ENERGY-AWARE SECURE MULTICAST COMMUNICATION IN AD-HOC NETWORKS USING GEOGRAPHIC LOCATION INFORMATION

Loukas Lazos, Radha Poovendran[*]
Network Security and Cryptography Laboratory
University of Washington, Seattle, WA 98195
radha@ee.washington.edu, llazos@u.washington.edu

ABSTRACT

The problem of securing multicast communications in an energy-constrained ad-hoc network requires the efficient management of cryptographic quantities. We show that existing efficient key distribution techniques for wired networks that rely on logical hierarchies are extremely energy inefficient. We also show that the consideration of the physical location of the members is critical for developing energy-efficient key distribution schemes. By exploiting the spatial correlation between the members of the multicast group, we construct an energy-aware key distribution scheme. We present simulation results to illustrate the improvements achieved by our proposed algorithm.

1. INTRODUCTION

When an identical message has to be sent to multiple receivers, multicast communications model reduces the sender as well as the network management overhead. Many applications that make use of single-sender-multiple-receiver communication model can benefit from multicast mode. In order to secure the communication channel, the multicast communication should be encrypted [1]. The use of symmetric key cryptography allows the sender to perform one encryption (in broadcast mode) and every user to perform one decryption per message, thus reducing the computational and communication overhead. However, the use of a single key known to all members, requires its update each time a group member joins or leaves the group, in order to provide backward and forward traffic protection. Since every member holds the data encryption key also known as session encryption key (SEK), when a member leaves the group, a secure channel to reach the remaining valid members for the update of the SEK is required. Hence, the group has to have additional keys called Key Encrypting Keys (KEKs) [1].

The *key management problem* is to ensure that only valid members hold the valid keys at any time. The key management problem can be reduced to the *key distribution problem*, which involves the secure and efficient distribution of the SEK and the KEKs to valid members. In case of wired networks, the rooted tree based hierarchical key distribution schemes are known to be optimal [1]. In [2], these results were directly used for energy-constrained sensor networks. However, as we show in this paper, such models are not energy-efficient. We present an energy-aware key distribution scheme that makes use of the geographical location information about the multicast members, for achieving efficient key distribution.

2. THE AD-HOC NETWORK ENVIRONMENT

We assume that omni-directional antennas are used for transmission and reception of the signal. The required power P_d for reaching a receiver at a distance d is proportional to the γ^h power of that distance, with $2 \le \gamma \le 4$. Assuming the proportionality constant to be one, we have $P_d=d^{\gamma}$.

We now demonstrate how transmission power (a quantity defined in the physical layer), affects the way the routing procedure is realized at the network layer. The wireless medium along with the omni-directional antennas, offer the unique characteristic of the *broadcast advantage* [3]. In Fig. 1(a), sender S transmits a message to node

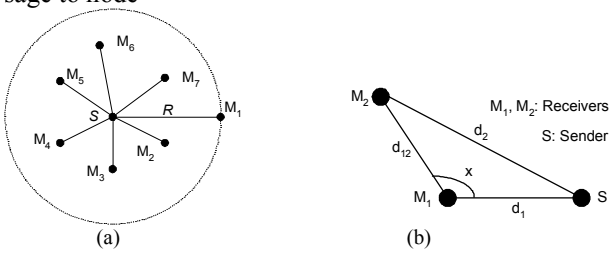


Fig. 1. (a) Broadcast advantage for members M_1-M_7. (b) S transmits an identical message to both receivers

M_1, all nodes that lie within the circle of radius $|SM_1|$ receive the message for "free".

We now show the impact of this physical layer property on the routing decision. In Fig. 1(b), assume that

[*] This research was funded by NSF grant ANI-0093187 and ARO grant DAAD-190210242

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
2003		**00-00-2003 to 00-00-2003**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Energy-Aware Secure Multicast Communication in Ad-Hoc Networks Using Geographic Location Information	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **University of Washington,Department of Electrical Engineering,Seattle,WA,98195**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**		**4**	

$d_2>d_1$ and that the sender S needs to transmit an identical message to nodes M_1 and M_2. A simple strategy would be to use unicast transmissions requiring a total energy expenditure of $(d_1^{\gamma}+d_2^{\gamma})$. However, the broadcast nature of the wireless medium can reduce this expenditure with the use of one of the two following strategies: (a) transmit to M_1 and let M_1 relay the message to M_2. (b) transmit to M_2 and let M_1 receive the message for free, since $d_2>d_1$ (due to broadcast). This leads to the following rule: if $d_2^{\gamma}>(d_1^{\gamma}+d_{12}^{\gamma})$ then the sender chooses the strategy (a), otherwise strategy (b) is preferred. Note that the routing decision relies solely on the nodes' physical location if identical path loss model is assumed.

