Source Addressing and Medium Access Control in Bacterial Communication Networks

Bhuvana Krishnaswamy, Raghupathy Sivakumar
School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA, USA

ABSTRACT

In this work, we focus on the problem of *source addressing* in multiple source single receiver bacterial communication network. We propose *amplitude-addressing*, where the amplitude of transmitted signal is assigned as address of the source. We analyse the performance of the network with different addressing mechanisms and propose an optimum address sequence for a given network design. We also show that amplitude-addressing implicitly solves the problem of medium access control.

1. INTRODUCTION

Nano-scale communication strategies can be categorized into two broad domains depending upon their target environment: *electromagnetic communication (EM)* at the nanoscale involves the extension of traditional EM based communication techniques for use in non-biological applications [2,8]; and *molecular communication* involves strategies (typically bio-inspired) for use in biological applications [1,9,13, 20]. In recent years, bacteria have emerged as promising candidates for nano-machines in biological applications [7].

Genetic engineering of bacteria to introduce or delete DNA for specific traits (e.g., bioluminescence, motility, adhesion, etc.) has enabled recent advancements in synthetic biology [12]. Many bacteria utilize a process called quorum sensing, whereby bacterial cells behave as transceivers that interact with one another, relaying signals by transmitting and receiving chemical signal molecules [4, 17]. Using the power of synthetic biology and the inherent transceiver properties, bacterial nano-machines hold much promise to be used in biological applications such as toxicology, biofouling, and biosensing. For e.g, receiver bacteria have been used as biosensors to detect the presence of metals [6], and to detect arsenic pollution [19]. *The context for this work is thus molecular communication between bacterial populations.* Specifically, we consider a system in which bacterial populations are used as transceivers connected through microfluidic pathways for molecular signals. Molecular communication between bacterial population has been the topic of study in many nano-communication works. [16] proposes modulation techniques to improve the throughput performance, [10, 15] analyses the capacity of a bacterial communication network, [3, 18] proposes mathematical models for the transceiver and channel.

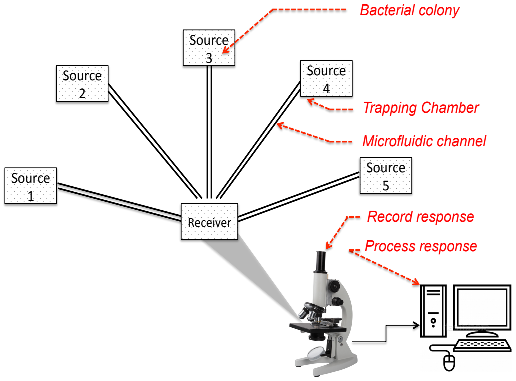


Figure 1: Network Setup

All these works focus on a single source single receiver topology. In practice, a network will have more than one transceiver. The focus of this work is a multiple source single receiver topology shown in Figure 1. Such a topology is most common in sensing applications where multiple sources communicate to a single sink/receiver. The sources share the channel and the resources at receiver. While the sources can broadcast their information, receiver must know the address of each source to make use of the information it receives. In this paper, we make the following contributions:

First, we propose an addressing mechanism for a multiple source single receiver topology. We propose a source addressing mechanism called $Amplitude-addressing$, which uses the amplitude of the signal transmitted as the address of source. Each source is assigned a unique amplitude. The sources transmits signals with the assigned amplitude. When multiple sources transmit simultaneously, receiver receives the sum of amplitudes, from which receiver identifies the components of the sum and hence the address of the sources.

Second, we analyse the performance of different addressing sequences and propose an optimum set of addresses to minimize address resolution error and maximize the number of sources accommodated in the network.

The rest of the paper is organized as follows. Section 2 presents the advantages and disadvantages of different source addressing mechanisms, Section 3 presents the address allocation in amplitude addressing with 100% interference cancellation, Section 4 analyses address allocation to minimize interference in a dense network, Section 5 presents

NANOCOM'15, September 21 - 22, 2015, Boston, MA, USA

http://dx.doi.org/10.1145/2800795.2800801 ...$15.00.

the simulation results of the performance of amplitude sequences proposed. Section 6 discusses some of the issues and future work and concludes the paper.

