This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSC.2021.3103105, IEEE Transactions on Services Computing

Adaptive Data Dissemination Algorithm based on Storing-Discarding Equilibrium for OUSNs

Linfeng Liu, *Member, IEEE*, Zhiyuan Xi, Jiagao Wu, *Member, IEEE*, Jia Xu, *Member, IEEE*

***Abstract*—Opportunistic underwater sensor networks (OUSNs) are deployed for various underwater applications, such as underwater creatures tracking and tactical surveillance. The data dissemination in OUSNs differs significantly from those in terrestrial wireless sensor networks or delay-tolerant networks, due to the signal irregularity in underwater communications and the limited storage capacity of the nodes in OUSNs. To alleviate the storage overflows on nodes and make room for the newly arriving data packets, some stored data packets ought to be actively discarded by nodes. This research begins with the construction of a differential equation set to describe the propagation process of data packets in OUSNs, and the storing-discarding equilibrium is investigated such that each data packet is expected to propagate and disappear during the allowable dissemination time slots. After that, the optimal storing probabilities and discarding probabilities are obtained for the nodes with different in-degrees to maximize the delivery ratio of data packets. Then, we propose an Adaptive Data Dissemination Algorithm (ADDA) for the storage-limited OUSNs with signal irregularity, where at each time slot the newly arriving data packets are stored and the stored data packets are discarded by nodes according to the obtained storing probabilities and discarding probabilities, respectively. Simulation results demonstrate the excellent performance of ADDA, showing that it can enhance the delivery ratio of data packets and reduce the number of storage overflows.**

***Index Terms*—opportunistic underwater sensor network; data dissemination; storing-discarding equilibrium; delivery ratio; storage overflow.**

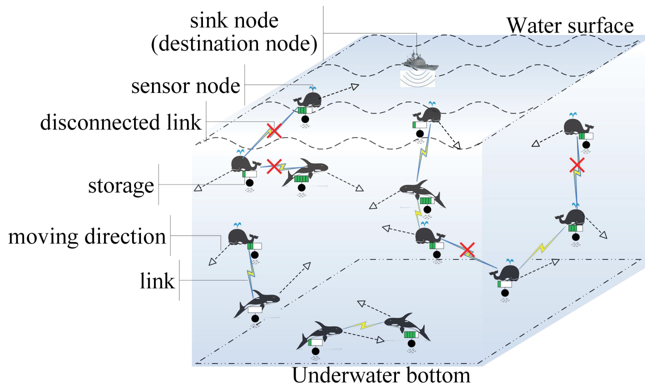


Fig. 1: Architecture of an OUSN.

Fig. 2: A biomimetic fish developed by CRIC (CNO Rapid Innovation Cell).

I. Introduction

The Opportunistic Underwater Sensor Networks (OUSNs) [1] technology enables various underwater applications, such as underwater creatures tracking [2] and tactical surveillance [3]. Due to the underwater mobility of nodes, the available contacts between nodes become scarce and short, and thus the data packets cannot be disseminated along stable paths. As shown in Fig. 1, environmental events are monitored by the sensors fastened on mobile underwater vehicles (such as the biomimetic fish in Fig. 2) and are encapsulated into some data packets, which are then transmitted to the sink node through intermittent communication links.

L. Liu, Z. Xi, J. Wu, and J. Xu are with the School of Computer Science and Technology, Nanjing University of Posts and Telecommunications, Nanjing 210023, China, and also with the Jiangsu Key Laboratory of Big Data Security and Intelligent Processing, Nanjing 210023, China. Email: {liulf,jgwu,xujia}@njupt.edu.cn, Xizy186@126.com.
(Corresponding author: Jia Xu.)

Due to the intermittent communication links, the deliveries of data packets from source nodes to sink node (destination node) cannot be guaranteed even though the epidemic forwarding manner [4] (each node disseminates the stored data packets to all the encountered nodes) is adopted for the data dissemination, and thus the improvement of delivery ratio remains an important issue in OUSNs.

Signal irregularity is a common phenomenon in OUSNs [5]. Signal irregularity is caused by various factors, such as antenna directions, transmitting power, antenna gains, battery status, signal-noise-ratio threshold and obstacles. Typically, a variety of obstacles are distributed in underwater environments, which makes the underwater signal easy to be reflected, diffracted, or scattered, and hence some communication links between nodes are potentially disconnected.

Besides, the data packets containing environmental events are generated frequently in OUSNs, and then are further copied and disseminated among all the nodes, which leads to a large number of storage overflows on nodes* in storage-limited OUSNs (the storage memory of nodes is usually measured in KB) [6], [7]. The storage overflows on nodes will prevent the newly arriving data packets from being stored. Thus, the propagations of these data packets are impaired, and the delivery ratio of data packets is depressed.

The signal irregularity phenomenon could make some communication links disconnected, and the data packets cannot be disseminated on the disconnected communication links. Besides, the limited storage capacity of nodes restricts the amount of data packets that can be stored by nodes, and some data packets stored by nodes will be discarded when the storage of the nodes is full. These phenomena could hinder the propagation process of data packets and reduce the delivery ratio of data packets seriously. Motivated by the above considerations, we conduct a study about the data dissemination problem in the storage-limited OUSNs with signal irregularity.

Naturally, the proper strategies of storing and discarding the data packets can be explored through formulating and analysing the propagation process of data packets. The main contributions of this paper are threefold. First, the propagation process of data packets is formulated by a differential equation set, and the storing-discarding equilibrium is specially investigated to improve the delivery ratio of data packets and reduce the number of storage overflows on nodes. Second, an Adaptive Data Dissemination Algorithm (ADDA) is proposed, and the optimal storing probabilities and discarding probabilities of the nodes with different in-degrees are derived from the storing-discarding equilibrium. Last but not least, ADDA is completely distributed with an acceptable complexity, i.e., each node makes the storing decision and discarding decision in a distributed manner. Thus, ADDA is practical and with a good scalability.

The remainder of this paper is organized as follows: Section II briefly surveys some existing related studies. Section III provides a problem formulation for the data dissemination in OUSNs. Section IV gives some analyses on how to achieve a storing-discarding equilibrium and obtain the optimal storing probabilities and discarding probabilities. Section V presents an Adaptive Data Dissemination Algorithm (ADDA), and then covers some further analyses on ADDA including the complexity and the expected delivery ratio. Simulation results for performance evaluation of ADDA are reported in Section VI. Finally, Section VII concludes the paper.

II. Related Work

Extensive studies have been carried out on the problem of data dissemination for Delay-Tolerant Networks (DTNs) or Wireless Sensor Networks (WSNs). The early representative algorithm proposed in [4] is Epidemic Forwarding (EF), where random pair-wise exchanges of data packets among mobile nodes ensure the maximum delivery ratio and the minimum propagation delay. However, numerous redundant data packet copies are generated in the transmissions. Niu *et al.* [8] conduct a theoretical analysis on the aggregated throughput capacity of opportunistic routing in WSNs with the consideration of lossy links and transmission fairness, and introduce the concept of the concurrent schedulable set to represent the constraints imposed by transmission conflicts, and then propose the Opportunistic Routing (OR). Reference [9] provides a smart insect pest detection technique with qualified underground wireless sensor nodes. The received signal strength and path loss parameters are examined to evaluate the performance of the underground network structure. In [10], the unlicensed users in idle mode are allowed to help sensor nodes as a cooperative relay, and the sensor nodes help unlicensed users for detecting the idle frequency bands while in sleep mode.

Some relevant research has been conducted on the data dissemination in OUSNs. In [11], a generic prediction assisted single-copy routing scheme which can be configured for multiple mobility models is investigated. This scheme differentiates the network mobility patterns according to the short-duration traces, and then defines the features of the best routing paths. Its outstanding advantage is the self-adaptivity for various node mobilities. However, it relies on the historical trajectories to instantiate the prediction. Such historical trajectories are difficult to be exploited in OUSNs due to the sophisticated mobility of nodes. Reference [12] proposes the Redundancy Based Adaptive Routing (RBAR) for underwater sensor networks. RBAR satisfies different delay requirements by explicitly controlling the replication process, but the delivery ratio is not very desirable. Zhang *et al.* develop a Beam width and Direction Concerned Routing protocol (BDCR) for the data dissemination problem [13], which can obtain a high delivery ratio by considering the beam width and three-dimensional direction. Recently, a routing protocol Mobile Sink (MobiSink) for underwater sensor networks is presented in [14] to balance the loads on intermediate nodes through deploying some mobile sinks in four horizontal regions. Obviously, the cost of these mobile sinks is extremely expensive, and thereby the availability of MobiSink is restricted. In [15], a data routing framework where the senders select the best forwarding relay is given, and then a data forwarding method is configured to consider channel conditions and route-wide residual energy.

Reference [16] proposes the GEographic and opportunistic routing with Depth Adjustment-based topology control Routing protocol (GEDAR). GEDAR is an any-cast, geographic and opportunistic routing protocol that routes data packets from nodes to sonobuoys deployed at the sea's surface. When a node is in a communication void region, GEDAR tries to recover it through the depth adjustments of other nodes. An opportunistic routing framework considering the characteristics of underwater sensor networks, such as the network density, traffic load, underwater environment and acoustic channels is proposed

*Storage overflow indicates that the storage of a node has been full, and the node cannot provide enough space to store the newly arriving data packets.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSC.2021.3103105, IEEE Transactions on Services Computing

TABLE I: Comparisons of Related Works

Algorithm	Feature	Performance	Scenario
EF [4]	Epidemic forwarding	High delivery ratio, large message complexity	Small-scale opportunistic networks
OR [8]	Lossy links, transmission fairness	Large throughput	WSNs with different traffic demands
Smart insect pest detection [9]	Measurement of received signal strength and path loss	Reveal of increase in path loss	WSNs for precision agriculture
Amplify & forward based cooperative communication [10]	Cooperative communication	Large throughput, low energy consumption and delay	WSNs
Generic prediction assisted single-copy routing [11]	Mobility pattern differentiation	Self-adaptivity for various mobilities	OUSNs with various mobilities
RBAR [12]	Control of replication process	Self-adaptivity for delay requirements	OUSNs with different delay requirements
BDCR [13]	Consideration of beam width and three-dimensional direction	High delivery ratio	OUSNs
MobiSink [14]	Balance the loads on intermediate nodes	Long network lifetime, and large throughput	UWSNs with mobile sinks
HyDRO [15]	Consideration of residual energy and foreseeable harvestable energy	Low energy consumption and latency, high delivery ratio	Energy harvesting-enabled underwater sensor networks
GEDAR [16], [17]	Recovery of communication void regions	Simple and scalable geographic routing protocol	Various OUSN scenarios of network density and traffic load
AREP [18]	Asymmetric link-based reverse routing	Short transmission delay, and high delivery ratio	Underwater acoustic sensor networks
Mobility assisted Geo-opportunistic routing [19]	Exploitation of geographic information	Long network lifetime, and high delivery ratio	Underwater wireless sensor networks with mobile sinks
TORA [20]	Avoidance of horizontal transmission	Long network lifetime, and high delivery ratio	Sparse and dense network scenarios
Automatic and intelligent monitoring system [21]	Pollution control	Effective cost	Underwater pollution monitoring

in [17]. In [18], an asymmetric link-based reverse routing is designed to ensure the bidirectional communications from source nodes to destination nodes. In this method, each node maintains a neighbor table in which the table items are utilized to analyze the link states, and the routing paths are established by prioritizing the utilization of symmetric links. Ahmed $et\ al.$ propose the mobility assisted geo-opportunistic routing paradigm based on interference avoidance for OUSNs [19], and the network volume is divided into logical small cubes to reduce the interference and make more informed routing decisions for efficient energy consumption. In [20], an anycast, geographical and Totally Opportunistic Routing Algorithm (TORA) for OUSNs is proposed to avoid horizontal transmission, reduce end to end delay, and maximize network throughput efficiency. In [21], a simulation model is proposed to define the intelligent sensor-based monitoring system that identifies and alarms the formation of underwater pollution. The main differences of these related works are listed in TABLE I.

