무선 센서 네트워크에서 에너지 효율적인 집단화를 위한 경험적 백트랙크 탐색 알고리즘

손석원*

Heuristic Backtrack Search Algorithm for Energy-efficient Clustering in Wireless Sensor Networks

Surgwon Sohn*

요 약

체약만족문제(Constraint Satisfaction Problem)의 연구에서 발현되었듯이, 제약최적화 문제(Constraint Optimization Problem)를 효율적으로 풀기 위해서는 변수 순서화의 경험적 방법이 매우 중요하다. 이 기존이 혼합된 무선 센서 네트워크의 에너지 효율적인 집단화같은 문제는 클러스터 헤드가 기지국과 가깝게 위치하려는 경향이 있다. 본 논문은 이 집단화 문제를 풀기 위해서 경적 우선순위 변수 순서화의 기반을 둔 새로운 접근방법을 제시하고, pnode 라는 새로운 알고리즘을 제안한다. 이 pnode 알고리즘은 우선순위가 가장 높은 변수를 다음 변수로 선택한다. 집단화 문제에 있어서, 우선순위가 높다는 것은 클러스터 헤드가 최적지역에 근접하게 된다는 것을 의미하며 이것은 문제의 성격상 미리 정해진다. 클러스터화된 센서 네트워크에서 클러스터 헤드는 에너지 소비가 가장 많이 일어나는 것에서 여러가지 원인 때문에 센서 노드뿐만 아니라 클러스터 헤드에서 발생하는 최대 에너지 소비를 최소화하도록 만드는 방법을 찾는 것이 본 논문의 목적이다. pnode 알고리즘을 사용하여 시뮬레이션한 결과 제안된 방법이 다른 방법들보다 우수함을 알 수 있었다.

Abstract

As found in research on constraint satisfaction problems, the choice of variable ordering heuristics is crucial for effective solving of constraint optimization problems. For the special problems such as energy-efficient clustering in heterogeneous wireless sensor networks, in which cluster heads have an inclination to be near a base station, we propose a new approach based on the static preferences variable orderings and provide a pnode heuristic algorithm for a specific application. The pnode algorithm selects the next variable with the highest preference. In our

* 제1저자 : 손석원
* 호서대학교 뉴미디어학과 부교수
* 이 논문은 2007년도 호서대학교의 재원으로 학술연구비 지원을 받아 수행된 연구임(20080022)
problem, the preference becomes higher when the cluster heads are closer to the optimal region, which can be obtained a priori due to the characteristic of the problem. Since cluster heads are the most dominant sources of power consumption in the cluster-based sensor networks, we seek to minimize energy consumption by minimizing the maximum energy dissipation at each cluster head as well as sensor nodes. Simulation results indicate that the proposed approach is more efficient than other methods for solving constraint optimization problems with static preferences.

- **Keyword**: Wireless Sensor Network, Heuristic Backtrack Search Algorithm, Constraint Optimization Problem

### 1. Introduction

The appropriate variable and value ordering heuristics are important to solve constraint optimization problems. Variable ordering heuristics explain a decision on which branches are searched, and value orderings describe a decision on which values are to be assigned to the variables. The characteristics of variable orderings are that they can change the tree shape of search space and guide the search process towards the regions that are likely to contain solutions. On the other hand, value orderings affect the quality of the solutions. Although several good heuristics have been developed in previous research[1], finding good heuristics for a particular problem has been often a matter of intuition or trial-and-error. Furthermore, it may be very hard to develop good application-independent heuristics, and it seems unlikely to find the universally best algorithm[2]. In order to deal with this difficulty successfully, Gent et al. proposed a definition of constrainedness[3] by which the construction of heuristics is motivated.

If a problem has a property that the variance of domain size for each variable is far larger than that of the connectivity of the variable, the domain size may change radically from branch to branch by constraint propagation. To this kind of problems, we apply the dynamic variable ordering(DVO) heuristics because the main consideration of DVO is domain size. If the connectivity property of variables is dominant in the problem and other properties such as the variance of domain size are insignificant, the static variable ordering(SVO) heuristics may have the best performance because DVOs, which consider mainly domain size, do not make an effective contribution to the finding of solutions.

