



# PARRoT: Predictive Ad-hoc Routing Fueled by Reinforcement Learning and Trajectory Knowledge

Benjamin Sliwa, Cedrik Schüler, Manuel Patchou, and Christian Wietfeld

Communication Networks Institute, TU Dortmund University, 44227 Dortmund, Germany

e-mail: {Benjamin.Sliwa, Cedrik.Schueler, Manuel.Mbankeu, Christian.Wietfeld}@tu-dortmund.de

**Abstract**—Swarms of collaborating Unmanned Aerial Vehicles (UAVs) that utilize ad-hoc networking technologies for coordinating their actions offer the potential to catalyze emerging research fields such as autonomous exploration of disaster areas, demand-driven network provisioning, and near field delivery in Intelligent Transportation Systems (ITSs). As these mobile robotic networks are characterized by high grades of relative mobility, existing routing protocols often fail to adopt their decision making to the implied network topology dynamics. For addressing these challenges, we present Predictive Ad-hoc Routing fueled by Reinforcement learning and Trajectory knowledge (PARRoT) as a novel machine learning-enabled routing protocol which exploits mobility control information for integrating knowledge about the future motion of the mobile agents into the routing process. The performance of the proposed routing approach is evaluated using comprehensive network simulation. In comparison to established routing protocols, PARRoT achieves a massively higher robustness and a significantly lower end-to-end latency.

## I. INTRODUCTION

Collaborating autonomous drones that coordinate their actions using Flying Ad-hoc Networks (FANETs) offer the potential to efficiently perform important disaster relief tasks — e.g., remote sensing and network provisioning — without risking the lives of human helpers [1]. A closely related emerging research field is the integration of small-scale UAVs into future ITSs [2] for applications such as aerial traffic monitoring [3] and UAV-aided near field delivery [4]. While the latter concept has been initially proposed for reducing the delivery time in inner cities, its inherent avoidance of direct human-to-human interaction also makes it a promising candidate for increasing the delivery safety during the COVID-19 pandemic [5]. An illustration about different applications of UAV-based FANETs is shown in Fig. 1. For enabling these novel use-cases, the provision of efficient and reliable means of communication even in challenging environments is an important prerequisite. However, established Mobile Ad-hoc Network (MANET) routing protocols can often barely cope with the small channel coherence time and the network topology dynamic implied by the high grade of relative mobility. Anticipatory mobile networking [6] has been proposed for explicitly addressing the interdependency of mobility and communication by integrating *context* knowledge into the corresponding decision processes. This novel communications paradigm has a strong relationship to the usage of machine learning for optimizing wireless communication networks [7]

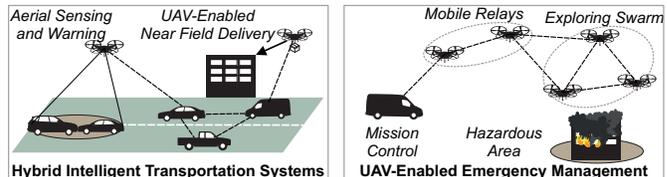


Fig. 1. Example applications of UAV-enabled wireless mesh networks.

which manifests in the trend of replacing complex mathematical models by learned representations of the corresponding phenomena [8]. Moreover, it has catalyzed the emergence of novel performance evaluation methods that are capable of replacing computationally expensive entity-based modeling with machine learning-based end-to-end models [9]. In this paper, we present PARRoT as a novel reinforcement learning-enabled cross-layer routing protocol which leverages application layer knowledge from the mobility control routines for proactively optimizing the robustness of vehicular routing paths. The remainder of the paper is structured as follows. After discussing the related work in Sec. II, the novel PARRoT protocol is presented in Sec. III. Afterwards, an overview about the methodological aspects of the simulative performance evaluation is given in Sec. IV. Finally, detailed simulation results are provided and discussed in Sec. V.

## II. RELATED WORK

A wide range of solution approaches for specific applications and different kinds of vehicular networks has been proposed by literature. Comprehensive summaries about existing protocols for highly mobile networks are presented by the authors of [10] and [11]. Moreover, Cavalcanti et al. [12] provide an empirical analysis of the popularity of existing routing protocols and performance evaluation methods in the context of vehicular networking. Classically, MANET routing protocols have been classified into *reactive* — e.g., Ad-hoc On-demand Distance Vector (AODV) and Dynamic MANET On Demand Routing Protocol (DYMO) — and *proactive* — e.g., Destination-Sequenced Distance Vector (DSDV), Optimized Link State Routing (OLSR), and Better Approach To Mobile Ad-hoc Networking (B.A.T.M.A.N.) — methods. However, the need to pay attention to the interdependency between mobility and communication has led to the rise of *geo-based* routing approaches such as Greedy Perimeter Stateless Routing in Wireless Networks (GPSR) which integrate position and velocity information into their decision making.

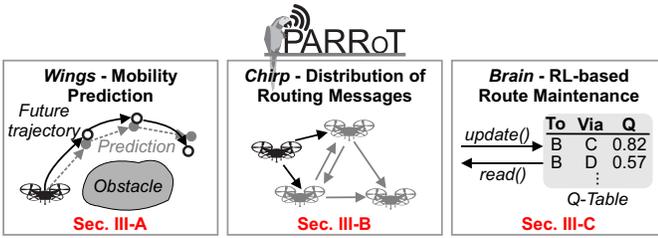


Fig. 2. Overall system architecture model of the PARRoT routing protocol.