3. IMPACT OF THE PHYSICAL LOCATION INTO THE KEY DISTRIBUTION SCHEME

We now demonstrate the need for consideration of the geographical location information in the construction of the key distribution scheme. Such information can be obtained using the Global Positioning System (GPS) [4]. In Fig. 2, we represent a wireless network of 7 nodes, with one of them being the sender, also known as the group controller (*GC*), and two intermediate nodes R_1, R_2 relaying traffic to four receiving nodes M_1-M_4, which form a multicast group. The energy required for sending a message from the GC to the two relay nodes is set to one unit and the energy required for sending a message from the relay nodes to the receiving nodes is also set to one unit. Hence, the GC needs to perform only one transmission to reach R_1, R_2 and, relay nodes R_1, R_2 need to perform one transmission each to reach $\{M_1, M_2\}$ and $\{M_3, M_4\}$, respectively.

Fig. 3 presents two different key distribution strategies for the multicast group in Fig. 2. The one in Fig. 3(a) is built according to the available geographical location information (close-by members are placed adjacently into the key tree), while the one in Fig. 3(b) is a result of a random placement of the members into the leaves of the tree (logical assignment). All the nodes share the key K_0.

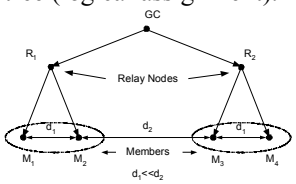


Fig. 2

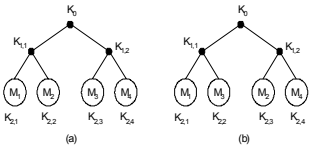


Fig. 3

Fig. 2. Routing tree of an energy constrained wireless network
Fig. 3. (a) A hierarchical tree based key distribution scheme based on geographical location information. (b) A logical hierarchical tree key distribution scheme.

Let's assume that key K_0 has been compromised and needs to be replaced by the new key K_0'. For scheme in Fig. 3(a), the GC generates encrypted messages $\{K_0'\}_{K_{1,1}}$ and $\{K_0'\}_{K_{1,2}}$ and transmits them to relay nodes R_1 and R_2, respectively. Node R_1 performs one transmission to M_1, M_2 and R_2 performs one transmission to M_3, M_4. The total energy expenditure is 4 energy units. For scheme in Fig. 3(b), the GC transmits two messages to both R_1, R_2. Both R_1 and R_2 need to transmit twice to reach nodes M_1, M_3 and M_2, M_4, since nodes that do not share common keys cannot be reached with a single transmission. The scheme in Fig. 3(b) requires 6 energy units. *Hence, for this example, the consideration of the physical location information in the realization of the key distribution scheme leads to energy savings of 33.*

4. ENERGY-AWARE KEY DISTRIBUTION USING GEOGRAPHICAL LOCATION INFORMATION

We now propose an energy-aware key distribution scheme, which uses geographical location information. We make the observation that members that are spatially close to each other can potentially be reached with broadcast, or use the same routing paths to receive data. If we represent members as points in the 2-dimentional plane, we can employ a clustering algorithm to cluster them into groups and construct a hierarchical key tree structure.

Though one can use any suitable clustering algorithm, we have developed a variant of *K-means* (a popular algorithm in pattern recognition and classification), algorithm for creating appropriate clusters. K-means algorithm is used due to its ease of implementation and the ability to control the number of steps in which it terminates [5]. The goal of K-means is to create K clusters out of N points ($K<N$) such that a "loss" cost function is minimized with respect to a dissimilarity measure. K-means uses the squared Euclidean distance as a dissimilarity measure. If the coordinates of point i are $x_i = (x_{i1}, x_{i2})$, the Euclidean distance is equal to:

$$d(x_i, x_i') = \sum_{j=1}^{2} (x_{ij} - x_{i'j})^2 = \|x_i - x_i'\|^2 \quad (1)$$

When $\gamma = 2$, the Euclidean distance is proportional to the transmission power and is suitable for our clustering procedure. In case where $\gamma \neq 2$, we can modify the dissimilarity measure to be proportional to the transmission power as in (2).

$$d(x_i, x_i') = \|x_i - x_i'\|^{\gamma} \quad (2)$$

In our analysis we will focus, without loss of generality, on the case where $\gamma = 2$. When $\gamma \neq 2$ instead of K-means, *K-medoids* can be used to solve the optimization problem at the expense of increased computational complexity [5]. The "loss" function that is minimized in K-means is:

$$W(C)=\frac{1}{2}\sum_{k=1}^{K}\sum_{C(i)=k}\sum_{C(i')=k}\left\|x_i-x_i'\right\|^2$$

$$=\sum_{k=1}^{K}\sum_{C(i)=k}\left\|x_i-\bar{x}_k\right\|^2 \qquad (3)$$

where $C(i) = k$ symbolizes the assignment of the i^{th} member to the k^{th} cluster, $\bar{x}_k=(\bar{x}_{1k},\bar{x}_{2k})$ is the mean vector of cluster k and C is the resulting cluster configuration (N members assigned into K clusters). The "loss" function is minimized by assigning the N members to K clusters in such a way, that within each cluster the average dissimilarity (distance between points assigned to the same cluster and the cluster mean) is minimized.