Nevertheless, the issue of storage overflows in storage-limited networks is not considered in the aforementioned works, which reduces the delivery ratio of data packets in storage-limited OUSNs. Moreover, some communication links are likely to be disconnected due to the signal irregularity, which hinder the propagation process of data packets, and hence the issue of signal irregularity should be taken into account as well. In this paper, the data propagation process are first formulated and analyzed. To alleviate the storage overflows, some stored data packets are allowed to be actively discarded to reduce the number of storage overflows on nodes and make room for the newly arriving data packets, which can be realized through achieving the storing-discarding equilibrium in data dissemination. Then, the optimal storing probabilities and discarding probabilities are obtained to maximize the delivery ratio of data packets.

III. Problem Formulation

In this section, we describe the OUSN model and present the problem formulation of data dissemination.

A. OUSN Model

Suppose that N mobile nodes are uniformly deployed in a convex space $\boldsymbol{D}$, $\boldsymbol{D} \in \mathbb{R}^3$. The time is divided into discrete time slots with an equal length τ_{s}. Each data packet needs to be delivered from source node to destination node during the allowable dissemination time slots (t^* time slots).

The communication range and storage size of each node are denoted by R_{c} and S, respectively. At each time slot, each node generates a new data packet (the size of each data packet is L_{s}) with the probability ρ. The coordinate of the node V_i at the t-th time slot is denoted by $C(i)^{(t)}$.

The distance between two nodes V_i and V_j at the t-th time slot is denoted by $d(i,j)^{(t)}$. If $d(i,j)^{(t)} \leq R_{\mathrm{c}}$, and then $(i,j)^{(t)}$ is taken as a potential link. Due to the signal irregularity in underwater communications, the existence of a potential link $(i,j)^{(t)}$ is determined by the probability $P(i,j)^{(t)}$. As introduced in [22], [23], there is a power-function relation between the transmission distance and path loss (when the absorption coefficient is equal to 1.0),

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSC.2021.3103105, IEEE Transactions on Services Computing

and hence we give the following expression of $P(i,j)^{(t)}$:

$$P(i,j)^{(t)} = \begin{cases} 0, & if\ d(i,j)^{(t)} > R_c, \\ c_1 \cdot \Omega(j)^{-\zeta} \cdot d(i,j)^{(t)^{-\eta}}, & otherwise, \end{cases} \quad (1)$$

where c_1 is a constant, and $\Omega(j)$ denotes the signal irregularity around V_j which is caused by various factors, such as antenna directions, antenna gains, battery status, signal-noise-ratio threshold and obstacles [5]. Equation (1) implies that the existence probability of a potential link is decreased with the increase of link length or the increase of signal irregularity. ζ and η are the exponents reflecting the impacts of signal irregularity and link length on this existence probability, respectively.

An example of signal irregularity due to obstacles is given in Fig. 3, which indicates that $\Omega(j)$ will become larger when there are more obstacles around V_j. Since the underwater environment is extremely sophisticated, we assume the signal irregularity around each node obeys a uniform distribution $U(\Omega_{min}, \Omega_{max})$ [24]. Thus, the in-degrees of nodes follow a power law distribution.

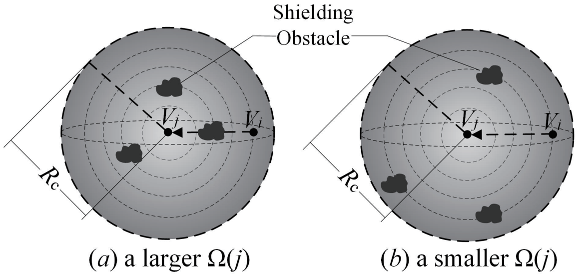


Fig. 3: Signal irregularity.

Besides, the underwater mobility pattern of nodes is comprised of autonomous movement and external movement, as introduced in our early work [25].

B. Data Propagation Formulation

To formulate the propagation process of data packets in OUSNs, the explanations of main notations are first presented in TABLE II.

The propagation process of a data packet $data(s,d)$ is illustrated in Fig. 4, where $data(s,d)$ needs to be disseminated from the source node V_s to the destination node V_d. Some nodes could store and relay $data(s,d)$, and the stored $data(s,d)$ could be discarded by some nodes as well. With regard to each data packet, there are three states for each node: (i) $Initial\ state$. Each node except the source node is in the initial state at the birth of the data packet; (ii) $Data\ holding\ state$. The node having received and stored a data packet is in the data holding state; (iii) $Data\ discarding\ state$. The node having discarded the stored data packet is in the data discarding state. The three states are depicted in Fig. 5.

The in-degree of a node indicates the likelihood of receiving a data packet, i.e., a node with a larger in-degree will receive a data packet more easily. Each node with the in-degree k is assumed to store the received data packets

TABLE II: Main Notations

Parameter	Description
$I_k(t)$	Proportion of initial state nodes (with in-degree k) at the t-th time slot
$H_k(t)$	Proportion of data holding nodes (with in-degree k) at the t-th time slot
$H(t)$	Proportion of data holding nodes at the t-th time slot
$D_k(t)$	Proportion of data discarding nodes (with in-degree k) at the t-th time slot
$D(t)$	Proportion of data discarding nodes at the t-th time slot
v_k	Storing probability of nodes with in-degree k
$\langle v \rangle$	Expectation of storing probabilities
μ_k	Discarding probability of nodes with in-degree k
$\langle \mu \rangle$	Expectation of discarding probabilities
$\langle k \rangle$	Expectation of in-degrees of nodes
$P(k)$	Proportion of nodes with in-degree k
$\mathcal{R}(t)$	Proportion of data packets delivered during t time slots
$\mathcal{A}(t)$	Probability of data packets delivered at the t-th time slot

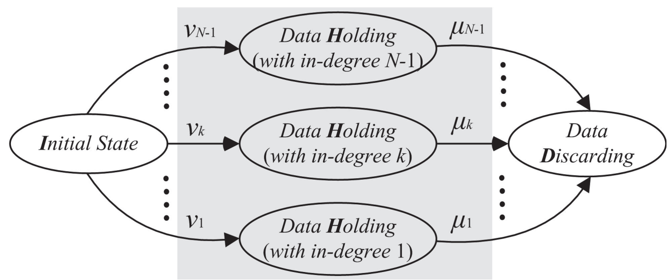


Fig. 5: State transitions of nodes.

with the probability v_k and discard the stored data packets with the probability μ_k. To analyze the optimal storing probabilities and discarding probabilities of the nodes with different in-degrees, the propagation process of each data packet over time is formulated as a differential equation set:

$$\begin{cases} I_k(t) + H_k(t) + D_k(t) = 1, \\ \frac{\mathbf{d}I_k(t)}{\mathbf{d}t} = -k \cdot v_k \cdot I_k(t) \cdot \theta(t), \\ \frac{\mathbf{d}H_k(t)}{\mathbf{d}t} = k \cdot v_k \cdot I_k(t) \cdot \theta(t) - \mu_k \cdot H_k(t), \\ \frac{\mathbf{d}D_k(t)}{\mathbf{d}t} = \mu_k \cdot H_k(t), \end{cases} \quad (2)$$

where the in-degree k falls into interval $[0, N-1]$. $\theta(t) = \frac{\sum_{k=1}^{N-1} k \cdot P(k) \cdot H_k(t)}{\langle k \rangle}$, which indicates the probability of a communication link with at least one end carrying the data packet, and $\langle k \rangle = \sum_{k=1}^{N-1} k \cdot P(k)$. Thus, $k \cdot v_k \cdot I_k(t) \cdot \theta(t)$ denotes the proportion variation of initial state nodes (with in-degree k) at the t-th time slot, and $\mu_k \cdot H_k(t)$ denotes the proportion variation of data discarding nodes (with in-degree k) at the t-th time slot.

