Applied Soft Computing xxx (2017) xxx–xxx

Contents lists available at ScienceDirect

Applied Soft Computing

journal homepage: www.elsevier.com/locate/asoc

Energy efficient and QoS-aware routing protocol for wireless sensor network-based smart grid applications in the context of industry 4.0

M. Faheem*, V.C. Gungor

Department of Computer Engineering, Abdullah Gül University, Kayseri 38039, Turkey

ARTICLE INFO

Article history:
Received 4 September 2016
Received in revised form 8 May 2017
Accepted 21 July 2017
Available online xxx

Keywords:
Internet of things
Industry 4.0
Smart grid industry 4.0
Smart grid
Wireless sensor network
Routing
Bio-inspired
Bird mating optimization

ABSTRACT

Recently, there have been great advances in internet of things (IoT) and wireless sensor networks (WSNs) leading to the fourth industrial revolution in power grid, namely, Smart Grid Industry 4.0 (SGI 4.0). In the Smart Grid Industry 4.0 framework, the WSNs have the potential to improve power grid efficiency by cable replacement, deployment flexibility, and cost reduction. However, the smart grid (SG) environment that the WSNs operate in is very challenging because of equipment noise, dust, heat, electromagnetic interference, multipath effects and fading, which make it difficult for current WSNs to provide reliable communication. For SGI 4.0 to come true, a WSN-based highly reliable communication infrastructure is essential for successful operation of the next-generation electricity power grids. To address this need, in this paper a novel dynamic clustering based energy efficient and quality-of-service (QoS)-aware routing protocol (called EQRP), which is inspired by the real behavior of the bird mating optimization (BMO), has been proposed. The proposed distributed scheme improves network reliability significantly and reduces excessive packets retransmissions for WSN-based SG applications. Performance results show that the proposed protocol has successfully reduced the end-to-end delay and has improved packet delivery ratio, memory utilization, residual energy, and throughput.

1. Introduction

The rise of new digital industrial technology known as Industry 4.0, has lately gained a lot of interest from researchers, manufacturers, and application developers [1,2]. The objective of Industry 4.0 is to connect and integrate all objects of the traditional factory world for enabling faster, more flexible, and more efficient processes to produce higher-quality goods at low costs. Industry 4.0 will bring great opportunities for promoting industrial upgrades to increase manufacturing productivity, shift economics, faster industrial growth, and integration of supply and demand processes between the factories. This will modify the profile of the workforce-ultimately changing the competitiveness of companies and regions [3]. In Industry 4.0, the wireless or wired connected systems located in different remote places can interact with one another using standard internet-based protocols and analyze data to predict failure, configure themselves, and adapt to changes [4]. Currently, most devices within factories are connected based on wired infrastructure working over industrial protocols; however the wireless solutions are increasingly playing a complimentary role to wired solutions [5]. To this end, industrial wireless networks (IWNs) are the key technology enabling the deployment of Industry 4.0, since it can offer a reduction in energy consumption, increase economic benefits with least maintenance and breakdowns, and enable smart production. This does not only affect machine-to-machine communication, but will also have far-reaching consequences for the interplay of engineers and embedded system and wireless technologies [6].

In the context of Smart Grid Industry 4.0 (SGI 4.0), the cooperative communication requirement is focused on multiple factors, such as reliability, latency, scalability especially for a very large area of coverage and longevity of communicating devices [7]. In this respect, WSNs can significantly improve product quality, streamline operations, speed up production, make installation easier, increase the flexibility and reduce expenditure for the infrastructure in the SGI 4.0. The current and envisioned applications of WSNs in the SGI 4.0 span a wide range, including substation automation, overhead transmission line monitoring, energy management, advanced metering infrastructure, outage management, distribution automation, demand response and dynamic pricing,

* Corresponding author.
E-mail addresses: smfaheem@ymail.com (M. Faheem), cagri.gungor@agu.edu.tr (V.C. Gungor).

http://dx.doi.org/10.1016/j.asoc.2017.07.045
1568-4946/

load control and energy [8]. Importantly, all these applications would lead to new products, processes and services, improving industrial efficiency and use of sustainable energy resources while providing a competitive edge for in the fourth generation global market place. At the same time, it would ensure the reliability of the electric power infrastructure, helping to improve the daily lives of ordinary citizens. However, the realization of all these currently designed and envisioned smart grid applications directly depends on reliable and efficient communication capabilities of the deployed network.

Recent field tests show that wireless channel in power grid has many unique challenges, which do not generally exist in other communication systems. These challenges, include high packet loss rates, electromagnetic interference, equipment noise, multipath effects and fading [7]. This leads to time and location dependent delay and link quality variations of wireless links in harsh nature smart grid environments. Thus, the key design issue in smart grid is to provide reliable link quality aware data delivery under adverse wireless communications conditions [9].

Prior to the SGI 4.0, many other advanced manufacturing schemes have been presented in research studies [10–12] to empower electricity industry. Some of them focus on improving reliability and end-to-end delay performance for efficient data collection in smart grid. In [13], the authors present a WSN-based reliable routing protocol (called ETL-AODV) for smart metering systems that is capable of overcoming a false indication caused by temporary loss of data, signal interference, or invalid data. In [14], the authors design a self healing routing mechanism (called HRL-AODV) that provides energy efficient and reliable two way communications of smart meters in the power grid. In [15], the authors study the issue of routing criterion and provide airtime cost metric that considers the data size and transmission rate for reliable communication in smart grid. The authors in [16] try to solve the issue of end-to-end transmission delay for providing data communication in smart grid. To guarantee the reliability of the system, an optimal sensor computational overhead mechanism for estimating the tree delay is presented, where a sensor with the highest capacity is scheduled to transmit its sensed information to the sink. Recent filed tests and measurements have shown that organizing nodes in several clusters and specifying a particular node in each cluster to perform cluster head (CH) task, not only allows aggregation of data, but also limits data transmission primarily within and among the clusters [17]. Moreover, the coordination provided by the CH allows sensor nodes to sleep for extended period and aid to save more energy in each node. Thus, clustering improves network scalability and longevity by reducing both the traffic and the contention for the channel [18]. However, highly dynamic nature of the WSN topology requires frequent re-clustering in smart grid due to the variations in the environment conditions. In recent, some other cluster based routing schemes [19–22] have tried to improve the energy efficiency and packet delivery ratio for smart grid applications to empower electricity industry. In these schemes, the network is partitioned into clusters with a node elected as a cluster-head (CH) within each cluster. All the sensor nodes sense data and send it to their corresponding CH, which finally send it to the sink for further processing for smart grid applications. Although the existing studies focusing on routing issues provide valuable insights and guide design decisions for WSN-based smart grid applications, they generally ignore the impact of external interference and noise on transmission reliability during cluster formation in harsh smart grid environments. Moreover, in these schemes due to inappropriate inter-cluster and intra-cluster architecture between sensor nodes and CHs, a CH consumes available resources more rapidly than others nodes in the network. This unbalanced distribution of heavy data traffic load among the CHs consumes more energy and may lead to early death of nodes, which partitions the network and thereby degrade the overall performance of the WSNs in smart grid. In addition, the proposed schemes due to their excessive maintenance, high costs, energy wastage, and periodic service interruptions, they cannot distinguish the damage state of the network structure and the scope of clusters dynamics change fail to fulfill diverse QoS requirements for smart grid applications.

The existing studies provide valuable insights and guide design decisions for WSN-based smart grid applications. However, these traditional proactive methods are based on previously known global information, which is hard to obtain or update in the SGI 4.0. Moreover, none of the abovementioned studies consider the effective use of limited node memory and careful resource management in highly dynamic network topological smart grid environments. In addition, the existing routing solutions do not take into account the region affected by sparse or densely network deployment and are unaware of the changing data path opportunity due to lack of localization information. Therefore, these approaches are not fully able to optimize data path performance for reliable data transmission in an effective manner. Also, an effective neighbor discovery mechanism, which is important for robust routing is absent in most of these schemes. The omission leads to an excessive delay at hops in the network if a relay node dies in the routing path. Additionally, they generally ignore the impact of external interference and noise on transmission reliability in harsh SG environments. Therefore, these limitations of the previous studies motivate us to develop a new QoS-aware communication framework that improves memory utilization, throughput and provides reliable packet delivery with low network delay and network energy consumption for SGI 4.0 applications.

The contributions of this study can be summarized as follows: First, a novel BMO-based dynamic clustering algorithm is proposed to balance the data traffic and energy consumption among cluster members while constructing even size clusters in a fully distributed manner. Considering highly reliable links, the proposed algorithm performs cluster formation and cluster head selection stages concurrently which leads to low number of control packets. Second, an innovative BMO-based clustering routing algorithm is proposed to solve the imbalanced energy consumption among cluster heads caused by the non-uniform node distribution. The designed algorithm balances the energy consumption among CHs by adjusting the intra-cluster and inter-cluster energy consumption of CHs. Therefore, it can achieve the balance of energy consumption among nodes and prolong the network lifetime of WSN-based SG applications. Residual energy, proximity degree, number of hops through the path to the sink and actual distance each data traverses to reach the sink are such parameters involved in the selection of relay nodes. The designed scheme is fully capable to find robust data paths, load balancing and avoids data path loops in the network. Moreover, it significantly decreases the probability of packet loss and preserves high link quality among CHs in the smart grid in both sparse and densely deployment. Third, extensive simulations are carried out to evaluate the performance of the proposed protocol by comparing its performance with the existing approaches under realistic smart grid channel conditions. Performance results show that the proposed protocol has successfully reduced the end-to-end delay and has improved packet delivery ratio, memory utilization, residual energy, and throughput.

The organization of this paper is as follows: The following section briefly introduces the architecture of the bird mating optimization. Section 3 explains the proposed routing approach. Section 4 reports the network and path loss model, and simulation parameters. Section 5 provides experimental results where the effectiveness of the proposed routing scheme is compared to the

Please cite this article in press as: M. Faheem, V.C. Gungor, Energy efficient and QoS-aware routing protocol for wireless sensor network-based smart grid applications in the context of industry 4.0, Appl. Soft Comput. J. (2017), http://dx.doi.org/10.1016/j.asoc.2017.07.045

existing routing schemes. Finally, section 6 concludes this paper by summarizing our results, significance, limitation and open issues for the potential future work.

2. Bird mating optimization (BMO) working principal

BMO is a new member of the nature inspired metaheuristic approach is composed of population-based and evolutionary-based searching mechanisms, initially introduced by A. Askarzadeh in [23]. Since then it has been successfully applied to solve highly nonlinear combinatorial optimization problems in several science and engineering domains [24,25]. In BMO by observing the mating and offspring producing procedures of the bird life helps to develop the algorithm by mapping them into mathematical components. In BMO, birds are observed as social insects where the male, female and broods are the most important members of the population called society and each of them represents a feasible solution is called bird. At each generation, in society female birds are categorized into two groups based on their most promising gene structure called polyandrous and parthenogenetic. In each society, a predefined number of parthenogenetic birds having the worst fitness values are replaced by using a chaotic sequence mechanism with the new ones. The key role of the female birds is to produce new individuals and behave as mother of the current broods. On the other hand, male birds of the society are considered as father of the current broods. They only act as an amplifier for its female genome and can be categorized into main three classes, namely, monogamous, polygynous as well as promiscuous. In the population, monogamous are the birds containing superior fitness values than polygynous and promiscuous types. The main purpose of the male birds is to provide security and search food sources. In society, the polyandrous and promiscuous type birds have low percentage while the monogamous and polygynous types birds have great portion of the society which can be determined manually.

