

Brain-Inspired Stigmergy Learning

XING XU¹, ZHIFENG ZHAO^{1,2}, RONGPENG LI¹, AND HONGGANG ZHANG¹

¹College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou 310027, China

²Zhejiang Lab, AI Town, Hangzhou 311121, China

Corresponding author: Zhifeng Zhao (zhaozf@zju.edu.cn; zhaozf@zhejianglab.com)

This work was supported in part by the National Key R&D Program of China under Grant 2017YFB1301003, in part by the National Natural Science Foundation of China under Grant 61701439 and Grant 61731002, and in part by the Zhejiang Key R&D Development Plan under Grant 2019C01002.

ABSTRACT Originating from entomology, stigmergy has provided an effective framework for swarm collaboration. Based on new discoveries on astrocytes in regulating synaptic transmission in the brain, this paper has mapped stigmergy mechanism into the interaction mediated by the propagation of calcium waves in astrocytes between synapses and investigated its characteristics and advantages. Particularly, we have divided the short-range interaction between synapses that are not directly connected by neurons into three phases and proposed a stigmergic learning model. In this model, the state change of an agent will expand its influence to affect the states of others. The strength of the interaction is determined by the level of neural activity as well as the distance between the agents. Inspired by these findings, we propose a model to help solve task assignment and coordination problems. The simulation results have verified the effectiveness of the proposed model.

INDEX TERMS Stigmergy, astrocytes, synapses, calcium waves, neural networks, artificial intelligence, machine learning.

There is a crack in everything, that's how the light gets in - Leonard Cohen

I. INTRODUCTION

Stigmergy was first introduced by French entomologist Pierre-Paul Grassé in 1950s [1] [2] when studying the behavior of social insects. The word stigmergy is a combination of the Greek words “stigma” (outstanding sign) and “ergon” (work), indicating that some activities of agents are triggered by external signs, which themselves may be generated by agent activities [3]. Stigmergy allowed Grassé to explain why insects of very limited intelligence, without apparent communications, can collaboratively tackle complex tasks, such as building a nest. Combined with computer science, stigmergy has inspired many effective swarm intelligence algorithms, such as Ant Colony Optimization (ACO) [4].

Originating from entomology, stigmergy has been widely studied in the aspect of social insects [5] and applied in different scenes [6]–[8]. But this concept has been rarely mentioned in the study of the brain, which is assumed to be the most complex system. As important glial cells in the Central Nervous System (CNS), astrocytes have been traditionally placed in a subservient position in the past decades, which only supports the physiology of the

The associate editor coordinating the review of this manuscript and approving it for publication was Shubhajit Roy Chowdhury.

associated neurons. However, recent experimental neuroscience evidences indicate that astrocytes also interact closely with neurons and participate in the regulation of synaptic neurotransmission [9]. These evidences have motivated new perspectives for the research of stigmergy in the brain.

There is a large number of complicated biochemical reactions between synapses and astrocytes to support the implementation of various neuromodulations [10]. In the CNS, each astrocyte might contain hundreds or thousands of branch microdomains, each of which encloses a synapse [11]. The synaptic activity can elevate the concentration of Ca^{2+} in the corresponding branch microdomain [12]. Ca^{2+} and inositol-1,4,5-triphosphate (IP_3), as important messengers within astrocytes, are believed to expand the influence of synaptic activities [13]. The range of influence is relevant to the level of synaptic activity as well as the distance between coupled branch microdomains [12]. Besides, these branch microdomains with the elevated concentration of Ca^{2+} can provide a regulation for the wrapped synapses [14]. Therefore, as explained in Section II, general stigmergy can be mapped into this neural process in which astrocytes play the role of medium carriers to provide regulations for the involved synapses.

The propagation of Ca^{2+} in cytosol within astrocytes can comprise both short-range and long-range interactions

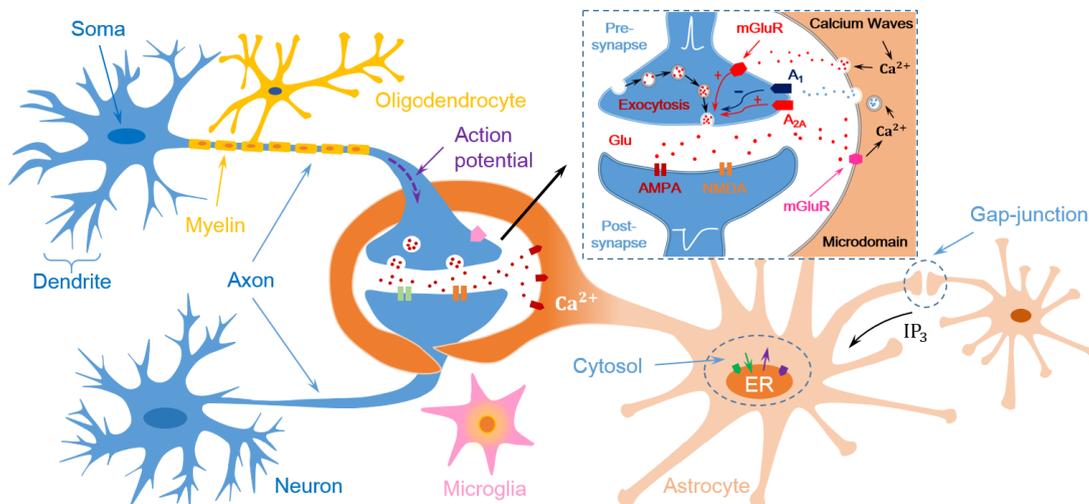


FIGURE 1. An intuitive diagram of the tripartite synapse.

among different synapses. We have preliminarily divided the short-range interaction into three phases in this paper. These phases will be concretely described in Section III so as to formulate the stigmergic system model. In this fundamental model, the strength of interaction between agents is determined by the level of stimulations as well as the distance between them, which is consistent with the strength of interaction between synapses via astrocytes. Inspired by the morphological and functional changes in astrocytes during the environmental enrichment [15], it is likely that the cross regulations of distance between agents is critical to obtain stigmergy learning gain. Accordingly, we have used two different task assignment and coordination scenarios to verify the effectiveness of the proposed model considering the distance adjustment between agents. In the first problem, with the existence of the cross regulation between agents, the proposed model can finish the same task more quickly compared with the method in [4]. In the second problem, we have considered the cross regulation on the basis of stigmergy in the process of moving agents to form a specified shape. Simulation results have indicated that the distance-based stigmergy model can improve the quality of the formed shape.

The remainder of this paper is organized as follows. In Section II, new discoveries on astrocytes in regulating synaptic transmission will be introduced and the existence of general stigmergy in the brain will be explored. In Section III, three important phases within the short-range interaction between synapses will be concretely described. The interaction is referred to set up the stigmergic system model and learning algorithm. In Section IV, simulations on the performance of the proposed stigmergy learning model are carried out on the problem of task assignment and coordination as well as the problem of moving agents, in order to verify its effectiveness and advantages. Finally, we conclude this paper with a summary.

II. STIGMERGY IN THE BRAIN

In the CNS, there are many similarities between general stigmergy and the interaction between synapses which is mainly mediated by the propagation of Ca^{2+} within astrocytes. As important medium carriers, astrocytes are coupled together by the gap-junction to comprise an additional neuromodulatory system that acts in complement to the neuronal ones.

A. GLIAL CELLS IN THE CNS

In the process of nerve conduction, action potentials represented by the purple dotted arrow in Fig. 1 are conducted along the axon to the pre-synaptic terminal. Then a large quantity of neurotransmitters will be released into synaptic cleft through exocytosis. These molecules can bind with various receptors on the surface of post-synaptic terminal to change the terminal membrane potential. Besides, they can also diffuse out from synaptic cleft and bind with receptors of surrounding glial cells, which will typically release neuromodulators in return [12]. In essential, there are three types of glial cells in the CNS: microglia, oligodendrocytes, and astrocytes [16].

Microglia, as illustrated in Fig. 1, are macrophages in the CNS. Their key roles are immune surveillance as well as responding to infections or other pathological states such as neurological diseases or injuries [17] [18]. For the synaptic activity, microglia can play the role of supervision and protection.

Oligodendrocytes can contribute to the plasticity of nervous system in the above-mentioned process of nerve conduction. An action potential normally needs to spend a certain amount of time reaching the pre-synaptic terminal. Many factors affect the conduction velocity, such as the thickness of myelin sheath, the axon diameter and the spacing and width of the Ranvier nodes [19]. Increasing the thickness of myelin sheath can significantly improve the velocity, which helps to

form the salutatory conduction [20]. Oligodendrocytes play a critical role in this process because they can regulate the production of lecithin, which is an important substance for the compound of myelin [20], as illustrated in Fig. 1. Therefore, oligodendrocytes can adjust the arrival times of different nerve pulses through continuously changing the thickness of myelin sheath on each axon branch.

Astrocytes can support the implementation of many neuromodulations because of the enrichment of various receptors [21]. The phenomenon that synaptic terminals and cleft are wrapped by branch microdomains of surrounding astrocytes gives rise to the discovery of tripartite synapse [22], which is illustrated with details in Fig. 1. In Fig. 1, the pre-synaptic and post-synaptic terminals are represented by the blue parts. The branch microdomain within astrocytes is represented by the yellow part. As action potentials reach the pre-synaptic terminal, a large quantity of glutamate (Glu) will be released into synaptic cleft. These molecules can diffuse and bind with metabotropic glutamate receptors (mGluRs) to induce transient calcium elevations in branch microdomains of surrounding astrocytes, as illustrated in Fig. 1. In response, astrocytes will normally release neuromodulators to regulate synaptic transmissions. For example, as gliotransmitters, ATP/adenosine can act on purinergic A_{2A} (or A_1) receptors on the pre-synaptic terminal to increase (or reduce) the number of exocytosis. Besides, due to transient calcium elevations, astrocytes can also release Glu acting on mGluRs on the pre-synaptic terminal to increase the synaptic efficiency [10].

B. REGULATION WITH TWO DIFFERENT TYPES

There are two main types of neuronal activity-dependent Ca^{2+} responses observed in astrocytes [23]: (1) transient calcium elevations that are restricted to the scale of several microdomains and (2) calcium elevations propagating along these microdomains as regenerative calcium waves, eventually reaching the cell soma or neighboring astrocytes [24].