Recall that $P(k)$ follows a power law distribution, let $P(k) = \alpha \cdot k^{-\gamma}$, where $\gamma = \frac{1}{\zeta} + 1$ $(2 < \gamma \leq 3)$ and $\alpha = \frac{1}{\sum_{k=1}^{N-1} k^{-\gamma}}$. Besides, the expectations of μ_k and v_k are computed as:

$$\begin{cases} \langle \mu \rangle = \sum_{k=1}^{N-1} \mu_k \cdot P(k), \\ \langle v \rangle = \sum_{k=1}^{N-1} v_k \cdot P(k). \end{cases}$$

At each time slot, $N \cdot \rho$ new data packets are generated in an OUSN. The total number of stored data packets (copies) is restricted by the storage limitation of nodes. To alleviate

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSC.2021.3103105, IEEE Transactions on Services Computing

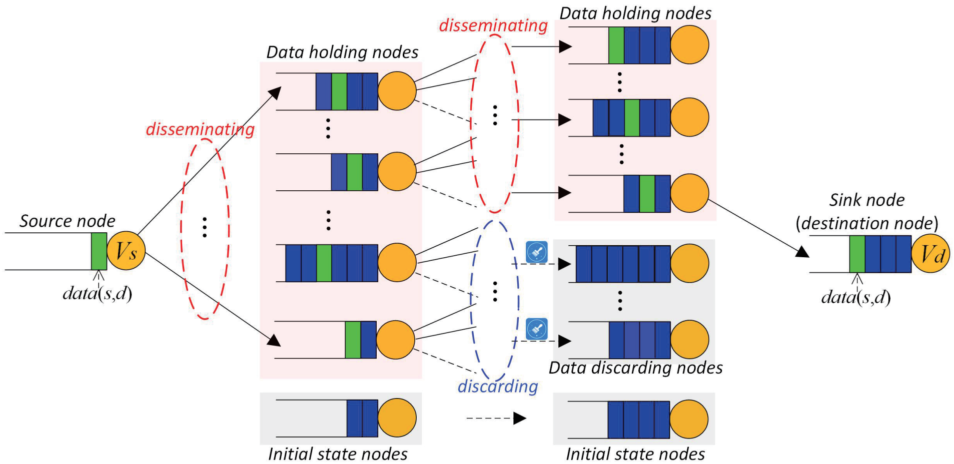


Fig. 4: Propagation process of a data packet.

the storage overflows, the proportion of data holding nodes for each data packet is expected to gradually reduce to 0 by the final time slot (the t^*-th time slot). Thus, the objective function of delivery ratio enhancement and storage overflow alleviation is formulated as:

$$\max \mathcal{R}(t^*),$$
$$s.t. \begin{cases} H_k(t^*) = 0, \quad \forall\, k = 1, \ldots, N-1, \\ N \cdot \rho \cdot \sum_{t=0}^{t^*-1} H(t) \leq \frac{S}{L_s}, \end{cases} \tag{3}$$

where $\mathcal{R}(t^*)$ denotes the delivery ratio of data packets, i.e., the proportion of data packets which can be delivered to the destination node during t^* time slots. The second constraint of (3) indicates that the number of data packets (copies) should be limited by the total storage capacity of all nodes. Note that the values of $\mathcal{R}(t)$ and $H(t)$ are determined by the storing probabilities and discarding probabilities of nodes. The objective function (3) is to obtain the optimal storing probabilities and discarding probabilities of nodes so that the delivery ratio of data packets is maximized while a storing-discarding equilibrium is achieved.

IV. Problem Analysis

In this section, we present an analysis regarding the data propagation process to achieve a storing-discarding equilibrium, and then obtain the optimal storing probabilities and discarding probabilities of nodes.

A. Analysis of Propagation Process

We first construct the expression of $D(t)$ as $D(t) = \frac{\sum_{k=1}^{N-1} k \cdot P(k) \cdot D_k(t)}{\langle k \rangle}$, and the first-order derivative of $D(t)$ with respect to t is expressed as:

$$\begin{aligned}\frac{\mathbf{d}D(t)}{\mathbf{d}t} &= \frac{\sum_{k=1}^{N-1} k \cdot P(k) \cdot \frac{\mathbf{d}D_k(t)}{\mathbf{d}t}}{\langle k \rangle} = \frac{\sum_{k=1}^{N-1} k \cdot P(k) \cdot \mu_k \cdot H_k(t)}{\langle k \rangle} \\ &\approx \langle \mu \rangle \cdot \frac{\sum_{k=1}^{N-1} k \cdot P(k) \cdot H_k(t)}{\langle k \rangle} = \langle \mu \rangle \cdot \theta(t) \\ &= \langle \mu \rangle \cdot \frac{\sum_{k=1}^{N-1} k \cdot P(k) \cdot [1 - I_k(t) - D_k(t)]}{\langle k \rangle},\end{aligned}$$

where $I_k(t)$ is expressed in (4), because the second sub-equation of (2) is a Bernoulli equation, and thus we have that:

$$I_k(t) = [1 - H_k(0)] \cdot e^{-k \cdot \frac{v_k}{\langle \mu \rangle} \cdot D(t)} \approx e^{-k \cdot \frac{v_k}{\langle \mu \rangle} \cdot D(t)}, \tag{4}$$

where the approximation is made since it is obviously $I_k(0) \approx 1$, $H_k(0) \approx 0$, and $D_k(0) \approx 0$.

When N is large enough, $\frac{\mathbf{d}D(t)}{\mathbf{d}t}$ can be rewritten as:

$$\begin{aligned}\frac{\mathbf{d}D(t)}{\mathbf{d}t} &= \langle \mu \rangle \cdot \left[1 - D(t) - \frac{\sum_{k=1}^{N-1} k \cdot P(k) \cdot e^{-k \cdot \frac{v_k}{\langle \mu \rangle} \cdot D(t)}}{\langle k \rangle}\right] \\ &\approx \langle \mu \rangle \cdot \left[1 - D(t) - \frac{\alpha \cdot \int_0^{N-1} e^{-k \cdot \frac{v_k}{\langle \mu \rangle} \cdot D(t)} \mathbf{d}k}{\langle k \rangle^{\gamma}}\right] \\ &\approx \langle \mu \rangle \cdot \left[1 - D(t) + \frac{\alpha \cdot \langle \mu \rangle}{\langle k \rangle^{\gamma} \cdot \langle v \rangle} \cdot \frac{e^{-\frac{\langle v \rangle \cdot (N-1)}{\langle \mu \rangle} \cdot D(t)} - 1}{D(t)}\right] \\ &\approx \langle \mu \rangle \cdot \left[1 - D(t) - \frac{\alpha \cdot \langle \mu \rangle}{\langle k \rangle^{\gamma} \cdot \langle v \rangle \cdot D(t)}\right].\end{aligned} \tag{5}$$

$H(t) = \sum_{k=1}^{N-1} P(k) \cdot H_k(t)$ denotes the proportion of data holding nodes at the t-th time slot. According to objective function (3), each data packet is expected to be disseminated and discarded adaptively during the allowable dissemination time slots through setting appropriate storing probabilities and discarding probabilities for nodes. Especially, note that a larger delivery ratio can be obtained with a larger $H(t)$, while $H(t)$ should gradually reduce to 0 by the final time slot of each data packet. To this end, the value of $\frac{\mathbf{d}H(t)}{\mathbf{d}t}$ is analyzed to infer the value variation of $H(t)$, as described in Theorem 1.

Theorem 1: The proportion of data holding nodes can first increase and then decrease over time slots by setting appropriate storing probabilities and discarding probabilities of nodes.

Proof: Because $\frac{\mathbf{d}H(t)}{\mathbf{d}t} = \sum_{k=1}^{N-1} P(k) \cdot \frac{\mathbf{d}H_k(t)}{\mathbf{d}t}$. By the third sub-equation of (2), we have

$$\begin{aligned} H_k(t) &= e^{-\mu_k \cdot t}\left[\int k \cdot v_k \cdot I_k(t) \cdot \theta(t) \cdot e^{\mu_k \cdot t}\mathbf{d}t + c_2\right] \\ &= e^{-\mu_k \cdot t}\left[\frac{k \cdot v_k}{\langle \mu \rangle} \cdot \int e^{-k \cdot \frac{v_k}{\langle \mu \rangle} \cdot D(t)} \cdot \frac{\mathbf{d}D(t)}{\mathbf{d}t} \cdot e^{\mu_k \cdot t}\mathbf{d}t + c_2\right] \\ &\approx -e^{-k \cdot \frac{v_k}{\langle \mu \rangle} \cdot D(t)} + c_2 \cdot e^{-\mu_k \cdot t}. \end{aligned} \tag{6}$$

As a special case, we have $H_k(0) \approx 0$, which implies that $c_2 \approx 1$. Hence, $\frac{\mathbf{d}H_k(t)}{\mathbf{d}t}$ is further derived as:

$$\frac{\mathbf{d}H_k(t)}{\mathbf{d}t} = \frac{k \cdot v_k}{\langle \mu \rangle} \cdot e^{-\frac{k \cdot v_k}{\langle \mu \rangle} \cdot D(t)} \cdot \frac{\mathbf{d}D(t)}{\mathbf{d}t} - \mu_k \cdot e^{-\mu_k \cdot t}. \tag{7}$$

Then, by (5) and (7) $\frac{\mathbf{d}H(t)}{\mathbf{d}t}$ is expressed as:

$$\begin{aligned} \frac{\mathbf{d}H(t)}{\mathbf{d}t} &= \sum_{k=1}^{N-1} \frac{\alpha}{k^\gamma} \cdot \frac{\mathbf{d}H_k(t)}{\mathbf{d}t} \\ &= \sum_{k=1}^{N-1} \frac{\alpha}{k^\gamma} \cdot \left[\frac{k \cdot v_k}{\langle \mu \rangle} \cdot \frac{\mathbf{d}D(t)}{\mathbf{d}t} \cdot e^{-\frac{k \cdot v_k}{\langle \mu \rangle} \cdot D(t)} - \mu_k \cdot e^{-\mu_k \cdot t}\right] \\ &= \alpha \cdot \sum_{k=1}^{N-1}\left[\frac{v_k}{k^{\gamma-1} \cdot \langle \mu \rangle} \cdot \frac{\mathbf{d}D(t)}{\mathbf{d}t} \cdot e^{-\frac{k \cdot v_k}{\langle \mu \rangle} \cdot D(t)} - \frac{\mu_k}{k^\gamma} \cdot e^{-\mu_k \cdot t}\right] \\ &\approx \alpha \cdot \langle \mu \rangle \cdot \left[\frac{1 - D(t) - \frac{\alpha \cdot \langle \mu \rangle}{\langle k \rangle^\gamma \cdot \langle v \rangle} \cdot D(t)^{-1}}{\langle k \rangle^{\gamma-1} \cdot D(t)} - \frac{e^{-\langle \mu \rangle \cdot t}}{\gamma - 1}\right], \end{aligned}$$

where

$$\frac{1 - D(t) - \frac{\alpha \cdot \langle \mu \rangle}{\langle k \rangle^\gamma \cdot \langle v \rangle} \cdot D(t)^{-1}}{D(t)} = \frac{\langle k \rangle^\gamma \cdot \langle v \rangle}{4\alpha \cdot \langle \mu \rangle} - 1 - \left(\sqrt{\frac{\alpha \cdot \langle \mu \rangle}{\langle k \rangle^\gamma \cdot \langle v \rangle}} D(t)^{-1} - \frac{1}{2\sqrt{\frac{\alpha \cdot \langle \mu \rangle}{\langle k \rangle^\gamma \cdot \langle v \rangle}}}\right)^2.$$

Typically, there is $D(t) \geq 2\frac{\alpha \cdot \langle \mu \rangle}{\langle k \rangle^\gamma \cdot \langle v \rangle}$, and there will be $\frac{\mathbf{d}H(t)}{\mathbf{d}t} < 0$ when $D(t) > \frac{\sqrt{\frac{\alpha \cdot \langle \mu \rangle}{\langle k \rangle^\gamma \cdot \langle v \rangle}}}{\frac{1}{2\sqrt{\frac{\alpha \cdot \langle \mu \rangle}{\langle k \rangle^\gamma \cdot \langle v \rangle}}} - \sqrt{\frac{\langle k \rangle^\gamma \cdot \langle v \rangle}{4\alpha \cdot \langle \mu \rangle} - 1 - \frac{e^{-\langle \mu \rangle \cdot t} \cdot \langle k \rangle^{\gamma-1}}{\gamma - 1}}}$.

