传感器网络远程网络重编程服务安全认证机制研究*)

张 羽 周兴社 蒋泽军 王丽芳
(西北工业大学计算机学院 西安 710072)

摘 要 传感器网络远程网络重编程服务不仅需要有效和可靠的分发机制,而且还需要高效的安全认证机制。提出了传感器网络远程网络重编程服务认证安全需求和性能评价标准,介绍了典型的网络重编程服务安全认证方案,分析了当前已有方案并进行了比较,最后指出了传感器网络远程网络重编程服务安全认证机制存在的问题及研究方向。

关键词 传感器网络,网络重编程,代码分发,安全,认证机制

Security Authentication Mechanisms Study of Remote Network Reprogramming Service for Sensor Networks

ZHANG Yu ZHOU Xing-she JIANG Ze-jun WANG Li-fang
(School of Computer Science, Northwestern Polytechnical University, Xi'an 710072, China)

Abstract Network reprogramming services need to be not only reliable and efficient, but secure as well. The secure and performance evaluation criterion of network reprogramming services for sensor networks was proposed, and the classic secure network reprogramming schemes were introduced. The current research progresses were analyzed and compared. The open research problems in this area were also pointed out.

Keywords Sensor networks, Network reprogramming, Code dissemination, Security, Authentication mechanisms

1 引言

网络重编程(Network Reprogramming),也称作无线重编程(Wireless Reprogramming)或代码更新(code update)是解决大规模无线传感器网络(Wireless Sensor Networks,简称WSNs)难于管理和维护的有效途径[1,2]。它已作为WSNs这类新型计算形式能否实用化的一个必要和关键服务[3-10]。网络重编程提供在多跳网络环境下向全网节点有效和可靠分发大数据对象(如程序代码)的功能。它可以对远程的原地(in situ)传感器节点实施任务再分配,所以避免了使用传统手工方法去收集网络节点再编程的低效方式。

初始的网络重编程研究主要集中于解决高度动态、有损的无线通信环境下,更新代码可靠和有效的广播分发问题,而缺乏对安全性问题的考虑。由于WSNs的固有特性使其比传统网络更加脆弱[11-13];加之,为了让更新代码最终抵达网络中的所有传感器节点,现有网络重编程服务大多具有"流行病"特征[14],这使得攻击者只要俘获网络中任意一个节点,就能将一个局部的危害迅速"传染"到整个网络造成严重的全局后果。因此,网络重编程服务不仅需要可靠和有效的代码分发机制,而且还需要高效的安全认证机制来确保更新代码的真实性(authenticity)和完整性(integrity),防止篡改、冒充等主动攻击。

本文全面介绍了WSNs网络重编程服务安全认证机制方面的研究进展。第2节阐述了WSNs网络重编程服务的安全需求和性能评价标准;第3节综述了典型WSNs无线重编程安全服务认证方案;第4节分析了已有的认证方案,并做出了相应的比较;最后指出了需要解决的问题以及未来的研究方向。

2 网络重编程服务安全需求和性能评价

为了确保网络重编程过程的安全性,使可信更新代码有效分发的同时阻止恶意更新代码传播和安装,WSNs网络重编程服务认证机制必须满足以下安全需求:

2.1 安全需求

(1)真实性。传感器节点必须能验证更新代码来自于可信的分发源(如,基站)。

(2)完整性。传感器节点必须能验证接收的更新代码和从可信源发出时一样,没有在传输过程中被其他节点篡改。

(3)抗俘获性。即使网络中有一个节点或多个传感器节点被俘获,也不会导致网络其他部分的不安全。

(4)正确性。除非传感器节点已被俘获,否则任何未经验证的更新代码都不能被转发或安装。

(5)完成性(Completeness)。每个行为良好的传感器节点应该最终能安装一个被验证通过的更新代码。

(6)可用性。具有抗DoS攻击能力,它使传感器节点能及时验证更新代码。

(7)新鲜性。更新代码本身具有时效性,传感器节点必须能够判断接收到的更新代码是最新的版本或是全新的程序映像。

此外,安全认证机制还应根据网络重编程和传感器节点本身的固有特性[15,12],满足以下性能评价标准。

2.2 性能评价

(1)低安全开销。网络重编程安全认证机制应该是资源敏感、轻量和有效的。因此在设计时最小化其计算开销、通信

*)国家自然科学基金(编号:60573161),陕西省自然科学基金(编号:2006F08)资助项目。张 羽 博士研究生,主要研究方向为传感器网络和分布计算;周兴社 教授,博士生导师,主要研究方向为普适计算和嵌入式网络计算;蒋泽军 教授,主要研究方向为嵌入式系统和网络安全;王丽芳 教授,主要研究方向为网络安全和电子商务。

开销和存储开销。

(2)低端到端时延。为了网络重编程过程快速完成,使其减少对 WSNs 主应用程序的影响,应该尽可能缩短验证更新代码所需的时间。

(3)可兼容性。网络重编程安全认证机制应该对下层的网络重编程分发机制提出最少的假设,使其尽可能与现有的协议兼容,并能充分利用已有的有效性机制(如 Pipelining,抑制机制)。

(4)可扩展性。网络重编程安全认证机制应该能够适应不同规模的 WSNs。

3 典型网络重编程服务安全认证机制

从近年来的研究进展看,根据对散列链的链首进行验证所使用的密码体制,WSNs 网络重编程的安全认证方式目前可以分成两类。一类是基于公钥密码的数字签名认证方式;一类是基于对称密钥密码的安全认证方式。

3.1 基于公钥密码的数字签名认证方式

该方式的主要思想是使用单向散列函数和数字签名混合方法来认证网络中的更新代码。即,一个可信的基站有一个私钥,同时每个传感器节点预置这个基站相应的公钥。基站用它的私钥对每个更新包进行签名,传感器节点用公钥来验证每个接收更新包的真实性和完整性,因为任何节点在没有获得私钥的情况下,都无法冒充更新代码的合法签名。

3.1.1 Lanigan 等人的 Sluice 安全认证方案[16]

Lanigan,Grandhi 和 Narasimhan 在 WSNs 中最先提出了网络重编程安全认证方案。该方案的基本思想是基于下层的网络重编程服务,将更新代码分成固定大小的页(page),计算每页 p_i 的散列值 h_i,并将其附加到它前一页 p_{i-1} 的负载中($p_{i-1}|h_i$),以此方式从最后一页 P_{n-1} 向前创建一条单向散列链,最后对该链第一个散列值 h_1 及首页 p_0 的其他内容(如程序版本号等)做一次数字签名。该数字签名 σ 用来认证更新代码的合法性(因为只有基站才拥有唯一的私钥),同时单向散列链可以保证一旦链首的 h_1 通过验证后,其余的页也可依次得到验证,进而确保更新代码的完整性,如图 1 所示。

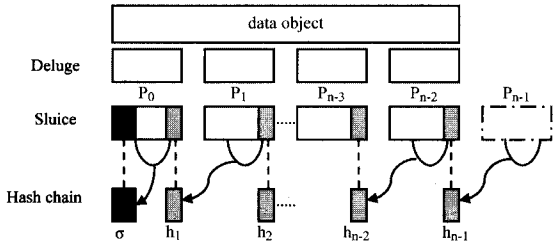

图 1 Sluice 方案的散列链构建

Sluice 方案实现了网络中即便存在传感器节点被俘获的情况,也能阻止被俘节点向网络传播恶意代码,同时不影响可信代码的快速分发。由于该方案使整个更新代码只做一次数字签名,因此大大降低了公钥运算量和相应的计算开销,在一定程度上,满足了 WSNs 的特点。Sluice 方案的主要缺点是会出现当一个散列值认证失败后,节点将被迫重新请求传输大量的数据(48 个包)。原因是该方案采用粗粒度的"页级"散列方法使一个节点无法确定页中具体哪个包被毁,因而只能被迫请求重传整个页。这个问题将会给攻击者提供实施 DoS 攻击的机会。

