

An Information Theoretic Model of Molecular Communication Based on Cellular Signaling

Jian-Qin Liu

Kobe Advanced ICT Research Center,
National Institute of Information and
Communications Technology, Japan

liu@nict.go.jp

Tadashi Nakano

Department of Computer Science,
Donald Bren School of Information and Computer
Sciences, University of California, Irvine

tnakano@ics.uci.edu

ABSTRACT

In this paper, a class of molecular communication media is formalized, in which senders and receivers communicate by second messengers. Second messengers in the communication medium are released by senders of communication, propagated, amplified, removed in the medium, and sensed by receivers of communication that in turn decode the second messengers. Based on generalization of intracellular molecular communication processes in terms of information theory, an information theoretic model is proposed to understand and characterize the signaling mechanism of the molecular communication medium.

Keywords

Molecular communication, information theory, channel capacity, calcium signaling, kinase/phosphatase, phosphorylation/dephosphorylation

1. INTRODUCTION

Molecule-based communication, or *molecular communication*, is universally found at all levels of biological systems ranging from the molecular level to the cellular and tissue/organ levels, where molecular components of corresponding biological systems communicate by means of signal molecules [1]. In addition to naturally existing biological systems, molecular communication is also used in engineered biological systems, where engineered biological devices exchange signal molecules that carry specific information [7, 19, 20, 21].

Molecular communication is important for numerous cellular processes and functioning including cellular signaling and signal transduction, and therefore the biological mechanisms of molecular communication have been long studied in cell biology. However, the corresponding informatics of molecular communication has not been paid much attention in the literature. In this paper, we discuss informatics approaches to molecular communication from a view point of information theory and communication engineering. Informatics approaches employed in this paper would possibly reveal new aspects of molecular communication, including the origin of the robustness and

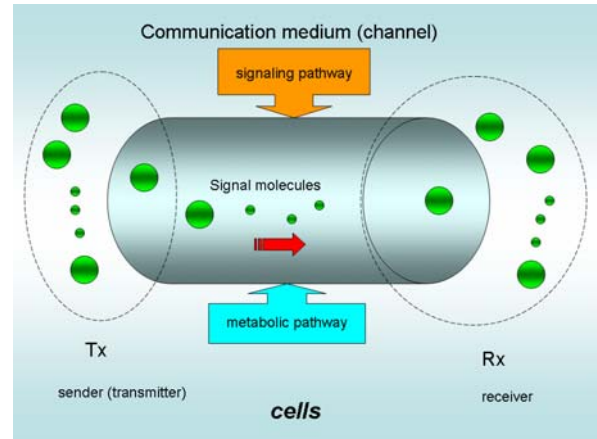


Figure 1: Molecular Communication

efficiency of molecular communication.

A general form of molecular communication is schematically described in Figure 1. Here, the transmitter (i.e., sender) and receiver are the two communicating entities, which could be molecules, cellular organelles, cells, or organs. The transmitter and receiver in molecular communication often represent a group of entities performing *n-to-n* communication (as opposed to *1-to-1* communication) (or MIMO -- multiple input multiple output communication). The signal molecules are transmitted by a sender(s) of communication, transported or propagated over the communication medium, and received by a recipient(s) of communication (i.e., receivers), along which various chemical reactions may occur (e.g., one of the effector molecules involved in MAPK cascades amplify signals and increase the concentration of target molecules.) The communication medium that propagates signal molecules from transmitters to receivers is regarded as the communication channel in a dynamical environment. The environment contains various noise sources such as white noise or colored noise, owing to the features of thermal dynamics in the environment.

In this paper, we model an instance of molecular communication based on Ca^{2+} signaling without a feedback loop (cf. section 3). Ca^{2+} signaling is the “ubiquitous” signaling mechanism for cellular organisms [3, 4]. The temporal and spatial dynamics patterns of Ca^{2+} signaling, termed as Ca^{2+} spikes, oscillations and waves, are known to regulate and control various biological processes such as fertilization, differentiation, proliferation, and death.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

Bionetics'07, December 10-13, 2007, Budapest, Hungary.

