MITSUBISHI ELECTRIC RESEARCH LABORATORIES
http://www.merl.com

Molecular Fountain: A Robustness Enhancement Framework for Diffusive Molecular Communication

Egashira, H.; Suzuki, J.; Koike-Akino, T.; Nakano, T.; Fukuda, H.; Orlik, P.V.

TR2017-177 December 2017

Abstract

Molecule loss is a critical reliability issue in diffusive molecular communications. This paper proposes a communication framework that allows biologically-enabled machines (bionanomachines) to transmit and receive information-carrying molecules (information molecules) in a robust manner against molecule losses. The proposed framework, called molecular fountain, employs deoxyribonucleic-acid (DNA) molecules as information carriers and leverages molecular fragmentation (i.e., packetization) between transmitter (Tx) and receiver (Rx) bio-nanomachines. It performs feedback-aided rateless erasure coding that considers biochemical constraints in DNA synthesis and sequencing to generate molecular packets. The Tx bio-nanomachine repeatedly generates molecular packets with Luby transform codes and transmits them to the Rx bionanomachine until it receives an acknowledgment from the Rx bio-nanomachine. The Rx bio-nanomachine can reconstruct lost molecular packets from other packets that have been successfully transmitted. Simulation results show that molecular fountain enhances robustness against molecular packet losses and in turn improves communication performance such as transmission latency, jitter, error rate, and coding overhead.

IEEE Global Communications Conference (GLOBECOM)

Molecular Fountain: A Robustness Enhancement Framework for Diffusive Molecular Communication

Hiroaki Egashira*§, Junichi Suzuki*, Toshiaki Koike-Akino†, Tadashi Nakano‡, Hiroaki Fukuda§ and Philip Orlik†
*University of Massachusetts, Boston, Department of Computer Science, Boston, MA 02125, USA
†Mitsubishi Electric Research Laboratories (MERL), 201 Broadway, Cambridge, MA 02139, USA
‡Osaka University, Graduate School of Frontier Biosciences and Institute for Academic Initiatives, Osaka 565-0871, Japan
§Shibaura Institute of Technology, Department of Information Science and Engineering, Tokyo 135-8548, Japan
Email: {hiroaki, jxs}@cs.umb.edu, {koike, porlik}@merl.com, tadasi.nakano@fbs.osaka-u.ac.jp, hiroaki@shibaura-it.ac.jp

***Abstract*—Molecule loss is a critical reliability issue in diffusive molecular communications. This paper proposes a communication framework that allows biologically-enabled machines (bio-nanomachines) to transmit and receive information-carrying molecules (information molecules) in a robust manner against molecule losses. The proposed framework, called molecular fountain, employs deoxyribonucleic-acid (DNA) molecules as information carriers and leverages molecular fragmentation (i.e., packetization) between transmitter (Tx) and receiver (Rx) bio-nanomachines. It performs feedback-aided rateless erasure coding that considers biochemical constraints in DNA synthesis and sequencing to generate molecular packets. The Tx bio-nanomachine repeatedly generates molecular packets with Luby transform codes and transmits them to the Rx bio-nanomachine until it receives an acknowledgment from the Rx bio-nanomachine. The Rx bio-nanomachine can reconstruct lost molecular packets from other packets that have been successfully transmitted. Simulation results show that molecular fountain enhances robustness against molecular packet losses and in turn improves communication performance such as transmission latency, jitter, error rate, and coding overhead.**

I. Introduction

Diffusive molecular communication is known inherently unreliable due to stochastic molecular propagation, molecule-to-molecule collisions, and environmental noise. They cause extremely long latency, large jitter, high molecule loss rate, and low capacity [1–3]. This paper addresses the reliability issue, particularly molecule losses in information transmission, and proposes a robustness enhancement framework, called *molecular fountain*, for biologically-enabled machines (or *bio-nanomachines*) to reliably transmit and receive information-encoded molecules (or *information molecules*) through diffusive transports in aqueous environments.

Molecular fountain employs deoxyribonucleic-acid (DNA) molecules as information molecules because of their high information density (petabytes of data per gram), high information integrity and evolutionarily optimized machinery for replication, fragmentation and reassembly. In molecular fountain, the transmitter (Tx) bio-nanomachine fragments a large DNA molecule to smaller DNA molecules, called *molecular packets*. The receiver (Rx) bio-nanomachine reassembles the original DNA molecule from received packets. Such molecular fragmentation (i.e., packetization) and reassembly can lead to higher arrival probability at the Rx [3] because smaller DNA molecules diffuse faster [4] in general.

Molecular fountain performs feedback-aided rateless erasure coding [5] that respects biochemical constraints in DNA synthesis and sequencing (e.g., long homopolymer runs) to generate stable molecular packets in a redundant manner against packet losses. The Tx repeatedly generates molecular packets with Luby transform (LT) codes [5] and propagates them to the environment until it receives an acknowledgment (ACK) molecule from the Rx. An ACK is a receipt of a sufficient number of packets that the Rx requires to reassemble an information molecule. The Rx can reconstruct lost packets from other packets that have been successfully transmitted. Simulation results demonstrate that the proposed molecular fountain enhances robustness against molecular packet losses and in turn improves communication performance such as transmission latency, jitter, and error rate. The impact of coding overhead on communication performance is also analyzed. The key contribution of this paper is summarized below.