3.1.2 Dutta 等人的 SecureDeluge 方案[17]

Dutta,Hui,Chu 和 Culler 提出了与 Sluice 类似的网络重编程安全认证方案,即,也采用了基于单向散列链的一次数字签名方法。而不同于 Sluice 方案之处在于 SecureDeluge 选用了细粒度的"包级"散列。其优点在于一个坏包不会触发整个页的重传。然而,SecureDeluge 方案的缺点是下层更新代码包尺寸被迫增大,散列值也要被迫截短,因此降低了散列函数的安全强度。该方案与下层网络重编程协议 Deluge 的兼容性并不理想,没有提供容忍包乱序到达的机制。然而,由于 MAC 层冲突或传输路径上传感器节点缓冲区溢出等问题会使 WSNs 很容易出现丢包和包乱序到达现象,加之,Secure-Deluge 包尺寸的增加意味着会出现更多包丢失的情况。此外,Deluge 中的 NACK 可靠重传机制使发送节点要等到收到从接收节点发来的反馈确认消息后,才决定它是否需要重传丢失的包,这样就阻碍了接收节点对更新代码包进行及时验证,从而带来更大的时间延迟,这也使其不具有抗 DoS 攻击能力。

3.1.3 Deng 等人的 Deng-tree 方案[18]

Deng,Han 和 Mishra 为解决包的乱序到达问题采用了一个基于 Merkle 树的网络重编程安全认证方案。该方案的基本思想是提前传播更新代码包的散列值,从而加快更新代码的验证过程。在 Deng-tree 方案中,提前发送的包被称为索引包(index packets),它包含更新代码包的散列值。基站将更新代码分成多个包,并对每个包计算散列值,用这些散列值作为输入创建新一级的散列值,这样一直向上到树顶。即,如果散列树有 $m+1$ 层(从 0 到 m),在第 i 层的包含第 $i+1$ 层 ω 个包的散列值。散列树以此方式构建,直到第 1 级刚好有一个包。树顶第 0 级的根值是用基站私钥对第 1 级散列值加密后产生的一个签名。该签名用于第 1 级包的真实性和完整性检验。由于散列树将该层与下一层的所有包相关联,因此这个数字签名能依次验证散列树中所有包的真实性和完整性,如图 2 所示。

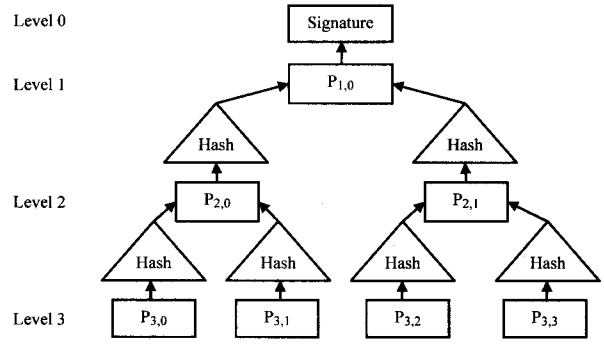

图 2 Deng-tree 方案的散列树构建

Deng-tree 方案能使每个传感器节点及时验证更多数量的乱序包。因为一旦所有的索引包被接收并验证通过,更新代码包就可以以任意顺序到达,并且能够被及时验证。该方案的主要缺点是为给更新代码这样的大数据对象构建一棵散列树需要在代码分发的过程中保留大量的内存空间来存储索引包。即便以每页为单位构建多个规模相对较小的散列树也会因每棵散列树都要做一次数字签名这样的公钥运算,而带来巨大的计算开销。

3.1.4 Deng 等人的 Deng-hybrid 方案[18]

Deng,Han 和 Mishra 为解决多个散列树的多次签名问题又提出了一种结合散列树和散列链的混合方案。该方案利

用下层的 Deluge 将一个更新代码分成若干个页，并以页为单位，构建一棵散列树，然后再用每页的根级包构建一条散列链。混合方案试图结合散列树和散列链方案中各自的优势，使其具有散列树方案能容忍包乱序到达的能力，同时具有散列链方案中整个更新代码只需一次数字签名的优点。混合签名散列树方案的缺点是由于每页要传输一颗散列树而引入了可观的通信开销，加之传感器节点内存的限制，额外的索引包需要存储在 EEPROM 中，进而产生大量的能量消耗[19]。此外，散列树虽不需要完全顺序发送，但部分顺序是必需的，因此它没有最终解决容忍包乱序到达的问题，如图 3 所示。