$$C^*=\arg\min_{C}\sum_{k=1}^{K}\sum_{C(i)=k}\left\|x_i-\bar{x}_k\right\|^2 . \qquad (4)$$

For a set of points S,

$$\bar{x}_S=\arg\min_{m}\sum_{i\in S}\left\|x_i-m\right\|^2 , \qquad (5)$$

where m is the mean point of the set S. We can obtain C^* by solving the enlarged optimization problem

$$\min_{C,\{m_k\}_1^K}\sum_{k=1}^{K}\sum_{C(i)=k}\left\|x_i-m_k\right\|^2 , \qquad (6)$$

where m_k is the mean of the k^{th} cluster. The iterative algorithm for solving (5) is as follows [5]:

K-means algorithm:

Step 1: Have an arbitrary assignment C of points into the specified number of clusters K (initialization can be done by assigning the i^{th} point to the i mod K cluster). Compute the mean vector for each cluster.

Step 2: For the given assignment C, the cluster variance as expressed in (5) is minimized with respect to $\{m_1,\ldots,m_K\}$ yielding the means of the currently assigned clusters (5).

Step 3: Given a current assignment of means $\{m_1,\ldots,m_K\}$, (6) is minimized by assigning each point to the closest cluster mean. That is,

$$C(i)=\arg\min_{1\le k\le K}\left\|x_i-m_k\right\|^2 \qquad (7)$$

Step 4: Iterate step 2 and 3 until the cluster memberships do not change.

We now propose a method that utilizes K-means algorithm in order to construct a hierarchical tree structure. Using K-means, N members are assigned into *2* clusters. Since K-means does not guarantee that equal number of members will be assigned to each cluster or N might be odd, a *refinement procedure* for assigning equal number of members to each cluster takes place. This refinement leads to a construction of a balanced key tree when $N = 2^n$ and forces a structure as close to the balanced as possible otherwise. At every following step all clusters are further divided into two new clusters and the refinement procedure for each pair of clusters is repeated. The algorithm terminates when we have created clusters of two members (after $\log_2 N$ splits when $N = 2^n$).

Key distribution scheme using refined K-means:

Step 1: Assign all points to an initial global cluster.

Step 2: Divide each cluster into two clusters using the K-means algorithm.

Step 3: Use the refinement procedure detailed below to balance the number of points that are assigned to each cluster, i.e. assign the same number of points to each cluster.

Step 4: Iterate step 2 and 3 until clusters of two or one points have been created.

Step 5: Merge single points, if possible, with the use of K-means for only single points.

Step 6: Map the cluster hierarchy into tree hierarchy.

If cluster C_1 has $|C_1|$ points and cluster C_2 has $|C_2|$ points with $|C_1|>|C_2|$, the refinement procedure moves $\left\lfloor \left\|C_1\right|-\left|C_2\right\|/2 \right\rfloor$ points from C_1 to C_2. The criterion by which the points to be moved are selected, is the minimum Euclidean distance of members belonging to cluster C_1 from the point expressing the mean vector m_{C_2} of cluster C_2.

Refinement Procedure:

$C_{\min}= min\{|C_1|, |C_2|\}$, $C_{\max}= max\{|C_1|, |C_2|\}$

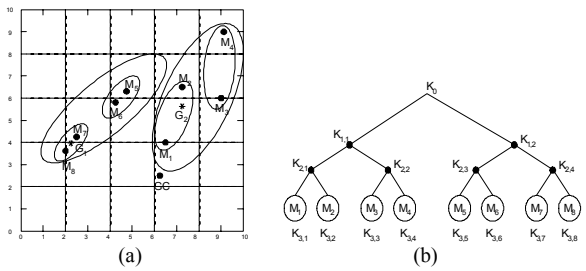
for k=1 to $\left\lfloor \left(\left|C_{\max}\right|-\left|C_{\min}\right|\right)/2 \right\rfloor$ *do*

find $i^*\in C_{\max}$ *such that*

$$d(x_{i^*},m_{C_{\min}})=\min_{i\in C_{\max}}\left\|x_i-m_{C_{\min}}\right\|^2 \qquad (8)$$

and move it to cluster $C_{\min}$.

recalculate means m_{C_1}, m_{C_2} .



(a) (b)

Fig 4: Application of refined K-means algorithm to an eight-node network.

