Performance of SW-ARQ in Bacterial Quorum Communications

Chenyao Bai[a], Mark. S. Leeson[a], Matthew D. Higgins[a,*]

[a]*School of Engineering, University of Warwick, Coventry, CV4 7AL, UK*

Abstract

Bacteria communicate with one another by exchanging specific chemical signals called autoinducers. This process, also called quorum sensing, enables a cluster of bacteria to regulate their gene expression and behaviour collectively and synchronously, such as bioluminescence, virulence, sporulation and conjugation. Bacteria assess their population density by detecting the concentration of autoinducers. In *Vibrio fischeri*, which is a heterotrophic Gram-negative marine bacterium, quorum sensing relies on the synthesis, accumulation and subsequent sensing of a signalling molecule (3-oxo-C6-HSL, an N-acyl homoserine lactone or AHL). In this work, a data link layer protocol for a bacterial communication paradigm based on diffusion is introduced, considering two populations of bacteria as the transmitter node and the receiver node, instead of employing two individual bacteria. Moreover, some initial results are provided, which concern the application of the Stop-N-Wait Automatic Repeat reQuest (SW-ARQ) schemes to the proposed model. The performances of the system are later evaluated, in terms of the transmission time, frame error rate, energy consumption and transmission efficiency.

Keywords: Bacterial Communication, Quorum Sensing, SW-ARQ, Critical Distance, Diffusion, AHL, Molecular Communication, Nano Communications

1. Introduction

Molecular communication is a new and interdisciplinary field that combines aspects of the nanoscale world and the life sciences, with communication engineering [1]. One of the most promising paradigms of molecular communication is the use of molecular signalling inspired by the naturally occurring communication between bacteria via a process called Quorum Sensing (QS) [2]. Bacteria use QS to coordinate certain tasks based on the local density of the bacterial population. They communicate with each other using signalling molecules to perform complex tasks [3] such as light production and attacking suitable hosts. In particular, if the concentration of molecular signals in the medium exceeds a certain threshold, an individual bacterium in a population releases more molecules into the environment [4], which will in turn increase the density of signalling molecules over time. This produces what can be considered as a positive feedback process. Thus, the local density of bacteria can be measured by sensing the concentration of molecules present and bacteria perform tasks when the concentration exceeds a certain threshold. The output of the QS mechanism can be in various forms and one example is the production of Green Fluorescent Protein (GFP) [4]. A well-studied bacteria species, *Vibrio fischeri*, which is most famous for its bioluminescence and has been used to study toxicity of aquatic environments, has been employed to investigate the bacterial communication network in this paper.

For a molecular communication system, errors may be caused by inter symbol interference (ISI) which will result in data packet corruption and out-of-sequence delivery. This makes it necessary to apply error detection rules and ARQ mechanisms, which have been extensively proposed in conventional wireless systems, for reliable transmission [5]. The term ARQ was first introduced by Chang [6], after which three widely used ARQ schemes, including Stop-N-Wait (SW-ARQ), Go-Back-N (GBN-ARQ) and Selective-Repeat (SR-ARQ), have been presented and developed [7]. The contributions of this paper are as follows. Unlike most research in bacterial communications which concentrates on the physical mechanisms and channel models, this paper takes ARQ protocols in bacterial communication to enhance reliability, which is believed to be the first attempt in this field. Specifically, this research maps existing protocol concepts to biological quorum sensing processes and shows how different parameters can be fit to different modes of bacterial communication. Also, the transmission delay, frame error rate, power consumption and transmission efficiency are investigated to analysis the channel performance, which can be used to evaluate different protocol designs, taking into account the biological characteristics of quorum sensing. It is expected that this work will make a significant contribution to the design of biosensors, drug

☆Date of First Submission 3rd July 2014. Revisions Submitted 2nd October 2014.

*Corresponding author

Email addresses: Chenyao.Bai@warwick.ac.uk (Chenyao Bai), mark.leeson@warwick.ac.uk (Mark. S. Leeson), m.higgins@warwick.ac.uk (Matthew D. Higgins)

delivery systems and water toxicity detection mechanisms.

The rest of paper is organized as follows: In Section 2 the basic bacteria communication scheme is introduced. In Section 3, the transmitter and receiver models are developed, followed by the establishment of the channel model in Section 4. In Section 5, the SW-ARQ and CRC codes are used to enhance the system performance, followed by the results and discussions in Section 6. Finally, Section 7 gives the conclusions and possible future work.

2. Bacteria Communication Scheme

Because of the high degree of randomness and limited capabilities of a particular bacterium, communication between two individual bacteria can be unreliable. In addition, the delay in the communication process can be fairly large due to biological actions such as transcription and translation. Hence, to achieve reliability of the communication system, here, the communications model between two populations of bacteria which is proposed in [4] is taken into consideration. In this model, a cluster of bacteria trapped in a chamber is considered as a node. The model consists of the transmitter node, the receiver node and the communication channel. Both the transmitter and receiver nodes are considered to be genetically modified bacteria [8] [9]. Molecular communication between two bio-nodes can be made up of three procedures. The transmitter node produces the signalling molecules by adequate stimulation, then these molecules propagate through the medium undergoing Brownian motion and finally the receiver node senses the concentration of the local signalling molecules and takes appropriate actions. The transmitted information is encoded via the concentration of signalling molecules, i.e. the embedding of the information is by alteration of the concentration of the molecules and its transmission relies on diffusion. The output of the receiver node, in the form of luminescence, is measured in steady-state to estimate the concentration of signalling molecules at the vicinity of the node, and hence decode the transmitted information [4].