In extension, *geo-predictive* approaches such as B.A.T.Mobile consider the anticipated future relative mobility of the vehicles within the routing process [13]. Similarly, Song et al. [14] present an extended variant of OLSR which uses Kalman filter-based mobility prediction for optimizing the Multipoint Relay (MPR) determination process.

*Reinforcement learning* can be regarded as a step towards zero touch optimization of wireless communication systems. Hereby, an *agent* learns to autonomously perform favorable *actions* in a defined *environment* through observation of the *rewards* of previously taken actions. Machine learning-enabled routing methods have been proposed by different researchers. The authors of [15] utilize an Artificial Neural Network (ANN)-enabled centralized routing approach using Software-defined Networking (SDN) for Vehicular Ad-hoc Networks (VANETs) delay minimization. Similar to our work, their proposed routing method Centralized Routing Scheme with Mobility Prediction (CRS-MP) takes into account mobility predictions of the mobile vehicles. Oddi et al. [16] use geo-based routing metrics jointly with Q-Learning-enabled reinforcement learning. Similar to B.A.T.Mobile, and as further discussed in Sec. III, this method represents important groundwork for the novel PARRoT protocol. Due to the availability of data analysis tools such as LighTweight Machine learning for IoT Systems (LIMITS) [17] which allow to automatically derive C++ implementations of trained prediction models, it can be expected that pervasive machine learning will be one of the key enablers for future network generation.

Another ongoing development is the partial convergence of cellular and ad-hoc networking paradigms, illustrated by a growing interest in integrating single- and multi-hop device-to-device communication into cellular networks and manifested by concepts such as Multi-hop Cellular Networks (MCNs) [18]. As a consequence, novel developments in the ad-hoc networking domain might also have an impact on future cellular network generations [8].

### III. MACHINE LEARNING-ENABLED WIRELESS MESH ROUTING WITH PARRoT

The overall system architecture model of the proposed PARRoT is shown in Fig. 2. PARRoT consists of three logical core components which are further explained in the following paragraphs.

#### A. PARRoT Wings - Cross-Layer Mobility Prediction

PARRoT utilizes a cross-layer approach that leverages knowledge from the mobility control domain for anticipating

Byte Offset	
0	1 2 3
0	Originator Address
4	
8	Current Position $\mathbf{p}(t)$
12	
16	
20	Predicted Position $\tilde{\mathbf{p}}(t + \tau)$
24	
28	Reward $V$
32	Cohesion $\Phi_{\text{Coh}}$
36	SEQ   TTL

Fig. 3. Structure of the chirp data packet with a total size of 40 Byte.

the relative mobility between the different agents for a defined prediction horizon  $\tau$ . For this purpose, each agent estimates its own future position  $\tilde{\mathbf{p}}(t + \tau)$  based on the current position  $\mathbf{p}(t)$ . The result is then propagated to the other nodes of the network via the *chirp* distribution process (see Sec. III-B). If the PARRoT agent owns information about the planned trajectory as a sequence of  $k$  waypoints, an iterative prediction method consisting of  $N = \lfloor \tau / \Delta t \rfloor$  steps is applied whereas  $\Delta t$  represents the mobility update interval and the final result is given by  $\tilde{\mathbf{p}}(t + \tau) = \tilde{\mathbf{p}}_N$ . In each iteration  $i$ , the agent virtually moves towards the current waypoint  $\mathbf{w}_k$  as

$$\tilde{\mathbf{p}}_{i+1} = \tilde{\mathbf{p}}_i + \frac{\mathbf{w}_k - \tilde{\mathbf{p}}_i}{\|\mathbf{w}_k - \tilde{\mathbf{p}}_i\|} \cdot v \cdot \Delta t \quad (1)$$

with  $\tilde{\mathbf{p}}_0 = \mathbf{p}(t)$  and  $v$  being the current velocity. After each iteration, it is checked if the vehicle is now within the radius  $r_w$  of the waypoint sphere: If  $\|\mathbf{w}_k - \tilde{\mathbf{p}}_{i+1}\| \leq r_w$  is fulfilled, the waypoint is considered *reached* and  $k$  is incremented.

As a fallback mechanism for cases where no waypoint information is available, the average slope of the previous  $h$  positions is computed and extrapolated for  $\tau$ . This method is non-iterative and allows to immediately compute the final result as

$$\tilde{\mathbf{p}}(t + \tau) = \mathbf{p}(t) + \frac{\tau}{h-1} \sum_{i=1}^{h-1} \frac{\mathbf{p}_i - \mathbf{p}_{i-1}}{\Delta t}. \quad (2)$$

#### B. PARRoT Chirp - Dissemination of Routing Messages and Context Knowledge

PARRoT relies on an exchange of User Datagram Protocol (UDP)-based routing messages — which are referred to as *chirps* in the following — for adopting the local routing knowledge to the highly dynamic network topology conditions. The structure of the 40 Byte chirp packet is illustrated in Fig. 3.