Let us now consider another problem in which the variance of domain size and the connectivity of variables are so small that they may be negligible but there are static preferences for the variables. Static preferences for the variables suggest that new heuristics called static preferences variable orderings(SPVOs) are more effective than pure SVO heuristics or DVOs because SPVOs consider variable preferences instead of domain size or connectivity. A cluster-based heterogeneous wireless sensor network[4] is known to be a typical problem where cluster heads need to be installed in some optimal region so as to be energy-efficient. For this specific problem, we need SPVOs other than DVOs or SVOs because there are location preferences of the cluster heads. Hence, we propose a pnode algorithm in which the order of the variables is predetermined according to variable preference, that is, cluster heads closer to the optimal region are selected first.

Like the warehouse location problem[5] and the base station location problem[6], the energy-efficient clustering problem in the wireless sensor networks can also be modeled as a constraint satisfaction optimization problem[7] because the clustering problem is very similar to the warehouse location problem and it is NP-hard[8]. Kang and Poovendran[9] investigated energy-efficient broadcast routing mainly from the viewpoint of MINMAX optimization criteria to maximize
the network lifetime, and they demonstrated that the MINMAX strategy can be a more effective load-balancing technique and achieve higher performance than the MINTOTAL strategy. Therefore, we provide the constraint optimization model of energy-efficient clustering with the MINMAX method in heterogeneous wireless sensor networks, and solve it by constraint propagation with the proposed prnodo backtracking heuristics.

II. Related Work

The theoretical aspect of clustering problems in homogeneous sensor networks applied to energy optimization has been widely studied nowadays[10]. Related studies proposed an algorithm for balanced $k$-clustering problems, where $k$ signifies specific clusters in the system. Unlike the previous studies, the optimal assignment of nodes to cluster heads was performed in cluster-based ad-hoc networks[11] where focus was placed on maximizing the lifetime of cluster heads because they are the most critical network elements from the viewpoint of energy. Inevitably, the nodes closer to the cluster heads experience higher traffic and energy consumption rate. To avoid this unbalanced energy problem, a cluster-based energy balancing scheme[12] has been introduced in heterogeneous wireless sensor networks where a strong node acts as a cluster head. However, this algorithm does not use a centralized algorithm but a distributed one, thus it cannot obtain the optimal energy consumption. On the other hand, the method of reducing the amount of data transmitted in each sensor node and cluster head was proposed in order to lengthen the lifetime of the heterogeneous sensor network[13].

If the domain size of a problem changes radically from branch to branch by constraint propagation, we apply the DVO heuristics to the problem because the main consideration of DVO is domain size. The $k$-coloring of a complete graph is a typical example of this kind of problems[14]. One of the best known dynamic variable ordering heuristics is dom that chooses the next variable with the smallest remaining domain. Therefore, it is also called smallest domain first principle. This ordering heuristic was introduced by Golomb and Baumert[15], and Haralick and Elliott[16] analytically showed that dom minimizes the depth of the search trees.

For branches with the same domain size, a tie-breaker was used by Brelaz to order the variables for graph coloring problems. It is called saturation degree algorithm(Dsatur[14]), which selects the variable with the smallest domain size and breaks ties by selecting the variable with the highest degree. Thus, it can be called $<\text{dom}, \text{deg}>$ as the tuple name. Smith also improved Dsatur by adding to it a third tie breaker. If more than one variable has the smallest current domain size and the largest future degree, choose the one for which the domain size amongst its future neighbors is least. It is called BZ2[15], and she also proposed an alternative heuristic, which reverses the second and third tie-breakers. This heuristic is called BZ3 which chooses the variable with the smallest remaining domain, and break ties by choosing the variable which has the smallest remaining domain amongst its future neighbors, and breaks remaining ties by choosing the variable with the largest future degree.