- **Packet fragmentation and reassembly:** Packet fragmentation and reassembly are employed for DNA-based molecular communications to achieve high packet arrival probability with increased diffusion speed.
- **LT erasure coding:** Fountain coding based on LT [5] is introduced to molecular communications in order to improve the robustness against molecular packet loss with a smallest possible overhead even without any knowledge of channel conditions in advance.
- **DNA screening:** Encoded packets are designed to eliminate biochemically unstable DNA molecules having high/low guanine-cytosine content and long homopolymer runs.
- **Multi-type data:** The advantage of fountain coding in transmission latency, jitter, and error rate performance is demonstrated with two different types of information, big image data and small text data.

II. Related Works

While various research efforts have focused on the physical layer's characteristics such as channel capacity, latency, signal attenuation, and energy requirements (e.g., [1, 2, 6–8]),

this paper investigates a higher-layer issue to realize reliable molecular transmission via robustness enhancement against molecule losses. There exist several relevant work to enhance the reliability of short-range molecular communications in aqueous environments [9–13]. Nakano et al. [9] and Felicetti et al. [10] study feedback-based rate control schemes for diffusive molecule propagation. Those schemes are designed to ensure delivering a given number of information molecules to the Rx bio-nanomachine while preventing the Tx bio-nanomachine from transmitting molecules faster than the Rx reacts. Nakano et al. examine both positive and negative feedback schemes [9], whereas Felicetti et al. study a negative feedback scheme [10].

While in-sequence delivery of information molecules is out of the scope of [9, 10], Wang et al. [11], Bai et al. [12] and Mitzman et al. [13] investigate in-sequence and at-least-once delivery schemes with positive feedbacks. The Rx is designed to explicitly acknowledge information molecules and request the Tx to retransmit lost ones based on stop-and-wait automatic repeat request (SW-ARQ) [14]. (Note that implicit acknowledgment is used in [9, 10].) Diffusive molecular transports are considered in [11] and [12], whereas Mitzman et al. [13] consider directional and diffusive-directional hybrid transports in addition to diffusive transports.

This paper enhances reliability in molecule transmission by means of forward error correction (FEC) instead of retransmission-based schemes such as SW-ARQ. We use erasure coding which allows the Rx to recover lost molecular packets without requesting the Tx to retransmit them. Such an approach is similar to [12, 15–22] which make use of FEC schemes including Hamming codes, cyclic redundancy check (CRC) codes, self-orthogonal convolutional codes (SOCCs), and minimum energy codes (MECs). Although those FEC schemes can detect and correct bit errors in information molecules, recovery of lost molecules is out of their scope.

Molecule losses are addressed in [21, 22], which perform erasure coding that is driven by a given overhead. Overhead indicates the level of redundancy in packet transmissions. However, it is often difficult to estimate the best, or even a reasonable, overhead for molecular communications due to inherent randomness of diffusion. In order to minimize the overhead even without knowledge of channels in advance, this paper leverages *rateless* codes (LT codes [5]). The proposed molecular fountain allows the Tx to adaptively determine a best overhead rate as it transmits molecular packets in a given environment until the Rx has required to reconstruct the original data. Another key contribution of this paper compared to [21, 22] includes a fact that we take a careful consideration of biochemical constraints in DNA synthesis and sequencing to generate packets, so that generated packets (DNA molecules) are less error-prone.

The similar method is experimentally demonstrated in [23], where LT fountain code is employed for error control of DNA storage. DNA allows extremely high storage capacity; e.g., a human cell, with a mass of roughly 3 picograms, accommodates 6.4 GB of information. This paper makes an analysis of packetization effects for image and text data transmissions in the context of communications rather than storage applications.

Furubayashi et al. propose packet fragmentation and reassembly in molecular communications [3]. Our paper shares the same assumptions for packetized molecular communications, while several new findings are obtained by considering three-dimensional molecular diffusion in the presence of molecule-to-molecule collisions. The paper in [3] assumes a one-dimensional and collision-free environment. In addition, although Furubayashi et al. consider packet reassembly at the Rx, recovery of lost packets is not investigated.

III. Molecular Fountain

This section describes the communication model and protocol in the proposed molecular fountain.

A. Diffusive Communication Model

This paper assumes a bounded three-dimensional aqueous environment in which the Tx and Rx bio-nanomachines exchange molecules for data transmission. The Tx propagates information molecules, each of which encodes a certain message and carries it to the Rx. The Rx is assumed to capture an information molecule when the molecule arrives at a surface of the Rx. The Rx propagates back ACK molecules, each of which encodes a receipt of the message to make a notification towards the Tx. The Tx is assumed to capture the ACK molecules when those are successfully arrived.

This paper assumes a pure random walk for diffusive molecule propagation. Diffusive movement is governed by the diffusion coefficient D on each dimension: $D = \partial x^2/(2\times\partial t)$. x denotes the distance of molecular movement during an amount of time t. When a molecule collides with another molecule, it randomly moves to another position with D.

