Poster:Measuring Location Privacy with Process calculus

GUO Yunchuan, YIN Lihua, LI Chao
Institute of Information Engineering
Chinese Academy Sciences
Beijing, China
{guoyunchuan,yinlihua,lichao}@iie.ac.cn

CUI Xiang
Institute of Computing Technology
Chinese Academy Sciences
Beijing, China
cuixiang@ict.ac.cn

***Abstract*—In the mobile wireless Internet, location privacy is serious concerns. As a response to these concerns, many location-privacy protection mechanisms (LPPMs) have been proposed. However, the existing work doesn't integrate formal models into assessments, which leads a huge gap: after designing a LPPM, we have to select formal methods to formalize them, and select another evaluation metric to measure them. In this paper, we propose a probabilistic process calculus to model the LPPMs and use the relative entropy to measure the degree of location privacy LPPMs can leak. Our work decreases the gap between the formalization and the measurement for LPPMs.**

Keywords—Location Privacy; Process Calculus;Measurement ;

I. Introduction

In the mobile wireless Internet, the wide usage of personal communication devices equipped with high-precision localization capabilities brings more convenience, but it also causes serious privacy risks, location privacy in particular, for their owners. In order to relieve the risk of location privacy, researchers make a large number of studies, roughly classified into 2 categories: (1) Analyzing the threats and formalizing the attacks on location privacy, designing the corresponding LPPMs for the different contexts[1], (2) Designing an appropriate evaluation metric for location privacy based on a sound theoretical model[2]. Nowadays, the need for applying formal tools to privacy has been widely recognized. To our knowledge, however, formalizing the LPPMs and designing an appropriate evaluation metric work independently: when formalizing the LPPMs, quantitative evaluation metric is not considered and vice versa. This leads to a huge gap: after designing the LPPMs, we have to select an appropriate formal method to formalize them, and select another evaluation metric to measure them. Worse, the accuracy of measuring location privacy, if formal models don't consist with measurements, will decrease. This increases the difficult of guaranteeing location privacy.

In this paper, we propose a probabilistic process calculus, called δ-calculus, to model the LPPMs and use the relative entropy as a metric to measure the degree of privacy LPPMs can guarantee. δ-calculus is obtained by extending π calculus via adding *location calculus* which models the location(maybe a false location) of nodes, *probabilistic choice* of location which models the probability distribution of locations of a node, *communication radius* which denotes the communication range of a node, *movement operator* which models the movement of physical location of nodes.

II. δ-Calculus

The mobile wireless Internet comprises a set of communication devices, called *nodes*, each of which runs a *process* at the location of being randomly distributed and may move to another location. We use N and P to denote the sets of nodes and all processes of nodes, respectively, with *M*, *N* ranging over nodes and *P*, *Q* ranging over processes.

The syntax of δ-calculus, describing communication between nodes, is defined as follows.

$$N,M ::= z[|P|](\sum_i p_i l_i, r) \quad | \quad M \mid N \quad | \quad (\nu l)N \quad | \quad (\nu x)N$$

Its informal semantics is as follows. In $z[|P|](\sum_i p_i l_i, r)$, *z* is for the node name (for example node ID) and *r* represents the *communication radius.* Note that if *P* is not the communication between nodes, *r* is meaningless; the p_i's represents positive probabilities, that is , $p_i \in (0,1]$ and $\sum_i p_i = 1$; the l_i's are all possible locations of node *z.* $z[|P|](\sum_i p_i l_i, r)$ represents that node *z* runs process *P* at location l_i with probability p_i, and the *z*'s communication distance is less than *r*. *M*|*N* represents the *parallel composition* of node *M* and node *N*, The symbols ν is the *restriction* operator, (νl) and (νx) are used to restrict the scope of locations and variables, respectively.

The syntax of δ-calculus processes is defined as the following grammar:

$$P,Q ::= \bar{S}T.P \quad | \quad S(x).P \quad | \quad \sum_i p_i P_i \quad | \quad P|Q \quad | \quad !P \quad | \quad (\nu x)P \quad | \quad \mathrm{MV}\ l.P \quad | \quad nil$$

Process $\bar{S}T.P$ and process $S(x).P$ mean "sending *T* along channel *S* before running *P*" and "receiving *x* along channel *S* before running *P*", respectively. $\sum_i p_i P_i$ is the *probabilistic choice* operator, meaning that P_i is selected with probability p_i, where $p \in (0,1]$ and $\sum_i p_i = 1$. Operators | , ! and (νx) are for *parallel composition*, *replication* and *restriction*, respectively. MV *l.P* makes a given node move into location *l* and then executes *P*. *nil* represents an empty process.

In δ-calculus processes, *S* and *T* range over terms and are defined as the following syntax.

$$S,T ::= x \quad | \quad a$$

Where x ranges over a countable set of variables and a ranges over a countable set of *channel names*..

III. Measuring Location Privacy

In order to provide location privacy, many LPPMs add location noises - a certain set of discrete locations (including false locations and true location), denoted by LOC. In order to measure the degree of location privacy, we define a renaming function $f_{LOC}: M \to M$, which permutes elements in LOC and identity elsewhere. That is, for each location in LOC, function f is executed and identity elsewhere, where $f: LOC \to LOC$ such that $f(l) \neq l$ and $l_1 \neq l_2$ implies $f(l_1) \neq f(l_2)$. We use F_{LOC} to denote the set of all renaming functions f_{LOC} on LOC.