If a problem has mainly a static characteristic and the variance of domain size is quite small, we have an intuition that the best algorithms of variable orderings for this problem are SVO heuristics. Location problems[16] are typical problems in which SVOs are effective. Static variable ordering heuristics order the variables from the highest to the lowest degree by repeatedly selecting a variable in the constraint graph, and the typical algorithm is deg presented by Freuder[17]. For a special problem of cluster-based heterogeneous wireless sensor networks, conventional heuristics based on DVO or SVO may not result in good solutions. In this case, we instantiate the values to preferred variables first.
We name these algorithms SPVO heuristics. As one of the SPVOs, we propose the pnode algorithm, in which the order of the variables is predetermined according to variable preference.

III. Description of a Problem with Static Preferences

For motivational purpose, consider LEACH(18) as a typical cluster-based hierarchical routing protocol in homogeneous networks, in which all sensor nodes have the same function and initial energy. In LEACH, cluster heads are selected according to their probability threshold value. That is, cluster heads are distributed quite evenly over the service area. The optimal number of cluster heads is estimated to be 5% of the total number of nodes. During the setup phase, the predetermined fraction of sensor nodes (e.g., 5%) elect themselves as cluster heads. Cluster heads change randomly over time in order to balance the energy dissipation of sensor nodes. This protocol offers no guarantee of the placement and/or number of cluster heads. To overcome this problem, LEACH-Centralized(19) was introduced to use information about the location and energy level of sensor nodes, and good cluster formation was obtained. This attempts to minimize the amount of energy for the sensor nodes to transmit their data to the cluster heads by minimizing the total sum of squared distances between all the sensor nodes and the closest cluster heads.

Consider the same radio model discussed in LEACH for the radio hardware energy dissipation, in which the transmitter spends energy to run the radio electronics and the power amplifier, and the receiver consumes energy to run the radio electronics. To accurately model the attenuation of radio waves between the sensor nodes and cluster heads, radio engineers(20) typically use a free-space model that attenuates the power of a signal to $1/r^2$ at short distances in which $r$ is the distance between the sensor node and the cluster head. At longer distances, the multi-path model is used where the signal power is attenuated to $1/d^4$ in which $d$ is the distance between the cluster head and the base station. If the distance is less than a crossover point that is called reference distance $r_0$, the free space model is used. Otherwise, the multi-path model is used. The reference distance is typically around 100 meters for outdoors low gain antennas 1.5m above the ground plane operating in the 1-2 GHz band. Thus, to send a $k$-bit message through distance $r$, radio energy needed in the free space channel is shown below.

$$E_{tx}\left(k, r\right) = E_{elec}(k) + E_{amp}(k, r) \quad \text{(1)}$$
$$= kE_{elec} + kP_fs r^2$$

where $E_{elec}$ is the power consumed by the transmitter circuit and a distance-independent term, whereas $P_fs r^2$ is transmitter amplifier energy and a distance-dependent term. In the multi-path fading channel, the total transmission energy is

$$E_{tx}\left(k, d\right) = E_{elec}(k) + E_{amp}(k, d) \quad \text{(2)}$$
$$= kE_{elec} + kP_mp d^4$$

where $P_mp d^4$ is energy for the transmission amplifier in the multi-path fading channel. In both free-space and multi-space fading channels, the radio consumes energy to receive the message.

$$E_{rx}(k) = E_{elec}(k) = kE_{elec} \quad \text{(3)}$$

From the perspective of transmission energy expressed in Equation (1) and Equation (2), energy consumption by cluster heads in multi-path channels is far higher than that by sensor nodes in free-space channels. The optimal region of cluster heads is the upper square area in Figure 1, where the service area of the sensor network is displayed by a square. Cluster heads have an inclination to be located in
the upper region of the network service area in the figure for energy-efficient positioning as the number of cluster heads increases. Where minimum distance \( d_1 = 76 \text{ m} \) and maximum distance \( d_2 = 132 \text{ m} \).

