje česte promene u mrežnoj topologiji. Ovo značajno otežava održavanje optimalne putanje za slanje podataka, zbog čestih prekida linkova koji su deo trenutnih putanja. Rešenje ovog problema može biti u uključivanju mašinskog učenja u proces izbora optimalne putanje. Tip mašinskog učenja koji najviše odgovara ovako dinamičnim mrežama je učenje potkrepljivanjem (*Reinforcement Learning*, RL) [6], koje podrazumeva stalnu interakciju agenta učenja sa okruženjem kroz preduzimanje određenih akcija prema okruženju, na šta okruženje odgovara odgovarajućom nagradom (ili kaznom) za preduzetu akciju i informacijom o novom stanju okruženja nakod preduzete akcije.

Kod VANET mreža agent učenja može biti određeni čvor (vozilo) u mreži, akcije koje preduzima ovaj agent ogledaju se u izboru sledećeg čvora za slanje paketa ka odredištu, a okruženje predstavljaju svi ostali čvorovi u mreži. Vozila dobijanjem periodičnih povratnih informacija od okruženja prate njegove promene, prilagođavaju im se i imaju mogućnost da uvek izaberu najbolju putanju za slanje podataka. Jedan od

najčešće korišćenih tipova RL u VANET mrežama je Q-učenje (*Q-Learning*, QL) [7], u kom čvorovi na osnovu interakcija sa okruženjem računaju Q-vrednosti za sve putanje do odgovarajućih odredišta, preko svojih susednih čvorova. Ove Q-vrednosti čvorovi čuvaju u svojoj Q-tabeli, na osnovu koje biraju sledeći čvor za slanje podataka ka odredištu.

Protokole bazirane na RL je moguće primeniti i u brojnim drugim mrežama, kao što su bežične senzorske mreže (*Wireless Sensor Networks*, WSN), mobilne *ad hoc* mreže (*Mobile Ad-hoc Networks*, MANET) ili bežične *ad hoc* mreže za bespilotne letelice (*Flying Ad-hoc Networks*, FANET). Takođe, pored QL koji je opisan u ovom radu, moguće je primeniti brojne druge tipove RL, kao što su SARSA algoritam, DRL (*Deep Reinforcement Learning*), DDRL (*Dueling DRL*) ili MBRL (*Model-Based RL*). Testiranje složenijih tipova RL u scenarijima sa većim brojem čvorova zahteva računare izuzetno velike procesorske moći, kako bi simulacije bile obavljene u razumnom vremenskom periodu.

3. QLAODV protokol

Kako bi se protokol rutiranja što bolje prilagodio dinamičkoj prirodi VANET mreža, razvijene su brojne modifikacije postojećih protokola koje uz pomoć učenja potkrepljivanjem pokušavaju da unaprede mrežne performanse stalnom interakcijom sa mrežnim okruženjem [8]. QLAODV protokol predstavlja modifikaciju jednog od najpoznatijih reaktivnih protokola rutiranja, AODV protokola. Princip otkrivanja putanje kod QLAODV protokola je sličan kao kod AODV protokola. Glavna promena je u uvođenju konstantnog periodičnog slanja *Hello* poruka ka susedima. Putem ovih poruka čvorovi razmenjuju svoje znanje o okruženju, na osnovu čega se ažuriraju Q-vrednosti koje određuju kojom putanjom će se slati paketi podataka. Svaki čvor održava tabelu Q-vrednosti za putanje ka svakom odredištu u mreži, preko svih susednih čvorova. Kada šalje pakete ka nekom odredištu, čvor uvek bira onu putanju preko svog suseda koja ima najveću Q-vrednot.

<i>IP adresa pošiljaoca</i>
<i>Skv. broj Hello paketa</i>
<i>Faktor mobilnosti</i>
<i>Faktor propusnog opsega</i>
<i>Broj maks. Q-vrednosti</i>
<i>Maks. Q-vrednost suseda [1]</i>
<i>...</i>
<i>Maks. Q-vrednost suseda [n]</i>