Generally, during modelling the mating process, several birds employ a variety of intelligent behaviors, including, singing, tail drumming or dancing to attract potential mates and try to reach partner sooner. Here, it is important to note that the quality of a bird is determined by its ability to get more food and memorize data paths over long distance towards the destination. Then, using a probabilistic approach each bird evaluates the quality of his/her interested bird and mates with the one having most promising genes than others. In the case of fruitful mating, male sperm is stored in the female spermatheca where the individual genome is attached to one solution to find the best solution in a d-dimensional search space for the studied optimization problem. In the beginning, both male and female birds energy to find partner is very high, which gradually decreases after each mating. This entire process repeats until a feasible solution of the problem or energy is diminished. Subsequently, during mating process each bird of the society selects and attempts to pass on better genes to their broods probabilistically. In the next stage, to enhance the next generation each bird is responsible to feed the elite selected brood with the best food. At last, a best individual (brood) converts to best feasible solution and replaces the old one based on the parent's decision in case of its supremacy to the present' solution. Otherwise, the new born brood dies and the entire process repeats iteratively. Consequently, the superiority of the bird's society improves as the algorithm progresses.

Generally, BMO consists of five types of bird species and each of them is equipped with its own updating pattern to provide a candidate solution as described below in detail.

2.1. Monogamy mating model $\left(\boldsymbol{M}^3\right)$

In this system a male bird $\vec{\mathscr{B}}^{\mathrm{m(i)}}$ tends to mate once his own interested female $\vec{\mathscr{B}}^{\mathrm{f}}$ from the parthenogenetic and polyandrous society [23]. The probabilistic approach is used to generate resultant brood. In the first phase of this approach male combines his genes with the genes of his own interesting female while in the second phase male employ the mutation probability function $P_b\left(M^3\right)$ to improve the quality of the brood is given as follows

$$\begin{aligned}&\vec{\mathscr{B}}^{b(i)} = \vec{\mathscr{B}}^{m(i)} + \phi_w \times \vec{\phi_r} \times (\vec{\mathscr{B}}^{f(i)} \times \vec{\mathscr{B}}^{m(i)})\\ &if \phi_{r1} > \phi_{mcf}\\ &\vec{\mathscr{B}}^{b(i)}(c) = \phi_l(c) - \phi_{r2} \times (\phi_l(c) - \phi_u(c)) \in P_b(M^3);\\ &end\end{aligned} \tag{1}$$

2.2. Polygyny mating model $\left(\boldsymbol{PM}^2\right)$

In this system a male bird $\vec{\mathscr{B}}^{\mathrm{m(i)}}$ tends to mate with two or more females $\left(\vec{\mathscr{B}}^{\mathrm{f(i)}}, \vec{\mathscr{B}}^{\mathrm{f(j)}}, +\ldots+\vec{\mathscr{B}}^{\mathrm{f(n)}}\right)$ from the parthenogenetic and polyandrous society. The probabilistic approach is used to generate resultant brood in which a male shares his gene with multiple females to raise quality broods. However, herein, the behavior of BMO is considered as metaphorically which means that only one brood with better fitness value will be raised in the system can be described numerically as:

$$\begin{aligned}&\vec{\mathscr{B}}^{b(i)} = \vec{\mathscr{B}}^{m(i)} + \phi_w \times \sum_{j=1}^{n_i} \vec{\phi}_{rj} \times (\vec{\mathscr{B}}_j^{\,i} \times \vec{\mathscr{B}}^{m(i)}) \in P_b(PM^2)\\ &if \phi_{r1} > \phi_{mcf}\\ &\vec{\mathscr{B}}^{b(i)}(c) = \phi_l(c) - \phi_{r2} \times (\phi_l(c) - \phi_u(c))\\ &end\end{aligned} \tag{2}$$

2.3. Promiscuity mating model $\left(\boldsymbol{PrM}^2\right)$

This system follows the chaotic social structure in which there is no stable relationships between the male $\vec{\mathscr{B}}^{\mathrm{m(i)}}$ and female $\vec{\mathscr{B}}^{\mathrm{f(i)}}$ birds in one time event. In this type of mating most likely male will not see the female for another brief visit and will almost certainly never see his brood or the nest. Although, the promiscuous birds generation follows chaotic sequence, however, the promiscuous bird breeding process is similar to that of monogamous birds.

2.4. Polyandry mating model $\left(\boldsymbol{PoM}^2\right)$

In this model a female $\vec{\mathscr{B}}^{\mathrm{f(i)}}$ tends to mate with two or more than two males $\left(\vec{\mathscr{B}}^{\mathrm{m(i)}}, \vec{\mathscr{B}}^{\mathrm{m(j)}}, +\ldots+\vec{\mathscr{B}}^{\mathrm{m(n)}}\right)$ from the monogamy, polygyny or promiscuity society. Similar to M^3, in PoM^2, probabilistic approach is used to generate resultant brood in which a female share her gene with multiple males to raise better fitness quality broods. The resultant brood creation is similar as in Eq. (2).

2.5. Parthenogenesis mating model $\left(\boldsymbol{PrM}^2\right)$

in this system a female tends to raise brood without the help of males by passing on better genes to her brood to make a small

Please cite this article in press as: M. Faheem, V.C. Gungor, Energy efficient and QoS-aware routing protocol for wireless sensor network-based smart grid applications in the context of industry 4.0, Appl. Soft Comput. J. (2017), http://dx.doi.org/10.1016/j.asoc.2017.07.045

change in her genes probabilistically. The new generated brood can be numerically expressed by the following process as:

$$\begin{aligned} &for i = 1 : n \\ &if \phi_{r1} > \phi_{pmcf} \\ &\vec{\mathcal{B}}^{b}(i) = B^{f}(i) + \phi_{\sigma} \times (\phi_{r2} - \phi_{r3}) \times B^{f}(i);\ \in P_b(PrM^2) \\ &else \\ &\vec{\mathcal{B}}^{b}(i) = B^{f}(i) \\ &end \\ &end \end{aligned} \tag{3}$$

This entire natural phenomenon has been mapped into mathematical components where the number of matings is the number of iterations, a bird with highest fitness value is the best solution and birds sperm is the decision variable to study the optimization problem of the developed algorithms. In each mating, the birds energy gradually decreases which can be calculated by taking into account the following equation as:

$$\phi_{E(t+1)} = \propto \times E(t) \tag{4}$$

After completing the mating process, all died society members position is swapped by the new breeding ones and assumed to be the responsible one for feeding and improving the next generation of the broods. In addition, the simulated functions limit include the prior define number of society bird types, female spermatheca capacity in which the information of utmost number of mating's in each mating, and the number of broods related to each individual bird (Table 1).

3. Proposed routing protocol (EQRP)

This section introduces a new clustering routing protocol for WSN-based smart grid applications inspired by the real process of bird mating. In society, the male, female and brood birds have their own genome composed of genes. This genome is attached to each individual solution during modelling the mating process of our proposed scheme to study the optimization problem. A list of numerical values is used to describe one genome where each value is attached to a decision variable that indicates an unknown of the problem solutions. The fitness function is the one who is responsible to present the genome associated solution in a stronger or weaker manner. This genotype representation presents a solution with several unknowns equivalent to the total number of clusters that form the essential for designing algorithms studied as clustering solution at the lowest cost (the best performance clustering solution). Here, the best performance clustering means to generate a high quality clustering solution, which exploits the fitness function for each individual cluster generated in the network. The superior quality solution adopts the minimum Euclidian distance information between each set of pair of node to find effective and energy efficient clustering in the smart grid. This includes the low transmission energy consumption cost and low clustering complexity as well as preserves the link quality between sensor nodes in the harsh smart grid environment. The entire process of our scheme consists of seven different phases are shown below:

Phase1. Defining Input Data

In this phase, we define the input data, including the number of monogamous ($\mathcal{B}_{\mathrm{m}}$), polygynous ($\mathcal{B}_{\mathrm{pg}}$), promiscuous ($\mathcal{B}_{pc}$), polyandrous ($\mathcal{B}_{pa}$), parthenogenetic ($\mathcal{B}_{pag}$), broods ($\vec{\mathcal{B}}^{b(i)}$), female bird's spermatheca capacity (δB^{f}), bird's allure factor ($\mathcal{B}_{\mathrm{Amax.||Amin.}}$), initial energy ($\mathcal{B}_{E(i)}$), decreasing factor ($\propto$), number of iterations ($\mathcal{S}_{maxi}$), transmission loss matrix coefficients and other different parameters as shown in Table 2.

Phase2. Network Initialization

Table 1
Notations used in EQRP

Notation	Description
$\vec{\mathcal{B}}^{m(i)}$	male bird i from the monogamy mating system
$\vec{\mathcal{B}}^{f(i,j)}$	female bird i or j from the parthenogenetic and polyandrous birds
$\vec{\mathcal{B}}^{b(i)}$	resultant brood i generated after mating ϕ_{mi}
ϕ_w	time varying weight to adjust the importance of the interested female
$\vec{\phi}_r$	$1 \times d$ vector whose each element are distributed randomly between 0 and1
ϕ_{mcf}	mutation control factor varying between 0 and1
c	is the random number between [0,n], where n denotes the problem dimension
m	is the mutation control factor between 0 and 1
r_i	is the random number whose value is between [0,1]
ϕ_l, ϕ_u	upper and lower bound of the element, respectively
$P_b\left(M^3\right)$	$P_b\left(M^3\right) = 1 - \phi_{mcf}$ is the probability of mutation in one brood's gene
n_i	is the number of interested birds
$\vec{\mathcal{B}}_j^i$	denotes the $j$thinteresting bird for mating
P_b, M	P_b and M indicate the probability and mutation process
$P_b\left(PM^2\right), P_b\left(PrM^2\right)$	is the probability of a bird whose value is selected randomly between [0,1]
$\phi_{pmcf}, \phi_{\sigma}$	is the parthenogenetic mutation control factor and ϕ_{σ} denotes the step size
$\phi_{E(t+1)}$	indicates the bird's energy decreases at each event time t
$\phi_{P_b\left(\vec{\mathcal{B}}^{m(i)}, \vec{\mathcal{B}}^{f(i)}\right)}$	denotes the sperm adding probability of a male $\vec{\mathcal{B}}^{\mathrm{m(i)}}$ to female $\vec{\mathcal{B}}^{f(i)}$spermatheca
$\propto, \beta, \gamma$	shows the decay rate whose value is randomly selected between[0, 1]
$t, \mathcal{B}_A(t), A(t)$	represents the allure factor at time event t $\in$ [0, 1, 2, …, n]
A, E	shows the allure and energy of a male $\vec{\mathcal{B}}^{\mathrm{m(i)}}$ and female $\vec{\mathcal{B}}^{\mathrm{f(i)}}$ bird, respectively.
$\mathcal{S}_i^{max}, \mathcal{S}_i^{min}$	allure factors whose maximum value is set to 1 and minimum value to 0
$\phi_{P_b\left(\vec{\mathcal{B}}^{m(i)}, \vec{\mathcal{B}}^{f(i)}\right)}$	illustrates the successfully mating probability $\in$ random [0, 1]
$f\left(\vec{\mathcal{B}}^{(i)}\right)$ $and f\left(\vec{\mathcal{B}}^{m(i)}\right)$	the performance functions of $\vec{\mathcal{B}}^{\mathrm{f(i)}}$ and $\vec{\mathcal{B}}^{\mathrm{m(i)}}$ $\in$ (i = 1, 2, …, n) are the ith fitness function values for the female and male bird, respectively.
β	indicates the decay coefficient belongs to [0.1,0.5]
$\gamma \mathcal{B}^i E$	is the energy loss after each iteration
$\aleph$	is bird attracting factor whose value is set to 1 for both male and female birds
DT_{MOD}	is the defined distance threshold value
CT_{MOD}	is the calculated value of threshold distance in the deployment field
$2\pi/3 and\ \pi$	the switching angles, $\mathcal{P}_{\mathrm{low}}$ and $\mathcal{P}_{\mathrm{high}}$ denotes the low and high transmission power
$E_{\max}$ & E_{rd}	indicates the maximum available energy and remaining energy of the node
RN_i	is a relay node i
Rd	represents the satisfied routing destinations and R is the routing table in a node j
C_i	is a directed chain rooted at the source node S

The designed scheme takes into account the advantage of a centralized control mechanism to start the clustering process for partitioning the set of nodes into multiple clusters. In the beginning, a sink is responsible to initiate the cluster formation process by disseminating initiation message (*init_msg*) in the network. After receiving information from the sink each sensor node is responsible to disseminate *init_msg* to its neighboring nodes by taking into account the *CSMA* (Carrier Sense Multiple Access) mechanism. This *init_msg* contains information about the sender sensor node identity (*id*), residual energy (R_e) and angle of transmit-

ting (*AOT*). Then after receiving *init _ msg* successfully each receiver sensor node replies to the sender node by sending an acknowledge message(*ack_msg*), which guarantees that a message has been received successfully. This *ack _ msg* contains the information of the receiver sink node identity, residual energy and angle of receiving (*AOR*) measured during receiving information. In our scheme, it is also possible that a sink node may receive multiple *init _ msg* copies from the same or different senders sink neighbors. In this case, each receiver sink node follows first come first serve (FCFS) based policy, which means that a receiver node only replies to that sender node to whom it receives the message first while the rest of the *init _ msg* will be ignored. Initially, throughout the sending and receiving process each node is responsible to maintain information about the *AOT*, *AOR*, neighboring node *id*, residual energy and distance information (D_i) in its routing table in decreasing order.