Glu which are released from pre-synaptic terminal can bind on mGluRs located at adjacent branch microdomains, evoking the production of a fix amount of IP_3 [14]. This process is schematically illustrated in the upper part of Fig. 2. As shown in Fig. 2, IP_3 is considered as the second messenger to trigger the release of Ca^{2+} from endoplasmic reticulum (ER) [25]. ER can be considered as a reservoir with higher concentration of Ca^{2+} than that in cytosol.

A basic model in [26] has been used to describe the dynamics of Ca^{2+} in cytosol due to the binding of IP_3 with IP_3 receptors (IP_3Rs) in ER. There are three main flows within the model which are respectively shown in the ER area in Fig. 2. J_{Leak} represents the leakage-flux of Ca^{2+} from ER into cytosol which is largely proportional to the concentration gradient of Ca^{2+} between ER and cytosol. J_{Pump} represents the pump-flux from cytosol into ER which needs to consume energy to maintain a concentration gradient. $J_{Channel}$ represents the channel-flux from ER into cytosol which is generated due to the binding of IP_3 with IP_3Rs . The elevated concentration of Ca^{2+} in cytosol will further

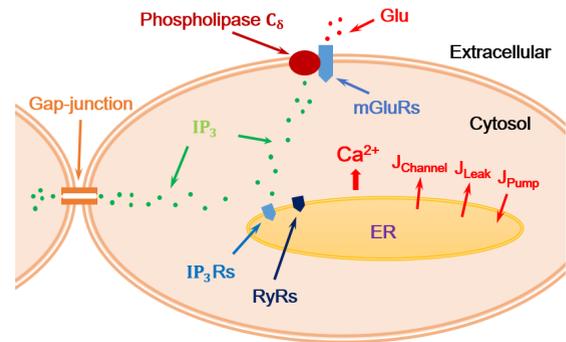


FIGURE 2. Intracellular Ca^{2+} dynamics in astrocytes due to IP_3 .

increase the open probability of IP_3Rs and ryanodine receptors (RyRs) [27], comprising of the mechanism known as Calcium-Induced Calcium-Release (CICR). Nevertheless, excessive concentration of Ca^{2+} in cytosol will bring down the open probability of IP_3Rs and RyRs, and the pump-flux J_{Pump} will become the main factor until a concentration gradient is re-established.

Calcium waves can propagate between astrocytes to incur calcium oscillations [28]. There are many studies trying to describe and model the properties of the gap-junction between various astrocytes [24], as illustrated in Fig. 2. A large number of observations indicate that the gap-junction between astrocytes has a smaller conductance for Ca^{2+} , but a larger one for IP_3 [29]. Therefore, the above-mentioned IP_3 might be the main factor to promote the propagation of calcium waves between astrocytes. Besides, the activation of phospholipase C_δ is also required for the regeneration and propagation of calcium waves [29], as illustrated in Fig. 2.

Typically, calcium elevations occurring in branch microdomains are much more frequent and transient than those in the cell soma [30]. Researchers in [31] indicated that there should be Transient Receptor Potential Ankyrin type 1 (TRPA1) or receptor-gated Ca^{2+} -permeable ions channels in the astrocyte membrane, through which Ca^{2+} could flux into the cell from the extracellular matrix. As mentioned above, two main different types of neuronal activity-dependent Ca^{2+} responses are observed in astrocytes, and thus they may correspond to different regulations: (1) transient calcium elevations which provide the short-range regulation within the scale of several synapses locally, in which Ca^{2+} influx through the receptor-gated ions channels can be the main factor and (2) calcium waves which provide the long-range regulation among neighboring astrocytes, in which IP_3 can be the main factor.

C. ASTROCYTES AS REGULATION NETWORKS

Astrocytes occupy a fundamental position in the synaptic activity. It is suggested that the efficiency of synaptic transmission through the pre-synaptic terminal will be greatly decreased without the calcium signal [10]. Many researchers tried to decode the calcium signal [32] [14] which can

TABLE 1. The main symbols and acronyms.

Acronym	Description
CNS	Central Nervous System
Glu	Glutamate
IP ₃	Inositol-1,4,5-triphosphate
ER	Endoplasmic Reticulum
IP ₃ Rs	Inositol-1,4,5-triphosphate receptors
RyRs	Ryanodine receptors
CICR	Calcium Induced Calcium Release
mGluRs	Metabotropic glutamate receptors
AMPA	α-Amino-3-hydroxy-5methyl-4-isoxazolepropionic acid
NMDA	N-methyl-D-aspartic acid
TRPA1	Transient Receptor Potential Ankyrin type 1
SOM	Self-Organizing Mapping

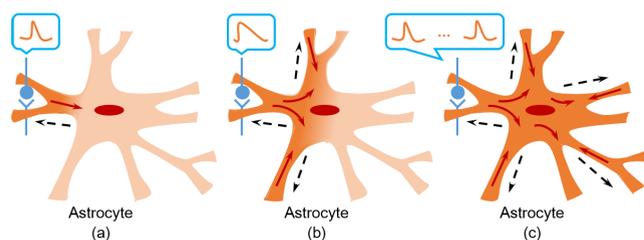


FIGURE 3. Different levels of calcium elevations caused by different levels of synaptic activities.

provide a regulation for the wrapped synapse. Most receptors of neurotransmitters on the membrane of post-synaptic terminal have low affinity. But the biochemical reaction between synapses and astrocytes is largely granted by receptors with high affinity and slow desensitization [12]. Therefore, the influence might not disappear immediately whether it comes from synapses or astrocytes.

In general, consecutive action potentials at a certain synapse can be regarded as a discrete-time pulse sequence. Each of them can change the synaptic state into excitatory or inhibitory. The synaptic state change can further generate calcium elevations in the corresponding branch microdomains which will provide a feedback regulation in return. As a result, the synapse will gradually recover to its original state until the arrival of the next action potential. In this situation, the duration of the calcium elevation depends on the interval of two consecutive action potentials, and the shorter one will produce a longer duration. Therefore, the level of synaptic activities can be measured by the level of calcium elevations. Researchers in [12] found that increasing the level of synaptic activities would lead Ca^{2+} diffusing into the adjacent microdomains, and a persistent high level would eventually make Ca^{2+} full of the whole astrocyte, which is depicted in Fig. 3. In Fig. 3, (a) represents the response of astrocytes under low intensity stimulus. The red solid arrow represents the diffusion direction of Ca^{2+} while the black dotted arrow represents the feedback effect. Fig. 3 (b) is the intermediate result of increasing the intensity of stimulus. Fig. 3 (c) shows the final diffusion effect caused from a synapse which is stimulated by consecutive action potentials.

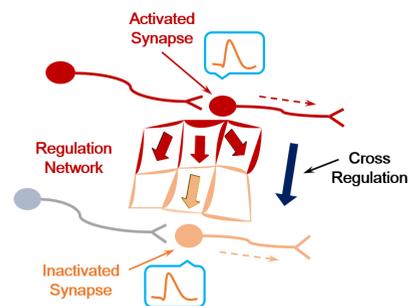


FIGURE 4. A cross regulation provided by astrocytes for nervous system.

In the classical ACO algorithm, each ant in the colony communicates indirectly with others through the pheromone left in a certain area. For each ant, the distribution of the pheromone could be regarded as a pheromone map which can influence the selection of current actions. In the nervous system, different levels of calcium elevations could be generated in the corresponding branch microdomains within astrocytes due to different synaptic activities. Furthermore, these calcium elevations can diffuse into other microdomains comprising an important interaction between different synapses. Therefore, we could say that the distribution of Ca^{2+} within astrocytes could be regarded as a Ca^{2+} concentration map for involved synapses. Researchers in [12] supposed that astrocytes might act as time and space integrators, decoding neuronal information occurring in a large number of synaptic activities. This integration encompasses faster and more local changes based on the rapid activation of small compartments along the astrocytic microdomains up to complex multi-astrocytic and neuronal interactions that are induced by intense and sustained activities resulting in long-term changes in the synaptic network properties.

Astrocytes can be involved in a large number of synaptic events because of their own morphological characteristics. Researchers in [33] suggested that astrocytes could play the role of an additional neuromodulatory system which acts in complement to the neuronal ones, but with its own time and space domains based upon the particular intrinsic properties of Ca^{2+} signaling which encode and integrate incoming inputs from neurons and other environmental sources. Astrocytes can also be regarded as a spatial regulation network, in which synapses could interact with others through the intrinsic Ca^{2+} properties even if there is no direct neuronal connection between them. As described in Fig. 4, this regulation network can provide a cross regulation for nervous system.

D. STIGMERGY IN THE BRAIN

In the hippocampal stratum radiatum, the detailed 3D reconstruction work shows that 80% synapses are coupled with the branch microdomains, and astrocytes almost completely wrap synapses which are rich in docked vesicles [11]. A large number of synapses with certain functions are coupled together through astrocytes to form a potentially collaborative nervous system, in which calcium elevations

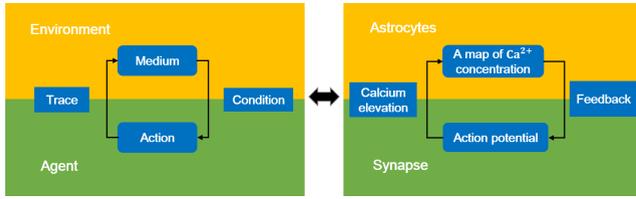


FIGURE 5. A comparison between general stigmergy and the mechanism of stigmergic interactions between synapses through astrocytes.

act as an important communication method between different neurons. Accordingly, general stigmergy can be mapped into the mechanism of cooperative interaction between synapses.

In the nervous system, various synapses can be regarded as different agents, and a map of Ca^{2+} concentration within astrocytes can be regarded as the medium. Action potentials can change the synaptic state into excitatory or inhibitory, which can further generate different levels of calcium elevations in corresponding branch microdomains. This process can be regarded as leaving traces in the medium as in general stigmergy. Astrocytes can be regarded as significant medium carriers, which maintain the map of Ca^{2+} concentration. Astrocytes can also provide the regulation for the involved synapses, whose implementation benefits from a large number of receptors with different types between synapses and astrocytes. This effect can be regarded as the condition provided by the medium for agents. An illustrative comparison between general stigmergy and the mechanism of stigmergic interactions between synapses through astrocytes is illustrated in Fig. 5.