Copyright 2007 ICST 978-963-9799-11-0

In modeling the above mentioned molecular communication, we employ an information theoretic approach to address a set of fundamental questions regarding cellular signaling, including (1) molecular codes - how biological information may be encoded to and decoded from signal molecules, (2) performance of molecular communication - how the channel capacity that transmits biological information can be quantitatively measured, calculated or estimated, (3) efficient design schemes of molecular communication systems - how the channel capacity can be maximized for synthetic molecular communication systems (e.g., under the condition of a discrete memory-less channel), and (4) dynamic features of molecular communication - what the impact of environmental noise on channel capacity is and how noise can be optimally filtered out for signal detection in the noisy environment.

Information theoretic research of biological communication media is still premature. To the best of our knowledge, there are only a few published papers addressing channel capacity of biological communication media. Those papers found in literature [6, 19, 22, 23] have so far modeled and quantified channel capacity of (a) a free diffusion medium over which a sender communicates with a receiver using molecules that randomly move based on Brownian motion [6], (b) a free diffusion media in which a sender communicates with a receiver using frequency encoding (e.g., concentration change of signal molecules) [23], (c) transmembrane of neural systems where frequency encoded extracellular signals are transduced into intracellular Ca^{2+} spikes [22], and (d) an active transport medium that transport signals molecules by molecular motors [19].

The remaining sections of this paper are organized as follows. Section 2 examines cellular signaling from a molecular communication view point relative to the general representation of molecular communication. Section 3 describes an information theoretic model of molecular communication media based on cellular signaling and Section 4 concludes this paper.

2. MOLECULAR COMMUNICATION BASED ON CELLULAR SIGNALING

Ca^{2+} signals are the ubiquitous second messengers for cellular communication and are involved in a number of signal transduction pathways. In [3], the Ca^{2+} signaling process in cells is generalized as the Ca^{2+} signaling “toolkit” that is composed of the following four units: (1) stimuli that generate Ca^{2+} mobilizing signals activating the “on”-mechanisms, (2) the “on”-mechanisms that release Ca^{2+} signals for increasing the concentration of cytosolic Ca^{2+} , (3) Ca^{2+} sensitive processes (e.g., Ca^{2+} sensitive proteins) that activate various Ca^{2+} sensitive signaling pathways in response to the increased cytosolic Ca^{2+} concentration, and (4) the “off”-mechanisms that decrease the cytosolic Ca^{2+} concentration. From the viewpoint of communication engineering, we consider the stimuli as information sources, cellular components equipped with on-mechanisms as transmitters, Ca^{2+} sensitive components that trigger downstream signaling pathways as receivers, and the cytosolic medium containing the “off”-mechanisms as the communication medium.

Three specific examples (Figure 2) are provided below to illustrate how Ca^{2+} signaling is mapped to the general representation of molecular communication (Figure 1).

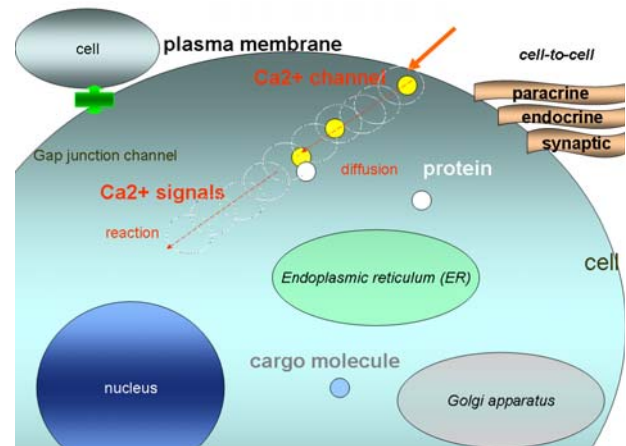


Figure 2: Cellular Signaling

Example 1: Ca^{2+} signaling as a communication channel

In response to various stimuli, Ca^{2+} signals are released from Ca^{2+} sources via Ca^{2+} channels to the cytoplasm. For example, when extracellular signal molecules (e.g., agonists) bind to Ca^{2+} entry channels on the cell surface, the channels open, allowing extracellular Ca^{2+} signals to enter the cytoplasm. The increased Ca^{2+} signals within the cell can then activate Ca^{2+} binding proteins such as CAM (calmodulin) that is used to regulate numerous cellular processes; TnC (troponin C) that controls muscle contraction; and PYK2 (proline-rich tyrosine kinase 2) that can initiate the MAPK cascading event [3]. In this example, Ca^{2+} entry channels are viewed as transmitters, Ca^{2+} binding proteins that decode Ca^{2+} signals as receivers, and free diffusion space in the cytosolic medium between transmitters and receivers as the communication medium.