2. SOURCE ADDRESSING

In this work, we focus on a network topology with multiple sources and single receiver as illustrated in Figure 1. Sources and receiver are bacterial populations. The source transmits molecules of a given concentration(amplitude) for a given duration. The transmitted signal $m(t)$ represented in equation 1 is a rectangular pulse with concentration A transmitted for T seconds. The transmitted molecules propagates through a microfluidic channel and trigger the receiver to generate Green Fluorescent Protein (GFP) in response. The receiver samples GFP response to find the transmitted signal. In this work, we assume zero sampling and demodulation error i.e., we assume that the receiver can detect the rectangular pulse with zero error. Extremely small sampling and demodulation error can be achieved by increasing the number of bacteria per transceiver [11]. The modulation considered is On-Off-Keying (OOK); the presence and absence of signal represents bits *1* and *0* respectively.

Addressing in most cases is used to identify/specify the destination. For e.g., in postal services and e-mails, address is used to find the receiver and identify sender/source. *In this work, we focus on source addressing, where receiver utilizes the addressing mechanism to identify the source and all the sources communicate to the same receiver.*

Bits as Address

In electromagnetic communication, a pre-determined number of bits is encoded in the header of every packet. On receiving a packet, the receiver identifies the sender from these address bits. This same addressing mechanism is inefficient in bacterial communication. Due to high delays, the data-rate is very low (order of 10^{-5} bits per second, [16]) and adding more bits as address will increase per-frame delays and decrease per-user throughput.

Bits can be used as global address but has high delays, reduced network throughput and will need a MAC protocol.

An addressing that does not affect the throughput of the network is needed in such high-delay networks i.e., address embedded in the signal. Such a mechanism does not require extra bits for address and hence will be an efficient communication with reduced delay and improved system throughput. We refer to such an addressing as *Embedded Addressing.* In this section, we explore different characteristics of the transmitted signal to embed address and present the advantages and disadvantages of each addressing mechanism.

Embedded Addressing cannot be used as a global address as the characteristics of the signal are local to the network. Hence, a global and local address is needed, in which the signal characteristics can be used for local address. Assigning global address is out of the scope of this work. The 3 characteristics of a signal to embed local address are molecule type, signal duration and signal amplitude.

$$m(t) = \begin{cases} A, & 0 \leq t \leq T. \\ 0, & \text{otherwise} \end{cases} \tag{1}$$

where A is the amplitude of the signal, T is the duration. We consider each characteristic of the signal as address and analyse the feasibility of each mechanism.

Type of Molecule as Address

The amplitude and duration of a signal is same across all sources and address is embedded in the *type* of molecule. Each source is assigned a unique molecule. The receiver must be capable of receiving all the molecules assigned as address. Based on the molecule received, receiver assigns the bits received to the corresponding source. A receiver designed to receive more than one molecule can receive all molecules simultaneously and hence can also act as a MAC protocol that allows all sources to use the channel simultaneously without any MAC overhead.

But, the number of distinct molecules that can be identified by a receiver is limited and thus limits the number of addresses. *Molecule address* also requires the receiver to know all the sources during receiver design.

MAC is a by-product of molecule type addressing but is not scalable and transceiver designs are complex.

Embedding address in duration and amplitude allows all sources to transmit the same molecule. The receiver is designed to receive only one type of molecule simplifying receiver and source designs.

Signal Duration as Address

The amplitude of signal and type of molecule is fixed across all sources and address is embedded in signal duration. Each source is assigned a unique duration. When a source has data, it transmits a signal with a given amplitude and the duration assigned to it as address. When only one source is transmitting at a time, the receiver identifies the address from the duration. When multiple sources have data to send, the receiver receives the sum of signals. From the received sum, receiver has to identify the individual components. A source can begin transmission at any time and hence the time between signals cannot be controlled. In order for the receiver to identify all durations, no duration can be the sum of any 2 or more other durations assigned. If this condition is not satisfied, due to random delays, signals can overlap and be decoded with error. By assigning increasing durations, per-frame delay of each source is different and leads to unfair throughput. Delay in the network will increase by assigning increasing duration as address.

Random delays increases error and network delays. Throughput is unfair due to varying bit periods.

