IEICE TRANS. INF. & SYST., VOL.E98–D, NO.5 MAY 2015

LETTER

On the Probability of Certificate Revocation in Combinatorial Certificate Management Schemes*

Dae Hyun YUM[†a)], ***Member***

SUMMARY To enhance the privacy of vehicle owners, combinatorial certificate management schemes assign each certificate to a large enough group of vehicles so that it will be difficult to link a certificate to any particular vehicle. When an innocent vehicle shares a certificate with a misbehaving vehicle and the certificate on the misbehaving vehicle has been revoked, the certificate on the innocent vehicle also becomes invalid and is said to be covered. When a group of misbehaving vehicles collectively share all the certificates assigned to an innocent vehicle and these certificates are revoked, the innocent vehicle is said to be covered. We point out that the previous analysis of the vehicle cover probability is not correct and then provide a new and exact analysis of the vehicle cover probability.

key words: *vehicular communications, public key infrastructure, privacy, anonymity, certificate revocation*

1. Introduction

In vehicular networks, there is a strong correlation between a vehicle's identity and that of the driver [1]. To encourage drivers' participation in the vehicular networks, privacy-preservation techniques that conceal vehicles' identity should be employed [2]. However, a conventional public-key certificate (e.g., the ITU-T X.509 standard and IETF RFC 5280) includes plaintext information about the subject of the certificate which adversaries can use to track a vehicle and to determine which messages are sent from the vehicle.

The challenge for designing a privacy-preserving public key infrastructure is to make public-key certificates anonymous while meeting other design goals such as achieving high scalability and robustness [3]. For example, the highest level of anonymity can be easily achieved when all vehicles use the same certificate. However, revoking the certificate will require changes to all vehicles, making the system unscalable.

Combinatorial certificate management schemes assign each certificate to a large enough group of vehicles so that it will be difficult for adversaries to link a certificate to any particular vehicle [3]–[6]. After a certificate authority (CA) creates a shared certificate pool, each certificate in the pool is assigned to multiple vehicles and each vehicle is assigned multiple certificates. When a certificate needs to be revoked, the CA revokes the certificate by posting its identifier (e.g., a serial number) on a certificate revocation list (CRL) and making the CRL available to all vehicles. When a certificate assigned to a misbehaving vehicle is revoked, all the other vehicles sharing this certificate will also not be able to use it.

When an innocent vehicle shares a certificate with a misbehaving vehicle and the certificate has been revoked, this certificate on the innocent vehicle is said to be covered. A group of misbehaving vehicles may collectively share all the certificates assigned to an innocent vehicle. When these certificates are revoked, all the certificates assigned to the innocent vehicle will also be revoked and the innocent vehicle is said to be covered. When all the certificates assigned to a misbehaving vehicle have been revoked, this vehicle is said to be revoked [3].

Assume that n certificates are assigned to each vehicle and m misbehaving vehicles are revoked. It is trivial to compute the vehicle cover probability for $n = m = 1$. However, if n is larger than one (or both n and m are larger than one), then probabilistic events are no longer independent and the computation of the vehicle cover probability becomes complicated. We show that the previous analysis of the vehicle cover probability is not correct and then provide a new and exact analysis of the vehicle cover probability by considering all the dependency between various probabilistic events.

2. Preliminaries

The binomial coefficient $\binom{a}{b}$ can be defined recursively by $\binom{a}{b} = \binom{a-1}{b-1} + \binom{a-1}{b}$ for all integers $a,\ b > 0$ with initial conditions $\binom{a}{0} = 1$ for $\forall a \in \mathbb{N}$ and $\binom{0}{b} = 0$ for all integers $b > 0$. If $b \notin \{0, 1, \ldots, a\}$, then $\binom{a}{b} = 0$. The factorial formula for the binomial coefficient is given by $\binom{a}{b} = \frac{a!}{b!(a-b)!}$ for $0 \le b \le a$, where $a! = \Pi_{i=1}^{a} i$ and $0! = 1$.

Let $J \subset \mathbb{N}^+$ be a set of positive integers. The principle of inclusion-exclusion states that for finite sets B_j where $j \in J$, the number of elements in the union of B_j can be obtained by the following formula:

$$\Big|\bigcup_{j\in J} B_j\Big| = \sum_{\emptyset\neq K\subset J} (-1)^{|K|-1} \Big|\bigcap_{k\in K} B_k\Big| \tag{1}$$

If the size of the intersection sets in Eq. (1) depend only on

Manuscript received January 13, 2015.
Manuscript publicized February 18, 2015.

