基于知识积累的告警相关方法*)

龚发根 秦 拯 张大方
(湖南大学软件学院 长沙410082)

摘 要 黑客的入侵是一个逐步积累、逐步深入的过程。在入侵过程中,黑客所积累的有关目标系统的信息越多,越有利下一步入侵的成功。现有的告警相关方法不能识别带分支的入侵过程,也不能识别属于某个入侵过程的失败的入侵步骤。该文针对这两种情况提出了一种基于知识积累的告警相关方法,这种方法不仅能识别完整入侵过程,而且能对入侵过程的相关程度及其入侵结果进行评估。

关键词 入侵检测,告警相关

An Alert Correlation Method Based on Knowledge Accumulation

GONG Fa-Gen QIN Zheng ZHANG Da-Fang
(Software College,Hunan University,Changsha 410082)

Abstract Hacker's intrusion is a process to accumulate information from the target system step by step. The more information hacker collect from the target system,the more probability to succeed of the next intrusion step. The existing alert correlation method can't identify the intrusion process which have several embranchment,and also can't recognize the unsuccessful intrusion step belonging to certain intrusion process. Considering this two condition,this paper propo ses an alert correlation method based on knowledge accumulation. This kind of method not only can identify more intact intrusion process,but also can evaluate the correlation degree of the intrusion process and result of the intrusion process.

Keywords Intrusion detection,Alert correlation

1 引言

目前的入侵检测系统主要注重对单个的入侵事件或异常状态进行检测[1],当检测到入侵事件或异常状态时,就产生相应的告警信息。系统安全管理员只能通过分析告警信息来识别黑客的入侵过程及其入侵策略。这种通过人工的方式来识别入侵过程的方法存在很多问题:第一,对系统安全管理员的要求太高;第二,在告警信息量很大或者存在误告警的情况下,这种方法是行不通的[2,3]。

通过对黑客入侵的分析,可以发现:黑客要达到其入侵的目的,往往需要经过多个入侵步骤,不同的步骤所采用的入侵方法及其想达到的目的也不一样,前一阶段的入侵为后一阶段的入侵做准备[4]。例如,分布式拒绝服务攻击(DDOS),可以利用 Solaris 主机上的 Sadmind 漏洞先进行缓冲区溢出攻击,以取得该主机上的根用户权限,然后再发起 DDOS 攻击。它的一般入侵步骤如下:

(1) 利用 IPsweep 扫描目标网络,以探测活动主机信息。

(2) 利用 SadmindPing 来探测活动 Solaris 主机是否运行了有漏洞的 sadmind 后台进程。

(3) 利用含有 sadmind 漏洞的主机发起缓冲区溢出攻击(SadmindBOF),以取得该主机的根权限。这个攻击步骤需进行多次,以取得多台主机的根权限。

(4) 在这些取得了根权限的主机上分别安装 DDOS 攻击程序,并在其中一台主机上安装控制程序。

(5) 黑客从控制机上控制 DDOS 攻击程序进行攻击。

不同的入侵方法会引起不同的告警信息,这些告警信息同属某个入侵过程,它们之间存在一定的逻辑关系,称为相关关系。通过识别告警信息之间的相关关系,就可以识别出入侵者的入侵逻辑步骤及其入侵策略[5]。而且通过孤立误告警信息还可识别误告警,从而降低入侵检测系统的误警率[6]。

告警相关是最近两年才提出来的新的研究内容,国内尚未见这方面公开报道。在文[7]中,Peng Ning 等提出了一种基于因果关系的告警相关方法,他们为每种入侵定义了前置条件和后置状态,前置条件是成功实施这个入侵必需满足的前提,而后置状态是成功实施这个入侵后所能得到的结果。根据前一个入侵的后置状态的实例是否与当前入侵的前置条件的实例相匹配,可识别出前后告警信息之间的相关关系。Alfonso Valdes 和 Keith Skinner 在文[6]中提出了一种基于概率的告警相关,主要思想是比较告警信息之间的相似性,比较的属性包括目标主机信息、入侵类型等,用0～1之间的数来表示相似程度,0表示没有相似性,1表示完全相似,并定义了相关的最小相似值,告警信息之间的相似性低于这个值的则认为是不相关的,只有大于或等于这个值才有可能相关。在文[9]中提出了一种用贝叶斯的多目标跟踪方法来对告警信息进行分类,以识别入侵过程。

这些告警相关方法均存在一些问题,第一,对于基于滥用

*)国家自然科学基金资助项目(No. 60273070)、国家863资助项目(No. 2003AA118201)、湖南省2004年科技攻关资助项目(No. 04gk3022)、东莞科研发展基金资助。龚发根 硕士研究生,主要研究方向:网络与信息安全。秦 拯 博士,主要研究方向:网络与信息安全。张大方 教授,博士生导师,主要研究方向:容错计算、网络安全、网络测试。