**Chirp message distribution:** Each PARRoT node periodically generates *chirp* messages based on a fixed interval  $\Delta t_{\text{chirp}}$  which are propagated through the network. Sequence numbering is used in order to allow the assessment of the data freshness. Upon reception of a PARRoT chirp message by node  $i$  via forwarder  $j$  and originated from  $d$ , the following steps are performed:

- If the Sequence Number (SEQ) for  $d$  is not newer than the one of the last received chirp from  $d$ , the message is discarded and the following steps are omitted.
- Otherwise, the received information is utilized to update the local knowledge maintained in the Q-Table based with

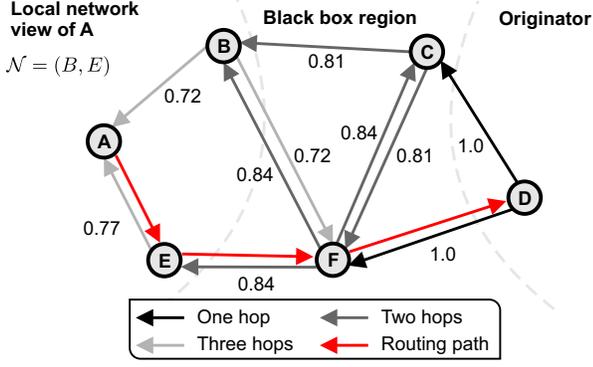


Fig. 4. Example for the propagation of chirp messages from originator  $D$  and its usage for reverse path routing from  $A$  to  $D$ .

the extracted values of  $V$  and  $\Phi_{\text{Coh}}$ . The corresponding process is further described in Sec. III-C.

- After the local handling, the chirp message is forwarded. For this purpose, the contained mobility information  $\mathbf{p}(t)$  and  $\tilde{\mathbf{p}}(t + \tau)$  of the forwarder  $j$  is replaced by the corresponding values of node  $i$ . The Time to Live (TTL) is reduced by 1 and  $V$  as well as  $\Phi_{\text{Coh}}$  are updated according to Sec. III-C.

An example for the propagation of chirp messages from node  $D$  and its usage for routing data packets from  $A$  to  $D$  is shown in Fig. 4.

- Node  $D$  generates a chirp and initializes the reward  $V_D$  with the highest possible value 1.0.
- The message is received by the one hop neighbor nodes  $C$  and  $F$  which update their Q-Table entries according to Eq. 3. Both then forward the chirp message with an updated reward which is calculated as  $V_{C/F} = Q(D, D)$ .
- The forwarded messages are received and handled by  $B$  and  $F$ . The rewards are updated as  $V_{B/F} = Q(D, \arg \max_j Q(D, j))$ .  $C$  and  $F$  also mutually receive the chirp message but discard it as the sequence number is outdated.
- $A$  finally receives the message originated from  $D$  via  $B$  and  $E$  and updates the corresponding table entries.
- For routing messages from  $A$  to  $D$ , node  $A$  only knows about its direct neighbors  $\mathcal{N} = (B, E)$  and about the existence of the destination node which is hidden beyond a black box network region. In each forwarding step, the message is propagated to neighbor  $j = \arg \max_j Q(D, j)$ .

### C. PARRoT Brain - Reinforcement Learning-Based Route Maintenance

The routing process of PARRoT is inspired by existing approaches and distills initial ideas by the authors of [16] and [13]. It consists of two core components which are explained in the following.

**Online routing process:** Similar to decentralized approaches such as B.A.T.M.A.N. and in contrast to path planning-based protocols such as OLSR, each node has only a partial view on the overall network topology. In order to route

messages to given destination  $d$ , each node  $i$  only assesses the suitability of its direct neighbors  $\mathcal{N}$  for reaching  $d$  — the intermediate network is treated as a black box. For this purpose, the numeric values for the end-to-end link quality  $Q(d, j)$  are maintained in a  $Q$ -Table. The online routing decision process can then be formulated as  $\arg \max_j Q(d, j)$  whereas the one-hop neighbor  $j$  with the highest  $Q$  value is chosen as a message forwarder. The implied maintenance of multiple paths for each destination inherently provides the protocol with self healing capabilities that allow PARRoT to quickly recover after failure of nodes.

**Update procedure:** Upon reception of a chirp message, the receiver node updates its local knowledge about the reverse path to the originator using a modified  $Q$ -Learning method as

$$Q(d, j) = Q(d, j) + \alpha [\gamma(j) \cdot V_j - Q(d, j)] \quad (3)$$

whereas  $\alpha$  is the learning rate and  $V_j$  represents the received reverse path score to the originator  $d$  via forwarder  $j$  which is extracted from the chirp message. The variable *discount factor*  $\gamma(j)$  serves as a *multidimensional routing metric* and is computed as

$$\gamma(j) = \gamma_0 \cdot \Phi_{\text{LET}}(i, j) \cdot \Phi_{\text{Coh}}(j) \quad (4)$$

with  $\gamma_0$  being a constant value for ensuring loop-free routing through a guaranteed metric degradation per hop.

$\Phi_{\text{LET}}(i, j)$  utilizes an estimation of the Link Expiry Time (LET) between  $i$  and  $j$  which takes into the account the results of the mobility prediction process (see Sec. III-A). For a given prediction horizon  $\tau$ , the metric is computed as

$$\Phi_{\text{LET}}(i, j) = \begin{cases} \sqrt{\frac{\text{LET}(i, j)}{\tau}} & : \text{LET}(i, j) < \tau \\ 1.0 & : \text{else} \end{cases} \quad (5)$$

