

A Distributed Smart Routing Scheme for Terrestrial Sensor Networks with Hybrid Neural Rough Sets

Frank Jiang and Michael Frater
Department of Engineering and IT
The University of New South Wales
Email: Frank.Jiang@unsw.edu.au

Steve S.H. Ling
Faculty of Engineering and IT
University of Technology, Sydney.
Email: Steve.Ling@uts.edu.au

Abstract— The limited power consumption, as a major constraint, presents challenges in improving the network throughput for Wireless Sensor Networks (WSNs). Due to the limited computational power, the applications of WSNs in Terrestrial Networks require the capability to pre-process the observation data so as to remove irrelevant features or factors from multi-dimensional dataset. This paper proposes a intelligent distributed energy efficient routing algorithm inspired from natural learning and adaptation process with the aid of hybrid Neural Rough Sets theory, which is used to efficiently reduce the dimensionality of input dataset. The algorithmic implementation and experimental validation are described in this paper. Details of the algorithm and its testing procedures are presented in comparison with the other power-aware protocols, e.g., mini-hop. The validation of the proposed model is carried out via a wireless sensor network test-bed implemented in Castalia Simulator. The experimental results show the network performance measurements such as delay, throughput and packet loss that have been greatly improved as the outcome of applying this integration with Neural Rough Sets.

Keywords—*Neural Rough Set; Wireless Sensor Networks (WSNs); Biologically Inspired System; Ant Colony Optimisation (ACO); Swarm Intelligence (SW)*

I. INTRODUCTION

Wireless Sensor Networks (WSNs) have drawn considerable attention recently in both terrestrial and underwater network research community [1-2], [14-15]. Specifically, WSNs consist of a number of sensor devices (MOTES), which are normally inexpensive, portable, having communication and limited computational power. It potentially enlists our daily life with such characteristics as required by pervasive computing, ubiquitous communication and monitoring, which enables to establish on-demand interactions with data devices, services and applications without spacial and time limitations. Portable wireless communication devices, microprocessors as well as communication protocols and wireless techniques including WiMAX all contribute to the construction of WSNs.

The MOTES are interconnected to each other via the network grid. Routing schemes play an important role in maintaining efficient data communications across this network grid. Whereas, routing algorithms are required to consider the constraints of MOTES such as limited energy supply, limited computational power. Huge amount of data are produced to be

processed in this regard, it is desirable to have the reduced data dimensions in a pre-processing stage. In this paper, we explore the possibility of using a new Neural Rough Set process to fulfill this purpose.

Current and future computing environments will be of massive scale and complexity. They have heterogeneous products, services, devices and platforms of ever growing scale. Intensive research has recently gone into developing a robust distributed smart algorithms and protocols for decentralised cases.

At the meantime, there are a number of routing protocols that have been proposed so far, such as Low Energy Adaptive Clustering Hierarchy (LEACH), Data-centric SPIN and Direct diffusion. LEACH relies heavily on the regular routing structure, which cannot flexibly adapt to changeable network environment. While the Direct diffusion does not build up a regular routing structure and SPIN even does not maintain a structural hierarchy at all, even no state information. These protocols cannot satisfy the practical requirements with these disadvantages under various circumstances. In fact, there are two major aspects for a routing scheme to consider (1) energy efficiency and (2) self-healing. Based on this consideration, a *self-governed* bio-inspired routing scheme is proposed in this paper. This algorithm is designed to solve the *packet* routing problem with minimal energy consumption for each hop communication and aims to maximise the lifetime of nodes in the network. Meanwhile, the new self-healing algorithm well reduce the possibility of node failures and satisfy the need for various levels of dynamic packets routing when routing decisions can be made locally at each node.

The remainder of the paper is organized as follows: Section II reviews the background of Neural Rough Set theory and the new bio-inspired scheme initiated by the swarm behavior. Section III presents the new routing algorithm details including algorithmic steps, and its implementation details. Simulation results of the new algorithm are presented in Section IV. As a validation experiment, the performance analysis shows the efficiency of our proposed routing scheme that potentially to small sized networks in WSNs. Finally, we conclude the contributions of this paper and look into the future work.

