Efficient Assignment of Multicast Groups to Publish-Subscribe Information Topics in Tactical Networks

Shobha Erramilli, Shrirang Gadgil, Narayanan Natarajan
Telcordia Technologies
331 Newman Springs Road, Red Bank, NJ 07701

ABSTRACT

A key challenge in a network-centric tactical environment is disseminating information related to sensors, situational awareness, and command and control to the appropriate receivers in a timely, reliable and efficient fashion. IP Multicast is the transport mechanism used for information dissemination. It is typically the case in tactical networks that the number of IP multicast addresses available for use is limited and it becomes necessary to assign multiple information topics to each multicast address. This necessitates use of an efficient assignment algorithm during network planning to minimize the dissemination of unwanted information in order to decrease the unnecessary use of network resources and reduce the information load on the receiving nodes. This paper describes an efficient algorithm for assigning information topics to multicast groups taking into consideration the overlapping information needs of subscriber nodes and network bandwidth that will be consumed in the dissemination of each topic based on its size and frequency of publication. The assignment algorithm supports unique requirements of tactical networks: for an information topic, the subscribers may not be related by echelon hierarchy and the criticality of information may vary with subscribers. The assignment algorithm uses the simulated annealing technique to compute an optimal address assignment.

1. INTRODUCTION

A key challenge in a network-centric tactical environment is disseminating information related to sensors, situational awareness, and command and control to the appropriate receivers in a timely, reliable and efficient fashion. This information dissemination is typically accomplished using the publish-subscribe (pub-sub) communication paradigm, with IP (Internet Protocol) multicast services as the transport mechanism at the network level. In tactical networks, a factor that motivates the need to minimize the number of multicast groups used for publish-subscribe communications is that secure multicast requires association of a key with each multicast group. Having fewer multicast groups simplifies the key management problem. Further, many tactical radios impose a limit on the number of multicast groups that a node can subscribe.

This is done to reduce the size of security association data stored in the radios. These factors effectively limit the number of multicast groups that can be used for pub-sub communications. Hence, it becomes necessary to assign the same multicast address to multiple information topics. This necessitates the use of an efficient multicast address assignment algorithm during mission network planning to minimize the dissemination of unwanted information which both decreases the unnecessary use of network resources and reduces the information load on the receiving nodes. Several factors need to be considered in coming up with an efficient assignment algorithm, such as comparing the set of receivers for each information topic and the relative bandwidth consumed by each information topic. Frequently, in the battlefield, different subscribers to the same information topic require varying qualities of delivery. This paper describes an efficient algorithm for assigning information topics to multicast groups taking into consideration all of the above factors.

A previous paper that we presented at MILCOM 2007 for efficient multicast address assignments for publish-subscribe groups considered only the subscription group overlap factor in the assignment algorithm [1]. As in the case of the scheme described in [1], the scheme described in this paper also supports information subscriptions that are based on echelon hierarchy as well as subscriptions that are not limited by echelon hierarchy. The former category enables echelon based information dissemination. The latter supports collaboration and mission task grouping that cut across the echelon hierarchy. The assignment scheme presented in this paper is an improvement over the previous paper in the following ways:

- It takes into account the publication frequency and message size of information topics. Consideration of this factor leads to an improved characterization of wastage of information in networks that arises when unwanted information is delivered to nodes.
- In addition, the varying criticality of an information topic to different subscribers is taken into account. This requires such an information topic to be published in different multicast addresses, one for each level of criticality. The algorithm described in this
paper assigns such information topics to multiple addresses without sacrificing efficiency.

- The assignment algorithm uses the simulated annealing technique to compute an optimal address assignment. Simulated annealing is a well known technique in optimization literature and has been successfully used to solve many types of optimization problems [2]. We believe that this approach of using simulated annealing is a significant improvement over the problem-specific heuristic approach that was described in [1]. Our experimental results (see Section 4) show that the simulated annealing algorithm generates more efficient assignments than those generated by the heuristics based algorithm described in [1]. Simulated annealing lends itself better to the introduction of additional constraints and enhanced efficiency metrics that are required to support the new features described above. This is discussed further in Section 3.