Because the value of $D(t)$ will keep increasing over time slots. Therefore, $H(t)$ will first increase, and then decrease after the value of $D(t)$ having reached the numerical value:

$$\frac{\frac{\alpha \cdot \langle \mu \rangle}{\langle k \rangle^\gamma \cdot \langle v \rangle}}{\frac{1}{2} - \sqrt{\frac{\alpha \cdot \langle \mu \rangle}{\langle k \rangle^\gamma \cdot \langle v \rangle} \cdot \left\{\frac{\langle k \rangle^\gamma \cdot \langle v \rangle}{4\alpha \cdot \langle \mu \rangle} - 1 - \frac{e^{-\langle \mu \rangle \cdot t} \cdot \langle k \rangle^{\gamma-1}}{\gamma - 1}\right\}}}. \ \square$$

The conclusion of Theorem 1 validates the feasibility of the data dissemination manner where the proportion of data holding nodes for each data packet first increases and then decreases over time slots, and this manner can be realized through setting appropriate storing probabilities and discarding probabilities of nodes. In the following sub-sections, the storing-discarding equilibrium is analyzed to reduce the number of storage overflows, and then the optimal storing probabilities and discarding probabilities of nodes with different in-degrees are obtained for the delivery ratio enhancement.

B. Storing-Discarding Equilibrium

The two constraints in objective function (3) is further analyzed in this section to achieve a storing-discarding equilibrium:

(I) The first constraint $H_k(t^*) = 0$ ($\forall\, k = 0, \ldots, N-1$). The first constraint expresses that the proportion of data holding nodes for each data packet is expected to gradually reduce to 0 by the final time slot, which implies that $\frac{\mathbf{d}H_k(t^*)}{\mathbf{d}k} = 0$. Besides, the first-order derivative of $H_k(t)$ with respect to k is expressed as:

$$\frac{\mathbf{d}H_k(t)}{\mathbf{d}k} = \frac{D(t)}{\langle \mu \rangle} \cdot \left(v_k + k \cdot \frac{\mathbf{d}v_k}{\mathbf{d}k}\right) \cdot e^{-\frac{k \cdot v_k}{\langle \mu \rangle} \cdot D(t)} - t \cdot \frac{\mathbf{d}\mu_k}{\mathbf{d}k} \cdot e^{-\mu_k \cdot t}. \tag{8}$$

According to expressions $H_k(t^*) = 0$ and $\frac{\mathbf{d}H_k(t^*)}{\mathbf{d}k} = 0$, we obtain the following equations:

$$\begin{cases} -e^{-k \cdot \frac{v_k}{\langle \mu \rangle} \cdot D(t^*)} + e^{-\mu_k \cdot t^*} = 0, \\ \frac{D(t^*)}{\langle \mu \rangle} \cdot \left(v_k + k \cdot \frac{\mathbf{d}v_k}{\mathbf{d}k}\right) \cdot e^{-\frac{k \cdot v_k}{\langle \mu \rangle} \cdot D(t^*)} = t^* \cdot \frac{\mathbf{d}\mu_k}{\mathbf{d}k} \cdot e^{-\mu_k \cdot t^*}, \end{cases}$$

which yields that:

$$\frac{D(t^*)}{\langle \mu \rangle} \cdot \left(v_k + k \cdot \frac{\mathbf{d}v_k}{\mathbf{d}k}\right) = t^* \cdot \frac{\mathbf{d}\mu_k}{\mathbf{d}k}. \tag{9}$$

(II) The second constraint $\sum_{t=0}^{t^*-1} H(t) \leq \frac{S}{N \cdot \rho \cdot L_s}$. Firstly, $\sum_{t=0}^{t^*-1} H(t)$ can be rewritten as:

$$\sum_{t=0}^{t^*-1} H(t) = \sum_{t=0}^{t^*-1}\sum_{k=1}^{N-1} P(k) \cdot H_k(t) = \sum_{k=1}^{N-1}\sum_{t=0}^{t^*-1} P(k) \cdot H_k(t).$$

When N is large enough, we have:

$$\begin{aligned} \sum_{t=0}^{t^*-1} P(k) \cdot H_k(t) &= \sum_{t=0}^{t^*-1} \frac{\alpha}{k^\gamma}\left[e^{-\mu_k \cdot t} - e^{-k \cdot \frac{v_k}{\langle \mu \rangle} \cdot D(t)}\right] \\ &\approx \frac{\alpha}{k^\gamma} \cdot \int_{t=0}^{t^*}\left[e^{-\mu_k \cdot t} - e^{-k \cdot \frac{v_k}{\langle \mu \rangle} \cdot D(t)}\right]\mathbf{d}t \\ &\leq \frac{\alpha}{k^\gamma} \cdot \left[\frac{1 - e^{-\mu_k \cdot t^*}}{\mu_k} - e^{-k \cdot \frac{v_k}{\langle \mu \rangle} \cdot D(t^*)} \cdot t^*\right] \\ &= \frac{\alpha}{k^\gamma} \cdot \left[\frac{1}{\mu_k} - \left(\frac{1}{\mu_k} + t^*\right) \cdot e^{-\mu_k \cdot t^*}\right], \end{aligned}$$

which gives rise to the following inequality:

$$\begin{aligned} \sum_{t=0}^{t^*-1} H(t) &\leq \int_1^{N-1} \frac{\alpha}{k^\gamma} \cdot \left[\frac{1}{\mu_k} - \left(\frac{1}{\mu_k} + t^*\right) \cdot e^{-\mu_k \cdot t^*}\right]\mathbf{d}k \\ &\approx \frac{\alpha}{\gamma - 1} \cdot \left[\frac{1}{\langle \mu \rangle} - \left(\frac{1}{\langle \mu \rangle} + t^*\right) \cdot e^{-\langle \mu \rangle \cdot t^*}\right]. \end{aligned}$$

C. Optimal Storing Probabilities and Discarding Probabilities

Based on the storing-discarding equilibrium, the optimal storing probabilities and discarding probabilities are investigated to maximize the delivery ratio of data packets. The delivery ratio is maximized when the storage of nodes is full, and thus we have $\sum_{t=0}^{t^*-1} H(t) = \frac{S}{N \cdot \rho \cdot L_s}$, which gives that:

$$\frac{\alpha}{\gamma - 1} \cdot \left[\frac{1}{\langle \mu \rangle} - \left(\frac{1}{\langle \mu \rangle} + t^*\right) \cdot e^{-\langle \mu \rangle \cdot t^*}\right] = \frac{S}{N \cdot \rho \cdot L_s}. \tag{10}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSC.2021.3103105, IEEE Transactions on Services Computing

Then, the optimal value of $\langle\mu\rangle$ is obtained.

In (9), the values of $D(t^*)$, $\langle\mu\rangle$ and t^* will not be changed with the value variation of k. Besides, for any k ($k \in [0, N-1]$), (9) must be satisfied, which indicates that $\frac{v_k + k\cdot\frac{\mathbf{d}v_k}{\mathbf{d}k}}{\frac{\mathbf{d}\mu_k}{\mathbf{d}k}}$ must be independent of the value of k. Thus, the form of (9) motivates us to construct the expressions of v_k and μ_k as:

$$\begin{cases} v_k = \begin{cases} 0, & if\ k = 0, \\ \frac{\lambda}{k^{\delta+1}}, & if\ k > 0, \end{cases} \\ \mu_k = \begin{cases} 0, & if\ k = 0, \\ \frac{\beta}{k^{\delta}}, & if\ k > 0, \end{cases} \end{cases} \tag{11}$$

where $\delta \geq 1$. Moreover, λ and β should satisfy that:

$$\frac{\beta}{\lambda} = \frac{D(t^*)}{\langle\mu\rangle \cdot t^*}. \tag{12}$$

Note that the optimal value of δ is analyzed in Section V.C, and the computation of $D(t^*)$ is derived as follows:

By a variable separation method, (5) can be expressed as:

$$\frac{-D(t)}{D(t)^2 - D(t) + \frac{\alpha\cdot\langle\mu\rangle}{\langle k\rangle^{\gamma}\cdot\langle v\rangle}}\mathbf{d}D(t) = \langle\mu\rangle\,\mathbf{d}t. \tag{13}$$

The left part of (13) can be rewritten as:

$$\frac{-D(t)}{D(t)^2 - D(t) + \frac{\alpha\cdot\langle\mu\rangle}{\langle k\rangle^{\gamma}\cdot\langle v\rangle}}\mathbf{d}D(t) = -\frac{1}{B}\cdot\left[\frac{\frac{D(t)}{B} - \frac{1}{2B}}{(\frac{D(t)}{B} - \frac{1}{2B})^2 + 1} + \frac{\frac{1}{2B}}{(\frac{D(t)}{B} - \frac{1}{2B})^2 + 1}\right]\mathbf{d}D(t),$$

where $B = \sqrt{\frac{\alpha\cdot\langle\mu\rangle}{\langle k\rangle^{\gamma}\cdot\langle v\rangle} - \frac{1}{4}}$. Thus,

$$\begin{aligned} &\int \frac{-D(t)}{D(t)^2 - D(t) + \frac{\alpha\cdot\langle\mu\rangle}{\langle k\rangle^{\gamma}\cdot\langle v\rangle}}\mathbf{d}D(t) \\ &= -\frac{1}{B}\cdot\int\left[\frac{\frac{D(t)}{B} - \frac{1}{2B}}{\left(\frac{D(t)}{B} - \frac{1}{2B}\right)^2 + 1} + \frac{\frac{1}{2B}}{\left(\frac{D(t)}{B} - \frac{1}{2B}\right)^2 + 1}\right]\mathbf{d}D(t) \\ &= -\frac{1}{B}\cdot\left\{\begin{array}{l} \frac{B}{2}\cdot\ln\left[\left(\frac{D(t)}{B} - \frac{1}{2B}\right)^2 + 1\right] \\ +\frac{1}{2}\cdot\arctan\left(\frac{D(t)}{B} - \frac{1}{2B}\right) \end{array}\right\} \\ &= \int\langle\mu\rangle\,\mathbf{d}t = \langle\mu\rangle\cdot t + f(B). \end{aligned} \tag{14}$$

As a special case, $D(0) = 0$, which yields that:

$$f(B) = -\frac{1}{B}\cdot\left\{\frac{B}{2}\cdot\ln\left[\left(\frac{1}{2B}\right)^2 + 1\right] + \frac{1}{2}\cdot\arctan\left(\frac{1}{2B}\right)\right\},$$

and thus the numerical result of $D(t^*)$ can be obtained from the following equation:

$$-\frac{1}{B}\cdot\left\{\begin{array}{l} \frac{B}{2}\cdot\ln\left[\left(\frac{D(t^*)}{B} - \frac{1}{2B}\right)^2 + 1\right] \\ +\frac{1}{2}\cdot\arctan\left(\frac{D(t^*)}{B} - \frac{1}{2B}\right) \end{array}\right\} = \langle\mu\rangle\cdot t^* + f(B). \tag{15}$$

V. Adaptive Data Dissemination Algorithm

A data packet $data(s,d)$ is assumed to be disseminated from the source node V_s to the destination node V_d, and the node V_i stores $data(s,d)$ at the t-th time slot. Each node maintains a list regarding its stored data packets. TABLE III provides symbols used in the description of the proposed algorithm ADDA.