Example 2: Amplification and relay of Ca^{2+} signals

Ca^{2+} signals locally induced as in Example 1 can be amplified and relayed along the ER, propagating globally in the cytoplasm. In one form of neural Ca^{2+} signaling [2], Ca^{2+} influx across the plasma membrane can initiate Ca^{2+} waves that propagate through CICR (Calcium Induced Calcium Release) along the ER (Endoplasmic Reticulum). The ER extends throughout the neuron toward the nuclear envelop, thereby propagating Ca^{2+} signals into the nucleus. Owing to the phosphorylation function, Ca^{2+} entry into the nucleus can in turn phosphorylate CREB to regulate gene transcription. In this example, membrane Ca^{2+} channels act as transmitters, the Ca^{2+} sensitive proteins in the nucleus as receivers, and the ER in the cytosol as the communication medium.

Example 3: Gap junction channels for cell-cell Ca^{2+} signaling

Cell-to-cell communication can be carried out by various mechanisms including paracrine, endocrine, synaptic and gap junctional communication. In the case of gap junctional communication, neighboring cells are connected through physical channels called gap junction channels, and Ca^{2+} signals propagate cell to cell known as intercellular Ca^{2+} waves [3]. Ca^{2+} waves may become regenerative if Ca^{2+} signals are amplified and relayed recursively while they are spatially diffusing. Intercellular Ca^{2+} waves allow connected cells to coordinate various cellular activities such as cell proliferation and death. In this example,

senders and receivers can be located in different cells, and the communication medium includes gap junction channels.

3. INFORMATION THEORETIC MODEL

The goal of this paper is to describe an information theoretic model of entire molecular communication processes illustrated in previous section. This goal is motivated by authors' previous work on intercellular communications [20, 21] and intracellular communications [13, 14] and is geared toward establishing a unified communication capacity theory of integrated molecular communication systems built based on intercellular and intracellular communications.

It is noted that Eckford [6] and Thomas et al. [23] have successfully formulated diffusion-based (or Brownian motion-based) molecular communication media, in which Gaussian distribution and Poisson distribution are applied for calculating and estimating the channel capacity. Specifically in their models, molecular communication channels are quantified as a random process in molecular concentrations or some observed values (e.g., time in the Eckford's model).

In our model, we generalize communication processes of a different type of molecular communication media using a unified measurement, that is, *point process* in probability theory. In the molecular communication media, Ca^{2+} signals diffuse spatially and Ca^{2+} sensitive processes contribute to the channel capacity of the communication media, owing to its role as the universal second messenger. In this model, the amount of messages (bits) transmitted in the molecular communication channel directly derived from Ca^{2+} signaling can be calculated and estimated to infer the upper bound of channel capacity, $\sup(C)$. A method of calculating $\sup(C)$ is under development and is not presented in this paper¹. The model to be presented in this paper is instead designed for understanding a network protocol of cellular communications (e.g., how Ca^{2+} signals are translated into a signaling event in a pathway network). In the present model, the communication performance (i.e., channel capacity) is therefore limited to the topological structure of the signal transduction pathway network (and not impacted by diffusion-reaction processes of Ca^{2+} signals); and accordingly, we discuss the lower bound of channel capacity, $\inf(C)$.

It is also noted that Thomas et al. has estimated channel capacity as 0.087 bits/second under the condition given in [23] and K. Prank et al. has used bit/spike as the unit for quantitative description of information transmission process in [22]. In a generalized description of molecular communication, the form of a nonlinear function is expected in order to unify different molecules in homogenous and heterogeneous media.

3.1 Ca^{2+} Signals Involved in Signal Transduction Networks

According to Kraus et al. [9], the temporal representation of Ca^{2+} is realized by Ca^{2+} oscillators, which is the basis of temporal codes of Ca^{2+} signals. As shown in the reaction-diffusion equation, the spatial distribution of Ca^{2+} is one of the important keys to understand the diffusion mechanism of the Ca^{2+} signaling processes in cells.

From [9], the reaction process is formulated in stochastic modeling and the diffusion process is formulated in a similar way:

$$\langle p_N(t), \Lambda(\alpha) \rangle$$

where the notations are defined in above-mentioned reference.