[†]The author is with the Department of Information and Communication Engineering, Myongji University, Yongin, Gyeonggido, 449–728, Republic of Korea.

*This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning (NRF-2013R1A1A1009630) and by the Dual Use Technology Program.

a) E-mail: dhyum@mju.ac.kr

DOI: 10.1587/transinf.2015EDL8012

the number of sets in the intersections and not on which sets appear, i.e., $|K_1| = |K_2|$ implies $|\bigcap_{k\in K_1} B_k| = |\bigcap_{k\in K_2} B_k|$, then we have

$$\Big|\bigcup_{j\in J} B_j\Big| = \sum_{\emptyset\neq K\subset J} (-1)^{|K|-1}\Big|\bigcap_{k\in K} B_k\Big| = \sum_{s=1}^{|J|}(-1)^{s-1}\binom{|J|}{s}d_s \quad (2)$$

where d_s is the cardinality of the intersection of s finite sets B_k for $|K| = s$, i.e., $d_s = |\bigcap_{k\in K} B_k|$ for every s-element set K.

More information about the binomial coefficient and the principle of inclusion-exclusion can be found in [7].

3. Vehicle Cover Probability

3.1 Previous Analysis

We review the previous analysis of the vehicle cover probability in the combinatorial certificate management scheme [3, Chapter 16], [4], [6].

For a set of indices $I = \{1, 2, \ldots, N\}$, the CA creates a shared certificate pool of $C = \{C_i \mid i \in I\} = \{C_1, C_2, \ldots, C_N\}$ where C_i denotes a certificate. Each vehicle is given $n\ (\leq N)$ certificates with their associated private and public keys from the shared certificate pool. If m misbehaving vehicles have been revoked, the probability that any given certificate on an innocent vehicle is not covered will be $(1 - \frac{n}{N})^m$. Therefore, the probability that an innocent vehicle is covered is

$$\Pr_{\text{cover}}(N, n, m) = \left(1 - \left(1 - \frac{n}{N}\right)^m\right)^n. \quad (3)$$

Remark 1. In the literature, Eq. (3) is sometimes described as an approximation. However, we checked that the previous analysis of Eq. (3) was always derived as an exact formula rather than an approximation. For example, [4] cites [8] as the origin of Eq. (3) and one can find that Eq. (3) was derived as an exact formula in [8]; specifically, see the second equation of Lemma 6 in [8].

3.2 Counterexample

We revisit the validity of Eq. (3) with a simple counterexample of $(N, n, m) = (3, 2, 1)$. Without loss of generality, suppose that the CA generates a shared certificate pool $C = \{C_1, C_2, C_3\}$ and gives $\{C_1, C_2\}$ to an innocent vehicle. For vehicle cover probability, we compute the probability that both C_1 and C_2 are covered when a misbehaving vehicle is revoked. Let α be the event that the innocent vehicle is covered and α_i $(i = 1, 2)$ be the event that the certificate C_i on the innocent vehicle is covered when a misbehaving vehicle is revoked. Then, the vehicle cover probability is

$$\Pr[\alpha] = \Pr[\alpha_1 \wedge \alpha_2] = \Pr[\alpha_1]\Pr[\alpha_2|\alpha_1] \quad (4)$$

where $\Pr[\alpha_2|\alpha_1]$ denotes the conditional probability of α_2 given α_1.

Let (C_j, C_k) be the certificates assigned to the misbehaving vehicle. As (C_j, C_k) is chosen randomly from the shared certificate pool $C = \{C_1, C_2, C_3\}$, (C_j, C_k) will be one of three pairs (C_1, C_2), (C_1, C_3), (C_2, C_3). Therefore, the probability that C_1 on the innocent vehicle is covered is $\frac{2}{3}$, i.e., $\Pr[\alpha_1] = \frac{2}{3}$; similarly we can get $\Pr[\alpha_2] = \frac{2}{3}$. We now turn to $\Pr[\alpha_2|\alpha_1]$. What is the probability that C_2 on the innocent vehicle is covered when C_1 has already been covered? Since C_1 has been covered (or the event α_1 has already occurred), we know that (C_j, C_k) will be either (C_1, C_2) or (C_1, C_3) and thus $\Pr[\alpha_2|\alpha_1] = \frac{1}{2}$. Note that two events α_1 are α_2 are not independent because $\left(\Pr[\alpha_2] = \frac{2}{3}\right) \neq \left(\Pr[\alpha_2|\alpha_1] = \frac{1}{2}\right)$. Finally, the vehicle cover probability is $\Pr[\alpha] = \Pr[\alpha_1]\Pr[\alpha_2|\alpha_1] = \frac{2}{3}\cdot\frac{1}{2} = \frac{1}{3} \simeq 0.33$.

