一种有效的匿名分析算法*)

郭亚军[1,2] 何炎祥[1]

(武汉大学计算机学院 武汉 430079)[1] (华中师范大学计算机科学系 武汉 430079)[2]

摘 要 匿名是保护用户隐私的主要方法。当前的研究主要集中在设计具体匿名方案,较少涉及如何评估匿名机制的匿名性。本文根据匿名与不可关联性具有紧密的关系,设计了一套完备的匿名性推理系统,在此基础上给出了匿名分析算法。该算法能够发现隐藏的匿名漏洞,有效地评估现有匿名机制的安全性,并为设计一个新的匿名保护方案提供支持。

关键词 匿名,不可关联性,隐私

An Efficient Anonymity Analysis Algorithm

GUO Ya-Jun[1,2] HE Yan-Xiang[1]

(School of Computer Science, Wuhan University, Wuhan 430079)[1]

(Department of Computer Science, Central China Normal University, Wuhan 430079)[2]

Abstract Anonymity is the main approach to protect users' privacy. Currently researches mainly concentrate on the concrete anonymity schemes, but seldom concern how to evaluate the security of an anonymity mechanism. In this paper, a complete anonymity reasoning system is addressed that it relies on the closely relationship between anonymity and unlinkability. Based on the anonymity reasoning system, an anonymity analysis algorithm is designed, which it can find out the concealing anonymity exposure, can evaluate and analyze the security of the anonymous mechanism, and can help design a new anonymity protection scheme.

Keywords Anonymity, Unlinkability, Privacy

1 引言

匿名是保护用户隐私的一个重要手段。匿名确保一个用户在使用资源或者服务时,没有显示自己的身份。目前已经提出了许多匿名解决方案,这些方案主要分为三类。一个主要方法是通过隐藏用户的身份来实现匿名,如 MIX[1]、Crowd[2]等。第二种方法是借助盲签名来实现匿名[3,4]。第三种方法是使用假名[5,6],当使用服务时,每次改变假名,类似给一个用户定义多个身份,用户根据具体情况使用不同的身份。但是很少涉及如何分析用户匿名的安全性,如何设计一个新的匿名机制等。隐私钻石模型[7]是一个简单的隐私分析模型,它通过是否存在一个传递闭包来判断隐私保护性,当攻击者从多方面的信息推出一个用户的身份,该模型将无法判断隐私漏洞。

2 匿名与不可关联性

匿名表示无法从一组主体集合中确切定出其中某一个主体与某行动相关联。例如一组主体都可能是某信息的发送者,如果无法定出到底谁是信息的真实发送者,就称该主体是发送者匿名。

匿名性与关联性是紧密耦合的两个概念。关联性(linkability)是指系统中的对象(如主体、消息、事件、行动等)之间在某一行为发生前与发生后的联系。匿名性往往以关联的形式表述。不可关联性(Unlinkability)可以是系统中任何对象之间的不可关联性的。对一个用户而言,用户和它的行动不可关联性表示为不能将用户与它的行动(如访问某个资源)联系起来。一个对象的匿名性是指它与任何身份标记不相关联。

在一个匿名通信系统中的不可关联性可分为消息对发送者的不可关联性、消息对接收者的不可关联性、消息对发送者与接收者联系的不可关联性,对应的匿名性可表述为发送者匿名(sender anonymity)、接收者匿名(recipient anonymity)、通信关系匿名(relationship anonymity)。这三种类型中,发送者匿名、接收者匿名要强于通信关系匿名,具有以下关系:

发送者匿名⇒通信关系匿名

接收者匿名⇒通信关系匿名

如果发送者匿名或接收者匿名成立,那么通信关系匿名也是成立的,但通信关系匿名不一定是发送者或接收者匿名。比如有些情况下,发送者和接收者本身是相互知道身份的,但它们之间的通信关系对其他成员是匿名的。不可关联性是匿名性的充分条件,但不是必要条件,即不可关联性⇒匿名性