In this proposed model, both the transmitter and receiver nodes contain N bacteria, specifically *Vibrio fischeri*, which is a species of bioluminescent bacterium. These bacteria are motile, gram-negative rods, $0.8-1.3\mu m$ in diameter and $1.8-2.4\mu m$ in length [10]. As a marine bacterium, *V. fischeri* exists at low cell densities when free living and at high cell densities when colonising the light organ. The luminescence is governed by the expression of certain genes, called the *lux* operon, in the cell, which in turn is controlled by the density of cells in a population. The regulation of the luminescence genes, named *luxCDABE*, depends on the production and detection of the signal (3-oxo-C6-HSL, an N-acyl homoserine lactone or AHL), which is synthesised by the protein LuxI and sensed by the protein LuxR. Fig. 1 [4] illustrates the process with the structure of a bacterium used in such a node shown in Fig. 1 (a). The receiver node senses the surrounding

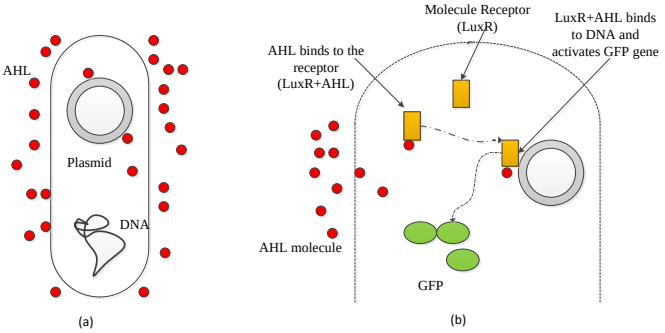


Figure 1: Bacterial communication [4]: (a) bacterium structure; (b) GFP production.

concentration of AHLs, which will trigger the production of GFP and the process is shown in Fig. 1 (b).

In *Vibrio fischeri*, the signal AHL is synthesised by the protein LuxI and sensed by the protein LuxR which is displayed in Fig. 1 (b). At low cell densities when only a small number of bacteria are present, the signal is produced by the bacteria at a low level. Then the molecules diffuse out of the bacteria cells and propagate into the surrounding environment. When the bacteria population increases, the concentration of AHLs around the node will grow. If the concentration of signal reaches a critical threshold, it is able to interact with the LuxR protein, which is acted as the ligand receptor for AHL. The LuxR/AHL complex binds to a region of DNA called the *lux* box, causing the luminescence genes to switch on. In addition, the LuxR/AHL complex also triggers the AHL (via LuxI) to be produced at a higher level. Thus the AHL is said to auto-induce its own synthesis.

Fig. 2 shows the schematic for the communication between two populations of bacteria. The specifics of the nodes will be discussed in the next section. In this work, the number of bacteria in each node is assumed to be constant. The bacteria inside the node can grow, divide and die to maintain the constant population through the process of gene regulation [11]. It is assumed that each bacterium can sense and produce two different types of AHL molecules, denoted as type I and type II [12]. The transmitted information which will be transmitted at the transmitter is encoded into the concentration of signals, denoted by A_0. The bacteria inside the transmitter node can produce various concentrations of type II molecules to be transmitted through the channel by the stimulation of different levels of concentration of type I molecules surrounding them [4]. At the receiver, each bacterium senses the concentration of type II molecules through type II receptors, followed by the production of GFP by bacteria, which is used to decode the input signal concentration A_0.

3. Transmitter and Receiver Model

As discussed in section 2, each bacterium must be equipped in general with two distinct receptor types (type

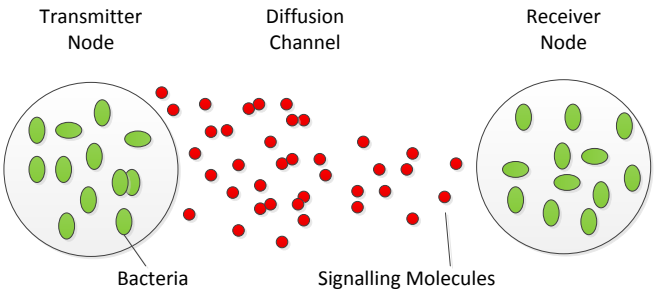


Figure 2: Bacterial communication setup consisting of the transmitter node, the diffusion channel and the receiver node.

I and type II molecules). For the bacteria in the transmitter node, only type I receptors are enabled, while for bacteria in the receiver node, only type II receptors are enabled. It is assumed that for each type of molecules, there are M ligand receptors. The bacteria will produce type II molecules when stimulated by type I molecules, but will produce GFP when stimulated by type II molecules.

Both the transmitter and receiver node have a probabilistic nature. For the transmitter and receiver nodes, the process of reception is governed by a set of equations, but with possibly different coefficients [4]. The response of different kinds of bacteria to different levels of concentration of signalling molecules has been well studied in the literature. In this work, the model in [13] is adopted, which uses a series of linear differential equations, which describes the average dynamic behaviour of bacteria and also their steady state, to explain the process of luminescence in response to AHL concentration difference.

According to [13], the probability p of AHL+LuxR binding is given by:

$$\dot{p} = -\kappa p + A\gamma(1-p) \tag{1}$$

where A is the concentration of signalling molecules surrounding the bacterium, γ is the input gain and κ is the dissociation rate of captured molecules in the cell receptors. Equation 1 shows that each cell receptor is activated by capturing one signalling molecule with a probability p, which depends on the molecular concentration around it. Also, it can be inferred that the steady state value is denoted by:

$$p^* = \frac{A\gamma}{A\gamma + \kappa} \tag{2}$$

Equation 2 shows that with higher concentration of molecules, the capture probability is higher. Also, it approaches 1 for very high concentrations.

4. Channel Model

In the channel model, the signalling molecules propagate through the channel via diffusion process in a three dimensional medium. To effectively represent the transmitted symbols, the propagation time is divided into time slots, also called symbol durations, which have the equal length. Only one symbol propagates in single time slot which is denoted by t_s. The motion of information molecules is inspired by the forces produced by the constant random thermal motion of the molecules within the fluid medium. The information molecule is at a distance r away from the receiver which has a radius of R. The radius of the receiver node R is related to the number of bacteria N in the receiver node. Moreover, the diffusion coefficient is $4.9 \times 10^{-6}\text{cm}^2\text{s}^{-1}$, which is settled as a conservative value for AHL in water at 25°C [14].