The remainder of the paper is organized as follows. Section 2 discusses related work. Section 3 describes the multicast addresses assignment problem formulation and the solution based on the simulated annealing technique. Section 4 presents experimental results. Section 5 concludes the paper by outlining directions for further work.

2. RELATED WORK
To the best of our knowledge, there is no published scheme for assigning pub-sub topics to multicast (MC) groups with a limited MC address space other than our previous paper [1]. As described in [1], current day tactical network planning systems define multicast groups based on echelon structure and assign memberships accordingly. This does not support dynamic ad hoc collaboration needs of future force. In the future force environment, information needs of warfighters are not determined only by echelon structure, but also by tasks that are assigned to them. The proposed MC address assignment scheme supports both echelon based communications and ad hoc collaboration. Geographic multicasting (Geocast) schemes have been proposed, especially for situational awareness. Here, nodes dynamically join a multicast mesh based on their current location in a geographic region [3]. The concept of Geocast has been generalized to a context-aware pub-sub concept and routing protocols have been proposed [4]. In contrast our work is not focused on new routing protocols, but on the problem of planning pub-sub communications over multicast transport.

In order to solve the additional problems stated earlier, we considered optimization techniques described in literature such as linear programming and simulated annealing. Linear programming was not found to be a good fit as the problem being solved here has an inherently non-linear objective function. The cost of assigning an information topic to a multicast group is not independently set, but depends on the other information topics assigned to the same multicast group, which makes it non-linear. Techniques are described in literature for linear approximations to non-linear problems such as solving the quadratic knapsack problem using branch-and-bound algorithms [5]. These algorithms solve a linear approximation of the non-linear objective function and then apply heuristics to get a solution closer to the original problem.

Simulated annealing is another optimization technique in use in many practical situations where the more traditional optimization techniques are intractable. It provides a flexible means to achieving a good solution to large combinatorial problems [2] without imposing any constraints on the objective function definition. Simulated annealing was chosen as the algorithm to provide a solution to the pub-sub group assignment problem subject to multiple constraints because of its relative simplicity when compared to other methods.

3. PUB-SUB TO MC GROUP ASSIGNMENT SCHEME
This section gives a formal definition of the multicast address allocation problem and describes the optimality criterion for such allocation. A detailed description of the proposed address assignment algorithm is then presented.

3.1 PROBLEM STATEMENT
Consider a set $S$ of $k$ subscribers $S = \{S_1,.. S_k\}$, a set $C$ of $m$ information categories $C = \{C_1,.. C_m\}$, and a set $A$ of $n$ multicast addresses $A = \{A_1,.. A_n\}$. Information needs of each subscriber are expressed by a boolean function $N(s, c)$. $N(s, c) = 1$ implies that subscriber $s$ needs information of category $c$. $N(s, c) = 0$ implies that subscriber $s$ is not interested in information category $c$.

Given the above inputs, the goal of the algorithm is to find an optimal assignment function $M(C, A)$ that assigns a multicast address from $A$ to each category in $C$. $M(c, a) = 1$ if the multicast address $a$ is assigned to the category $c$. $M(c, a) = 1$ implies that a source publishing information belonging to the category $c$ uses the multicast address $a$ for dissemination. It is generally expected that the number of categories will be far greater than the number of available multicast addresses and the mapping function will result in assigning the same multicast address to several categories.
3.2 OPTIMIZATION CRITERIA
Since a multicast address may be used to disseminate several information categories, a subscriber desiring information belonging to one of these categories may end up receiving all categories that are assigned the same address. This results in unnecessary information being transported in the network using up network resources. We also need to consider the fact that wastage increases with the bandwidth consumed by the unnecessary information. We seek to find an address assignment function \( M \) that is efficient, i.e., one that minimizes network resources consumed in delivering unwanted information to subscribers. Since different categories may be published at different rates (varying publication frequency) and the data size of categories vary, these factors also must be considered in determining network resources wasted in delivering unwanted information.