Given wet laboratory implementations of molecular communication (e.g., [24]), the molecular fountain assumes DNA molecules, as information and ACK molecules, which can contain data by means of nucleotide sequences. DNA is structurally defined as a linear chain of repeating units of deoxyribonucleotide. Each deoxyribonucleotide is composed of a nucleobase, either adenine (A), guanine (G), cytosine (C), or thymine (T), as well as a five-carbon sugar called deoxyribose, and a phosphate group. A nucleobase of DNA encodes 2 bits of information because it has four choices (i.e., A, G, C, and T). It is approximately 0.34 nm long [25]; thus, up to $5{,}882$ bits/μm can be achieved. Therefore, molecular communications using large (heavy) molecules like DNA have a significant advantage in the high-density information over conventional molecular communications with small (lightweight) molecules (e.g., [1, 2]).

Promising approaches to engineer bio-nanomachines with communication capability include the modification of biological cells and the production of artificial cell-like structures using biological materials (e.g., a vesicle embedded with proteins) [8]. This paper assumes that bio-nanomachines are

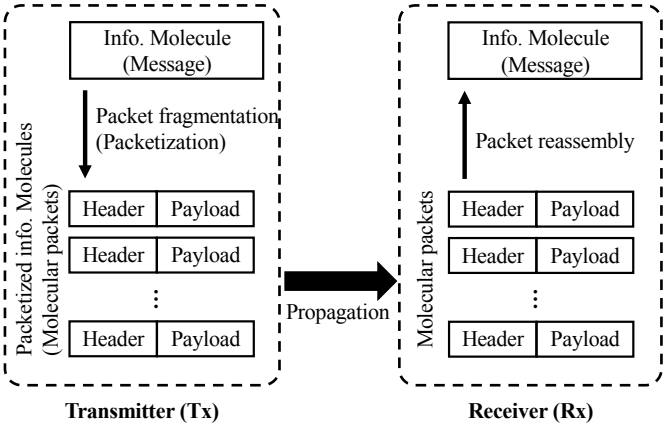


Fig. 1. Packet fragmentation and reassembly.

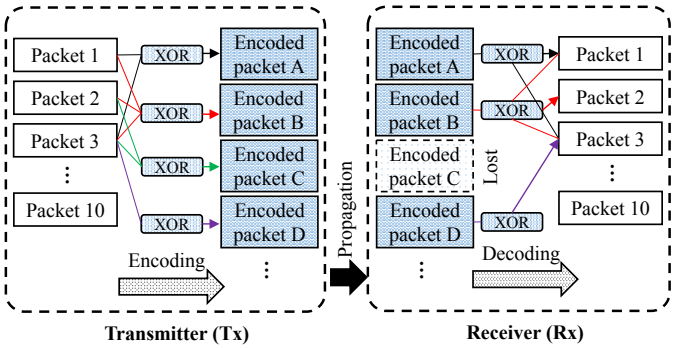


Fig. 2. Luby transform (LT) coding in molecular fountain.

realized as modified biological cells, which are known to potentially possess various communication-related functions including a transmission function to synthesize and release specific types of molecules, a reception function to capture specific types of molecules, logic gates to trigger programmed chemical responses upon receiving molecules, toggle switches (i.e., 1-bit memories) to retain communication-related states (e.g., ready-to-transmit and in-transmission/waiting states), and oscillators (i.e., clocks) to control the temporal timing of releasing molecules.

B. Packetization of Information Molecules

Molecular fountain employs the notion of packet fragmentation and reassembly [3]. As shown in Fig. 1, the Tx *packetizes* a large information molecule (i.e., a large DNA molecule containing a long nucleobase chain) into smaller pieces (i.e., smaller DNA molecules containing shorter nucleobase chains) and propagates the packetized information molecules in the environment. The Rx receives these packetized information molecules (or *molecular packets*) and reassembles the original information molecule. Packetization of information molecules is motivated by a fact that smaller DNA molecules diffuse faster [4] and arrive at the Rx with a higher probability [3].

Each molecular packet is assumed to be uniquely identifiable by a header (Fig. 1). A payload in the molecular packet contains a fragment of the original message. The header contains control information such as a receiver address to which the molecular packet is delivered, an identifier (or sequence number) for the original message that the molecular packet belongs to, and an identifier for the molecular packet.

The packet fragmentation and reassembly may be implemented by exploiting enzymes from biological cells, e.g., restriction enzymes to cut a DNA molecule into smaller fragments, and DNA ligases to join two DNA fragments into a larger one. Restriction enzymes may be embedded in the Tx and DNA ligases in the Rx. Chemical reactions to implement the packet fragmentation and reassembly are simple with these specific enzymes as well as a few small cofactors. The biochemical reaction to cut a DNA molecule into fragments is a hydrolysis reaction, which requires no energy. On the other hand, the biochemical reaction to concatenate DNA molecules requires chemical energy. The energy cost required for the proposed molecular communication scheme increases linearly with respect to the number of fragments.

C. Rateless Erasure Coding

In order to enhance reliability in molecular communications against molecule packet losses, the molecular fountain allows the Tx to propagate packetized information molecule (molecular packets) with a redundancy. It leverages the Luby transform (LT) [5] to take k packets and encode more than k packets. The Tx transmits those *encoded* packets toward the Rx. Fig. 2 illustrates a simple example where molecular fountain takes 10 packets, which are fragmented from an information molecule ($k = 10$), and propagates more than 10 encoded packets.

