International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-9 Issue-1, October, 2019

Simulation Framework on Developing Neural Network Model of C. Elegans

Raihaan Kamarudin, M.N. Shah Zainudin, R.H. Ramlee, M.I. Idris

Abstract: ***The nervous system is a complex yet efficient structure – with superior information processing capabilities that surely surpass any man-made high-performance computer. However, without having a complete architectural blueprint of this "technology," understanding its underlying mechanisms is a challenging task. Current research has focused on investigating the interactions between neurons. However, this is not sufficient to explain each neuron's functionality – as they might be physiologically different. One of the approaches to understanding the role of a specific neuron is to observe how its behaviour changes as a network develops. Notwithstanding, observations of biological changes in behaviour accompanying network developmental are experimentally challenging. This therefore creates a new possibility for research exploration: using a network developmental model that can simulate the behaviour of the biological network during the development process. In this research, the biological network of C. Elegans is used as the foundation for the design of the model. Although composed of only with 302 neurons, the nematode's network has the capability of handling complex biological processes, for instance, locomotory and sensory behaviour. In addition, the invariant developmental cell lineage and available network connectivity information provide advantages in terms of analysing the developmental pattern. This creates the possibility of defining a mathematical description of the developmental trajectory. This work aim to design a network trajectory model that is governed by specific rules – a model that has the ability to self-generate a single cell for the multicellular network. For this purpose, a high-processing event-based neural simulator with a 2D virtual environment is developed – to be used for the simulation of a complex multi-stimulus environment. This paper describes the approach and initial design of the network developmental trajectory model.***

Keywords **: *C. Elegans, Development Trajectory Model Neural Network, Nervous System.***

Revised Manuscript Received on October 30, 2019.

* Correspondence Author

Raihaan Kamarudin*, CeTRI, Faculty of Electronics and Computer Engineering, Universiti Teknikal Malaysia Melaka, Melaka, Malaysia. Email: raihaan@utem.edu.my

M.N. Shah Zainudin, CeTRI, Faculty of Electronics and Computer Engineering, Universiti Teknikal Malaysia Melaka, Melaka, Malaysia. Email: noorazlan@utem.edu.my

R.H. Ramlee, CeTRI, Faculty of Electronics and Computer Engineering, Universiti Teknikal Malaysia Melaka, Melaka, Malaysia. Email: radihusin@utem.edu.my

M.I. Idris, CeTRI, Faculty of Electronics and Computer Engineering, Universiti Teknikal Malaysia Melaka, Melaka, Malaysia**.** Email: idzdihar@utem.edu.my

I. INTRODUCTION

Understanding how the nervous system functions is a popular topic that it is extensively discussed in the research world. The nervous system's superior ability to memorize, learn and analyse complex information makes is a valuable system to study. Although plenty of knowledge has been uncovered in this area, numerous mysteries remain unanswered. One of the constraints is the lack of physiological data on the neural network itself. For instance, although the nervous system network of *C. Elegans* is composed of only with 302 neurons and its connectivity is known in detail, researchers still struggle to fully understand the mechanisms of biological behaviour in this organism. One of the approaches to uncovering such mechanisms is to observe the behavioural and neural activity changes during the network's developmental process. Studying network development might give us a clue about the functionality of the sub-circuit and each individual neuron through correlating network development with behavioural changes. The current approach usually involves a mutant *C. Elegans* that lacks certain key neurons; its behaviour is compared with that of a normal nematode in order to understand the functionality of those neurons. Observing the network's interaction with the outside world during the development process is also important in order to reveal and understand the neural mechanism – although this is experimentally difficult to implement. Thus, it would be beneficial to have a system model capable of self-generating the trajectory of the developing network, as this could enhance our understanding of the dynamics of neural activity. The invariant network developmental process of *C. Elegans*, together with the availability of a complete map of the organism's cell lineage and network connectivity, provide the ability to model this process. This paper is organised as follows. Section two, current understanding of the *C. Elegans* nervous system development process is explained; equally important is realizing the limitations we face in designing the development trajectory model. Section three discussed current approaches used in research on the mechanisms of nervous system development, with a focus on work done on *C. Elegans*. Section four briefed the overview on infrastructure of neural simulator which specifically developed to simulate the behaviour of the generated networks in a virtual environment. In section five, the initial idea for designing the development trajectory model of *C. Elegans* is proposed.

This idea explains the research process for designing and validating the development model and details the approaches used to define the algorithm(s) or mathematical definition(s) governing development rules. In last two sections, discussion and conclusion are made regarding the current work and the future research trajectory is discussed.

II. INVARIANT NETWORK DEVELOPMENT IN C. ELEGANS

A. Overview on *C. Elegans* Nervous System Development

Development of a complex nervous system with 302 neurons in *C. Elegans* involves important biological processes, from cell division to cell fate specification, cell migration and synaptogenesis. In *C .Elegans*, all of these processes occur within two different life phases – the embryonic and post-embryonic phases, although most of the neurons are born during the former phase. From a single egg cell (known as a zygote), a process of repeated cell division generates thousands of cells of different cell types and classes. The mechanism of cellular division is complex, as it involves asymmetric division especially in the early development stages, where one cell might differentiate and give birth to two cells of different sizes and "cell cycles" (period starting from birth before the cell division process). Furthermore, not all cells survive during development; some die and are removed from the system under a cell death programme (CDP). The nervous system, however, is generated over period of interaction between neurons during the developmental process. This interaction creates synaptogenesis or the formation of connections (known as synapses). Statistically, at the end of the development process, of the 1,090 of cells born, 131 of them have died under a CDP, and more than one-third of the remaining (302 neurons) form the nervous system. Despite of the advantages of the invariant developmental process and simple network in *C. Elegans*, the lack of essential biological information hinders us from fully understanding and accurately modelling this mechanism.

B. Complete Map of the *C. Elegans* Developing Network: What do we know*?*

The main reason that *C. Elegans* is widely use in neural network development research is because of its invariant development pattern. Each cell has a unique but invariant development trajectory; thus we can trace its origin from the early division process.