Signal Amplitude as Address

In electro-magnetic communication, amplitudes have been used for modulating the signal in Amplitude-Shift Keying. Here, we consider the use of amplitudes as address. Similar to duration as address, each source is assigned a unique amplitude. A source transmits its signal with the assigned amplitude for a duration that is fixed for the network. The receiver, on receiving an amplitude, maps the amplitude to the respective source. When multiple sources transmit at the same time, receiver receives the sum of amplitudes at every instant. To identify all sources, receiver must find the individual amplitudes given the sum of amplitudes. Therefore, the amplitudes must be assigned such that the receiver can identify the individual components given the sum.

Simple to implement and scalable. Throughput is fair but the number of sources is limited by receiver design.

Following the above discussions, we choose amplitude-addressing in this work. In the next section we present an optimum address allocation mechanism that can provide 100% interference cancellation.

3. ADDRESS ALLOCATION WITH INTERFERENCE CANCELLATION

The amplitudes assigned determine the correctness of address resolution at receiver. Incorrect identification of addresses by the receiver leads to address-resolution error (ARE). For the receiver to identify every source with zero ARE, the receiver should be capable of finding all the components given a sum. For e.g., if the amplitudes chosen are 1,2,3 the above condition cannot be satisfied; when sources with amplitude 1 and 2 transmit at the same time, it can be misinterpreted as source with amplitude 3. Thus, for zero ARE, sum of every combination of assigned amplitudes must be unique. This problem has been studied in number theory as "Distinct subset sum (SSD)". A sequence is called an SSD if and only if the sum of every subset of the sequence is unique [5]. An example SSD is,

$$S = \{1, 2, 4, 8\} \tag{2}$$

The set of amplitudes assigned as address is the ***Sequence S***. The elements of the set are the *amplitudes* assigned and the different combinations of these elements are the subsets. Inorder to achieve zero error, the subset sums must be distinct. Majority of works on SSD in the field of Number Theory focuses on finding the limit of maximum value in a subset sum sequence [14]i.e.,

$$f(n) = min\{max\ S : S \text{ has distinct subset sums and } \mid S \mid= n\} \tag{3}$$

In bacterial communication, the maximum amplitude a receiver can identify is determined by the receiver design.

The first constraint in *amplitude addressing* is thus the maximum receivable amplitude R_{max} beyond which the receiver saturates. If an amplitude greater than R_{max} is transmitted, it is received by the receiver as R_{max}.

The second constraint is the step size of the levels of amplitudes that can be distinguished by the receiver. R_{max} and step size of the amplitudes determines the number of amplitude levels that can be distinguished by the receiver. To accommodate the constraint on the step size, we assume integer amplitudes that can be multiplied by the step size. Therefore, in *Amplitude-addressing* we use integers to analyse the performance of the network. The addresses proposed here can be used in a network with a step size greater or less that "1" by multiplying the proposed addresses with the appropriate step size.

To achieve zero ARE, the requirements are, 1. sum of each subset is unique 2. maximum subset sum $\leq R_{max}$

Theorem 3.1 *If maximum sum is limited, Binary sequence is an optimum sequence that renders zero ARE.*

PROOF. In **S** with n elements, there are $2^n - 1$ subsets and hence $2^n - 1$ sums. To get zero ARE, these $2^n - 1$ sums must be distinct. Each sum must be different from other by atleast *1* and hence the sum of all n elements in the sequence is atleast $2^n - 1$. *Binary sequence*, a sequence of powers of 2 satisfies the above condition and is thus an optimum address sequence if the maximum sum is the constraint. □

An element a_i in binary sequence $S : \{2^0, 2^1, 2^2, ..\}$ is the smallest number that is not the sum of any combinations of intergers $< a_i$. Figure 2(b) shows the number of sources a binary sequence can support for increasing R_{max}. As expected, the logarithmic dependence of number of sources N_s on R_{max} limits the number of sources a binary sequence can allow in a network. A binary sequence can allow upto $N_{log} = \log_2 R_{max}$ sources. Binary sequence amplitude addressing implicitly solves medium access. MAC is needed when a channel or receiver is shared by multiple sources to allow a fair chance for sources to access the channel. Amplitude-addressing allows multiple sources to use the channel simultaneously. Even though signal from multiple sources collide at receiver, binary sequence addressing helps receiver recover all signals without any error from the collided signal. Thus, it reduces interference noise at the receiver to zero, achieving an interference free and fair MAC.