Similar to the pheromone left by ants in a certain area, we assume that calcium elevations which are represented by Ca^{2+} ions with different concentrations can superpose linearly, which might comprise a positive feedback loop potentially. Besides, calcium elevations in astrocytes will normally decay with time, which comprises a negative feedback loop and provides stability for the nervous system with controlled cycles. Because of a limited range of influence, only state changes reflecting the right condition of the nervous system will superpose and have a longer duration. Through this kind of stigmergic process, astrocytes can integrate the neuronal changes from different sources and provide cross-regulation in return for various individual synapses in the nervous system.

III. STIGMERGY LEARNING MECHANISM AND MODEL

A. THE MECHANISM OF INTERACTION BETWEEN SYNAPSES

Though there are two different types of neuronal activity-dependent Ca^{2+} responses observed in astrocytes, the stigmergic model established in this paper is largely based on the short-range regulations within the scale of several synapses which are mainly provided by the transient calcium elevations. This is partly due to that there are still many controversies about the substances involved in the propagation

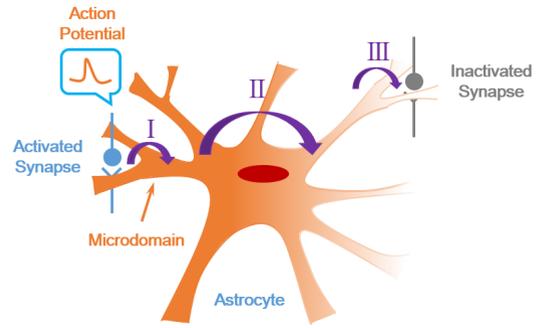


FIGURE 6. Three phases included in the interaction between synapses. They are respectively numbered by I, II and III.

of calcium waves across neighboring astrocytes through the gap-junction [25]. And this part will continue to be studied in the future work. In this section, the process in which a synapse influences the synaptic state of another one within the same astrocyte through the passive diffusion of Ca^{2+} ions is modeled. Specifically, this process has been divided into three phases, which are illustrated in Fig. 6.

The first phase represents the generation of calcium elevations in the corresponding branch microdomain resulting from the arrival of action potentials, which is indicated by I in Fig. 6. At first, the concentration of Glu in synaptic cleft released by the pre-synaptic terminal due to the arrival of an action potential is modeled by [34]:

$$[T_{Neur}] = \frac{T_{max}}{1 + \exp(-\frac{V_d - V_{base}}{K_N})} \quad (1)$$

where $[T_{Neur}]$ is the concentration of Glu in synaptic cleft, and T_{max} represents its maximum value. V_d is the voltage of dendrite in the Pinsky-Rinzel model [35], and it can be used to roughly describe the strength of input in three phases. V_{base} and K_N are parameters used to modify the sigmoid function curve. Glu can diffuse and some of them may act on mGluRs of neighboring branch microdomain to increase the intracellular concentration of Ca^{2+} [10]:

$$\frac{d[Ca^{2+}]}{dt} = \frac{v_{Ca} \cdot [T_{Neur}]^n}{k_{Ca} + [T_{Neur}]^n} - \frac{1}{\tau_{Ca}} ([Ca^{2+}] - [Ca^{2+}]^*) \quad (2)$$

where v_{Ca} is the parameter used to regulate the amplitude of the change rate of Ca^{2+} concentration and k_{Ca} is a scalar factor. n is the power exponent, whose sign can determine two opposite amplification effects. For example, the increase of $[T_{Neur}]$ would improve the change rate of Ca^{2+} concentration if $n > 0$, but would reduce that if $n < 0$. τ_{Ca} is a decay constant. $[Ca^{2+}]^*$ represents the concentration of Ca^{2+} at equilibrium in cytosol. The first item in Eq. (2) expresses the increment of Ca^{2+} concentration in the branch microdomain. The second item indicates that the exorbitant concentration of Ca^{2+} also decreases with time because of the concentration gradient of Ca^{2+} between cytosol and the extracellular matrix.

The second phase considers the passive diffusion of Ca^{2+} ions across microdomains locally, which is indicated

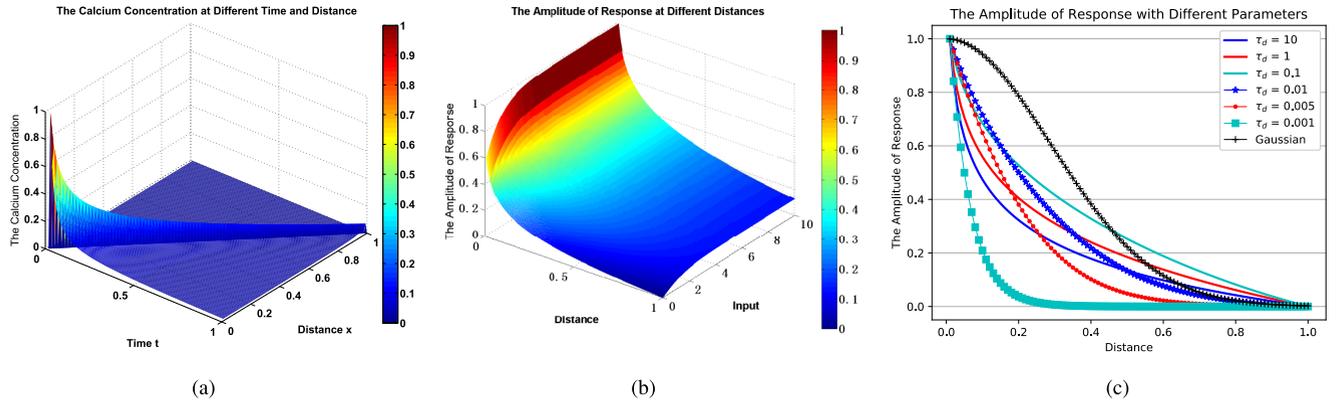


FIGURE 7. (a) The calcium concentration at different time and distance. (b) The amplitude of response at different distances. (c) The amplitude of response with different parameters.

by II in Fig. 6. The telegraph equation which considers a finite speed of propagation on the basis of the second Fike’s law is used here to describe the passive diffusion process of Ca^{2+} [36]:

$$\tau_d \frac{\partial^2 c(x, t)}{\partial t^2} + \frac{\partial c(x, t)}{\partial t} = D \nabla^2 c(x, t) + b(x_0, t) \quad (3)$$

where τ_d is the relaxation factor accounting for a finite propagation speed. $c(x, t)$ is the concentration of Ca^{2+} at location x and time t . D is the diffusion coefficient which is determined by properties of the solution. Furthermore, $b(x_0, t)$ representing the change rate of concentration is given by [37]:

$$b(x_0, t) = \frac{dc(x_0, t)}{dt} \quad (4)$$

where x_0 is the location of the initial point. $b(x_0, t)$ can be represented by the concentration change rate of Ca^{2+} in Eq. (2). Since the input concentration rate is a non-zero value only at the initial point, researchers in [37] have utilized an analytical solution for the telegraph equation to establish a physical end-to-end molecular communication model where $b(x_0, t)$ is assumed to be a Dirac delta function both in time and space (i.e., $b(x_0, t) = \delta(x_0)\delta(t)$):

$$c(x, t) = U(t - ||x||/c_d) \exp\left(-\frac{t}{2\tau_d}\right) \frac{\cosh(\sqrt{t^2 - (||x||/c_d)^2})}{\sqrt{t^2 - (||x||/c_d)^2}} \quad (5)$$

where $U(\cdot)$ is the step function. $||x||$ is the distance from the initial point and t represents the interval from the beginning. c_d is the propagation speed of wavefront which expresses as $c_d = \pm\sqrt{D/\tau_d}$. An intuitive diagram of the calcium passive diffusion process is illustrated in Fig. 7 (a), in which the propagation speed is set as 1 and the concentration range is normalized between 0 and 1. In Fig. 7 (a), the concentration of Ca^{2+} decays quickly with respect to both time t and distance x after the arrival of the wavefront.

The third phase considers the regulation process provided by branch microdomains with elevated calcium for wrapped

synapses, which is indicated by III in Fig. 6. As there are many different methods to decode the properties of intracellular calcium into the effect provided for synapses [32] [14], we apply the result presented in [14] in which a natural logarithmic function was used to model the relationship between the concentration of Ca^{2+} in the corresponding branch microdomain and the amplitude of slow inward currents in the pre-synaptic terminal:

$$\begin{aligned} I_{current} &= k_I \Theta(\ln y) \ln y \\ y &= [Ca^{2+}] - I_{th} \end{aligned} \quad (6)$$

with regard to Eq. (6), there is a threshold value I_{th} for the concentration of Ca^{2+} before astrocytes providing a regulation. k_I is a scale factor. Θ represents the Heaviside function. $I_{current}$ is the amplitude of slow inward currents in the pre-synaptic terminal which can represent the amplitude of response between synapses. The threshold value I_{th} sets a diffusion range of Ca^{2+} in Eq. (3), as calcium elevations with low concentration provide no regulations for neighboring synapses.

In Fig. 7 (a), there are two factors affecting the calcium concentration in the diffusion process: (1) time t and (2) distance x . To simplify the analysis, we model the amplitude of response between different synapses as a distance-dependent function whose value is obtained by sampling as the wavefront arrives. It is intuitively expressed as the ridge line in Fig. 7 (a). Furthermore, we can integrate the above three phases together to describe the amplitude of response between synapses with different distances, as illustrated in Fig. 7 (b). In Fig. 7 (b), the amplitude of response decreases with the increase of synaptic distance and it also increases with the increase of input.

The amplitude of response in the calcium diffusion process can exhibit different properties when the relaxation factor τ_d takes different values, as illustrated in Fig. 7 (c). As a contrast, the Gaussian function which is widely used as the neighborhood function in Self-Organizing Mapping (SOM) of neural networks has also been described. The neighborhood

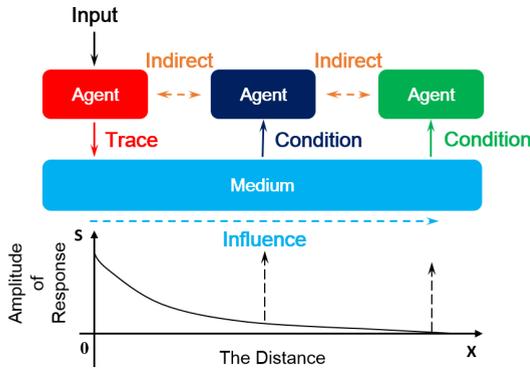


FIGURE 8. The stigmergic learning mechanism.

function used in SOM can represent the similarities between different neurons in the process of feature mapping, which might be analogous to the role played by the diffusion process between different synapses. In general, the amplitude of response in the calcium diffusion process decreases faster at smaller distances compared with the Gaussian function. We will take advantage of this relationship between synapses to coordinate the behaviors of agents.