Temporal and Spatial Information of Individual Neurons

The transparent body of *C. Elegans* gives it an advantage as an adequate subject in which to study the nervous system development process; it is possible to trace each cell's position and time of birth, allowing the creation of a complete developmental plan. As reported in [1], an asymmetric and asynchronous cell differentiation process during the early embryonic stage creates six founder cells (AB, MS, E, C, D and P_4) that are crucial in determining cell fate. Generally, most of the neurons are created within the AB lineage and during the embryonic phase, although a small number are born from other founder cell lineages. Additionally, cell differentiation process for the first four cells is critical, as this asymmetric and asynchronous division process creates the 3D body plan in the nematode, including the establishment of the following:

- ***Anterior-Posterior (AP) axis:*** Formed during the 1st asymmetrical cell division in the one-cell stage, with AB located at the anterior and P_1 at the posterior of embryo.
- ***Dorsal-Ventral (DV) axis:*** Formed during the 2nd cell division process in the transition between the 2- to 4-cell stage. The difference in cell division timing between AB and P_1 and the fact the AB cell is larger than P_1 create dynamic movement inside the fixed-size embryo.
- ***Left-Right (LR) axis*:** Formed during 3rd cell differentiation process between the 4- and 6-cell stage. ABa and ABp cells divide and both create two daughters that determine the left-right of the embryo.

C. Limitations: Unanswered Questions

This section will discuss important aspects of the network development process that neuroscientists are still struggling to fully understand. These limitations should be taken into consideration in designing the developing network model.

Mechanism of Asymmetric and Asynchronous Cell Differentiation Process

Remarkably, the invariant map and fixed cell fate of lineages in *C. Elegans* involves an asymmetric and asynchronous cell division process, especially during early cell division. For instance, during the first cell division, the parent cell (zygote) creates two children (AB & P_1) that differ in molecular size, with AB being larger than P_1. In addition, the developmental potential of the daughter cells are not equal; the cell cycle of AB is shorter and thus the next cell division occurs asynchronously, with AB dividing 2 minutes earlier than P_1 [2]. One of the factors affecting the cell cycle is PAR polarity proteins, which regulates the timing of mitotic entry and the rate of DNA replication.

Additionally, the statistical analysis done in [3] shows that these variations are dependent on i) the global developmental clock of the embryo (depending on the experimental conditions); and ii) the variation of the local individual cell cycle relative to the global embryo clock. Therefore, to model the timing variation in cell division, we might make an assumption that two important "variables" are essential: i) the cell cycle of the parent cell; and b) the size factor of the daughter cell.

Cell Death Programme

Not all cells generated during the development process survive; 131 cells will "die" and disappear from the system under a cell death programme (CDP) – to be more precise, 113 cells in the embryonic phase and 18 cells in the post-embryonic phase [4]. The CDP mechanism consists of three important processes: *the specification phase*, *the killing phase* and *the execution phase*, where the processes are regulated by more than 24 genes/proteins. The *specification phase* is the period in which the cell is destined for death depending on the presence of specific genes. Three important genes (egl-1, ced-3, ced-4) promote "dying" behaviour in the *killing phase.* In the *execution phase,* cells start "shrinking" due to DNA degradation, before being eaten by neighbouring cells.

One interesting fact is that most of the "death cells" have a neuronal fate, meaning that there is high probability that those cells that do not die will become neurons. However, the big question here is how the cell knows it going to die, and more importantly, why it happens at an exact time in its life cycle. The CDP in *C. Elegans* generally involves one of two mechanisms: either the cell is pre-programmed to "die" or the death occurs due to an external signal (e.g., γ-irradiation) [4]. In former mechanism, the cell acts autonomously and seems to know when it should die; the cell lineage is the best indicator of the spatial-temporal information of the CDP. For the development trajectory model, it might not necessary to specify the biological process of "*killing*" and "*execution*"; however, determination of which cells should "die" and be removed from the system is essential. Cells that have a survival destiny show low or absent activity of the egl-1 gene, indicating that it might be important to include this parameter in our model. However, the mechanism that regulates egl-1 gene expression in the cell remains unresolved, leaving the question of how to define a specific rule to express this mechanism.

Ratio of Excitatory to Inhibitory Neurons in the Developing Network

It is well understood that a balance between excitatory and inhibitory neurons is essential for keeping the nervous system functioning normally. For instance, evidence suggests that the co-existence of synaptic excitation and inhibition in the brain's cortex is essential in shaping spontaneous neural activity [5]. In general, inhibition in the cortex originates from neurons that produce the GABA neurotransmitter; GABA inhibits action potential generation at post-synaptic neurons. These neurons (mainly interneurons), despite making up only around 20% of the neuron population in the cortex, keep cortical neurons from firing constantly.

Stages	Sub-stages	Details of Biological Behaviour	Timeline
EMBRYONIC	**Proliferation** *28 Cells* *140 Neurons*	• Most cell division happens in this sub-stage; first asymmetric cell differentiation happens after 40 minutes. • In the initial 150 minutes, only founder cells are created. • At 150 minutes onward, there are different sub-stages with periods of ***synchronous stem cell division*** and ***specific cell migration***. Cell fate specification occurs, in which a massive number of neuronal and non-neuronal cells are born. • At the end of this stage, the majority of neurons are residing in the ***head and along ventral cord***; only a few neurons are found in the tail.	0 min 40 min *1st cell differentiation* 150 min *Start of Cell migration*
	Organogenesis *558 Cells* *201 Neurons*	• This stage occurs between ***350 minutes and 800 minutes***, and only ***terminal differentiation (with no cell division)*** happens during this period. • In the late 3-fold stage (between 550 and 840 minutes), there is an ***indication of locomotory development,*** where the worm starts moving around its longitudinal axis. • There are also a ***small number of neurons*** born within this period that reside in the head. • By the end of this stage and before hatching, the total number of neurons is ***approximately 201.***	350 min 550 min *1st locomotion behaviour*
POSTEMBRYONIC	**L1** *270 Neurons*	• Developmental program restarts with ***cell division at 180 minutes*** after hatching in the presence of food. • If there is no food present, development is arrested and might enter the *Dauer* stage if conditions are not pleasant. • Another ***five classes of motor neurons*** (VAn, VBn, VCn, ASn, VDn) are generated along the ventral cord at the end of this stage. In addition, pre-existing ***DDn motor neurons complete their synaptic connections***. • ***Synaptic pattern of DDn*** neurons changes after appearance of VD motor neurons. • Number of neurons at the end of this stage is ***approximately 270.***	840 min *Hatching* 1000 min *Developmental restart*
	L2 ~Adult *959 Cells* *302 Neurons*	• Little ***cell division*** occurs during the L2 stage. • The adult *C. Elegans* it has ***302 neurons*** (***381 neurons*** in males, with most of the extra neurons related to mating behaviour)	≈30 hours ≈60 hours