Unlike the above analysis, the previous analysis Eq. (3) gives $\Pr_{\text{cover}}(3, 2, 1) = \left(1 - \left(1 - \frac{2}{3}\right)\right)^2 = \left(\frac{2}{3}\right)^2 = \frac{4}{9} \simeq 0.44$. What is wrong with Eq. (3)? The previous analysis in Sect. 3.1 computes $\Pr[\alpha_1] = \frac{2}{3}$ and concludes $\Pr[\alpha] = \left(\frac{2}{3}\right)^2$. In other words, the previous analysis implicitly assumes that the events α_1 and α_2 are independent (i.e., $\Pr[\alpha_2] = \Pr[\alpha_2|\alpha_1]$) and wrongly computes the vehicle cover probability as $\Pr[\alpha] = \Pr[\alpha_1]\Pr[\alpha_2]$.

Remark 2. Actually, the above counterexample can be computed directly; if the innocent vehicle has (C_1, C_2) and the misbehaving vehicle has one of (C_1, C_2), (C_1, C_3), (C_2, C_3), the innocent vehicle is covered only in the case of $(C_j, C_k) = (C_1, C_2)$ and therefore $\Pr[\alpha] = \frac{1}{3}$.

3.3 New Analysis

To derive the exact formula for the vehicle cover probability, we must consider the dependency between probabilistic events. Recall that each vehicle is given $n\ (\leq N)$ certificates from the shared certificate pool $C = \{C_i \mid i \in I\} = \{C_1, C_2, \ldots, C_N\}$ where $I = \{1, 2, \ldots, N\}$. Suppose that an innocent vehicle $\mathcal{V}$ is given n certificates $\{C_{\lambda_1}, C_{\lambda_2}, \ldots, C_{\lambda_n}\}$ where $\lambda_i \in I$ for $1 \leq i \leq n$. Let α be the event that the innocent vehicle $\mathcal{V}$ is covered. We compute the vehicle cover probability $\Pr[\alpha]$ that all n certificates $\{C_{\lambda_1}, C_{\lambda_2}, \ldots, C_{\lambda_n}\}$ are covered when m misbehaving vehicles are revoked.

Let β_J be the event that the m misbehaving vehicles collectively possess $\{C_j \mid j \in J\}$ for $J \subset I$. As each vehicle is assigned n certificates and some certificates can be repeatedly assigned, the size of J is $n \leq |J| \leq mn$. When m misbehaving vehicles are revoked, all certificates in $\{C_j \mid j \in J\}$ are revoked. In this case, the innocent vehicle $\mathcal{V}$ is covered if and only if $\{C_{\lambda_1}, C_{\lambda_2}, \ldots, C_{\lambda_n}\} \subset \{C_j \mid j \in J\}$. Therefore, we have $\Pr[\alpha|\beta_J] = \frac{\binom{|J|}{n}}{\binom{N}{n}}$ where $\binom{|J|}{n}$ is the number of ways of choosing n innocent certificates $\{C_{\lambda_1}, C_{\lambda_2}, \ldots, C_{\lambda_n}\}$ from $\{C_j \mid j \in J\}$ and $\binom{N}{n}$ is the total number of ways of choosing n innocent certificates from the shared certificate pool $C = \{C_1, C_2, \ldots, C_N\}$. With conditional probabilities, the vehicle cover probability can be expressed as follows.

$$\Pr[\alpha] = \sum_{J} \Pr[\alpha|\beta_J] \Pr[\beta_J] = \sum_{J} \frac{\binom{|J|}{n}}{\binom{N}{n}} \Pr[\beta_J] \tag{5}$$

where $J \subset I$ and $n \le |J| \le mn$.