Rooted in the above three interactive phases, a stigmergic learning mechanism is proposed and illustrated in Fig. 8, in which the communications between different agents (i.e. synapses) labeled by different colors are indirect. When getting a stimulus input, an agent will leave traces (i.e. calcium elevations) which are expressed by the red solid arrow in the outside environmental medium to affect the state of other agents. As illustrated in Fig. 8, the amplitude of response for the interactive influence indicated by the blue dotted arrow is determined by the inter-synapse distance x between agents as well as the intensity of initial stimulus s . The intensity of initial stimulus is consistent with the level of synaptic activity in the nervous system while the synaptic distance can be relevant to the actual distance between the coupled branch microdomains within astrocytes. Especially, the inter-synapse distance x between agents can help to set up **the cross regulation** which can bring more effective collaborations among individuals.

The traces left by different agents in the medium can be regarded as various messages to individuals which can influence the next step selection. Traditionally, agents correct their actions mostly according to the received feedbacks from surrounding environment. In the absence of information exchange with each other, an agent must experience approximately the same amount of trial-and-error process in order to reach the same level of “intelligence” as his workmates. However, with the existence of the cross regulation, an agent may shorten the “learning” process through receiving useful messages implied in the traces. As the traces may contain lots of noise signals, the inter-synapse distance x between agents can help to decide the importance of these messages from different sources. For example, a smaller distance between two agents can implicitly indicate that the messages implied in

their traces are more useful to each other. Therefore, the inter-synapse distance adaptation between agents in the stigmergic learning mechanism can be leveraged and well-regulated to produce more mutually efficient collaborations.

B. SYSTEM MODEL BASED ON THE STIGMERGY IN THE BRAIN

In [5], researchers proposed a strategy to solve the problem of task assignment and coordination among insects colonies based on the mechanism of stigmergy without considering the impact of distance between insects. In this part, we redesign this strategy by taking advantage of the regulation of inter-synapse distance. For ease of representation, we first discuss how the classical stigmergy model works.

For a certain task, several available agents are assumed to be continuously selected out in a probabilistic manner from the group to form a batch. A batch acts as a unit to take actions at different times. This process will continue until the objective of this task is reached. In particular, I and J denotes the set of agents and tasks respectively, and the agent i in I is selected to perform task j in J at index t obeying the following probability $p_{i,j}(t)$, that is,

$$p_{i,j}(t) = \frac{s_j^n(t)}{s_j^n(t) + \alpha \theta_{i,j}^n(t) \cdot \beta \varphi_{i,j}^n(t)} \quad (7)$$

where $s_j(t)$ is the emergency degree of j_{th} task. α and β are weight factors. n is the power exponent. $\theta_{i,j}(t)$ is the state value of i_{th} agent for j_{th} task. $\varphi_{i,j}(t)$ is a heuristic factor. Here, we have changed “+” in [5] into “ \cdot ” in Eq. (7) between $\theta_{i,j}(t)$ and $\varphi_{i,j}(t)$ in order to bring a faster convergence rate. Accordingly, the advantage of this modification will be presented in Fig. 10 and 11. Please refer to the difference between “Classical Stigmergy v.1.0” and “Classical Stigmergy v.2.0”.

In each turn, each agent in a selected batch \mathcal{S} will take an action and then receives a reward for this action. After each turn, $s_j(t)$ will be updated, which is expressed as:

$$R_j(t) = R_j(t - 1) + \sum_{m \in \mathcal{S}} r_{m,j}(t) \quad (8)$$

$$s_j(t) = R_j(t) / T_j \quad (9)$$

where $r_{m,j}(t)$ is the reward that m_{th} agent obtains in j_{th} task at time t from the outside environment. $R_j(t)$ is the sum of all received rewards at time t . T_j is the expected objective for task j . As the accumulation of received rewards for a certain task, $s_j(t)$ will approach to 1 and thus $p_{i,j}(t)$ will approach to 1 which can provide a stimulus with higher intensity for each agent to ensure the task accomplishment.

The state value $\theta_{i,j}(t)$ of various agents for the same task can be different in Eq. (7). After taking an action, this state value will be updated according to the received reward through the following equations:

$$\theta_{i,j}(t) = \theta_{i,j}(t - 1) + \Delta \theta_{i,j}(t) \quad (10)$$

$$\Delta \theta_{i,j}(t) = \rho_1 \cdot \left(\frac{1}{|\mathcal{S}|} \sum_{m \in \mathcal{S}} r_{m,j}(t) - r_{i,j}(t) \right) \quad (11)$$

where ρ_1 is a scale factor. $\Delta\theta_{i,j}(t)$ can be positive or negative, which corresponds to a low or high reward respectively. In Eq. (10) and (11), for each agent, the update of $\theta_{i,j}(t)$ can be regarded as an independent learning process, in which the state value is adjusted according to the “trial-and-error” interaction with the environment. The classical stigmergy learning approach to the problem of task assignment and coordination without considering the impact of distance between agents is presented in Algorithm 1.

Algorithm 1 The Classical Stigmergy Learning Approach to the Problem of Task Assignment and Coordination

- 1: **input** : the objective \mathbf{T}_j , the batch size;
 - 2: **initialize** the task emergency degree \mathbf{s}_j as 0;
 - 3: **initialize** the state value θ_j and the heuristic factor φ_j of each agent;
 - 4: **while** ($\mathbf{s}_j < \mathbf{1}$) **do**
 - 5: calculate the selection probability $p_{i,j}(t)$ of each agent according to Eq. (7);
 - 6: select a certain number of agents in a probabilistic manner to form a batch according to the batch size;
 - 7: agents in the batch take actions and get corresponding rewards;
 - 8: update \mathbf{s}_j through \mathbf{T}_j , \mathbf{R}_j using Eq. (8) and (9);
 - 9: **for each agent in the batch do**
 - 10: update the state value $\theta_{i,j}(\mathbf{t})$ as well as $\varphi_{i,j}(\mathbf{t})$ according to Eq. (10) and (11);
 - 11: **end for**
 - 12: **end while**
 - 13: **output** : the probability of each agent
-

Obviously, in Algorithm 1, the interactions among agents do not incorporate the inter-agent distance factor. Inspired by the findings concerning inter-synapse distance, in this paper, we propose to utilize the mechanism of interaction between different synapses in the brain to regulate the cooperation between agents. Concretely, the stigmergic learning mechanism discussed in Section III-A is used to improve the update of state values in Eq. (10) and (11). The stigmergy learning approach to the problem of task assignment and coordination considering the cross regulation is presented in Algorithm 2. The distance matrix χ in Algorithm 2 contains the distances between any pair of agents in the group. Conditions and traces in the algorithm may have different representations in different scenarios which will be described in details in the following simulations.

IV. NUMERICAL SIMULATION AND RESULTS

In order to verify the effectiveness and advantages of the proposed stigmergy learning model, a number of numerical simulations with different settings have been carried out in this section.

A. THE DISTANCE CROSS REGULATION GAIN

In this part, we primarily assume that there is only one task and various agents in the group may have different rewards

Algorithm 2 The Stigmergy Learning Approach to the Problem of Task Assignment and Coordination Considering the Cross Regulation

- 1: **input** : the objective \mathbf{T}_j , the batch size;
 - 2: **initialize** the task emergency degree \mathbf{s}_j as 0;
 - 3: **initialize** the state value θ_j and the heuristic factor φ_j of each agent;
 - 4: **initialize** the distance matrix χ between agents;
 - 5: **while** ($\mathbf{s}_j < \mathbf{1}$) **do**
 - 6: each agent gets its condition which may have different meanings in different scenarios from the medium;
 - 7: calculate the selection probability $p_{i,j}(t)$ of each agent according to Eq. (7);
 - 8: select a certain number of agents in a probabilistic manner to form a batch according to the batch size;
 - 9: agents in the batch take actions and get corresponding rewards;
 - 10: update \mathbf{s}_j through \mathbf{T}_j , \mathbf{R}_j using Eq. (8) and (9);
 - 11: **for each agent in the batch do**
 - 12: update the state value $\theta_{i,j}(\mathbf{t})$ according to its condition;
 - 13: update the state value $\theta_{i,j}(\mathbf{t})$ as well as $\varphi_{i,j}(\mathbf{t})$ according to Eq. (10) and (11);
 - 14: update the distance matrix χ between agents according to each piece of traces received;
 - 15: leave the trace which may have different meanings in different scenarios in the medium to provide conditions for other agents;
 - 16: **end for**
 - 17: **end while**
 - 18: **output** : the probability of each agent
-

for the task. Moreover, the aim of the simulation is to select agents which can better finish the task so as to examine the distance cross regulation gain. In particular, a random reward (*Agent Reward*) ranging from 1 to 10 is assigned to each agent and these assigned values remain unchanged during the whole simulation process. Moreover, several agents are allowed to take actions together as a batch in each turn. Besides, there is a fixed cost (*Cost*) for each action which is the same for all agents. And an ability value (*Agent Ability*) is randomly assigned to each agent indicating the number of actions to be taken. The real-time ability value can also be utilized as a heuristic factor in Eq. (7). The state values of all agents are set to 0.5 in the initialization, and can change between 0 and 1. Furthermore, the distance between any pair of different agents is set to 5 in the initialization, whose range is between 1 and 10. The main parameters used in this subsection is illustrated in TABLE 2.

As mentioned in section II, the synaptic state change can elevate the concentration of Ca^{2+} in the branch microdomain as an input to affect the states of other synapses. Therefore, we use the state value $\theta_{i,j}(t)$ to represent the synaptic state of i_{th} agent to j_{th} task. And the synaptic state change $\Delta\theta_{i,j}(t)$

TABLE 2. The main parameters in section IV-A.