Fig. 1.Details of developmental process of *C. Elegans*
(The development process happened in invariant pattern within 4 different stages/life cycles)

In addition, besides having high connectivity between themselves, individual interneurons in the cortex also establish connections with a large number of principal neurons (usually excitatory neurons); thus, these cells also influence the neuronal activity in different cortical regions [5]. Furthermore, the number and distribution of excitatory and inhibitory inputs onto single neurons may vary significantly between neighbouring neurons. Therefore, differences in this excitation-inhibition ratio could lead to different spatial activity patterns in the cortex. However, it is worth noting that the excitation-inhibition ratios in cortical networks are not fixed; somehow, they proportionally change according to local and/or incoming excitation [5]. There is also evidence showing that during network development, the balance between excitation and inhibition governs the establishment of sensory system projections, including the onset of the critical period for visual system plasticity [5]. Despite the clear importance of balanced excitation and inhibition in brain function, the mechanism of developmental control of this balance remains unknown.

III. RESEARCH ON THE DEVELOPMENT OF THE *C. ELEGANS* NERVOUS SYSTEM

Two models are required to define the trajectory of the nervous system development process: i) a model of cell lineage tree development, and ii) a model of network establishment. This section reviews related work that might provide insights into how these models could be developed.

A. Reconstruction of Cell Lineage

Cell Lineage Model

The cell lineage of *C. Elegans* is invariant, making it possible to define the nematode's developmental trajectory, perhaps using a simple model. In [6], the author proposed a DRGN (Dynamic Recurrent Gene Network) model to control the developmental trajectory of a cell. The model defines the transcription factor of each cell using a simple genetic network, where the network is constructed from N_S structural nodes, N_R regulatory nodes and one Boolean input. The asymmetry of child cells is dependent on the cell's spatial position during cell division – providing an input of either 0 or 1 to the network. The model is reported to be capable of accurately generating all six founder cells after the 4-cell division stage and maintaining the accuracy of the first five cell divisions. It is capable of finding at least one solution that represents the unique pattern of each individual neuron, although the connection coefficient between the nodes needs to be optimized in advance. Additionally, the model is capable of finding the solution that generates an identical sub-lineage in *C. Elegans* with 100% accuracy, with the exception of a complex MS-lineage due to the lineage's size and irregular structure. Nevertheless, the model evolves in a fixed time step – cell division happens in fixed intervals that are not biologically plausible and not suitable if information on the neuron's birth time is required to accurately generate the nervous system. A different approach used to model the *C. Elegans* tree lineage is reported in [7]. The model introduced in this work is claimed to have the ability to define asymmetric and symmetric cell division. The model defines the cell lineage tree development using the *factors* that contribute to any process, for instance, the *factor* that determines the cell type and the *factor* that initiates cell division. The advantages of this model are firstly, that the mechanism of asymmetric cell division considers the intercellular and extracellular factors – the transcription rate and cell-cell signalling, respectively. Secondly, in contrast with [6], cell division does not occur at a fixed simulated time; instead, cell division is initiated by a *"Division"* factor, which is closer to the biological process. Thirdly, the model is capable of specifying different cell types within the cell lineage. The developers claim that the model is applicable to any organism for generating a "tree" lineage, although the list of *factors* involved in the development process is needed in advance. Additionally, we might inquire as to what the distribution pattern of the *C. Elegans* cell lineage development is. By using a similar network structural analysis as reported in [8], development motifs of the *C. Elegans* cell lineage can be identified. From the biological data, it can be seen that *C. Elegans* has repeated motifs or a tree pattern, which is impossible to generate under a stochastic developmental trajectory. Additionally, using only DRGN, as explained above, creates little diversity in tree lineage – different from the *C. Elegans* lineage. Therefore, it is suggested that *C. Elegans* appears to share structural characteristics with both the developmental and stochastic patterns.

Analysis of Division Timing

Although the cell lineage and each cell division during the development process are invariant, the division timing is asynchronous between cells. Biological studies show that the cell division time is dependent on the cell cycle. The question here is what factors regulate the cell cycle. One of the factors is DNA replication rate, as reported in [2]; it regulates asynchronous cell division, especially in the early embryonic stage. In [3], it is suggested that the timing of *C. Elegans* cell division is reproducible, and depends on the following factors: i) the cell cycle of founder cell (depends on lineage development pace), ii) the cell cycle ratio between parent and child cell; and iii) the fate of the individual cell, or its specific role, based on its structural and physiological functionality. However, the cell cycle overall is adjusted proportionally according to the embryo developmental speed, which itself is affected by external circumstances (e.g., temperature). Additionally, one important issue is highlighted in [9]: If a cue from another cell regulates cell fate, cell division is not a deterministic process but depends on external factors.

B. Mechanism of Neural Network Establishment

In modelling the network establishment process, one key question needs to be answered: What mechanism underlies the establishment of the synapse between any two neurons?

Spatio-Temporal Connectivity Changes

The network in adult *C. Elegans* has fixed and pre-determined connections. However, similar to any nervous system, the connectome of *C. Elegans* also undergoes several stages in its life cycle and changes temporally. In [10], Varier and Kaiser published a statistical study on the development of *C. Elegans* that used real biological data to look at the temporal and spatial connectivity patterns. The authors analysed the connection pattern in *C. Elegans* according to the birth time and 3D position of each neuron in different discrete time frames during embryonic burst and post-embryonic cycle stages. Their observations show that early interactions between connected neurons may be essential, as most of the neurons that have long distance connections are born around the same time. The observations also suggest that early-born neurons have the tendency to create more synapses compared to neurons that appear at later stages.

In [11], the connection mechanism based on the neuron development time window was studied. The paper reviewed the effect of axon growth time on network topological changes, focusing on i) connection probability, ii) number of connections, and iii) direction of connections. The authors offer a model of axon outgrowth, in which growth occurs in a straight line and in a random direction, and connections are established with other neurons within a specific/connectible range. The simulation result was then compared with the network connectivity pattern in *C. Elegans*.

The work suggested that neurons with different growth times have a tendency to hold more connections and long-range connections. In contrast, the same birth time window contributes to the establishment of reciprocal connections between two neurons with less connection probability.