II. PROPOSED NEURAL ROUGH SETS AND ANT COLONY BIOLOGICAL MODEL

A. Related Rough Set Preliminaries

The rough set theory is an extension of set theory for the study of intelligent system with indiscernible, ambiguous and imperfect data information. Rough-set theory appeared two decades ago. Its main thrust is used in attribute reduction, rule-generation and prediction. For instance, by applying a mathematical approach to solve the data reduction problem, it is good at finding out the discernity, vagueness and uncertainty in data analysis. The rough-set concept was introduced by Pawlak and his colleagues [16]. This theory became very popular among scientists around the world in the last decade, and it is now one of the most developed AI methods. The primary goal of this new emerging theory is to classify the information from incomplete information. Its successful application to a variety of problems has shown its usefulness and versatility [18].

As we know, the decision table in a rough set scheme plays an important role in clarifying a discretized feature values. Applying rough-neural bundled computation has proved to be more efficient in discovering emergence and hidden order of control area as an approach to knowledge discovery and data mining. The integration of such an approach with telecommunication complex scenario, the goal of having a more robust internet tolerating the nonstationary anomalies can be achieved. Below are some definitions and domain descriptions [18].

The information system is The information system is formed as a pair of $IS=(U, A, V, f)$, where U is a *non-empty* finite set of objects (formed as an universe), and A is a set of finite attributes, V is the domain limitation of A , and $f:U \times A \rightarrow V$ is an information function. $\forall a \in A, x \in U, f(x,a) \in V$. Each non-empty subset $B \subseteq A$, Let $H \subseteq U \times U$ denote an indiscernibility relation on U that can be defined as:

$$H = \{(x,y) \in U \times U : b(x) = b(y), \forall b \in B\} \quad (1)$$

Objects x and y satisfying the relationship H are indiscernible by attributes from B .

Definition 1 Given an information system S , let $B \subseteq A$, and $x(i), x(j) \in U$, a binary relation $[x]_B \triangleq \{(x(i), x(j)) \in U \times U / \forall a \in B, f(x(i), a) = f(x(j), a)\}$ is an indiscernibility relationship.

Let U/H denote the equivalent class of H . For any $x \in U$, we denote the set, $[x]_B \triangleq \{(x(i), x(j)) \in U \times U / \forall a \in B, f(x(i), a) = f(x(j), a)\}$ is called the equivalent class in B . The empty set $\emptyset$ and the elements in U/H are called elementary sets. Any subsets $X \subseteq U$ from the given data can be approximated by the elementary sets. Let $X \subseteq U$, We define two sets $SLB(X)$ and $SHB(X)$ as the *lower approximation* and *upper approximation* of the set of objects X as:

$$SLB(X) = \{x \in U : [x]_B \subseteq X\} \quad (2)$$

$$SHB(X) = \{x \in U : [x]_B \cap X \neq \emptyset\} \quad (3)$$

where $[x]_B$ denotes the set of all objects indiscernible with x by attributes from B . In Pawlak's definition of indiscernibility relations, for $x, y \in U$, it is called that y is indiscernible to x with respect to $B \subseteq A$ if $f(x,a) = f(y,a), \forall a \in B$. Where $f(x,a)$ means a mapping of value x with respect to attribute a [18].

B. Neural Rough Sets and BP Training

The concept of lower and upper approximation has been applied in many applications of computational intelligence. Successive research shows that the general concepts of lower and upper bound offer a broader application field in various types of industrial fields. In particular, rough set theory has demonstrated the power of using these bounds in rule generation [18].

Although there are plenty of classifications of recurrent neural networks, Zhang, et al [19] make a good performance comparison among the main algorithm i.e. BP, Perceptron-Back propagation Hybrid (PBH), Fuzzy ARTmap and Radial-Basis Function Network (RBF), their results reveal that BP and PBH nets outperform Perceptron, Fuzzy ARTmap and RBF by the criteria of classification errors. Hence we will apply BP nets in our design. The input datasets with all features are filtered by rough-sets logic and subsequently the models based on the minimum feature sets are built with the integration of rough neurons into Neural Rough networks. In this part, our network architecture is Ninput : Nhidden:1 three layer feed-forward rough neural network for processing. During this process, the error in the output neuron is propagated back from one layer to the previous layer by altering the weights of the connections, i.e. Back-Propagation (BP) algorithm. A rule of thumb, known as the Baum-Haussler rule [17], is used to determine the number of hidden neurons to be used:

$$N_{hidden} = \frac{N_{train} \times E_{tolerance}}{N_{pts} + N_{output}} \quad (4)$$

where N_{train} is the number of hidden neurons, $E_{tolerance}$ is the error tolerance, N_{pts} is the number of data points for training example, N_{output} is the number of output neurons, in our architecture, $N_{output} = 1$. In our case, for example, Let $E_{tolerance} = 0.001$; $N_{train} = 4000 \times N_{pts}$; then the above equation becomes:

$$N_{hidden} \leq \frac{4N_{pts}}{N_{pts} + 1} \quad (5)$$

Hence, the number of neurons in the hidden layer should be no more than 4. Now, we are ready to start the rough neural multilayer sets process to reduce the dimensionality of inputs datasets [18].

C. Proposed Bio-inspired Algorithm

We apply the metaphor of ant food foraging behavior into network routing system. As stated in the literature, ant colony optimization metaheuristic (ACO) was first systematically

introduced by Dorigo in 1992 in his PhD dissertation [6]. ACO is a probabilistic technique for solving computational problems which can be reduced to finding good paths through graphs. They are inspired by the behavior of ants in finding paths from the colony to food. The early attempts of biological researches on ants were conducted by Dorigo and Gambardella [7] for routing in wired and wireless mobile environments.

In addition, Stutzle and Hoos proposed MAX-MIN ant system, namely MMAS, it is an improvement version of ACO. MMAS differs from standard ACO in that (i) only the best ant adds pheromone trails, and (ii) the minimum and maximum values of the pheromone are explicitly limited [8]. Agassounon 9], [10] proposed interesting work based on ACO to solve information retrieval and resource allocation issues. A good review of available ACO basis, models and new trends can be found in Oscar's work [11].

ACO algorithms can be applied to find the near-optimal multihop routing path in mobile ad hoc networks, namely AntHocNet, the experimental results indicate AntHocNet algorithm outperforms the other popular Ad hoc On-Demand Distance Vector Routing (AODV) algorithm in terms of *end-to-end delay* and *delivery* ratio. Related problems such as load-balance and routing in computer networks are also investigated in the work by Hsiao et. al, by use of ACO [12]. In sum, performance comparisons between ACO-based algorithm and other routing algorithm show better performance by use of ACO metaheuristics.

The most important feature of ACO metaheuristic is inspired from the *stigmergy*, which mimics the indirect communication behaviors adopted by social insects. The proposed algorithm, termed iACO, especially for WSNs, each packet are regarded as an individual ant, communicating with each other via "pheromone" values stored in each wireless sensor node's pheromone table.

Three main steps are included in the flowchat of the proposed algorithms for the routing scheme of WSNs. In brief, the steps include (1) Generate local solutions based on the path (i,j) selection probability; (2) Pheromone update; (3) Decision making. Maintaining the Integrity of the Specifications

Figure 2 depicts the overall integration process where Rough set process and the proposed algorithm are attached together to solve the routing path in network environment.

III. ROUTING ALGORITHM IN WSNs

Suppose the ant is randomly placed on node i , the probability of an ant k choosing the next adjacent node j is as depicted in Equation 6. And the pheromone table helps to choose the next neighbor in the process.

$$P_{ij} = \begin{cases} \frac{[\tau(i,j)]^\alpha \cdot [\eta(j)]^\beta}{\sum_{u \notin M_k} [\tau(i,u)]^\alpha \cdot [\eta(u)]^\beta} & \text{if } j \notin M_k \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

where P_{ij} = probability of packet going from node i to j ,
 $\tau(i,j)$ = pheromone level in the table of node i for node j ,
 $\eta(j)$ = energy level of node j ,
 M_k = identifier of every visited node,
 α = parameter that denotes pheromone level,
 β = parameter that denotes energy level heuristics.

Two categories of ants are in use, they are termed as forward ants and backward ants. We adopted this concept from the relating work of ant algorithm routing such as T-ant, antNet [2] and EEABR [3]. We further develop the definitions as used in the literature. The forward ants are defined as the ants that are sent from the source. The forward ants will carry the information such as (1) the required sensor type, (2) the sensor region of interest, (3) the data rate and (4) the time period of the interest. More details can be seen in table I where both categories of ants are described in functions.