Item	Description
Agent Number	30
Objective	1100
Batch Size	5
Agent Reward	[1, 10]
Agent Ability	[50, 120]
Cost	10
α	1
β	1
n	2
ρ_1	1
ρ_2	1
σ	0.1

can thus play a role of traces, as illustrated in Fig. 8. Before providing a condition which can influence the state value of an agent, each piece of trace needs to be processed through the diffusion process, whose final value may be determined by the diffusion distance between two agents. Furthermore, in a group, the influenced state value $\theta'_{i,j}(t)$ of an agent will be determined by several traces from other different members, which can be expressed as:

$$\theta'_{i,j}(t) = \theta_{i,j}(t) + \overline{\Delta\theta_{i,j}(t)} \quad (12)$$

$$\overline{\Delta\theta_{i,j}(t)} = \sum_{k \in \pi_i(t-1)} D(d_{k,i}(t-1)) \cdot \Delta\theta_{k,j}(t-1) \cdot \rho_2 \quad (13)$$

where ρ_2 is a scale factor. $\pi_i(t-1) = \{X_k | k \neq i, d_{k,i}(t-1) < d_{th}\}$. $d_{k,i}(t-1)$ is the inter-synapse distance between k_{th} and i_{th} agent at time $t-1$, which corresponds to the distance in Fig. 7(c). d_{th} is a threshold value for the inter-synapse distance. $D(\cdot)$ represents the distance-dependent diffusion function which is sampled from several curves in Fig. 7(c). $\Delta\theta_{k,j}(t-1)$ represents the trace left by k_{th} agent at time $t-1$ for i_{th} agent, which will be discounted by their synaptic distance $d_{k,i}(t-1)$.

The distance between different agents, which is relevant to the amplitude of response or condition, is very important for the propagation of traces. Accordingly, we further put forward the following scheme to regulate the distance between two agents after each turn according to the change $\Delta\theta_{i,j}(t)$:

$$d_{k,i}(t) = \begin{cases} d_{k,i}(t-1) - \eta, & \text{if } \phi > 0 \\ d_{k,i}(t-1) + \eta, & \text{otherwise} \end{cases} \quad (14)$$

$$\phi = \Delta\theta_{i,j}(t) \cdot \Delta\theta_{k,j}(t-1) \quad (15)$$

where η is a constant which decays with time. $\Delta\theta_{k,j}(t-1)$ represents a piece of trace. To some extent, the distance between agents can also represent the similarity of these agents participating in the same task. Therefore, the amplitude of response will be larger if the similarity between two agents is higher. The systematic stigmergy learning gain can be obtained through the adjustment of inter-synapse distance between agents. The pseudo code for the stigmergy learning approach to select agents which can better finish a specified task is presented in Algorithm 3.

Algorithm 3 The Stigmergy Learning Approach to the Problem of Task Assignment and Coordination Considering the Distance Regulation

```

1: input : the objective  $\mathbf{T}_j$ , the batch size;
2: set the number of iterations  $\mathbf{N}$ ;
3: initialize the distance matrix  $\chi$  between agents;
4: for  $n = 1 : \mathbf{N} + 1$  do
5:   if ( $n \leq \mathbf{N}$ ) then
6:     initialize the objective  $\mathbf{T}_j$  as the re-given objective  $\mathbf{T}'_j$ ;
7:   end if
8:   initialize the task emergency degree  $\mathbf{s}_j$  as 0;
9:   initialize the state value  $\theta_j$  and the heuristic factor  $\varphi_j$  of each agent;
10:  while ( $\mathbf{s}_j < 1$ ) do
11:    calculate the influenced state value  $\theta'_{i,j}(t)$  of each agent according to Eq. (12) and (13);
12:    calculate the selection probability  $p_{i,j}(t)$  of each agent according to Eq. (7);
13:    select a certain number of agents in a probabilistic manner to form a batch according to the batch size;
14:    agents in the batch take actions and get corresponding rewards;
15:    update  $\mathbf{s}_j$  through  $\mathbf{T}_j, \mathbf{R}_j$  using Eq. (8) and (9);
16:    for each agent in the batch do
17:      if ( $n = \mathbf{N} + 1$ ) then
18:        update the state value  $\theta_{i,j}(t)$  as  $\theta'_{i,j}(t)$ ;
19:      end if
20:      update the state value  $\theta_{i,j}(t)$  as well as  $\varphi_{i,j}(t)$  according to Eq. (10) and (11);
21:      if ( $n \leq \mathbf{N}$ ) then
22:        update the distance matrix  $\chi$  between agents according to each piece of trace received using Eq. (14) and (15);
23:      end if
24:      leave the synaptic state change  $\Delta\theta_{i,j}(t)$  as traces in the medium to provide conditions for other agents;
25:    end for
26:  end while
27: end for
28: output : the probability of each agent

```

In Algorithm 3, the cross regulation between agents is determined by the distance matrix χ . The distance matrix χ needs to be calculated in advance through $\mathbf{N} = 500$ iterations. At each iteration, the task is re-given one objective and the distance between agents is regulated according to the synaptic state change as well as the traces. As the result of iterations, the state values of all agents as well as the distance matrix χ in the group are given in Fig. 9.

In Fig. 9, agents are arranged in a descending order according to their state values. This order also corresponds to the agent index in the distance matrix χ . Each point with a certain color in the distance matrix χ represents the value of the

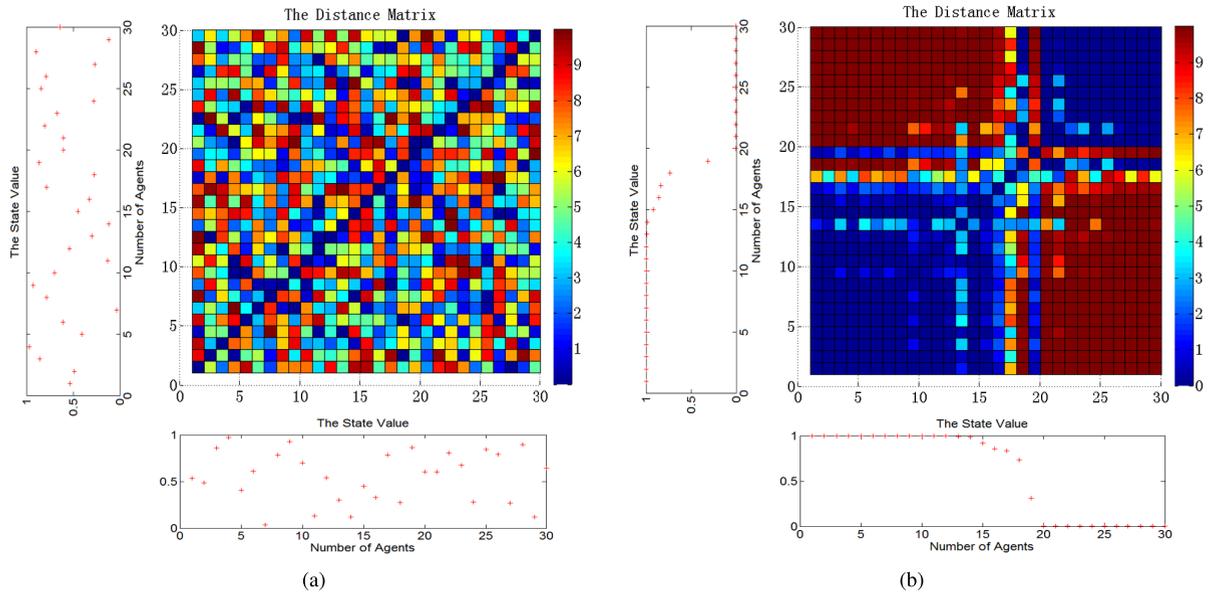


FIGURE 9. (a) The state value of agents and the distance matrix χ before simulations. (b) The state value of agents and the distance matrix χ after simulations.

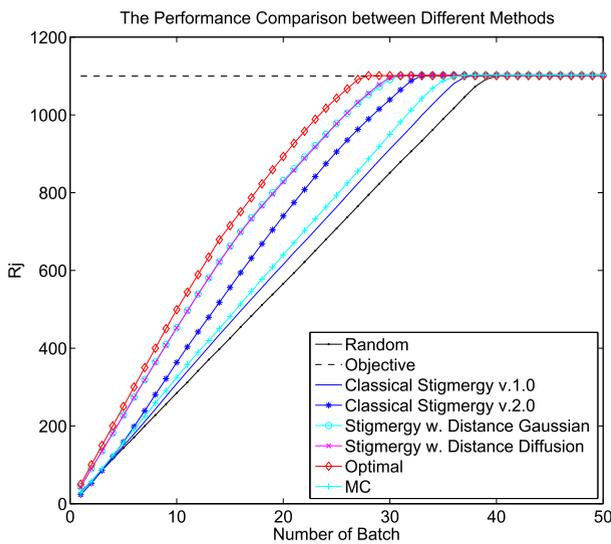


FIGURE 10. The performance comparison between different methods.

distance between two specified agents. It can be observed that compared with the initial results in Fig. 9 (a), there are roughly two clusters in the group according to the final distribution of distance values between agents in Fig. 9 (b) after 500 iterations. And it is basically consistent with the distribution of their state values. As mentioned above, the state value can implicitly indicate the size of reward an agent can obtain in the implementation of the task. For example, a small state value normally corresponds to a large size of reward. Therefore, in Fig. 9 (b), agents with similar rewards are largely added into the same cluster. This relationship between agents can bring the stigmergy learning gain for the system, which will be verified in the following results.

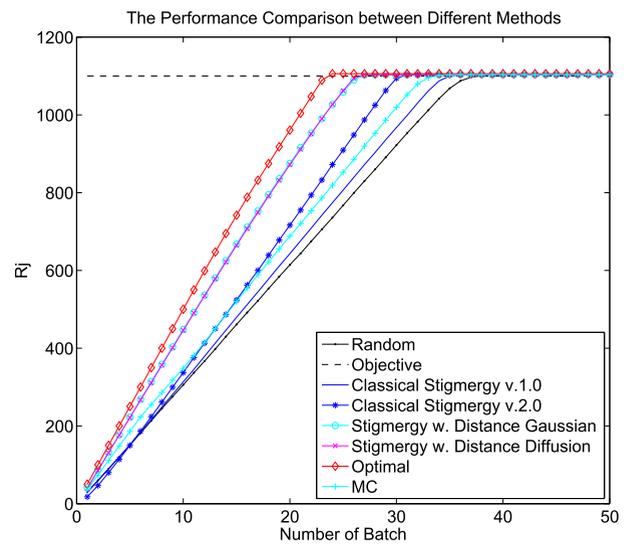


FIGURE 11. The performance comparison between different methods by changing the number of agents to 50.