With  $\Delta \mathbf{p} = \mathbf{p}_j - \mathbf{p}_i$  being the relative position and  $\Delta \mathbf{v} = \mathbf{v}_j - \mathbf{v}_i$  being the relative velocity, the relative trajectory can be written as  $\Delta \tilde{\mathbf{p}} = \Delta \mathbf{p} + t \cdot \Delta \mathbf{v}$ . The  $\text{LET}(i, j)$  between nodes  $i$  and  $j$  represents the time  $t$  where the distance between  $i$  and  $j$  exceeds the maximum communication radius  $r_{\text{TX}}$ . Thus,  $r_{\text{TX}} = \sqrt{\Delta \tilde{\mathbf{p}}_x^2 + \Delta \tilde{\mathbf{p}}_y^2 + \Delta \tilde{\mathbf{p}}_z^2}$  needs to be solved for  $t$  in order to determine  $\text{LET}(i, j)$ . Substituting

$$a = \Delta \mathbf{v}_x^2 + \Delta \mathbf{v}_y^2 + \Delta \mathbf{v}_z^2 \quad (6)$$

$$b = 2(\Delta \mathbf{p}_x \Delta \mathbf{v}_x + \Delta \mathbf{p}_y \Delta \mathbf{v}_y + \Delta \mathbf{p}_z \Delta \mathbf{v}_z) \quad (7)$$

$$c = \Delta \mathbf{p}_x^2 + \Delta \mathbf{p}_y^2 + \Delta \mathbf{p}_z^2 - r_{\text{TX}}^2 \quad (8)$$

allows to derive

$$t_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \quad (9)$$

Due to the square root, three different cases can be distinguished:

$$\text{LET}(i, j) = \begin{cases} 0 & : t_1 \leq 0 \wedge t_2 \leq 0 \\ t_2 & : t_1 \leq 0 \wedge t_2 > 0 \\ 0 & : t_1 > 0 \wedge t_2 > 0 \end{cases} \quad (10)$$

These conditions can be interpreted as follows. In the first case, the link is and will stay unavailable. In the second case,

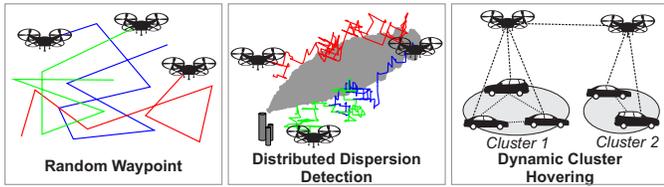


Fig. 5. Example trajectories and network topologies for the considered mobility algorithms.

the link is currently available and will expire at  $t_2$ . In the last case, the link is expected to become available at  $t_1$  and will expire at  $t_2$ .

$\Phi_{\text{Coh}}(j)$  is a measure for the neighbor set *coherence* of message forwarder  $j$  based on the difference between the neighbor sets  $\mathcal{N}$  at the times  $t$  and  $t - t_h$ . According to [16], it is derived as

$$\Phi_{\text{Coh}}(j) = \sqrt{1 - \frac{\mathcal{N}(t) \Delta \mathcal{N}(t - \Delta t)}{\mathcal{N}(t) \cup \mathcal{N}(t - \Delta t)}}} \quad (11)$$

with operator  $\Delta$  being the symmetrical difference between the two considered node sets.

#### IV. METHODOLOGY OF THE SIMULATIVE PERFORMANCE EVALUATION

In this section, the methodological aspects of the performance evaluation are presented. All simulations are performed using the Objective Modular Testbed in C++ (OMNeT++) framework [19] jointly with INETMANET which provides implementation of various MANET protocols. For the performance analysis, the Packet Delivery Ratio (PDR) and the end-to-end latency of a video streaming application modeled as UDP Constant Bitrate (CBR) is considered. Within each evaluation run, sender and receiver are randomly chosen from the total set of vehicles. The source code of the developed OMNeT++ implementation for PARRoT is provided in an open source manner<sup>1</sup>.

**Mobility models:** For analyzing the routing behavior of PARRoT, a mixture of generic and realistic mobility models is considered. Visualizations of corresponding example trajectories and network topologies are shown in Fig. 5.

- **Random waypoint** is used as an abstract reference scenario within the parameter selection process. In order to allow PARRoT to exploit the trajectory knowledge and in contrast to conventional implementations, all future waypoint locations are computed at the beginning of the simulation run.
- **Distributed Dispersion Detection (DDD)** [20] is an algorithm for coordinated swarm-based plume source exploration in disaster scenarios which ensures intra-swarm connectivity through communication-aware mobility maneuvers.
- **Dynamic cluster hovering** is a method for UAV-based network provisioning for hybrid vehicular networks initially presented in [21]. Hereby, multiple UAVs dynamically adjust their locations for providing network

coverage for clusters of ground-based vehicles. Within the evaluation scenario (which corresponds to the default scenario of the Lightweight ICT-centric Mobility Simulation (LIMoSim) mobility simulator [21]), a total number of 10 UAVs operates at a flying height of 30 m. From a total number of 50 cars, a random subset of 10 vehicles is chosen to be equipped with communication interfaces.

**Reference routing protocols:** In the next section, PARRoT is compared to established routing protocols which implement different routing philosophies. All of the considered methods are configured according to their default parameter specification for mobile networking.

- **AODV** is a well-established *reactive* routing protocol.
- **OLSR** is a *proactive* protocol which uses a path planning approach that involves information about the whole network topology. Moreover, the protocol utilizes so-called MPRs for minimizing the amount of broadcast messages within the routing message distribution process.
- **GPSR** is a *geo-based* routing method which bases its greedy routing process on minimizing the geo-distance to the destination in each packet forwarding step.
- **B.A.T.Mobile** is a *geo-predictive* extension to B.A.T.M.A.N. and represents an immediate groundwork for PARRoT.

According to the empirical analysis of [12], the first three protocol represent the most commonly used approaches for VANET routing. A summary about the simulation parameters is provided by Tab. I.