Phase3: Initialization of population

In this phase, after receiving *init _ msg* by all birds (sensor nodes) in the problem space an initial set of the population is randomly generated by using a random number generator between 0 and 1 for the potential solutions of the problem. A population consists of a group of individuals (sensor nodes of size = 300) that exemplify an entire solution to a defined problem where each individual is represented by a sequence of 0 s or 1s. During initializing, the generational Genetic algorithm (GA) and steady-state GA are the two popular methods used to generate a new population of the individuals. The generational GA replaces all members of the population, whereas the steady-state GA swaps one or two population members at each generation of the evolution. In this paper, we adopt the generational GA method in which designed algorithm is responsible to preserve a certain amount of the best individuals at each generation and copies them to new generations as studied in [26–30]. This results a new population that will have the current individuals as a result of crossover and mutation as well as numerous individuals from the previous generation to optimize the solution in problematical search space. The initial population of individuals (ϕ_i) generated for the potential solutions of the problem can be numerically formulated as:

$$\phi_{initial\_society} = \begin{bmatrix} \phi_1 \\ \phi_2 \\ . \\ \phi_n \end{bmatrix} \quad (5)$$

$$\phi_i = \left[\phi_i^3 + \phi_i^2 + \phi_i^5, \ldots, +\phi_i^n\right], \forall i = 1, 2, \ldots, n \quad (6)$$

Each genome randomly builds within admissible range of the variables by considering its genome list of similarity of the unknowns. Then, based on the fitness function values an initial population of the birds (ϕ_i) is ranked decreasingly can be written as:

$$\phi_i = \left[\phi_i^1 + \phi_i^2 + \phi_i^3, \ldots, +\phi_i^n\right] \quad (7)$$

Thus, entire society containing various types of bird can be numerically indicated as

$$\phi_{society} = \begin{bmatrix} \phi_m 1 & \phi_{pg} 1 & \phi_{pr} 1 & \phi_{pa} 1 & \phi_{pag} 1 \\ \phi_m 2 & \phi_{pg} 2 & \phi_{pr} 2 & \phi_{pa} 2 & \phi_{pag} 2 \\ . & . & . & . & . \\ . & . & . & . & . \\ \phi_{m_n} & \phi_{pg} & \phi_{pr} & \phi_{ar} & \phi_{aggr} \end{bmatrix} \quad (8)$$

The one with the best quality solution is selected as initial individual monogamous ($\phi^i_{m_i}$) birds in the existing population for the current round *i* is formulated as:

$$\phi_{monogamous} = \left[\phi^1_{m_i}, \phi^2_{m_i}, \phi^2_{m_i}, \ldots, +\phi^n_{m_i}\right] \quad (9)$$

After employing a set of monogamous birds some of the remaining solution forms a list of polygynous birds $\left(\phi^i_{pg_i}\right)$ are ranked in decreasing order can be numerically shown as:

$$\phi_{polygynous} = \left[\phi^1_{pg_i} + \phi^2_{pg_i} + \phi^3_{pg_i}, \ldots, +\phi^n_{pg_i}\right] \quad (10)$$

The remaining solution with lower fitness value forms a list of promiscuous birds $\left(\phi^i_{pr_i}\right)$ are ranked in decreasing order can be shown as:

$$\phi_{promiscuous} = \left[\phi^1_{pr_i} + \phi^2_{pr_i} + \phi^3_{pr_i}, \ldots, +\phi^n_{pr_i}\right] \quad (11)$$

In society, the best quality solution among the female types is called polyandrous $\left(\phi^i_{pa_i}\right)$ can be numerically formulated in decreasing order as:

$$\phi_{promiscuous} = \left[\phi^1_{pa_i} + \phi^2_{pa_i} + \phi^3_{pa_i}, \ldots, +\phi^n_{pa_i}\right] \quad (12)$$

while the parthenogenetic female birds $\left(\phi^i_{pag_i}\right)$ with lower fitness value can be numerically expressed as:

$$\phi_{parthenogenetic} = \left[\phi^1_{pag_i} + \phi^2_{pag_i} + \phi^3_{pag_i}, \ldots, +\phi^n_{pga_i}\right] \quad (13)$$

In addition to its genome, each female bird is characterized by her spermatheca capacity ($\mathcal{Q}_n$) as well as her allure values $((\mathcal{B}_{Amax})/(\mathcal{B}_{Amin}))$ and/or energy $\left(\mathscr{P}_{E(i)}\right)$. The spermatheca capacity is equal to the maximum number of males that may mate with the female is set as constant during all mating and can be numerically expressed as:

$$\phi_{femalebirds\_spermatheca} = \begin{bmatrix} \phi_{\delta\mathscr{F}} 1 \\ \phi_{\delta\mathscr{F}} 2 \\ . \\ . \\ \phi_{\delta\mathscr{F}\, n(sprem)} \end{bmatrix} \quad (14)$$

$$\phi_{\delta\mathscr{F}} = \left[\phi^1_{\delta\mathscr{F}\, i} + \phi^2_{\delta\mathscr{F}\, i} + \phi^3_{\delta\mathscr{F}\, i}, \ldots, +\phi^n_{\delta\mathscr{F}\, i}\right] \quad (15)$$

where δ is the spermatheca capacity set to constant for all the female birds $\phi^i_{\delta\mathscr{F}\, i}$.

Phase4: Mating Process

In the beginning of the mating process, each individual bird $\mathcal{B}_i$ based on the available energy $\left(E_i^{max}, E_i^{min}\right)$ and allure $\left(\mathcal{S}_i^{max}, S_i^{min}\right)$ parameter values is initialized as:

$$\phi_{E\mathscr{B}_i} = rand(\cdot) \times (E_i^{max} - E_i^{min}) \quad (16)$$

and

$$\phi_{S\mathscr{B}_i} = \text{rand}(\cdot) \times (\mathcal{S}_i^{\max} - \mathcal{S}_i^{\min}) + \mathcal{S}_i^{\min} \cdot A \quad (17)$$

Initially, during the mating process a number of polyandrous birds are generated randomly and monogamous, polygynous, and promiscuous birds are positioned over those females. The transition made by each bird depends on his/her allure values. If the allure is higher then a bird can make extremely large steps in the search space, which signifies the flipping probability of individual gene in the partner bird allele. At each step in the search space, we use simulated annealing method to choose the best male at random to mate with the female at that step. Then, according to the fitness function value of the male and elected female mating probability is evaluated. After that, we generate a random number between 0 and 1 to compare with the calculated probability. If generating probability is less than the computed probability, then male can successfully mate the female and his genome stored in the female's

Please cite this article in press as: M. Faheem, V.C. Gungor, Energy efficient and QoS-aware routing protocol for wireless sensor network-based smart grid applications in the context of industry 4.0, Appl. Soft Comput. J. (2017), http://dx.doi.org/10.1016/j.asoc.2017.07.045

spermatheca otherwise discarded. The probability of a male $\vec{\mathscr{B}}^{\mathrm{m(i)}}$ successful mate with the female $\vec{\mathscr{B}}^{\mathrm{f(i)}}$ can be calculated as:

$$\phi_{P_b\left(\vec{\mathscr{B}}^{m(i)},\vec{\mathscr{B}}^{f(i)}\right)} = exp\left(|f\left(\vec{\mathscr{B}}^{(i)}\right) - f\left(\vec{\mathscr{B}}^{m(i)}\right)|/\mathscr{B}_A(t)\right) \tag{18}$$

$$|\left(f(\vec{\mathscr{B}}^{(i)}\right) - f\left(\vec{\mathscr{B}}^{m(i)}\right))| = \sqrt{\left(f^1_{\mathscr{B}_i} - f^1_{\mathscr{B}_{m(i)}}\right)^2 + \left(f^2_{\mathscr{B}_i} - f^2_{\mathscr{B}_{\mathscr{B}(i)}}\right)^2 +, \ldots, +\left(f^n_{\mathscr{B}_i} - f^n_{\mathscr{B}_{m(i)}}\right)^2} \tag{19}$$

The above function is highly associated with the bird allure factor and offers superior values when bird allure factor is greater or when the mating partner fitness function is closed to the bird fitness function. During the first mating process, among the available the best male $Best\vec{\mathscr{B}}^{m(i)}$, with the largest annealing functions of probability $P_b\left(\vec{\mathscr{B}}^{\mathrm{f(i)}}, \vec{\mathscr{B}}^{\mathrm{m(i)}}\right)$ are selected randomly from the male list ranked in decreasing order as the object of mating for the female $\vec{\mathscr{B}}^{\mathrm{f(i)}}$ can be formally written as:

$$\phi_{Best\_male\_society} = \begin{bmatrix} \phi_{Best.\vec{\mathscr{B}}^{m(1)}} \\ \phi_{Best.\vec{\mathscr{B}}^{m(2)}} \\ . \\ . \\ \phi_{Best.\vec{\mathscr{B}}^{m(n)}} \end{bmatrix} \tag{20}$$

$$\phi_{Best.\vec{\mathscr{B}}^{m}} = \left[\phi^1_{B.\vec{\mathscr{B}}^{m(i)}} + \phi^2_{B.\vec{\mathscr{B}}^{m(i)}} + \phi^3_{B.\vec{\mathscr{B}}^{m(i)}}, \ldots, +\phi^n_{B.\vec{\mathscr{B}}^{m(n)}}\right], \quad letBest = B \tag{21}$$

For each individual best quality female bird $\vec{\mathscr{B}}^{\mathrm{m(i)}}$, the mating process $\mathfrak{m}^i$ with randomly selected male $\vec{\mathscr{B}}^{\mathrm{m(i)}}$ from the list can numerically written as:

$$\phi_{\delta.\mathscr{B}^{f(n)}} = [\phi^1_{B.\vec{\mathscr{B}}^{m(1)}} \times rand(\cdot) + \phi^2_{B.\vec{\mathscr{B}}^{m(1)}} \times rand(\cdot) +, \ldots, +\phi^n_{B.\vec{\mathscr{B}}^{m(1)}} \times rand(\cdot)] \tag{22}$$

For entire mating process the best set of males $\phi_{B.\vec{\mathscr{B}}^{\mathfrak{m}}}$ and female $\phi_{\delta.\mathscr{B}}{}^{f(\mathfrak{m})}$ birds can be numerically denoted as:

$$\phi_{B.\vec{\mathscr{B}}^{m^j}}\left[\phi^1_{B.\vec{\mathscr{B}}^{m(i)}} + \phi^2_{B.\vec{\mathscr{B}}^{m(i)}} + \phi^3_{B.\vec{\mathscr{B}}^{m(i)}}, \ldots, +\phi^n_{B.\vec{\mathscr{B}}^{m(i)}}\right]^n, \quad = \forall j = 1, 2, \ldots, n \tag{23}$$

$$\phi_{\delta.\mathscr{B}}f(\mathfrak{m}) = [\phi^1_{\delta.\mathscr{B}^{f(i)}} + \phi^2_{\delta.\mathscr{B}^{f(i)}} + \phi^3_{\delta.\mathscr{B}^{f(i)}}, \ldots, +\phi^n_{\delta.\mathscr{B}^{f(i)}}]^n \tag{24}$$