Fig. 10 provides the performance comparison between the following methods:

- Random: select agents randomly in each turn whose results are represented as “Random” in Fig. 10.
- Classical Stigmergy v.1.0: the method illustrated in Algorithm 1 which uses the addition operation in Eq. (7) between $\theta_{i,j}(t)$ and $\varphi_{i,j}(t)$.
- Classical Stigmergy v.2.0: the method illustrated in Algorithm 1 which uses the multiplication operation in Eq. (7) between $\theta_{i,j}(t)$ and $\varphi_{i,j}(t)$.
- Stigmergy w. Distance Diffusion: the method illustrated in Algorithm 3, which has utilized the distance

matrix χ presented in Fig. 9(b) to regulate the amplitude of the cross regulation between different agents. The distance-dependent diffusion function utilized in Eq. (13) is sampled from Fig. 7(c) with $\tau_d = 0.01$.

- Stigmergy w. Distance Gaussian: the method illustrated in Algorithm 3, which has utilized the distance matrix χ presented in Fig. 9(b) to regulate the amplitude of the cross regulation between different agents. The distance-dependent diffusion function utilized in Eq. (13) is sampled from the Gaussian function.
- Optimal: select agents optimally in each turn whose results are represented as “Optimal” in Fig. 10.
- MC: Monte Carlo estimation method [38] which tries to directly estimate the actual reward of each agent. The estimated reward of each agent can normally be calculated by the average value of all received rewards. Agents with larger estimated reward will be given preference during the selection in each turn. In the Monte Carlo estimation method, estimation of the reward will become more accurate as the number of repetitions increases.

In Fig. 10, R_j represents the sum of all received rewards in the group. It can be observed that all methods can finally complete the task. “Classical Stigmergy v.2.0” seems to perform better than “Classical Stigmergy v.1.0”, which has verified the effectiveness of the modification in Eq. (7). And “Classical Stigmergy v.2.0” seems to perform better than “MC” as the former has utilized the feedback information more adequately through the adjustment of the state value. There seems to be no difference between “Stigmergy w. Distance Diffusion” and “Stigmergy w. Distance Gaussian”, but they both obtain significant system gains through the cross regulation between agents compared with “Classical Stigmergy v.2.0”. Finally, to make the results more convincing, we have increased the number of agents in the group to 50 without changing other conditions and give the results in Fig. 11. It can be observed that the result still hold. As more agents participate in this group, “Stigmergy w. Distance Diffusion” or “Stigmergy w. Distance Gaussian” can obtain even a larger system gain compared with “Classical Stigmergy v.2.0”, which has verified the effectiveness of the cross regulation between agents.

B. THE IMPACT OF THE CROSS REGULATION

The purpose in this simulation is to move a fixed number of moving agents in a certain area to form a specified shape, where agents will share some distance-based information so as to verify the effectiveness of the cross regulation.

Different from the last simulation, agents in this part are all the same. An action of each agent is designed to move a block towards one of four directions: (1) up (2) down (3) left (4) right. As illustrated in Fig. 12 (a), the white block represents an agent which currently has four directions to move. The black block represents a location where an agent can move, but if this position is already occupied by agents, this

agent will stop beside the position. There are totally two different areas in the overall area classified by the target shape: (1) the labeled area which represents the locations where agents need to move (2) the unlabeled area which represents the other locations.

Algorithm 4 The Coordination Method Of Moving Agents To Form A Specified Shape

```

1: input : the objective  $T_j$ , the batch size;
2: initialize the task emergency degree  $s_j$  as 0;
3: initialize the labeled area;
4: initialize the digital pheromone in the total area;
5: initialize the sensing radius and locations of agents;
6: initialize the state value  $\theta_j$  and the heuristic factor  $\varphi_j$  of each agent;
7: while ( $s_j < 1$ ) do
8:   each agent updates the concentration of digital pheromones in the sensing area;
9:   calculate the selection probability  $p_{i,j}(t)$  of each agent according to Eq. (7);
10:  select a certain number of agents in a probabilistic manner to form a batch according to the batch size;
11:  for each agent in the batch do
12:    agent selects an attraction according to Eq. (18);
13:    agent moves one step in the direction of the attraction.
14:    agent gets the reward for this moving according to TABLE III;
15:  end for
16:  update  $s_j$  through  $T_j$ ,  $R_j$  using Eq. (8) and (9);
17:  for each agent in the batch do
18:    update the state value  $\theta_{i,j}(t)$  as well as  $\varphi_{i,j}(t)$  according to Eq. (10) and (11);
19:    update the distance between agents according to their locations;
20:    leave digital pheromones as traces in the current location to provide conditions for other agents according to Eq. (16);
21:  end for
22: end while
23: output : locations of agents

```

In order to make the moving of agents more effective, we need to determine the priority of their movements. For example, the moving of an agent located at the unlabeled area is normally prior to the moving of an agent located at the labeled area. We have used the probability $p_{i,j}(t)$ in Eq. (7) to determine the priority of each agent. And an agent with larger probability normally corresponds to a higher priority. Several agents are selected out from the group to form a batch in each turn. These agents in the batch can choose to move a step based on their own selections. And after each turn, each agent in the batch will receive a predefined reward based on its current location and the number of neighboring occupied positions as illustrated in TABLE 3. Agents which have less

TABLE 3. The reward values in Section IV-B.

Number of neighboring occupied positions	In the labeled area	Reward
4	<i>Yes or No</i>	0
3	<i>Yes</i>	4
3	<i>No</i>	12
2	<i>Yes or No</i>	8
1	<i>Yes</i>	8
1	<i>No</i>	12
0	<i>Yes or No</i>	12

number of neighboring occupied positions and are not in the labeled area are designed to receive larger rewards. This received reward can be used to update the state value $\theta_{i,j}(t)$ of each agent in the batch according to Eq. (10) and (11). Different from its utilization in the last simulation, $R_j(t)$ in Eq. (8) is not updated by the received rewards. It is directly determined by the number of agents in the labeled area. Therefore, T_j in Eq. (9) is determined by the number of agents needed to fill in the labeled area.

In addition to the ability moving in a specified direction, each agent can also have the following four abilities: (1) leaving digital pheromone in the current location (2) sensing the concentration of digital pheromone within a certain range (3) identifying whether the current location is labeled or unlabeled (4) sensing the existence of neighbors in the up, down, left and right directions. Similar to the pheromone left by ants, we use the concept of digital pheromone to represent the trace left at one particular location by agents. After each turn, each agent in the group will leave the digital pheromone in its current location according the following conditions:

$$ph(m, n)(t) = \begin{cases} ph(m, n)(t - 1) + 1, & \text{if } (m, n) \text{ is labeled} \\ ph(m, n)(t - 1) \cdot 0.5, & \text{otherwise} \end{cases} \quad (16)$$

where $ph(m, n)(t)$ represents the concentration of digital pheromone in the location of (m, n) at time t . Digital pheromones can be superimposed adding the remaining concentration left by different agents at previous times. Besides, the concentration of digital pheromone will decline with time to provide a negative feedback for the system.

As each agent can sense digital pheromones with different concentrations and locations within its sensing area, a moving direction need to be determined in order to approach the block with digital pheromone, as illustrated in Fig. 12 (b). In Fig. 12 (b), the white dotted circle represents the sensing area of an agent. And the yellow blocks represent the locations which have digital pheromones with different concentrations. These yellow blocks will attract agents to approach and can also be occupied. An agent normally needs to make a selection from several blocks with digital pheromones which can also be regarded as attractions. During the simulation, each agent will select its attraction in a probabilistic manner,

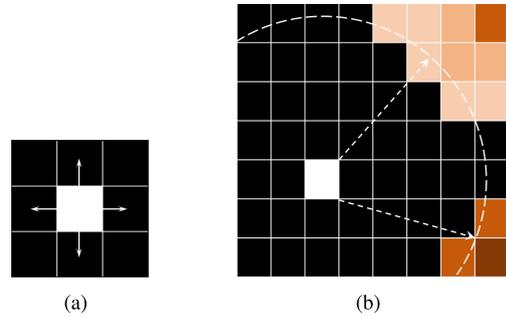


FIGURE 12. (a) The moving method of each agent. (b) The digital pheromones within the sensing area of an agent.

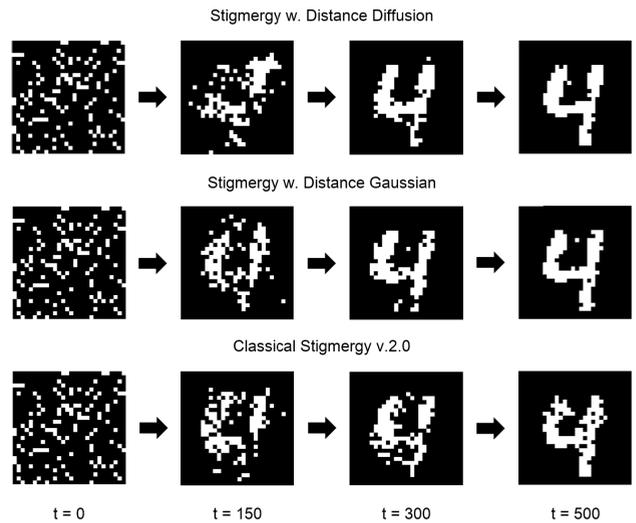


FIGURE 13. The performance comparison in different methods.

that is:

$$C_{i,j} = \frac{\varepsilon_j}{\sum_{j \in \xi_i} \varepsilon_j} \quad (17)$$

where $C_{i,j}$ represents the probability of agent i selecting block j . ε_j represents the concentration of digital pheromone in block j . ξ_i represents the set of blocks with digital pheromones within the sensing area of agent i . Considering the cross regulation, we add a distance-dependent function to the selection of attraction, that is:

$$C_{i,j} = \frac{D(d_{i,j}) \cdot \varepsilon_j}{\sum_{j \in \xi_i} D(d_{i,j}) \cdot \varepsilon_j} \quad (18)$$

where $d_{i,j}$ represents the distance between agent i and block j . In the simulation, as blocks with digital pheromones are normally occupied by other agents, $d_{i,j}$ can thus be regarded as the distance between two agents.