TABLE I  
DEFAULT SIMULATION PARAMETERS

Parameter	Value
OMNeT++ version	5.6.1
INETMANET version	4.x
MAC	IEEE 802.11g
Path loss exponent $\eta$	2.75
Channel model (Rural, Urban)	Friis, Nakagami (m=2)
Scenario size (General)	500 m × 500 m × 250 m
Scenario size (Cluster Hovering)	750 m × 600 m × 50 m
Number of runs per configuration	25
Simulation time	900 s
Number of routing hosts	10 s
Mobility update interval $\Delta t$	100 ms
Prediction horizon $\tau$	2.5 s
Waypoint radius $r_w$	10 m
Velocity $v$	50 km/h
Transport protocol	UDP
Traffic load per video stream	2 MBit/s
Learning rate $\alpha$	0.5
Basic discount factor $\gamma_0$	0.8
Chirp interval $\Delta t_{\text{chirp}}$	0.5 s

#### V. SIMULATION RESULTS

In this section, the results of the OMNeT++-based simulations are presented. All errorbars show the 0.95 confidence interval of the mean value over the different simulation runs. In the following, the impact of different parameters on the behavior of PARRoT is analyzed before the performance of the novel protocol is compared to established methods in defined reference scenarios.

<sup>1</sup>Source code available at <https://github.com/cedrikschueler/PARRoT>

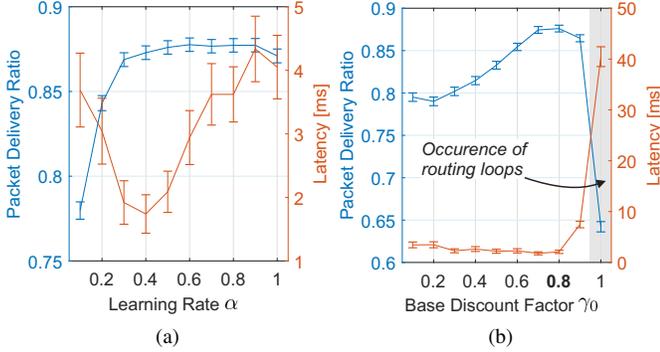


Fig. 6. Impact of the reinforcement learning parameters on the end-to-end performance of PARRoT.

### A. Parameterization of PARRoT

For all following parameter variations, the performance of PARRoT is analyzed using the rural channel model and random waypoint mobility.

**Learning parameters:** An evaluation of different values of the key reinforcement learning parameters of PARRoT is shown in Fig. 6. The learning rate  $\alpha$  corresponds to the information gain per received chirp message. Thus, it is dependent to the relative mobility of the agents and the chirp interval  $\Delta t_{\text{chirp}}$ . If  $\alpha$  is chosen too small, the agent fails to adopt its decision making to the dynamics of the network topology. As a consequence of the resulting choice of sub-optimal routing paths, the end-to-end delay is increased and the PDR is reduced. If  $\alpha$  is chosen too large, the impact of single chirp messages — which might be impacted by short-term effects such as local queuing — is overemphasized. In Fig. 6 (a), it can be seen that  $\alpha$  tolerates a certain grade of derivation from the optimal value without significantly reducing the end-to-end behavior of PARRoT.

The basic discount factor  $\gamma_0$ , which is shown in Fig. 6 (b), represents an implicit hop punishment for ensuring that the propagated reverse path quality to the originator decreases if the number of hops increases. Since the individual metrics that jointly form  $\gamma(j)$  (see Eq. 4) are multiplied with each other, it also acts as a scaling factor for the latter. Similar to  $\alpha$ ,  $\gamma_0$  needs to be chosen large enough such that the information gain converges with the network dynamics. For  $\gamma_0 = 1$ , the loop free routing condition  $\gamma(j) < 1$  cannot be guaranteed. As a consequence, a massive drop of the end-to-end routing performance can be observed.

**Mobility prediction:** For improving the communication robustness through consideration of the relative mobility of the agents, PARRoT uses mobility prediction for deriving estimates of the corresponding LETs. Due to this dependency, the error of the mobility prediction should be minimized. Fig. 7 shows the resulting errors for the considered mobility prediction schemes with respect to the prediction horizon  $\tau$ . As a reference, the behavior for a naive prediction method which assumes the vehicle position to stay constant is shown. In this case, the resulting error can be derived as  $e_{\text{naive}} = v \cdot \tau$ . In the worst case, the prediction points in the exact opposite direction

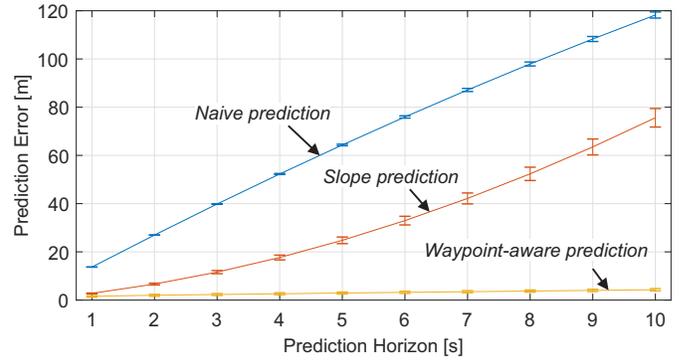


Fig. 7. Accuracy of different mobility prediction methods.