As a corollary of Theorem 3.1, when the number of sources is $> R_{max}$, ARE is non-zero. In turn, interference at the receiver is non-zero. The choice of amplitudes determines ARE and hence interference at receiver. In the next section, we present an algorithm to choose an address sequence that can allow more than N_{log} users with minimum ARE.

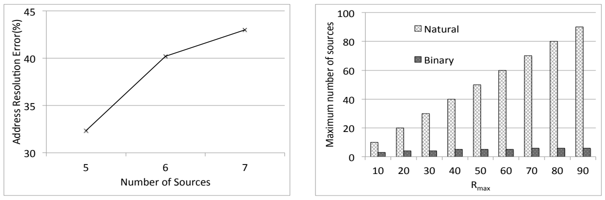
4. ADDRESS ALLOCATION WITH PARTIAL INTERFERENCE CANCELLATION

Following Theorem 3.1, the maximum number of sources a bacterial communication network can allow with zero ARE is $log_2 R_{max}$. When the number of sources N_s is greater than N_{log}, a simple address sequence is to use all integers in the range $[1, R_{max}]$ i.e., the set of natural numbers. The maximum number of sources a natural number sequence can support is R_{max}. *Natural sequence*, the set of integers from 1 to R_{max} uses all amplitudes that can be distinguished by the receiver and supports maximum N_s for a given R_{max}. Maximum number of sources for increasing R_{max} supported by binary and natural sequences is plotted in Figure 2(b). N_s by natural sequence increases linearly with R_{max} whereas that of binary increases logarithmically.

Though *Natural sequence* can support maximum number of sources, the address resolution error increases with increasing number of sources. We illustrate the ARE performance of natural sequence using the following example.

Consider a network where the receiver saturates above $R_{max} = 5$. The network can assign address to a maximum of 5 sources. Let us assume sources T1,T2,T3,T4 and T5 are assigned amplitudes 1,2,3,4 and 5 respectively i.e., $S : \{1, 2, 3, 4, 5\}$. When T3 transmits, receiver can interpret 3 as T3 or T1 and T2 transmitting simultaneously. Since all signals have the same duration, signals from 2 sources can add up and appear to be a single signal. Each possible combination that can add up to a sum is referred as a *configuration*. For e.g., the sum 5 can be obtained in 3 ways $\{(2, 3), (1, 4), (5)\}$ and $(1, 4), (2, 3)$ and (5) are configurations of 5. Figure 2(a) shows the ARE performance of natural sequence. In the simulations used to obtain Figure 2(a), every source was assumed to have 50% probability of transmitting a *1* and always had data to transmit. It can be observed that the address resolution error increases with increasing N_s. With increasing N_s, number of configurations that add up to a sum increases thus reducing the probability of choosing the right configuration. At $N_s = 5$, ARE is 33%.

In the rest of this section, we study the parameters that affect ARE, derive probability of success and propose an



(a) Address resolution error (b) Number of sources

Figure 2: Performance of Natural sequence

algorithm to choose an address sequence that can allow more than N_{log} sources and minimize ARE.

To analyse the parameters that affect ARE, we derive the probability of success per source from which probability of address error is calculated.

$$\Pr(T_i \text{ success}) = \sum_{x=0}^{x=Maxsum} \Pr(sum = x) * \{\Pr(T_i \in \mathbf{Rx} \mid sum = x) * \Pr(T_i \in \mathbf{Tx} \mid sum = x) + \Pr(T_i \notin \mathbf{Rx} \mid sum = x) * \Pr(T_i \notin \mathbf{Tx} \mid sum = x)\} \tag{4}$$

$$\Pr(T_i \in \mathbf{Rx} \mid sum = x) = \frac{N_{T_i}(x)}{N(x)} \tag{5}$$

where, $N(x)$ is the number of configurations for sum x and $N_{T_i}(x)$ is the number of configurations of x thats has T_i as an element. **Rx** is the set of addresses received and **Tx** is the set of of addresses transmitted. $\Pr(T_i \in \mathbf{Tx} \mid sum = x) * \Pr(sum = x)$ is the sum of probabilities of configurations of x that contains T_i. The total number of successful frames received is bounded from above by the sum of probability of successes of individual sources. We make the following observations based on the probabilities in equation 4.