As shown in Fig. 13, the size of the total area is 28×28 . In the original picture, the white part is set as the labeled area and the black part is set as the unlabeled area. The labeled area contains 119 locations which need to be occupied by 119 agents. The main parameters of the simulation are shown in TABLE 4. In TABLE 4, *Sensing Radius* indicates

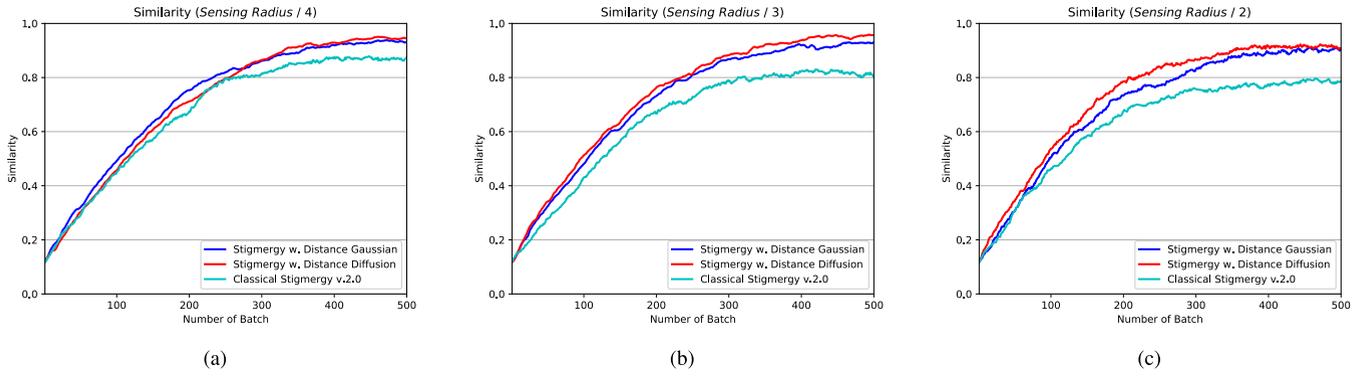


FIGURE 14. (a) Simulation results with *SensingRadius/4* in the labeled area. (b) Simulation results with *SensingRadius/3* in the labeled area. (c) Simulation results with *SensingRadius/2* in the labeled area.

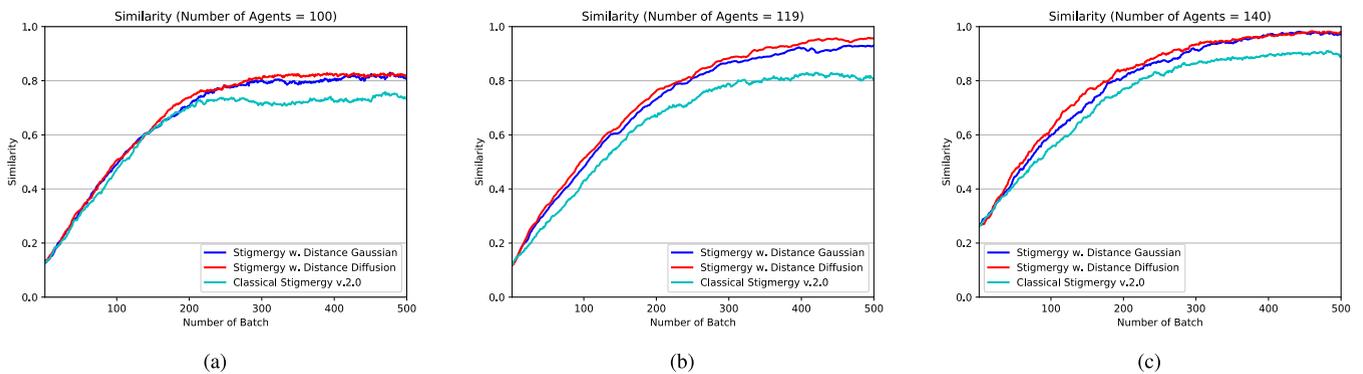


FIGURE 15. (a) Simulation results with 100 agents in the total area. (b) Simulation results with 119 agents in the total area. (c) Simulation results with 140 agents in the total area.

the sensing area of each agent in the unlabeled area. Here, we assume that the distance between two neighboring blocks is 1. The pseudo code for the coordination method of agents to form a specified shape is shown in Algorithm 4.

The performance comparison in different methods are shown in Fig. 13. t represents the number of batches. The results in each method is averaged after 5 runs. We use the similarity between the image formed by agents and the original picture to represent the quality of results, which can be expressed by the value of $s_j(t)$ in Eq. (9). The similarity in the method of “Stigmergy w. Distance Diffusion”, “Stigmergy w. Distance Gaussian” and “Classical Stigmergy v.2.0” are 95.8%, 94.1% and 86.5% respectively. It can be observed that “Stigmergy w. Distance Diffusion” or “Stigmergy w. Distance Gaussian” performs better than “Classical Stigmergy v.2.0”. With the same number of iterations, “Stigmergy w. Distance Diffusion” or “Stigmergy w. Distance Gaussian” can get better moving results. Besides, compared with “Stigmergy w. Distance Gaussian”, “Stigmergy w. Distance Diffusion” gets small improvements in performance.

In Fig. 13, the sensing radius of an agent in the unlabeled area is set 10, while the sensing radius is set 3.33 when an agent is in the labeled area. The reason is that a small sensing radius in the labeled area can prevent an agent

TABLE 4. The main parameters in Section IV-B.

Item	Description
<i>Agent Number</i>	119
<i>Objective</i>	119
<i>Batch Size</i>	10
<i>Sensing Radius</i>	10
<i>Agent Ability</i>	[50, 120]
<i>Cost</i>	1
α	1
β	1
n	2
ρ_1	1
ρ_2	1

continuously jumping between several attractions and can avoid potential ping-pang effects. Accordingly, the simulation results with different sensing radius in the labeled area are presented in Fig. 14. In Fig. 14, the value of Sensing Radius is 10. It can be observed that the performance of “Classical Stigmergy v.2.0” improves as an agent has a smaller sensing radius in the labeled area. Compared with “Classical Stigmergy v.2.0”, “Stigmergy w. Distance Diffusion” or “Stigmergy w. Distance Gaussian” has a better performance and is less sensitive to the change of sensing radius in the labeled area. There are almost no differences between the performance of “Stigmergy w. Distance Diffusion” and

“Stigmergy w. Distance Gaussian” when the sensing radius in the labeled area is set 2.5 in Fig. 14(a). The reason is that an agent will focus more on the nearest few locations regardless of the distance-dependent function. Consistent with results in Fig. 13, it can be observed in Fig. 14(b), “Stigmergy w. Distance Diffusion” performs better than “Stigmergy w. Distance Gaussian”. In Fig. 14(c), “Stigmergy w. Distance Diffusion” and “Stigmergy w. Distance Gaussian” are both limited to a large value of sensing radius in the labeled area, as too many blocks with digital pheromones are presented before an agent which is easy to jump among different attractions.

We have also changed the number of agents in the total area but left the size of the labeled area unchanged. The simulation results are presented in Fig. 15. The sensing radius of an agent in the labeled area is also set 3.33. In Fig. 15 (a), the maximal value of the similarity is 84.03%, and it is 100% in Fig. 15(b) and 15(c), which are consistent with their upper limits. It can be observed that “Stigmergy w. Distance Diffusion” or “Stigmergy w. Distance Gaussian” still performs better than “Classical Stigmergy v.2.0” when the number of agents is different.

The above results have proved that the cross regulation plays an important role for the cooperation of agents in the stigmergy learning mechanism. It can help focus the attention of an agent on more important attractions and avoid irrelevant noise. In the simulation, the cross regulation can improve the quality of agents forming a specified shape and speed up the convergence.

V. CONCLUSIONS

Stigmergy phenomena are widely discovered in natural colonies and perform well through the way of collective collaboration. Inspired by the new discoveries on astrocytes in synaptic transmission, we have explored and mapped stigmergy in the regulation of synaptic activities in the brain. In particular, the short-range interaction between agents (synapses) has been thoroughly studied and a stigmergic learning system model has been put forward. We have found that the regulation of distances between agents plays an important role in the proposed model. The well-regulated distances between agents can bring gain for the system. We have verified its importance in two different simulations by exploiting different cross regulation definitions. However, there is still a long way to fully model the interaction between synapses. For example, for the long-range regulation, the participation of IP_3 must be taken into account, which will be our future research work.

