On attractant scheduling in networks based on bacterial communication

Yunlong Gao  
Shanghai Jiaotong University, Shanghai, China

Sriram Lakshmanan, Raghupathy Sivakumar  
Georgia Institute of Technology, Atlanta, GA, USA

Abstract—In this paper, we explore the problem of attractant scheduling in networks based on bacterial communication. Bacterial communication is a communication paradigm between biological cells which involves the physical motion of flagellated bacteria such as E. coli from the transmitter to the receiver using chemical attractants. Although bacterial communication occurs in nature, engineering such communication links is non-trivial because the number of attractants is limited in practice. Thus scheduling among bacterial communication links is required when there are multiple information transmitter-receiver pairs in the same vicinity. We analyze this problem and model the delay and information loss in such environments. We validate our model using simulations. Our study provides new insights for the design of attractant scheduling algorithms in multi-node bacterial communication networks.

I. INTRODUCTION

Nanotechnology is the study of manipulating matter at the atomic and molecular scales. At such extremely small dimensions (of the order of nanometers), quantum mechanical effects govern the operation of materials. Understanding the operation of nanomachines promises new solutions for applications in several fields including medicine [2], environmental research, military technology [4], industrial and consumer goods [3]. At the nanoscale, the most basic functional unit is called a nanomachine. A nanomachine is equipped to perform very simple tasks such as computing, data storing, sensing and actuation. A nanonetwork is an interconnection of nanomachines, which can expand the capabilities of single nanomachines by providing them a way to cooperate and share information. Thus, communication among nanomachines becomes critical to realizing the true benefits of nanotechnology.

Traditional communication technologies are not directly applicable in nanonetworks due to the unique characteristics of the environment and the physical laws operating at nanoscales. Hence recent research has begun to study potential communication technologies for nanonetworks and can be classified into two main categories: Electromagnetic communication and Molecular communication. Electromagnetic communication in the context of nanonetworks focuses on implementing nano-scale antennas and actuators by using graphene or nanotubes [1]. Bio-molecular communication draws inspiration from nature and uses biological entities such as molecules or micro-organisms, e.g., bacteria, to transmit information between transmitters and receivers. In this work, we focus on networks based on bacterial communication. In such networks a nanomachine acts as a transmitter or receiver. Such nanomachines can be either naturally occurring biological cells or modified biological cells. The transmitter encodes its information into the DNA of a flagellated bacterium such as an Escheria Coli [8], which moves to the receiver to transfer information as illustrated in Figure 1.

In such networks, simultaneous communication between multiple transmitters and receivers is important. This communication paradigm forms the prototype of nanonetworks in our work. As an example, endothelial cells excrete Nitric Oxide (NO) which serves as a second messenger for many important cellular functions, particularly for signaling smooth muscle relaxation [5]. Endothelial cells also discharge a protein called Interleukin-6 (IL-6) [6], which plays an important role in activating B cells and thus enhancing the immune systems of a human being [7]. Another example is macrophage, which can also excrete both NO and IL-6 and lay effect on different cells. Thus, enabling communication from multiple transmitters to multiple receivers is useful to improve overall communication efficiency. In such a network with multiple transmitters and receivers, ensuring that the correct information reaches the correct receiver is important. This is accomplished by the release of specific chemicals called attractants by the receiver. The concentration gradient directs the bacteria to move toward the correct receiver. Since the number of attractants is limited [9], correct scheduling of multiple transmitter-receiver pairs is required to minimize the information delivery time in the network and forms the main focus of this paper. We develop analytic models for the delay and loss probability in a bacterial communication environment with two transmitter-receiver (T-
that considers scheduling aspects of bacterial communication to validate our model and to evaluate the performance of our algorithm. To the best of our knowledge, this is the first work that considers scheduling aspects of bacterial communication networks with multiple T-R pairs.

The rest of the paper is organized as follows: Section II presents a brief background of bacterial communication, while section III presents our analysis of the two T-R-pair scenario. Section IV presents our directions for algorithm design in a general network scenario and Section V presents our conclusions.

II. BACKGROUND AND PROBLEM

A. Network architecture

A bacterial communication network is composed of five components: the transmitter, receiver, message, information carrier and the medium. In this section, we describe these components in detail.

1. Transmitter: We define a transmitter as a nanomachine, which is in the size of most eukaryotic cells and can encode information by modifying molecules by means of chemical reactions. A transmitter emits certain chemicals called attractants that attract information carriers. It can also transfer messages to information carriers via plasmid conjugation.