Insight 4.1 *Higher the number of configurations per sum, lower is the probability of success on receiving the sum.*

Insight 4.2 *At low probability of transmitting a* 1*, the number of colliding signals is* ≤ 2 *with high probability.*

When the number of collisions is ≤ 2, an address sequence that has high probability of success for configurations with 1 or 2 elements will be an optimum sequence. A shifted natural sequence can be a good option. Shifted (natural) sequence is a natural sequence with the first element shifted by $\frac{R_{max}}{2}$, $S : \{\frac{R_{max}-1}{2}, \frac{R_{max}-1}{2}+1, ..., R_{max}-1\}$. A maximum of $\frac{R_{max}}{2}$ sources can be supported using this sequence. An address in shifted sequence cannot be written as a sum of any 2 addresses i.e., $a_i + a_j \neq a_k$ where $a_i, a_j, a_k \in S$ and $i \neq j \neq k$ Therefore, when one source transmits a *1*, it will be received without any error and when 2 sources transmit *1*, the error is smaller compared to other sequences since error in one address will not affect others.

A shifted natural sequence allows $\frac{R_{max}}{2}$ number of sources. If the number of sources is greater than $\frac{R_{max}}{2}$, we extrapolate the shifted sequence i.e., $S : \{R_{max} - N_u - 1, R_{max} - N_u - 2, ...R_{max} - 1, R_{max}\}$. The extrapolated sequence does not hold the property of shifted sequence. But, the number of elements that can be written as sum of 2 other elements is the lowest in this sequence since its elements are decreasing integers from R_{max}.

Insight 4.3 *At high probability of transmitting a* 1*, the number of colliding signals is* $\geq N_{max}$ *with high probability.*

When more than N_{max} sources collide, Natural sequence will have a high probability of success. N_{max} is the number of sources such that $\sum_{x=1}^{N_{max}} x = R_{max}$. When the number of sources transmitting bit *1* at a given time increases, the probability of receiving configurations that sum to $\geq R_{max}$ increases. All sums greater than R_{max} is received as R_{max} by the receiver. On receiving a sum R_{max}, the receiver chooses one of the configurations with sum $\geq R_{max}$. The probability of choosing a particular configuration is$\frac{1}{N(x \geq R_{max})}$, where $N(x \geq R_{max})$ is the number of configurations that sums up to R_{max} or higher. The lesser the number of configurations with sum $\geq R_{max}$, higher is the probability of success. A sequence that can minimize $N(x \geq R_{max})$ will have a good ARE performance. Natural sequence has maximum number of configurations $< R_{max}$ and hence minimum $N(x \geq R_{max})$ among all sequences.

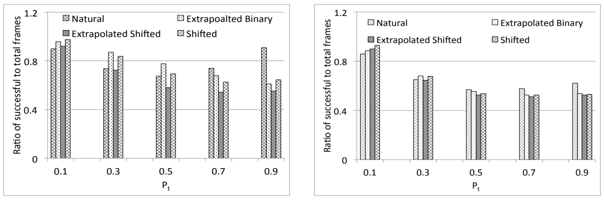
Insight 4.4 *At intermediate probability of transmitting a* 1*, number of colliding signals is* $\leq N_{log}$ *with high probability.*

At intermediate values of P_t, per-user probability of transmitting a *1*, up to N_{log} signals collide. A sequence that has maximum number of sums with unique configurations with up to N_s elements will have a high probability of success. Extrapolated binary sequence $S : \{2^0, 2^1, 2^2, ..N_{log}, R_{max} - N_u - N_{log}, R_{max} - N_u - N_{log} - 1, ..., R_{max}\}$ combines binary sequence and shifted sequence. $N_{log} = log_2 R_{max}$ is the number of sources a binary sequence can support. Binary sequence is an optimum solution to achieve zero address resolution error. Extrapolated binary sequence utilizes the maximum number of distinct sums when using a binary sequence. These distinct sums are obtained from the binary elements. The rest of the elements are integers decreasing from R_{max}, which reduces the number of overlapping sums. As the number of collisions approach N_{log}, the distinct sums contributed by the binary elements will improve the probability of success. Algorithm 1 summarizes the insights gained from the sequences observed and the probability of success derivation.