If the male bird's genes are selected randomly for mating from the female's spermatheca, then selection can be written as:

$$\phi^{\delta\vec{\mathfrak{n}}}_{B.\vec{\mathscr{B}}^{m}}j = [\phi^1_{B.\vec{\mathscr{B}}^{m(i)}} \times rand(\cdot) + \phi^2_{B.\vec{\mathscr{B}}^{m(i)}} \times rand(\cdot) +, \ldots, +\phi^n_{B.\vec{\mathscr{B}}^{m(i)}} \times rand(\cdot)]^n, \quad rand(\cdot) \in [0, 1] \tag{25}$$

Since the number of mating males are constant in all mating process thus as a sum entire mating process can be formulated as:

$$\forall\phi(\mathscr{B}^i) = \mathscr{B}^{f(1)}\sum_{i=1}^{n}(\phi_{B.\vec{\mathscr{B}}^{m(i)}} \times rand(\cdot))^i + \mathscr{B}^{f(2)} \sum_{i=1}^{n}(\phi_{B.\vec{\mathscr{B}}^{m(i)}} \times rand(\cdot))^i +, \ldots, +\mathscr{B}^{f(n)}\sum_{i=1}^{n}(\phi_{B.\vec{\mathscr{B}}^{m(i)}} \times rand(\cdot))^i \tag{26}$$

The above equation in sum may be formulated as:

$$\phi[\forall\mathscr{B}^i] = \mathscr{B}^{f(i)}\sum_{i=1}^{n}(\phi_{B.\vec{\mathscr{B}}^{m(i)}} \times rand(\cdot))^i \tag{27}$$

The bird's energy decay iteratively depending on the time t_i even if the mating processes succeed or not can be numerically indicated as:

$$\phi\mathscr{B}^i(t_{i+1}) = \beta \times \mathscr{B}^i\mathcal{S}(t_i) and \phi\mathscr{B}^iE(t_{i+1}) = (\mathscr{B}^iE(t_i) - \gamma\mathscr{B}^iE) \tag{28}$$

The entire process repeats either until each bird allure/energy factor reaches to the minimum given value of $\mathscr{B}^iS_{\min}$ and $\mathscr{B}^iE_{\min}$ or until the female's spermatheca is full (maximum capacity $\mathscr{B}^i\left(\delta_n\right)$).

Phase5: Breeding Process

After completing the mating process all birds start breeding process where each female bird is mated with an arbitrarily chosen male sperm stored in her spermatheca with her own genome for the desired number of broods. The genome of each individual female bird $\mathscr{B}^{\mathrm{f(i)}}$ can be numerically expressed as:

$$\phi^1_{\delta.\mathscr{B}^{f(i)}} = \left[X_1\phi^1_{\delta.\mathscr{B}^{f(i)}} + X_2\phi^2_{\delta.\mathscr{B}^{f(i)}} + X_3\phi^3_{\delta.\mathscr{B}^{f(i)}} +, \ldots, +X_n\phi^n_{\delta.\mathscr{B}^{f(i)}}\right], \quad forX \in \mathscr{B}^{f(i)}(\text{genome}) \tag{29}$$

The male sperm stored in the female $\mathscr{B}^{f(i)}$ spermatheca can be formally expressed as:

$$\phi^1_{\delta.\mathscr{B}^{f(i)}} = \left[Y_1\phi^2_{B.\vec{\mathscr{B}}^{m(i)}} + Y_2\phi^1_{B.\vec{\mathscr{B}}^{m(i)}} + Y_3\phi^3_{B.\vec{\mathscr{B}}^{m(i)}} +, \ldots, +Y_n\phi^n_{B.\vec{\mathscr{B}}^{m(i)}}\right], \quad forY \in \mathscr{B}^{m(i)}(\text{sperm}) \tag{30}$$

In decreasing order above equation can illustrated as:

$$\phi^1_{\delta.\mathscr{B}^{f(i)}} = \left[Y_2\phi^1_{B.\vec{\mathscr{B}}^{m(i)}} + Y_1\phi^2_{B.\vec{\mathscr{B}}^{m(i)}} + Y_3\phi^3_{B.\vec{\mathscr{B}}^{m(i)}} +, \ldots, +Y_n\phi^n_{B.\vec{\mathscr{B}}^{m(i)}}\right] \tag{31}$$

For each individual female $\mathscr{B}^{\mathrm{f(i)}}$, the number of mating males may be formally indicated as:

$$\phi^1_{\delta.\mathscr{B}^{f(i)}} = \sum_{i=1}^{n} X * (Y\phi^n_{B.\vec{\mathscr{B}}^{m(i)}}) \tag{32}$$

For each individual male $\mathscr{B}^{m(i)}$, the number of mating females may be formally indicated as:

$$\phi^1_{\mathscr{B}^{m(i)}} = \sum_{i=1}^{n} Y * (X\phi^n_{B.\vec{\mathscr{B}}^{f(i)}}) \tag{33}$$

For each individual female $\mathscr{B}^{f(i)}$ more than one mating process can be expressed as:

$$\phi^i_{\delta.\mathscr{B}^{f(i)}} = \sum_{j=1}^{k}\sum_{i=1}^{n} X * (Y\phi^n_{B.\vec{\mathscr{B}}^{m(i)}}), \quad where\forall j = 1, 2, \ldots, k \tag{34}$$

Please cite this article in press as: M. Faheem, V.C. Gungor, Energy efficient and QoS-aware routing protocol for wireless sensor network-based smart grid applications in the context of industry 4.0, Appl. Soft Comput. J. (2017), http://dx.doi.org/10.1016/j.asoc.2017.07.045

Similarly, for each individual male $\mathscr{B}^{m(i)}$ more than one mating process can be expressed as:

$$\phi^{i}_{\mathscr{B}^{m(i)}} = \sum_{j=1}^{k}\sum_{i=1}^{n} Y * (X\phi^{n}_{B.\vec{\mathscr{B}}^{f(i)}}) \tag{35}$$

In sum, for all the female birds in the network the equation can be formed as:

$$\phi_{\mathscr{B}^{f}} = \sum_{i=1}^{n}(\phi^{n}_{\mathscr{B}^{f(i)}}) \cdot [\sum_{i=1}^{n} X * (Y\phi^{n}_{B.\vec{\mathscr{B}}^{m(i)}})] \tag{36}$$

Similarly, for male birds above equation can be modified as below

$$\phi_{\mathscr{B}^{m}} = \sum_{i=1}^{n}(\phi^{n}_{\mathscr{B}^{m(i)}}) \cdot [\sum_{i=1}^{n} Y^{*}(X\phi^{n}_{B.\vec{\mathscr{B}}^{f(i)}})] \tag{37}$$

In the problematic search space, to produce a pair of offspring strings, crossover operator is responsible for partial swapping of the bits between two individual strings with a single heuristic crossover operator numerically can be shown as:

$$\phi_{\mathcal{NB}_{b(i)}} = \mathscr{B}^{f(i)} + P_c(r(\mathscr{B}^{f(i)} + \mathscr{B}^{m(i)})) \tag{38}$$

Generally, crossover is also known as recombination of component materials is a process of choosing a random position in the string and swapping the characters either left or right of this point with another similarly partitioned string. For more simple, a crossover is a binary genetic operator acting on two parents at random points to alter the genome of individuals in the current population. This random position is called the crossover point. It is a simulation of the sexual reproductive process where the outcome of crossover heavily depends on the selection of chromosomes made from the population for the transfer of genetic inheritance. In recent, a number of crossover operators have been proposed. However, in this study we use the simplest one called single-point crossover operator whereby a point is selected at random and the two parent chromosomes exchange information after that point. To generate a new population of individuals we define a decreasing individual crossover rate probability as $P_c = 0.99$, $P_c = 0.95$, $P_c = 0.90$ is the parameters that affect the rate at which the crossover operator is applied. This high crossover rate introduces new clusters more quickly into the population by eliminating low quality clusters then selection can produce improvements rather than low crossover rate which may cause stagnations due to the lower exploration rate. Furthermore, this decreasing probability helps to find the optimal solution in more robust and accurate manner even when a set of sensor nodes die in the ending rounds of the deployed network in the smart grid. To generate a new bird $\mathcal{NB}_{b(i)}$ (a new clustering solution including cluster heads), the male birds $\mathscr{B}^{m(i)}$ is randomly ($rand \in [0, 1]$) selected from the spermatheca where η helps the rounding solution toward the nearest integer. For a given crossover probability each pair of strings from the new brood population is randomly picked out by proposed algorithm. Then a uniformly distributed random number is generated in the predefined range between 0 and 1. If the generated random number is less than the predefined value, then the developed algorithm is responsible to apply the crossover operator for producing swapping new bird generation. Otherwise, it does not apply crossover operator on these two particular strings. Through applying the crossover operator along with the strings a crossover point is also randomly generated. The crossover operator plays an important role for swapping individual bird values from one string to another pairwise string in the population.

First, a single or a list of male genome $\left(Y_1\phi^{1}_{B.\vec{\mathscr{B}}^{m(i)}} | Y_2\phi^{2}_{B.\vec{\mathscr{B}}^{m(i)}}, Y_3\phi^{3}_{B.\vec{\mathscr{B}}^{m(i)}}, Y_4\phi^{4}_{B.\vec{\mathscr{B}}^{m(i)}}, Y_5\phi^{5}_{B.\vec{\mathscr{B}}^{m(i)}}\right)$ are randomly selected from the female's spermatheca with genome Y_1, Y_2, Y_3, Y_4 and Y_5 where $Y_1 \neq Y_2 \neq Y_3 \neq Y_4 \neq Y_5$. Then, an improved solution in the entire problem search space by applying the crossover operator can be numerically expressed as:

$$\phi^{i\ldots n}_{\delta\mathscr{B}^{f}} = \left[\left(X\phi^{1}_{\delta\mathscr{B}^{f(i)}} + X\phi^{2}_{\delta\mathscr{B}^{f(j)}} +, \ldots, +X\phi^{n}_{\delta\mathscr{B}^{f(n)}}\right) \circledast \left(Y\phi^{1}_{B.\vec{\mathscr{B}}^{m(i)}} + Y\phi^{3}_{B.\vec{\mathscr{B}}^{m(j)}} +, \ldots, +Y\phi^{n}_{B.\vec{\mathscr{B}}^{m(n)}}\right)\right] \tag{39}$$

subjected to

$$\phi^{1}_{\delta\mathscr{B}^{f}} X^{1}_{improved} = Y_1\phi^{1}_{B.\vec{\mathscr{B}}^{m(i)}} + rand(\cdot) \times (Y_2\phi^{2}_{B.\vec{\mathscr{B}}^{m(j)}} - Y_3\phi^{3}_{B.\vec{\mathscr{B}}^{m(k)}}) \times (Y_4\phi^{4}_{B.\vec{\mathscr{B}}^{m(l)}} - Y_5\phi^{5}_{B.\vec{\mathscr{B}}^{m(e)}}) \tag{40}$$

$$\phi^{1}_{\delta\mathscr{B}^{f}} = \left[\left(X_1\phi^{1}_{\delta\mathscr{B}^{f(i)}} + X_2\phi^{2}_{\delta\mathscr{B}^{f(j)}} +, \ldots, +X_2\phi^{5}_{\delta\mathscr{B}^{f(e)}}\right) \circledast \left(Y_1\phi^{1}_{B.\vec{\mathscr{B}}^{m(i)}} + Y_2\phi^{2}_{B.\vec{\mathscr{B}}^{m(j)}} +, \ldots, +Y_5\phi^{5}_{B.\vec{\mathscr{B}}^{m(e)}}\right)\right] \tag{41}$$

