Applied Soft Computing Journal 73 (2018) 83–95

Contents lists available at ScienceDirect

Applied Soft Computing Journal

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

A genetic algorithm based solution to the Minimum-Cost Bounded-Error Calibration Tree problem

Hüseyin Akcan

Department of Software Engineering, Izmir University of Economics, Izmir, Turkey

HIGHLIGHTS

- The MBCT problem is a spanning tree problem with two objectives.
- MBCT minimizes the spanning tree cost and bounds the maximum post-calibration skew.
- GAWES is proposed as a novel genetic algorithm based solution to the MBCT problem.
- Main novelty of GAWES is using extreme efficient solutions in the genetic algorithm.
- Experiments confirm that GAWES is superior to the state of the art.

ARTICLE INFO

Article history:
Received 11 October 2017
Received in revised form 20 June 2018
Accepted 9 August 2018
Available online xxxx

Keywords:
Genetic algorithm
Wireless sensor networks
Energy efficiency
Calibration tree

ABSTRACT

Sensors in wireless sensor networks are required to be self-calibrated periodically during their prolonged deployment periods. In calibration planning, employing intelligent algorithms are essential to optimize both the efficiency and the accuracy of calibration. The Minimum-Cost Bounded-Error Calibration Tree (*MBCT*) problem is a spanning tree problem with two objectives, minimizing the spanning tree cost and bounding the maximum post-calibration skew. The decision version of the *MBCT* problem is proven to be NP-Complete. In this paper, the *GAWES* algorithm is presented as a novel genetic algorithm based solution to the optimization version of the *MBCT* problem. *GAWES* adopts extreme efficient solution generation within the genetic algorithm to improve the search quality. It is demonstrated through experimentation that *GAWES* is superior to the existing state of the art algorithm, both in energy efficiency and calibration accuracy.

1. Introduction

Technological advances in Micro-Electro-Mechanical Systems (MEMS) have brought the network of self-configurable widely deployed sensor nodes, as wireless sensor networks, in our daily lives. Equipped with low power radios, nodes in wireless sensor networks are able to perform various sensing tasks and facilitate an ad hoc network to aggregate and extract useful data from the deployed environment. As a consequence of their flexible design and wireless operability, wireless sensor networks are capable of performing tasks that are not suitable or affordable for humans, such as remote area monitoring [1], underwater monitoring [2], and deployments in hazardous environments [3,4]. To tackle these deployment constraints, wireless sensor networks are designed to operate in a self-configurable manner where manual configuration is not a viable option. As each individual sensor unit operates on low power battery, energy efficiency is the most essential constraint for all the algorithms that need to be developed for sensor networks. The total lifetime of the sensor network depends on the lifetime of each individual sensor node in the network. Therefore, algorithms deployed on wireless sensor networks should not only use less power, but also be well distributed to avoid energy depletion on a single node.

Periodic calibration of each individual sensor is a critical problem in wireless sensor networks. As manual calibration is not an option after deployment, these sensors should self-calibrate themselves using nearby sensors as references. However, sensors need to communicate during calibration and wireless communication is one of the most energy consuming tasks for a sensor node. Therefore, an efficient and accurate self-calibration algorithm is essential for sensors that are deployed in remote areas for extended periods of time.

Calibration in wireless sensor networks poses many challenges [5–9]. A list of these challenges includes the inability to physically access sensors in most scenarios, their massive number,

E-mail address: huseyin.akcan@ieu.edu.tr.

https://doi.org/10.1016/j.asoc.2018.08.013
1568-4946/

and their energy constraints. To overcome these difficulties, researchers have proposed methods to calibrate sensors without any human intervention using peer based iterative calibration [10–14]. However, iterative calibration algorithms also introduce a new set of challenges to the sensor network community. One such major challenge is to calibrate the network to achieve minimum calibration error using minimum energy in exchange.

Minimum-Cost Bounded-Error Calibration Tree (*MBCT*) problem [15] is based on iterative calibration of nodes in wireless sensor networks. In these networks, energy usage is a tight constraint. Therefore, calibrating the sensor network by using the minimum communication cost and yet get a reasonable accuracy on the calibration is a critical problem. However, *MBCT* problem is generic enough to be applied to other domains. In its abstract form, the problem optimizes the cost of the spanning tree, as well as the cost of reaching from each vertex to the root of the tree, where each vertex has an associated cost value.

The main contribution of this paper is a novel genetic algorithm based solution to the optimization version of the *MBCT* problem. In this work, a method to find the extreme efficient solutions after the crossover stage of the proposed genetic algorithm is employed. Consequently, the search is more efficiently directed to the ideal point that minimizes both the energy usage and the calibration error. As a result, through experimentation, this paper demonstrates that the proposed algorithm outperforms the existing state of the art in terms of both energy efficiency and calibration accuracy.

The rest of the paper is organized as follows, related work is presented in Section 2. In Section 3 *MBCT* problem definition is presented. Section 4 outlines extreme efficient solution calculation. In Section 5 the proposed genetic algorithm based solution is given, and experimental results are presented in Section 6. Finally Section 7 concludes the paper.