To capture this notion of efficiency of address assignment, we introduce the concept of the bandwidth factor for categories. For a category, \( c \), the bandwidth factor \( B_c \) is a number (a relative weight) that denotes the network bandwidth consumed for the dissemination of the category \( c \) to a single subscriber. \( B_c \) is a function of the size and the publication frequency of \( c \). The efficiency of the address assignment function \( M \) is defined as the ratio of the useful (needed) information delivered to the total information transmitted. Thus, the efficiency factor \( E \) is a number from 0 to 1 that is formulated as follows. This formulation assumes that a multicast message contains only one information item belonging to only one category; i.e., no packing of multiple items in one message. Further, it assumes that a category is assigned only a single multicast address. To easily express the efficiency of address assignment, we use an auxiliary function \( M_t \) \((S,C)\) that is derived from \( M(C,A) \). \( M_t \) expresses a receiver relationship between subscribers and categories. \( M_t(s,c) = 1 \) if subscriber \( s \) receives \( c \) as a result of the assignment function \( M \) irrespective of whether \( s \) subscribed to \( c \) or not. That is, \( M_t(s,c) = 1 \) if and only if for some \( a \) and \( c1 \) \((N(s,c1) = 1 \wedge M(c1,a) = 1) \). Note that in the above, \( c1 \) may be \( c \).

The efficiency factor \( E \) for the address assignment function \( M \) is defined as

\[
E = \sum_{c \in C} \left( B_c \sum_{s \in S} N(s,c) \right) / \sum_{c \in C} \left( B_c \sum_{s \in S} M_t(s,c) \right)
\]

3.3 Multiple Dissemination Priority for an Information Topic
Quite often, in tactical networks, there is an operational need to deliver the same information with varying quality of service (QoS) to different subscribers. For example, when a soldier in a platoon detects a mine, this information should be disseminated to other soldiers in the battlefront with higher criticality compared to a commander in the rear. Current day multicast schemes do not support information delivery with varying QoS using a single multicast address. Thus, the above operational need can be realized only by assigning multiple multicast addresses to the same topic (event), one for each QoS class, and publishing the same topic on each assigned multicast address. To support this capability, the assignment scheme needs to be extended to consider each topic with varying degree of QoS as multiple categories and generate address assignments satisfying the constraint that each of these categories should be mapped to a different multicast address.

3.4 ASSIGNMENT ALGORITHM
Simulated annealing is a technique used to solve large combinatorial problems by searching through the problem space in a manner that reduces the defined objective function. An initial solution is perturbed repeatedly in small steps and the new solution is accepted or not accepted based on the algorithm’s rules. All steps that result in a better solution or equally good solution are accepted. Steps that result in a worse solution are also accepted based on a probability (see below).

In an analogy to the metallurgical annealing method, where a metal is cooled gradually from a high temperature state in order to achieve its lowest energy state, this approach starts from a high temperature and reduces the temperature in stages according to a cooling rate until the desired low temperature is reached. At each temperature stage, \( n \) solutions are generated and evaluated. In our implementation, \( n \) is set to 500. At each step \( i \), the objective function \( P_i \) is evaluated. Steps that result in a worse solution are accepted based on a probability using the inequality

\[
r < e^{-\Delta P/T},\quad \Delta P = (P_i - P_{i-1}),\quad T
\]

is the temperature at that stage

The inequality results in the algorithm accepting worse solutions with a higher probability at high temperatures, thus preventing the algorithm from converging to a local minimum. This basic algorithm is combined with specific objective functions and constraints to achieve an approximately optimal solution.

The objective function used in the annealing algorithm for multicast address assignment is the wastage in the network and looks to minimize the wastage. The wastage is the sum of wastages for each category. The wastage attributed to each category is a function of the bandwidth used in dissemination of information to subscribers that are not
The wastage factor is defined as follows. We use another auxiliary function $M_2(s,c)$ that expresses the wastage delivery relationship between subscribers and categories. $M_2(s,c) = 1$ if subscriber $s$ has not subscribed to $c$ but receives $c$ as a result of the address assignment. That is, $M_2(s,c) = 1$ if and only if for some $a$ and $c1$ ($N(s, c) = 0 \land M_2(c, a) = 1 \land N(s, c1) = 1 \land M_2(c1, a) = 1$).