Molecular fountain applies an exclusive-or (XOR) operation to one or more packets to generate an encoded packet. The number of packets used in the XOR is called the *degree* of the encoded packet. All packets used to generate an encoded packet are called *neighbors* of the encoded packet. The encoded packets follow a certain degree distribution $\Omega(g)$. This paper uses the robust soliton distribution (RSD) [5]. The encoding process can be broken down into three steps: (1) Randomly choose a degree g by sampling $\Omega(g)$. (2) Choose g of the k packets uniformly at random. (3) Perform XOR of the g chosen packets. The output of this XOR operation is an encoded packet. The Tx repeats this encoding process with DNA screening (described in depth later) and propagates them one by one with a given transmission interval until it receives an ACK molecule from the Rx. Fig. 2 shows an example where the Tx has encoded four packets (packets A to D) so far. Their degrees are 2, 3, 2, and 1. For example, packet B is encoded as XOR combination of original packets 1, 2, and 3.

For the Rx to decode transmitted packets, molecular fountain uses a belief propagation decoder whose complexity is low in particular for erasure channels [26]. First, it identifies all degree-1 encoded packets and recovers their corresponding packets. These are moved to a storage referred to as the *ripple*. Packets in the ripple are *processed* one by one, which means they are XOR'ed with all encoded packets, who have them as neighbors. Once a packet has been processed, it is removed

TABLE I
PARAMETER SETTINGS

Parameter	Value
Size of the environment	150 μm x 150 μm x 150 μm
Diameter of Tx and Rx	5 μm
Tx to Rx distance (d)	30, 60, and 90 μm
Length of a nucleobase	0.34 nm
Length of a bigger info. molecule ($L_{\mathrm{m}}^{\mathrm{b}}$)	66.07 μm (48,579 bytes)
Length of a smaller info. molecule ($L_{\mathrm{m}}^{\mathrm{s}}$)	1.28 μm (940 bytes)
Length of an ACK molecule	0.14 μm (40 bits)
Length of a packet payload (L_{payload})	0.44 μm (32 bytes)
Length of a packet header (L_{header})	0.14 μm (40 bits)

from the ripple and considered decoded. The processing of packets in the ripple will potentially reduce the buffered packets to degree one, in which case the neighboring packet is recovered and moved to the ripple. Such an iterative decoding process can be explained in two steps: (1) Identify all degree-1 encoded packets and add their corresponding packets to the ripple. 2) Process a packet from the ripple, remove it afterwards and go to Step 1. Decoding succeeds when all packets are recovered. When the ripple size becomes zero, decoding has failed. Once decoding is successfully completed to reassembly the original message 1, the Rx transmits ACK molecules toward the Tx as a receipt of the message.

An example in Fig. 2 assumes that the Tx has propagated four encoded packets and the Rx has received three of them (packet C has been lost). Packet 3 can be directly recovered from packet D because packet D's degree is one. Then, packet 1 can be recovered with packets 3 and A. Given packets 1, 3, and B, molecular fountain recovers packet 2 despite the loss of packet C. As a result, we can recover three packets with four encoded packets by introducing one extra (redundant) packet in the presence of packet losses.

D. Biochemical Screening of DNA Molecules

As described above, we use packet screening for DNA molecules before propagating them in order to reduce the number of error-prone DNA molecules. Our scheme is designed to eliminate DNA molecules with high/low GC content and long homopolymer runs. High/low GC content and long homopolymer runs are major sources of errors in DNA synthesis and sequencing [27–30]. GC content is the ratio of guanine (G) and cytosine (C) nucleobases on a DNA sequence. For example, the GC content of a DNA sequence ACCTGCGAAT is 50% ($5/10$). Long homopolymer runs are repetitive DNA sequences such as AAAAA, GGGGG, CCCCC, and TTTTT. Schwartz et al. found that DNA molecules with 60% or higher GC content exhibit high dropout rates [27]. According to Ross et al. [28], when a DNA sequence contains four or more consecutive nucleotides of the same type, sequencing error increases significantly. Therefore, the proposed scheme carries out DNA screening in molecular packets such that the GC content will be within 45% and 55% and the length of a homopolymer run will be below four.

IV. PERFORMANCE EVALUATION

In this section, we evaluate molecular fountain through simulations based on the MolComKit simulator[1]. Table I shows simulation parameter settings, which follow realistic values found in biomedical engineering (e.g., [4]). Every result is shown based on $1{,}000$ independent Monte-Carlo simulations.

This paper considers two types of information: bigger and smaller ones. The former is simulated to contain a "big" message, $48{,}579$ bytes of data, assuming a full-color image with 640×480 pixels. The latter contains a "small" message, $4{,}000$ bytes of data, assuming text data. Regardless of information molecule sizes (i.e., message sizes), molecular fountain utilizes fixed-length molecular packets. The length of each packet (L_{packet}) is denoted as follows:

$$L_{\mathrm{packet}} = L_{\mathrm{payload}} + L_{\mathrm{header}}, \tag{1}$$

where L_{payload} indicates the length of a payload (Fig. 1), which is fixed to 0.44 μm. It is intended to contain 32 bytes of data. L_{header} denotes the length of a header, which is 0.14 μm long, capable of containing 40 bits of data. A "big" message is packetized to $1{,}519$ packets, and a "small" message is packetized to 30 packets when no redundancy is used for erasure coding. An ACK molecule is propagated as a single packet with no payload. It is 0.14 μm long and capable of containing 40 bits. The Rx releases 30 duplicated ACK molecules once the reconstruction of the message is accomplished.