IV. OVERVIEW OF SIMULATOR ARCHITECTURE

The research aim is to design a developmental trajectory model for the *C. Elegans* nervous system. For that purpose, one neural simulator is specially designed consisting of two main components: i) a network developer and ii) a neural simulator with s 2D environment. The network developer is designed to self-generate the nematode's neural network from a single cell according to a specific developmental trajectory defined by a set of rules. Although the main objective of this research work is to accurately generate a network that shares similar morphology and physiology to the *C. Elegans* network, the result might be not the ideal case. Most worms with an artificially generated network might not accurately produce the same neural behaviour. Therefore, it is essential to have an evaluation method to analyse the interaction of the simulated worms with the environment. For this reason, the neural simulator was built as an evaluation tool with an embedded 2D virtual environment. A full diagram of the simulator is shown in figure 2. In the 2D environment neural simulator, two main modules are included – ANSWER and Neuroscape. ANSWER is a simple simulator designed to handle the flow of events between the worms and the environment. As an extension to that, Neuroscape encapsulates the 2D environment and the activity of a set of worms. The user initially defines specific conditions in the environment – food sources, temperature gradient and obstacles. All these conditions are set by the user using simple a ASCII description decrypted by the *Command Reader* class. Additionally, the simulator is designed handle multiple simulation sessions, for instance, to simulate the behaviour of different development stages.

A. Simulator Components

The Worm

Multiple worm objects in Neuroscape are held by a nest class, where each worm object has a specific ID, which the user can define with any *string* type name. The worm is constructed from two main modules: i) the network that defines the topology, and ii) the locomotion class to update the position of the nematode.

Network Connectivity

The network topology of worm is defined as a *pdigraph* class, constructed from different types of node/device objects. Currently, *neuron, source, environment sensor, and motor output* devices are available in this simulator; the behaviour of each node can be defined using various models in the *model* class.

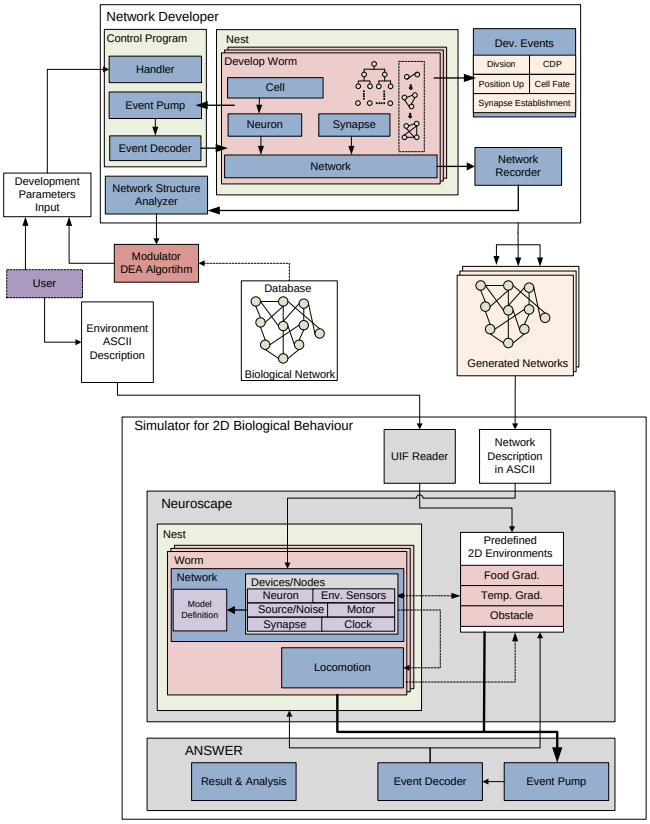

Fig. 2.Simulators Infrastructure. (Above is the simulator to generate network topology according to neural development trajetory algrotihms; below is the simulator for simulating the behaviour of the generated network in 2D virtual environment)

Defining the Environment

In Neuroscape, three different stimuli in the virtual environment can be pre-defined by the user using different, which are specified within the environment file before initiate the simulation. Each of these stimuli can be defined in multiple location within the 2D virtual environment. For food sources, the distribution is defines using following equation; where, variable H is for the food's maximum height, r_0 for distribution radius and n for exponential decay.

$$f(r) = \begin{cases} H(\frac{r_0}{r})^n & r > r_0 \\ H & r \leq r_0 \end{cases} \quad (1)$$

Temperature distribution is added using the `temp` command and currently only supports two distribution types: *linear* and *circular*. The user can pre-set the cultivation temperature, T_C, minimum-maximum temperature, the steepness and the distribution type before initiate the simulation. Additionally, the users can define spatial x-y area of obstacles which changes the environment surface at a specific location, where the *soil gradient* varies randomly between 0~1 in exponential distribution.

Interaction with the Worm

The environment is a separate entity in Neuroscape. Thus, a special sensor is needed to interact with the environment during the run time.

These environmental sensors are one part of the *pdigraph* network and directly provide inputs to the corresponding sensory neuron in the nematode network. For worm movement, a special device called *motor* is embedded in the *pdigraph* network to process the outgoing signal from the nematode network. This *motor* device has a fixed interval for internal state update, which consequently moves the nematode to a new position. During this *event*, the environmental sensor is called by the simulator to obtain the current environment status at the new nematode position.

B. ANSWER Module – Control Event Flow

In ANSWER, the state of any device is updated either at a fixed interval or upon arrival of an event. This creates a balance between flexibility and computational speed, giving the user freedom to define any specific model and reduce unnecessary processing. For instance, the event from the *clock* device is pumped at a fixed interval if the device is included in the network. The *clock event* is later called back by the decoder and processes the routines to update the internal state of the neurons. Currently, only the neuron device is included in this routine, although any device can easily be added later. ANSWER also gives the option for the neurons to update their state during an event – for example, if a neuron receives a spike from a connected neuron. Additionally, the *neuron* device is designed to pump the event if the internal state changes past a certain threshold and fires a spike. This event is held in the event pump at specific time delay before been sent to the connected neuron(s). The *source* device self-generates the event, with the event interval dependent on the definition given by the user. The *source* device is used as a generator pulse or as a noise signal that affects the internal state of the connected neuron(s). Figure 3 illustrates the mechanism of the event flow process in ANSWER, which is controlled by two modules: i) *Event Pump* and ii) *Event Decoder*.