TABLE III: Symbols in Description of ADDA

Symbol	Definition
$data\_list(i)^{(t)}$	List of data packets stored by V_i
$dis\_data\_list(i)^{(t)}$	List of data packets discarded by V_i
$\mathcal{N}(i)^{(t)}$	Set of V_i's neighbours at the t-th time slot
$\mathcal{M}(i)^{(t)}_{(s,d)}$	Set of V_i's neighbours storing $data(s,d)$ at the t-th time slot
$\mu(i)^{(t)}_{(s,d)}$	Discarding probability of V_i at the t-th time slot
$v(i)^{(t)}_{(s,d)}$	Storing probability of V_i at the t-th time slot
$S(i)^{(t)}$	Residual storage of V_i at the t-th time slot

A. ADDA Stages

In ADDA, each node makes the storing decision and discarding decision in a distributed manner, and the operation of ADDA is described in terms of five stages: initialization, inquiry and reception, probability settings, data dissemination and data discarding, as illustrated in Fig. 6, where the symbol t^+ denotes some (small) time after the start of t-th time slot, and $(t+1)^-$ denotes some (small) time before the end of t-th time slot, such that $t < t^+ < (t+1)^- < t+1$.

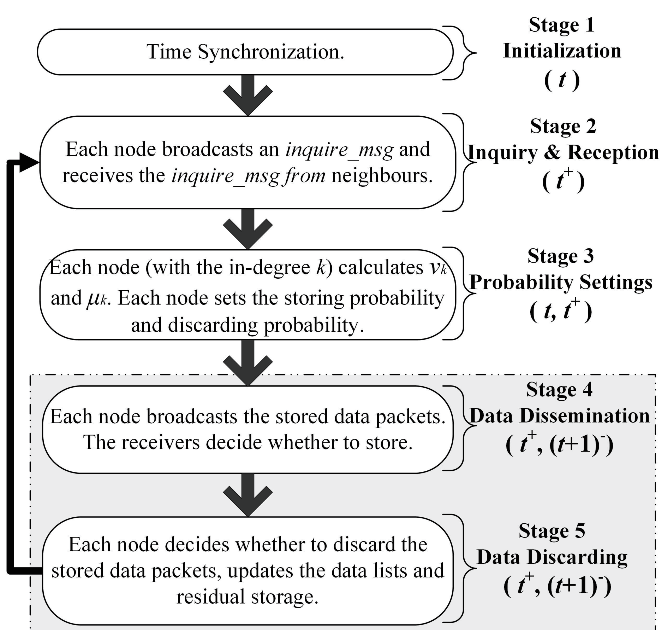


Fig. 6: The stages of ADDA.

***Stage* 1**. *Initialization*. At the start of any time slot, the time synchronization is accomplished among each node

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSC.2021.3103105, IEEE Transactions on Services Computing

and its neighbours.

Stage 2. *Inquiry and Reception*. In Stage 2, each node interacts with neighbours for the information exchanges, and there are two sub-stages as follows:

Sub-Stage 2-1. At the t-th time slot, each node V_i broadcasts an $inquire\_msg$ packet, including a quadruplet $(V_i, t, C(i)^{(t)}, data\_list(i)^{(t)})$.

Sub-Stage 2-2. Each node V_i determines the set of neighbours $\mathcal{N}(i)^{(t)}$ on basis of the received $inquire\_msg$ packets, and thus the in-degree of V_i is $\left|\mathcal{N}(i)^{(t)}\right|$.

Besides, the set of neighbours storing $data(s,d)$ is marked as $\mathcal{M}(i)^{(t)}_{(s,d)}$ and is determined by:

$$\mathcal{M}(i)^{(t)}_{(s,d)} = \bigcup \left\{ V_j \middle| \; V_j \in \mathcal{N}(i)^{(t)} \wedge data(s,d) \in data\_list(j)^{(t)} \right\}.$$

Stage 3. *Probability Settings*. Each node V_j sets the storing probability and discarding probability as:

$$\begin{cases} v(j)^{(t)}_{(s,d)} \leftarrow \left|\mathcal{M}(j)^{(t)}_{(s,d)}\right| \cdot v_{|\mathcal{N}(j)^{(t)}|}, \\ \mu(j)^{(t)}_{(s,d)} \leftarrow \mu_{|\mathcal{N}(j)^{(t)}|}, \end{cases} \tag{16}$$

where the value of $v_{|\mathcal{N}(j)^{(t)}|}$ and $\mu_{|\mathcal{N}(j)^{(t)}|}$ can be obtained from (11). Note that there are always several data packets propagated in an OUSN simultaneously, and these probabilities for different data packets should be individually calculated.

Stage 4. *Data Dissemination*. Each node V_i having stored $data(s,d)$ will broadcast the $data(s,d)$ to neighbours.

Sub-Stage 4-1. If $V_i \in \mathcal{N}(j)^{(t)}$, there are three cases as follows:

- **Case** 1: If $data(s,d) \in data\_list(j)^{(t)}$ or $S(j)^{(t)} < L_s$, the received $data(s,d)$ will be ignored by V_j;
- **Case** 2: If $data(s,d) \in dis\_data\_list(j)^{(t)}$, the received $data(s,d)$ will be ignored by V_j;
- **Case** 3: If $data(s,d) \notin data\_list(j)^{(t)}$ and $S(j)^{(t)} \geq L_s$, the received $data(s,d)$ will be stored by V_j with the probability $v(j)^{(t)}_{(s,d)}$.

Sub-Stage 4-2. If the $data(s,d)$ is stored by V_j, and then the list of stored data packets and residual storage of V_j are updated by: $data\_list(j)^{(t+1)} \leftarrow data\_list(j)^{(t)} \bigcup data(s,d)$, $S(j)^{(t+1)} \leftarrow S(j)^{(t)} - L_s$.

Stage 5. *Data Discarding*. In Stage 5, each node discards the stored data packets and updates the residual storage.

Sub-Stage 5-1. For each stored data packet $data(s,d) \in data\_list(i)^{(t)}$, the stored $data(s,d)$ will be discarded with the probability $\mu(i)^{(t)}_{(s,d)}$. Especially, all the copies of $data(s,d)$ will be discarded if the allowable dissemination time slots have been expired. If the stored $data(s,d)$ is discarded by V_i, the lists of stored data packets and discarded data packets are updated by:

$$data\_list(i)^{(t+1)} \leftarrow data\_list(i)^{(t)} \setminus data(s,d),$$
$$dis\_data\_list(i)^{(t+1)} \leftarrow dis\_data\_list(i)^{(t)} \bigcup data(s,d).$$

Sub-Stage 5-2. The residual storage of V_i is updated by $S(i)^{(t+1)} \leftarrow S(i)^{(t)} + L_s$.

The above stages will be repeated until V_d receives $data(s,d)$, or the allowable dissemination time slots have been expired. If $data(s,d)$ has been delivered to V_d before the deadline (t^* time slots), an announcement message originated from V_d will be propagated to all the nodes storing $data(s,d)$ to stop the further dissemination and discard all the stored $data(s,d)$. The size of announcement message is very small and can be piggybacked with other packets, and thus the effect of broadcast is negligible.

B. Complexity of ADDA

TABLE IV shows the packet amount and time consumption of each stage in ADDA. The packets of ADDA are mainly generated in Stage 2 and Stage 4. In Stage 2, each node broadcasts an $inquire\_msg$. Each node is supposed to store at least one data packet in the worst case, therefore the total packet amount will reach $O(N)$; in Stage 4, there are at most $\frac{S}{L_s}$ data packets (when the storage of each node is full) to be disseminated by each node, leading to a total number of $N \cdot \frac{S}{L_s}$ data packets.

With regard to the time complexity, in Stage 1 the time synchronization results in $O(N)$ computations; in Stage 2, the in-degree of each node is calculated, and there is a total of $O(N)$ computations; Stage 3 makes the storing probability and discarding probability be set for each node; in Stage 4 and Stage 5, the residual storage, the lists of stored data packets and discarded data packets will be updated for each node, and there are at most $O(N)$ updates.

Moreover, the above stages will be executed at most t^* time slots, where t^* is a constant. Thus, both message complexity and time complexity of ADDA are $O(N)$.