The wastage factor $W$ for the address assignment function $M$ is defined as

$$W = \sum_{c \in C} B_c \sum_{s \in S} M_2(s,c)$$

Note that our formulation of wastage and efficiency is from an end-to-end perspective and does not take into account routing of packets to the multicast groups.

The annealing algorithm starts with an initial random assignment of categories to addresses and then repeatedly generates new assignments that are slightly different from the previous one. In our implementation, we used one of two basic moves in generating the new assignments. In one move, we take two random categories and swap the addresses assigned to them. In the other move, we take one random category and assign it to a different randomly selected address. The only constraint applied here is to not swap out the last category assigned to an address. Note that it is not necessary to apply this constraint as the algorithm would also only pick a solution where all the addresses were being utilized because of the objective function.

To solve our second problem related to the differing priority of delivery for different subscribers, we added a new constraint to the assignments generated by the algorithm. The problem requires that categories that need different priorities of delivery be assigned to multiple multicast addresses, one for each priority level, making sure that the subscribers to the category only receive at the desired priority. In order to simplify the assignment algorithm, we defined a constraint that is stricter than what is required. We split the categories such that the high priority categories were in a separate group. We further constrained the assignment so that all categories assigned to a multicast address are of the same priority. Thus, all information topics published on a single multicast address are of the same priority. The algorithm was then applied to find the most efficient (least wastage) solution while applying the new constraint. The next section describes the details of the experiments and the results obtained.

4. EXPERIMENTAL SETUP

In order to evaluate the proposed multicast address assignment algorithm for realistic tactical environments, we constructed test data sets containing a large number of information categories and subscribers with varied set of information needs using a multi-echelon organization within a brigade combat team. Subscribers in this context are either soldiers or C2 (Command and Control) applications in vehicles in the various echelons within the brigade. For every generated data set, the organizational structure of a brigade was randomly created in the following manner. A brigade can have three to five battalions. A battalion can have three to five companies. Each company is comprised of three to five platoons. For every echelon, between four to seven subscribers were randomly created.

Information subscriptions were created in the following manner. For every echelon within the brigade, an information category was created and all subscribers under that echelon hierarchy were deemed to be interested in this information category. Thus, for instance, a category associated with the brigade has all subscribers in the brigade interested in it. These subscriptions account for echelon based subscriptions for Situational Awareness and C2 information. In addition to echelon based categories, a specified number of cross-echelon categories were also added. Subscriptions to these categories are not based on echelon hierarchy. They correspond to information subscriptions supporting collaboration and mission task grouping across multiple echelons. In the experiments, subscriptions to these categories are randomly generated. Every subscriber is associated with a number of randomly chosen cross-echelon categories. The input data set was very similar to the one used in [1], with the addition of parameters that capture the new constraints, i.e., bandwidth and priority.

In a tactical environment, it is likely that categories that serve a larger portion of the echelon hierarchy have low bandwidth and the categories that are exclusively of interest in the lower echelons have higher bandwidth. We devised experiments by assigning different bandwidth levels for the different echelon levels. The bandwidth for all categories was set relative to the highest echelon category’s (brigade level) bandwidth which was set at 1.

Using data sets of categories and subscribers as described above, the assignment algorithm was run in the following conditions:

Experiment 1: Equal bandwidth for all categories

For this experiment, the bandwidth of all categories was set to 1. This problem then reduces to the same one that was solved in the heuristics based algorithm, which ignored bandwidth [1]. A data set of categories and subscribers was randomly generated as described above. Total number of available addresses was specified as a percentage of the number of categories generated. The
algorithm was run and the efficiency of the algorithm was measured for different values of address pool size. This provided a comparison of the new algorithm with the heuristics based algorithm to establish a baseline.