The set of all states of $\{N\}$ is denoted as U_N . Owing to the eq.1-3 in [9], it is inferred as follows:

$$p_N(t+1) \leftarrow p_N(t) \quad \infty \quad \text{the Markov process}$$

where the corresponding transition probability can be estimated e.g., by the Monte Carlo algorithm in [9].

With the lattice coordinates in two dimensional space, the measurement $\Lambda(\alpha)$ connects the locations in $X(t)$ and $X(t+1)$, which the spatial measurement can be synchronously calculated with the comparison to the ones $p_N(t)$ and $p_N(t+1)$.

The unified measure proposed in this paper is the coefficients of a two dimensional simultaneous auto-regression (2D SAR) model:

$$\beta_{\Xi} = \langle \beta(1,1), \beta(-1,1), \beta(1,0), \beta(-1,0), \beta(0,0), \beta(0,1), \beta(0,-1), \beta(1,-1), \beta(-1,-1) \rangle$$

which is defined in a 8-neighborhood.

This measurement describes the diffusion factor by $\beta(1,1), \beta(-1,1), \beta(1,0), \beta(-1,0), \beta(0,1), \beta(0,-1), \beta(1,-1), \beta(-1,-1)$ and the reaction factor by $\beta(0,0)$.

From the observed samples in the set of β_{Ξ} , the Ca^{2+} signaling process is represented by the extracted vectors

$$\beta(x(t-1), y(t-1)) \rightarrow \beta(x(t), y(t)) \rightarrow \beta(x(t+1), y(t+1)) \quad \text{s.t. } P_{\beta}(t)$$

where the non-causal constraint is assumed within the observed domain of the sample signals.

Let T_{β} be the threshold of the extracted vectors in 2D SAR, it is inferred that:

- the input to the LNC (linear network coding) model for describing the Ca^{2+} signaling process:

$$1, p(\beta(x(t-1), y(t-1))) \geq T_{\beta}$$

$$0, p(\beta(x(t-1), y(t-1))) \leq T_{\beta}$$

- the output of this signaling process:

$$1, p(\beta(x(t+1), y(t+1))) \geq T_{\beta}$$

$$0, p(\beta(x(t+1), y(t+1))) \leq T_{\beta}$$

The average value of the capacity of the individual Ca^{2+} signaling channel is formulated as

$$\min \arg k \quad \text{s.t. } (p(\beta(x(t), y(t)))) \geq T_{\beta}$$

where k refers to the bit number reflecting the number of the Ca^{2+} signals in the corresponding activated pathway networks.

Ca^{2+} signals are universal intracellular second messengers involved in complex cellular signal transduction networks. Numerous Ca^{2+} sensitive processes act as switch-like units that perform cross-talks in signal transduction networks. To model Ca^{2+} signaling in signal transduction networks, we first introduce the following two axioms.

Axiom 1: Provided that there is a mutual relationship between Ca^{2+} signals and phosphorylation/dephosphorylation of particular Ca^{2+} sensitive proteins, we have the predicate forms of logical representation for describing the building-blocks of cellular signal transduction networks as follows:

¹ To be presented in [18].

Calcium-to-phosphorylation (X),
 Calcium-to-dephosphorylation (X),
 Phosphorylation-to-calcium (Y),
 Dephosphorylation-to-calcium (Y),

0 0 0 1 0 0 0
 0 0 0 0 1 0 0
 0 0 0 0 0 1 1
 0 0 0 0 0 0 0
 0 0 0 0 0 0 0

where X refers to a set of signaling molecules located in upstream pathways (e.g., Ca^{2+} signaling pathways activated by membrane receptors); Y refers to a set of signaling molecules located in the downstream pathways, (e.g., mitosis pathways); and the predicates take a value from the set of {T, F} in which T and F denote true and false values, respectively.

Axiom 2: Provided that the conditions of axiom 1 are satisfied, cross-talks among Ca^{2+} signals and phosphorylation/dephosphorylation of particular Ca^{2+} sensitive proteins can be formulated as follows:

Feedback (+, \rightarrow , Calcium-to-phosphorylation (X), Z),
 Feedback (+, \rightarrow , Calcium-to-dephosphorylation (X), Z),
 Feedback (-, \rightarrow , Phosphorylation-to-calcium (Y), Z),
 Feedback (-, \rightarrow , Dephosphorylation-to-calcium (Y), Z),

where + refers to a positive feedback; - refers to a negative feedback; \rightarrow refers to the direction of information flow from a pathway to another pathway according to the appeared order in the predicate form; Z refers to a pathway in a signal transduction network.