Slika 1. Struktura Hello paketa kod QLAODV protokola

Čvorovi periodično šalju *Hello* pakete ka svojim susedima, kako bi im dostavili podatke neophodne za ažuriranje Q-vrednosti. Izbor podataka koje čvor šalje u ovim

porukama od ključnog je značaja za kreiranje adekvatne slike agenta učenja o stanju mreže. Struktura *Hello* paketa prikazana je na slici 1. IP adresa je neophodna za identifikaciju čvora pošiljaoca na prijemu, sekvencijalni broj prati redosled pristiglih *Hello* paketa, dok ostali parametri služe za računanje Q-vrednosti. Faktor mobilnosti i faktor propusnog opsega oslikavaju uticaj relativnog kretanja susednih čvorova i opterećenosti veze preko čvora pošiljaoca. Broj maksimalnih Q-vrednosti odgovara broju odredišta do kojih čvor koji šalje *Hello* paket ima putanju, nakon čega sledi niz maksimalnih Q-vrednosti za putanje (preko najboljeg suseda) do svakog odredišta.

Po prijemu *Hello* paketa od nekog suseda x , čvor s ažurira Q-vrednosti za putanje ka svim odredištima d koje idu preko ovog suseda, primenjujući sledeću formulu:

$$Q_s(d, x) = (1 - \alpha) * Q_s(d, x) + \alpha[R + \gamma_x * \max_{y \in N_x} Q_x(d, y)] \quad (1)$$

gde N_x predstavlja skup svih suseda čvora x , α i γ_x označavaju stepen učenja i diskontni faktor, R predstavlja nagradu, a $\max_{y \in N_x} Q_x(d, y)$ predstavlja maksimalnu Q-vrednost koju ima čvor x za putanju ka odredišnom čvoru d preko nekog od svojih suseda. Stepem učenja može biti u opsegu od 0 do 1 i on određuje nivo uticaja promena u mrežnom okruženju na izbor putanje za slanje podataka. Ovaj koeficijent ne bi smeo da bude preveliki, jer bi čvorovi prebrzo učili i svaka promena u mreži bi previše brzo menjala Q-vrednosti i dovela do nestabilnosti mreže. Ali stepen učenja ne treba da bude ni premali kako čvorovi ne bi presporo pratili promene u mrežnom okruženju. U ovom protokolu α ima fiksnu vrednost od 0,8, do čega se došlo eksperimentalnim putem.

Diskontni faktor određuje važnost budućih Q-vrednosti i u QLAODV protokolu se računa na sledeći način:

$$\gamma_x = \gamma * \sqrt{(MF_x * BF_x)} \quad (2)$$

gde γ ima fiksnu vrednost od 0,9, a MF_x i BF_x predstavljaju faktor mobilnosti i faktor uticaja propusnog opsega za čvor x , respektivno. Oba faktora se računaju u čvoru x i šalju čvoru s u okviru *Hello* paketa. Faktor mobilnosti se računa na sledeći način:

$$MF_x = \begin{cases} \sqrt{\frac{|N_x \cap N_x^p|}{|N_x \cup N_x^p|}}, & \text{za } N_x \cup N_x^p \neq \emptyset \\ 0, & \text{inače} \end{cases} \quad (3)$$

gde N_x predstavlja trenutni skup suseda čvora x , dok N_x^p predstavlja skup suseda čvora x u trenutku slanja prethodnog *Hello* paketa. Očigledno je da se na ovaj način pednost daje stabilnijim putanjama, gde nema velikih promena u okruženju posmatranog čvora x . Faktor uticaja propusnog opsega računa se preko sledeće jednačine:

$$BF_x = \frac{\text{Dostupan propusni opseg čvora } x}{\text{Maksimalni propusni opseg za čvor } x} \quad (4)$$

Dostupni propusni opseg čvora x može se dobiti oduzimanjem iskorišćenog propusnog opsega od maksimalnog propusnog opsega za čvor x . Iskorišćeni propusni

opseg (BW) u nekom vremenskom periodu T moguće je izračunati na osnovu sledeće jednačine:

$$BW(bps) = \frac{n \cdot S_B \cdot 8}{T} \quad (5)$$

gde n predstavlja ukupan broj paketa poslatih i primljenih od strane čvora x tokom vremenskog intervala T , a S_B je veličina paketa u bajtima. Maksimalni propusni opseg je konstantni parametar koji zavisi od karakteristika upotrebene bežične veze i identičan je za sve čvorove u mreži. Uzimajući u obzir faktor propusnog opsega, prednost se daje putanjama koje nisu preopterećene, pa se tako može izbeći zagušenje mreže i povećanje procenta izgubljenih paketa.