Experiment 2: Bandwidth set differently for each category
A data set of categories and subscribers was randomly generated as described above. The bandwidth for each category was set as described earlier. The experiment first generated a solution where the bandwidth was ignored in determining the assignment and then generated a solution which takes into account the bandwidth. The efficiencies achieved by the two algorithms were compared. The experiment was repeated with a fresh set of data where the ratio of the bandwidth of the platoon level categories to that of the brigade level category was varied in steps of 50:1, 100:1, 250:1 and 500:1. The number of available addresses in all the experiments was set at 50% of the total number of categories. The cross-echelon categories were fixed at 50% of the total. The experiment was repeated with different fractions of cross-echelon categories to the total categories.

Experiment 3: Categories with different Priorities
The data set was generated as described earlier. We considered two levels of priority in our experiments. The cross-echelon categories, which are more likely to have subscribers with differing priority, were each split into two based on the set of subscribers. One set of categories was assigned the high priority and the other set was assigned the low priority. The algorithm was run by varying the number of available addresses as a percentage of categories. For the initial assignment in the algorithm, we identified a fraction of addresses and assigned the high priority categories to these addresses. The number of addresses initially available for high priority categories was set to a percentage of the total addresses, varying from 5% to 25% in steps of 5. We removed the constraint that the algorithm reject any assignment that removes the last category assigned to an address. By doing this we allowed the algorithm to move away from the initial address mix and find a potentially better mix of high vs low priority addresses. The assignment solution was evaluated to ensure that the categories of different priorities were not assigned to the same address and the efficiency was evaluated.

4.1 EXPERIMENTAL RESULTS
Experiment 1 showed that the assignment algorithm based on simulated annealing technique is effective at choosing an efficient assignment. The algorithm started from a random assignment which was changed in each step repeatedly until the low temperature was reached. The value of the objective function at the end of each temperature stage was plotted and the results are shown in Figure 1. It shows clearly that the algorithm converges to a good solution based on the objective function. Also, about 50% of the optimization is achieved very quickly in the very first temperature step. This pattern was observed in all our experiments.

Figure 1 Optimization Progress

Figure 2 compares the efficiency achieved by the simulated annealing solution with that achieved by the heuristic solution described in [1]. The two assignment algorithms were applied to the same data set 10 times, where the address pool size was varied from 10% to 100% in steps of 10 and the same efficiency metric was calculated in both cases. The new algorithm produces a better solution at all address pool sizes, with a more marked improvement in the range of address pool sizes of 30% to 70%. At large address pool sizes both algorithms achieve good efficiencies as there are enough resources.

Figure 2 Comparison of efficiency
Having established that the new algorithm produces better solutions consistently we then ran Experiment 2 where the optimization algorithm was applied to categories that had different bandwidth factors as described earlier. Figure 3 shows the efficiencies achieved. The line marked “Efficiency Ignoring b/w” is the result achieved when the assignment algorithm ignored bandwidth in computing wastage and efficiency. However, when we computed the efficiency with our improved formula taking bandwidth into account, the resulting efficiency was quite different. It was seen to get worse as the ratio in bandwidth factors between platoon and brigade levels was increased, i.e. the cost of ignoring bandwidth increases with bandwidth. The bandwidth aware algorithm is seen to achieve efficiencies that are well above the bandwidth unaware algorithm and the benefits increase as the bandwidth ratio increases. By taking the bandwidth factor of each category as a factor in the objective function, the bandwidth aware assignment algorithm tries to minimize the number of subscribers receiving an unwanted category with a large bandwidth factor. It distinguishes between two categories with the same subscriber size but different bandwidth factors in its decision to group it with a non-overlapping category. As the bandwidth factor increases, the impact of bandwidth is much more dominant and therefore the assignment algorithm does a better job of staying away from assignments that reduce efficiency.