V. THE DEVELOPER – DESIGNING THE DEVELOPMENT TRAJECTORY MODEL

The main idea behind this project is to develop a model with the capability of self-creating the neural network according to set of development trajectory rules that approximately fit the biological process in *C. Elegans.* The nervous system development process in *C. Elegans* consists of a complex process, for which knowledge is still limited. This section discusses the initial idea of the "Developing Network" model. The first part explains the workflow process for designing and verifying development rules. In the second part, the algorithm is explained in detail. Nevertheless, our incomplete knowledge regarding the biological development process in *C. Elegans* might hinder an accurate definition of development rules and process model.

A. Defining Development Trajectory Rules

In general, four research stages are required in order to define and evaluate the accuracy of the development trajectory rules, as illustrated in figure 4. In the first stage shown in figure 4(a), the network is self-generated by the *Developer*, which has been designed i) to generate thousands of cells from a single cell through a repeated division process in which the timing of each division event is defined by specific rules; and ii) to specify neuron type and establish the connections. As we are only interested in the development of the nervous system, the *Developer* was designed only to classify generated cells either as *neurons* or *muscle cells*, although other cell type information is stored, as it is essential for the calculation of dynamic cell movement during the developmental process.

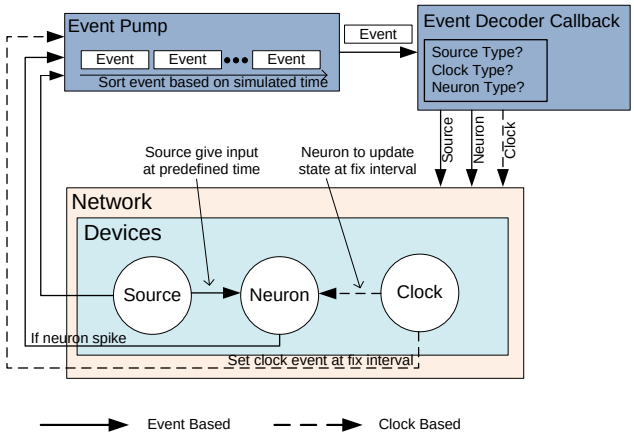

Fig. 3.Event flow in the Neuroscape simulator. (Event flow is controlled by *Event Pump* class. Each time the devices generate the event, it will store at the *Event Pump* until it callback at a specific time)

The structure of the generated network is then validated using **Network Structural Analysis**, which consists of three different criteria described below. The calculated value for each criterion is compared with the result published in [12], where same analysis was done on the biological network of *C. Elegans.* These criteria are used as a reference to generate new adjusted parameter values using differential evolution Algorithms (DEAs). It might be necessary to repeat the procedure hundreds to thousands of times to find the optimal value. It is also possible that the model may not be capable of generating an accurate network structure only using the optimization method, and thus the validation of each rule is essential as a prior condition.

Analysis 1- Small world index

Small world index is defined as the ratio of the clustering coefficient, C, to characteristic path length, L. The clustering coefficient indicates how well information is distributed locally; it can be calculated using the proportion of connections actually present out of all possible connections among network nodes directly connected to a node.

$$\sigma_{SW} = (C/L)/(C_{Rand}/L_{Rand}) = \mathbf{5.37} \qquad (2)$$

It takes the average value from all individual neurons in the network. Characteristic path length shows the level of external influence on the local network; it is mathematically defined as the average number of connections that have to be passed on the shortest paths between all pairs of network nodes. To get the small world index, this value can be compared with benchmark network (random network).

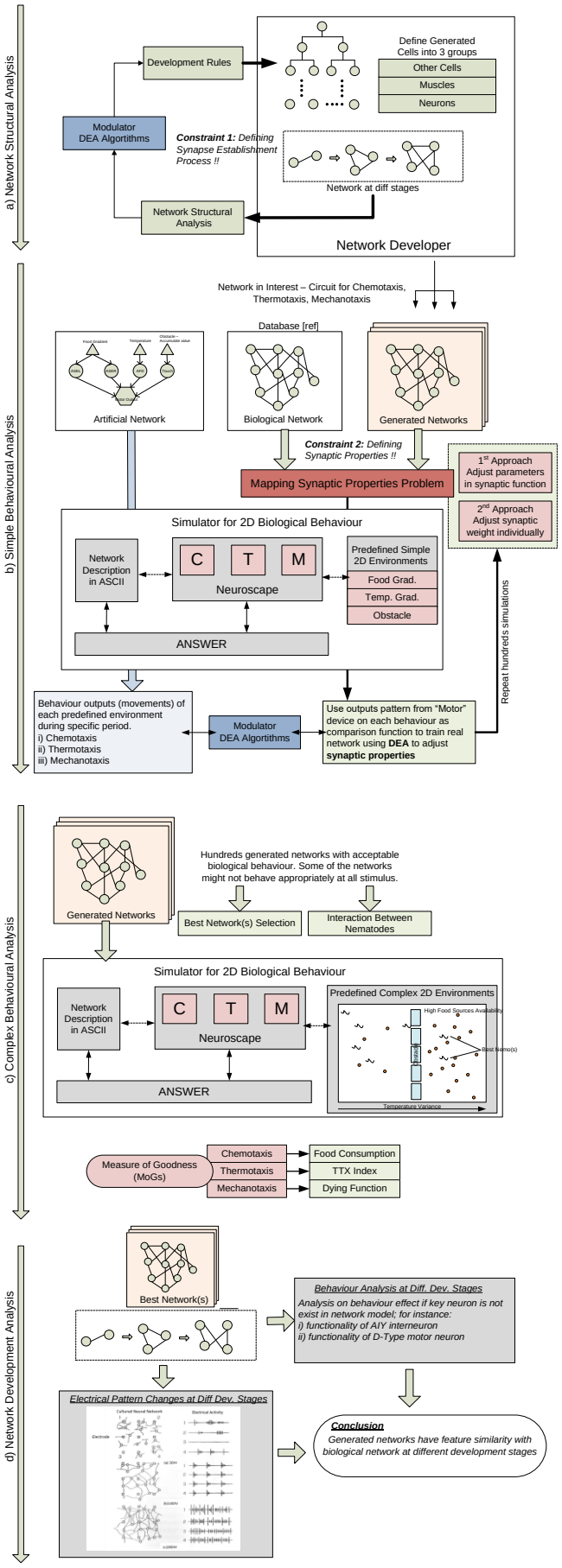

Fig. 4.Proposed Framework for Designing the Development Trajectory Model (4 different analysis is suggested a) Network Structural Analysis, b) Simple Behavioural Analysis, c) Complex Behavioural Analysis, d) Network Development Analysis)