In essence, the information molecules propagate through the fluid medium undergoing Brownian motion which is a random procedure and a probabilistic behaviour, which means that the molecules are not ensured to arrive at the receiver. In other words, there is a probability that the molecule will hit the receiver at a time slot. According to [15], the capture probability $P(r,t)$ can be calculated by:

$$P(r,t) = \frac{R}{r}\text{erfc}\left\{\frac{r-R}{2\sqrt{Dt}}\right\} \tag{3}$$

where $\text{erfc}\{\cdot\}$ is the complementary error function [14]. Equation 3 shows the probability that a molecule arrives at the receiver at a time slot from mathematical approach. There is no need to use the more laborious and time consuming Monte Carlo simulation approach to analyse the point to point link.

To achieve the hit time probability, which refers to the probability that an information molecule arrives at the receiver at a certain time t, equation 3 is differentiated with respect to time, obtaining the hit time distribution, which is calculated by:

$$h(t) = \frac{R}{r}\frac{d}{2\sqrt{\pi D}}\frac{1}{t^{3/2}}\exp\left(-\frac{d^2}{4Dt}\right) \tag{4}$$

where $d = r - R$ is the distance between the information molecule and the boundary of the receiver.

In this work, the communication channel is a Binomial one, where each molecule arrives at the receiver or does not. Assuming that in the current time slot, the number of received molecules is N_C when n molecules are sent at the start of the time slot. N_C can be described as follows according to the binomial distribution:

$$N_C \sim \text{Binomial}\left(n, P_{\text{hit}}(d,t_s)\right) \tag{5}$$

where $P_{\text{hit}}(d,t_s)$ represents the hit probability with transmission distance d and symbol duration t_s. The previous bits can have an influence on the current bit due to inter symbol inference (ISI). Here, only the previous one time slot will be taken into consideration since this has been shown to be a reasonable approximation [16].

Thus the number of molecules received in a time slot which is denoted by N_{hit} is made up of the molecules sent at the start of the current time slot and the start of the previous symbol duration. Molecules which are received

	$s_p s_c$(input)			
Y(output)	00	01	10	11
0	$P_{R(0,0)}$	$1-P_{R(0,1)}$	$P_{R(1,0)}$	$1-P_{R(1,1)}$
1	$1-P_{R(0,0)}$	$P_{R(0,1)}$	$1-P_{R(1,0)}$	$P_{R(1,1)}$

Table 1: Binary channel model (possibility representation).

from the previous time slot during the current one are called residual molecules denoted by N_p, which is also a random variable. It is always obtained by calculating the difference between two binomial distributions, the number of molecules received during two time slots and the number of molecules received by the current time slot.

The binomial distribution can be approximately replaced by a normal or Gaussian distribution which has the same mean and standard deviation. Thus, by replacing the binomial distribution with normal distribution, the number of residual molecules N_p can be described as:

$$\begin{aligned} N_p \sim & N(nP_{\text{hit}}(d,2t_s), nP_{\text{hit}}(d,2t_s)\left[1-P_{\text{hit}}(d,2t_s)\right]) \\ & - N(nP_{\text{hit}}(d,t_s), nP_{\text{hit}}(d,t_s)\left[1-P_{\text{hit}}(d,t_s)\right]) \end{aligned} \quad (6)$$

Given that the one-bit information of the current intended symbol is s_c and that of the previous time slot is s_p, the current received symbol depends on both of them. For this binomial channel model, both s_p and s_c can be taken as zero or one. The possibilities of outputs with respect to different inputs, which are represented by the combinations of s_p and s_c, for this binary channel model are shown in table 1.

As shown in table 1, there are four different cases for the binary channel model which are denoted by bit pairs $\{00, 01, 10, 11\}$ for received symbol decoding, according to the different values of s_c and s_p. Y is the received symbol in the current time slot. The probability $P_{R(p,c)}$ represents the probability of successfully receiving the current intended symbol in the current time slot, where p is the one-bit information represented by the previous intended symbol and c is that of the current one. The different four cases are displayed below.

Case {11}: Both the one-bit information represented by the previous intended symbol and the current one are "1". Thus the current received symbol is affected by both the previous and current time slot. Assuming that $P_1 = P_{\text{hit}}(d,t_s)$ and $P_2 = P_{\text{hit}}(d,2t_s)$, the number of molecules received at the current symbol duration can be described as:

$$N_{\text{hit}} \sim N(nP_2, n\left[P_2(1-P_2) + 2P_1(1-P_1)\right]) \quad (7)$$

The probability of success is the case when the number of received molecules exceeds the chosen threshold which is denoted by τ. It can be calculated by:

$$\begin{aligned} P_{R(1,1)} &= P(N_{\text{hit}} \geq \tau) \\ &\approx Q\left(\frac{\tau - nP_2}{\sqrt{n\left[P_2(1-P_2) + 2P_1(1-P_1)\right]}}\right) \end{aligned} \quad (8)$$

where $Q(\cdot)$ is the tail probability of the standard normal distribution with zero mean and unity variance, often referred to as Q-function, which can be calculated by the tail integration of normal distribution.

Similarly, the expressions of the other three cases are given by [16]:

$$\begin{aligned} P_{R(1,0)} &= P(N_{\text{hit}} < \tau) \\ &\approx 1 - Q\left(\frac{\tau - n(P_2 - P_1)}{\sqrt{n\left[P_2(1-P_2) + P_1(1-P_1)\right]}}\right) \end{aligned} \quad (9)$$

$$P_{R(0,1)} = P(N_{\text{hit}} \geq \tau) = I_{P_1}(\tau, n-\tau+1) \quad (10)$$

$$P_{R(0,0)} = 1 \quad (11)$$

where $I_{P_1}(\tau, n-\tau+1)$ is the regularized incomplete beta function.

The bit error rate (BER) is considered as a key parameter which is often employed to assess the performance of communication systems that transmit information from one position to another. Various kinds of noise, interference and phase jitter may cause degradation of the transmitted signal. Here, the BER refers to the probability of one bit error when information symbols are transmitted in the diffusion based communication channel. The total average BER can be obtained by calculating the average BER of all the four states stated above. According to [16], most molecules arrive at the receiver in a relatively short time while only a few molecules arrive after a very long period of time, which will lead to the unsatisfied increasing average hitting time. In this model, the symbol duration t_s is chosen as the time before 60% of the molecules arrive at the receiver. The optimized BER versus molecules per bit for different transmission distance and different number of bacteria in the receiver node is shown in Fig. 3.