Nagrada R služi za nagrađivanje direktnih putanja ka odredišnom čvoru i definisana je na sledeći način:

$$R = \begin{cases} 1, & \text{ako } s \in N_d \\ 0, & \text{inače} \end{cases} \quad (6)$$

gde je N_d skup suseda odredišnog čvora d . Drugim rečima, ukoliko čvor primi *Hello* poruku od odredišnog čvora, nagrada će biti 1, inače nagrada je 0. Nagrađivanjem direktne putanje smanjuje se broj hopova i kašnjenje paketa s kraja na kraj mreže.

Pored dinamičkog ažuriranja putanja putem razmene *Hello* paketa, QLAODV protokol obezbeđuje mehanizam provere nove putanje slanjem *unicast* RCNG-REQ i RCNG-REP paketa. Ovi paketi nisu postojali u AODV protokolu. Ukoliko čvor koji koristi određenu putanju za slanje podataka ka odredištu, otkrije novu putanju sa većom Q-vrednošću, pre promene putanje vrši proveru nove putanje slanjem RCNG-REQ kontrolnog paketa. Međučvorovi takođe prosleđuju ovaj paket preko suseda sa najvećim Q-vrednostima ka odredištu. Po prijemu RCNG-REQ paketa, odredišni čvor odgovara slanjem RCNG-REP paketa. Tek po prijemu RCNG-REP paketa međučvorovi i čvor pošiljalac ažuriraju svoju tabelu rutiranja i menjaju putanju kojom šalju pakete podataka.

3. Testiranje i analiza protokola u simulacionom okruženju

Testiranje protokola rutiranja u realnim uslovima bilo bi veoma komplikovano i skupo za realizaciju, pa se zato u tu svrhu najčešće koriste mrežni simulatori. Posebnu pogodnost daju simulatori otvorenog koda, koji istraživačima omogućavaju menjanje postojećih i implementaciju novih modela. U ovom radu je za testiranje protokola izabran NS-3 simulator, u okviru kog je već implementiran model AODV protokola. Za testiranje QLAODV protokola, bila je neohodna njegova implementacija u simulaciono okruženje. Da bi protokol nesmetano funkcionisao, bile su neohodne određene korekcije originalne verzije protokola. Takođe je izvršeno i poređenje sa ARPRL protokolom, opisanim i implementiranim u [5].

Protokoli su testirani u dva različita scenarija, koji odgovaraju scenarijima iz [3] i [4]. Prvi scenario podrazumeva simulaciono područje dimenzija 1000m x 1000m, sa po 4 horizontalne i vertikalne ulice i po jednom trakom u svakom smeru. Unutar ovog područja kreće se 80 vozila od kojih 30 (nasumično izabranih) generiše CBR (*Constant Bit Rate*) saobraćaj brzine 32 kb/s. Maksimalna dozvoljena brzina čvorova varira od 1 m/s do 25 m/s. Drugi scenario čini simulaciono područje dimenzija 2000m x 2000m, sa

po 5 horizontalnih i vertikalnih ulica, svaka ulica ima po dve trake za kretanje vozila u oba smera, gde svaka raskrsnica ima semafor. Maksimalna dozvoljena brzina za kretanje vozila je 15 m/s. Ukupan broj vozila varira od 50 do 350, pri čemu uvek 20 nasumično izabranih vozila generiše CBR saobraćaj brzine 4 kb/s.

Ostali parametri simulacije su identični za oba scenarija. Za generisanje mobilnosti vozila korišćen je SUMO (*Simulation of Urban MObility*) simulator [9]. Korišti 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*, a na transportnom sloju je korišćen UDP (*User Datagram Protocol*) protokol. Aplikacioni sloj generiše pakete veličine 512B. Ostali parametri imaju podrazumevana podešavanja za mrežni simulator. Pregled najvažnijih parametara prvog i drugog simulacionog scenarija prikazani su u tabelama 1 i 2, respektivno.