Algorithm 1 Address Allocation

```
$R_{max} \leftarrow$ Maximum Receivable amplitude
$N_s \leftarrow$ Number of simultaneous sources
$N_u \leftarrow$ Total number of sources
$S \leftarrow$ Set of addresses
if $\Pr(N_s \leq 2) > \Pr(N_s > 2)$ and $N_u \leq \frac{R_{max}}{2}$ then
    $S \leftarrow \{\frac{R_{max}-1}{2}, \frac{R_{max}-1}{2} + 1, ..., R_{max} - 1\}$
else if $\Pr(N_s \leq 2) > \Pr(N_s > 2)$ and $N_u > \frac{R_{max}}{2}$ then
    $S \leftarrow \{R_{max} - N_u, R_{max} - N_u - 1, ..., R_{max} - 1\}$
else if $\Pr(N_s \leq N_{log}) > \Pr(N_s > N_{log})$ then
    $S \leftarrow \{2^0, 2^1, .., 2^{N_{log}}, R_{max}, R_{max} - 1, ..\}$
else if $\Pr(N_s \geq N_{max}) > \Pr(N_s < N_{max})$ then
    $S \leftarrow \{1, 2, 3, ..., N_u\}$
else
    $S \leftarrow \{2^0, 2^1, .., 2^{N_{log}}, R_{max}, R_{max} - 1, ..\}$
```

5. PERFORMANCE EVALUATION



(a) N_s =5, $R_{max} = 15$ (b) N_s =7, $R_{max} = 15$

Figure 3: Performance of Proposed Sequences Vs. P_t

We built NS_{BC} (Network Simulator - Bacterial Communication), a python based simulator, to evaluate the performance of bacterial commuication network. Channel is assumed to be noise free i.e., zero channel noise and hence zero modulation and zero sampling error. Every source has uninterrupted supply of data to transmit. The probability of transmitting a *1* is the same across sources and is tunable. Every source transmits a *1* with probability P_t and *0* with probability $1 - P_t$.

NS_{BC} uses a maximum likelihood receiver. On receiving the samples, the receiver decodes the components of each sum. If there is only one possible configuration for that sum, receiver directly assigns the samples to respective addresses. When a received sum has more than one configuration, receiver chooses one of the many configurations at random i.e., on receiving 3, (3) and $(1, 2)$ are equally likely to be chosen. Once the receiver assigns each sample received to a transmitter, it uses *maximum likelihood* to identify the bit transmitted. The number of samples per bit is pre-determined. For e.g., in the experimental setup we consider, the duration of the transmitted signal is 50 minutes and a measurement is taken every 10 minutes at receiver. Thus, every bit has 5 sample points.*Maximum likelihood* principle is used to map samples to bits per transmitter.

We evaluate the performance of the algorithm proposed in section 4 using NS_{BC}. From equation 3.1, the parameters that influence address resolution error are 1. per-user probability of transmitting *1*, P_t 2. Maximum receivable amplitude R_{max} 3. Number of sources N_s. We vary these 3 parameters to analyse the address resolution error performance of the 4 sequences presented in algorithm 1. Unless otherwise mentioned, R_{max} is set to 15 and P_t to 0.5.

Figures 3(a) and 3(b) plots the ratio of successful frames to total transmitted frames as a function of increasing P_t for different address sequences. R_max is set to 15 and hence $N_{log} = 4$. Number of sources greater than 4 cannot have 100% success rate.

It can be observed that at lower probabilities, *shifted sequence* has the highest success rate and at higher probabilities *natural sequence* has the best success rate. In the intermediate probabilities *extrapolated binary sequence* has the highest probabilities as learnt from the insights in Section 4. The above behavior can be explained using the distribution of configurations. The number of configurations per sum for each sequence is plotted in Figure 4 for R_{max} and N_s. Natural sequence has least number of configurations at R_{max}. Therefore, the probability of success due to configurations greater than R_{max} is highest for natural sequence. The figure within is a zoomout of sum 1 through 14. Shifted has minimum configurations and hence has best performance in a 2 collision domain. Figures 5(a) and 5(b) plots the ratio