The nonlinear dynamics features of network reconstruction described in Axiom 2 may play an important role in understanding and characterizing the role of Ca^{2+} signals in complex signaling networks. One of the latest evidence about the nonlinear features in cell biology is the double punch effect that links Ca^{2+} signals and phosphorylation/dephosphorylation pathways [8].

3.2 Network Coding for Signal Transduction Networks

3.2.1 Network Coding for Dynamical Networks

Linear network coding (LNC for short) [24] is a widely accepted tool that can be used to formulate a complex network structure. LNC is, for example, used to formulate communication networks at the protocol level and to understand the dynamic behavior of communication networks.

Here we introduce necessary notations of dynamical networks in order to apply LNC to signal transduction networks. Corresponding to Fig.1 (d) in [24], we have the following notations:

$$G(\text{LNC}) = \langle V, E \rangle$$

where

$$V = \{x1, u(1,1), u(1,2), u(2,1), u(3,1), y1, y2\},$$

$$E = \{(x1 \rightarrow u(1,1)), (x1 \rightarrow u(1,2)), (u(1,1) \rightarrow u(2,1)),$$

$$(u(1,2) \rightarrow u(2,1)), (u(2,1) \rightarrow u(3,1)), (u(3,1) \rightarrow y1),$$

$$(u(3,1) \rightarrow y2)\}.$$

The network-coding scheme is explained as follows:

Input $X = \{x1\}$
 Output $Y = \{y1, y2\}$
 System $F(X, Y) = G(X)$

The structural matrix of the system $F(X, Y)$ is

0 1 1 0 0 0 0
 0 0 0 1 0 0 0

where the index is with the order of $\{x1, u(1,1), u(1,2), u(2,1), u(3,1), y1, y2\}$.

For nodes in G, the operations take the form of

$$Q(v(i) \rightarrow v(j)), \quad i, j \in N.$$

For dynamical networks, the operations (operators) take the form of

$$Q = \{Q(v(i) \rightarrow v(j)), \quad i, j \in N\},$$

in which the concrete forms of related operators can be realized in terms of (hyper-)graph rewriting. Dynamically reconfiguring signal transduction networks needs the following major operators:

hyper-edge replacement ::= delete or insert linked edges;
 hyper-vertex replacement ::= delete or insert aggregated vertexes

Here notice that in order to unify the notations of (hyper-)graph rewriting, the ‘‘hyper-vertex’’ is defined as a set of the nodes (vertexes) located in the same sub-set in the related graph, i.e., the input set, output set, and the reactant sets in signal transduction networks. The readers can easily find the difference between the above-mentioned term and the well-known one by Courcelle [5].

3.2.2 Network Coding for Phosphorylation/Dephosphorylation

The operation set of LNC is used to formulate signal transduction processes of phosphorylation and dephosphorylation pathways under the regulation of kinases and phosphatases.

For the system (here we use the notations in [15, 16])

$$Y = \text{GKP}(X)$$

where

$X = \{x1\}$
 $Y = \{y1, y2\}$
 x1 – the variable (kinase or phosphatase);
 y1 – the phosphorylated protein;
 y2 – the protein with the dephosphorylation state;
 u(1,1) – kinase;
 u(1,2) – phosphatase;
 u(2,1) – the activation state of kinase or phosphatase;
 u(2,2) – the input to the phosphorylation/dephosphorylation pathway.

3.2.3 Embedding Ca^{2+} Pathways into the LNC Model

Considering diverse building blocks of functional pathways, the signal transduction processes are modeled in the terms of system of systems, in which Ca^{2+} signaling pathway and phosphorylation/dephosphorylation pathway interact. Consequently, a generalized form of several network-coding models is formalized as follows:

$$\text{NC}(\Phi_1, \Phi_2, \dots, \Phi_L) = \Phi_1 \Theta \Phi_2 \Theta \dots \Theta \Phi_L$$

where $\Phi_1, \Phi_2, \dots, \Phi_L$ refers to each network-coding model corresponding to each building block of pathways; Θ refers to the operator of network reconstruction for the interaction of the building blocks; and L is the number of the building blocks.