Tabela 1. Pregled parametara prvog simulacionog scenarija

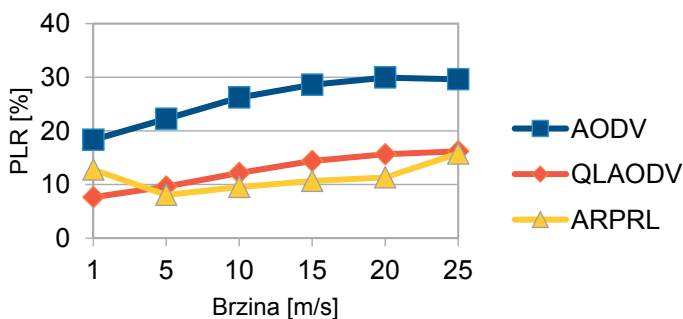
Parametar	Vrednost
Dimenzije prostora	1000 m x 1000 m
Ulice	Po 4 horizontalne i vertikalne ulice
Trajanje simulacije	600 s
Ukupan broj vozila	80
Maksimalna brzina vozila	1 m/s, 5 m/s, 10 m/s, 15 m/s, 20 m/s, 25 m/s
Model mobilnosti	<i>Manhattan Grid</i>
Propagacioni model	<i>Two Ray Ground</i>
MAC	IEEE 802.11p, širina kanala 10 MHz, protok 6 Mb/s
Protokol rutiranja	AODV, ARPRL, QLAODV
Transportni protokol	UDP
Aplikacioni protok	32 kb/s (CBR, 30 vozila)
Veličina paketa	512 B

Tabela 2. Pregled parametara drugog simulacionog scenarija

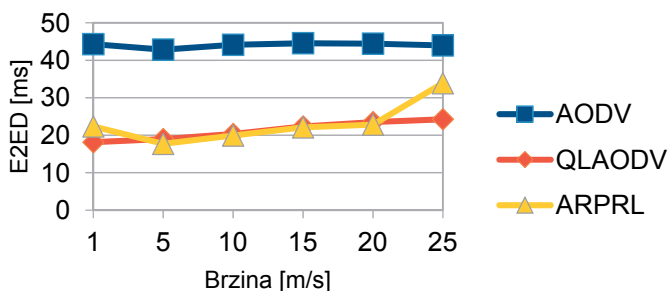
Parametar	Vrednost
Dimenzije prostora	2000 m x 2000 m
Ulice	Po 5 horizontalnih i vertikalnih ulica
Trajanje simulacije	600 s
Ukupan broj vozila	50, 100, 150, 200, 250, 300, 350
Maksimalna brzina vozila	15 m/s
Model mobilnosti	<i>Manhattan Grid</i>
Propagacioni model	<i>Two Ray Ground</i>
MAC	IEEE 802.11p, širina kanala 10 MHz, protok 6 Mb/s
Protokol rutiranja	AODV, ARPRL, QLAODV
Transportni protokol	UDP
Aplikacioni protok	4 kb/s (CBR, 20 vozila)
Veličina paketa	512 B

Poređenje protokola je izvršeno na osnovu tri relevantne mrežne performanse, procenta izgubljenih paketa (PLR), prosečnog kašnjenja s kraja na kraj mreže (E2ED) i procenta korisnog saobraćaja (UTR). Posmatrajući prvi scenario, može se zaključiti da

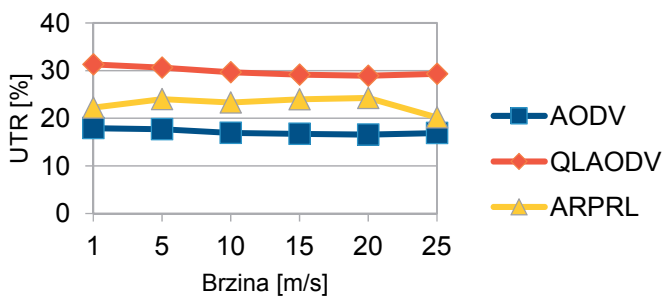
QLAODV, slično kao ARPRL, pokazuje značajno bolje performase u pogledu PLR (slika 2) i E2ED (slika 3) u odnosu na AODV protokol, kako pri manjim tako i pri većim brzinama kretanja vozila. Ovo je definitivno posledica boljeg izbora putanja i brzog otkrivanja promena u topologiji mreže, što je omogućeno stalnom interakcijom čvorova (koji šalju pakete) sa okruženjem. Posmatrajući UTR u prvom scenariju (slika 4), primetno je da QLAODV pokazuje najbolje rezultate, što znači da se koristi manje kontrolnih paketa, koji samim tim unose manje dodatno opterećenje mreže.