Fig. 3 shows the bit error probabilities for this communication channel, with respect to different number of bacteria in the receiver node and different transmission distance, which can be considered as a noise for the channel. It demonstrates that the communication system has a better performance with a smaller BER if large numbers of molecules are sent during one time slot. In addition, the channel performs better with a smaller transmission distance and a larger population of bacteria in the receiver node.

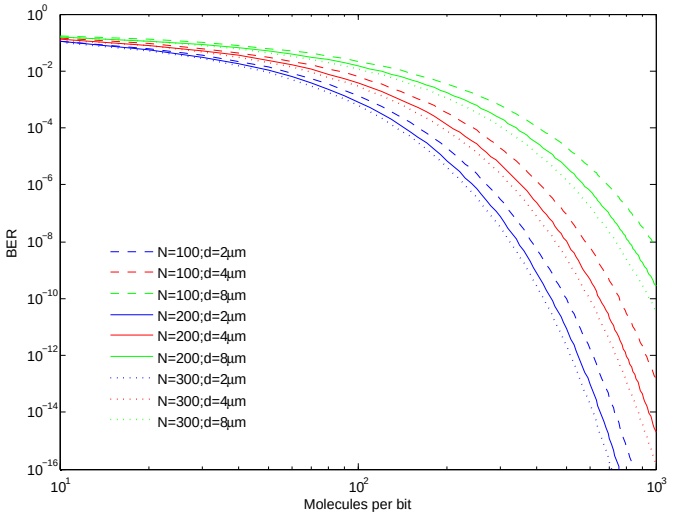


Figure 3: Bit error rate vs. molecules per bit for different number of bacteria in the receiver node ($N = 100, 200, 300$) and different transmission distance ($d = 2\mu m, 4\mu m, 8\mu m$).

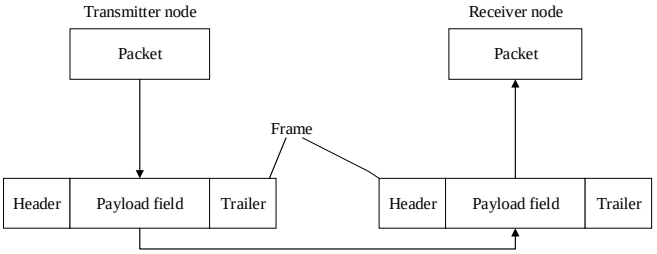


Figure 4: Relationship between packets and frames.

Name	Polynomial	Used in
CRC-8	$x^8 + x^2 + x + 1$	ATM header error check
CRC-10	$x^{10} + x^9 + x^5 + x^4 + x + 1$	ATM AAL CRC
CRC-12	$x^{12} + x^{11} + x^3 + x^2 + x + 1$	Bisync
CRC-16	$x^{16} + x^{15} + x^2 + x + 1$	Bisync
CCITT-32	$x^{32} + x^{26} + x^{23} + x^{22} + x^{16} + x^{12} + x^{11} + x^{10} + x^8 + x^7 + x^5 + x^4 + x^2 + x + 1$	IEEE802, DoD, V.42, AAL5

Table 2: Standard generator polynomials [19].

5. Stop-N-Wait ARQ

A time varying channel with a relatively high BER level causes frequent packet corruptions and out-of-sequence delivery, which need error check codes and Automatic Repeat reQuest (ARQ) mechanisms for effective error detection and recovery, respectively [17]. ARQ is a technique which has been used to ensure that a data stream is delivered accurately to the user despite errors in transmission. In terms of the open system interconnection (OSI) reference model for layered network architectures [18], an ARQ protocol is usually located at the data link layer, which takes the packets it gets from the network layer and encapsulates them into frames for transmission [19]. Each frame contains a frame header, a payload field for holding the packet, and a frame trailer which is shown in Fig. 4. ARQ forms the basis for peer-to-peer protocols that provide for the reliable transfer of information.

The advance of synthetic biology, particularly the foundation of the BioBricks database [20], enables many types of capabilities based on genetically engineered bacteria, including timing, counting, clocking, logic processing, pattern detection and intercellular communication [21], making it possible for substantial complex computational operations such as ARQ mechanisms and error detection techniques. In addition, genetically modified bio-nodes can harvest energy from biological systems and require no external energy sources, which is therefore expected to be energy efficient [22]. In this model, the simplest ARQ scheme, the Stop-N-Wait (SW-ARQ) scheme, is used to improve the channel performance. In SW-ARQ, the transmitter node sends a frame to the receiver and waits for its acknowledgment.

Due to the noise in the communication channel, error detection techniques will be used in the model here to improve the error-rate performance. There are three kinds of error detection codes, parity check codes, the Internet checksum and polynomial codes which are also known as cyclic redundancy check (CRC) codes. The choice of method is an open question for nanoscale communication systems but should be energy efficient given their likely energy storage capabilities. Here, CRC is adopted for error checking, which is widely used in data communication systems. CRC has a good error sensing performance, fast encoding and decoding capabilities and applicability to varying message lengths [23]. In CRC the information symbols, the codewords and the error vectors are represented by polynomials with binary coefficients [24]. CRC is specified by its generator polynomial $g(x)$. Table 2 gives some generator polynomials that have been endorsed in a number of standards.

In this channel model, the transmitter node generates a sequence of information frames for transmission. The ARQ mechanism requires the frame to contain a header with control information that is essential for proper operations. In addition, CRC check bits will also be appended in the frame to determine if error occurs during transmission. Fig. 5 shows the basic elements of ARQ protocols. It contains the information frames that transfer the information packets, the acknowledgment frames (ACKs) and the time-out mechanisms. The ACK frame signifies the receipt of a given frame. The time-out mechanism is required to maintain the flow of frames. In this model, it is assumed that the information flows only in one direction, from the

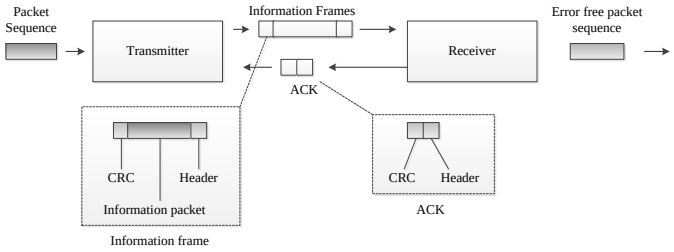


Figure 5: Basic elements of ARQ.

