Secret Dispersion: Secure Data Delivery in Cyber Physical System

Rui Hu and Yanmin Gong
School of Electrical and Computer Engineering
Oklahoma State University
Stillwater, Oklahoma 74075
Email: {rui.hu@,yanmin.gong@}okstate.edu

***Abstract*—In cyber physical systems, the communication infrastructure is faced with multiple challenges such as cyber attacks. In this paper, we propose a scheme for secure data communication in CPS against eavesdropping attacks. Our basic idea is combining multipath routing and secret sharing based on the special condition of CPS. This scheme achieves that even if an adversary compromises some nodes, it would not be able to recover the original data as long as the shares on compromised nodes are less than a predefined threshold. The simulation results demonstrate the effectiveness of our scheme.**

***Index Terms*—CPS, data delivery, security, secret sharing, multipath routing.**

I. Motivation

In a large-scale CPS, networks contains dense networks and sparse networks like the Internet and wireless sensor networks shown in Fig.1. In order to achieve secure data transmission, our basic idea is to split the data into multiple independent shares and then spread them into multiple independent paths which results in much lower probability of being attacked.

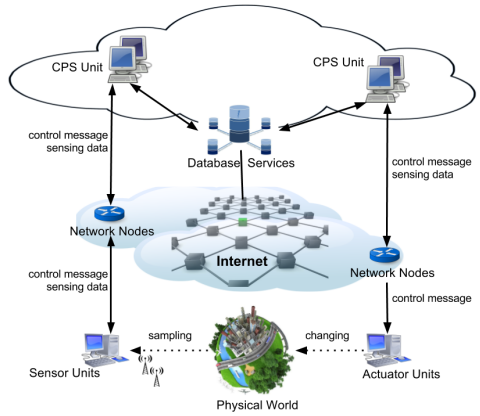

Fig. 1. A Prototype Structure of CPS

However, the number of usable independent paths is limited by sparse networks in the large CPS and a feasible mathematical model is required. Thus we utilize anchor node and graphical model to solve these problems.

II. System Model

Attack: Considering the eavesdropping attack under the assumption that a node will be eavesdropped once it is compromised and a path is eavesdropped once one node of it is compromised [1]. Given a network $G = (V, E)$ with nodes V and edges E, random variables X_i and Y_j refers to the secure state of node i and path j which are subject to the binomial distribution. Assume the probability that node i is compromised is p_i. Then,

$$X_i \sim \text{BIN}(1, p_i), i = 1, ..., V \tag{1}$$

$$Y_j \sim \text{BIN}(1, 1 - \prod_{X_i \in v} (1 - p_i)), j = 1, ..., m \tag{2}$$

where v is the set of nodes in path j.

Defense: We combine multipath routing and secret sharing to disperse the information [2] [3]. Multipath routing uses multiple alternative paths through a network which can be node-disjoint that are mutually independent and braided that are partially independent. Threshold secret sharing splits a secret into multiple independent shares, and any group of T or more shares can reconstruct the secret but a group of fewer than T shares reveal nothing. Accordingly, we split the data and allocate them into multiple paths in an optimal way.

III. Multipath Discovery: Find More Paths

Simple multipath routing is infeasible for the mixed network in CPS because that few node-disjoint paths exist in sparse network [4]. We proposed to choose a set of anchor nodes as trusted nodes to maximize the number of alternative paths [5]. In practice, we protect these anchor nodes from attacking and then the p_i of anchor nodes are zero. So our goal is to

$$\begin{aligned} \max \quad & f(I, G), \quad \text{where} \quad I_i \in \{0, 1\} \\ \text{subject to} \quad & \sum_{i=1}^{V} I_i \leq T_{anc} \\ & cap(X_i) \leq 1 + I_i \times T_{cap} \end{aligned} \tag{3}$$

where $f(\cdot)$ is the number of paths found by using path discovery algorithm, and T_{anc} refers to the maximum number anchor nodes we can set under our abilities and T_{cap} refers to the maximum frequency of anchor nodes that can be used in different paths. Thus, we can at most obtain $\triangle m$ more paths where $\triangle m = T_{anc} \times (T_{cap} - 1)$.

978-1-5386-4586-4/18/$31.00 ©2018 IEEE

IV. Defense Design: Achieve Required Security

In our scheme, we are aimed to achieve a flexible security requirement since different importance of delivered data. Specifically, we define the security requirement on how many paths adversaries need to compromise in order to success. Thus, following related constraints should be satisfied:

- The number of shares transmitted in each path should be less than T.
- The success of attack is limited: setting that it is impossible for adversaries to obtain T shares unless k paths are compromised.
- The total number of shares in all paths equals to N.