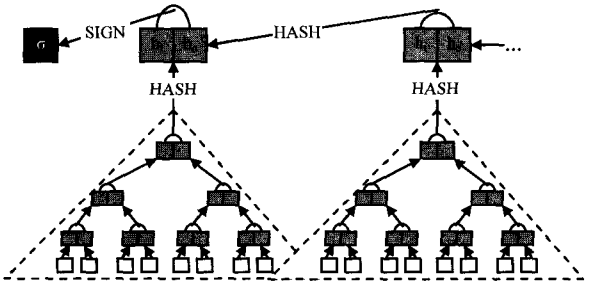

图 3　Deng-hybrid 方案构建

3.2　基于对称密钥密码体制的认证方案

尽管前面介绍的几种基于非对称密码体制的安全认证方案对整个更新代码仅使用一次数字签名来最小化公钥运算本身巨大的计算、存储开销，但对于像传感器节点这样资源极度受限的计算环境，若采用基于对称密钥加密方法来取代数字签名可以进一步减少更新代码的认证开销，缩短端到端传输的时间延迟。然而，使用对称密钥加密方法，需要在发送者和接受者之间事先建立一个共享密钥。在网络重编程过程中，只要有一个节点被俘获就意味着共享密钥被泄露，导致整个网络都不再安全。此外，即便使用基站与传感器节点之间的配对密码(pairwise keys)方案，也会因 WSNs 规模增大所产生的显著开销，而变得难以实用。因此如何在节点被俘的情况下，仍能利用对称密钥加密方法进行更新代码的安全认证是研究人员们所要面对的一个巨大挑战。

3.2.1　Kim 等人的 Castor 安全认证方案[20]

Kim，Gandhi 和 Narasimhan 通过采用基于消息认证码(message authentication code，简称 MAC)的对称密钥加密广播认证方法[21]来避免对散列链的链首进行数字签名。该方案将单向散列链、MAC 序列(MACs)、单向密钥链和延迟密钥公布相互结合实现更新代码的认证。在基站处，Castor 以更新代码的页为粒度构建一条单向散列链和一条单向密钥链，用单向密钥链中的每个密钥 k_i 计算首页 p_0 散列值 h_0 的 MAC，从而生成一个 MACs。基站先分发全部的 MACs，而后再分发更新代码的各页给所有传感器节点。按照预先设定的密钥公布时延 T，基站逐个发布密钥保证节点验证接收的 MACs 和之后的更新代码。如图 4 所示，由于采用计算开销更低的对称密钥加密方法，Castor 降低了端到端的传输时延。同时，该方案借鉴了安全的组播源认证思想[22]解决广播认证中，当密钥公布时延过小时，会出现的“节点掉队”问题(node-left-behind)。然而该方案的缺点是密钥延迟公布的正确实施要求网络必须具有安全的时间同步能力。目前大多数的网络重编程服务不对更新代码抵达节点的传输时间设置边界，只是保证更新代码最终能抵达所有的传感器节点。另外，MAC 序列方法在连续多次更新代码需要突发验证的情况下(由于传感器节点要多次使用单向密钥函数验证大间隔密钥的合法性)，会变得异常低效。并且尽管 Castor 保证一个未被俘获的节点不会转发未经验证的更新代码，但它却允许转发 MAC 序列，而一旦 MAC 被攻击者伪造，将使网络重编程过程不安全。