2. Related works

Calibration in sensor networks is an essential task and each sensor needs to be calibrated periodically [6,8]. Results of real world tests for calibration are reported in [7,9]. As sensors are expected to operate prolonged amounts of time after deployment, efficient periodic calibration is a critical task for the lifetime of the network. The challenges with respect to periodic calibration in wireless sensor networks are reported in [5].

Researchers have proposed parametric calibration methods [10–14] as an alternative to traditional calibration methods. In parametric calibration, a calibration function is used to map the reported output values of the reference sensor to the input values of the current sensor. Therefore, each sensor can self-calibrate parametrically based on a presumed reference sensor nearby, without any physical interaction.

There are various representations in the literature for encoding spanning trees in evolutionary algorithms. These methods can be listed as Characteristic Vectors [16], Predecessor Coding [17], Prüfer Numbers [18], Blob Code [19], Link-and-Node Biasing [17], Network Random Keys [20,21], and Edge-Sets [22]. A comparative analysis of these methods are presented in [22] on various criteria including locality, heritability, feasibility, and time-and-space complexity. It is reported in [22] that Edge-Sets is superior to the rest of the methods listed. [22] also discusses three different methods in order to create random spanning trees for initializing the population and performing the crossover. The methods are listed as PrimRST, KruskalRST, and RandomWalkRST. PrimRST uses a modified version of Prim's spanning tree algorithm and is biased towards creating star topologies. RandomWalkRST uses a random walk based strategy and is biased to create path like topologies. KruskalRST uses a modified version of Kruskal's spanning tree algorithm and creates trees that are in between star and path like topologies. The *GAWES* algorithm proposed in this paper uses Edge-Sets representation as chromosome encoding, and KruskalRST method to populate the initial population and perform crossover.

The first definition and complexity result of *MBCT* problem appeared in [15]. A genetic algorithm based heuristic algorithm (*GA*) is also proposed in [15] to solve the optimization version of the *MBCT* problem. The efficiency of *GA* is evaluated using various fitness functions on a set of randomly generated graphs. *GA* uses Edge-Sets representation for chromosome encoding and a modified version of Kruskal minimum spanning tree algorithm (KruskalRST) is used to create random chromosomes and perform crossover. The suggested parameters for *GA* algorithm is reported in [15] as an iteration size of 50 000, mutation rate of 0.1, and initial population size of 400.

MBCT problem seeks an efficient answer to the bicriteria spanning tree problem, where one needs to minimize both edge cost and maximum post-calibration skew of the spanning tree. In this sense, *MBCT* has similarities to hop constrained [23] and rooted distance constrained [24] spanning tree problems. In hop constrained spanning tree problem, the objective is to minimize the spanning tree cost given that the hop distance of each node to the root is less than some predefined constant. Similarly, in rooted distance constrained spanning tree problem, the objective is to find the minimum cost spanning tree given that delay of each node, associated with each edge, is less than a predefined distance value. In both hop constrained and rooted distance constrained spanning tree problems, both of the costs that needs to be minimized are on the edges. In *MBCT*, having the second cost on the vertex changes the definition and the characteristic of the problem. Therefore, *MBCT* problem clearly distinguishes itself from existing hop constrained and rooted distance constrained spanning tree problems.

3. The problem definition

The Minimum-Cost Bounded-Error Calibration Tree problem was first defined in [15]. Formal definition of the *MBCT* problem was stated in [15] as:

Definition 3.1 (*MBCT [15]*). "Given a wireless sensor network modeled as an undirected graph $G(V, E)$, and a designated reference node $r \in V$, where each $e \in E$ is assigned distance values $d_e > 0$, and each $v \in V$ is associated with a maximum random measurement error ϵ_v, the *MBCT* problem is defined as finding a spanning tree over G rooted at r with total edge cost not greater than a constant $C > 0$, while the post-calibration skew of each sensor $v \in V$ is bounded by a positive constant k".

The *MBCT* problem was shown to be NP-complete in [15], and a genetic algorithm based heuristic was proposed in the same work for the optimization version of the problem. The optimization version of the *MBCT* problem minimizes both the total cost value and the post-calibration skew.

Fig. 1 (a) presents an example graph, where each node i has an associated maximum calibration error (ϵ_i), and each edge j has an associated cost c_j. S_0 is marked as the pre-calibrated sensor with zero calibration error. Post-calibration skew of each node is the sum of the absolute maximum errors of each node along the path to node S_0. For example, in Fig. 1 (b), node S_2 is connected to S_0 through a path passing through node S_3. Therefore, the post-calibration skew of node S_2 is then the sum of the calibration errors of nodes S_2, S_3, and S_0, which is equal to $|\xi_2| = 2 + 8 + 0 = 10$. Fig. 1 (b) shows the minimum spanning tree based on edge costs, where the total cost is 5, post-calibration skew of each node is given as $|\xi_1| = 7$, $|\xi_2| = 10$, $|\xi_3| = 8$, $|\xi_4| = 9$, $|\xi_5| = 1$, and the maximum post-calibration skew is 10. Fig. 1 (c) shows