For the event β_J to occur, two conditions should be satisfied; (1) n certificates of each misbehaving vehicle should belong to $\{C_j \mid j \in J\}$ and (2) all certificates in $\{C_j \mid j \in J\}$ should be assigned to the misbehaving vehicles. With the two conditions, how many ways of choosing certificates of m misbehaving vehicles do we have? Let W_J be the set of ways of choosing certificates of m misbehaving vehicles only from $\{C_j \mid j \in J\}$, i.e., $|W_J| = \binom{|J|}{n}^m$. Even though any case belonging to W_J satisfies the first condition, W_J also includes the case that some certificates in $\{C_j \mid j \in J\}$ are not assigned to the misbehaving vehicles. Therefore, we should subtract from W_J the cases that do not satisfy the second condition. The cases (in W_J) where there is at least one certificate in $\{C_j \mid j \in J\}$ that is not assigned to the misbehaving vehicles can be expressed as $\bigcup_{j\in J} W_{J-\{j\}}$. The probability $\Pr[\beta_J]$ can be computed as follows.

$$\begin{aligned}\Pr[\beta_J] &= \frac{|W_J| - |\bigcup_{j\in J} W_{J-\{j\}}|}{|W_I|} \\ &= \frac{\binom{|J|}{n}^m - |\bigcup_{j\in J} W_{J-\{j\}}|}{\binom{N}{n}^m}\end{aligned} \tag{6}$$

where W_I is the set of ways of choosing certificates of m misbehaving vehicles from the shared certificate pool $C = \{C_1, C_2, \ldots, C_N\}$ and the size of W_I is $|W_I| = \binom{|I|}{n}^m = \binom{N}{n}^m$.

The probability $|\bigcup_{j\in J} W_{J-\{j\}}|$ can be computed by using the principle of inclusion-exclusion (i.e., Eq. (2)) as follows.

$$\begin{aligned}\Big|\bigcup_{j\in J} W_{J-\{j\}}\Big| &= \sum_{\emptyset\neq K\subset J} (-1)^{|K|-1} \Big|\bigcap_{k\in K} W_{J-\{k\}}\Big| \\ &= \sum_{s=1}^{|J|} (-1)^{s-1} \binom{|J|}{s} \binom{|J|-s}{n}^m\end{aligned} \tag{7}$$

where $\big|\bigcap_{k\in K} W_{J-\{k\}}\big|$ is equal to the number of ways of choosing certificates of m misbehaving vehicles only from $\{C_i \mid i \in J - K\}$ and we get $\big|\bigcap_{k\in K} W_{J-\{k\}}\big| = \binom{|J|-s}{n}^m$ for any s-element set K (i.e., $|K| = s$).

Finally, we can get the following formula for the vehicle cover probability from Eq. (5), Eq. (6), and Eq. (7).

$$\begin{aligned}&\Pr^{\text{new}}_{\text{cover}}(N, n, m) \\ &= \Pr[\alpha] \\ &= \sum_{J} \frac{\binom{|J|}{n}}{\binom{N}{n}} \Pr[\beta_J] \qquad (\text{for } J \subset I \text{ and } n \le |J| \le mn) \\ &= \sum_{J} \left(\frac{\binom{|J|}{n}}{\binom{N}{n}} \cdot \frac{\binom{|J|}{n}^m - |\bigcup_{j\in J} W_{J-\{j\}}|}{\binom{N}{n}^m} \right) \\ &= \sum_{J} \left(\frac{\binom{|J|}{n}}{\binom{N}{n}} \cdot \frac{\binom{|J|}{n}^m - \sum_{s=1}^{|J|} (-1)^{s-1} \binom{|J|}{s} \binom{|J|-s}{n}^m}{\binom{N}{n}^m} \right)\end{aligned}$$

Table 1 Sample values of $err(10000, 5, m)$.

m	Eq. (3)	Eq. (8)	$err(10000, 5, m)$
1	3.125E-17	1.201E-18	2501.6
2	9.988E-16	3.023E-16	230.4
3	7.575E-15	3.598E-15	110.5
4	3.188E-14	1.855E-14	71.8
5	9.717E-14	6.350E-14	53.0
6	2.415E-13	1.701E-13	42.0
7	5.213E-13	3.870E-13	34.7
8	1.015E-12	7.835E-13	29.6
9	1.827E-12	1.453E-12	25.7
10	3.090E-12	2.517E-12	22.8