Analysis 2 - Network Dispersion

This analysis measures the degree of synapse distribution of the individual neuron to different modules of the network. The mathematical definition is as follows, where R_i is the number of modules connected to each individual neuron. For the *C. Elegans* network, R = 10 modules and N = 279 neurons.

$$D = \sum \frac{D_i}{N} = \mathbf{0.46} \; ; where \; D_i = \frac{R_i}{R} \qquad (3)$$

Analysis 3 - Network Modularity

This measures the degree of modularity of an unknown network by calculating the fraction of the links connecting two nodes within one module and its deviation from the case where the links are distributed at random. The equation is defined as follows, where A_{ij} is elements of the adjacency matrix, k_i is the number of connections, or the degree of node i, and δ is Kronecker's delta function.

$$Q = \frac{1}{2E}\sum_{ij} A_{ij} - \frac{k_i k_j}{2E}\delta(q_i, q_j) = \mathbf{0.15} \qquad (4)$$

From the **Network Structural Analysis**, the parameters governing development rules are optimized; hence, it is assumed that the generated network approximates a biological network – at least in terms of its structural organization. However, the behaviour of the generated network might differ from that of the biological network for two reasons: i) the network connectivity is not completely similar, and ii) differences in synapse properties. These constraints are expected due to our limited knowledge of biological development processes, especially the synapse establishment process, preventing us from accurately modelling the network connections. Additionally, there is the limitation of determining the synaptic properties (usually the synaptic weight) between neurons in the network model. The current approach, as reported in [13] and [14], uses the optimization technique to adjust the synaptic weight to generate the desired output. The question here is whether this really maps the synaptic properties of a biological network? Currently, besides the availability of the full map of the *C. Elegans* connectome, the only available data is the synapse number between two connected neurons and the neurotransmitter type produced by each neuron. The neurotransmitter type might be used as an indicator to determine the connection sign, either excitatory or inhibitory. However, is this information sufficient to map the synaptic weight in the network model? Or does the synaptic value need to be individually adjusted?

To determine the value synaptic weight value, two approaches are proposed (Figure 4(b)): i) model a synaptic function, where weight is determined mainly from synapse number; or ii) adjust the synaptic weight individually. Both methods require an optimization technique in which DEA (one of the genetic evolution techniques) is used to adjust either the parameters of synaptic transfer function or synaptic weight.

The former approach is preferable, as the synaptic properties of generated networks can be automatically mapped, although the lack of currently available information might prevent us from accurately defining the synaptic transfer function. This optimization technique requires a prerequisite step – finding the correlation between the synapse properties of the biological network and the synaptic weight in the network model. For this problem, the output behaviour from the artificial network is used as a reference from which to modulate the parameters in the synaptic transfer function. By using **Simple Behavioural Analysis**, the synaptic properties of generated networks are determined, and the networks that are not capable of producing chemotaxis, thermotaxis and mechanotaxis behaviour are filtered and removed.

Under **Simple Behavioural Analysis**, generated networks with acceptable biological behaviour are selected. However, some of the networks might not respond accurately to all type of stimuli (food, temperature or obstacles). Therefore, further analysis is necessary to sort out the best networks with reliable biological behaviour, as shown in figure 4(c). In the figure, **Complex Behavioural Analysis"** is performed whereby all the networks are placed within same complex 2D environment – with the coexistence of multiple food sources, temperature gradients and selected areas with a high soil gradient. For network evaluation, three indexes are introduced to measure the goodness of the networks and to analyse the network response to multiple stimuli.

The question becomes does the selected network have feature similarity with biological network during the development process? Or is it alike only at end product? To answer this, **Network Developmental Analysis** (Figure 4(d)) is carried out to analyse network features during the development process and to verify that the development trajectory model is applicable at different network stages. The analysis aims focuses on two aspects: i) changes in electrical activity patterns and ii) behaviour changes at different development stages. For the former, the following can by analysed:

- the effect of the refractory period on the activity of the selected neuron when the number of synapses is increase during the development process.
- the role of inhibitory neurons in later stages of the development process
- depletion of neurotransmitter

In analysing biological behaviour changes during the development process, the functionality of existing of key neurons is observed by comparing the worm's behaviour changes with experimental data. An instance of this is as follows:

- functionality of AIY interneuron
- functionality of D-Type motor neuron

B. Rules and Algorithms Underlying the Development Process

An Overview

The nervous system development process involves multiple important events that are governed by specific rules or algorithms. In order to design the model, two important factors need to be considered:

i) **Mechanism of repeated division** – needed to generate thousands of cells from a single parent cell, where after each division;
   - the daughter cell can either become i) a terminal cell (death or specified into one of three different cell types); or ii) a non-terminal cell (will undergo cell division at a specific time in the future).
   - division timing varies among cells, as determined by factor(s) from the parent cell.
   - each time a new cell is created, the overall position of each cell within the embryo is affected.

ii) **Mechanism of synapse establishment** – *C. Elegans* has invariant connectivity, and generated networks should
   - have similar network organization to the biological network.
   - be capable of establishing the different types of synapses.

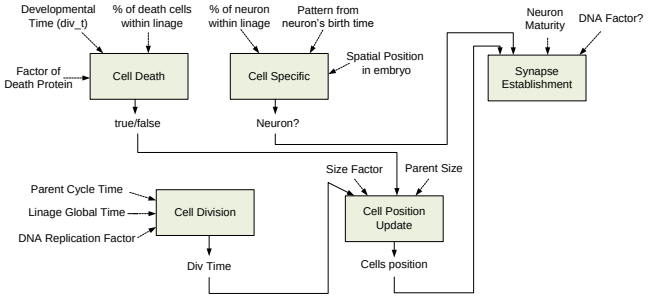

Fig. 5.Input-Output Relation between Event Class. (All event class is interaleted with each others – the output(s) from each event have significant effect on other event)

The flowchart in figure 6 shows the programming flow for network generation according to the development rules. The green box indicates a class event that holds at least one rule specified by mathematical definition(s) or algorithm(s). Details on the approach used to design the mathematical definition(s) are given in the following section. At the end of the process, the identity of each node or neuron in the generated network is identified according to its cell division history or cell lineage via the *C. Elegans* neuron database. The development process begins with the initialization of the first event, in which the parent cell is created and the parameters regulating the developmental rules are initialized; these parameters are initiated by the user and automatically updated by the modulator according to the **goodness** of previous network structure. The objective here is to repeatedly generate the network based on pre-defined developmental rules and optimize these parameters until the *Developer* is capable of generating networks structurally similar to the biological network. During the initialization process, the parent cell gets its first division time from the *Division Time* Class; the time solely depends on an outside parameter set by the user that indicates the overall embryo development speed.