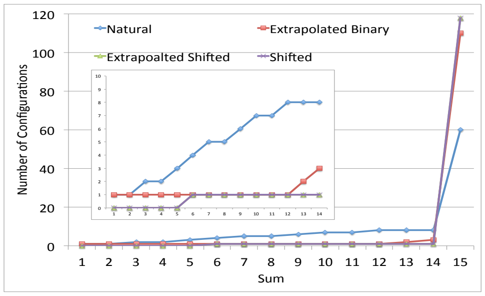
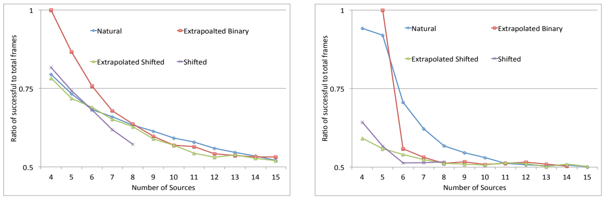


Figure 4: Distribution of configuration sums



(a) $P_t = 0.3$ (b) $P_t = 0.9$

Figure 5: Performance of Proposed Sequences Vs. N_s

of successful frames to total transmitted frames as a function of increasing number of sources for different address sequences. With increasing N_s, number of successful frames decreases for all sequences. The rate of fall in successful frames is also a function of P_t. At low P_t, extrapoalted binary sequence performance exponentially while at high P_t it is approximately linear. Figures 6(a) and 6(b) plots the ratio of successful frames to total transmitted frames as a function of increasing R_{max} for different address sequences. The number of successful frames depends on the maximum number of sums that can be distinguished by the receiver. At low R_{max}, for the same number of sources, ARE is higher than that at high R_{max}

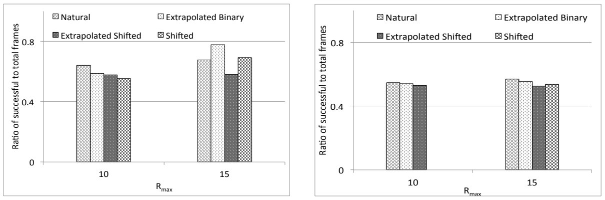
To verify the correctness of our algorithm, we used the probability of success derived in equation 4 and performed an exhaustive search over all possible sequences to find the minimum address resolution error for a given R_{max} and P_t. For a given R_{max}, all possible address sequences by choosing amplitudes from the range $[1, R_{max}]$ was generated. There are $2^{R_{max}} - 1$ sequences for each P_t. The percentage of successful frames transmitted for each sequence is calculated. Figures 7(a) and 7(b) shows the performance of the algorithm proposed when $R_{max} = 15$ and $P_t = 0.5$ with 5 and 7 sources respectively. Given R_{max} and P_t as input, the algorithm chooses the optimum sequence and the ratio of successful frames to the total transmitted frames is calculated from simualtions. The theoretical result is the probability of success derived from equation 3.1 evaluated over all possible sequences. It can be observed that the algorithm performs close to 99% of the theoretical result at all P_t and N_s presented.

6. ISSUES AND CONCLUSIONS

The following assumptions were made in deriving the best address sequence for a network. We discuss in detail each of the assumptions below.

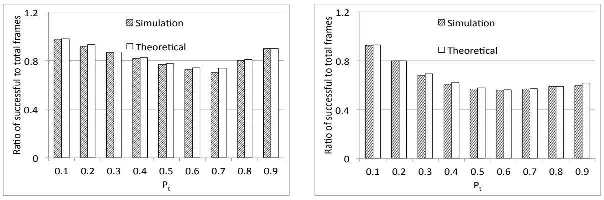
P_t *is known*

As seen in Figures 3(a) and 3(b), the performance of an addressing sequence depends on the load. Inorder to choose the best sequence, knowledge of an approximate range of



(a) $N_s = 5$ (b) $N_s = 7$

Figure 6: Performance of Proposed Sequences Vs. R_{max}



(a) 5 sources (b) 7 sources

Figure 7: Performance of Source Addressing Algorithm

P_t is important. We had assumed that P_t for every user is known before transmitting. This assumption can be removed by providing the sources with multiple sequences and the source choose the sequence based on its P_t. Since one of the above 4 sequences have a good performance in most cases, we can let the source choose the sequence.

P_t is same across sources

The insights developed rely on P_t being same for all sources. Finding the probability of success and the best sequence for different P_t across sources is a challenging problem that we are currently working on.

Non empty data queue

We assumed that every source has uninterrupted supply of data to transmit. Absence of signal is assumed as *0*. In practice, we need to differentiate between a zero and *no-data*. We propose the use of start-of-frame and end-of-frame sequences to solve this problem. A pre-assigned bit sequence can be used to define the start and end of a frame. Anything outside this start and end of frames is considered as no-data.