The diffusion coefficient D and the radius R of a molecular packet are derived from its length L_{packet} with the following equations [4]:

$$D = \alpha L_{\mathrm{packet}}^{-\beta}, \tag{2}$$

$$R = \gamma / D, \tag{3}$$

where $\alpha = 2.8$, $\beta = 0.6$, and $\gamma = 0.34$ according to experimental values in [4].

Figs. 3 and 4 show how molecular fountain impacts the average round trip time (RTT) to accomplish a transmission of "big" and "small" messages, respectively, for a distance from Tx to Rx of $d = 30$, 60, and 90 μm. A number in parentheses indicates a standard deviation of RTT. RTT measures the amount of time since a (packetized) message leaves the Tx and until one of redundant ACK molecules first hits the Tx. The bottom of each figure shows the RTT results without using molecular fountain (denoted by "w/o Erasure Coding"). Fig. 3 illustrates that RTT improves simply by packetizing a message ($1.07\times$ speedup when $d = 90$ μm). This verifies that smaller molecules diffuse faster and arrive at the Rx with a higher probability than bigger molecules. In Fig. 4, packetization does not help in an improvement of RTT because the ratio of packet size to message size is not sufficiently low.

As depicted in Figs. 3 and 4, molecular fountain successfully improves RTT for both "big" and "small" messages. When packet transmission interval is set to 60 seconds, molecular fountain provides $2.8\times$ and $7.4\times$ speedups over

[1] http://www.cs.umb.edu/~jxs/molcomkit/

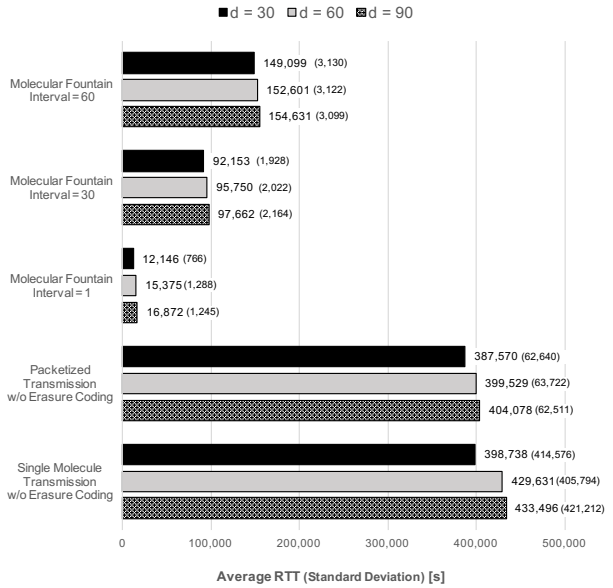


Fig. 3. RTT with and without molecular fountain ("big" message).

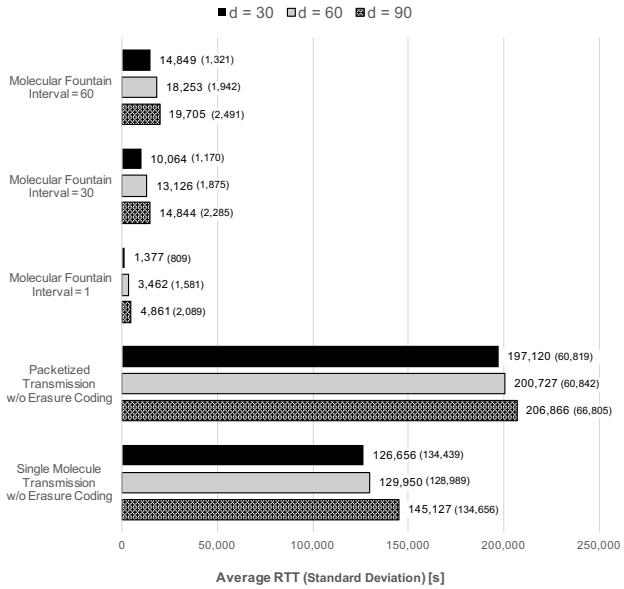


Fig. 4. RTT with and without molecular fountain ("small" message).

non-packetized transmission of "big" and "small" messages, respectively (for $d = 90$ μm). The performance gain increases as packet transmission interval decreases. With the interval of one second, molecular fountain provides $25.7\times$ and $29.9\times$ speedups over non-packetized transmission of "big" and "small" messages, respectively (for $d = 90$ μm). These results demonstrate that the proposed method is robust against molecule losses and improves RTT performance significantly.

Figs. 5(a) through 5(f) illustrate the cumulative probability of RTT for various simulation settings (different message types and Tx-to-Rx distances). When a message is not packetized, RTT distribution is long-tailed. This finding is consistent with the standard deviation results in Figs. 3 and 4. When the packet transmission interval is set to 60 seconds, the molecular fountain yields 99% and 98% lower jitter (standard deviation), compared to non-packetized transmission of "big" and "small" messages, respectively (at $d = 90$ μm). The results show that molecular fountain improves RTT jitter performance and makes RTT characteristics more predictable/reproducible.