2. Receiver: A receiver is a nanomachine that can extract a message from the information carriers and react to it accordingly. A receiver can also emit attractants in order to guide information carriers directionally. Moreover, each receiver is equipped with a unique molecular tag, which helps to identify that receiver. The molecular tags can be of two types: The first type is gemini-peptide-lipids. The receivers embedded with certain type of gemini-peptide lipids react with the information carrier with the same tag in response to an external stimulus (e.g., light, ions, and temperature). This allows a selective reception of information molecules at a receiver. The other type of tags is antibody/antigen. Each antibody binds to a specific antigen and the interaction is similar to the match between a lock and a key.

3. Information Carrier: An information carrier is a genetically modified bacterium which is able to accept a message from a transmitter and directionally transport the message to a receiver. Information carriers have two basic functions. First, they can carry information and transport it from the transmitter to an appropriate receiver through the medium. Furthermore, carriers act as information containers that encapsulate information during the propagation from a transmitter to a receiver and thus prevent the message (e.g., an DNA sequence) from degrading during the transmission. Among all bacteria, we focus on E. coli, because of the following reasons: (1) it has 4 to 10 flagella, which enable it to propel itself [12] and (2) An E. coli has small circular DNA sequences, called plasmids [13], which are widely used in genetic engineering for gene manipulation [14] and therefore E. coli is more amenable to engineering than other bacteria. (3) By modifying the genes of E. coli, we can make them generate specific tags on their membrane [21], and thus enable the selective reaction with the receiver who has the same tag.

4. Message: The message is an encoded DNA sequence. Since a DNA is a long polymer made from four nucleotides: Adenine, Thymine, Cytosine and Guanine (A, T, C and G) [16], the message is an encoded string of these bases.

5. Medium: The medium is the environment where the transportation of information takes place. e.g., in pheromone-based nanonetworks, the medium is the air and in bacteria-based nanonetworks, the medium is biological-friendly ambiance.

B. Bacterial Communication

In this section, we present the foundations of bacterial communication.

1. Chemotaxis: Bacteria such as the E. coli are equipped with sensory devices called “chemoreceptors”, which can measure the changes of concentration of certain chemicals. These changes are reported to the flagella that rotate in response and propel the E. coli [8]. The types of chemicals determine whether the bacteria move towards or away from the chemical. Bacteria are attracted by some chemical due to positive chemotaxis and repelled by other chemicals due to negative chemotaxis. In the case of E. coli, 12 types of chemoreceptors have been identified for positive chemotaxis and about 10 for negative chemotaxis [9]. Further, one chemoreceptor can detect more than one types of chemicals and one chemical can also be detected by more than one chemoreceptors [10]. This property makes E. coli an excellent information carrier in a multi-transmitter multi-receiver environment. Though, the strength of attraction varies with the attractants [17], we assume that they have the same effect on the E. coli for simplicity in this paper.

2. Attractant properties: In addition to the basic performance of each attractant in isolated conditions, related research has also performed experiments with multiple attractants in the same medium. Suppose one type of chemoreceptor can detect two different chemicals, e.g., chemical A and chemical B. If chemical A is present in high enough concentration to saturate its chemoreceptor, it will completely block the response to chemical B of the chemoreceptor [9]. Therefore, in order to avoid the situation mentioned above from happening, each chemoreceptor must detect only one kind of attractant. Since only 12 types of chemoreceptors have been identified so far, the number of attractants in the same medium cannot exceed 12 for the bacterial communication networks discussed in this paper. Additionally, related work has established that the E. coli can be genetically programmed to be attracted by a single attractant [18]. Note that due to the limitation of the present development of nanomachines, we assume that receivers and transmitters can only emit one type of attractant at any time instant.

3. Communication process: To begin with, each transmitter emits an attractant which causes one or more E. coli to move toward it. By plasmid conjugation, the transmitter embeds
its information into the DNA of the E. coli and releases it. The E. coli starts its movement away from the transmitter and alternates between the run state and the tumble state. In the run state, the E. coli moves at a constant velocity in a certain direction and drifts due to rotational diffusion. In the tumble state, the E. coli slows down or stops to change its direction. The E. coli constantly senses the concentration. When it moves up the gradient, it will increase the run length time. Thus, the E. coli could pick a random direction to begin with, but eventually adjusts its trajectory toward the direction of higher concentration of the attractant [24]. Simultaneously, the receiver releases the right attractant which causes the E. coli to be propelled toward it. In this manner, the E. coli reaches the receiver. If the molecular tag on the E. coli matches the receiver’s tag, the receiver extracts the information from the DNA of the E. coli by plasmid conjugation. Otherwise, the E. coli is released back to the medium. We discuss the exact mathematical models of this process in the next section and model the delay and loss probability for the messages.