Unfortunately, this protocol does not consider the communication energy of cluster heads, in which most of energy consumption takes place, at the setup phase of selecting cluster heads. Thus, cluster heads are distributed evenly over the service area, and energy dissipation is not optimized to be minimal in LEACH-Centralized. It is worthwhile to note that the locations of the cluster heads are inclined to be near the base station of a clustered network. This happens because the objective function or soft constraint of determining cluster heads minimizes the energy consumption of the clusters in which both cluster heads and sensor nodes are considered at the same time.

IV. Constraint Optimization Model and Algorithms of Energy-efficient Clustering

Let us assume that a set of possible cluster heads \( CH = \{1, \ldots, m\} \) is already given a priori and is fixed in a static sensor network. In a dynamic network, this cluster head selection is performed every round. Cluster head \( j \) can be installed and \( j \in CH \). A set of sensor nodes \( SN = \{1, \ldots, n\} \) is also given. Each sensor node \( i \in SN \) is assigned to cluster head \( j \).

Perfect power control is assumed both in the sensor nodes and the cluster heads, and traffic load is uniformly distributed throughout the networks. Now suppose that \( X \) is the consumed energy vector of cluster heads. \( X_i \) is the total amount of energy consumed at cluster head \( j \) to send to the base station a \( q \)-bit data message which comes from the sensor nodes \( i \), and we observe that

\[
X_i = q \eta (X_{tx} + X_{rx}) \tag{4}
\]

where \( X_{tx} \) is a distance-dependent term, and equal to \( E_{elec} + P_{mp}d^3 \), and \( X_{rx} \) is a distance-independent term and equal to \( E_{ds} + E_{elec} \). \( n_j \) is the number of sensor nodes assigned to cluster head \( j \). Therefore, Equation (4) yields

\[
X_j = q \eta (2E_{elec} + E_{ds} + P_{mp}d_j^4) \tag{5}
\]

where \( d_j \) is the distance between cluster head \( j \) to the base station, and \( E_{ds} \) is the consumed energy for data aggregation from the sensor nodes within cluster \( j \). We now consider sensor energy matrix \( Y \), whose dimension is equal to \( |CH| \times |SN| \). \( Y_{ij} \), the \( i,j \) entry of the matrix, represents energy dissipation at sensor node \( i \) to transmit a \( q \)-bit data message to cluster head \( j \) and is expressed as
\[ Y_o = q(E_{elec} + P_{id}^2) \] .................................(6)

where \( E_{elec,i} \) is electronics energy for sensor node \( i \) to transmit data to cluster head \( j \), and is the same as \( E_{elec} \) for every \( i \). \( r_{ij} \) is the distance between sensor node \( i \) and cluster head \( j \). Thus, we get the total dissipated energy of sensor nodes assigned to cluster \( j \) as follows.

\[ Y_j = \sum_{i \in I_j} Y_{ij} = qn_j \sum_{i \in I_j} (E_{elec,i} + P_{id}^2) \] .................................(7)

where \( I_j \) is a set of sensor nodes assigned to cluster \( J \). Therefore, the total sum of energy spent on the transmission in cluster head \( j \) and their assigned sensor nodes is as follows.

\[ Z_j = X_j + Y_j \] .................................(8)

Once matrix \( Y \) is computed, the model becomes very similar to the warehouse location problem of industrial engineering or the base station location problem in mobile networks. According to many papers[6, 21], this problem is NP-hard and we need good heuristic search methods to get approximate solutions for large networks within a reasonable length of time. When we seek to minimize the maximum energy of cluster heads as well as sensor nodes, the objective function using Equation (8) can be represented as

\[ \min \{ \max_{j \in CH} X_j + \max_{j \in CH} Y_j \} \] .................................(9)

where \( X_j \) is Equation (5) and \( Y_j \) is Equation (7). Let us define \( x_{ij} \) as a decision variable, which is a binary variable. It is equal to 1 if sensor node \( i \) is assigned to cluster head \( j \) otherwise it is equal to 0. \( y_{ij} \) is equal to 1 if a cluster head is installed in \( j \); otherwise it is 0. For the constraint optimization model using objective function (9), we can rewrite Equation (9) above as follows. The first term in the curly braces in Equation (9) becomes

\[ \max_{j \in CH} \left\{ q \sum_{i \in I_j} (E_{elec} + E_{da} + P_{id}^2 r_{ij}^2) \right\} \] .................................(10)