$$\phi^{1}_{\delta\mathscr{B}^{f}} = \left[\left(Y_1\phi^{1}_{B.\vec{\mathscr{B}}^{m(i)}} \circledast X_1\phi^{1}_{\delta\mathscr{B}^{f(i)}} + Y_2\phi^{1}_{B.\vec{\mathscr{B}}^{m(i)}} \circledast X_3\phi^{1}_{\delta\mathscr{B}^{f(i)}} + Y_4 \phi^{1}_{B.\vec{\mathscr{B}}^{m(i)}} \circledast X_5\phi^{1}_{\delta\mathscr{B}^{f(i)}} +, \ldots, +Y_3\phi^{1}_{B.\vec{\mathscr{B}}^{m(i)}} \circledast X_4\phi^{1}_{\delta\mathscr{B}^{f(i)}}\right)\right] \tag{42}$$

$$\phi^{1}_{\delta\mathscr{B}^{f}} = \left[\left(Y_1 \otimes X_1\phi^{1}_{B.\vec{\mathscr{B}}^{m(i)}}\phi^{1}_{\delta\mathscr{B}^{f(i)}} + Y_2 \otimes X_3\phi^{1}_{B.\vec{\mathscr{B}}^{m(i)}}\phi^{1}_{\delta\mathscr{B}^{f(i)}} + Y_4 \otimes X_5\phi^{1}_{B.\vec{\mathscr{B}}^{m(i)}}\phi^{1}_{\delta\mathscr{B}^{f(i)}} +, \ldots, +Y_3 \otimes X_4\phi^{1}_{B.\vec{\mathscr{B}}^{m(i)}}\phi^{1}_{\delta\mathscr{B}^{f(i)}}\right)\right] \tag{43}$$

$$\phi^{1}_{\delta\mathscr{B}^{f}} = \left[\left(\overline{Y_1X_1}\phi^{3}_{B.\vec{\mathscr{B}}^{m(k)}}\phi^{1}_{\delta\mathscr{B}^{f(i)}} + \overline{X_3Y_2}\phi^{5}_{B.\vec{\mathscr{B}}^{m(e)}}\phi^{2}_{\delta\mathscr{B}^{f(j)}} + \overline{Y_4X_5}\phi^{1}_{B.\vec{\mathscr{B}}^{m(i)}}\phi^{3}_{\delta\mathscr{B}^{f(k)}} +, \ldots, +\overline{Y_3X_4}\phi^{n}_{B.\vec{\mathscr{B}}^{m(n)}}\phi^{n}_{\delta\mathscr{B}^{f(n)}}\right)\right] \tag{44}$$

Finally, we get

$$\phi^{1}_{\delta\mathscr{B}^{f}} = \left[\left(\overline{Y_1X_1}\phi^{3}_{B.\vec{\mathscr{B}}^{m(k)}}\phi^{1}_{\delta\mathscr{B}^{f(i)}} + \overline{X_2Y_3}\phi^{5}_{B.\vec{\mathscr{B}}^{m(e)}}\phi^{2}_{\delta\mathscr{B}^{f(j)}} + \overline{Y_5X_4}\phi^{1}_{B.\vec{\mathscr{B}}^{m(i)}}\phi^{3}_{\delta\mathscr{B}^{f(k)}} +, \ldots, +\overline{Y_4X_2}\phi^{n}_{B.\vec{\mathscr{B}}^{m(n)}}\phi^{n}_{\delta\mathscr{B}^{f(n)}}\right)\right] \tag{45}$$

The new generation introduced as a result of the crossover will only have the traits of the parents, which sometimes lead to a problem where no new genetic material is introduced in the offspring. Mutation is the process of random modification of the value of a string position with a small probability by allowing new genetic patterns in the new generation [31]. It is basically a secondary search operator, which increases the diversity in the population with which each bit position of each string in the new population undergoes a random change after the selection process. Mutation introduces a new sequence of genes into a chromosome by ensuring that the probability of searching any region in the problem space is never zero. This prevents complete loss of genetic material occurs through reproduction and crossover. However, it does not give guarantee to produce desirable features in the new chromosome. The selection process will retain it if the fitness of the mutated chromosome is higher than the general population. Otherwise, selection will ensure that the chromosome will not live to

Please cite this article in press as: M. Faheem, V.C. Gungor, Energy efficient and QoS-aware routing protocol for wireless sensor network-based smart grid applications in the context of industry 4.0, Appl. Soft Comput. J. (2017), http://dx.doi.org/10.1016/j.asoc.2017.07.045

mate in the future. Unlike crossover, the mutation rate is considered very low which prevents any bit position from getting stuck to a single value whereas a high mutation rate results in essentially random search and depends on how often it is applied. In this paper, we consider a uniform mutation operator with the probability of ($P_m = 0.01$), to perform the mutation operation with the assertion that no important genetic material is lost. The new generation, which is the result of the crossover, the designed algorithm takes into account each string bit by bit (i.e. node by node), by generating a uniformly distributed random number between 0 and 1. If the random number value is less than the predefined value, then the developed algorithm is responsible to apply the mutation operator to the new individual else it does not ponder mutation operator to that particular node. This genomic type representation trickily eases the creation of dynamic clusters in the whole network in the smart grid. In a uniform mutation operator where the value of a gene for mutation is modified can be mathematically expressed as:

$$\phi^1_{\delta\mathscr{B}f} = [(\overset{\cdots}{\overline{Y_1X_1}}\phi^3_{B\vec{\mathscr{B}}m(k)}\phi^1_{\delta\mathscr{B}f(i)} + \overset{\cdots}{\overline{X_2Y_2}}\phi^5_{B\vec{\mathscr{B}}m(e)}\phi^2_{\delta\mathscr{B}f(j)} + \overset{\cdots}{\overline{Y_4X_4}}\phi^1_{B\vec{\mathscr{B}}m(i)}\phi^3_{\delta\mathscr{B}f(k)} + , \ldots, + \overset{\cdots}{\overline{Y_3X_3}}\phi^n_{B\vec{\mathscr{B}}m(n)}\phi^n_{\delta\mathscr{B}f(n)})] \quad (46)$$

Then after applying crossover operators the population of the broods can be indicated as

$$\phi^1_{\delta\mathscr{B}f}(\mathscr{B}) = \left[\left(\phi^1_{\mathscr{B}_{b(i)}} + \phi^2_{\mathscr{B}_{b(j)}} + \phi^3_{b(k)} + , \ldots, + \phi^n_{b(n)}\right)\right] \quad (47)$$

The best broods in the population of broods can be sorted according to their fitness as

$$\phi^1_{\delta\mathscr{B}f}(B)_{best} = \left[\left(\phi^1_{\mathscr{B}_{b(i)}} + \phi^2_{\mathscr{B}_{b(j)}} + \phi^3_{b(k)} + , \ldots, + \phi^n_{b(n)}\right)\right] \quad (48)$$

A new generated brood, which has a great probability to compare with the existing parent birds can be numerically written as

$$\phi^1_{\delta\mathscr{B}f(\mathscr{B}_{b(i)})} = \phi^1_{\mathscr{B}_{b(i)}} + \omega \times (\phi^2_{\mathscr{B}_{b(j)}} - \phi^3_{\mathscr{B}_{b(k)}})(\phi^4_{\mathscr{B}_{b(l)}} - \phi^5_{b(e)}) \quad i = 1, 2, \ldots, N_{\mathscr{B}} \text{and} \omega \in rand[0, 1] \quad (49)$$

Phase6: Fitness Function

The fitness function interprets the chromosome in terms of physical representation and evaluates its fitness to identify how good a potential solution is relative to other potential solutions. It is responsible to perform this evaluation and return a fitness value that reflects the performance of an individual with regard to the current optimum, so that different individuals can be compared. This fitness value is used in the process of selection to choose which potential solutions with better quality will survive and continue on to the next generation. However, the designed fitness function must guarantee that the individuals can be differentiated according to their suitability for solving optimization problems by accurately measuring the quality of the chromosomes in the population. Therefore, the definition of the fitness function is very critical in most of optimization scenarios. Usually, an extent of solutions exists, ranging in fitness from extremely poor to good where the chance of survival is higher for good fitness values. In this scheme, in each generation a fittest bird (sensor node) receives a higher fitness value and thus has a higher chance of surviving in the resulting generation. Here, the better fitness function means the low clustering cost, which includes the stable link quality for reliable data transmission with low energy consumption as well as proficient in load balancing. Therefore, the probability of fitness function that involves computational efficiency and accuracy at low cost for a cluster C_i is defined as follows

$$\phi_{\rho(\mathcal{C}_j)} = \psi_{\mathcal{C}_j} / \sum_{i=1}^{n} \psi_{\mathcal{C}_i} \quad (50)$$

in which $\phi_{\rho(\mathcal{C}_j)}$ is the probability of a cluster $\mathcal{C}_j$, $\psi_{\mathcal{C}_j}$ is the is the fitness value of an individual cluster $\mathcal{C}_j$, and $\sum_{i=1}^{n}\psi_{\mathcal{C}_i}$ is the fitness value of a set of clusters $j \in [1, 2, \ldots, n]$ in the smart grid. In our scheme, the probability of an individual sensor node i being selected as a member in a cluster $\mathcal{C}_j$ with the highest fitness value is given as:

$$\phi_{(Sn_j)} = \psi_{Sn_j} / \sum_{i=1}^{n} \psi_{Sn_i} \quad (51)$$

where $\phi_{(Sn_j)}$ is the probability of an individual (sensor node) Sn_j, ψ_{Sn_j} is the fitness value of an individual Sn_j, and $\sum_{i=1}^{n}\psi_{Sn_i}$ is the set of sensor nodes in the smart grid.

Phase7: Termination criterion

In the last phase, the new generated broods population is ordered according to the calculated fitness values and a superior brood is selected based on its rank in each region R_i. Thus, the new generated best broods replace the worst parents having low fitness values until no new broods that are better than any of the parent in the current population of individuals. The remaining best broods are then marked for the new mating process by ensuring that the broods with lower fitness values have been slayed. This helps to minimize computation and energy consumption cost, and delay in the network. If the improved solution belongs to an individual in a region R_i cannot be found in a given time interval or over a predefined number of cycles then that individual is supposed to be abandoned. Then, the existing best bird solution will be considered as a better solution in a region R_i. This entire process (from phase4 to phase7) repeats iteratively until all defined mating process is completed or a convergence criterion is satisfied. Consequently, by modifying the above Eq. (48) and Eq. (49), we can conclude for the society birds as:

$$\phi^1_{\mathscr{B}^i}(B)_{best} = \left[\left(\mathscr{B}_m\phi^1_{\mathscr{B}_{b(i)}} + \mathscr{B}_{pg}\phi^2_{\mathscr{B}_{b(j)}} + \mathscr{B}_{pc}\phi^3_{b(k)} + \mathscr{B}_{pa}\phi^2_{\mathscr{B}_{b(j)}} + \mathscr{B}_{pag}\phi^3_{b(k)} + , \ldots, + \phi^n_{b(n)}\right)\right] \quad (52)$$

The decision variables to select the best bird$i$for the current round j from the current population in the entire network region $R_{1,2,3,\ldots,n}$ can be expressed as:

$$\mathscr{B}^j_i(n) = \begin{cases} \mathscr{B}^j_{i(new)} & \mathscr{B}^j_{b(i)(previous)} < \phi^{i,\ldots,n}_{\mathscr{B}_i} \\ \mathscr{B}^j_{i(previous)} & otherwise \end{cases} \quad (53)$$

in which $\mathscr{B}^j_{i(new)}$is the new generated brood,$\mathscr{B}^j_{b(i)(previous)}$ is the old brood and $\phi^{i,\ldots,n}_{\mathscr{B}_i}$ is the sum of broods. Rewriting the Eq. (52), for the current solution based on the above mentioned decision variables in Eqs. (1), (2) and (3), we have the following equation.

$$\phi^1_{\mathcal{B}^i}\left(\mathscr{B}_{parent} \| \mathscr{B}_{brood}\right)_{best} = \left[\left(\mathcal{B}_m\phi^1_{\mathcal{B}_{b(i)}} \cup \mathscr{B}_{pg}\phi^2_{\mathscr{B}_{b(j)}} \cup \mathscr{B}_{pc}\phi^3_{b(k)} \cup \mathscr{B}_{pa}\phi^2_{\mathscr{B}_{b(j)}} \cup \mathscr{B}_{pag}\phi^3_{b(k)} + , \ldots, + \phi^n_{b(n)}\right)\right] \quad (54)$$

Please cite this article in press as: M. Faheem, V.C. Gungor, Energy efficient and QoS-aware routing protocol for wireless sensor network-based smart grid applications in the context of industry 4.0, Appl. Soft Comput. J. (2017), http://dx.doi.org/10.1016/j.asoc.2017.07.045

in which $\mathcal{B}_m$, $\mathcal{B}_{pg}$, $\mathcal{B}_{pc}$, $\mathcal{B}_{pa}$ and $\mathcal{B}_{pag}$, are the parent birds while $\phi^1_{\mathcal{B}_{b(i)}}$, $\phi^2_{\mathcal{B}_{b(j)}}$, $\phi^3_{b(k)}$, $\phi^2_{\mathcal{B}_{b(j)}}$ and $\phi^3_{b(k)}$, are the resultant broods. After a best bird $\mathcal{B}_i$(CH) selection in a region R_i, it broadcasts association message to its neighboring birds (sensor nodes) for completing clustering process by taking into account the CSMA mechanism. The sensor nodes having high residual energy and small distance joined the neighboring cluster heads by sending cluster head joining (*msg_ch_join*) message. Herein, each cluster head in region R_i assigns a unique time to its member nodes using TDMA (Time Division Multiple Access) mechanism in the smart grid.