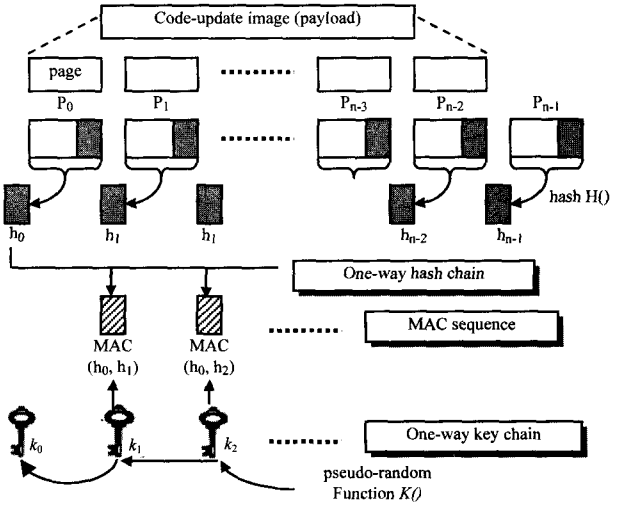

图 4　Castor 方案构建

3.2.2　Krontiris 等人的 r-times 签名认证方案[23]

Krontiris 和 Dimitriou 通过借鉴一次签名[24]和 Merkle 散列树[25]思想，用对称密钥加密方法实现对散列链链首(Root block)的数字签名。该方案的基本思想是将私钥分布到多个 Merkle 散列树中，获得公钥大小与签名长度的平衡。r-times 签名方案将公钥大小和签名长度最小化，使其适用于 WSNs。由于在传感器节点处的验证过程只进行散列和比较运算，因此减少了计算开销和验证时间，如图 5 所示。方案的缺点是采用了“页级”的散列粒度，致使其不具有抗 DoS 攻击的能力。

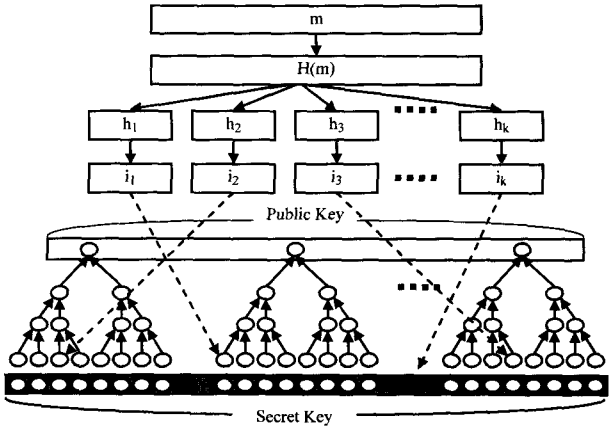

图 5　使用多 Merkle 树方案构建消息签名

4　网络重编程安全认证机制分析比较

从研究现状看，现提出的基于非对称密钥的更新代码安全认证方案的主要区别在于各自的混合结构(如，散列链 VS 散列树)、散列粒度(如，单个包 VS 一组包或页)和散列强度(如，全散列值 VS 截短散列值)的不同[26]。如表 1 所示，我们从各方案的安全性和性能方面给出了相应的比较。在表

2，表 3 中，“↓”表示小或弱，“—”表示中或一般，而“↑”表示大或强。

从前面的介绍可以看出基于非对称密钥的更新代码认证方案虽取得了一定的研究进展，但仍存在许多需要解决的问题。具体如下。

表 1　四种非对称密钥更新代码安全认证方案的主要区别

Schemes and protocols	Structure	Granularity	Strength
Sluice(CMU)[16]	Hash Chain	Page	Full
SecureDeluge(UCB)[17]	Hash Chain	Packet	Truncated
Deng-tree(Colorado)[18]	Hash Tree	Packet	Truncated
Deng-hybrid(Colorado)[18]	Hash Tree + Hash Chain	Packet	Truncated

表 2　不同结构和粒度下的安全认证方案安全性比较

Structure& Granularity	Security of Hash	DoS-attack Resilience
Hash Chain + Page	↑	↓
Hash Chain + Packet	↓	—
Hash Tree + Packet	↓	↑
Hash Tree + Hash Chain + Packet	↓	↑