Correspondingly, assuming path j transmits C_j shares, the mathematical formulations for the constraints is given as follows

$$\begin{cases} C_j < T \\ N - \sum_{i=k}^{m} C_j < T \\ \sum_{i=1}^{m} C_j = N \end{cases} \tag{4}$$

Then, we could derive a necessary condition for limiting the range of c according to Equation(4), i.e.,

$$\begin{cases} \frac{N-T}{m-k+1} < C_j < T \\ \sum_{i=1}^{m} C_j = N \end{cases} \tag{5}$$

As we mentioned, our goal is to design optimal scheme that minimize the probability of eavesdropping original information by adversaries. In other words, it is to maximize the probability of eavesdropping less than T shares by adversaries. This optimal problem is expressed as,

$$\begin{aligned} \max \quad & P_r[\sum_{j=1}^{m} Y_j C_j < T] \\ \text{subject to} \quad & \frac{N-T}{m-k+1} < C_j < T \\ & \sum_{i=1}^{m} C_j = N \\ & N - T \geq 0 \\ & N, M, T \in \mathbb{Z}^+ \\ & C_j \in \mathbb{Z} \end{aligned} \tag{6}$$

In order to solve this optimization problem, we use the Bayesian graphical model to express the objective function in (6). Let $Z_j = \sum_{j=1}^{m} Y_j C_j$, the joint distribution of all random variables is

$$\begin{aligned} P(X, Y, N, T, C, Z | k, p_i, G) = P(Z|Y, C) \prod_{j=1}^{m} P(C_j | N, T, k) \\ P(T|N) P(N|T_N) \prod_{j=1}^{m} P(Y_j | X, G) \prod_{i=1}^{V} P(X_i | p_i) \end{aligned} \tag{7}$$

where we assume that $N \sim \text{UNI}(2, T_N)$ and T_N is the maximum value of N, then $T \sim \text{UNI}(3, T)$. Thus, We can obtain the conditional probabilities $P(C_j|N, T, k)$ and $P(Z|Y, C)$ by using Monte Carlo simulations. Thus $P(Z)$ is the marginal probability of $P(X, Y, N, T, C, Z|k, p_i, G)$, and the objective function should be $P(Z < T)$.

V. Performance Evaluation

We evaluate the performance of our scheme in a sparse network $G(V, E)$ which is generated randomly with $|V| \in [20 : 100]$. Set $T_{anc} = 4$ and $T_{cap} = 3$, the result is shown in Fig.2.

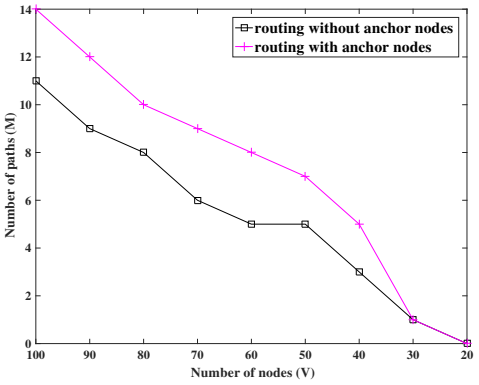

Fig. 2. Capability of Path Discovery

Our path discovery method can generally identify at least two more paths than the normal path discovery method.

Then, we simulate the whole data delivery process based on Monte Carlo method with 13 alternative paths of a network with 100 nodes, and analyze the performance of our final scheme. Assume that the probability that a node is compromised is $p_i = 0.1$, $T_N = 50$, $k = m$, then we obtained the average probability of $(Z < T)$ is 0.95 which means the probability of being eavesdropped decreases from 0.1 to 0.05 generally even though we do not choose the optimal design.

However, in a mathematical way, the objective function is based on the product of all the conditional density functions. So the challenge is to calculate all the conditional density functions and then do integration mathematically rather than by simulations, and in this way the result will be accurate.

References

[1] T. N. Dinh and M. T. Thai, "Network under joint node and link attacks: Vulnerability assessment methods and analysis," *IEEE/ACM Transactions on Networking*, vol. 23, no. 3, pp. 1001–1011, 2015.

[2] W. Lou, W. Liu, and Y. Fang, "Spread: Enhancing data confidentiality in mobile ad hoc networks," in *INFOCOM 2004. Twenty-third AnnualJoint Conference of the IEEE Computer and Communications Societies*, vol. 4. IEEE, 2004, pp. 2404–2413.

[3] A. Beimel, "Secret-sharing schemes: A survey." *IWCC*, vol. 6639, pp. 11–46, 2011.

[4] Z. Ye, S. V. Krishnamurthy, and S. K. Tripathi, "A framework for reliable routing in mobile ad hoc networks," in *INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications. IEEE Societies*, vol. 1. IEEE, 2003, pp. 270–280.

[5] S. Tian, X. Zhang, X. Wang, P. Sun, and H. Zhang, "A selective anchor node localization algorithm for wireless sensor networks," in *Convergence Information Technology, 2007. International Conference on*. IEEE, 2007, pp. 358–362.