A binary tree is the basic pattern for unifying multiple network-coding models, and we have that

$T(x) ::=$ a binary tree consisting of a root x and two leafs $\{\text{leaf1}, \text{leaf2}\}$,

$T^{-1} ::=$ a inverse binary tree consisting of two leafs and a root z where the order of graphical description is inverse compared with the normal diagram illustration of a binary tree.

$T(y1, y2) ::=$ a binary tree consisting of two leafs $y1$ and $y2$ and a root in a binary tree.

$T^*(m) ::=$ a binary tree consisting of a root x and two leafs $\{\text{leaf1}, \text{leaf2}\}$ where $\text{leaf1} = \text{leaf2} = m$ or $y1 = y2$.

where $x, y1, y2, y = y1$ or $y2, z, m \in A$, refers to the alphabet set for the entire domain of the underlying problem in this paper.

The notation $G(\text{LNC})$ is modified as

$$\mathcal{K}(x, y) = T(x) \Xi T^*(m) \Xi T^*(m') \Xi T^{-1}(y1, y2) \Xi T(y)$$

where Ξ refers to the partial order relation in algebraic theory to connect the upstream pathway with the downstream pathway in a hierarchical structure for network-coding model. With the operator Ξ , $T(y)$ is added to the classical LNC model. The reduction rules of logical reasoning are given as follows:

$$u(1,1) \rightarrow y1, u(1,2) \rightarrow y2, u(2,1) \rightarrow z, u(3,1) \rightarrow m$$

Notice the related logical inference processes may obey to the constraint of λ - calculus in mathematical logic and/or π - calculus by R. Milner (owing to the limitation of space, we omit the details here).

It is natural that the investigation on the coupling mechanism of multiple indivisible pathways can start from the interaction of Ca^{2+} signaling pathway and phosphorylation/ dephosphorylation pathway:

$$\mathcal{K}(\text{Ca}^{2+}, \text{kinase}) \Phi_c \mathcal{K}(p, p')$$

where p and p' refer to a phosphorylated protein and a dephosphorylated protein, respectively.

$$\Phi_1 \mathcal{K}(x, y) \Phi_1$$

With the MAPK cascade in general, we can get that

$$\mathcal{K}(\text{Ca}^{2+}, \text{kinase}) \Phi_c \mathcal{K}(\text{MAPKKKK}) \Phi_c \mathcal{K}(\text{MAPKKK}) \Phi_c \mathcal{K}(\text{MAPKK}) \Phi_c \mathcal{K}(\text{MAPK})$$

Here the dephosphorylation is omitted (as a default variable in the predicate) in order to obey the tradition of cell biology for representing pathway.

Corresponding to the evidence reported in [8], it is inferred that

$$\mathcal{K}(\text{Ca}^{2+}, \text{kinase}) \Phi_c \mathcal{K}(\text{MAPKKK}) \Phi_c \mathcal{K}(\text{MAPKK}) \Phi_c \mathcal{K}(\text{MAPK}) \text{ s.t. Feedback } (+, \rightarrow, \text{ Calcium-to-dephosphorylation } (\text{Ca}^{2+}), \text{ MAPKKK}),$$

The linear constraint of networking-code model in our hands is "relaxed" and nonlinearity is emerged mainly owing to the feedback mechanism added into the signal transduction network. Readers may reveal the partial order condition we mentioned before, it is reasonably claimed that from the model

$$\mathcal{K}(\text{Ca}^{2+}, \text{kinase}) \Phi_c \mathcal{K}(\text{MAPKKK}) \Phi_c \mathcal{K}(\text{MAPKK}) \Phi_c \mathcal{K}(\text{MAPK}) \text{ s.t. Feedback } (+, \rightarrow, \text{ Calcium-to-dephosphorylation } (\text{Ca}^{2+}), \text{ MAPKKK}),$$

The robustness performance of the underlying signal transduction network can be analyzed by a general algebraic network-coding model as defined in [10]. The arising question becomes how to reconstruct the topological structure of the pathway network as a synthetic way to explain the robustness observed from it.