表 3　不同结构和粒度下的安全认证方案性能比较

Structure& Granularity	Computation overhead	Communication overhead	Store overhead
Hash Chain + Page	↓	↑	↓
Hash Chain + Packet	—	—	—
Hash Tree + Packet	↑	↑	↑
Hash Tree + Hash Chain + Packet	↑	↑	↑

(1)抗 DoS 攻击问题。目前已有的网络重编程认证方案都没有提供良好的抵御 DoS 攻击的能力。而 DoS 攻击却是 WSNs 中一个难以避免的威胁类型[27]。

(2)网络重编程的攻击与对抗模型有待进一步完善。目前已有的网络重编程认证方案都集中在解决更新代码数据包的认证问题上，而针对网络重编程协议 Deluge 传播和抑制机制本身存在的问题仍没有得到解决(如 Signature，ADV，SNACK 等特定包的攻击问题)。

(3)多基站网络重编程安全认证问题。目前已有的网络重编程认证方案，不管是基于公钥密码还是基于对称密钥密码方案，都是针对单基站 WSNs 设计的。而单基站 WSNs 有时难于满足实际的应用需求。如何将现有的随机密钥预分布方案[13]与网络重编程过程相结合也是一个需要解决的问题。

(4)基于混合结构的数字流认证问题。从目前已有的网络重编程认证方案可以看出，单纯基于散列链和单纯基于散列树的验证结构都难以获得安全与性能上的真正平衡。

结束语　WSNs 远程网络重编程服务已作为 WSNs 的一个必要和关键服务，受到国外众多大学和学术机构研究人员的广泛关注和研究。而网络重编程过程自身特有的要求，使其给合法用户提供方便性和有效性的同时却也给 WSNs 引来了新的安全威胁，并且一旦使攻击者得手将会给整个网络带来严重的后果。在本文中，我们介绍了 WSNs 网络重编程服务安全认证机制方面的研究进展，重点介绍了基于公钥密码体制的安全认证机制，给出了 WSNs 网络重编程的安全需求和性能评价标准，分析了已有的认证方案，并做出了相应的比较，最后指出了需要解决的问题以及未来的研究方向。

参 考 文 献

[1] Han S, Rengaswamy R, Shea R S, et al. Sensor Network Software Update Management: A Survey. Internation Journal of Network Management, July 2005

[2] Wang Q, Zhu Y Y, Cheng L. Reprogramming Wireless Sensor Networks: Challenges and Approaches. In IEEE Network, May 2006

[3] Stathopoulos T, Heidemann J, Estrin D. A Remote Code Update Mechanism for Wireless Sensor Networks. CENS Tech. Report #30. Centre for Embedded Networked Sensing, UCLA, 2003

[4] Hui J, Culler D. The Dynamic Behavior of a Data Dissemination Protocol for Network Reprogramming at Scale // Proc. of 2nd ACM SenSys'04. Baltimore, Maryland, USA, 2004: 81-94

[5] Kulkarni S S, Wang L. MNP: Multihop Network Reprogramming Service for Sensor Networks // Proc. of IEEE ICDCS 05. 2005: 7-16

[6] Kulkarni S S, Arumugam M. Infuse: A TDMA Based Data Dissemination Protocol for Sensor Networks. Technical Report MSU-CSE-04-46. Dept. of Computer Science and Engineering, Michigan State University, MI, 2004

[7] Naik V, et al. Sprinkler: A Reliable and Energy Efficient Data Dissemination Service for Wireless Embedded Devices // 26th IEEE Real-Time Sys. Symp. Dec. 2005

[8] Wang L, Kulkarni S S. Gappa: Gossi PBased Multi-Channel Reprogramming for Sensor Networks // International Conference on Distributed Computing in Sensor Systems (DCOSS). San Francisco, CA, June 2006

[9] Xiao Z, Sarikaya B. Code Dissemination in Sensor Networks with MDeluge. Sensor and Ad Hoc Communications and Networks (SECON '06), 2006