In Fig. 4(a), the results of the application of the refined K-means algorithm for an eight-node network are shown. Initially two clusters are constructed with centers G_1, G_2

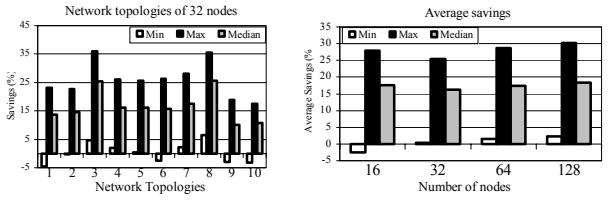
containing members $\{M_7, M_8\}$ and $\{M_1, M_2, M_3, M_4, M_5, M_6\}$. After the application of the refinement procedure, members M_5 and M_6 are moved from G_2 to G_1 in order to form two clusters of equal size. A second application of K-means results in the shown final four clusters and the tree in Fig. 4(b) is built.

5. SIMULATION RESULTS AND DISCUSSION

In this section we present the results of the application of our algorithm in a large number of simulation experiments.

Simulation was performed in randomly generated network topologies confined in a 10x10 square grid region. After the network generation, *BIP* algorithm [3] was applied in order to provide the routing tree. The routing tree is used to calculate the consumed energy for transmitting to each member or group of members from the GC an updated key. In our network model, members have the ability to transmit with infinite power and the propagation loss factor is set to γ=2. Mobility is not considered in our network model.

For the comparison of the performance of our algorithm, 10,000 key distribution trees are randomly generated. The energy for re-keying the key tree constructed from the refined K-means algorithm is compared to the minimum, maximum, median energy required for the update of the 10,000 trees. The size of the multicast group takes different values N=16, 32, 64, 128 plus the GC.



(a) (b)

Fig 5: (a) Performance of K-means compared to 10,000 random key trees for 10 different network topologies. (b) Average savings using the K-means algorithm for networks with different number of nodes.

In Fig. 5(a) the savings for 10 different randomly generated network topologies are shown. In Fig. 5(b) the average savings are shown. The savings are averaged over 100 different network topologies for each value of N and 10,000 trees are used for comparison for each topology.

We can observe that refined K-means algorithm significantly out-performs the Logical Key Hierarchy (LKH) [1] key distribution scheme in energy efficiency. *Refined K-means procedure provides 15%-37% savings from the worst possible assignment, 15%-26% savings from the median case and does almost as good as the best possible case out of 10,000 trees.*

However, there are cases where the application of our algorithm results in higher energy expenditure, compared to a randomly generated tree. This fact reveals the non-optimality of our algorithm. There are cases that refined K-means fails to accurately exploit the broadcast advantage. The reason for this failure lies in the omni-directionality of the broadcast advantage. If two nodes are almost at the same distance from the GC, but in an opposite direction, refined K-means procedure will fail to capture the benefit from clustering them together, since those nodes will not be spatially correlated.

The performance of the refined K-means algorithm relies also on the homogeneity of the environment where the ad-hoc network is deployed. If the path loss model is identical throughout the network terrain, then the physical location of the nodes is directly connected with the energy spent for reaching them. Hence, the observation that close by members will receive information through the same routing paths is valid. However, when the environment is not homogeneous but different path loss models must be assumed for the description of different regions, high spatial correlation between nodes does not necessarily imply the use of similar routing paths for directing traffic to those nodes.

Hence, we can partition the area where the network is deployed, by identifying regions that are described by the same path loss model. This partition is followed by the application of refined K-means to each region, resulting in the construction of sub-clusters. The combination of those sub-clusters according to an energy minimization criterion can provide a desired key distribution scheme.

6. CONCLUSION

We showed that the secure multicast in ad-hoc networks must consider the physical location of the members in order to be energy-efficient. In particular, we showed that the results [1] do not generalize to ad-hoc networks. Recent past work had implied this generalization was feasible [2]. We also presented an energy-aware key distribution scheme that relies on the spatial correlation of the members.

7. REFERENCES

[1] D. M. Wallner, E. C. Harder and R. C. Agee, "Key Management for Multicast: Issues and Architectures," *INTERNET DRAFT*, September 1998.

[2] D. Carman, P. Kruus, B. Matt, "Constraints and Approaches for Distributed Sensor Network Security," NAI Labs Technical Report #00-010 September 2000.

[3] J.E. Wieselthier, G.D. Nguyen, A. Ephremides, "On the Construction of Energy Efficient Broadcast and Multicast Trees in Wireless Networks," in Proceedings *IEEE INFOCOM* 2000, pp. 586-594.

[4] Educational Observatory Institute GPS page, available via WWW at URL: http://www.edu-observatory.org/gps/gps.html.

[5] T. Hastie, R. Tisbshirani, J. Friedman "The Elements of Statistical Learning, Data Mining, Inference and Prediction," Springer Series in Statistics, NY, 2001.