的网络型入侵检测系统来说，它所检测的数据包来自于整个网络，其工作原理是首先针对所有滥用行为进行建模，组成一个滥用模式库，检测时采用某种匹配算法对每个网络数据包进行分析，以判断其是否含有某种滥用模式，如果有则认为存在入侵并产生相应的告警信息，否则就认为没有入侵。从这可以看出，基于滥用的网络型入侵检测系统产生告警信息只能说明存在入侵的行为，入侵的结果是成功还是失败却是未知的。以上述 DDOS 入侵过程来说，在进行第二步入侵后，可能会发现远程主机虽然运行了 Sadmind 后台进程，但不存在漏洞，这时黑客仍有可能会试着继续进行攻击；或者黑客可能会采用事先编制好的脚本进行攻击，也会出现这种情况。在这两种情形下都会产生相应的告警信息，但入侵却是失败的。现有的告警相关方法无法识别属于某个入侵过程的失败的入侵步骤。第二，某些入侵步骤可能会存在多个必须满足的条件，对应于前置条件中的多个谓词表达式，要满足这些条件，就必须在实施这个入侵步骤之前，实施其它的入侵步骤，以便为该入侵做好准备。准备工作有时可能只需一个入侵步骤就可以完成；也有可能需要多个入侵步骤才能完成，每个入侵步骤只满足后续入侵的部分条件。也就是说，会存在带有分支的入侵过程。另外，入侵检测系统进行检测时，存在误肯定与误否定两种情况[8]。这样在进行告警信息的相关性分析处理时，不能简单地将满足部分条件的告警信息认为不相关，而应该考虑前后的告警信息可能存在部分相关、完全相关两种情况。本文针对这些问题提出了一个改进的告警相关方法。

2 一种基于知识积累的告警相关方法

黑客的每个入侵步骤，都是为了获得目标系统中某种状态信息或某种权限。随着入侵过程的不断深入，黑客所了解到的有关目标系统的知识越多，越有利下步入侵的实施。因此，黑客的入侵是一个逐步积累、逐步深入的过程，前一阶段的所有入侵步骤都可为后一阶段的入侵步骤或者是最终目标服务。为了方便评价每步的相关程度以及整个入侵过程的相关程度，本文将给出相关度这个概念。

2.1 相关知识库

入侵过程可以看成是一个有计划的序列活动，在这个过程中的每个入侵步骤对应于一个入侵动作。每个入侵动作的成功执行，必须满足一定的前提条件，称为前置条件。对入侵 A 来说，用 $PRE(A)$ 表示其前置条件。前置条件的实例如果与相应的目标系统的实际状态一致，则对应的入侵步骤就是可以成功的，否则不可能成功。同样，每个入侵动作成功执行后，也可获得一定的结果，称为后置状态。对入侵 A，用 $POST(A)$ 表示其后置状态。后置状态的实例表示在相应的目标系统上成功执行该入侵步骤后所能获得的系统状态，对黑客来说，就是其入侵的成果。黑客可以把这些信息积累起来，作为实施下步入侵的准备知识。因此，在进行告警信息是相关性分析处理时，就应该检测所有已经积累起来的知识是否能为当前的入侵步骤做准备。对于每种入侵模式，定义一种相应的相关模式，由这些相关模式组成相关知识库。每个相关模式用一个三元组来描述：

$intrusion\text{-}name(fact, prerequisite, consequence)$

其中 $fact$ 对应于所入侵的主机的 IP 地址、端口号等，在进行相关处理时用初级告警信息进行实例化；$prerequisite$ 是成功实施这个入侵动作必须满足的最小逻辑条件组合，在进行相关处理时用初级告警信息进行实例化；$consequence$ 是这个入侵动作成功实施后所能得到的结果的最小逻辑组合。在进行相关处理时用初级告警信息及现场知识库的信息进行实例化。之所以要用到现场知识库信息，主要是因为有的入侵步骤是用来探测目标主机信息的，其入侵结果对黑客是已知的，但对于告警相关处理部件来说却是未知的。例如，上述例子中的第二步的入侵，只能检测到存在 $SadmindPing$ 入侵，但告警信息中没有包含相应的主机上是否含有 $Sadmind$ 漏洞的信息，所以有必要用现场知识库的信息进行实例化，以便让告警相关处理部件知道黑客已经掌握的信息。相关模式实例化以后就形成了一条该入侵动作的超级告警。对于上面所举的 DDOS 攻击来说，它的每步入侵所对应的相关模式如下：

(1) $Ipsweep(\{IP\}, Null, \{ExistHost(IP)\})$

(2) $SadmindPing = (\{IP, PORT\}, \{ExiistHost(IP)\}, \{VulnerableSadmind(IP)\})$

(3) $SadmindBufferOverflow = (\{IP, Port\}, \{ExistHost(IP), VulnerableSadmind(IP)\}, \{GainRootAccess(IP)\})$