Figs. 6 and 7 show overhead rates under different packet transmission intervals. An overhead rate indicates redundancy in packet transmissions; it is denoted as n/k where $k = L_{\mathrm{m}}/L_{\mathrm{payload}}$ and n is the total number of packets that the Tx sends until it receives an ACK. In both "big" and "small" message transmissions, overhead increases as packet transmission interval decreases. Figs. 8 and 9 reveal the trade-off relationship between overhead and RTT. The Tx-Rx distance has limited impact on the trade-off relationship, except for the case where packet transmission interval is set large to one second.

Figs. 10 and 11 show the trade-off relationship between overhead and message transmission failure (error). A message transmission is considered failed if a sufficient number of packets did not arrive at the Rx to decode the original message within a certain timeout period: $R_{\mathrm{avg}} + \frac{1}{5}R_{\mathrm{std}}$ where R_{avg} is the average RTT and R_{std} is the standard deviation of RTT. In "big" message transmissions (at $d = 90$ μm), failure rate reaches 30% when packetization and erasure coding are disabled. Molecular fountain decreases failure rate to 0% by increasing overhead to 171% or greater. In "small" message transmissions (at $d = 90$ μm), failure rate reaches 65% when packetization and erasure coding are disabled. Molecular fountain decreases failure rate to 0% by increasing overhead to $1{,}145\%$ or greater.

V. Conclusion

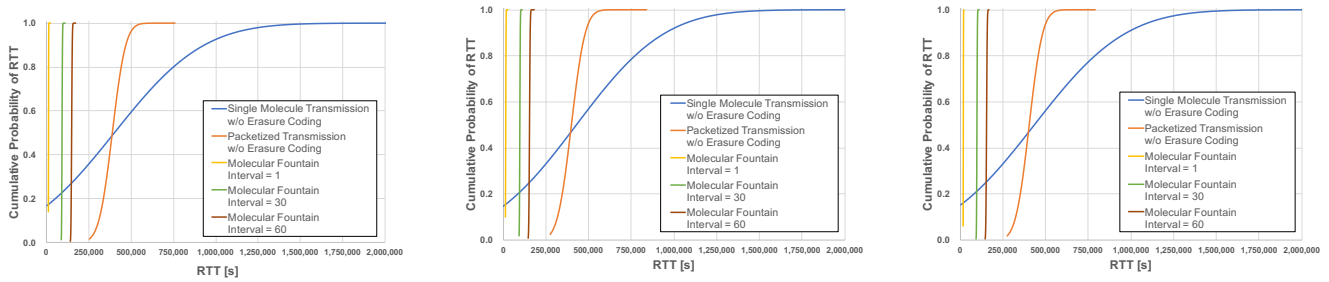
This paper proposes a new framework called molecular fountain, which enhances robustness against molecule losses for reliable transmission of information-encoded molecules between bio-nanomachines. Molecular fountain employs DNA molecules as information carriers and leverages molecular fragmentation (i.e., packetization) between Tx and Rx bio-nanomachines. It performs feedback-aided rateless erasure coding that respects biochemical constraints in DNA synthesis and sequencing to generate stable molecular packets. Simulation results show that molecular fountain significantly improves communication performance such as transmission latency, jitter, and error rate.

References

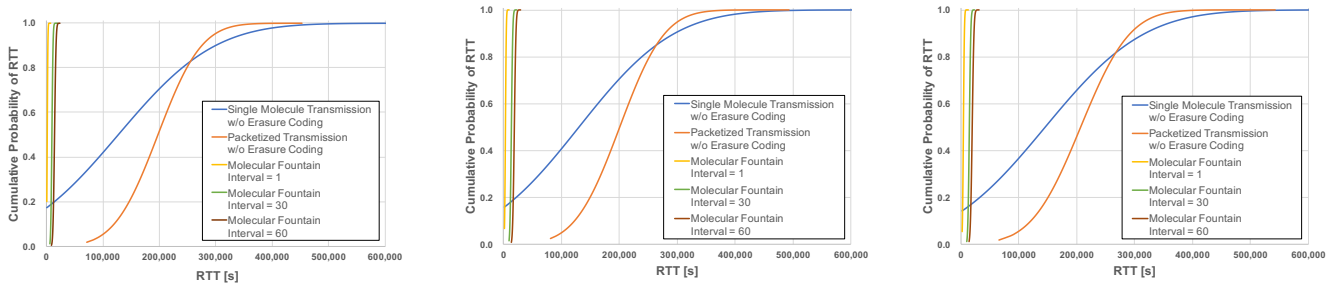
[1] M. Pierobon and I. F. Akyildiz, "Diffusion-based noise analysis for molecular communication in nanonetworks," *IEEE Trans. Signal Process.*, vol. 59, no. 6, 2011.

[2] K. V. Srinivas, R. S. Adve, and A. W. Eckford, "Molecular communication in fluid media: The additive inverse gaussian noise channel," *IEEE Trans. Inf. Theory*, vol. 58, no. 7, 2012.

[3] T. Furubayashi, T. Nakano, A. Eckford, Y. Okaie, and T. Yomo, "Packet fragmentation and reassembly in molecular communication," *IEEE Trans. NanoBiosci.*, vol. 15, no. 3, 2016.



(a) "Big" message: $d = 30\,\mu$m (b) "Big" message: $d = 60\,\mu$m (c) "Big" message: $d = 90\,\mu$m



(d) "Small" message: $d = 30\,\mu$m (e) "Small" message: $d = 60\,\mu$m (f) "Small" message: $d = 90\,\mu$m

Fig. 5. Cumulative probability of RTT for "big" and "small" messages.