Slika 2. Zavisnost procenta izgubljenih paketa od maksimalne dozvoljene brzine čvorova u mreži

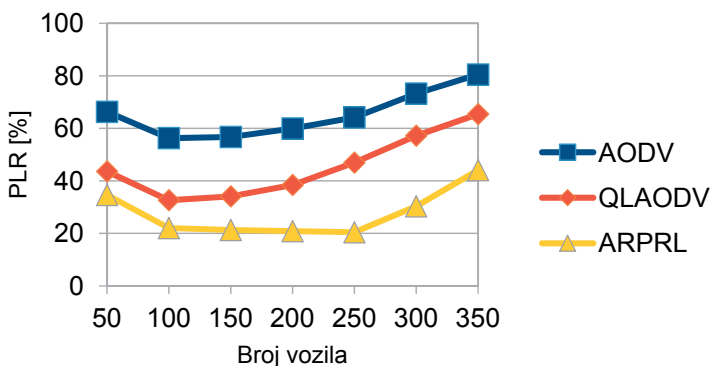


Slika 3. Zavisnost prosečnog kašnjenja paketa s kraja na kraj mreže od maksimalne dozvoljene brzine čvorova u mreži

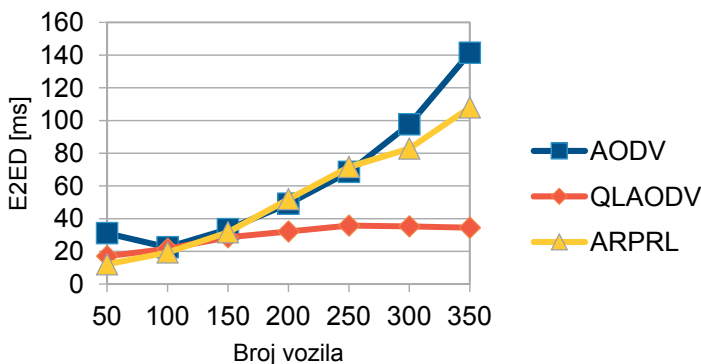


Slika 4. Zavisnost procenta korisnog saobraćaja od maksimalne dozvoljene brzine čvorova u mreži

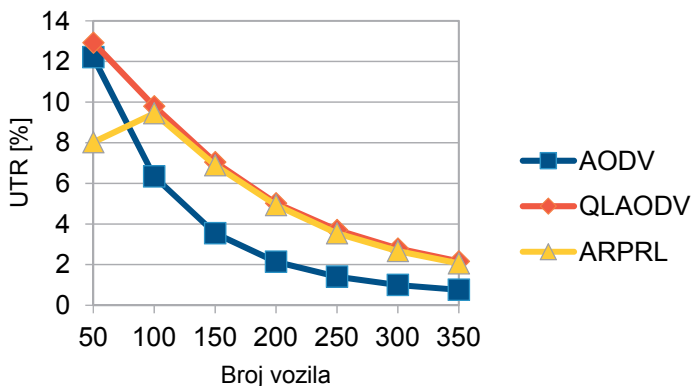
U drugom scenariju analizirani su mrežni parametri uz varijaciju ukupnog broja čvorova u mreži. Može se primetiti da i u ovom scenariju QLAODV protokol pokazuje značajno bolje performanse u pogledu PLR (slika 5) i UTR (slika 7) u odnosu na AODV protokol, a u pogledu E2ED (slika 6) ubedljivo najbolje performanse pri većem broju čvorova u mreži. Sa druge strane, ARPRL protokol pokazuje najbolje performanse u pogledu PLR u ovom scenariju, što govori da je nešto pogodniji za pokrivanje većih oblasti sa većim brojem vozila. U svakom slučaju, oba protokola bazirana na primeni učenja potkrepljivanjem pokazuju značajno poboljšanje performansi u poređenju sa AODV protokolom, bez obrira na veličinu simulacione oblasti, ukupan broj vozila i brzinu kojom se ova vozila kreću.