These results indicate the effectiveness of the algorithm in optimizing against this new bandwidth factor that we added. They also show quantitatively that the penalty of ignoring bandwidth given by poorer efficiency of the network is greater where the variation in bandwidths of categories is large. Table 1 shows improvement achieved in efficiency compared with the old algorithm [1] that did not consider bandwidth. An improvement of 48% is achieved when the ratio of highest bandwidth at the platoon level is 500 times that of the brigade level category.

While Experiment 2 demonstrated the effectiveness of the algorithm in optimizing the efficiency of the network where the only constraint was the number of addresses available, Experiment 3 shows that the algorithm is capable of capturing additional constraints in finding an optimal solution. We set up experiments by constraining a feasible solution to be one where a set of subscribers identified as high priority subscribers could not share a multicast group with lower priority subscribers. Earlier the assignments generated at each annealing step were completely random and unconstrained. Here we constrained the assignment to prevent the mixing of high priority categories with low priority categories. Again the algorithm was run for a range of data sets. We varied the number of categories, the fraction of categories that require priority treatment, and the number of addresses. The algorithm was effective in segregating the two types of subscribers while finding a good assignment solution. The additional constraint was added to the existing optimization criteria from the previous experiment. Table 2 shows some results of the multiple runs of the experiment with different data sets for the case where the address pool size is 40% of the number of categories.

The results indicate that addition of the priority related constraint did not negatively impact the efficiency of the assignment. In fact, as the proportion of high priority categories increases we see that the efficiency improves. The right most two columns show the addresses used for
high priority categories in the initial assignment and the final assignment. The results indicate that the algorithm resulted in more addresses assigned for high priority categories while producing the final assignment.

5. SUMMARY AND FURTHER WORK
In this paper, we described the problem of efficient assignment of multicast addresses to pub-sub topics in an environment where the available multicast address pool is only a fraction of the number of information topics, the network bandwidth consumed in the dissemination of different information topics are different based on their publication frequency and data size, and the same information needs to be disseminated to different subscribers with different QoS. The metric for efficiency is minimization of reception of unwanted information by subscribers and minimization of network resources consumed by this unnecessary information dissemination. We described a scheme for efficient multicast address assignment based on simulated annealing technique. We evaluated the efficiency of the proposed algorithm for different data sets that involve subscriptions based on echelon hierarchy as well as subscriptions that cross the hierarchy. The experimental results show that the simulated annealing approach generates efficient multicast address assignments and that these assignments are more efficient than those generated by the problem specific heuristics approach we presented in an earlier work. The results also show that consideration of the bandwidth factor of categories offers significant improvements in the utilization of network resources consumed in information dissemination. Simulated annealing provides a flexible and general-purpose optimization method that has the potential to be applied to solve other problems in this area.

There are several issues that merit further study. Some are listed below:
Dynamic replanning is often required in tactical networks. In the context of pub-sub communications, replanning may require subscription changes to information categories to accommodate retasking of warfighters during mission execution. Further work on multicast address assignment needs to consider at least a common replanning scenario called unit reorganization where a unit under one echelon is moved to be under another echelon; e.g., a platoon under one company is moved to be under another company. The problem of computing multicast address reassignments for unit reorganization is to compute new address assignments to categories satisfying changes in subscriptions such that they cause minimal changes in multicast group membership of subscribers. This minimization objective is motivated by the desire to minimize changes to existing multicast trees thereby minimizing disruptions to ongoing multicast traffic.

Another area for further research is to expand the solution space considered by the optimization criterion and the algorithm to include solutions where an information category may be associated with multiple MC addresses. In some cases, it may be more efficient to assign multiple MC addresses to a category, even in cases where there is no varying QoS requirement for the category. See [1] for an example. This suggests that the algorithm can be expanded to explore solutions where a category can be assigned multiple addresses. This requires optimization criterion to be expanded to cover wastage due to multiple transmissions.

Another direction for further work is refinement of the concept of bandwidth factor for categories capturing network resources consumed in each hop in the dissemination of categories. Since the proposed scheme works in network planning phase where network and multicast tree topology information are yet unknown, this cost can only be estimated. One approach is to assign costs that are proportional to the geographical distance between information sources and receivers.

6. REFERENCES