Now, we briefly address the routing problem where multiple data path routing optimization problem is decomposed into several single path routing optimization problems. In our approach, cluster heads nodes form a network among themselves where each fittest bird (CH) in the initial population corresponds to a valid routing solution. Here, our main objective is to figure out a chain like optimal route from a starting location (origin) to an end location (sink). The developed scheme selects the next-hop relay node in such a way that it requires less energy consumption by offering distributed data paths. Further, it considers the node over-assignment cost and loop free delay aware data path between each set of pair of relay nodes in harsh SG environment with obstacles under various kinds of constraints such as noise, interference, etc. The construction of the initial routing is based on the measured angles (*AOT*, *AOR*), residual energy, hop-to-hop distance (E_d), and over assignment cost (*OAC*), information of the relay nodes as well as the total length of all the routing (*MOD*), without repetition in routing table. Initially, throughout the routing process each relay node creates a list (i.e., routing table), N_i, $1 \le i \le n$ that holds all the one-hop neighbors j such that the link $i \to j, \forall j \in N_i$ satisfy the hop restriction that $Hop(i) < Hop(j)$, can be used to route data packets from*i*through j towards the destination in the chain like structure. Based on the given information for the potential routing solutions (best bird, i.e., the best next hop relay node), an initial population of individuals (CHs) is generated using a greedy approach and fitness of each individual is measured. Then, based on the fitness function value each individual in ranked in the bird list from decreasing to increasing order. The best bird (i.e., the best next hop relay node)*i*for the current round j is selected in greedy manner from the society, the decision variable for choosing the best bird can be expressed as:

$$R\mathcal{Q}^j_i(R_i) = \begin{cases} \mathcal{B}^j_{i(new)} & \mathcal{B}^j_{i(new)} > \phi^i_{\mathcal{B}_i(List)} \\ 0 & otherwise \end{cases} \tag{55}$$

Herein, we used a non-uniform crossover probability called as decreasing crossover probability ($P_c = 0.98$, $P_c = 0.95$), on average crossover in each individual generation are given as $P_{c_1} = 100x0.98 = 98$, and$P_{c_2} = 100x0.95 = 95$ and mutation operator with the probability of ($P_m = 0.001$)to perform the mutation operation with the assertion that no important genetic material is lost. Here, it is important to note that this crossover probability varies over time, which means that in the beginning probability is assumed to be very high and decreased with the number of relay nodes die in the network. This mechanism leads to more robust and accurate routing solutions in the harsh smart grid environment.

In nature, a bird is represented as a genotype marker, which can be randomly employed to mark half of the genes by ensuring that the rest of the leaving genes are half unmarked. In our illustration, only those unmarked genes that create sperms are arbitrarily used in the mating process. After mating, each bird is responsible to take care of the brood and among the available the best brood is selected and compared with the existing bird. The decision variable for selecting a best bird (i.e., relay node)*i* as a valid routing solution for the current round j from the population randomly between 0 and 1 in region R_i can be expressed as:

$$R\mathcal{B}^j_i(R_i) = \begin{cases} \mathcal{B}^j_{i(new)} & \mathcal{B}^j_{i(previous)} < \phi^i_{\mathcal{B}_i} \\ \mathcal{B}^j_{i(previous)} & otherwise \end{cases} \tag{56}$$

This entire process repeats until the defined criteria satisfied or best solution becomes abandoned. The sum of the fitness function used in routing can be numerically indicated as:

$$F_{\mathcal{B}_i} = F_{1\mathcal{B}_i} + F_{2\mathcal{B}_i} + F_{3\mathcal{B}_i} + F_{4\mathcal{B}_i} + F_{5\mathcal{B}_i} \tag{57}$$

$$F_{1\mathcal{B}_i} = Fit_{\mathcal{B}_i} / \sum_{j=1}^{n} Fit_{\mathcal{B}_j}(OAC) \tag{58}$$

$$F_{2\mathcal{B}_i} = Fit_{\mathcal{B}_i} / \sum_{j=1}^{n} Fit_{\mathcal{B}_j}(E_d) \tag{59}$$

$$F_{3\mathcal{B}_i} = Fit_{\mathcal{B}_i} / \sum_{j=1}^{n} Fit_{\mathcal{B}_j}(AOT + AOR), \forall RN_i(\theta) \in {}^{2\pi}/_{3} \| \pi \tag{60}$$

$$F_{4\mathcal{B}_i} = \mathrm{Fit}_{v_i} / \sum_{j=1}^{n} Fit_{v_j}(MOD), DT_{MOD} \ge \| \cong CT_{MOD} \tag{61}$$

$$F_{5\mathcal{B}_i} = Fit_{\mathcal{B}_i} / \sum_{j=1}^{n} Fit_{\mathcal{B}_j}(E), whereE = \frac{E_{rd}}{E_{max}} \in [1, 0], \quad \forall RN_i \in \mathcal{TP}_{low} \| \mathcal{TP}_{high} \tag{62}$$

Moreover, in routing table each entry is assigned a calculated probability in order to find an optimal routing solution in robust manner can be expressed as:

$$P_{ij} = \frac{F_{\mathcal{B}_i}}{\sum_{Rd \in R}^{E_{Rd}} / F_{\mathcal{B}_j}} \tag{63}$$

In EQRP, for measuring CH_i over-assignment cost, the entire network is divided into internal and external traffic by considering M/M/1 and M/M/N queuing models as discussed in [32], respectively. The sensor nodes in the same cluster generate internal traffic while the external data traffic cost is from the other cluster heads can be numerically indicated as:

$$\mathcal{C}_i = \sum_{r=1}^{n} \mathcal{C}_j \cdot \lambda_r (1 - P_b) \tag{64}$$

where λ_r is the packets arrival rate from a neighboring cluster $\mathcal{C}_j$ with probability P_b. The packets arrival rate from k number of clusters to a cluster $\mathcal{C}_i$ can be indicated as:

$$\mathcal{C}_i = \sum_{k=1}^{n} \sum_{r=1}^{n} \mathcal{C}_k \cdot \lambda_r (1 - P_b), \quad k \in 1, 2, \ldots, n \tag{65}$$

4. Network model

In this section, we provide a definition of the considered smart grid environment where a power grid station under 500KV is considered as the local power distribution area (LPDA). The primary substation including a base station and primary data management center reveals an important role for regional power management and controls substations. A power distribution network in which fewer than 300 transmission grids are involved is considered to test the effectiveness of proposed scheme. The distance between

Please cite this article in press as: M. Faheem, V.C. Gungor, Energy efficient and QoS-aware routing protocol for wireless sensor network-based smart grid applications in the context of industry 4.0, Appl. Soft Comput. J. (2017), http://dx.doi.org/10.1016/j.asoc.2017.07.045

each individual transmission grid is assumed to be 15–20 meters in local power distribution area. In this scenario, the performance of the randomly deployed wireless communication among the sensor nodes located in the transmission area brings an ultimate challenge to provide more reliable, scalable and flexible cost-effective solution with quality and accuracy by providing timely fault detection. The deployed sensor network is divided into basic three layers. The lower layer consists of sensor nodes and cluster heads. In the lower layer, nodes transmit observed data to their cluster heads. Cluster heads pass the data to the next hop cluster head relay nodes and this process continues till the data is transmitted to the sink at the middle layer. The middle layer is comprised of a sink which is connected to the internet. The IoT considers the various objects installed at the application layer. This allows a user to interact and collect the data from various regions, and must have unlimited power supply. Therefore, the upper layer includes the base stations which are interconnected and have unlimited power. In our performance evaluations, we consider some assumptions given as follows: First, we assumed that all sensor nodes are deployed in an ad hoc manner and are incapable to change their position. Second, all deployed sensor nodes are aware of their positions which can be estimated using localization mechanism presented in [33]. Third, each randomly deployed sensor node has the same level of initial energy. Fourth, omnidirectional antenna model is considered. Fifth, we assumed that the arrival time follows Poisson distribution while the service time is considered exponentially distributed.

4.1. Path loss model

In the current study, we used log-normal shadowing path-loss model and smart grid channel model parameters presented in [34]. The main reason to adopt this model is due to its more accurate multipath channel estimation results compared to the Nakagami and Rayleigh models in a smart grid environment. The mathematical form of the log normal shadowing model can be numerically expressed as:

$$\gamma(d)_{dB} = P_t - PL(d_0) - 10\eta \log_{10}\frac{d}{d_0} - X_\sigma - P_\eta \tag{66}$$

4.2. Simulation parameters

The network simulation tool called EstiNet9.0 [35] is used to evaluate the performance of this study. In our simulation study, we consider the total number of 300 sensor nodes equipped with the physical layer standard IEEE802.11 g deployed in 2D(length × width) with values (550 ×700) meters. The performance evaluations consist of 51sets of simulations. The further parameters used in this simulation are given in the following Table 2.

Table 2
Simulation and log normal shadowing model parameters values and description.

Simulation Model Parameters	Values
Initial sensor node energy	2.3 J
Maximum hop distance	80 m
High transmission power	0.87 W
Low transmission power	0.81 W
Packet receiving power	0.035 watts
Ideal listening	0.013watts
Sleeping power	3×10^{-6} W
Data aggregation	0.011 W
Antenna Beam width (Omni direction)	360°
Packet length	43bytes
Transmission grid distance	12 to 15 m
Antenna Beam width	360°
Topology	random
θ_{low}	$2\pi/3$
θ_{high}	π
Path loss Model Parameters	**Values**
Grid Station (outdoor)	500 kV
Noise power measured in dBm	P_η
Transmit power in dBm	P_t
Path-loss exponent	η
Shadowing deviation (σ) for both LOS, NLOS	3.12, 2.95
Signal to noise ratio	γ(d)
Path loss exponent (n) for both LOS, NLOS	2.4
Zero mean Gaussian random variable with standard deviation	X_σ
Line of sight	LOS
Non-line of sight	NLOS
Noise floor for both LOS, NLOS	−88, −93
Path loss at reference distance d_o	PL (d_O)
Shadowing deviation (σ) for both LOS, NLOS	3.12, 2.95
BMO Model Parameters	**Values**
Number of monogamous	130
Number of polygynous	80
Number of promiscuous	50
Number of parthenogenetic	25
Number of polygynous	15
Maxi. capacity of spermatheca	10, 4
Allure factor of the birds	[1,0]
The breeding ratio	0.90,0.87,0.85
Mutation variation	0.001
Random selection values (rand)	[0,1]
Maxi. number of iterations	15

5. Simulation results and discussion

The smart electricity industry is expected to be a game changer for the developing countries, in the growth of their power sector to be more manageable, reliable, and scalable. In this section, we present the effectiveness of our proposed scheme against the existing routing schemes, such as ETL-AODV [13] and HLR-AODV [14] by considering various metrics in the smart grid environment which are defined below:

5.1. Packet delivery ratio (PDR)

In our experimental studies, packet delivery ratio can be numerically as:

$$PDR = \frac{DP_{received}}{DP_{send}} \tag{67}$$

where DP_{send} and $DP_{received}$ are the total number of data packets send by a source node and data packets received successfully at the sink, respectively.