3.3 Operation for Reconstruction of Pathways Based on a Nonlinear Dynamic Networking Mechanism

As implied in [1], the analog model of molecular concentration of Ca^{2+} signals can be quantified by the discrete measure of kinase in signal transduction. The open problem is how to establish an algorithm to generate kinase-centered cascades for topologically configuring a network by quantitative criteria. Efforts on solving the problem may be assisted by using the algorithm for generating the fixed-point phenomena in a MAPK cascade (reported in [17]), that is, one of the criteria could be formulated as:

IF Δ satisfies the condition Λ

THEN $\mathcal{K}(\cdot)$ can be (re)constructed by the algorithm Υ

where the parameter set Δ refers to the set of the coefficients of the Michaelis-Menten equation, condition Λ refers to the existence of fixed-point in signal cascade, the algorithm in this sub-section is denoted as Υ .

Fixed point has been intensively studied in nonlinear systems, but its application in advanced telecommunication [11] and bioinformatics [12] are only reported in recent years. Our idea is to designate the stable topological network structure by quantifying the coefficients of the biochemical reactions from the signal transduction networks. From the fixed-point constrained configuration, the probability of the network configuration denoted as G is the basis of estimation of the mutual information of the cellular communication processes in Shannon theory.

3.4 Remarks on Capacity of Cellular Communications

There could be various ways to describe the capacity of molecular communication processes. As an equivalent result², the capacity of the cellular communication channel is defined based on the maximum number of the activated second messengers in the pathway network. Therefore, we have that

$$C = G_0(t) + \int_0^t L(t-s) dG_1(s)$$

where G refers to the probability of the constrained network configuration, $G_0(t)$ and $G_1(s)$ refer to the distribution of G under T_0 and T_1 , respectively. T_0 is the initial configuration of the pathway, i.e., joint probability of the parameters from the indivisible pathways; T_1 is the probability of the configuration of the pathway in the reconstruction processes, i.e., joint probability of the parameters from interacting building blocks in pathway networks.

Here the activated number of the second messengers is modeled as a renewal process in probability theory under the constraints we discussed above. The amount of the transmitted information inferred from the maximum configuration of the signal transduction network is equivalent to the maximum of mutual information in Shannon theory. The lower bound of C is

² The above-mentioned definition is an empirical concept or physical term rather than a rigorous mathematics-rooted one. The complete method of calculating C is under development [18].

estimated as $\inf(C) = \kappa(n,m,k) \times f(n,m)$ where n refers to the number of signaling molecules, m refers to the number of the cross-talked signaling molecules, k refers to the layer of MAPK cascade; κ is unction dependent on n , m , and k , from which the “side information” is expected to be extracted, $f(n,m)$ is a polynomial function that is regarded as a measure of the computational complexity of the reconstructed pathway networks.

Based on the information extraction (encoding/decoding) from the “wet” cellular signaling in living cells, cellular pathway networking provides us a bioinformatics framework for identifying the cellular communication protocols in nature based on the self-organizing mechanism in dynamics networks.

4. CONCLUSIONS

How biological and cellular information is encoded, communicated, and decoded with spatial and temporal dynamics of signal molecules is of great interest to cell biologists as well as molecular communication engineers. As the first step to understanding and characterizing the molecular communication media that propagate Ca^{2+} signals as carriers of information, we apply an information theoretic approach to modeling the molecular communication media. The presented model is currently under further refinement to possibly reveal the new aspects of molecular communication.