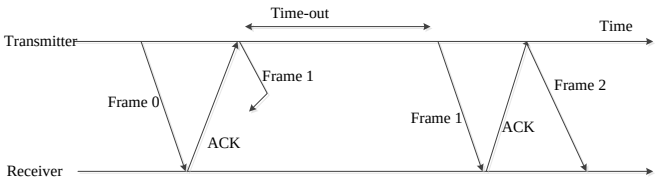


Figure 6: Transmission process.

transmitter to the receiver. The reverse communication channel is used only for the transmission of ACKs. This process is particularly needed since time varying channel with noise causes frequent packet corruptions.

Fig. 6 shows the process of transmitting a series of frames with ACKs and the time-out mechanism. At the initial point, frame 0 is transmitted and the transmitter will wait for a corresponding ACK frame. If the ACK frame is received without error, the transmitter will send frame 1 and reset the timer. Otherwise, when the time-out period expires, it resends frame 0. Each time the transmitter sends an information frame, it starts a timer. The retransmission continues until either the packet is received successfully or the number of retransmissions reaches a certain threshold. In addition, in order to avoid ambiguities, sequence number can be added to the frames. The protocol continues in this manner until all the frames are transmitted successfully.

6. Results

In this section the performance results for the SW-ARQ protocol are discussed in terms of transmission delay, average energy consumption per frame, frame error rate and transmission efficiency. The results show that the major aspects that affect the system performance are transmission distance, number of bacteria in the receiver node, frame length, CRC polynomial and time-out mechanism, see (12) below in this model. Here, the time-out is considered as the maximum number of transmissions per frame. In addition, the energy consumption for 1 bit is normalised to unity. All information frames are also supposed to be of the same length. The basic delay t_0, in the absence of errors, which transpires from the moment a frame is transmitted into the channel to the moment when the ACK is confirmed is calculated by:

$$\begin{aligned} t_0 &= 2t_{\text{prop}} + 2t_{\text{proc}} + t_f + t_{\text{ack}} \\ &= 2t_{\text{prop}} + 2t_{\text{proc}} + \frac{n_f}{R_b} + \frac{n_a}{R_b} \end{aligned} \tag{12}$$

In equation 12, the basic delay t_0 is made up of four components. The first bit of a frame that is input into the channel appears at the output of the channel after a propagation time t_{prop} (for both an information frame and a corresponding ACK frame) and then the end of the frame is received at the receiver after t_f (for and information frame) or t_{ack} (for an ACK frame) additional seconds. At both the transmitter and receiver, it requires t_{proc} seconds for processing, including CRC checking and preparing the next frame to be transmitted. In addition, n_f is the number of bits in the information frame and n_a is the number of bits for the acknowledgement frame. R_b is the bit rate of the transmission channel.

The effective information transmission rate of the protocol in the absence of errors is given by:

$$\begin{aligned} R^0_{\text{eff}} &= \frac{\text{Number of information bits delivered to receiver}}{\text{Total time required to deliver the information bits}} \\ &= \frac{n_f - n_0}{t_0} \end{aligned} \tag{13}$$

where n_0 is the number of overhead bits in a frame.

The effective information transmission rate of the protocol when error occurs is given by:

$$\begin{aligned} R_{\text{eff}} &= \frac{\text{Number of information bits delivered to receiver}}{\text{Average total time to transmit a frame}} \\ &= \frac{n_f - n_0}{t_{\text{ave}}} \end{aligned} \tag{14}$$

where t_{ave} is the average time to transmit a frame. The transmission efficiency is obtained by $R_{\text{eff}}/(R^0_{\text{eff}})$.

The number of bacteria in each node is first taken into consideration, ranging from $100 \sim 500$, which is shown in Fig. 7. Here the frame length is set to be 50 bits, the transmission distance is $4\mu m$, the data rate is 100bps, time-out is 12 and CRC-8 is employed. Results show that less time and energy will be consumed when information bits are transmitted through the channel if there is a larger population of bacteria in the receiver node. Moreover, less packet corruptions happen during the transmission process and the transmission efficiency is higher in this situation. In addition, compared with the performance where no CRC and ARQ are used at all, the channel performs better with a lower frame error rate, which improves the channel reliability significantly. In the following research, it is assumed that each node contains 300 bacteria.

Fig. 8 displays the results for different CRC codes. Here the frame length is set to be 50 bits, the transmission distance is $4\mu m$, time-out is 12 and the data rate is 100bps. It is clear that much more time and energy are needed when

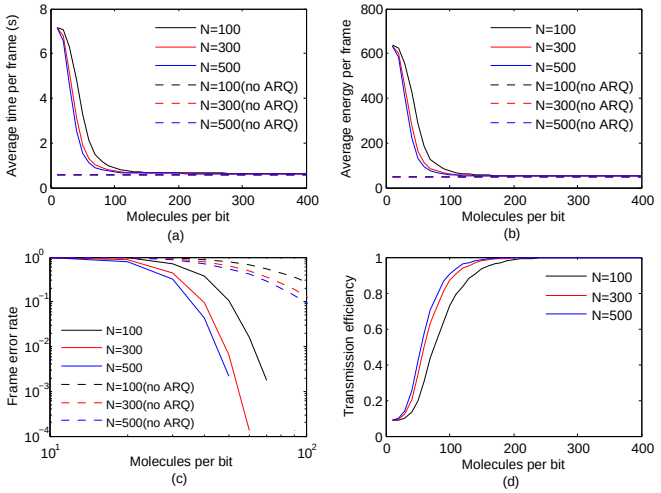


Figure 7: Channel performance, including (a) Average time consumption per frame; (b) Average energy consumption per frame; (c) Frame error rate; (d) Transmission efficiency, for different size of bacterial population in the receiver.