The second term in the curly braces in Equation (9) becomes

\[ \max_{j \in CH} \left\{ q \sum_{i \in I_j} (E_{elec} + P_{id}^2 r_{ij}^2) \right\} \] .................................(11)

subject to

\[ \sum_{j=1}^{m} x_{ij} = 1, \quad \forall i \in SN \] .................................(12)

\[ x_{ij} \leq y_{ij}, \quad \forall i \in SN, \forall j \in CH \] .................................(13)

\[ \sum_{j=1}^{m} y_{ij} = k \] .................................(14)

\[ \sum_{j=1}^{m} x_{ij} \leq C_j, \quad \forall j \in CH \] .................................(15)

\[ x_{ij}, y_{ij} \in \{0,1\}, \quad \forall i \in SN, \forall j \in CH \] .................................(16)

Constraint (12) makes sure that each sensor node \( i \) is assigned to a single cluster head. Constraint (13) requires that sensor nodes are only assigned to the installed cluster head. \( k \) is the number of clusters in Constraint (14). Constraint (15) makes the uncapacitated problem into a capacitated problem. \( C_j \) is the capacity of cluster head \( j \), that is, the maximum number of sensor nodes that can be handled within cluster \( j \). Intuitively, the value may be larger than \( N/k \), where \( N \) is the total number of sensor nodes, and \( k \) is the number of clusters. Owing to Constraint (14) and (15), the heuristic search becomes faster and the search space is gracefully diminished. To extend network lifetime, we need to minimize the maximum energy of each cluster head. In other words, we used the maximum energy of each cluster head as an evaluation criterion for the proposed algorithm because the energy of cluster heads was most important in
single-hop hierarchical sensor networks. When we tried the MINMAX method, the maximum energy of cluster heads decreased continuously as the number of clusters increased, and this is intuitively natural because the average energy per cluster head decreases as the number of clusters increases.

Although most variable ordering heuristics prefer the fail-first or smallest-domain-first rule, cluster heads should be selected from the upper region in Figure 1. When the variances of domain size and connectivity are sufficiently small but there are some preferences for selecting variables, SPVO heuristics should be used for the best solutions. As one of the SPVOs for our heterogeneous wireless sensor networks, we propose a pnode algorithm, in which the order of the variables is predetermined according to the location preference of the variables. The ordering that cluster heads closer to the optimal region are selected first embodies the intuition behind the energy consumption of cluster heads, which is most critical in total energy dissipation in the network. Generally speaking, we have preferences for selecting the next variable.

Once a cluster head has been selected, the algorithm generates a sensor node for it in a nondeterministic manner. The model specifies a heuristic for value ordering where the sensor node must be tried. To minimize the total amount of energy in the cluster, the model says to try first those values that have the closest sensor nodes to the cluster head. We name this closest-sensor node first (CSF) as a value ordering rule.

V. Simulation and Discussion

We consider a 100m square service area with a base station centered at (50,175) and 100 sensor nodes are chosen randomly from uniform distribution. Let us assume (0,0) as the lower-left corner and (100,100) as the upper-right corner in the service area. The unit of the coordinates is meter. Six cluster heads are chosen for installation out of 30 candidates based on Equation (9) and Heinzelman's research[19], and they have the same fixed initial energy. In our experiment, the locations of the 6 selected cluster heads and 100 sensor nodes are plotted in Figure 2.