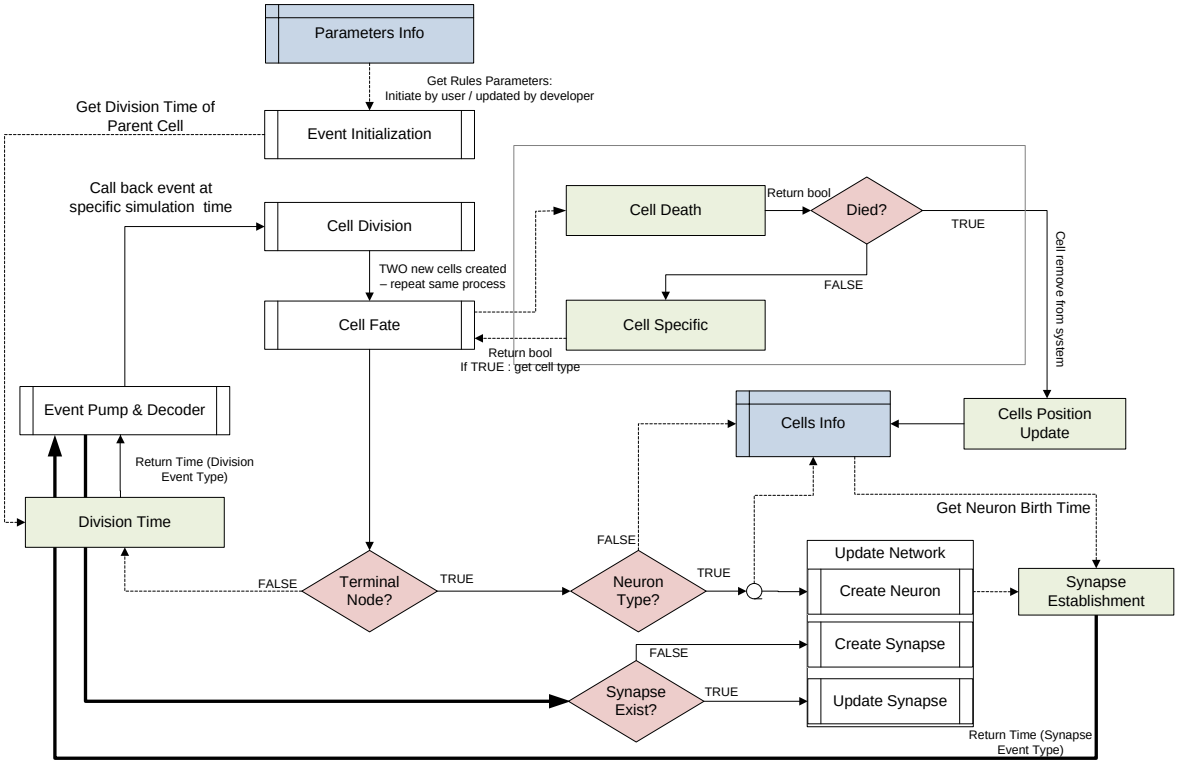

Fig. 6.Relation between Events and Flow of Developmental Process (Neural development trajectory consists of 2 important stages – cell division and synapses establishment. Main event in Cell Division process is *Division Time, Cell Specific, Cell Death and Cell Position Update*. Meanwhile in Synapses Establishment process consists only *Synapse Establishment* event)

The *Division Time* Class is designed to formulate the division time of each new cell created by the initialization / division process, but only if the cell is specified as a non-terminal cell. This class returns a specific simulation time to be used by the *Developer* to store the new division event in the memory to be called back in the future. Upon each cell division event, two new cells are created and undergo the *Cell Fate* process, where each cell is specified as a non-terminal or terminal cell. A terminal cell is defined if the cell is not destined for death or to become any cell type. In this process, two event classes – *Cell Death* and *Cell Specific* – are called in sequence. The *Cell Death* class return *Bool* type, where if the cell is calculated to die, the *Developer* removes it from system and a new cell position is re-calculated using the *Cells Position Update* class. For the other case, the *Cell Specific* class is called to determine the cell type. If cell type is returned *false*, the cell is retained as a non-terminal node and a new division event is created. In contrast, a terminal node cell specified as a neuron type will go through a network generation stage via the *Update Network* process. Three processes/functions are included to engage with different situations. If a new cell is created, then the *Create Neuron* process is called and the *Developer* calculates the probability that this cell will establish the synapse with existing neuron(s) via the *Synapse Establishment* class. If it returns *true*, the new synapse event is created and called back at a specified simulation time. Two neurons might already have established a connection between them, and so for this situation the feature of established synapse is updated – one feasible method is to change the synapse number.

Designing Event Rules and Algorithms

The main problem of modelling the biological process of *C. Elegans* is that it does not happen stochastically – it is somehow controlled by certain factors. However, designing a model solely based on these biological factors is almost impossible due to our limited knowledge. Nevertheless, as development process in *C. Elegans* is invariant, there development patterns can be generated using statistical data obtained from observations of real nematodes. For instance, number of cell deaths during the developmental period [4], birth time pattern during the embryonic stage [10], and percentage of cell deaths and number of each cell type within each lineage [15].

This pattern could surely be defined mathematically by a set of equations or algorithms, thus giving an approximate number of events and allowing the *Developer* to explicitly control the number of specific events happening within in a certain developmental period. The question here is whether it is sufficient to accurately model the development trajectory? Certainly, the individual cell also has a significant influence on the output of each event. For instance, the existence of a specific protein in an individual cell can determine the occurrence of an event (e.g., the egl-1 protein is the main factor determining cell death). Additionally, the cell-cell signalling factor reported in [16][17] needs to be included in the model, as it represents the influence of an individual cell on the development trajectory; the best approach to address this problem, however, is yet to be identified. In figure 5, the possible relation between inputs and output of each event class is shown.