C. Problem Formulation

We are interested in the total delay time to convey the information in an environment with multiple transmitters, receivers and attractants. We consider a two dimensional network where locations are expressed as $x$ and $y$ coordinates. For simplicity, we assume that the amount of information that each transmitter wants to send is identical in each pair. We also assume that each information carrier can carry finite amount of information, as a single plasmid in E. coli contains less than 1000 basepairs [15]. The total transmission time for all the T-R pairs can be expressed as

$$D = \sum_{i=0}^{n} t_i(x_T, x_R, y_T, y_R, m, N) \tag{1}$$

where $t_i$ is the transmission time for the $i$th T-R pair, and it is a function of the position of transmitters and the receivers, as well as the number of attractants $m$. $x_T$, $x_R$, $y_T$, and $y_R$ stand for the $x$ and $y$ coordinates of each transmitter and receiver, respectively, $n$ is the number of receivers in the nanonetwork and $N$ is the required number of packets to receive. Based on the above assumptions, our aim is to minimize $D$ by suitably assigning attractants to each of the T-R pairs.

III. TWO-LINK TOPOLOGY

In this section, we consider a topology with a single transmitter and two receivers. We analyze the communication process and model the information delivery time in this network. The transmitter wants to transmit two sets of message to two receivers. We consider the case where one and two attractants are used in this scenario. Using a single attractant for both receivers can cause interference. Hence we consider the case where the separation between the receivers is small and large. Based on the above conditions, we get four transmission scenarios as shown in Figure 2 to 5.

A. Discussion of cases

Case 1: In the first case, where the two receivers are far away from each other and there are two types of attractants in the networks, the information transmission is accomplished in two steps. The transmitter first emits one attractant (unique to $R_1$), attracting E. coli and embedding the information into their DNA. The E. coli that carry the message will then be attracted by the attractant emitted by $R_1$. Thus they move directionally to $R_1$’s vicinity and transfer the message to $R_1$. The transmitter performs a similar procedure with the other attractant to transmit message to $R_2$.

Case 2: In the second case where the two receivers are close to each other, the transmission procedure is similar to the first case.

Case 3: In the third case, since there is only one type of attractant and the receivers are separated, there can be interference between the transmissions of the two receivers i.e. the E. coli with the message intended for $R_1$, denoted as $M_1$, may go to $R_2$ instead because of the emission of the same (single) attractant from both $R_1$ and $R_2$. Likewise, the message for $R_2$ ($M_2$) may also get to $R_1$ by mistake. Thus, this information mis-delivery will lead to packet loss.

Case 4: Finally, in the fourth case, similar to the third one, the transmission to the two receivers will be simultaneous, as only one type of attractant is used. However, intuitively, in this case since the two receivers are near, an E. coli which goes to the unintended receiver may bounce back to the intended receiver due to the close proximity of the receivers. In practice, if the distance between $R_1$ and $R_2$ is at the range from 100nm to 1000nm, the gradient of attractants emitted by them will saturate within this area. If the E. coli with $M_1$ gets to $R_2$, it will find that the tags do not match and it cannot go to $R_1$ guided only by the traditional short range transmission paradigm [22] since the gradient is no longer existent.

B. System Delay and Packet Loss analysis

In this section, we analyze the delay and the packet loss for the nanonetwork in each of the above cases.
We begin by noting that with two different attractants, no packet losses happen since there is no interference among the two messages. Thus, the delay \( D \) for a single packet is the same for these two cases and can be expressed as follows:

\[
D = T_p(d_1) + T_p(d_2) + 2 \cdot T_c,
\]

where \( T_p(d_1) \) and \( T_p(d_2) \) stand for the transmission time for \( M_1 \) and \( M_2 \), respectively. Intuitively, the longer the distance, the larger the value of transmission time. In [22], Gregori et al. have presented an empirical equation for the transmission time as a function of distance based on their simulation:

\[
T_p(d) = 1.82d^2 + 4.49d + 0.17,
\]

Moreover, \( T_c \) is a fixed delay for one packet during transmission, and it contains DNA encoding delay, encapsulating and decapsulation delay, as well as DNA expression delay. Normally, it is equal to 122 minutes [23]. We call the \( 2 \cdot T_c \) as the fixed delay.