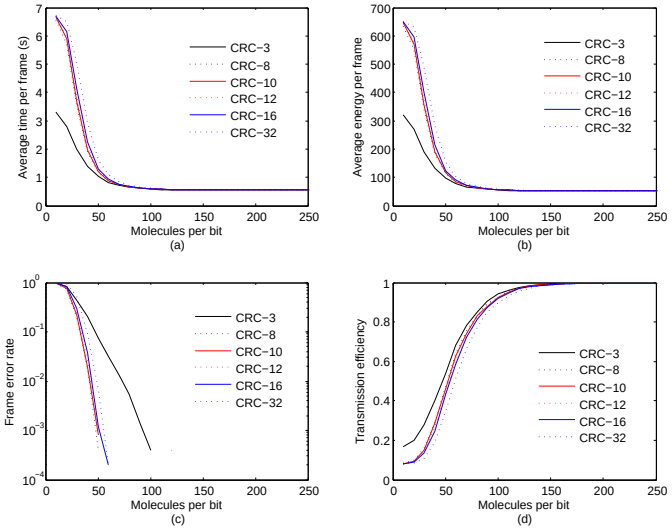


Figure 8: Channel performance, including (a) Average time consumption per frame; (b) Average energy consumption per frame; (c) Frame error rate; (d) Transmission efficiency, for different CRC polynomials.

more check bits are appended. Also, it should be noted that CRC-3 shows better transmission efficiency because not too many redundancy bits are added to the frames to occupy more time and energy. In addition, the frame error rate is almost the same for CRC-8, CRC-10, CRC-12, CRC-16 and CRC-32, while CRC-3 has a relatively higher error rate. Hence, CRC-8 will be adopted in the following investigations due to its relatively lower energy consumption and higher transmission efficiency.

Fig. 9 shows the system performances of different transmission distance. Here the frame length is 50 bits, the data rate is 100bps and CRC-8 is employed. The transmission delay per frame, the frame error rate and average energy

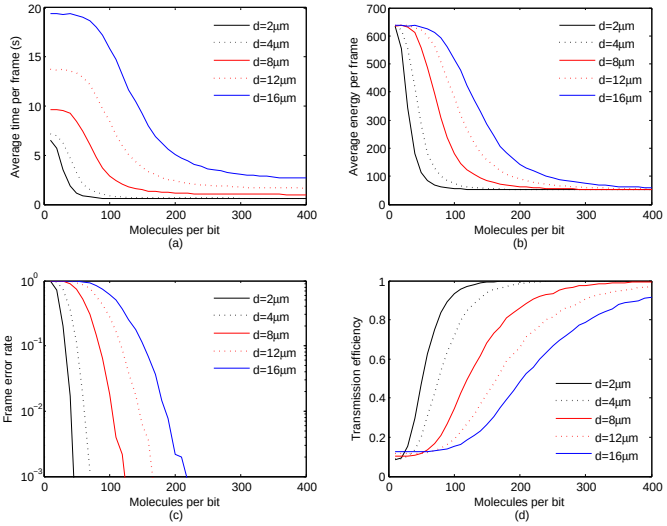


Figure 9: Channel performance, including (a) Average time consumption per frame; (b) Average energy consumption per frame; (c) Frame error rate; (d) Transmission efficiency, for different transmission distances.

consumption per frame are larger when there is a larger transmission distance. This is because over a larger distance, the bit error probability is higher according to Fig. 3, which will result to more transmission times per frame for average. Thus lower transmission efficiency should be observed with larger distances, which perfectly fits the results shown in Fig. 9 (d).

Fig. 10 shows the system performances when the frame length is different. Here the transmission distance is $4\mu m$, time-out is 12 and the data rate is again 100bps. Results show that the average time to transmit one frame, the frame error rate and average energy consumption per frame are larger when there is a larger frame length. This is because the probability of an error frame is calculated by

$$p_f = 1 - (1-p)^m \tag{15}$$

where p is the probability of one bit error and m is the frame length. It is clear that when frame length is larger, there is a larger probability that transmission error occurs in the frame. Thus, the corresponding transmission efficiency is lower.

In order to investigate the effect of the time-out mechanism to the system performance, additional research has been done which is shown in Fig. 11. The frame error rate and the transmission efficiency are essential aspects. It can be seen that the frame error rate is lower with a larger time-out, paying the price of much more energy consumption. Also, increasing the time-out has little effect on the transmission efficiency, which means that an increased time-out is not a good method to improve the system performance.

According to the investigations above, it is obvious that the four key parameters that have a big influence on the

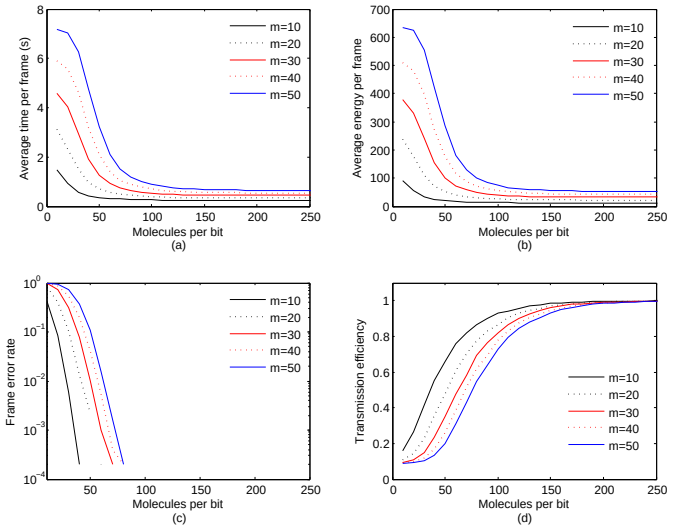


Figure 10: Channel performance, including (a) Average time consumption per frame; (b) Average energy consumption per frame; (c) Frame error rate; (d) Transmission efficiency, for different frame length.