We showed that amplitude of the transmitted signal can be used to address sources in a bacterial communication network. We proposed an optimum sequence of amplitudes for 4 scenarios and proved that the performance of these sequences are very close to the theoretical maximum.

7. REFERENCES

[1] Akyildiz, I. F., Fekri, F., Forest, C. R., Hammer, B. K., and Sivakumar, R. Monaco: Fundamentals of molecular nano-communication networks (invited paper). *IEEE Wireless Communications Magazine, Special Issue on Wireless Communications at the Nano-Scale* (2012).

[2] Akyildiz, I. F., and Jornet, J. M. Electromagnetic wireless nanosensor networks. *Nano Communication Networks* (2010).

[3] Austin, C. M., Stoy, W., Su, P., Harber, M. C., Bardill, J. P., Hammer, B. K., and Forest, C. R. Modeling and validation of autoinducer-mediated bacterial gene expression in microfluidic environments. *Biomicrofluidics 8*, 3 (2014), 034116.

[4] Bassler, B. L. How bacteria talk to each other: regulation of gene expression by quorum sensing. *Current Opinion in Microbiology* (1999).

[5] Bohman, T. A sum packing problem of erdös and the conway-guy sequence. *Proceedings of the American Mathematical Society 124*, 12 (1996), 3627–3636.

[6] Charrier, T., Chapeau, C., Bendria, L., et al. A multi-channel bioluminescent bacterial biosensor for the on-line detection of metals and toxicity. Part II: technical development and proof of concept of the biosensor. *Analytical and Bioanalytical Chemistry* (2011).

[7] Danino, T., Mondragon-Palomino, Octavio, Tsimring, et al. A synchronized quorum of genetic clocks. *Nature*.

[8] Deligeorgis, G., Dragoman, M., Neculoiu, D., et al. Microwave propagation in graphene. *Applied Physics Letters* (2009).

[9] Eckford, A. W. Molecular communication: Physically realistic models and achievable information rates.

[10] Einolghozati, A., Sardari, M., Beirami, A., and Fekri, F. Capacity of discrete molecular diffusion channels. In *Information Theory Proceedings (ISIT), 2011 IEEE International Symposium on* (2011), IEEE, pp. 723–727.

[11] Einolghozati, A., Sardari, M., Beirami, A., and Fekri, F. Data gathering in networks of bacteria colonies: Collective sensing and relaying using molecular communication. In *Computer Communications Workshops (INFOCOM WKSHPS), 2012 IEEE Conference on* (2012), IEEE, pp. 256–261.

[12] Endy, D. Foundations for engineering biology. *Nature*.

[13] Farsad, N., Eckford, A., Hiyama, S., and Moritani, Y. On-chip molecular communication: Analysis and design. *NanoBioscience, IEEE Transactions on* (2012).

[14] Guy, R. K. Sets of integers whose subsets have distinct sums. *North-Holland Mathematics Studies* (1982).

[15] Kadloor, S., Adve, R. S., and Eckford, A. W. Molecular communication using brownian motion with drift. *NanoBioscience, IEEE Transactions on*.

[16] Krishnaswamy, B., Austin, C. M., Bardill, J. P., Russakow, D., Holst, G. L., Hammer, B. K., Forest, C. R., and Sivakumar, R. Time-elapse communication: Bacterial communication on a microfluidic chip. *Communications, IEEE Transactions on*.

[17] Melke, P., Sahlin, P., Levchenko, A., and Jönsson, H. A cell-based model for quorum sensing in heterogeneous bacterial colonies. *PLoS Computational Biology* (2010).

[18] Pierobon, M., and Akyildiz, I. A statistical-physical model of interference in diffusion-based molecular nanonetworks.

[19] Stocker, J., Balluch, D., Gsell, et al. Development of a set of simple bacterial biosensors for quantitative and rapid measurements of arsenite and arsenate in potable water. *Environmental Science & Technology* (2003).

[20] Suda, T., Moore, M., Nakano, T., Egashira, R., Enomoto, A., Hiyama, S., and Moritani, Y. Exploratory research on molecular communication between nanomachines. In *GECCO 2005*.