As we have discussed in III, due to the constraints of the attractant, information mis-delivery leads to packet loss and time delay. Since the E. coli adjusts its moving direction according to the concentration change generated at the receiver, we assume the possibility that the E. coli moves to a certain receiver is proportional to the concentration gradient generated by that particular receiver. Furthermore, the concentration distribution of a molecule emitted at a constant rate of \( Q \) mol/s from a fixed point is

\[
U(r,t) = \frac{Q}{1000 \cdot 4\pi r} \operatorname{erf} \left( \frac{r}{\sqrt{4Dt}} \right)
\]

with \( d \) the distance from the point of release in \( m \), \( t \) the time during which the particles have been released in \( s \) and \( D \) the diffusion coefficient in \( m^2/s \). For simplicity, we only consider the concentration in stationary level, and thus \( U(r) = \frac{Q}{1000 \cdot 4\pi r} \).

We denote \( P_1 \) and \( P_2 \) as the possibility that E. coli is attracted by \( R_1 \) and \( R_2 \), respectively. Based on the discussion above, we can get \( P_1 = \frac{d_1}{d_1 + d_2} \) and \( P_2 = \frac{d_2}{d_1 + d_2} \). Therefore, in this case, we can derive the network delay as follows:

\[
D = \max\{T_p(d_1), T_p(d_2)\} + \frac{T_c}{P_1} + \frac{T_c}{P_2},
\]

Moreover, from the above analysis, we can easily derive the packet loss as follows:

\[
L = \frac{d_1^4 + d_2^4}{(d_1^2 + d_2^2)^2}.
\]

In case IV, where there is only one type of attractant and the two receivers are near each other, the E. coli’s movement pattern has not been characterized in related work. Hence, we use simulations to characterize the delay in the case.

As described in Section II, at any moment, an E. coli is either in the running or tumbling [24] state and alternates continuously between these states. According to [24], both running and tumbling time are exponentially distributed. During runs, bacteria swim in an approximately straight direction and in tumbles they spin on their position with respect to the former running direction. We assume that the duration of run and tumble are exponentially distributed with mean of 0.001s for runs and 0.1s for tumbles. However, in case IV, because the distance between the two receivers is small, the attractant concentration is nearly saturated within this area.

1) Runs: The E. coli drifts in the communication medium due to the rotational diffusion [24], which makes E. coli change their direction by a mean square angular deviation on each axis of \(< \theta^2 >= 2D_r t\), where \( D_r \) is the rotational diffusion coefficient and \( t \) the time. A typical value being \(< \theta^2 >= 1.12 \text{rad} \).

2) Tumble: After running for a while, an E. coli usually slows down or stops to change its direction, and then runs in the new direction until the next tumble state begins. Suppose the direction angle in time \( t \) is \( \theta_t \), then in the next time slot \( t+1 \), the angle will be as follows [22]:

\[
\theta_{t+1} = \theta_t + \gamma
\]

where \( \gamma \) is a random angle deviation whose p.d.f is subject to

\[
f(\gamma) = \begin{cases} 
\frac{1}{2}\cos(\gamma/2), & |\gamma| \leq \pi \\
0, & |\gamma| > \pi
\end{cases}
\]
We build the above model into our simulator. Our results are shown in Fig. 6 to Fig. 8. Since there is considerable randomness involved in the run and tumble process, the E. coli may not reach the destination within a finite time. Fig. 6 and 7 present two scenarios illustrating the success and failure to reach the destination, respectively. Note that the coordinates of the E. coli and the destination are (0,0) and (100,0), respectively. Moreover, we can easily get the packet loss in this direction at first is 1. Considering the symmetry of the network topology, we cannot get to the right destination if it goes to the wrong one. Therefore, we conclude that in case IV, an information carrier cannot get to the right destination within 1000s. If the E. coli does not get to the destination within certain time, we consider the system delay in this case:

\[ D = T_p(d) + 4 \cdot T_c, \]  

(9)

where \( d \) is the distance between the transmitter and the receiver. Moreover, we can easily get the packet loss in this case is 1/2.

From the above analysis, we can create the Table I to summarize the network delay and the packet loss for the four cases.