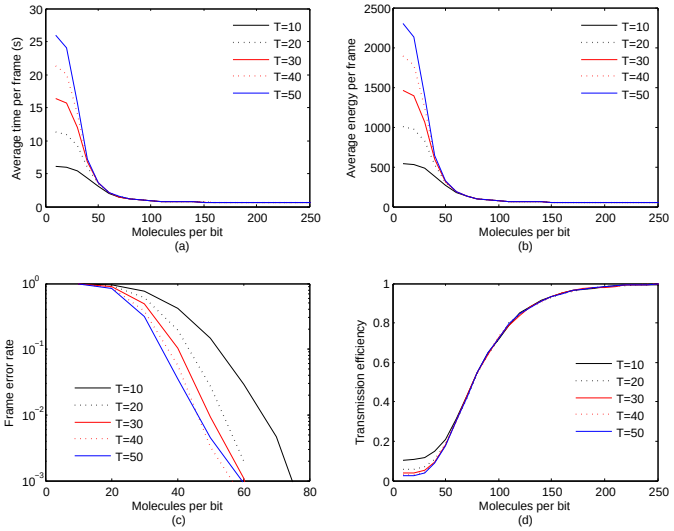


Figure 11: Channel performance, including (a) Average time consumption per frame; (b) Average energy consumption per frame; (c) Frame error rate; (d) Transmission efficiency, for different maximum transmission time per frame.

system performance are the number of bacteria in the receiver node, the CRC polynomial, the transmission distance and the frame length. Hence, as shown in Fig. 12, the performances of different combinations of these four factors are taken into consideration. Results show that for the 16 proposed conditions, the transmission delay per frame varies mainly because of the transmission distance. A smaller transmission distance achieves a smaller transmission time, regardless of the other three factors. While the average energy consumption per frame mainly depends on the frame length, which means that more energy will be consumed when the frame length is increasing. In addition, better transmission efficiency can be achieved with a smaller transmission distance, regardless of the other three factors. Moreover, according to the conditions which have the same transmission distance, the increasing of the number of bacteria in the receiver node leads to the increasing of transmission efficiency. Also, for the conditions which have the same propagation distance and population size of the receiver node, the system which has a smaller frame length has a better performance, with a higher transmission efficiency, regardless of the employed CRC polynomial. Hence, the significance to the transmission efficiency of the four factors will be in the decreasing order of transmission distance, number of bacteria in the receiver node, frame length and CRC polynomial.

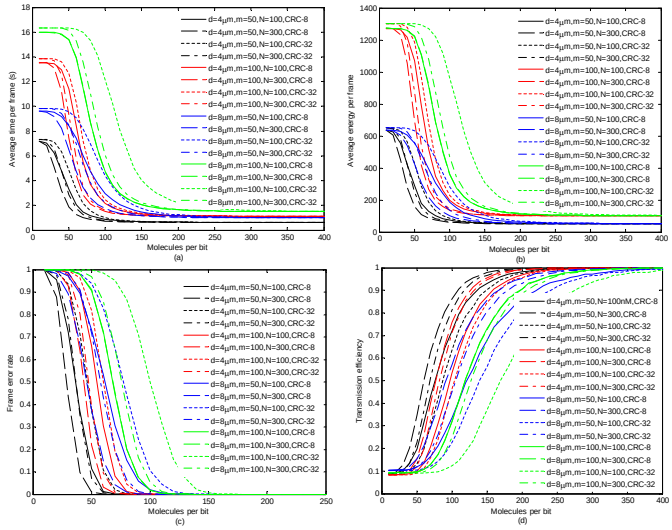


Figure 12: Channel performance, including (a) Average time consumption per frame; (b) Average energy consumption per frame; (c) Frame error rate; (d) Transmission efficiency, for different combinations of main factors.

7. Conclusions

In recent years, bacteria have been considered as one approach for molecular communication. They can communicate with each other through a process called quorum sensing. In this paper, a bacterial communication network model through a diffusion channel has been proposed, which considers two populations of bacteria as the transmitter node and the receiver node, respectively. A widely used protocol method has been employed that utilises CRC coding and ARQ dynamics. Changes to the number of bacteria in each node, the frame length, the transmission distance, the CRC polynomial and the time-out mechanism have been made according to varying network conditions to improve data link layer output. These showed that the transmission distance, the number of bacteria in the receiver node, the frame length and different kinds of error detection codes are four key parameters that will affect the system performance. The diffusion channel has better

transmission efficiency with smaller transmission distance and larger bacteria population in the receiver. In addition, smaller frame length and less complex error detection codes have better performance. The significance to the transmission efficiency of the four factors is in the decreasing importance order: transmission distance, number of bacteria in the receiver node, frame length and CRC polynomial. In addition, it is unnecessary to increase the time-out period, which will lead to relatively large energy consumption whilst having little improvement in transmission efficiency. It should also be noted that the ARQ dynamics and CRC coding are necessary, especially when large number of AHLs (more than 400) are sent at the start of each time slot, to achieve reliability. These properties may be used further to develop more powerful water contamination detection mechanisms with higher sensitivity.

References

[1] M. Moore, A. Enomoto, T. Nakano, R. Egashira, T. Suda, A. Kayasuga, et al. A design of a molecular communication system for nanomachines using molecular motors. In: Fourth Annual IEEE International Conference on Pervasive Computing and Communications Workshops, PerCom Workshops. IEEE; 2006, p. 554-559. doi:10.1109/PERCOMW.2006.4.

[2] B. L. Bassler. How bacteria talk to each other: regulation of gene expression by quorum sensing. Current Opinion in Microbiology 1999;2(6):582-587. doi:10.1016/S1369-5274(99)00025-9.

[3] H. B. Kaplan, E. P. Greenberg. Diffusion of autoinducer is involved in regulation of the vibrio fischeri luminescence system. Journal of Bacteriology 1985;163(3):1210–1214.

[4] A. Einolghozati, M. Sardari, F. Fekri. Design and analysis of wireless communication systems using diffusion-based molecular communication among bacteria. IEEE Transactions on Wireless Communications 2013;12(12):6096–6105. doi:10.1109/TWC.2013.101813.121884.