根据匿名和不可关联性的关系,下面将给出一个完备的匿名推理系统,它能够推断出对象之间是否存在隐含的相关联性。

3 匿名推理系统

在网络活动中,用户常常在某个时间、某个位置,使用一个设备进行一个行动(如访问某个资源)。这里定义几个与用

*)中国博士后基金(20070410953)、湖北省自然科学基金(2005ABA243)资助。郭亚军 副教授,博士(博士后),主要研究方向:信息安全、普适计算;何炎祥 博士,教授。

户相关的对象:行动(A)、设备(D)、用户(U)、位置(L)、时间(T)(用户完成行动的时间段)和用户的公钥(P)等。

定义 1(函数决定) 设有对象集合 R,X 和 Y 是 R 的子集,并且 $X \neq Y$,如果从 X 就能知道 Y,那么称 X 函数决定 Y,记为 $X \rightarrow Y$。

上面定义说明了 X 和 Y 是相关联的。

定义 2(函数决定集) 对于对象集合 R,攻击者知道的对象之间函数决定的集合 F 称为函数决定集。

定义 3(函数决定蕴涵) 如果 F 是一个攻击者知道的函数决定集,设有对象集合 R,X 和 Y 是 R 的子集,并且 $X \neq Y$,如果从 F 中能够推出 $X \rightarrow Y$,则称 F 蕴涵 $X \rightarrow Y$。

定义 4(函数决定闭包) 所有被 F 蕴涵的函数决定集称为 F 的闭包,记为 F^+。

通过判断在 F 蕴涵的函数决定集里是否存在对象之间的函数决定关系,就能判断用户匿名能否得到保护。

下面给出最基本的函数决定推理规则,它们构成一个完备的匿名函数决定推理系统。

对于对象集合 R,X,Y 和 Z 是 R 的子集,F 是攻击者知道的函数决定集。

(1) 规则Ⅰ:若 $Y \subseteq X$,则 $X \rightarrow Y$

(2) 规则Ⅱ:若 $X \rightarrow Y$,则 $XZ \rightarrow YZ$

(3) 规则Ⅲ:若 $X \rightarrow Y$,$Y \rightarrow Z$,则 $X \rightarrow Z$

定理 1 匿名函数决定推理规则是正确的。

根据函数决定定义很容易证明定理 1 的正确性。

从匿名函数决定推理规则可以得出下列推论:

(1) 规则Ⅳ:若 $X \rightarrow Y$,$X \rightarrow Z$,则 $X \rightarrow YZ$

(2) 规则Ⅴ:若 $X \rightarrow YZ$,则 $X \rightarrow Y$,$X \rightarrow Z$

(3) 规则Ⅵ:若 $X \rightarrow Y$,$YW \rightarrow Z$,则 $XW \rightarrow Z$

定理 2 函数决定推理规则的推论是正确的。

定理 3 设有对象集合 R,$u_i \in R$,其中 $i = 1, \cdots, n$。则 $X \rightarrow u_1 u_2 \cdots u_n$ 成立的充要条件是 $X \rightarrow u_i$ $(i=1,2,\cdots,n)$ 均成立。

由规则Ⅳ和规则Ⅴ很容易知道定理 3 的正确性。

求函数决定的闭包是比较困难的,但可以通过计算对象的闭包来判断某个函数决定是否属于函数决定闭包。

定义 5(对象闭包) 如果 F 是一个攻击者知道的函数决定集,设有对象集合 R,X 是 R 的子集,$u_i \in R$,所有从函数决定推理系统中推出的函数决定 $X \rightarrow u_i$ 中 u_i 的对象集合为对象 X 的闭包,记为 X^+。

定理 4 设 F 是一个攻击者知道的函数决定集,对象集合为 R,X 和 Y 是 R 的子集,则 $X \rightarrow Y$ 是用函数决定推理系统从 F 推出的充分必要条件是 $Y \subseteq X^+$。