5.2. Delay(D_e)

Delay can be numerically defined as:

$$T_{avg}^{P_i} = DP_{Sn_i}^{Source} + T\sum_{i=1}^{n} RDP + DP_s \tag{68}$$

in which $T_{avg}^{P_i}$, $DP_{Sn_i}^{Source}$, $T\sum_{i=1}^{n} RDP$ and DP_{sink} are, the average time to send a data or request packet, the time when a source sensor node Sn_i generates data or request packet, the time(T) taken by relaying data packet (RDP), and the time when data packet (DP) reached successfully at the sink (s), respectively.

5.3. Residual energy (R_e)

It is the amount of remaining energy after energy consumed during successful transmission of data packets by the nodes.

Please cite this article in press as: M. Faheem, V.C. Gungor, Energy efficient and QoS-aware routing protocol for wireless sensor network-based smart grid applications in the context of industry 4.0, Appl. Soft Comput. J. (2017), http://dx.doi.org/10.1016/j.asoc.2017.07.045

5.4. Throughput (T_p)

It is the number of data packets processed in a given amount of time (bits per second) at each sensor node and successfully transmitted to the sink.

During experimental studies, it is revealed that PDR in all routing schemes rapidly increases between round numbers 1 and 3000 as shown in Fig. 1(a). Initially, between round numbers 1 and 500 when network size is small, the PDR performance of HRL-AODV is observed superior compared to ETL-AODV routing scheme in the smart grid. However, the PDR of ETL-AODV becomes more superior compared to HRL-AODV routing scheme between round numbers 700 and 1500 and similar behavior is observed between round numbers 1600 and 3000. In general, the PDR performance of ETL-AODV is observed to be higher compared to the HRL-AODV routing scheme. Herein, it is witnessed that the PDR performance of the proposed EQRP is higher compared to both ETL-AODV and HRL-AODV routing scheme. This also leads to high network throughput of EQRP in the smart grid as shown in Fig. 1(b). There are a number of possible reasons which lead to low performance of both ETL-AODV and HRL-AODV routing schemes in the smart grid. First, one of the main challenging issues in both ETL-AODV and HRL-AODV routing schemes are their unstable link quality between each set of pair of nodes, since they are not considering the excessive interference in the smart grid. This brings data packet collision and data refutation issues due to extreme RF interference in the smart grid. Second, the data path looping in these routing schemes is another challenging issue causing packet invalidations in the smart grid. Therefore, a notable amount of packets generated becomes invalid because of not reaching at the sink/end user in a timely manner. Third, in most of the simulations HRL-AODV failed to find an appropriate next hop relay node having the minimum communication distance to convey data packets in the smart grid. If an appropriate next hop relay node is not found during conveying information then it simply stops to forward data packets towards the sink/end user. Thus, a significant amount of valuable data packets generated becomes invalid which consumes limited node energy resources rapidly for finding another appropriate relay node in the smart grid.

Fig. 1(c) shows that the delay profile in all routing schemes. It clearly shows that the delay rapidly increases between sensor nodes 1 and 100 and decreases when new sensor nodes between 100 and 200 are added to the network. When network size is small between 1 and 50 sensor nodes, the better delay performance of ETL-AODV is observed compared to the HLR-AODV. However, when the network size becomes large between 60 and 100 sensor nodes, the low delay performance of HLR-AODV is observed more significant compared to ETL-AODV routing scheme in the smart grid. Here, we have noticed that the delay performance of EQRP is outstanding in terms of small, medium and large size network compared to both ETL-AODV and HRL-AODV routing schemes in the smart grid. This is due to its adaptability network dynamics when finding optimal routes towards the sink. Moreover, it takes into account the memory overrun issues during forwarding data or request packets to the next hop node in appropriate way compared to ETL-AODV and HRL-AODV routing schemes as shown in Fig. 1 (d). Thus, it minimizes the data packet loss due to node buffer overflow. There are a few reasons which may cause increasing delay in both ETL-AODV and HRL-AODV routing schemes in the smart grid. First, both routing schemes are suffering from effective and efficient routing table management mechanism in terms of updating and finding optimal route to convey information. Second, both routing schemes are facing the problem of data path loops in the smart grid. Third, unnecessary multi-hop data packet transmissions also increase the network delay, which is found more in ETL-AODV compared to the HRL-AODV.

Fig. 1(e) shows that the residual energy in each scheme decreases rapidly with the increase in round numbers between 1 and 3000. At the beginning, when network size is small between round numbers 1 and 100, the performance of ETL-AODV is observed better for achieving a high level of residual energy profile compared to the HRL-AODV. When the network size grows up between round numbers 800 and 1400 and then between 1600 and 3000, the performance of minimizing energy consumption is found superior in ETL-AODV leading to high network residual energy compared to the HRL-AODV routing scheme. In the meanwhile, we observed that the performance of high residual energy in EQRP is found superior in terms of small, medium and large network size compared to all other routing schemes in the smart grid. This is due to its low energy consumption profile at each node leading to high network lifetime time as shown in Fig. 1(f). During experimental studies, we observed a number of possible reasons which may result in increasing energy consumption in both HRL-AODV and ETL-AODV routing schemes are: As mentioned earlier, one of the major challenging issues in both routing protocols is their lack of considering the efficient memory utilization during forwarding packets to the next hop relay node in the smart grid. Thus, a number of data packets lost due to memory over run issue which consumes a significant amount of network energy during retransmitting the same data packets with great accuracy in the network. Second, poor link quality in HRL-AODV compared to ETL-AODV brings several new challenges such as data packet collision and data refutation problems due to extreme high ecological interference in the network. Thus, a significant amount of energy is consumed when these corrupted packets travel through the network at each individual node in terms of receiving, processing and forwarding. Third, data path looping and unnecessary multi-hop data packet transmission is another challenging issue, which also consumes a signification of energy at each individual node during receiving and sending data packets in the network.

In EQRP, to attain high PDR profile, clustering mechanism offers a robust and stable load balancing architecture by providing reliable link quality among each pair of nodes in the smart grid. This provides strength to our designed scheme for robust data gathering at extremely low interference in the smart grid. As a result, small numbers of corrupted data packets are generated compared to other routing schemes in the smart grid. To provide robust data packet delivery from the source node towards the sink, the behavior of routing mechanism is remarkable in EQRP. The designed routing mechanism is responsible to route data packets by offering a robust rotating combinational structure of the cluster heads relay nodes in the chain like network. In a chain, a next hop relay cluster head node is selected based on its node assignment cost, inter distance and highly stable link quality based on distance information in a greedy manner in the smart grid. Moreover, the role of the AOT, AOR and MOD is remarkable and provides loop free robust data delivery from the source node towards the sink in the smart grid. In addition, the use of AOT and AOR guarantees that the data packets always movies towards the sink over minimum zonal area by exploiting location information and switching technique (min. and max. angle) of the relay nodes in the smart grid. Furthermore, in EQRP each relay node is equipped with dual transmission power mechanism to deal with the dense and sparse network deployment in the smart grid. In sparse network deployment, if a next hop relay node is not found in the range with the predefined probability then a sending relay node can choose an appropriate next hop relay node by utilizing its minimum angle information with higher transmission power. However, in the worst case, it can also utilize its high angle information with high transmission power for robust data delivery in the smart grid. During routing, the MOD behaves as a double check mechanism to avoid data path loops in the network. If the measured data path from the source node

Please cite this article in press as: M. Faheem, V.C. Gungor, Energy efficient and QoS-aware routing protocol for wireless sensor network-based smart grid applications in the context of industry 4.0, Appl. Soft Comput. J. (2017), http://dx.doi.org/10.1016/j.asoc.2017.07.045

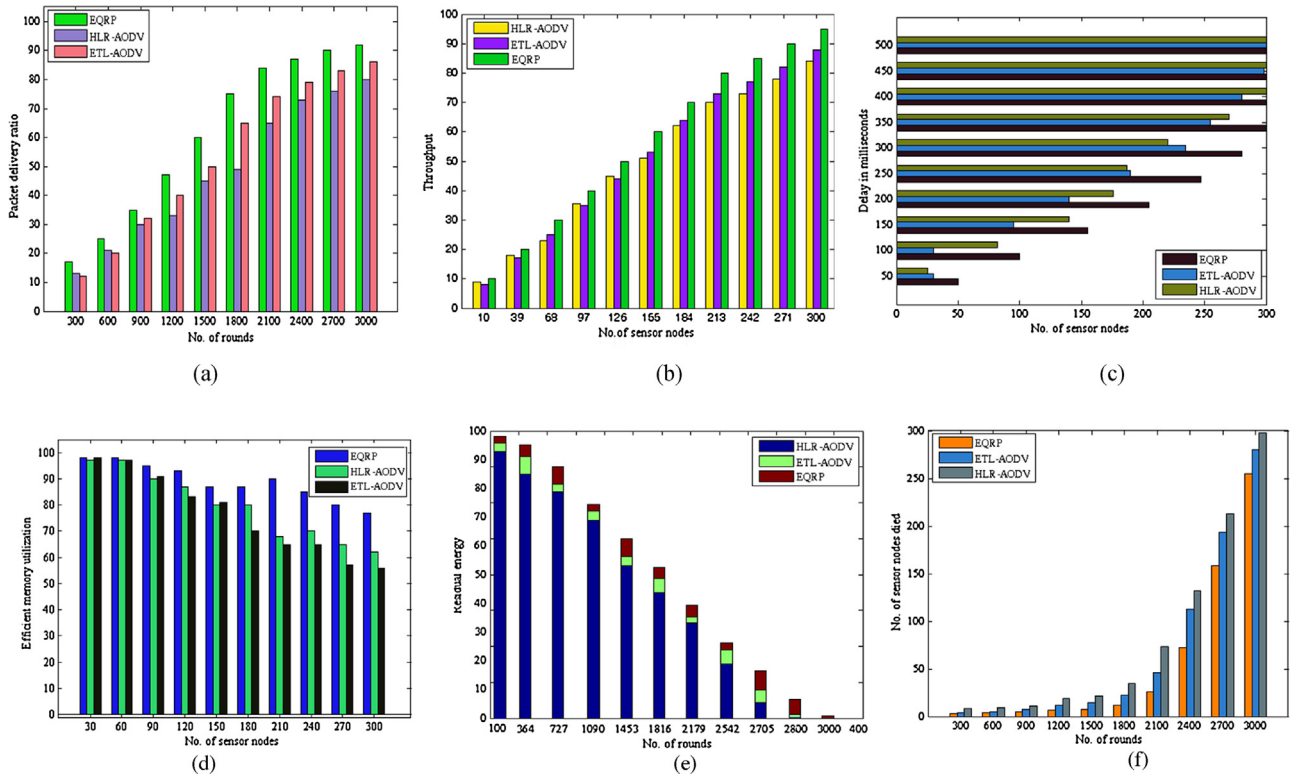

Fig. 1. (a) shows the packet delivery ratio vs total number of rounds between 1 and 3000 (b) indicates the throughput vs sensor nodes between 1 and 300 (c) denotes the delay vs number of sensor nodes between 1 and 300 (d) represents the efficient memory utilization vs sensor nodes between 1 and 300 (e) depicts the residual energy vs number of rounds between 1 and 300 and finally (f) illustrates the number of sensor nodes die vs number of rounds between 1 and 3000, respectively.

towards the sink is greater than the defined distance threshold value then entire chain formation repeats until the data path less than defined threshold is obtained. This entire mechanism guarantees that there is no path loop exit in the network. Thus, data packet generated in the network reached at the sink/end user in a timely manner leading to high PDR in the smart grid. In addition, routing table at each relay node plays an important role for robust data delivery in the network. Since each entry in the routing table is assigned a certain probability selection and listed in decreasing order, therefore after receiving information each relay node immediately looks into its routing table and selects an appropriate next hop relay node with highest probability. To cope with route failure issues, the designed routing scheme is equipped with a backup route finding mechanism. Thus, in a case of a route node failure a relay node with second higher priority in the routing table is selected to route information by avoiding delay in the smart grid. To sum up, our results show that the proposed protocol has successfully minimized the end-to-end delay and has improved packet delivery ratio, memory utilization, residual energy, and throughput with the expense of computational complexity. Thus, the designed scheme helps end users to achieve its effective control objectives in the smart grid. This promotes the use of the proposed approach in electric power industry for effectiveness and efficiency of smart grid operations.