<table>
<thead>
<tr>
<th>Case</th>
<th>Delay</th>
<th>Loss</th>
</tr>
</thead>
<tbody>
<tr>
<td>I/II</td>
<td>( T_p(d_1) + T_p(d_2) + 2 \cdot T_c )</td>
<td>0</td>
</tr>
<tr>
<td>III</td>
<td>( \max(T_p(d_1), T_p(d_2)) + \frac{d_1}{T_c} + \frac{d_2}{T_c} )</td>
<td>( d_1, d_2 )</td>
</tr>
<tr>
<td>IV</td>
<td>( D = T_p(d) + 4 \cdot T_c )</td>
<td>1/2</td>
</tr>
</tbody>
</table>

**TABLE I**

**SUMMARY TABLE FOR THE FOUR CASES**

IV. ALGORITHMIC DIRECTIONS FOR SCHEDULING IN M-RECEIVER NETWORKS

We describe how the analysis of the basic topology can be used to derive insights for algorithm development in more generic topologies. We consider the case where one transmitter wants to transmit specific messages to different receivers, as shown in Fig. 9. To begin with, we assume that each receiver is \( d \mu m \) away from the transmitter. Suppose the set of receivers is \( R = \{1, 2, \ldots, n\} \), and the set of attractants that are available in this network is \( A = \{1, 2, \ldots, m\} \), and \( m \leq n \). Assume that for one type of attractant \( i \in A \), \( m_i \) receivers emit it as the attractant to attract E. coli. Therefore, we derive the expression for the delay as follows:

\[ D = m \cdot T_p(d) + \sum_{i=1}^{m} m_i^2 \cdot T_c \]  

(10)

where \( T_p(d) \) can be obtained from Eq.3, and \( m_i \) is subject to \( \sum_{i=1}^{m_i} m_i = n \). Note that the term \( \sum_{i=1}^{m_i} m_i^2 \cdot T_c \) is the fixed delay for the network. For each receiver \( j \) that emits attractants \( i \), the expected number of attractants it receives is \( m_i \) since other \( m_i - 1 \) receivers are emitting the same kind of attractant as well. Therefore, the fixed delay for all the messages which contains the gene that can make carriers attracted by attractant \( i \) is \( m_i^2 \cdot T_c \). By summing this term, we can get the expression for the whole fixed delay, i.e., \( \sum_{i=1}^{m_i} m_i^2 \cdot T_c \). Given \( m, d, n \) and the constraint that \( \sum_{i=1}^{m} m_i = n \), it is easy to note that the assignment that minimizes the system delay is to assign \( m_i = \frac{n}{m} \) receivers to the same attractant.

**Insight 1:** For closely spaced receivers, the best strategy is to allocate the attractants equally among the receivers.

In order to optimize the system delay, we run the simulation for the system mentioned above, the result is shown in Fig. 10 and 11. ¹ In the former figure, where the number of receivers is less than 12, increasing the number of attractants does not exert any impact on the delay when it exceeds the number of receivers. The latter figure shows, however, there exists an optimized assignment of the number of attractants between 1 and 12, which can minimize the system delay.

Next, we consider the scenario where the distance between each receiver and the transmitter is different. Suppose the distance between receiver \( R_i \) and the transmitter is \( d_i \), and we rearrange the receivers in the decreasing order of \( d_i \), so that \( d_1 \leq d_2 \leq \ldots \leq d_i \leq \ldots \leq d_n \), as the Fig. 9 shows and group them into sets. Thus we can get that \( T_p(d_1) \leq T_p(d_2) \leq \ldots \leq T_p(d_i) \leq \ldots \leq T_p(d_n) \). Furthermore, we denote the link between the transmitter and \( R_i \) as \( l_i \). Assume that there are still \( m \) types of attractants in the network. Our

1In our simulation, we set \( d = 50 \mu m \) and \( N = 1 \).

Fig. 9. System with multiple receivers.
\begin{align*}
D &= T(\text{Propagation Delay}) + T(\text{Fixed Delay})
\end{align*}

where the expression of the fixed delay the same as the former case: $\sum_{i=1}^{m} m_i^2 \cdot T_c$. Since there is a trade-off between propagation delay and fixed delay in the different-distance case, solving the optimization problem is not trivial. However, our analysis in Section III leads us to a greedy solution for attractant scheduling in this case.

**Insight 2:** By partitioning the $n$ links into $m$ groups after rearranging the receivers in the decreasing order of distances, we can obtain a reasonable balance between propagation delay and fixed delay.

In summary, we have identified two key insights for attractant scheduling in multiple receiver networks that can be used for algorithm development. As part of future work, we intend to explore the optimal solution to this problem and compare its performance to an algorithm which uses the above insights.

\section*{V. CONCLUSION}

In this paper, we discuss the scheduling problem in bacterial communication networks with multiple T-R pairs. We derive the network delay and packet loss in four basic cases. We also characterize the movement pattern of E. coli when there is no concentration gradient in the medium. We then extend the simple case to a multi-receiver topology. Our future work will aim at expanding the basic scenario to generic topologies with multiple transmitters and receivers.

\section*{REFERENCES}