Parameters and variables formatted in the packet of the experiment are described here in Table I. The header and data are carried by each packet. The packet is sent from source node, taking the form as backward and forward ants

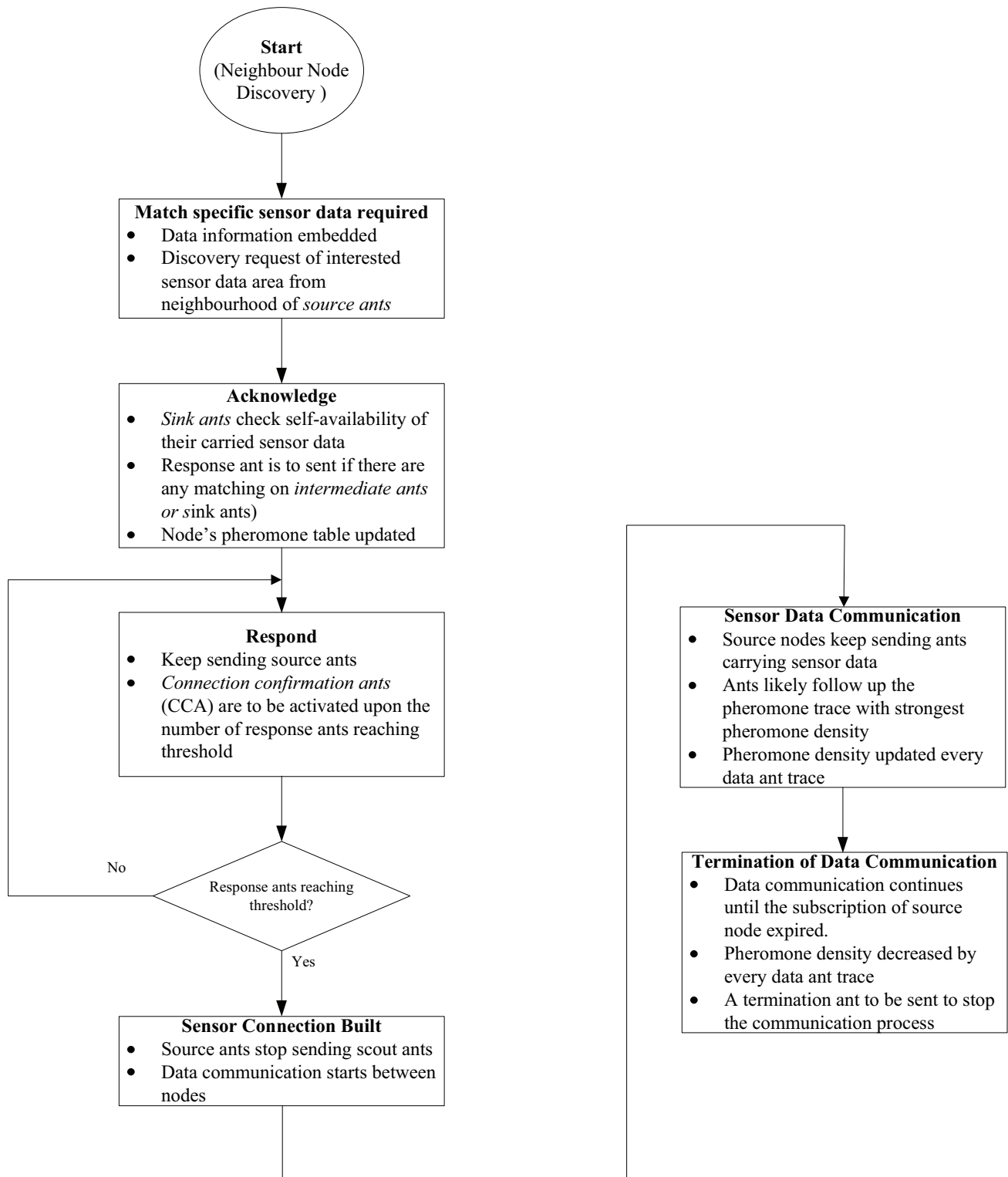


Figure 1. Flowchart of the proposed algorithms for the routing scheme of WSNs

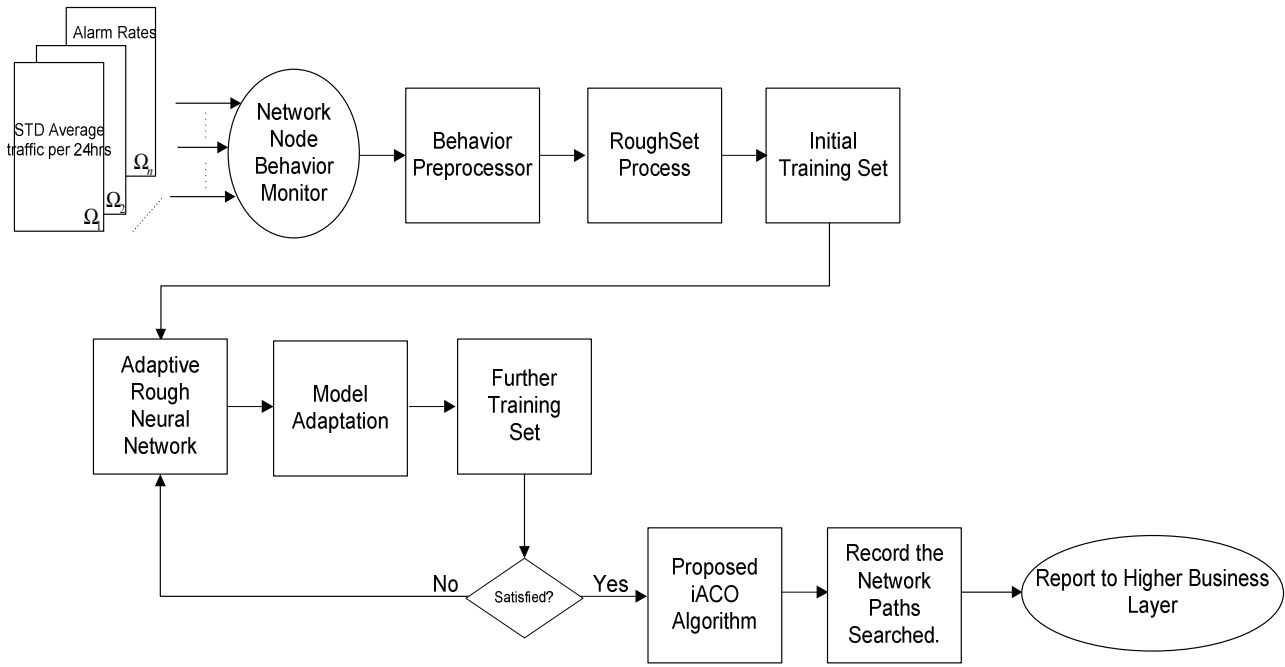


Figure 2. Illustration Schematic of The Overall Process

Pheromone Update:

Pheromone plays an important role in searching for a suitable node. In our algorithm, only the pheromone along our successfully recorded paths (usually more than one) is intensified, instead, pheromone in those incomplete paths evaporates. The pheromone increases and reductions will influence the searching process in next iteration. We adopt the conventional ACO pheromone update rule as [20]:

- Intensify pheromone: an absolute amount of pheromone $\Delta\tau_{i,j}^k > 0$ is added to the existing pheromone values on the paths completed by k -th ant

$$\Delta\tau_{i,j}^k = \begin{cases} \frac{Q}{Z_{\min(k)}} & \text{if } (i, j) \in T_k \\ 0 & \text{if } (i, j) \notin T_k \end{cases} \quad (7)$$

where T_k are the paths having been traversed by k -th ant. Q is a constant. More research on Q value have been done in Dorigo [6].

- Evaporate pheromone: pheromone on the edges of incomplete traversing paths is reduced in order to minimize the chances of being selected by other ants.

$$\tau_{i,j} = (1 - \rho) \times \tau_{i,j} \quad \forall i, j \in [1 : 52] \quad (8)$$

where $\rho \in [0, 1]$ is the pheromone evaporation rate. $\tau_{i,j}$ is the current pheromone density on the trail $a(i, j)$.

Mutate pheromone update process: In order to increase the convergence speed of the algorithm, we combine the *mutation* parameter into our pheromone update rules, where the concept of *threshold* is also involved. The mutation parameter in our algorithm are partly based on *Best-worst Ant Colony Algorithm* (BWACA) as well as *Max-min* Ant system.