[5] J. S. Lin, K. T. Feng, Y. Z. Huang, L. C. Wang. Novel design and analysis of aggregated ARQ protocols for ieee 802.11n networks. IEEE Transactions on Mobile Computing 2013;12(3):556–570. doi:10.1109/TMC.2012.34.

[6] S. Chang. Theory of information feedback systems. IRE Transactions on Information Theory 1956;2(3):29–40. doi:10.1109/TIT.1956.1056807.

[7] J. Li, Y. Q. Zhao. Resequencing analysis of stop-and-wait ARQ for parallel multichannel communications. IEEE/ACM Transactions on Networking 2009;17(3):817–830. doi:10.1109/TNET.2009.2020820.

[8] T. Danino, O. Mondragon-Palomino, L. Tsimring, J. Hasty. A synchronized quorum of genetic clocks. Nature 2010;463(7279):326–330. doi:10.1038/nature08753.

[9] M. A. Palacios, E. Benito-Peña, M. Manesse, A. D. Mazzeo, C. N. LaFratta, G. M. Whitesides, et al. Infobiology by printed arrays of microorganism colonies for timed and on-demand release of messages. Proceedings of the National Academy of Sciences 2011;108(40):16510–16514. doi:10.1073/pnas.1109554108.

[10] F. Cousteau. Ocean: the world's last wilderness revealed. Dorling Kindersley Ltd; 2011.

[11] I. F. Akyildiz, F. Fekri, R. Sivakumar, C. R. Forest, B. K. Hammer. Monaco: fundamentals of molecular nano-communication networks. Wireless Communications, IEEE 2012;19(5):12–18. doi:10.1109/MWC.2012.6339467.

[12] J. G. Cao, E. A. Meighen. Purification and structural identification of an autoinducer for the luminescence system of Vibrio harveyi. Journal of Biological Chemistry 1989;264(36):21670–21676.

[13] J. Muller, C. Kuttler, B. A. Hense. Sensitivity of the quorum sensing system is achieved by low pass filtering. Biosystems 2008;92(1):76–81. doi:10.1016/j.biosystems.2007.12.004.

[14] P. S. Stewart. Diffusion in biofilms. Journal of Bacteriology 2003;185(5):1485–1491.

[15] R. M. Ziff, S. N. Majumdar, A. Comtet. Capture of particles undergoing discrete random walks. The Journal of Chemical Physics 2009;130(20):204104. doi:10.1063/1.3137062.

[16] M. Ş. Kuran, H. B. Yilmaz, T. Tugcu, B. Özerman. Energy model for communication via diffusion in nanonetworks. Nano Communication Networks 2010;1(2):86–95. doi:10.1016/j.nancom.2010.07.002.

[17] E. Ustunel, I. Hokelek, O. Ileri. A cross-layer goodput enhancement considering CRC coding and ARQ dynamics. In: 2012 IEEE Symposium on Computers and Communications (ISCC) 2012, p. 000023–000028. doi:10.1109/ISCC.2012.6249263.

[18] W. Stallings. Data and computer communications; vol. 6. Prentice hall New Jersey; 1997.

[19] A. S. Tanenbaum. Computer networks 5th Edition. Pearson; 2011.

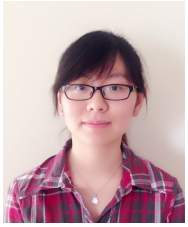
[20] R. P. Shetty, D. Endy, T. F. Knight Jr. Engineering BioBrick vectors from BioBrick parts. Journal of Biological Engineering 2008;2(5):1–12. doi:10.1186/1754-1611-2-5.

[21] Guopeng Wei, P. Bogdan, R. Marculescu. Efficient modeling and simulation of bacteria-based nanonetworks with BNSim. IEEE Journal on Selected Areas in Communications 2013;31(12):868–878. doi:10.1109/JSAC.2013.SUP2.12130019.

[22] T. Nakano, M. J. Moore, F. Wei, A. V. Vasilakos, J. Shuai. Molecular communication and networking: opportunities and challenges. IEEE Transactions on NanoBioscience 2012;11(2):135–148. doi:10.1109/TNB.2012.2191570.

[23] T. S. Baicheva. Determination of the best CRC codes with up to 10-bit redundancy. IEEE Transactions on Communications 2008;56(8):1214–1220. doi:10.1109/TCOMM.2008.070033.

[24] M. Hassner, R. Kotter, T. Tamura. Method and means for computationally efficient error and erasure correction in linear cyclic codes. 1999. US Patent No. 5,942,005.



Chenyao Bai received her B.S. degree in Optical Information Science and Technology from Shanxi University, China, in 2011 and her M.Sc. degree in Electronic Systems with Sensor Technology from University of Warwick, UK, in 2012. She is currently pursuing her Ph.D. study in School of Engineering at University of Warwick since 2012. Her research interests include molecular communication, quorum sensing, nanonetworks, mathematical modelling, error correction and detection codes and communication protocols.

Mark S. Leeson received the degrees of BSc and BEng with First Class Honors in Electrical and Electronic Engineering from the University of Nottingham, UK, in 1986. He then obtained a PhD in Engineering from the University of Cambridge, UK, in 1990. From 1990 to 1992 he worked as a Network Analyst for National Westminster Bank in London. After holding academic posts in London and

Manchester, in 2000 he joined the School of Engineering at Warwick, where he is now a Reader. His major research interests are coding and modulation, nanoscale communications and evolutionary optimization. To date, Dr. Leeson has over 220 publications and has supervised twelve successful research students. He is a Senior Member of the IEEE, a fellow of both the UK Institute of Physics and of the UK Higher Education Academy.

Matthew D. Higgins received his MEng in Electronic and Communications Engineering and PhD in Engineering from the School of Engineering at the University of Warwick in 2005 and 2009 respectively. Remaining at the University of Warwick, he then progressed through several Research Fellow positions with leading defence and telecommunications companies before undertaking two years as a Senior Teaching Fellow. As of July 2012 Dr. Higgins now holds the position of Assistant Professor. His major research interests are the modelling of optical propagation characteristics in underwater, indoor and atmospheric conditions. Dr. Higgins is a Member of the IEEE.