TABLE IV: Complexity of ADDA

Stage	Message Complexity	Time Complexity
1	0	$O(N)$
2	$O(N)$	$O(N)$
3	0	$O(N)$
4	$O(N)$	$O(N)$
5	0	$O(N)$

C. Expected Delivery Ratio of ADDA

The expected existence probability of a potential link is denoted by $\overline{P}$, which is calculated by:

$$\overline{P} = \frac{c_1 \cdot \int_{\Omega_{min}}^{\Omega_{max}} \omega^{-\zeta} \mathbf{d}\omega}{\Omega_{max} - \Omega_{min}} \cdot \frac{\sum_{k=1}^{\frac{R_c}{R_0}} \left[k^3 - (k-1)^3\right] \cdot (k \cdot R_0)^{-\eta}}{(\frac{R_c}{R_0})^3}$$
$$= \frac{c_1 \cdot ({\Omega_{max}}^{1-\zeta} - {\Omega_{min}}^{1-\zeta})}{(1-\zeta) \cdot (\Omega_{max} - \Omega_{min})} \cdot \frac{\sum_{k=1}^{\frac{R_c}{R_0}} (3k^2 - 3k + 1) \cdot (k \cdot R_0)^{-\eta}}{(\frac{R_c}{R_0})^3},$$

where R_0 denotes the minimum communication distance between two nodes. We can easily derive that $\mathcal{R}(t) = \mathcal{R}(t-1) + \mathcal{A}(t) \cdot [1 - \mathcal{R}(t-1)]$, where $\mathcal{A}(t)$ denotes the probability of $data(s,d)$ being delivered at the t-th time slot. After the movement of underwater nodes at any time slot, the nodes remain uniformly distributed, which has been proven in [1], and thus when N is large enough $\mathcal{A}(t)$ is expressed as:

$$\mathcal{A}(t) = \overline{P} \cdot \frac{4\pi \cdot {R_c}^3 \cdot N}{3|\boldsymbol{D}|} \cdot H(t) = \overline{P} \cdot \frac{4\pi \cdot {R_c}^3 \cdot N}{3|\boldsymbol{D}|} \cdot H(t).$$

There is approximately $\frac{\mathbf{d}\mathcal{R}(t)}{\mathbf{d}t} + \mathcal{A}(t+1) \cdot \mathcal{R}(t) = \mathcal{A}(t+1)$, which yields that:

$$\mathcal{R}(t) = 1 - c_4 \cdot e^{-\int \mathcal{A}(t+1)\mathbf{d}t} = 1 - c_4 \cdot e^{-\overline{P} \cdot \frac{4\pi \cdot {R_c}^3 \cdot N}{3|\boldsymbol{D}|} \cdot \int H(t+1)\mathbf{d}t}. \tag{17}$$

Hence, $\mathcal{R}(t^*)$ can be expressed as:

$$\begin{aligned}\mathcal{R}(t^*) &= 1 - c_4 \cdot e^{-\overline{P} \cdot \frac{4\pi \cdot {R_c}^3 \cdot N}{3|\boldsymbol{D}|} \cdot \int H(t+1)\mathbf{d}t}\Big|_{t=t^*} \\ &\approx 1 - c_4 \cdot e^{-\overline{P} \cdot \frac{4\pi \cdot {R_c}^3 \cdot N}{3|\boldsymbol{D}|} \cdot \sum_{t=0}^{t^*-1} H(t+1)} \\ &\approx 1 - c_4 \cdot e^{-\overline{P} \cdot \frac{4\pi \cdot \alpha \cdot {R_c}^3 \cdot N}{3(\gamma-1) \cdot |\boldsymbol{D}|} \cdot \left[\frac{1}{\langle\mu\rangle} - \left(\frac{1}{\langle\mu\rangle} + t^* + 1\right) \cdot e^{-\langle\mu\rangle \cdot (t^*+1)}\right]},\end{aligned} \tag{18}$$

a special case of which is $\mathcal{R}(0) = 0$, and hence there is $c_4 = e^{\overline{P} \cdot \frac{4\pi \cdot \alpha \cdot {R_c}^3 \cdot N}{3(\gamma-1) \cdot |\boldsymbol{D}|} \cdot \left[\frac{1}{\langle\mu\rangle} - \left(\frac{1}{\langle\mu\rangle} + 1\right) \cdot e^{-\langle\mu\rangle}\right]}$.

(18) indicates that $\mathcal{R}(t^*)$ is proportional to the part $\frac{1}{\langle\mu\rangle} - \left(\frac{1}{\langle\mu\rangle} + t^* + 1\right) \cdot e^{-\langle\mu\rangle(t^*+1)}$. There is

$$\begin{aligned}\langle\mu\rangle &= \sum_{i=1}^{N-1} \frac{\alpha \cdot \beta}{i^{(\delta+\gamma)}} \approx \frac{\alpha \cdot \beta}{\delta + \gamma - 1} \cdot \left[1 - \frac{1}{(N-1)^{\delta+\gamma-1}}\right] \\ &\approx \frac{\alpha \cdot \beta}{\delta + \gamma - 1}.\end{aligned}$$

Let $\mathcal{F}(\delta) = \frac{1}{\langle\mu\rangle} - \left(\frac{1}{\langle\mu\rangle} + t^* + 1\right) \cdot e^{-\langle\mu\rangle \cdot (t^*+1)}$, which can be rewritten as:

$$\mathcal{F}(\delta) = \frac{\delta + \gamma - 1}{\alpha \cdot \beta} - \left(\frac{\delta + \gamma - 1}{\alpha \cdot \beta} + t^* + 1\right) \cdot e^{-\frac{\alpha \cdot \beta \cdot (t^*+1)}{\delta + \gamma - 1}},$$

and the first order derivative of $\mathcal{F}(\delta)$ with respect to δ is expressed as:

$$\frac{\mathbf{d}\mathcal{F}(\delta)}{\mathbf{d}\delta} = \frac{1}{\alpha \cdot \beta} - \frac{\frac{1}{\alpha \cdot \beta} + \frac{\alpha \cdot \beta \cdot (t^*+1)}{(\delta+\gamma-1)^2} \cdot \left(\frac{\delta+\gamma-1}{\alpha \cdot \beta} + t^* + 1\right)}{e^{\frac{\alpha \cdot \beta \cdot (t^*+1)}{\delta+\gamma-1}}}.$$

The maximum of $\mathcal{F}(\delta)$ can be achieved when $\frac{\mathbf{d}\mathcal{F}(\delta)}{\mathbf{d}\delta} = 0$, from which the optimal δ is obtained by:

$$\delta = \frac{\alpha \cdot \beta \cdot (t^* + 1)}{\partial} + 1 - \gamma,$$

where ∂ satisfies that $e^{\partial} = 1 + \partial + \partial^2$.

The equation $\frac{\mathbf{d}\mathcal{F}(\delta)}{\mathbf{d}\delta} = 0$ also gives that:

$$\left(\frac{1}{\langle\mu\rangle} + t^* + 1\right) \cdot e^{-\langle\mu\rangle \cdot (t^*+1)} = \frac{1 - e^{-\langle\mu\rangle \cdot (t^*+1)}}{(t^* + 1) \cdot \langle\mu\rangle^2},$$

and thus the expected delivery ratio of ADDA is expressed as:

$$\begin{aligned}\mathbb{E}(\mathcal{R}(t^*)) &= 1 - c_4 \cdot e^{-\overline{P} \cdot \frac{4\pi \cdot \alpha \cdot {R_c}^3 \cdot N}{3(\gamma-1) \cdot |\boldsymbol{D}|} \cdot \left[\frac{1}{\langle\mu\rangle} - \frac{1 - e^{-\langle\mu\rangle \cdot (t^*+1)}}{(t^*+1) \cdot \langle\mu\rangle^2}\right]} \\ &= 1 - e^{\overline{P} \cdot \frac{4\pi \cdot \alpha \cdot {R_c}^3 \cdot N}{3(\gamma-1) \cdot |\boldsymbol{D}|} \cdot \left[\frac{1 - e^{-\langle\mu\rangle \cdot (t^*+1)}}{(t^*+1) \cdot \langle\mu\rangle^2} - \left(\frac{1}{\langle\mu\rangle} + 1\right) \cdot e^{-\langle\mu\rangle}\right]},\end{aligned}$$

which indicates the delivery ratio grows with the increase of N or R_c. Some numerical results are illustrated in Fig. 7.

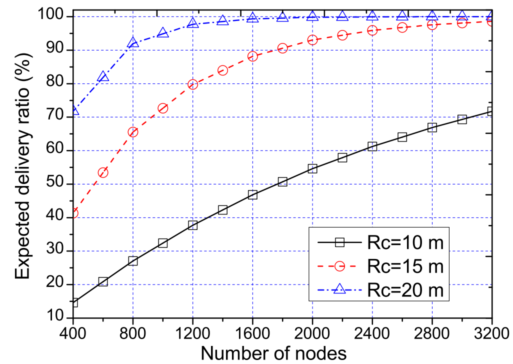


Fig. 7: Expected delivery ratio vs. N and R_c.

VI. Simulations

In this section, ADDA is evaluated by observing the performance variations with respect to different parameters and by comparing with other algorithms (EF [4], BDCR [13], and TORA [20]). The main parameter settings are shown in TABLE V.

TABLE V: Simulation Parameters

Parameter	Description	Value
N	Number of nodes	1,000
$\|\boldsymbol{D}\|$	Deployment space	400×400 ×100 m^3
S	Storage size of each node	120 KB
R_c	Communication range	20 m
R_0	Minimum communication distance	2 m
$\mathcal{V}_m$	Maximum autonomous speed	1.5 m/s
L_s	Size of each data packet	2,000 B
τ_s	Duration of each time slot	6.1 s
t^*	Number of allowable dissemination time slots	9
ρ	Probability of generating a data packet at each time slot	0.05
c_1	Coefficient in signal irregularity formula	0.679
ζ	Exponent in signal irregularity formula	0.77
η	Exponent in signal irregularity formula	2
γ	Exponent in power law distribution	2.1
β	Coefficient in discarding probability	0.601
λ	Coefficient in storing probability	0.852
Ω_{min}	Minimum signal irregularity	0.1
Ω_{max}	Maximum signal irregularity	0.9

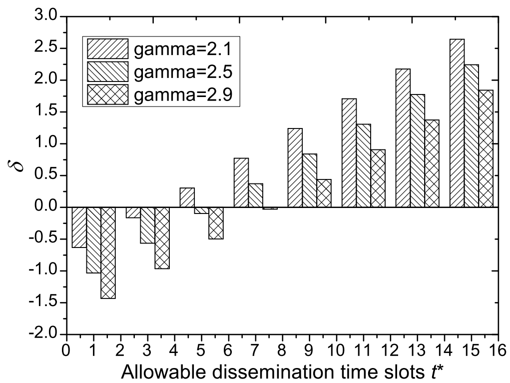
A. Simulation Setup

We develop a simulator using C++ language to assess the performance of ADDA. Besides the parameter values given in TABLE V, some other parameters (such as δ, $D(t^*)$, storing probabilities, and discarding probabilities) should be calculated for the simulations. The values of δ and $D(t^*)$ are first calculated, as shown in Fig. 8(a) and Fig. 8(b).