[10] Starobinski D, Xiao W Y, Qin X P, et al. Near-Optimal Data Dissemination Policies for Multi-Channel, Single Radio Wireless Sensor Networks // IEEE INFOCOM 2007. Anchorage, AK, May 2007

[11] Perrig A, Stankovic J, Wagner D. Security in wireless sensor networks. Communications of the ACM (Special Issue on Wireless Sensor Networks), 2004, 47(6): 53-57

[12] 孙利民，李建中，陈渝，等. 无线传感器网络. 北京：清华大学出版社，2005

[13] 苏忠，林闯，封富君，等. 无线传感器网络密钥管理的方案和协议. 软件学报，2007，18(5)：1218-1231

[14] Demers A, Greene D, Hauser C, et al. Epidemic algorithms for replicated database maintenance // Proceedings of the Sixth Annual ACM Symposium on Principles of Distributed Computing. ACM Press, 1987: 1-12

[15] Lanigan P E, Gandhi R, Narasimhan P. Disseminating Code Updates in Sensor Networks: Survey of Protocols and Security Issues. CMU-ISRI-05-122. School of Computer Science, Carnegie Mellon University, PA, 2005

[16] Lanigan P E, Gandhi R, Narasimhan P. Sluice: Secure dissemination of code updates in sensor networks // IEEE International Conference on Distributed Computing Systems. Lisbon, Portugal, July 2006

[17] Dutta P K, Hui J W, Chu D C, et al. Securing the Deluge network programming system // ACM/IEEE Conference on Information Processing in Sensor Networks. Nashville, TN, April 2006

[18] Deng J, Han R, Mishra S. Secure code distribution in dynamically programmable wireless sensor networks// ACM/IEEE Conference on Information Processing in Sensor Networks. Nashville, TN, April 2006: 292-300

[19] Barr K, Asanovic K. Energy aware lossless data compression// The First International Conference on Mobile Systems, Applications, and Services. San Francisco, CA, May 2003

[20] Kim D H, Gandhi R, Narasimhan P. Exploring symmetric cryptography for secure network reprogramming // International Workshop on Wireless Ad-hoc and Sensor Networks. New York, NY, June 2007

[21] Perrig A, Szewczyk R, Tygar J D, et al. SPINS: Security protocols for sensor networks. Wireless Networks, 2002, 8(5): 521-534

[22] Perrig A, Canetti R, Song D, et al. Efficient and secure source authentication for multicast// Proceedings of Network and Distributed System Security Symposium. San Diego, CA, February 2001: 35-46

[23] Krontiris I, Dimitriou T. Authenticated in-network programming for wireless sensor networks// International Conference on Ad-Hoc Networks and Wireless. Ottawa, Canada, August 2006

[24] Lamport L. Constructing digital signatures from a one-way function. Technical Report CSL-98. SRI International Computer Science Laboratory, Palo Alto, 1979

[25] Merkle R C. A certified digital signature// Proceedings on Advances in cryptology (CRYPTO '89). New York: Springer-Verlag, Inc., 1989: 218-238

[26] Lanigan P E, Narasimhan P, Gandhi R. Tradeoffs in Configuring Secure Data Dissemination in Sensor Network: An Empirical Outlook. CMU-CyLab-07-006. CyLab, Carnegie Mellon University, PA, 2007

[27] Wood A D, Stankovic J A. Denial of Service in Sensor Networks. IEEE Computer, Oct. 2002: 48-56

(上接第 14 页)

演算动态表达一个 MDP，并且通过一个称为决策理论回归(decision-theoretic regression)运算，精确产生最优值函数和最优策略的逻辑描述，它是 Reiter's 情景演算(situation calculus)的概率延伸，即是一阶决策理论回归。虽然情景演算语言富有表达力，然而它比较复杂。2003 年，Kristian Kersting 和 Luc De Raedt 提出逻辑马尔可夫决策过程，2004 年 Martijn Van Otterlo 提出关系马尔可夫决策过程，这些都是基于逻辑编程的系统。在某种意义上说，LOMDPs 和 RMDPs 和文献[1]中的思想是一致的，但比较简单些。也是基于文献[1]中的思想，Kersting, Van Otterlo & De Raedt，于 2004 年成功地把 Bellman 方程推广到关系领域，从而在逻辑层次上精确求解抽象值函数和抽象最优策略。