6. Conclusion

Recently, the proliferation of Internet of things (IoT) and wireless sensor networks (WSNs) introduces the fourth stage of industrialization, commonly known as Industry 4.0. One of the key features of Industry 4.0, is the wireless integration of various components within a factory to implement a flexible and reconfigurable manufacturing system for increasing manufacturing productivity. The Smart Grid Industry 4.0 (SGI 4.0) has recently been embracing the advances in WSNs as a promising candidate for efficient monitoring of the smart grid, given their capability to cover large geographic regions at low-cost. However, for an efficient SGI 4.0, reliable WSNs-based communication framework is essential. In this paper, we proposed a novel dynamic clustering based energy efficient and QoS-aware routing protocol inspired by the real behavior of the bird mating optimization (BMO) for smart grid applications. The proposed distributed scheme improves network reliability and reduces excessive packets retransmissions in both sparse and dense network deployment. We evaluate the proposed scheme by using a network simulation tool called EstiNet9.0 based on smart grid field measurements. Our results show that the proposed protocol has successfully minimized the end-to-end delay and has improved packet delivery ratio, memory utilization, residual energy, and throughput. To sum up, the proposed scheme is validated in comparison to the traditional WSN schemes and found to be more favored for various applications of SGI 4.0.

As for the future work, one direction is to further explore opportunities to improve the latency for robust data delivery of our routing scheme by studying the effects of mobile sinks. Another direction is to introduce new swarm intelligence based cross-layer communication framework equipped with parallel computation characteristics in order to provide robust data delivery for WSNs-based smart grid applications.

Acknowledgement

The work of V.C. Gungor was supported by the Turkish National Academy of Sciences Distinguished Young Scientist Award Program (TUBA-GEBIP) under Grand No: V.G./TUBA-GEBIP/2013-14.

Please cite this article in press as: M. Faheem, V.C. Gungor, Energy efficient and QoS-aware routing protocol for wireless sensor network-based smart grid applications in the context of industry 4.0, Appl. Soft Comput. J. (2017), http://dx.doi.org/10.1016/j.asoc.2017.07.045

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.asoc.2017.07.045.

References

[1] C.-C. Lin, D.-J. Deng, Z.-Y. Chen, K.-C. Chen, Key design of driving industry 4.0: joint energy-efficient deployment and scheduling in group-based industrial wireless sensor networks, IEEE Commun. Mag. 54 (2016) 46–52.
[2] J. Wan, S. Tang, Z. Shu, D. Li, S. Wang, M. Imran, et al., Software-defined industrial Internet of Things in the context of Industry 4.0, IEEE Sens. J. 16 (2016) 7373–7380.
[3] L. Da Xu, W. He, S. Li, Internet of things in industries: a survey, IEEE Trans. Ind. Inf. 10 (2014) 2233–2243.
[4] M.R. Palattella, N. Accettura, X. Vilajosana, T. Watteyne, L.A. Grieco, G. Boggia, et al., Standardized protocol stack for the internet of (important) things, IEEE Commun. Surv. Tutorials 15 (2013) 1389–1406.
[5] A. Mazak, C. Huemer, A standards framework for value networks in the context of Industry 4.0, IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), 2015 (2015) 1342–1346.
[6] S. Shamim, S. Cang, H. Yu, Y. Li, Management approaches for Industry 4.0: A human resource management perspective, IEEE Congress on Evolutionary Computation (CEC), 2016 (2016) 5309–5316.
[7] V.C. Gungor, D. Sahin, T. Kocak, S. Ergut, C. Buccella, C. Cecati, et al., Smart grid technologies: communication technologies and standards, IEEE Transactions on Industrial Informatics vol. 7 (2011) 529–539.
[8] V.C. Gungor, D. Sahin, T. Kocak, S. Ergut, C. Buccella, C. Cecati, et al., A survey on smart grid potential applications and communication requirements, IEEE Transactions on Industrial Informatics vol. 9 (2013) 28–42.
[9] V.C. Gungor, M.K. Korkmaz, Wireless link-quality estimation in smart grid environments, Int. J. Distrib. Sens. Netw. 2012 (2012).
[10] G.A. Shah, V.C. Gungor, O.B. Akan, A cross-layer QoS-aware communication framework in cognitive radio sensor networks for smart grid applications, IEEE Transactions on Industrial Informatics vol. 9 (2013) 1477–1485.
[11] S. Kim, Biform game based cognitive radio scheme for smart grid communications, J. Commun. Networks 14 (2012) 614–618.
[12] K. Kim, S. -i. Jin, Branch-based centralized data collection for smart grids using wireless sensor networks, Sensors 15 (2015) 11854–11872.
[13] H. Farooq, L. Tang Jung, Energy, traffic load, and link quality aware Ad Hoc routing protocol for wireless sensor network based smart metering infrastructure, Int. J. Distrib. Sens. Netw. 2013 (2013).
[14] H. Farooq, L.T. Jung, Health, link quality and reputation aware routing protocol (HLR-AODV) for Wireless Sensor Network in Smart Power Grid, 2012 International Conference on Computer & Information Science (ICCIS) (2012) 664–669.
[15] X. Deng, L. He, X. Li, Q. Liu, L. Cai, Z. Chen, A reliable QoS-aware routing scheme for neighbor area network in smart grid, Peer-to-Peer Netw. Appl. (2015) 1–12.
[16] R. Hou, C. Wang, Q. Zhu, J. Li, Interference-aware QoS multicast routing for smart grid, Ad Hoc Netw. 22 (2014) 13–26.
[17] M. Faheem, M.Z. Abbas, G. Tuna, V.C. Gungor, EDHRP: Energy efficient event driven hybrid routing protocol for densely deployed wireless sensor networks, J. Netw. Comput. Appl. 58 (2015) 309–326.
[18] M. Faheem, G. Tuna, V.C. Gungor, LRP: Link quality-aware queue-based spectral clustering routing protocol for underwater acoustic sensor networks, Int. J. Commun. Syst. (2016).
[19] M.E. Haque, U. Baroudi, Energy efficient routing scheme using leader election in ambient energy harvesting wireless ad-hoc and sensor networks, Sensors, 2015 Ieee (2015) 1–4.
[20] H. WANG, Z. GUAN, T. YANG, Y. XU, Top-K query framework in wireless sensor networks for smart grid, China Commun. 6 (2014) 011.
[21] H. Hong, P. Bao-ming, Y. Dong-sheng, L. Xiang-ze, A routing decision algorithm based on the trunk link for smart city, Control and Decision Conference (CCDC), 2016 Chinese (2016) 6305–6310.
[22] Runliang Dou, N. Guofang, Optimizing Sensor Network Coverage and Regional Connectivity in Industrial IoT Systems, 2015.
[23] A. Askarzadeh, Bird mating optimizer: an optimization algorithm inspired by bird mating strategies, Commun. Nonlin. Sci. Numer. Simulat. 19 (2014) 1213–1228.
[24] N. Behera, A. Routray, J. Nayak, H. Behera, Bird mating optimization based multilayer perceptron for diseases classification, in: Computational Intelligence in Data Mining-Volume, 3 ed., Springer, 2015, 2017, pp. 305–315.
[25] Q. Zhang, G. Yu, H. Song, A hybrid bird mating optimizer algorithm with teaching-learning-based optimization for global numerical optimization, Stat. Optim. Inf. Comput. 3 (2015) 54–65.
[26] Runliang Dou, C. Zong, M. Li, An interactive genetic algorithm with the interval arithmetic based on hesitation and its application to achieve customer collaborative product configuration design, Appl. Soft Comput. 38 (2016) 384–394.
[27] Runliang Dou, Chao Zong, M. Li, Application of an interactive genetic algorithm in the conceptual design of car console, WSEAS Transactions on Systems vol. 13 (2014) 596–605 (Art. #59).
[28] T.W. Liao, C.-F. Ting, P.-C. Chang, An adaptive genetic clustering method for exploratory mining of feature vector and time series data, Int. J. Prod. Res. 44 (2006) 2731–2748.
[29] F. Tao, Y. Feng, L. Zhang, T.W. Liao, CLPS-GA: A case library and Pareto solution-based hybrid genetic algorithm for energy-aware cloud service scheduling, Appl. Soft Comput. 19 (2014) 264–279.
[30] C.-F. Chien, F.-P. Tseng, C.-H. Chen, An evolutionary approach to rehabilitation patient scheduling: a case study, Eur. J. Oper. Res. 189 (2008) 1234–1253.
[31] C.-F. Chien, Y.-C. Huang, C.-H. Hu, A hybrid approach of data mining and genetic algorithms for rehabilitation scheduling, Int. J. Manuf. Technol. Manage. 16 (2008) 76–100.
[32] A.O. Allen, Probability, Statistics, and Queueing Theory, Academic Press, 2014, 2017.
[33] S. Pandey, S. Varma, A range based localization system in multihop wireless sensor networks: a distributed cooperative approach, Wireless Person. Commun. 86 (2016) 615–634.
[34] V.C. Gungor, B. Lu, G.P. Hancke, Opportunities and challenges of wireless sensor networks in smart grid, IEEE Transactions on Industrial Electronics vol. 57 (2010) 3557–3564.
[35] http://www.estinet.com/.

Muhammad Faheem received the B.Sc. Computer Engineering degree in 2010 from the Department of Computer Engineering at the University College of Engineering & Technology, Bahauddin Zakariya University Multan, Pakistan. In 2012, he received an MS degree in Computer Science from the Faculty of Computer Science and Information System at Universiti Teknologi Malaysia. Currently, he is a Ph.D. student at Abdullah Gul University, Kayseri, Turkey. His research interest includes the areas of smart grid communications, underwater acoustic communications, wireless ad hoc and sensor networks, and cognitive radio networks.

Vehbi Cagri Gungor received his B.S. and M.S. degrees in Electrical and Electronics Engineering from Middle East Technical University, Ankara, Turkey, in 2001 and 2003, respectively. He received his Ph.D. degree in electrical and computer engineering from the Broadband and Wireless Networking Laboratory, Georgia Institute of Technology, Atlanta, GA, USA, in 2007. Currently, he is an Associate Professor and Chair of Computer Engineering Department, Abdullah Gul University (AGU), Kayseri, Turkey. His current research interests are in smart grid communications, machine-to-machine communications, next-generation wireless networks, wireless ad hoc and sensor networks, cognitive radio networks. Dr. Gungor has authored several papers in refereed journals and international conference proceedings, and has been serving as an editor, reviewer and program committee member to numerous journals and conferences in these areas. He is also the recipient of the Scientist of the Year Award (Bilim Kahramanlari Dernegi) in 2016, the Editor of the Year Award, Ad Hoc Networks Journal (Elsevier Science) in 2015, Turkish Academy of Sciences Distinguished Young Scientist Award (TUBa-GEBIP) in 2014,the IEEE Trans. on Industrial Informatics Best Paper Award in 2012, the European Union FP7 Marie Curie RG Award in 2009, Turk Telekom Research Grant Awards in 2010 and 2012, and the San-Tez Project Awards supported by Alcatel-Lucent, and the Turkish Ministry of Science, Industry and Technology in 2010.

Please cite this article in press as: M. Faheem, V.C. Gungor, Energy efficient and QoS-aware routing protocol for wireless sensor network-based smart grid applications in the context of industry 4.0, Appl. Soft Comput. J. (2017), http://dx.doi.org/10.1016/j.asoc.2017.07.045