REFERENCES

- [1] P. P. Grassé, “La reconstruction du nid et les coordinations inter-Individuelles chez *Bellicositermes natalis* et *Cubitermes* sp. La théorie de la stigmergie,” *Insectes Sociaux*, vol. 6, no. 1, pp. 41–80, Mar. 1959.
- [2] U. Göllner, “Grassé, Pierre-P.: Fondation des sociétés—Construction. termitologia. 2. 624 S., 452 Fig., 28 Tab., Masson, Paris, New York, Barcelona, Milan, Mexico, Sao Paulo, 1984,” *Deutsche Entomologische Zeitschrift*, vol. 32, nos. 4–5, p. 379, 1985.
- [3] F. Heylighen, “Stigmergy as a generic mechanism for coordination: Definition, varieties and aspects,” *Evol., Complex. Cognition Group*, Vrije Univ. Brussel, 2011.
- [4] M. Dorigo, M. Birattari, and C. Blum, *Ant Colony Optimization and Swarm Intelligence* (Lecture Notes in Computer Science), vol. 49, no. 8. 2004, pp. 767–771.
- [5] M. Dorigo, E. Bonabeau, and G. Theraulaz, *Ant Algorithms and Stigmergy*. Amsterdam, Netherlands: Elsevier, 2000.
- [6] I. Kassabalidis, M. A. El-Sharkawi, R. J. Marks, P. Arabshahi, and A. A. Gray, “Swarm intelligence for routing in communication networks,” in *Proc. IEEE Global Telecommun. Conf.*, vol. 6, Nov. 2001, pp. 3613–3617.
- [7] J. Werfel and R. Nagpal, “Extended stigmergy in collective construction,” *IEEE Intell. Syst.*, vol. 21, no. 2, pp. 20–28, Mar./Apr. 2006.
- [8] A. L. Alfeo, M. G. C. A. Cimino, S. Egidi, B. Lepri, A. Pentland, and G. Vaglini, “Stigmergy-based modeling to discover urban activity patterns from positioning data,” in *Proc. Int. Conf. Social Comput., Behavioral-Cultural Modelling Predict. Behavior Represent. Modelling Simulation*, Jul. 2017, pp. 292–301.
- [9] P. G. Haydon and G. Carmignoto, “Astrocyte control of synaptic transmission and neurovascular coupling,” *Physiological Rev.*, vol. 86, no. 3, pp. 1009–1031, Jul. 2006.
- [10] M. Navarrete and A. Araque, “Basal synaptic transmission: Astrocytes rule!” *Cell*, vol. 146, no. 5, pp. 675–677, Sep. 2011.
- [11] R. Ventura and K. M. Harris, “Three-dimensional relationships between hippocampal synapses and astrocytes,” *J. Neurosci.*, vol. 19, no. 16, pp. 6897–6906, Aug. 1999.
- [12] A. Araque, G. Carmignoto, P. G. Haydon, S. H. R. Oliet, R. Robitaille, and A. Volterra, “Gliotransmitters travel in time and space,” *Neuron*, vol. 81, no. 4, pp. 728–739, Feb. 2014.
- [13] F. Mesiti, P. A. Floor, and I. Balasingham, “Astrocyte to neuron communication channels with applications,” *IEEE Trans. Mol., Biol. Multi-Scale Commun.*, vol. 1, no. 2, pp. 164–175, Jun. 2015.
- [14] S. Nadkarni and P. Jung, “Dressed neurons: Modeling neural–glial interactions,” *Phys. Biol.*, vol. 1, no. 1, p. 35, Feb. 2004.
- [15] G. G. Viola et al., “Morphological changes in hippocampal astrocytes induced by environmental enrichment in mice,” *Brain Res.*, vol. 1274, pp. 47–54, 2009.
- [16] L. Correia, A. M. Sebastião, and P. Santana, “On the role of stigmergy in cognition,” *Prog. Artif. Intell.*, vol. 6, no. 1, pp. 79–86, Mar. 2017.
- [17] D. Erny, A. L. H. de Angelis, and M. Prinz, “Communicating systems in the body: How microbiota and microglia cooperate,” *Immunology*, vol. 150, no. 1, pp. 7–15, Jan. 2017.
- [18] R. M. Ransohoff, “How neuroinflammation contributes to neurodegeneration,” *Science*, vol. 353, no. 6301, pp. 777–783, Aug. 2016.
- [19] S. Pajević, P. J. Bassar, and R. D. Fields, “Role of myelin plasticity in oscillations and synchrony of neuronal activity,” *Neurosci.*, vol. 276, pp. 135–147, Sep. 2013.
- [20] R. D. Fields, “A new mechanism of nervous system plasticity: Activity-dependent myelination,” *Nature Rev. Neurosci.*, vol. 16, no. 12, p. 756, Dec. 2015.
- [21] M. V. Sofroniew and H. V. Vinters, “Astrocytes: Biology and pathology,” *Acta Neuropathologica*, vol. 119, no. 1, pp. 7–35, Jan. 2010.
- [22] A. Araque, V. Parpura, R. P. Sanzgiri, and P. G. Haydon, “Tripartite synapses: Glia, the unacknowledged partner,” *Trends Neurosci.*, vol. 22, no. 5, pp. 208–215, May 1999.
- [23] M. Zonta and G. Carmignoto, “Calcium oscillations encoding neuron-to-astrocyte communication,” *J. Physiol.-Paris*, vol. 96, nos. 3–4, pp. 193–198, May/Jul. 2002.
- [24] M. Goldberg, M. D. Pittà, V. Volman, H. Berry, and E. Ben-Jacob, “Nonlinear gap junctions enable long-distance propagation of pulsating calcium waves in astrocyte networks,” *PLoS Comput. Biol.*, vol. 6, no. 8, Aug. 2010, Art. no. 1000909.
- [25] N. Bazargani and D. Attwell, “Astrocyte calcium signaling: The third wave,” *Nature Neurosci.*, vol. 19, no. 2, pp. 182–189, Feb. 2016.
- [26] Y.-X. Li and J. Rinzel, “Equations for $InsP_3$ receptor-mediated $[Ca^{2+}]_i$ oscillations derived from a detailed kinetic model: A Hodgkin–Huxley like formalism,” *J. Theor. Biol.*, vol. 166, no. 4, pp. 461–473, Feb. 1994.
- [27] I. Siekmann, P. Cao, J. Sneyd, and E. J. Crampin, “Data-driven modelling of the inositol trisphosphate receptor (IP_3R) and its role in calcium-induced calcium release (CICR),” in *Proc. Comput. Gliosci.*, 2019, pp. 39–68.
- [28] M. J. Berridge and A. Galione, “Cytosolic calcium oscillators,” *J. Off. Publ. Fed. Amer. Societies Exp. Biol.*, vol. 2, no. 15, pp. 3074–3082, Dec. 1988.

- [29] G. Ullah, P. Jung, and A. H. Cornell-Bell, "Anti-phase calcium oscillations in astrocytes via inositol (1, 4, 5)-trisphosphate regeneration," *Cell Calcium*, vol. 39, no. 3, pp. 197–208, Mar. 2006.
- [30] K. Kanemaru et al., "In vivo visualization of subtle, transient, and local activity of astrocytes using an ultrasensitive Ca_2^+ indicator," *Cell Rep.*, vol. 8, no. 1, pp. 311–318, Jul. 2014.
- [31] E. Shigetomi, X. Tong, K. Y. Kwan, D. P. Corey, and B. S. Khakh, "TRPA1 channels regulate astrocyte resting calcium and inhibitory synapse efficacy through GAT-3," *Nature Neurosci.*, vol. 15, no. 1, p. 70, Jan. 2012.
- [32] M. D. PittÄä, V. Volman, H. Levine, and E. Ben-Jacob, "Multimodal encoding in a simplified model of intracellular calcium signaling," *Cognit. Process.*, vol. 10, no. 1, p. 55, Feb. 2009.
- [33] P. Bezzi and A. Volterra, "A neuron–glia signalling network in the active brain," *Current Opinion Neurobiol.*, vol. 11, no. 3, pp. 387–394, Jun. 2001.
- [34] A. Destexhe, Z. F. Mainen, and T. J. Sejnowski, "Synthesis of models for excitable membranes, synaptic transmission and neuromodulation using a common kinetic formalism," *J. Comput. Neurosci.*, vol. 1, no. 3, pp. 195–230, Aug. 1994.
- [35] P. F. Pinsky and J. Rinzel's, *Intrinsic and Network Rhythmogenesis in a Reduced Traub Model for CA₃ Neurons*. New York, NY, USA: Springer, 1995.
- [36] Y. M. Ali and L. C. Zhang, "Relativistic heat conduction," *Int. J. Heat Mass Transf.*, vol. 48, no. 12, pp. 2397–2406, Jun. 2005.
- [37] M. Pierobon and I. F. Akyildiz, "A physical end-to-end model for molecular communication in nanonetworks," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 4, pp. 602–611, May 2010.
- [38] L. P. Kaelbling, M. L. Littman, and A. W. Moore, "Reinforcement learning: A survey," *J. Artif. Intell. Res.*, vol. 4, no. 1, pp. 237–285, May 1996.



XING XU is currently pursuing the master's degree with the College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou, China. His research interest includes collective intelligence.



ZHIFENG ZHAO received the bachelor's degree in computer science, the master's degree in communication and information system, and the Ph.D. degree in communication and information system from the PLA University of Science and Technology, Nanjing, China, in 1996, 1999, and 2002, respectively.

From 2002 to 2004, he was a Postdoctoral Researcher with Zhejiang University, China, where his works were focused on multimedia next-generation networks (NGNs) and soft-switch technology for energy efficiency. From 2005 to 2006, he was a Senior Researcher with the PLA University of Science and Technology, where he performed research and development on advanced energy-efficient wireless router, ad hoc network simulator, and cognitive mesh networking test bed. He is currently the Director of the Research Development Department at Zhejiang Lab, China. He is also with Zhejiang University. His research interests include cognitive radio, wireless multi-hop networks (Ad Hoc, Mesh, and WSN), wireless multimedia networks, and green communications.

Dr. Zhao is the Symposium Co-Chair of the ChinaCom 2009 and 2010. He is the Technical Program Committee (TPC) Co-Chair of the IEEE 10th IEEE International Symposium on Communication and Information Technology (ISCIT) 2010.



RONGPENG LI received the B.E. degree from Zhejiang University, Hangzhou, China, in 2010, and the Ph.D. degree from Xidian University, Xi'an, China, in 2015, both as "Excellent Graduates." From 2015 to 2016, he was a Research Engineer with the Wireless Communication Laboratory, Huawei Technologies Co. Ltd., Shanghai, China. He returned to academia, in 2016, initially as a Postdoctoral Researcher with the College of Computer Science and Technologies, Zhejiang

University, which is sponsored by the National Postdoctoral Program for Innovative Talents. He is currently an Assistant Professor with the College of Information Science and Electronic Engineering, Zhejiang University. His research interests currently focus on reinforcement learning, data mining, and all broad-sense network problems (e.g., resource management and security). He has authored/coauthored several papers in the related fields. He serves as an Editor for *China Communications*.



HONGGANG ZHANG was the International Chair Professor of Excellence with the Université Européenne de Bretagne and Supélec, France. He is currently a Full Professor with the College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou, China. He is also an Honorary Visiting Professor with the University of York, U.K. He was a coauthor and an editor of two books *Cognitive Communications–Distributed Artificial Intelligence (DAI)*, *Regulatory Policy and Economics, Implementation* (John Wiley & Sons) and *Green Communications: Theoretical Fundamentals, Algorithms and Applications* (CRC Press), respectively.

He served as the Chair for the Technical Committee on Cognitive Networks of the IEEE Communications Society, from 2011 to 2012. He is also active in the research on cognitive radio and green communications. He was a leading Guest Editor of the *IEEE Communications Magazine–Special Issues on Green Communications* as well as a Series Editor of the *IEEE Communications Magazine* for its Green Communications and Computing Networks Series. He is the Associate Editor-in-Chief of *China Communications*.

...