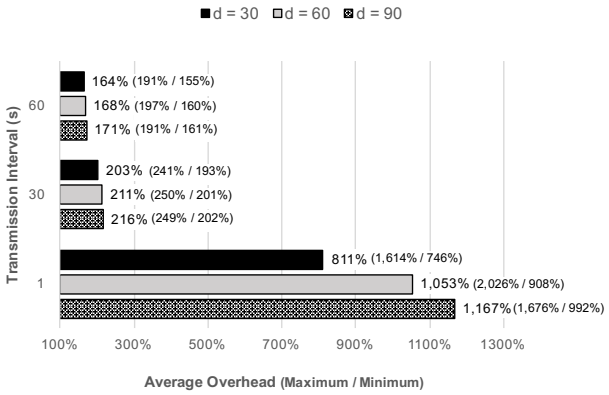


Fig. 6. Overhead for "big" message.

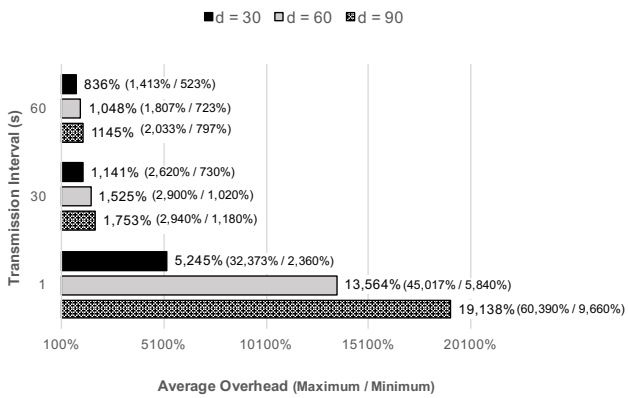


Fig. 7. Overhead for "small" message.

[4] D. E. Smith, T. T. Perkins, and S. Chu, "Dynamical scaling of DNA diffusion coefficients," *Macromolecules*, vol. 29, 1996.

[5] M. Luby, "LT codes," in *Proc. 43rd Annual IEEE Symposium on Foundations of Computer Science*, 2002.

[6] M. J. Moore and T. Suda, "Molecular communication: Modeling noise effects on information rate," *IEEE Trans. NanoBiosci.*, vol. 8, no. 2, 2009.

[7] N. Farsad, A. Eckford, S. Hiyama, and Y. Moritani, "A simple mathematical model for information rate of active transport molecular communication," in *IEEE INFOCOM Wksp Mol Nanosc Comm*, 2011.

[8] T. Nakano, M. Moore, F. Wei, A. V. Vasilakos, and J. W. Shuai, "Molecular communication and networking: opportunities and challenges," *IEEE Trans. NanoBiosci.*, vol. 11, no. 2, pp. 135–148, 2012.

[9] T. Nakano, Y. Okaie, and A. V. Vasilakos, "Transmission rate control for molecular communication among biological nanomachines," *IEEE J. Sel. Area Comm.*, vol. 31, no. 12, pp. 835–846, 2013.

[10] L. Felicetti, M. Femminella, G. Reali, T. Nakano, and A. V. Vasilakos, "TCP-like molecular communications," *IEEE J. Sel. Area Comm.*, vol. 32, no. 12, pp. 2354–2367, 2014.

[11] X. Wang, M. D. Higgins, and M. S. Leeson, "Simulating the performance of SW-ARQ schemes within molecular communications," *Simulation Modelling Practice and Theory*, vol. 42, 2014.

[12] C. Bai, M. S. Leeson, and M. D. Higgins, "Performance of SW-ARQ in bacterial quorum communications," *Nano Commun. Netw.*, vol. 6, no. 1, 2015.

[13] J. S. Mitzman, B. Morgan, T. M. Soro, J. Suzuki, and T. Nakano, "A feedback-based molecular communication protocol for noisy intrabody environments," in *17th IEEE Int'l Conference on E-health Networking Applications and Services*, 2015.

[14] H. O. Burton and D. D. Sullivan, "Errors and error control," *Proc. of the IEE*, vol. 60, no. 11, 1972.

[15] M. S. Leeson and M. D. Higgins, "Forward error correction for molecular communications," *Nano Commun. Netws*, vol. 3, no. 3, pp. 161–167, 2012.

[16] ——, "Error correction coding for molecular communications," in *2012 IEEE International Conference on Communications (ICC)*. IEEE, 2012, pp. 6172–6176.

[17] Y. Lu, M. D. Higgins, and M. S. Leeson, "Self-orthogonal convolutional codes (soccs) for diffusion-based molecular communication systems," in

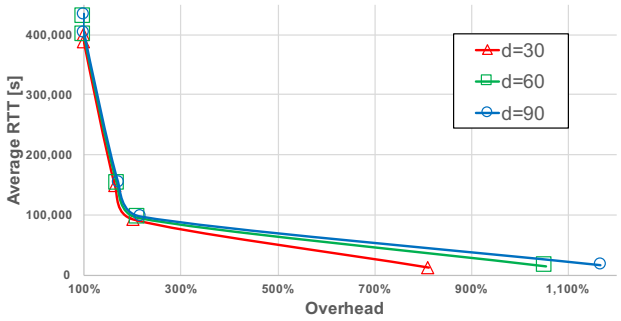


Fig. 8. Overhead-RTT trade-off for "big" message.