5. REFERENCES

- [1] Albert Bruce, et al., *Molecular biology of the cell*, the 4th edition, New York: Garland Science, 2002.
- [2] Berridge, M. J., “Neuronal calcium signaling,” *Neurons*, Vol. 21, 13-26, 1998.
- [3] Berridge, M. J., P. Lipp, and M. D. Bootman, “The versatility and universality of calcium signaling,” *Nature Reviews, Molecular Cell Biology*, Vol. 1, pp. 11-21, 2000.
- [4] Berridge, M. J., D. Bootman, and H. L. Roderick, “Calcium signaling: dynamics, homeostasis and remodelling,” *Nature Reviews, Molecular Cell Biology*, Vol. 4, pp. 517-29, 2003.
- [5] Courcelle, B., “Graph rewriting: an algebraic and logic approach,” *Handbook of Theoretical Computer Science, Volume B: Formal Models and Semantics*, pp. 193-242, 1990.
- [6] Eckford, A., “Nanoscale communication with Brownian motion,” in *Proc. 41st Annual Conference on Information Sciences and Systems (CISS 2007)*, 2007.
- [7] Hiyama, S., Y. Isogawa, T. Suda, Y. Moritani, and K. Suto, “A design of an autonomous molecule loading/ transporting/ unloading system using DNA hybridization and biomolecular linear motors in molecular communication,” in *Proc. European Nano Systems*, 2005.
- [8] Jessus, C. and O. Haccard, “Fertilization Calcium’s double punch,” *Nature*, Vo.449, 20, pp.297-298, 2007.
- [9] Klaus, M., Wolf, B. and Wolf, B., Crosstalk between cellular morphology and calcium oscillation patterns, *Cell Calcium*, 19 (6), 1996, pp. 461-472.
- [10] Koetter, R. and M. Médard. An Algebraic Approach to Network Coding,” *IEEE/ACM Transactions on Networking*, Volume 11, Issue 5, pp. 782-796, 2003.
- [11] Kumar, A., Altman, E., Miorandi, D., and Goyal, M., “New insights from a fixed-point analysis of single cell IEEE 802.11 WLANs,” *IEEE/ACM Transactions on Networking*, Volume 15, Issue 3, pp.588 – 601, 2007.
- [12] Li, J. Q. and Li, H., “Using fixed point theorems to model the binding in protein-protein interactions,” *IEEE Transactions on Knowledge and Data Engineering*, Vol. 17, Issue 8, pp.1079 – 1087, 2005.
- [13] Liu, J. Q., “On computational signal transduction in cells: a mathematical model theoretically capable of approaching the Shannon limit,” 1st International Symposium of Nanomedicine-from Basic to Applications (ISNM2007); and in *Proc. 2nd Molecule-Based Information Transmission and Reception (MB-ITR2007)*, p. 76, 2007.
- [14] Liu, J. Q. and K. Shimohara, *Biomolecular computation for bionanotechnology*, Boston & London: Artech House, 2007.
- [15] Liu, J. Q., “On Information theoretical formalization of intracellular communications based on linear network coding,” in *Proc. SICE Annual Conference (International Conference on Instrumentation, Control and Information Technology)*, pp. 909-912, 2007.
- [16] Liu, J. Q. and K. Shimohara, “Advanced control and signal processing paradigms in the new millennium for nano-bio-communication and information processing systems,” Tutorial 1, *Annual Conference (International Conference on Instrumentation, Control and Information Technology)*, pp 1-16, 2007.
- [17] Liu, J. Q., “On quantitative aspect of an information processing model inspired by signaling pathways in cells: an empirical study,” *IPSI MPS Technical Report*, Vol. 2007, No.43, pp.21-24, 2007.
- [18] Liu, J. Q. and T. Nakano, “On channel capacity of molecular communication systems based on cellular signaling networks,” (in preparation).
- [19] Moore, M., A. Enomoto, T. Nakano, Y. Okaie, T. Suda, A. Kayasuga, H. Kojima, H. Sakakibara, and K. Oiwa, “Molecular communication: simulation of microtubule topology,” 1st International Workshop on Natural Computing, 2007. (in submission).
- [20] Nakano, T., T. Suda, M. Moore, R. Egashira, A. Enomoto, and K. Arima, “Molecular communication for nanomachines using intercellular calcium signaling”, in *proc. 5th IEEE conference on nanotechnology*, 2005.
- [21] Nakano, T., Y. H. Hsu, W. C. Tang, T. Suda, D. Lin, T. Koujin, T. Haraguchi, and Y. Hiraoka, “Microplatform for Intercellular Communication,” *Third Annual IEEE International Conference on Nano/Micro Engineered and Molecular Systems*, 2008.
- [22] Prank, K., F. Gabbiani, and G. Brabant, “Coding efficiently and information rates in transmembrane signaling,” *BioSystems*, vol. 55, pp. 15-22, 2000.
- [23] Thomas, P.J., D. J. Spencer, S. K. Hampton, P. Park, and J. P. Zurkus, “The diffusion mediated biochemical signal relay channel,” in *Proc. 17th Annual Conference on Neural Information Processing Systems (NIPS ‘03)*, 2003.
- [24] Yeung, R., S.-Y.R.Li,N. Cai, and Z. Zhang, “Foundation and trends in communications and information theory,” vol.2, nos. 4 and 5, pp. 241-381, 2005.