$$= \sum_{\gamma=n}^{mn} \left(\binom{N}{\gamma} \cdot \frac{\binom{\gamma}{n}}{\binom{N}{n}} \cdot \frac{\binom{\gamma}{n}^m - \sum_{s=1}^{\gamma} (-1)^{s-1} \binom{\gamma}{s} \binom{\gamma-s}{n}^m}{\binom{N}{n}^m} \right) \tag{8}$$

where $\binom{N}{\gamma}$ in the last equation is the number of γ-element sets J with $J \subset I$ and $n \le |J| \le mn$. Note that the binomial coefficient in Eq. (8) should be evaluated as defined in Sect. 2. For example, $\binom{a}{0} = 1$ for $\forall a \in \mathbb{N}$ and $\binom{0}{b} = 0$ for all integers $b > 0$. If $b \notin \{0, 1, \ldots, a\}$, then $\binom{a}{b} = 0$.

3.4 Numerical Comparison

To evaluate the numerical difference between the previous formula (Eq. (3)) and the new formula (Eq. (8)), we define a function $err(N, n, m)$ as follows.

$$err(N, n, m) = \left(\frac{\text{Eq. (3)} - \text{Eq. (8)}}{\text{Eq. (8)}} \right) \times 100\ (\%) \tag{9}$$

In Table 1, we provide the numerical values for the parameter $N = 10000$ and $n = 5$ of the baseline system in [6].

Table 1 shows that the difference between two formulas is very large for small values of m and becomes smaller (but not negligible) as m grows. Therefore, Eq. (3) can only be used for a loose upper bound of the vehicle cover probability and Eq. (8) should be used to compute the exact probability.

4. Conclusion

When a misbehaving vehicle in the combinatorial certificate management scheme is revoked, shared certificates on an innocent vehicle also become invalid. Even though the computation of the vehicle cover probability may seem to be easy and simple at first, it turns out that the computation is relatively complicated. This is because various probabilistic events are not independent but affect each other. Therefore, care should always be taken on the computation of probabilities when various dependent events are considered. In addition to the vehicle cover probability, there also have been other problems in which old analysis was recently corrected by revisiting the dependency between probabilistic events; e.g., [9]–[11].

References

[1] G. Yan, S. Olariu, J. Wang, and S. Arif, "Towards providing scalable and robust privacy in vehicular networks," IEEE Trans. Parallel

Distrib. Syst., vol.25, no.7, pp.1896–1906, 2014.
[2] R. Lu, X. Lin, Z. Shi, and X.S. Shen, "A lightweight conditional privacy-preservation protocol for vehicular traffic-monitoring systems," IEEE Intelligent Systems, vol.28, no.3, pp.62–65, 2013.
[3] L. Delgrossi and T. Zhang, Vehicle Safety Communications: Protocols, Security, and Privacy, Wiley, 2012.
[4] E. van den Berg, T. Zhang, and S. Pietrowicz, "Blend-in: A privacy-enhancing certificate-selection method for vehicular communication," IEEE Trans. Veh. Technol., vol.58, no.9, pp.5190–5199, 2009.
[5] S. Tengler, S. Andrews, and R. Heft, "Digital certificate pool," United States Patent No.US 7734050 B2, June 8, 2010.
[6] R.G. White, S. Pietrowicz, E. van den Berg, G.D. Crescenzo, D. Mok, R. Ferrer, T. Zhang, and H. Shim, "Privacy and scalability analysis of vehicular combinatorial certificate schemes," the 6th IEEE Conference on Consumer Communications and Networking Conference (CCNC'09), pp.624–628, 2009.
[7] R.A. Brualdi, Introductory Combinatorics, 5th ed., Prentice Hall, 2009.
[8] J.A. Garay, J. Staddon, and A. Wool, "Long-lived broadcast encryption," CRYPTO 2000, pp.333–352, 2000.
[9] P. Bose, H. Guo, E. Kranakis, A. Maheshwari, P. Morin, J. Morrison, M.H.M. Smid, and Y. Tang, "On the false-positive rate of bloom filters," Inf. Process. Lett., vol.108, no.4, pp.210–213, 2008.
[10] K.J. Christensen, A. Roginsky, and M. Jimeno, "A new analysis of the false positive rate of a bloom filter," Inf. Process. Lett., vol.110, no.21, pp.944–949, 2010.
[11] D.H. Yum and P.J. Lee, "Exact formulae for resilience in random key predistribution schemes," IEEE Trans. Wireless Commun., vol.11, no.5, pp.1638–1642, 2012.