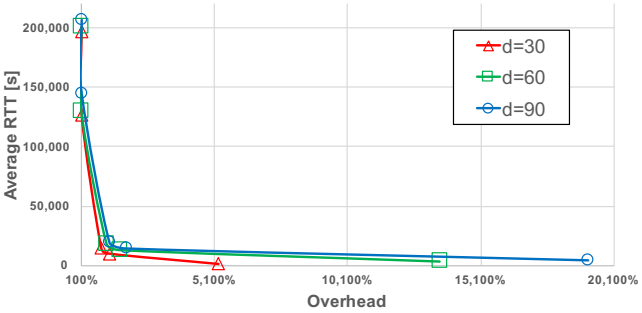


Fig. 9. Overhead-RTT trade-off for "small" message.

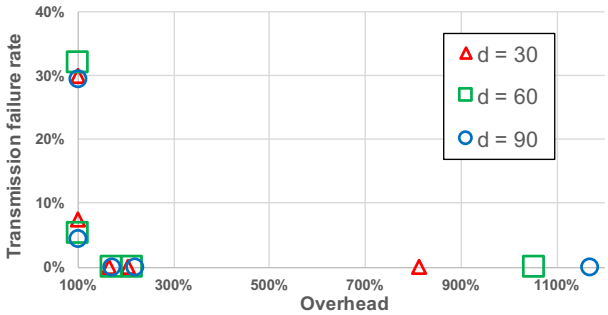


Fig. 10. Overhead-error trade-off for "big" message.

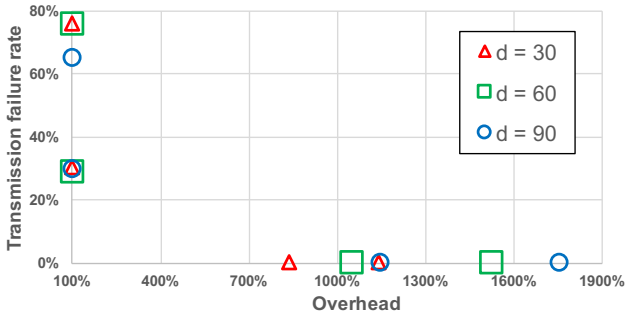


Fig. 11. Overhead-error trade-off for "small" message.

2015 IEEE International Conference on Communications (ICC). IEEE, 2015, pp. 1049–1053.

[18] C. Bai, M. S. Leeson, and M. D. Higgins, "Minimum energy channel codes for molecular communications," *Electronics Letters*, vol. 50, no. 23, pp. 1669–1671, 2014.

[19] Y. Lu, M. D. Higgins, and M. S. Leeson, "Comparison of channel coding schemes for molecular communications systems," *IEEE Transactions on Communications*, vol. 63, no. 11, pp. 3991–4001, 2015.

[20] ——, "Diffusion based molecular communications system enhancement using high order hamming codes," in *Communication Systems, Networks & Digital Signal Processing (CSNDSP), 2014 9th International Symposium on*. IEEE, 2014, pp. 438–442.

[21] H. Egashira, J. Suzuki, J. S. Mitzman, T. Nakano, and H. Fukuda, "Impacts of erasure coding on robustness against molecular packet losses in aqueous, collisional environments," in *Proc. 10th Int'l Conference on Bio-inspired Information and Communications Technologies*, 2017.

[22] ——, "Robust directional-diffusive hybrid molecular communication with parity-check erasure coding," in *Proc. Joint 17th World Congress of Int'l Fuzzy Systems Association and 9th Int'l Conference on Soft Computing and Intelligent Systems*, 2017.

[23] Y. Erlich and D. Zielinski, "DNA fountain enables a robust and efficient storage architecture," *Science*, vol. 355, no. 6328, pp. 950–954, 2017.

[24] M. E. Ortiz and D. Endy, "Engineered cell-cell communication via DNA messaging," *J. Biol. Eng.*, vol. 6, no. 16, 2012.

[25] T. Nakano, A. Eckford, and T. Haraguchi, *Molecular Communication*. Cambridge University Press, 2013.

[26] A. Shokrollahi, "Raptor codes," in *IEEE Trans. Info. Theory*, 2006, p. 25512567.

[27] J. J. Schwartz, C. Lee, and J. Shendure, "Accurate gene synthesis with tag-directed retrieval of sequence-verified dna molecules," *Nature methods*, vol. 9, no. 9, pp. 913–915, 2012.

[28] M. G. Ross, C. Russ, M. Costello, A. Hollinger, N. J. Lennon, R. Hegarty, C. Nusbaum, and D. B. Jaffe, "Characterizing and measuring bias in sequence data," *Genome biology*, vol. 14, no. 5, p. R51, 2013.

[29] N. Eroshenko, S. Kosuri, A. H. Marblestone, N. Conway, and G. M. Church, "Gene assembly from chip-synthesized oligonucleotides," *Current protocols in chemical biology*, pp. 1–17, 2012.

[30] B. C. Faircloth and T. C. Glenn, "Not all sequence tags are created equal: designing and validating sequence identification tags robust to indels," *PloS one*, vol. 7, no. 8, p. e42543, 2012.