证明:充分性。设 $Y \subseteq X^+$,并设 $Y = u_1 u_2 \cdots u_n$,根据对象闭包定义可知 $X \rightarrow u_i$ $(i=1,2,\cdots,n)$ 是用函数决定推理系统从 F 推出的,再根据规则Ⅳ可得 $X \rightarrow Y$。

必要性。设 $X \rightarrow Y$ 是用函数决定推理系统从 F 推出的,$Y = u_1 u_2 \cdots u_n$,根据规则Ⅴ有 $X \rightarrow u_i$ $(i=1,2,\cdots,n)$,由对象闭包定义知道 $u_i \in X^+$,所以 $Y \subseteq X^+$。

下面将讨论匿名函数决定推理系统的完备性问题。完备性保证能够推出所有的函数决定。

定理 5 函数决定推理系统是完备的。

证明:

(1)若 $V \rightarrow W$ 成立,且 $V \subseteq X^+$,由定理 4,$X \rightarrow V$ 成立;根据规则Ⅲ,$X \rightarrow W$,所以 $W \subseteq X^+$。

(2)设所有对象集合为 U,构造一个攻击者观察对象集合 *view*,攻击者不能从 X^+ 的对象集合中知道 $U - X^+$ 中的对象。现在证明 F 中的所有函数决定在 *view* 上成立。

假如 F 中的函数决定 $V \rightarrow W$ 在 *view* 不成立,由 *view* 的构成可以知道 V 是 X^+ 的子集,W 是 $U - X^+$ 的集合,即不是 X^+ 的子集,这与(1)中的 $W \subseteq X^+$ 矛盾,因此 F 中的所有函数决定在 *view* 上成立。

(3)若 $X \rightarrow Y$ 不能从 F 使用函数决定推理系统导出,由定理 4 可知,Y 不是 X^+ 的子集。因此存在 Y 的子集 Y' 满足 $Y \not\subset U - X^+$,则 $X \rightarrow Y$ 在 *view* 中成立,即 $X \rightarrow Y$ 不为 F 蕴含。

算法 1 求对象集 X 关于函数决定集 F 的闭包 X^+

输入:X, F

输出:X^+

步骤:

(1) $X^{(0)} = X$

(2) $X^{(i+1)} = X^{(i)} Y$ $(i=0,1,\cdots)$

其中 Y 是在 F 中寻找还没有用过的左边是 $X^{(i)}$ 子集的函数决定的对象子集。如果没有这样的 Y,则转(4)。

(3)判断 $X^{(i+1)} = X(i)$ 是否相等。若是,则转(4),否则转(2)。

(4)输出 $X^{(i)}$,即为 X^+。

4 匿名分析算法

对象的匿名性是指它与任何身份标记不相关联,也可能是一个对象与很多用户身份标记相关联。如:对于一个位置 l,如果攻击者知道的函数决定闭包中,不存在位置 l 和任何一个用户相关联。或者存在位置 l 与多个用户相关联,攻击者不能判断该位置与哪个用户有关,则称位置 l 是匿名的。相关联的用户数越大,表示攻击者越难知道位置与哪个用户相关。如果用 t 表示一个门槛值,则相关联的用户数大于或等于 t 时,对象位置是匿名的。

同样,对于一个行动 a,如果攻击者知道的函数决定闭包中不存在行动 a 和任何一个用户相关联,或者存在行动 a 与多个用户相关联,攻击者不能判断该位置与哪个用户有关,则称行动 a 是匿名的。

根据匿名推理系统,下面给出匿名分析算法。

算法 2:匿名保护方案的匿名性分析

输入:攻击者能够观察到的对象和对象之间的函数决定集合

输出:匿名性能否得到保证

步骤:

(1) 根据攻击者观察到的对象之间的函数决定集合,计算对象的闭包。

(2) 如果该对象闭包中包含一个用户身份标记,转(7)。

(3) 如果该对象闭包中包含的用户身份标记数小于规定的门槛值 t 时,转(7)。

(4) 如果对象闭包中不包含任何用户身份标记,转(6)。

(5) 如果该对象闭包中包含的用户身份标记数大于或者等于规定的门槛值 t 时,转(6)。

(6) 输出:该对象是匿名的。

(7) 输出:该对象不是匿名的。

应用举例:

Crowds 的目的是为用户提供匿名 Web 浏览,它使得用

(下转第 73 页)

使用资源的节点说真话，并且贡献资源的节点不能确定最正确的资源价格。

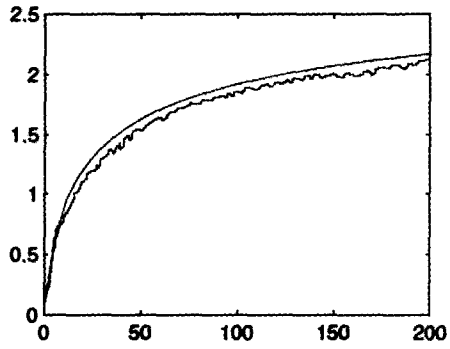

图 2

结论和将来的工作 本文提出了一种新的 P2P 激励机制，在这种激励机制中所有的资源都用虚拟货币进行标价。节点通过随机拍卖来确定那些节点是拍卖的获胜者，就可以使用资源并付出虚拟货币，而节点可以通过贡献自己的资源来得到虚拟货币。我们描述了拍卖的工作原理和资源贡献节点得到多少期望收益。通过分析实验结果，我们确定只有当 P2P 中的节点是诚实的，节点说的和行为一致时，激励机制才能发挥比较好的作用。所以在未来的工作中我们希望能够找到一种激励机制，鼓励节点说真话或者说真话要得到鼓励，并且说真话是节点收益最大化的重要保障。

参 考 文 献

1 Feldman M. FreeRiding and Whitewashing in PeertoPeer Systems. SIGCOMM'04 Workshop, Aug. 30+Sept. 3, 2004
2 Kenyon C, Cheliotis G, Buyya R. 10 Lessons from Finance for Commercial Sharing of IT Resources. In: Peer-to-Peer Computing: The Evolution of a Disruptive Technology, IRM Press, 2004
3 Hardin G. The Tragedy of the Commons. Science, 1968, 162: 1243～1248
4 Feigenbaum J, Shenker S. Distributed algorithmic mechanism design: Recent results and future directions. In: Proc. 6th Int'l Workshop on Discrete Algorithms and Methods for Mobile Computing and Communications, Atlanta, GA, Sept. 2002. 1～13
5 Golle P, Leyton-Brown K, Mironov I, et al. Incentives for sharing in peer-to-peer networks. In: Proc. 3rd ACM Conf. on Electronic Commerce, Tampa, FL, Oct. 2001
6 Milojicic D S, Kalogeraki V, Lukose R, et al. Peer-to-Peer Computing: [Technical Report HPL-2002-57]. HP Laboratories Palo Alto, 2002
7 Ratnasamy S, Francis P, Handley M, et al. A scalable content addressable network. In: Proc. of ACM SIGCOMM, 2001
8 Rowstron A, Druschel P. Pastry: Scalable, decentralized object location, and routing for large-scale peer-to-peer systems. Lecture Notes in Computer Science, 2001, 2218: 329～350
9 Stoica I, Morris R, Karger D, et al. Chord: A scalable peer-to-peer lookup service for internet applications. ACM SIGCOMM, 2001
10 Kenyon C, Cheliotis G, Buyya R. 10 Lessons from Finance for Commercial Sharing of IT Resources. In: Peer-to-Peer Computing: The Evolution of a Disruptive Technology. IRM Press, 2004
11 Adar E, Huberman B. Free Riding on Gnutella. First Monday, 2000, 5(10)
12 Yu Bin, Singh M P. Incentive Mechanisms for Peer-to-Peer Systems. http://www.csc.ncsu.edu/faculty/mpsingh/papers/mas/ap2pc-03-mechanism.pdf
13 Ma R T B. A Game Theoretic Approach to Provide Incentive and Service Differentiation in P2P Networks. SIGMETRICS/Performance'04, 2004
14 Bapna R, Goes P, Gupta A. Insights and analysis of online auctions. Communications of the ACM, 2001, 44: 42～50
15 Bhargava H K, Sundaresan S. Contigent Bids in Auctions: Availability, Commitment and Pricing of Computing as Utitily. In: Proc. of the 37th Hawaii International Conference on System Sciences, 2004
16 Lange J, Economides N. A parimutuel market microstructure for contingent claims. European Financial Management, 2005, 11(1): 25～49
17 Bapna A, Weber T. Efficient Dynamic Allocation with Uncertain Valuations. Stanford University
18 Bergemann D, V¨alim¨aki J. Dynamic Common Agency. Journal of Economic Theory, Forthcoming
19 Bergemann D, V¨alim¨aki J. Dynamic Price Competition. Journal of Economic Theory, 127: 232～263