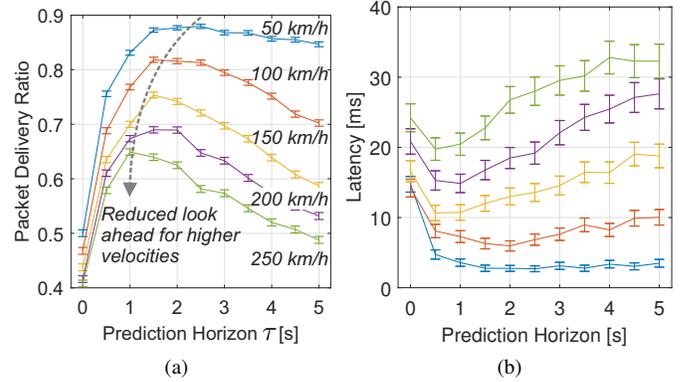


Fig. 8. Impact of the prediction horizon  $\tau$  on the behavior of PARRoT at different velocities.

of the actual movement which allows to derived an upper error bound  $e_{\text{max}} = 2 \cdot e_{\text{naive}}$ . It can be seen that the *waypoint-aware* prediction allows to provide comparable accurate prediction results. In contrast to the extrapolation-based *slope* method, it is able to integrate knowledge about the turning behavior of the vehicles into the prediction process. For completeness, it is remarked that the integration of waypoints also strengthens the robustness of the protocol against inaccurate location information [13]. As a consequence, PARRoT is configured to prefer the *waypoint-aware* mobility prediction method. If the mobility model does not provide waypoint information, the *slope* prediction is utilized as a fallback.

The impact of the prediction horizon  $\tau$  on the behavior of PARRoT for different velocities in the range of 50 km/h to 250 km/h is shown in Fig. 8. It can be seen that the link lifetime estimation of PARRoT is highly depending on the availability of mobility prediction results. In comparison to the non-predictive variant ( $\tau = 0$ ), the PDR is increased by up to 75 % if the protocol uses a meaningful prediction horizon. For higher speeds, the end-to-end routing performance decreases and smaller values of  $\tau$  should be preferred in order to adopt to the increased network dynamics. In addition, PARRoT becomes more sensitive to an optimal choice of  $\tau$ .

### B. Performance Comparison with Existing Routing Protocols

**Random mobility:** Fig. 9 show the performance of the considered routing methods using random waypoint mobility with the rural channel model. Due to the non-coordinated

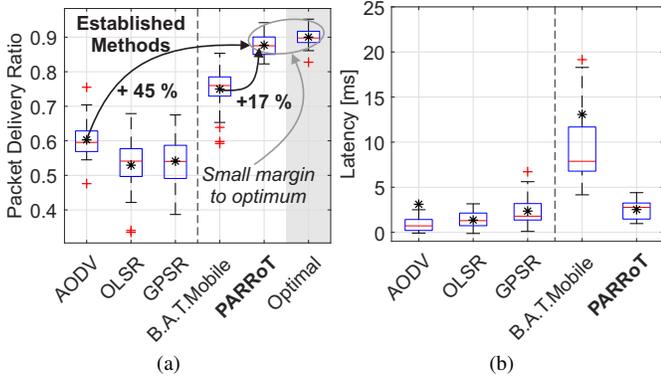


Fig. 9. Comparison PARRoT and existing routing methods. *Optimal* represents an upper topology-related bound for the path availability between sender and destination (Rural radio propagation model).

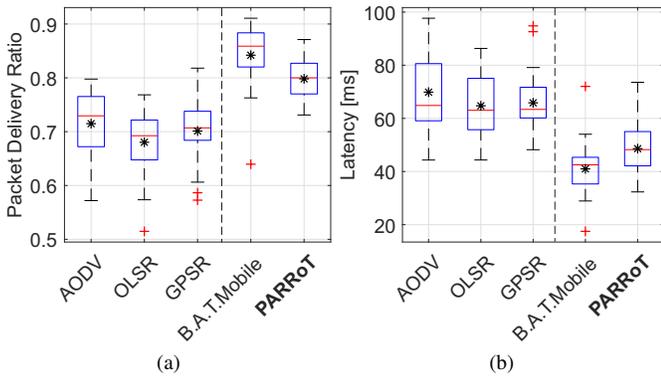


Fig. 10. Comparison PARRoT and existing routing methods (Urban radio propagation model).

motion, the existence of a routing path between sender and receiver is not always guaranteed. In order to pay attention to this aspect while classifying the performance of the different protocols, a theoretical upper bound — referred to as *Optimal* — for the PDR is provided based on post processing-based analysis of the network topology. However, it is remarked that this method is only able to consider the mobility-related aspects and does not account for load-related packet loss. In comparison to the well-established methods AODV, OLSR, and GPSR, PARRoT shows a massively higher — at least by 45 % — PDR which is close to the optimum robustness. Moreover, its reinforcement learning-based routing approach even outperforms the also prediction-based routing method B.A.T.Mobile by 17 %.

More challenging radio channel conditions are analyzed in Fig. 10. As the utilized channel model is subject to probabilistic effects, an upper bound for the PDR cannot be derived. Here, PARRoT shows a slightly lower PDR than B.A.T.Mobile. A plausible explanation for this observation is that the fast fading effects reduce the significance of the neighbor set coherence metric which then leads to sub-optimal routing decisions. Future extensions of PARRoT could explicitly address this issue through dynamic channel-dependent weighting the different metrics. All protocols suffer from a higher end-to-end delay due to an increased queuing time at the MAC layer due to sporadic link loss. The general tendency

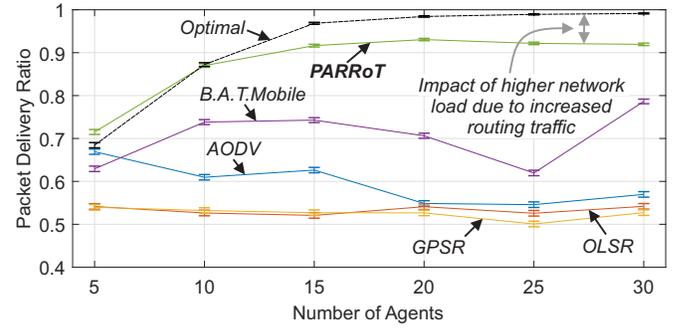


Fig. 11. Scalability analysis: Impact of the number of hosts on the end-to-end routing performance.