在近似求解和模型自由(model free)方面，Kerstian Kersting 和 Luc De Raedt 在 LOMDPs 中、Martijn Van Otterlo 在 RMDPs 中都提出了算法。但是这些算法都是基于传统的近似值函数算法，2006 年，Alan Fern, Sungwook Yoon & Robert Givan[7] 中运用类似于"可能学习近似正确假设"的机器学习方法，直接在偏置的策略空间中学习近似最优策略。文献[7]是值得阅读的，因为它提出的算法与传统不同，是直接在偏置(用一种表达策略的语言而得到)的策略空间里搜寻最好策略，从而为抽象近似策略迭代提供一个范例。当然，结构型里有关情景演算和因子化 MDPs 以及概括型里有关 LOMDPs 和 RMDPs 的文献及其最优策略的算法是较多的，但为了使我们的介绍既简单而又不流于空泛，在我们阅读许多文献后，根据自己的研究，把自己认为最主要的概念及代表不同风格和趋向的算法精选出来，加以较详细的介绍。至于其他作者的工作大都在我们所引的文献中有所涉及，故略而不述，有兴趣的读者可在我们所列的文献中找到有关他们的信息。

最后，对于 LOMDPs 和 RMDPs 今后的发展，我们简单提出几点看法。

①所有文献都指出抽象层次上的最优策略是有价值的，但这个论点至今尚未看到清晰的、完整的证明。我们认为，既然 LOMDPs 和 RMDPs 都基于这样一个事实，即逻辑层次上的表述能把基础层次上的状态和动作按某种类似性分类，那么我们就可以运用现代概率论中的条件数学期望概念来论证在抽象层次上的最优策略和实际层次上的具体最优策略之间的关系。我们认为，可以证明抽象层次上最优策略在平均意义上是基础层次上的最优策略。

②所有文献都把抽象层次和基础层次交织在一起，这为寻求抽象最优策略带来了很大麻烦。我们认为，也可以把这两个层次"完全"分离，转变为两个层次上的马尔可夫决策过程，这样在抽象层次上，利用它比实际状态数目较少的优点，较简单地运用传统方法求抽象最优策略，再按某种演算具体在实际状态空间实施，可能起到事半功倍之效。这个思路更符合人类思维特点，因为人们总是首先在大体上思考问题解决方案，然后再付诸实践，具体实现之。

参 考 文 献

[1] Boutilier C, Reiter R, Price B. Symbolic Dynamic Programming for First-order MDPs// Seventeenth International Joint Conference on Artificial Intelligence (IJCAI-01). Seattle: USA, 2001: 690-700

[2] Guestrin C, Koller D, Parr R, et al. Efficient Solution Algorithms for Factored MDPs. Journal of Artificial Intelligence Research, 2003, 19: 399-468

[3] Kersting K, Raedt L D. Logical Markov Decision Programs// Working Notes of the IJCAI-2003 Workshop on Learning Statistical Models from Relational Data (SRL-03). Acapulco, Mexico, 2003: 63-70

[4] Kersting K, Raedt L D. Logical Markov Decision Programs and the Convergence of Logical TD(λ)// Proceeding of The 14th International Conference of Inductive Logic Programming. Porto, Portugal, 2204: 180-197

[5] van Otterlo M. Reinforcement Learning for Relational MDPs// Proceedings of the Annual Machine Learning Conference of Belgium and the Netherlands. Brussels, Belgium, 2004: 138-145

[6] Kersting K, van Otterlo M, Raedt L D. Bellman goes to Relational// Proceedings of the 21st International Conference on Machine Learning. Banff, Canada, 2004

[7] Fern A, Yoon S, Givan R. Approxiamte Policy Iteration with a Policy Language Bias: Solving Relational Markov Decision Processes. Journal of Artificial Intelligence Research, 2006, 25: 75-118