Given a LPPM *M* modeled by δ-calculus, *M* 's behavior can be obtained via unfolding the δ-calculus. According to the semantics, its behavior is considered as as a trace distribution or a set of trace distributions, denoted by *tds*(*M*). Given a set *X* of trace distributions, a metric *D* on a set *X* is a function $D: X \times X \to R^+$ satisfying coincidence axiom, symmetry and subadditivity, where R^+ is the set of non-negative real numbers.

Definition 1. Given a metric *D* and a LPPM *M*, *M* is strong privacy-preservation under *D* on a set of locations *LOC* if $\forall f_{loc} \in F_{LOC}: D(M, f_{loc}(P)) = 0$; *M* is called ς-privacy *if* $\forall f_{loc} \in F_{LOC}: D(M, f_{loc}(P)) \leq \varsigma$.

Theorem. Given two metric D_1 and D_2, and a LPPM *M*, *M* is strong privacy-preservation under D_1 iff *M* is strong privacy-preservation under D_2. ς-privacy preservation of *M* under D_1 doesn't imply ς-privacy preservation under D_2.

In the information theory, although relative entropy is quasi-metric, it satisfies nonnegative and coincidence axiom, thus can be used to measure location privacy.

Definition 2. For discrete probability distributions u and u', the relative entropy of u' from u is defined to be

$$D_{KL}(u \| u') = \sum_i \log(\frac{u(i)}{u'(i)}) u(i)$$

where $0 \log \frac{0}{0} = 0$, $0 \log \frac{0}{q} = 0$, $0 \log \frac{q}{0} = \infty$ and $i \in I$ is an index set. Although the behavior of a node may be a set of trace distributions, only a trace distribution is considered in this paper, thus we have the following measurement for location privacy.

Measuring Location Privacy: Given node *M* under the protection of LPPM, if *M*'s behavior is a trace distribution, then the amount of leakage of local privacy is $\sup_{f_{loc} \in F_{LOC}} D_{KL}(tds(M) \| tds(f_{loc}(M)))$.

Example. Consider a wireless communication system, where node z_1 at location $l_1$sends information to base station z_b, and attacker z_a tries to obtain z_1's location by monitoring the communication. In order to provide location privacy, z_1 is protected by LPPM via adding one false location l'_1. The system can be modeled as:

$$M = Node|BaseStation|Attacker$$

$$Node = z_1[|\overline{send}(info)|](p_1 l_1 + (1-p_1)l'_1, r_1)$$

$$BaseStation = z_b[|send(x)|](l, r)$$

$$Attacker = z_a[|send(x)|](l, r)$$

That is, node z_1 sends information at location l_1 with probability p_1 and at location l'_1 with probability 1-p_1; Base station z_b receives information at location *l* and attacker z_a monitors information at location *l*. If z_a is in the range of communication radius of z_1 (that is, the distance between $l_1(l'_1)$ and l is less than communication radius r_1 of z_1), then z_a can receive the information sent by z_1 and infer z_1's location.

The only permutation function *f* on *LOC* is $f(l_1) = l'_1$ and $f(l'_1) = l_1$. Thus, $f_{loc}(Node) = z_1[|\overline{send}(info)|](p_1 l'_1 + (1-p_1) l_1, r_1)$, the amount of leakage of location privacy is $D_{KL}(tds(M) \| tds(f_{loc}(M)) = p_i \log \frac{p_i}{1-p_i} + (1-p_i) \log \frac{1-p_i}{p_i}$.

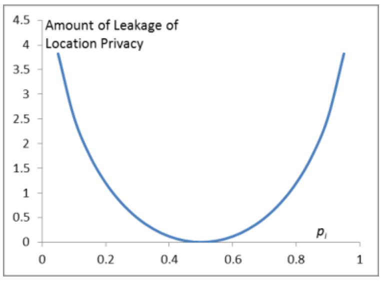

Fig. 1. Amount of Leakage of Location Privacy

Fig.1 gives the amount of leakage of location privacy of *M* with the change of p_i. The figure shows: the amount of location obtained by z_a is 0 when p_i*=0.5* (that is, z_a cannot infer the location of z_1) and the amount of leakage of location privacy is infinite (meaning that z_a can infer the true location of z_1) when $p_i \to 0$ or $p_i \to 1$. This shows that the result of measurements consists with the capability of the LPPM which can guarantee. Thus, our measurment is accurate.

IV. Conlusions

In this paper, we propose δ-calculus to measure location privacy which decreases the gap between the formalization and the measurement for LPPMs.

Acknowledgment

This research is supported by the National Natural Science Foundation of China (Grant No. 61070186, 61100186, 61202409).

References

[1] Shokri Reza, Theodorakopoulos George, Troncoso Carmela, Hubaux Jean-Pierre, Le Boudec Jean-Yves. Protecting Location Privacy: Optimal Strategy against Localization Attacks. Proceedings of the Conference on Computer and Communications Security. 2012: 617-627.

[2] Shokri Reza, Theodorakopoulos George, Le Boudec J, Hubaux J. Quantifying Location Privacy. IEEE Symposium on Security and Privacy. 2011: 247-262.