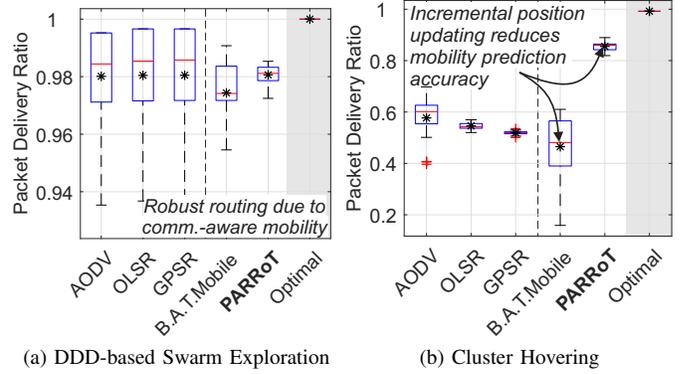


Fig. 12. Routing robustness in realistic mobility conditions (Rural channel model).

of these findings is also confirmed by the experiments of [22].

A scalability analysis of the considered routing protocols is shown in Fig. 11. As the size of the scenario is not varied, the increase of the number of agents corresponds to increasing the density of network nodes and the grade of routing opportunities. However, the higher amount of periodic routing traffic — especially for the proactive methods but also for neighbor discovery in reactive routing — leads to an increased probability for collision-related packet loss. It can be seen that the established routing methods fail to adapt to the network topology dynamics. Due to the high dependency of individual routing messages, they are unable to exploit the higher amount of possible routing paths. Due to the inherent maintenance of different routing paths per destination, PARRoT is robust against loss of individual chirp messages. In contrast to the reference protocols, it is able to leverage the increased network density for improving the robustness of the data transfer. However, for more than 20 active hosts, a slight decrease of the PDR can be observed.

**Application-driven mobility:** The behavior of the protocols in more realistic and application-specific mobility conditions is shown in Fig. 12. For the swarm mobility exploration method DDD, the corresponding results are visualized in Fig. 12 (a). Due to implemented *communication-aware* mobility approach which proactively adjusts the mobility behavior of the agents for ensuring swarm coherence, the PDR shows a high general level and does not vary significantly between the different routing protocols. Still, the predictive approach of PARRoT allows to proactively avoid some of the lower outlier values.

The results for the UAV-based cluster hovering analysis are

shown in Fig. 12 (b). This evaluation scenario is characterized by a complex network topology with a large number of mobile vehicles which show different mobility characteristics. In addition, the UAV cluster hovering is based on a ground traffic-related incremental position updating process which does not allow to accurately forecast future locations. As a consequence, the mobility-predictive routing methods are confronted with imprecise estimations for the relative mobilities. In this complex setting, B.A.T.Mobile is not able to maintain robust connectivity. In fact, the mobility predictions even reduce the routing performance of the protocol slightly beyond the PDR of the established methods. For completeness, it is remarked that these issues might be partially compensable through a scenario-specific parameter optimization. Although, PARRoT is also confronted with the same challenges, the resulting degradation of the routing performance is far less distinct. Due to the implemented multi-metric routing approach which does not only consider the link lifetime but also the neighbor set coherence, it is less vulnerable to imprecise mobility predictions. As a consequence, PARRoT is able to provide robust data delivery even in highly challenging hybrid vehicular networks.

## VI. CONCLUSION

In this paper, we presented the novel routing protocol PARRoT for highly mobile robotic networks which brings together mobility-predictive routing with reinforcement learning-based decision making. In a comprehensive simulation-based performance evaluation, it was shown that the consideration of the future relative mobility between the agents allows PARRoT to achieve robust and efficient data delivery even in challenging radio propagation conditions.

In future work, we will extend PARRoT for considering environment information within the link lifetime estimation process by integrating knowledge about the surrounding obstacles such as buildings. In addition, we will focus on analyzing the real world performance of the novel routing protocol.

## ACKNOWLEDGMENT

This work has been supported by the German Research Foundation (DFG) within the Collaborative Research Center SFB 876 "Providing Information by Resource-Constrained Analysis", projects A4 and B4 as well as by the German Federal Ministry of Education and Research (BMBF) in the project A-DRZ (13N14857).