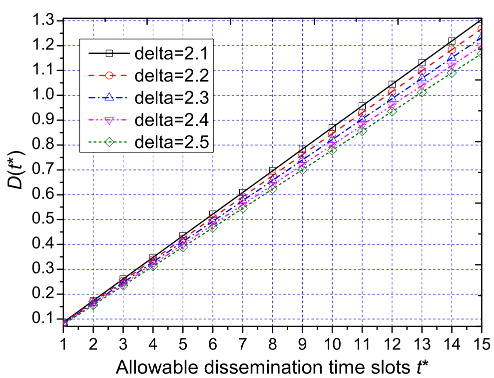
Based on the obtained values of δ and $D(t^*)$, the storing probabilities and discarding probabilities of nodes with different in-degrees are calculated according to (11). In Fig. 8(c), both the storing probability and discarding

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSC.2021.3103105, IEEE Transactions on Services Computing

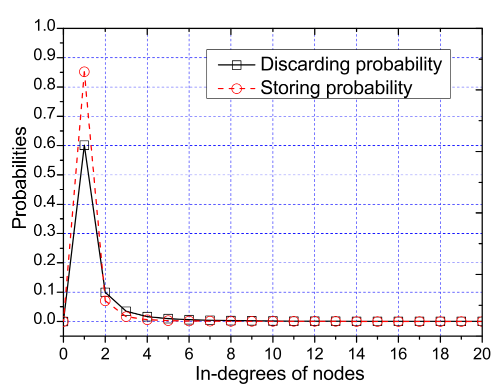
probability decrease with the increase of in-degree when the in-degree is not equal to 0. With the increase of in-degree, note that the ratio of the discarding probability to the storing probability is actually enlarged due to the expressions of μ_k and υ_k given in (11), although the two curves become closer to each other. Especially, the settings of β and λ are not unique (the values of β and λ satisfying the equation $\frac{\beta}{\lambda} = \frac{D(t^*)}{\langle\mu\rangle \cdot t^*}$ are available), and in our simulations β and λ are set to 0.601 and 0.852, respectively.



(a) Value of δ



(b) Value of $D(t^*)$

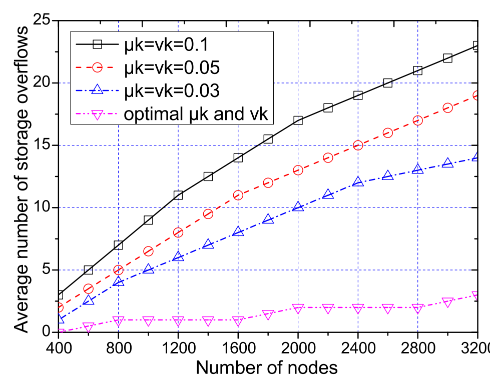


(c) Optimal Storing Probabilities and Discarding Probabilities

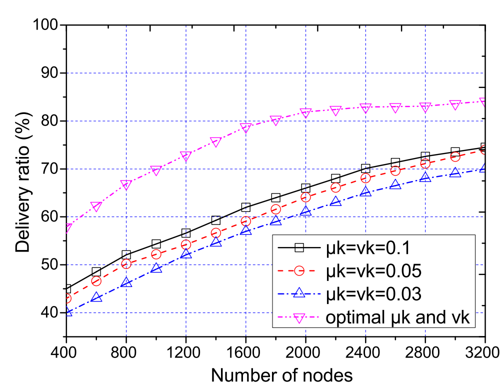
Fig. 8: Simulation precomputation.

B. Impacts of Storing Probabilities and Discarding Probabilities

The optimal storing probabilities and discarding probabilities of nodes with different in-degrees have been introduced in Section IV.C. ADDA with the optimal storing probabilities and discarding probabilities is compared with those with other storing probabilities and discarding probabilities, such as $\mu_k = \upsilon_k = 0.1$, $\mu_k = \upsilon_k = 0.05$, and $\mu_k = \upsilon_k = 0.03$. Fig. 9 indicates that the optimal storing probabilities and discarding probabilities help to reduce the number of storage overflows and improve the delivery ratio effectively: (i) In Fig. 9(a), the curve of $\mu_k = \upsilon_k = 0.1$ is much higher than other curves, which is attributed to the fact that with a larger υ_k more data packets can be stored by nodes, and the equivalency of μ_k and υ_k increases the number of storage overflows as well (the value of υ_k should be set slightly smaller than that of μ_k, as illustrated in Fig. 8(c)). (ii) Besides the storage overflows alleviation, the optimal storing probabilities and discarding probabilities of nodes also improve the delivery ratio, as evidently observed in Fig. 9(b).



(a) Average Number of Storage Overflows vs. μ_k and υ_k



(b) Delivery Ratio vs. μ_k and υ_k

Fig. 9: Impacts of storing probabilities and discarding probabilities.

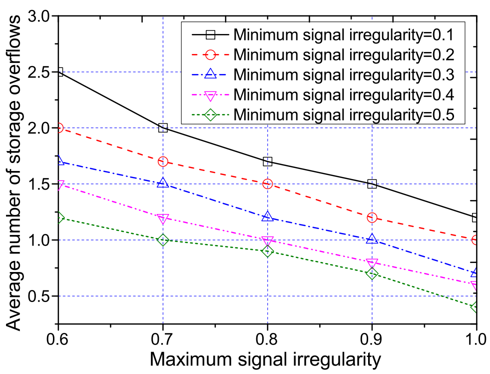
C. Impacts of Signal Irregularity

Fig. 10 shows the impacts of signal irregularity on the average number of storage overflows and the delivery ratio. The signal irregularity around nodes is determined by the values of Ω_{min} and Ω_{max}, i.e., the signal irregularity becomes larger when Ω_{min} or Ω_{max} becomes larger.

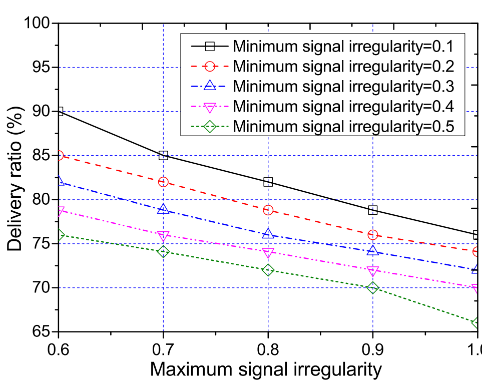
A larger signal irregularity will make more potential links disconnected and fewer data packets stored by nodes, and thus the average number of storage overflows is reduced, as illustrated in Fig. 10(a). However, a larger signal irregularity makes the data packets more difficult to be disseminated, and more data packets cannot be delivered to the destination node during the allowable dissemination time slots, and hence the delivery ratio is reduced as well, as shown in Fig. 10(b).

node increased over time slots. Besides, the average number of storage overflows becomes larger when the nodes are deployed more densely. This is due to the fact that more data packets are generated and disseminated in an OUSN with more nodes, and thus more storage overflows occur on nodes. The average number of storage overflows reaches 3.09 when N=3,200 at the 9-th time slot.

As shown in Fig. 11(b), the delivery ratio becomes larger when the observation period is prolonged, i.e., more data packets will be delivered to the destination node during a longer observation period. Moreover, the delivery ratio continues to increase with the growth of N, since more nodes can accommodate more data packets (copies), and hence the delivery ratio is raised, which tallies with the numerical results given in Fig. 7. Specifically, the highest delivery ratio at the 9-th time slot is 84.1% when N=3,200.

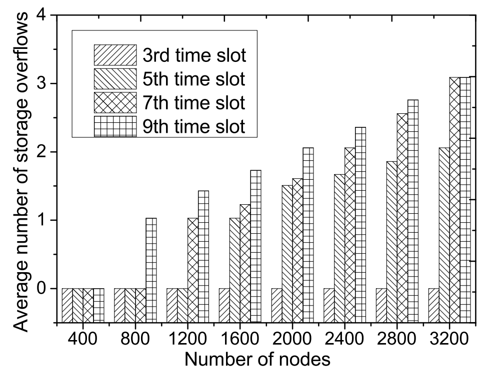


(a) Average Number of Storage Overflows vs. Ω_{min} and Ω_{max}

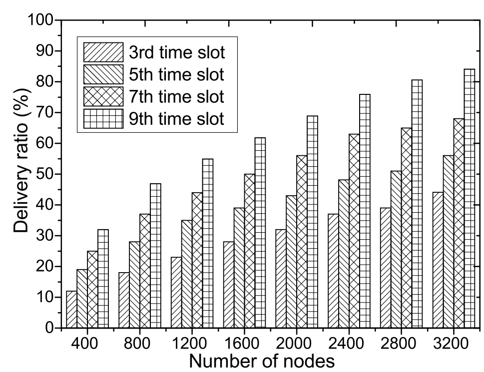


(b) Delivery Ratio vs. Ω_{min} and Ω_{max}

Fig. 10: Impacts of signal irregularity.



(a) Average Number of Storage Overflows



(b) Delivery Ratio

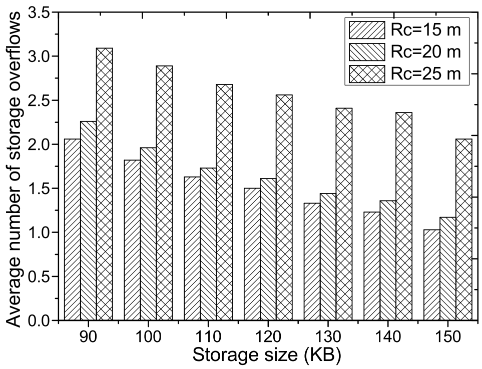
Fig. 11: Impacts of observation period.

D. Impacts of Observation Period

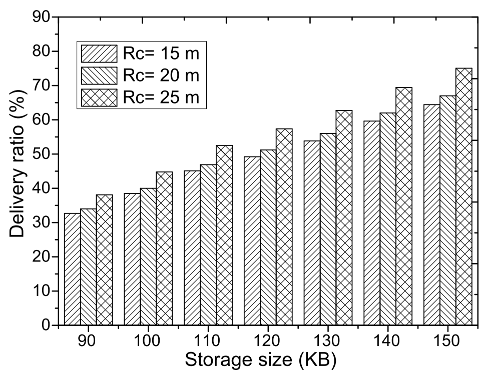
In Fig. 11(a), it is shown that the average number of storage overflows during a longer observation period is larger than those during shorter observation periods, and this is because the storage overflows cannot be completely avoided although some stored data packets are actively discarded by each node, making the number of storage overflows on each

E. Impacts of Storage Size and Communication Range

Fig. 12 illustrates the impacts of storage size and communication range of nodes. With the increase of storage size or communication range of nodes, the average number of storage overflows is generally reduced, and the delivery ratio is gradually increased. A larger storage size implies that each node can store more data packets, and thus these data packets can be delivered to the destination node more easily. Moreover, a larger communication range makes the nodes encounter more neighbours and have more opportunities to relay the stored data packets, and thereby the delivery ratio is enhanced.



(a) Average Number of Storage Overflows



(b) Delivery Ratio

Fig. 12: Impacts of storage size and communication range of nodes.

F. Algorithm Comparisons

We compare ADDA against EF, BDCR, and TORA under the same underwater mobility model, where the mobility of nodes is comprised of the autonomous movement controlled by underwater vehicles and the coordinate deviation impelled by underwater external forces. Given the simulation results in Fig. 13, we can observe that ADDA outperforms other algorithms (EF, BDCR, and TORA) in terms of the average number of storage overflows, and the delivery ratio. The reason is that in ADDA the stored data packets are allowed to be actively discarded to make room for the newly arriving data packets, and the storage overflows on nodes are accordingly alleviated. Moreover, ADDA seeks to achieve the storing-discarding equilibrium, and then sets the optimal storing probabilities and discarding probabilities to improve the delivery ratio.