IV. EXPERIMENT

A. Simulation Platform and Configuration Parameters

The simulation is conducted on the platform called Castalia network simulator [5]. Being a specially designed network simulator, Castalia contains models for realistic wireless channels, MAC layers, radio, physical sensor and resource management. Moreover, it contains modules that can simulate the different functionality and behaviours of wireless sensors and hence is regarded as a good simulation of the real behaviour in the research society for WSNs [4].

The dataset preprocessed by Neural Rough Set process is further processed in this simulation section.

The most commonly used algorithm for routing label witched Paths (LSPs) is the min-hop algorithm [13] The nature of this algorithm has been adopted by many routing protocols. Therefore, it is reasonable to compare our proposed algorithm with min-hop algorithm.

To evaluate two algorithms, we define the following parameter metrics in use in the simulation.

- 1) *Energy consumption*
- 2) *Delay*
- 3) *Throughput*

With regards to these metrics, Figure 3 and 4 depict the comparison view of *energy consumption* and *delay* exhibited by various node density levels. As shown, the small number of nodes would not make a significant impact on the energy consumption and delays.

When the interactive nodes increasing till 145, an energy peak occurs, and new routing algorithm proposed consumes less 1/3 energy than min-hop algorithm though the similar curve trend is observed. For the larger network, this difference tends to be more significant.

As shown in Figure 4, the impacts of delays for algorithms are trivial, though the proposed routing produces a slightly less delays than the min-hop. *Max-min* Ant system.

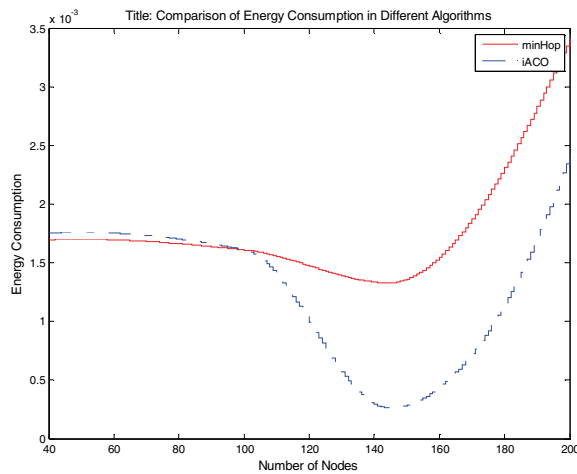


Figure 3 Comparison of Energy Consumption of Ant Population Levels

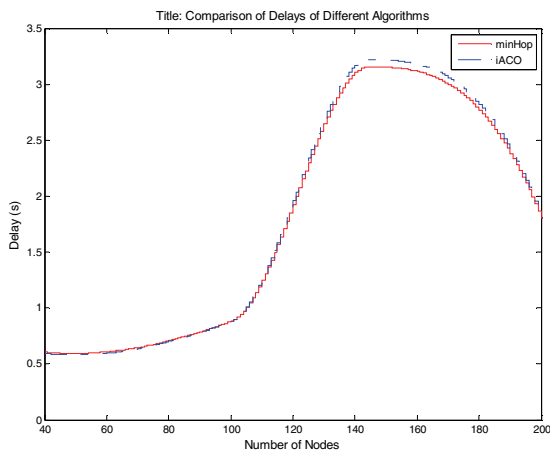


Figure 4 Comparison of Delays Between Two Algorithms

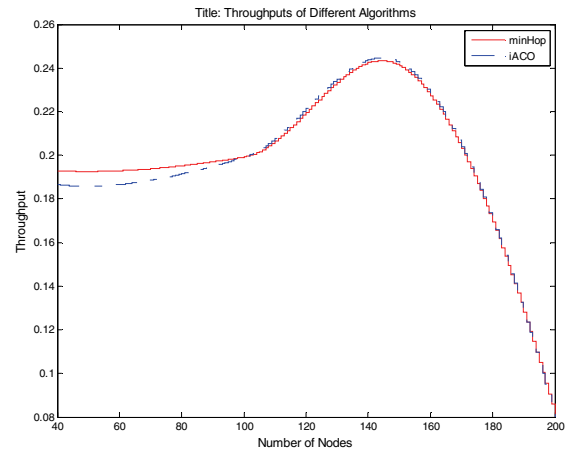


Figure 5 Comparison of Throughputs Between Two Algorithms

With regards to the throughput, as shown in Figure 5, when the number of nodes increase till 145, there are not much differences between min-hop algorithm and enhance ant-based protocol. The experience data and literature prove it is difficult for the traditional ACO achieve the marginal performance as min-hop routing scheme. The proposed protocol demonstrates the similar performance in this regard, so it can be regarded as a good improvement.