(上接第 57 页)

户能够匿名地从 Web 服务器取回信息而不对服务器和第三方泄漏用户自己的信息，其基本思想是把自己隐藏在群体中。

假设攻击者知道在 crowds 中有 n 台计算机($d_1 d_2 \cdots d_n$)参加访问了某个网页(该行动用 a_1 表示)，也知道某台计算机 d_k 访问了另一个网页(该行动用 a_2 表示)。攻击者知道用户使用哪台计算机，则攻击者的函数决定集合为 $F=\{(d_1 d_2 \cdots d_n)\rightarrow a_1, d_k\rightarrow a_2, u_1\rightarrow d_1, u_2\rightarrow d_2, \cdots, u_n\rightarrow d_n, u_k\rightarrow d_k\}$。

下面分析用户行动是否具有匿名性。

行动 a_i 的闭包为 a_i，用户 $u_i(i=1,2,\cdots,n)$ 的闭包为 $u_i d_i$，u_k 闭包为 $u_k d_k a_2$。

显然用户 u_i 的行动是匿名的，即不能从 u_i 推出某个行动。由于从 u_k 能够推出 a_2，因此用户 u_k 的行动不是匿名的。

小结 对于一个匿名系统而言，对象之间是否具有关联性不一定是明显的。如果攻击者根据获得的知识无法推断对象之间的关联性，则认为对象是匿名的。本文给出的完备的匿名推理系统能够推断对象之间的关系，从而能够发现隐藏的安全性问题。匿名分析算法给出了评估匿名机制的安全性方法，并对设计匿名机制具有很好的指导作用，与现有的隐私钻石模型相比具有分析能力强以及简单实用等优点。

参 考 文 献

1 Chaum D L. Untraceable electronic mail, return addresses, and digital pseudonyms. Communications of the ACM, 1981, 24(2): 84～88
2 Reiter M, Rubin A D. Crowds: anonymity for Web transactions. ACM Transactions on Information and System Security, 1998, 1(1): 66～92
3 Ren K, Lou W, Kim K, et al. A novel privacy preserving authentication and access control scheme for pervasive computing environments. IEEE Transactions on Vehicular Technology, 2006, 55(4): 1373～1384
4 Konidala D M, Duc D N, Kim K. A capability-based privacy-preserving scheme for pervasive computing environments. In: Proc. IEEE International Workshop on Pervasive Computing and Communication Security, 2005. 8～12
5 Wu X. Applying pseudonimity for anonymous data delivery in location-aware mobile ad hoc networks. IEEE Transactions on Vehicular Technology, 2006, 55(3): 1062～1073
6 Kobsa A, Schreck J. Privacy through pseudonymity in user-adaptive systems. ACM Transactions on Internet Technology, 2003, 3(2): 149～183
7 Zugenmaier A, Kreutzer M, Muller G. The freiburg privacy diamond: an attacker model for a mobile computing environment. In: Kommunikation in Verteilten Systemen (KiVS) '03, 2003
8 Iachello G, Truong K N, Abowd G D, et al. Prototyping and sampling experience to evaluate ubiquitous computing privacy in the real world. In: Proc. ACM Conference on Human Factors in Computing Systems, New York: ACM Press, 2006. 1009～1018