VI. DISCUSSION

The development of a self-generating network model requires four essential steps and a prior analysis of the developmental process in *C. Elegans.* Generally, the event in the model is assumed to be a black box, with a set of functions/algorithms that receive certain inputs (biological factors) and produce specific output(s). For that purpose, it is vital to statistically analyse each development event (cell division time, cell death, etc.) to find the correlation between these factors (or between events). Additionally, each event might have a set of rules, with the selection of these rules dependent on certain criteria (e.g., development time, lineage, etc.); this is similar to the *Boolean* selector, which selects the rule that is applicable to a certain situation. Each process of designing development rules is essential in order to answer the research question: **Step 1*:*** Adjustment of unknown variables of development rules by comparing the values from structural analysis. **Step 2*:*** Since the development rules include stochastic and deterministic processes, the networks are not identical. Thus, it is essential to filter the generated networks that have acceptable biological behaviour in response to different environment stimuli (food, temperature and obstacle). The idea is to use the connection of a real biological network (selected sub-network) and adjust its synaptic properties to behave exactly as observed in the artificial network. **Step 3*:*** Finding the networks most capable of handling a predefined complex environment within a specific simulation time. **Step 4*:*** Further analysis of the best networks in order to answer the main research question, which seeks to understand the features of biological networks during the development process.

VII. CONCLUSION

Indeed, *C. Elegans* is an important subject to study the mechanism of the neural system because of the simplicity of its architecture and invariant developmental process. By understanding the trajectory pattern of the developmental process which consists of cell division and synapse establishment, it is possible to define it using the mathematical model. This work proposed a framework to develop the model to that extend of currently available knowledge on *C. Elegans* nervous system. By having a simulator that can artificially generate the neural network, it should give new insight to the understanding of the current nervous system. Notwithstanding, this deign framework in future should be updated if there is a new knowledge gained from experimental studies.

ACKNOWLEDGMENT

The author would like to thank the Universiti Teknikal Malaysia Melaka (UTeM) and Centre for Telecommunication Research and Innovation (CeTRI) for supporting this research.

REFERENCES

1. J. E. Sulston, E. Schierenberg, J. G. White, and J. N. Thomson, “The embryonic cell lineage of the nematode Caenorhabditis elegans,” *Dev. Biol.*, vol. 100, no. 1, pp. 64–119, Nov. 1983.
2. N. Tavernier, J. C. Labb??, and L. Pintard, “Cell cycle timing regulation during asynchronous divisions of the early C. elegans embryo,” *Exp. Cell Res.*, vol. 337, no. 2, pp. 243–248, 2015.
3. Z. Bao, Z. Zhao, T. J. Boyle, J. I. Murray, and R. H. Waterston, “Control of cell cycle timing during C. elegans embryogenesis,” *Dev. Biol.*, vol. 318, no. 1, pp. 65–72, 2008.
4. G. Lettre and M. O. Hengartner, “Developmental apoptosis in C. elegans: a complex CEDnario.,” *Nat. Rev. Mol. Cell Biol.*, vol. 7, no. 2, pp. 97–108, 2006.
5. J. S. Isaacson and M. Scanziani, “How inhibition shapes cortical activity,” *Neuron*, vol. 72, no. 2, pp. 231–243, 2011.
6. N. Geard and J. Wiles, “A gene network model for developing cell lineages,” *Artif. Life*, vol. 11, no. 3, pp. 249–267, 2005.
7. J. ??ke Larsson, N. Wadstr??mer, O. Hermanson, U. Lendahl, and R. Forchheimer, “Modelling cell lineage using a meta-Boolean tree model with a relation to gene regulatory networks,” *J. Theor. Biol.*, vol. 268, no. 1, pp. 62–76, 2011.
8. J. Naish, “Developmental Motifs Reveal Complex Structure in Cell Lineages,” *Nurs. Manag. (Harrow).*, vol. 2, no. 1, p. 3, 2010.
9. R. W. Taylor, Y.-W. Hsieh, J. T. Gamse, and C.-F. Chuang, “Making a difference together: reciprocal interactions in C. elegans and zebrafish asymmetric neural development,” *Development*, vol. 137, no. 5, pp. 681–691, 2010.
10. S. Varier and M. Kaiser, “Neural development features: Spatio-temporal development of the Caenorhabditis elegans neuronal network,” *PLoS Comput. Biol.*, vol. 7, no. 1, 2011.
11. S. Lim and M. Kaiser, “Developmental time windows for axon growth influence neuronal network topology,” *Biol. Cybern.*, vol. 109, no. 2, pp. 275–286, 2015.
12. J. S. Kim, M. Kaiser, and M. Kaiser, “From Caenorhabditis elegans to the human connectome : a specific modular organization increases metabolic , functional and developmental efficiency,” 2014.
13. M. Suzuki, T. Tsuji, and H. Ohtake, “A model of motor control of the nematode C. elegans with neuronal circuits,” *Artif. Intell. Med.*, vol. 35, no. 1–2, pp. 75–86, 2005.
14. X. Deng and J. X. Xu, “A 3D undulatory locomotion model inspired by C. elegans through DNN approach,” *Neurocomputing*, vol. 131, pp. 248–264, 2014.
15. M. M. Metzstein, G. M. Stanfield, H. R. Horvitz, and V. M. Dixit, “Genetics of programmed cell death in C. elegans: past, present and future.,” *Trends Genet.*, vol. 14, no. 10, pp. 410–6, Oct. 1998.
16. L. T. Macneil and A. J. M. Walhout, “Gene regulatory networks and the role of robustness and stochasticity in the control of gene expression,” pp. 645–657, 2011.
17. Y. Arata, J.-Y. Lee, B. Goldstein, and H. Sawa, “Extracellular control of PAR protein localization during asymmetric cell division in the C. elegans embryo,” *Development*, vol. 137, no. 19, pp. 3337–3345, Oct. 2010.

AUTHORS PROFILE

Raihaan Kamarudin received the BE and ME degrees from Takushoku University, Japan. He is currently an academician at Faculty of Electronic and Computer Engineering, Universiti Teknikal Malaysia Melaka, Malaysia and working toward the PhD degree in electronics and computer science.

M. N. Shah Zainudin is an academician from Faculty of Electronic and Computer Engineering, Universiti Teknikal Malaysia Melaka, Malaysia. His research interest is in broadly areas of machine learning application, data mining and pattern recognition.

R.H. Ramlee is an academician from Faculty of Electronic and Computer Engineering, Universiti Teknikal Malaysia Melaka, Malaysia. His research interest is in broadly areas of FPGA applications and embedded system.

M.I. Idris is an academician from Faculty of Electronic and Computer Engineering, Universiti Teknikal Malaysia Melaka, Malaysia. His research interest is in broadly areas of microelectronic and VLSI design.