Figure 6, 7 and 8 shows the impacts of various node densities towards energy consumption, delays and throughput. It is shown there is an obvious effect on the efficiency of various node densities, but it seems not a linear relationship. There is a few optimal node densities at which the lowest energy consumption is observed. Similarly, it comes to the delay in Figure 7 and throughput in Figure 8. Specifically, the density is 0.005 when the minimum delay occurs, but the maximum delay occurs when the density is 0.1. This is also the case for throughput.

In particular, the minimum energy consumption, maximum

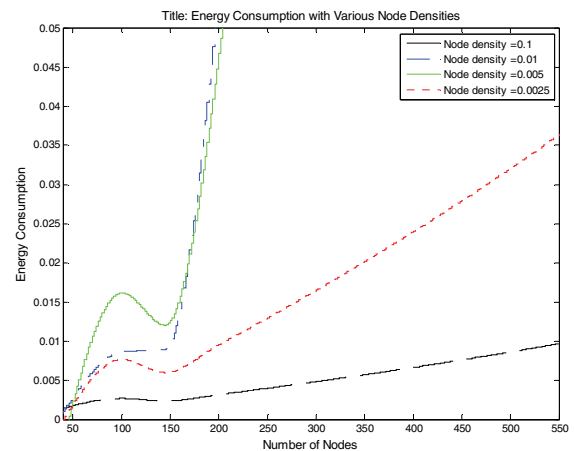


Figure 6 Energy Consumption with Various Node Densities

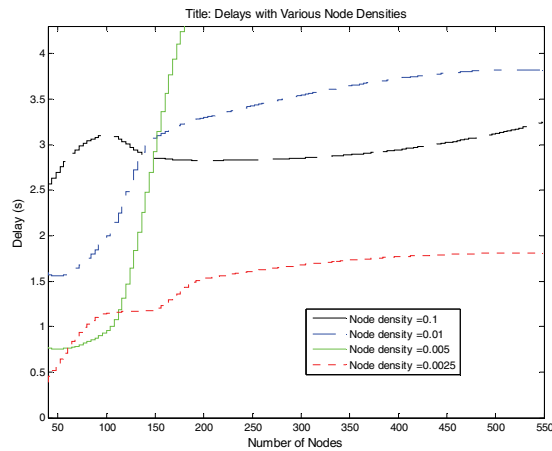


Figure 7 Delays with Various Node Densities

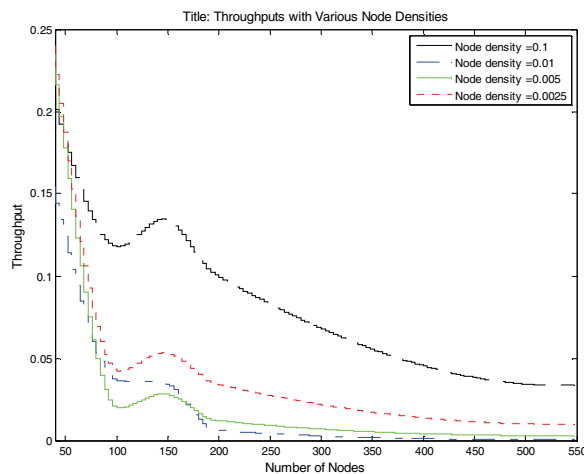


Figure 8 Throughputs with Various Node Densities

when the network has 50 nodes. The current good results are well compliance with the expected network size less than 100.

V. CONCLUSION AND FUTURE WORK

In this paper, a distributed, reliable nature-inspired routing algorithm for terrestrial sensor networks is presented. This routing process occurs after the dataset feature reduction via neural rough set theory. The swarm algorithm is partly based on the efficient *Max-Min* algorithm and it is suitable for flexible structure of wireless sensor networks. This new distributed scheme performs generally not worse than other standard routing algorithm, and in some occasions, it outperforms than conventional *min-hop* algorithm. More work in the future includes conducting further performance comparison work with other sensor network routing algorithms, as well as identifying and validating the best parameters configurations for the proposed algorithm under various network environment to guide real productive practices in the networking industry.