Slika 5. Zavisnost procenta izgubljenih paketa od ukupnog broja čvorova u mreži



Slika 6. Zavisnost prosečnog kašnjenja paketa s kraja na kraj mreže od ukupnog broja čvorova u mreži



Slika 7. Zavisnost procenta korisnog saobraćaja od ukupnog broja čvorova u mreži

4. Zaključak

U ovom radu je izvršena implementacija, testiranje i analiza primene AODV, QLAODV i ARPRL protokola rutiranja u VANET mrežama. Analizirajući rezultate primene pomenutih protokola, očigledno je da protokoli bazirani na učenju pokrepljivanjem značajno nadmašuju tradicionalni AODV protokol rutiranja posmatrajući ključne mrežne performanse u različitim simulacionim scenarijima. Posebno je značajno što su se ovi protokoli pokazali veoma dobro pri povećanju maksimalne dozvoljene brzine vozila, što govori o tome da uspevaju blagovremeno da isprate sve intenzivnije promene u mrežnoj topologiji i da adekvatno reaguju u slučaju prekida linkova između čvorova u mreži. Takođe, ovi protokoli su pokazali robusnost na povećanje saobraćajnog opterećenja, s obzirom da i u uslovima velikog broja vozila u mreži, a samim tim i veće opasnosti od zagušenja, pokazuju značajno bolje rezultate u poređenju sa AODV protokolom. Može se zaključiti da od testiranih protokola u slučaju velikih brzina vozila najbolje performanse pokazuje QLAODV protokol, dok je u slučaju veće gustine saobraćaja najbolje koristiti ARPRL protokol.

U budućim istraživanjima planiran je razvoj i implementacija novog protokola rutiranja, koji bi bio adekvatniji za primenu u VANET mrežama. Novi protokol bi efikasnije pratio promene u složenom mrežnom okruženju, uzimajući u obzir različite parametre kao što su rastojanje između čvorova, relativna promena rastojanja u vremenu, kašnjenje paketa s kraja na kraj mreže, pouzdanost veze, opterećenost propusnog opsega, kao i broj hopova od izvora do odredišta. Pored praćenja kvaliteta veze između susjednih čvorova, protokol treba da uzme u obzir i kvalitet celokupne putanje od izvornog do odredišnog čvora, čime bi se mrežne karakteristike dodatno poboljšale.

Literatura

- [1] NS-3, [Online]. Available at: <https://www.nsnam.org/>
- [2] C. Perkins, E. Belding-Royer, and Ss Das, "Ad Hoc On demand Distance Vector (AODV) routing" RFC 3561, IETF, October 2003.

- [3] 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
- [4] 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
- [5] N. Jevtić, P. Bugarčić. "Analiza protokola rutiranja baziranih na učenju potkrepljivanjem za VANET mreže", in *XL Simpozijum o novim tehnologijama u poštanskom i telekomunikacionom saobraćaju – PosTel 2022*, pp. 375-384, November 2021.
- [6] R. Sutton, A. Barto, *Reinforcement Learning: An Introduction, second edition*, Cambridge, Massachusetts, MIT Press, 2018.
- [7] M. L. Littman, "Reinforcement learning improves behaviour from evaluative feedback", *Nature*, vol. 521, pp. 445-451, May 2015. DOI: 10.1038/nature14540
- [8] P. Bugarčić, N. Jevtić, M. Malnar, "Reinforcement Learning-Based Routing Protocols in Vehicular and Flying Ad Hoc Networks—A Literature Survey", *Promet*, vol. 34, pp. 893-906, December 2022. DOI: 10.7307/ptt.v34i6.4159
- [9] SUMO, [Online]. Available at: <https://www.eclipse.org/sumo/>

Abstract: *In cases when the network topology is constantly changing, using the same route to send data for a long period leads to the degradation of various network performances. This problem occurs when traditional routing protocols are applied in vehicular ad-hoc networks (VANETs). Therefore, it is useful to monitor the state of the network environment and choose the best route for sending data accordingly. The mechanism that enables this in a very efficient way is called reinforcement learning (RL), which represents a special category of machine learning. This paper shows how the application of RL improves traditional routing protocols by comparing the well-known Ad hoc on-demand distance vector (AODV) protocol, Q-learning AODV (QLAODV) protocol, which is a modification of AODV based on RL, and Adaptive routing protocol based on reinforcement learning (ARPRL) that also uses RL for data routing. The simulation results proved that the RL-based protocols show significantly better performance compared to the AODV protocol.*

Keywords: *routing protocols, machine learning, VANET*

IMPROVEMENT OF ROUTING PROTOCOLS FOR VANETS USING MACHINE LEARNING

Nenad Jevtić, Marija Malnar, Pavle Bugarčić