## REFERENCES

- [1] Y. Zeng, Q. Wu, and R. Zhang, "Accessing from the sky: A tutorial on UAV communications for 5G and beyond," *Proceedings of the IEEE*, vol. 107, no. 12, pp. 2327–2375, 2019.
- [2] H. Menouar, I. Guvenc, K. Akkaya, A. S. Uluogac, A. Kadri, and A. Tuncer, "UAV-enabled intelligent transportation systems for the smart city: Applications and challenges," *IEEE Communications Magazine*, vol. 55, no. 3, pp. 22–28, March 2017.
- [3] B. Sliwa, T. Liebig, T. Vranken, M. Schreckenberg, and C. Wietfeld, "System-of-systems modeling, analysis and optimization of hybrid vehicular traffic," in *2019 Annual IEEE International Systems Conference (SysCon)*, Orlando, Florida, USA, Apr 2019.
- [4] M. Patchou, B. Sliwa, and C. Wietfeld, "Unmanned aerial vehicles in logistics: Efficiency gains and communication performance of hybrid combinations of ground and aerial vehicles," in *IEEE Vehicular Networking Conference (VNC)*, Los Angeles, USA, Dec 2019.
- [5] V. Chamola, V. Hassija, V. Gupta, and M. Guizani, "A comprehensive review of the COVID-19 pandemic and the role of IoT, drones, AI, blockchain, and 5G in managing its impact," *IEEE Access*, vol. 8, pp. 90 225–90 265, 2020.
- [6] N. Bui, M. Cesana, S. A. Hosseini, Q. Liao, I. Malanchini, and J. Widmer, "A survey of anticipatory mobile networking: Context-based classification, prediction methodologies, and optimization techniques," *IEEE Communications Surveys & Tutorials*, 2017.
- [7] J. Wang, C. Jiang, H. Zhang, Y. Ren, K. Chen, and L. Hanzo, "Thirty years of machine learning: The road to pareto-optimal wireless networks," *IEEE Communications Surveys Tutorials*, pp. 1–1, 2020.
- [8] S. Ali, W. Saad, N. Rajatheva, K. Chang, D. Steinbach, B. Sliwa, C. Wietfeld, K. Mei, H. Shiri, H.-J. Zepernick, T. M. C. Chu, I. Ahmad, J. Huusko, J. Suutala, S. Bhaduria, V. Bhatia, R. Mitra, S. Amuru, R. Abbas, B. Shao, M. Capobianco, G. Yu, M. Claes, T. Karvonen, M. Chen, M. Girnyk, and H. Malik, "6G white paper on machine learning in wireless communication networks," Apr 2020.
- [9] B. Sliwa and C. Wietfeld, "Data-driven network simulation for performance analysis of anticipatory vehicular communication systems," *IEEE Access*, Nov 2019.
- [10] O. S. Oubbati, M. Atiqzaman, P. Lorenz, M. H. Tareque, and M. S. Hossain, "Routing in flying ad hoc networks: Survey, constraints, and future challenge perspectives," *IEEE Access*, vol. 7, pp. 81 057–81 105, 2019.
- [11] R. A. Nazib and S. Moh, "Routing protocols for unmanned aerial vehicle-aided vehicular ad hoc networks: A survey," *IEEE Access*, vol. 8, pp. 77 535–77 560, 2020.
- [12] E. R. Cavalcanti, J. A. R. de Souza, M. A. Spohn, R. C. d. M. Gomes, and A. F. B. F. d. Costa, "VANETs' research over the past decade: Overview, credibility, and trends," *SIGCOMM Comput. Commun. Rev.*, vol. 48, no. 2, pp. 31–39, May 2018.
- [13] B. Sliwa, D. Behnke, C. Ide, and C. Wietfeld, "B.A.T.Mobile: Leveraging mobility control knowledge for efficient routing in mobile robotic networks," in *IEEE GLOBECOM 2016 Workshop on Wireless Networking, Control and Positioning of Unmanned Autonomous Vehicles (Wi-UAV)*, Washington D.C., USA, Dec 2016.
- [14] M. Song, J. Liu, and S. Yang, "A mobility prediction and delay prediction routing protocol for UAV networks," in *2018 10th International Conference on Wireless Communications and Signal Processing (WCSP)*, 2018, pp. 1–6.
- [15] Y. Tang, N. Cheng, W. Wu, M. Wang, Y. Dai, and X. Shen, "Delay-minimization routing for heterogeneous VANETs with machine learning based mobility prediction," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 4, pp. 3967–3979, 2019.
- [16] G. Oddi, D. Macone, A. Pietrabissa, and F. Liberati, "A proactive link-failure resilient routing protocol for MANETs based on reinforcement learning," in *2012 20th Mediterranean Conference on Control Automation (MED)*, 2012, pp. 1259–1264.
- [17] B. Sliwa, N. Piatkowski, and C. Wietfeld, "LIMITS: Lightweight machine learning for IoT systems with resource limitations," in *2020 IEEE International Conference on Communications (ICC)*, Dublin, Ireland, Jun 2020, Best paper award.
- [18] B. Coll-Perales, J. Gozalvez, and J. L. Maestre, "5G and beyond: Smart devices as part of the network fabric," *IEEE Network*, vol. 33, no. 4, pp. 170–177, July 2019.
- [19] A. Varga and R. Hornig, "An overview of the OMNeT++ simulation environment," in *Proceedings of the 1st International Conference on Simulation Tools and Techniques for Communications, Networks and Systems & Workshops*, ser. Simutools '08. ICST, Brussels, Belgium, Belgium: ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2008, pp. 60:1–60:10.
- [20] D. Behnke, P. Bök, and C. Wietfeld, "UAV-based connectivity maintenance for borderline detection," in *2013 IEEE 77th Vehicular Technology Conference (VTC Spring)*, 2013, pp. 1–6.
- [21] B. Sliwa, M. Patchou, and C. Wietfeld, "Lightweight simulation of hybrid aerial- and ground-based vehicular communication networks," in *2019 IEEE 90th Vehicular Technology Conference (VTC-Fall)*, Honolulu, Hawaii, USA, Sep 2019.
- [22] S. Saleh, M. Zuhairi, and H. Dao, "A comparative performance analysis of MANET routing protocols in various propagation loss models using ns3 simulator," *Journal of Communications*, pp. 537–544, 01 2020.