REFERENCES

- [1] S. Selvakkennedy, S. Sinnappan, and Yi Shang, "Data dissemination based on ant swarms for wireless sensor networks," in *Proc. Consumer Communications and Networking Conference*, 2006, pp. 132 – 136.
- [2] D. Caro and M. Dorigo, "AntNet: Distributed Stigmergetic Control for Communications Networks," *Journal of Artificial Intelligence Research*, vol. 9, pp. 317-365, 1998.
- [3] T. Camilo, C. Carreto, C. Silva, and J. S. Boavida, "An Energy-Efficient Ant-Based Routing Algorithm for Wireless Sensor Networks" *Lecture Notes In Computer Science*, pp. 49-59, 2006.
- [4] H. N. Pham, D. Pediaditakis, and B. Athanassios, "From Simulation to Real Deployments in WSN and Back," in *IEEE International Symposium on World of Wireless, Mobile and Multimedia Networks*, 2007, pp. 1-6.
- [5] NICTA. "Castalia: <http://castalia.npc.nicta.com.au/> [Feb. 6, 2010].
- [6] M. Dorigo, "Optimization, learning and natural algorithms (in italian)," Ph.D. dissertation, 1992.
- [7] R. Montemanni and L. Gambardella, "Swarm approach for a connectivity problem in wireless networks," in *Proc. IEEE Swarm Intelligence Symposium*, 2005, pp. 265 – 272.
- [8] M. Dorigo and T. Stuzle, *Ant Colony Optimization*. MIT Press, 2004.
- [9] R. Montemanni, L. M. Gambardella, "Exact algorithms for the minimum power symmetric connectivity problem in wireless networks," *Journal Computers and Operations Research*, vol. 32, issue 11, pp. 2891-2904, 2005.
- [10] W. Agassounon, "Distributed information retrieval and dissemination in swarm-based networks of mobile, autonomous agents," in *Proc. of IEEE Swarm Intelligence Symposium*, 2003, pp. 152–159.
- [11] O. Cordon, F. Herrera and T. Stuzle, "A review on the ant colony optimization metaheuristic: Basis, models and new trends," *Mathware and software computing*, vol. 30, no. 9, pp. 1 - 35, 2002.
- [12] Y. t. Hsiao, C. L. Chuang and C. C. Chien, "Computer network load-balancing and routing by ant colony optimization," in *12th IEEE International Conference on Networks*, 2004, pp. 313 - 318.
- [13] M. Zhu, M., Y. Wu, and F. Suili, "A New Dynamic Routing Algorithm Based on Minimum Interference in MPLS Networks," in *4th International Conference on Wireless Communications, Networking and Mobile Computing*, 2008, pp. 1-4.
- [14] M. Ismail and M.Y. Sanavullah, "Security topology in wireless sensor networks with routing optimization," in *4th International Conference on Wireless Communication and Sensor Networks*, 2008, pp. 7-15.
- [15] K. Yamasaki and T. Ohtsuki, "Design of energy-efficient wireless sensor networks with censoring, on-off, and censoring and on-off sensors based on mutual information," in *Vehicular Technology Conference*, vol. 2, 2005, pp. 1312 – 1316.
- [16] Z. Pawlak, "Why rough sets?," in *Proc. of the Fifth IEEE International Conference on Fuzzy Systems*, vol. 2, pp. 738 - 743, 1996.
- [17] E. Baum and D. Haussler, "What size net gives valid generalization?," *Neural Computation*, vol. 1, no. 1, pp. 151-160, 1989.
- [18] F. Jiang and R. Braun, "Intelligent Failure Domain Prediction in Complex Telecommunication Networks with Hybrid Rough Sets and Adaptive Neural Nets", in *International Information and Telecommunication Technologies Symposium*, pp. 45-52, 2004.
- [19] Z. Zhang., C. Manikopoulos, "Hide: a hierarchical network intrusion detection system using statistical pre-processing and neural network classification," in *Proc. the 2001 IEEE Workshop on information Assurance and security United States Military Academy*, 2001, pp. 85–90.
- [20] F. Jiang, and V. Mahadevan, "Towards the Distributed Autonomy in Complex Environments", in *Proc. International Conference on information and Multimedia Technology*, 2009, pp. 169-172.