![Graph](image.png)

*Fig 2. Locations of 6 Cluster Heads and 100 Sensor Nodes*

We assume that electronic device energy $E_{elec} = 50 \text{ nJ/bit}$, and the parameter of transmit amplifier energy $P_{t}= 10 \text{ pJ/bit/m}^2$ and $P_{mp} = 0.0013 \text{ pJ/bit/m}^4$. Energy for data aggregation is set to $E_{da} = 50 \text{ nJ/bit}$. The data message is 500 bytes long, and the packet header for each type of packet is 25 bytes long. Cluster heads just forward data from assigned sensor nodes to base station. For brevity's sake, cluster heads use multi-path model and sensor nodes use free-space model. All the experiments were done in Pentium IV/2.4GHz CPU, 2.6 GB RAM with ILOG OPL 3.7 programming solver. To illustrate the advantage of the proposed algorithm, we compare the results of the pnode algorithm with those of other variable orderings for one round transmission. Table 1 compares the performance of some variable ordering heuristics where the number of clusters is fixed at 6.
VI. Conclusion

For particular constraint optimization problems with static preferences such as energy-efficient clustering in heterogeneous wireless sensor networks, conventional static variable orderings have limitations in the quality of their solutions. Hence, we proposed the *pnode* variable ordering heuristic as one of static preferences variable orderings (SPVOs) to achieve higher performance. The *pnode* algorithm chooses cluster heads first according to location preference in the optimal region, in which cluster heads and sensor nodes minimize energy consumption. We also presented a constraint optimization model based on the MINMAX strategy, which minimizes the maximum energy of each cluster head and sensor nodes combined. Simple comparison with pure SVO or DVO shows that the performance of SPVO is higher than any other heuristics from the viewpoint of the power consumption in each cluster. Further improvements are expected in the SPVO in terms of choice points and execution time.

References


<table>
<thead>
<tr>
<th>Heuristics</th>
<th>maxX+ maxY/(mJ)</th>
<th>Choice Points</th>
<th>Time (sec.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SB</td>
<td>28.947</td>
<td>55,278</td>
<td>49.33</td>
</tr>
<tr>
<td>dom</td>
<td>28.249</td>
<td>2,627</td>
<td>2.67</td>
</tr>
<tr>
<td>MCCF</td>
<td>36.168</td>
<td>2,657</td>
<td>2.48</td>
</tr>
<tr>
<td>pnode</td>
<td>20.822</td>
<td>6,866</td>
<td>6.23</td>
</tr>
</tbody>
</table>

SB : simple backtracking (pure depth first search)

dom : smallest-domain-first variable ordering (DVO)

MCCF : most constrained cluster heads first variable ordering (SVO)

pnode : proposed static preferences variable ordering

From Table 1, we can see that pnode has the lowest maximum energy in each cluster but choice points and computing time need to be enhanced. As one of the pure SVOs, we tried the most constrained cluster heads first (MCCEF) variable ordering that chooses the cluster heads first out of 30 candidates when the connectivity is highest. dom as a DVO works better than MCCF since variance of domain size is larger than that of connectivity. Both dom and MCCF uses random value ordering because CSF value ordering makes the quality of solutions worse. But in the pnode experiment, CSF value ordering works better than random value ordering. From this point, we understand that each variable ordering and value ordering is closely related to the solutions of the problem.

As shown in Figure 2, the locations of cluster heads are inclined to be near the base station. This means that the cluster heads are positioned in a way that the energy use of communication is minimized. That is, the cluster heads need more energy than the sensor nodes. On account of this unbalanced energy consumption, heterogeneous wireless sensor networks should be the ultimate design architecture especially in the static sensor networks.


저 자 소개

손석원
1985년 2월 : 인하대학교 전자공학과 졸업
2007년 8월 : 인하대학교 컴퓨터공학과 졸업
1999년 ~ 현재 : 호서대학교 뉴미디어 엔지니어링과 부교수

이용분야 : 인공지능, 제어반도체학습, 무선통신학