EF adopts the epidemic dissemination manner, and thus EF achieve the shortest average delay among these algorithms, as shown in Fig. 13(c). However, in Fig. 13(a), many data packets cannot be stored and have to be overflowed during the period of t^* time slots, making the average number of storage overflows of EF much higher than others, which indicates that EF is not suitable for most of OUSN applications due to the extremely high communication complexity and large number of storage overflows. Especially, when the number of nodes is large enough (i.e., the number of data packets is very large), the delivery ratio of ADDA is larger than that of EF (depicted in Fig. 13(b)), and this is because the mechanism of casually discarding data packets worsens the delivery ratio when the storage of all nodes is full. Likewise, the storing-discarding equilibrium is not considered in BDCR and TORA, and hence they do not perform well in a storage-limited OUSN.

VII. Conclusions

This study explores the data dissemination problem for the delivery ratio enhancement and storage overflows alleviation in the storage-limited OUSNs with signal irregularity. A differential equation set describing the propagation process of data packets is provided to investigate the storing-discarding equilibrium, and then the optimal storing probabilities and discarding probabilities are obtained for the nodes with different in-degrees to maximize the delivery ratio. In the proposed Adaptive Data Dissemination Algorithm (ADDA), the newly arriving data packets are stored and the stored data packets are discarded by nodes according to the obtained storing probabilities and discarding probabilities, respectively. Simulation results show that ADDA outperforms other existing algorithms in terms of the average number of storage overflows and the delivery ratio, and it is specially suitable for the storage-limited or densely deployed OUSNs.

In this paper, the nodes are assumed to be with the same specification, and the communication range and storage size of each node are set the same. Our future research will focus on the data dissemination algorithm for the heterogeneous OUSNs, where the communication range and (or) storage size of each node could be heterogeneous.

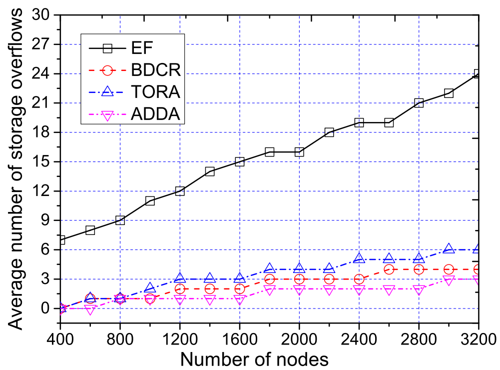
VIII. Acknowledgments

This research is supported by National Natural Science Foundation of China under Grant Nos. 61872191, 61872193; National Key R&D Program of China No. 2019YFB2101700; Six Talents Peak Project of Jiangsu Province under Grant No. 2019-XYDXX-247.

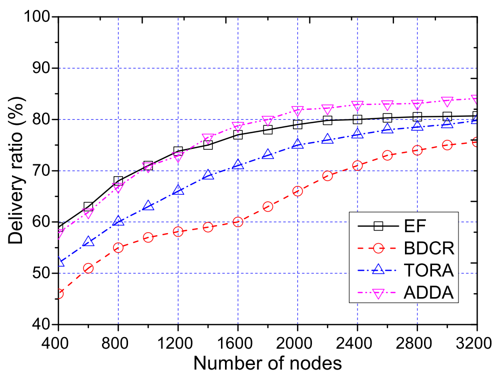
References

[1] L. Liu, R. Wang, G. Xiao, and D. Guo, "On the Throughput Optimization for Message Dissemination in Opportunistic Underwater Sensor Networks," *Computer Networks (Elsevier)*, vol. 169, pp. 1–16, 2020.

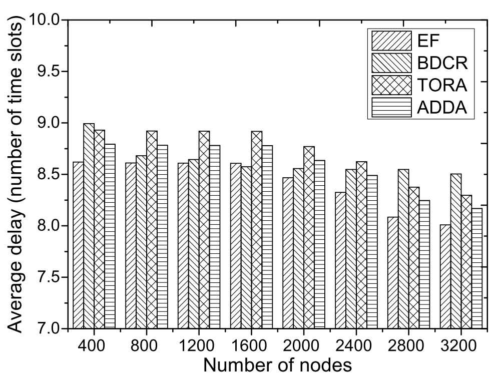
[2] R. Davis, M. Baumgartner, A. Comeau, et al, "Tracking Whales on the Scotian Shelf using Passive Acoustic Monitoring on Ocean Gliders," *IEEE OCEANS16*, Monterey, USA, 2016.



(a) Average Number of Storage Overflows



(b) Delivery Ratio



(c) Average Delay

Fig. 13: Algorithm comparisons.

[3] B. G. Ferguson, and K. W. Lo, "Acoustic Detection, Localization, and Tracking of Tactical Autonomous Aerial and Underwater Vehicles," *Journal of the Acoustical Society of America*, vol. 140, no. 4, 2016.

[4] F. Li, and J. Wu, "LocalCom: A Community-based Epidemic Disseminating Scheme in Disruption-tolerant Networks," *IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (IEEE SECON)*, pp. 574–582, Rome, Italy, 2009.

[5] L. Liu, Y. Liu, and N. Zhang, "A Complex Network Approach to Topology Control Problem in Underwater Acoustic Sensor Networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 25, no. 12, pp. 3046–3055, 2014.

[6] R. Martin, J. H. Cui, and S. Han, "Tracer Assisted Passive Named Data Routing for Underwater Sensor Networks," *OCEANS 2014-TAIPEI*, pp. 1–8, 2014.

[7] P. Rawat, K. D. Singh, H. Chaouchi, et al. "Wireless Sensor Networks: A Survey on Recent Developments and Potential Synergies," *The Journal of Supercomputing*, vol. 68, no. 1, pp. 1–48, 2014.

[8] X. Niu, and L. Cui, "Throughput Capacity of Opportunistic Routing in Wireless Sensor Networks," *International Journal of Distributed Sensor Networks*, vol. 2010, Article ID 148359, 2010.

[9] M. E. Bayrakdar, "A Smart Insect Pest Detection Technique with Qualified Underground Wireless Sensor Nodes for Precision Agriculture," *IEEE Sensors Journal*, vol. 19, no. 22, pp. 10892–10897, 2019.

[10] M. E. Bayrakdar, "Cooperative Communication based Access Technique for Sensor Networks," *International Journal of Electronics*, vol. 107, no. 2, pp. 212–225, 2020.

[11] Z. Guo, B. Wang, and J. Cui, "Generic Prediction Assisted Single-copy Routing in Underwater Delay Tolerant Sensor Networks," *Ad Hoc Networks (Elsevier)*, vol. 2013, no. 11, pp. 1136–1149, 2013.

[12] G. Zheng, Z. Peng, B. Wang, et al, "Adaptive Routing in Underwater Delay Tolerant Sensor Networks," *6th International ICST Conference on Communications and Networking in China*, pp. 1044–1051, Harbin, China, 2011.

[13] S. Zhang, and D. Li, "A Beam width and Direction Concerned Routing for Underwater Acoustic Sensor Networks," *2013 IEEE 9th International Conference on Mobile Ad-hoc and Sensor Networks*, pp. 17–24, Las Vegas, USA, 2013.

[14] P. M. Shah, I. Ullah, T. Khan, et al, "MobiSink: Cooperative Routing Protocol for Underwater Sensor Networks with Sink Mobility," *2016 IEEE 30th International Conference on Advanced Information Networking and Applications*, pp. 189–197, Switzerland, 2016.

[15] S. Basagni, V. D. Valerio, P. Gjanci, and C. Petrioli, "Harnessing HyDRO: Harvesting-aware Data Routing for Underwater Wireless Sensor Networks," *the Eighteenth ACM International Symposium on Mobile Ad Hoc Networking and Computing (Mobihoc'18)*, pp. 271–279, Los Angeles, USA, 2018.

[16] R. W. L. Coutinho, A. Boukerche, L. F. M. Vieira, et al, "Geographic and Opportunistic Routing for Underwater Sensor Networks," *IEEE Transactions on Computers*, vol. 65, no. 2, pp. 548–561, 2016.

[17] R. W. L. Coutinho, A. Boukerche, L. F. M. Vieira, and A. A. F. Loureiro, "Performance Modeling and Analysis of Void-handling Methodologies in Underwater Wireless Sensor Networks," *Computer Networks (Elsevier)*, vol. 126, pp. 1–14, 2017.

[18] G. Han, L. Liu, N. Bao, J. Jiang, W. Zhang, and J. Rodriguesc, "AREP: An Asymmetric Link-based Reverse Routing Protocol for Underwater Acoustic Sensor Networks," *Journal of Network and Computer Applications (Elsevier)*, vol. 92, pp. 51–58, 2017.

[19] F. Ahmed, Z. Wadud, N. Javaid, N. Alrajeh, M. S. Alabed, and U. Qasim, "Mobile Sinks Assisted Geographic and Opportunistic Routing Based Interference Avoidance for Underwater Wireless Sensor Network," *Sensors*, vol. 18, no. 4, 2018.

[20] Z. Rahman, F. Hashim, M. F. Rasid, and M. Othman, "Totally Opportunistic Routing Algorithm (TORA) for Underwater Wireless Sensor Network," *PLOS ONE*, vol. 13, no. 6, 2018.

[21] M. E. Bayrakdar, "Cost Effective Smart System for Water Pollution Control with Underwater Wireless Sensor Networks: A Simulation Study," *Computer Systems Science and Engineering*, vol. 35, no. 4, pp. 283–292, 2020.

[22] M. Stojanovic, and J. Preisig, "Underwater Acoustic Communication Channels: Propagation Models and Statistical Characterization," *IEEE Communications Magazine*, vol. 47, no. 1, pp. 84–89, 2009.

[23] H. Nouri, M. Uysal, and E. Panayirci, "Information Theoretical Performance Analysis and Optimisation of Cooperative Underwater Acoustic Communication Systems," *IET Communications*, vol. 10, no. 7, pp. 812–823, 2016.

[24] M. Hussain, and N. Trigoni, "Distributed Localization in Cluttered Underwater Environments," *the Fifth ACM International Workshop on UnderWater Networks (WUWNet'10)*, Article no. 8, Massachusetts, USA, 2010.

[25] L. Liu, P. Wang, and R. Wang, "Propagation Control of Data Forwarding in Opportunistic Underwater Sensor Networks," *Computer Networks (Elsevier)*, vol. 114, pp. 80–94, 2017.