(4) $Null$

(5) $DDOSDaemon = (\{IP\}, \{GainRootAccess(IP)\}, Null)$

第四步由于已经取得了该主机的根权限，在其上安装程序一般不产生告警信息。

2.2 相关处理方法

相关处理方法就是用来判断告警信息之间是否存在相关关系的方法。假设 E、F 两个逻辑谓词表达式形式如下：

$\cdot E = expr_{E1}, expr_{E2}, \cdots\cdots, expr_{Em}$

$\cdot F = expr_{F1}, expr_{F2}, \cdots\cdots, expr_{Fn}$

其中每一个表达式 $expr_{E_i}$（或 $expr_{F_i}$），或是谓词形式，或者是否定谓词的形式。例如：

$\cdot expr_{E_i} = pred$

$\cdot expr_{E_i} = not\ (\ pred\ )$

其中 $pred$ 表示一个谓词，比如，ExistHost，VulnerableSadmind 等。

定义1（相关） 在 E 与 F 两个表达式中不存在互斥项的情况下，如果在 E 中至少存在的一项 $expr_{E_i}$ 能与 F 中的某一项 $expr_{E_j}$ 相匹配，则认为 E 与 F 是相关的。

2.3 相关度

按照这个定义进行相关处理时，可能会存在多个正在识别的入侵过程满足某个告警信息的匹配条件，另外，考虑到存在多个入侵为某一个入侵做准备（称为多对一）这种情况。所以在进行相关处理时，还要从黑客的角度来考虑，黑客入侵的过程，就是一个不断获取新知识、积累知识的过程。对于正在识别的入侵过程来说，应该用黑客已经掌握的信息与新的超级告警信息进行相关处理，并且只有当新的超级告警的后置状态中包含有新的知识，才能成为该入侵过程一部分。对于有多个入侵过程相匹配的情况，我们选择相关度比较大的几个，并保证每个入侵总的相关度小于等于1。

定义2（相关度） 表示每步相关处理的相关程度。假设 $S(A_1, \cdots, A_i)$ 表示正在识别的入侵过程，并且已经识别了 i 个入侵步骤，B 表示新的超级告警。如果 S 与 B 相关，则它们之间的相关度用 $\omega(S,B)$ 来表示如下：

$$\omega(S,B) = \frac{Matched(S(A_1, \cdots, A_i), B)}{NUMPRE(B)}$$

其中 $MatchedS(A_1, \cdots, A_i), B)$ 表示 $ALL\_POST(S)$ 与 $PRE(B)$ 之间相匹配的表达式数，$ALL\_POST(S)$ 表示已经积累

的知识，对应于正在识别的入侵过程中已积累的后置状态，$NUMPRE(B)$表示$PRE(B)$中的表达式数。$\omega(A,B)=0$表示正在识别的入侵过程A与超级告警B不相关，$\omega(A,B)=1$表示完全相关，$\omega(A,B)<1$，则表示部分相关，但是要实施入侵B，还需要其它的入侵步骤来为B做准备。如果$NUMPRE(B)$为0，则表明$PRE(B)$为空，一般来说是一个探测类型的入侵步骤，同时也意味着是一个入侵过程的开始。

定义3(入侵过程相关度) 用来评价一个入侵过程整体的相关程度。对于一个入侵过程$S(A_1,\cdots,A_n)$，其中A_i表示入侵过程中的某个入侵步骤，且A_i与A_{i+1}是相关的，$g(A_1,\cdots,A_n)$用来表示一个入侵过程相关度，表示如下：

$$g(A_1,\cdots,A_n)=\left(\sum_{i=1}^{n-1}\omega(S(A_1,\cdots,A_i),A_{i+1})\right)/(n-1)$$

2.4 现场知识库

由于告警信息的产生只能说明存在入侵的行为，但是入侵的结果是成功了还是失败了却是未知的。为了方便告警相关的处理并使其能掌握入侵过程的结果，我们定义了一个现场知识库，用来对实际的网络环境进行建模。在这个库中，主要包括活动主机信息，以及这些主机实际运行了一些什么样的服务及其可能存在的漏洞等。现场知识库的建立可借助现有的漏洞扫描软件来实现。在进行相关处理时，按照先评估后相关的原则，对每个超级告警，首先利用现场知识库信息对其前置条件进行评估，以判断实际目标系统是否存在实施该入侵的前提条件，以判断该入侵步骤的成败。评估完成后再进行相关性分析处理。对于失败的入侵步骤，在进行相关性分析处理时，首先使其前置条件中不能满足的谓词表达式取反，然后再进行相关性处理。这样，就可识别属于某个入侵过程的失败的入侵步骤了。

3 实现框架

本实现框架主要由重复告警合并、过滤、预处理、相关性分析和相关图形成五部分组成。这部分内容已经在文[10]中进行了详细介绍，这里只给出总体框图，如图1所示。

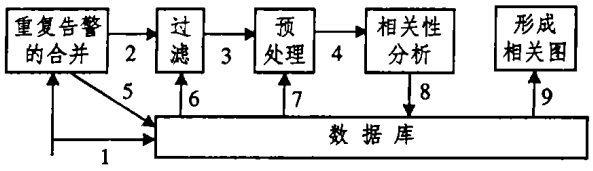

图1 告警相关部件结构图

4 实验结果

本文引用了两个评价告警相关有效性的参数：告警相关识别率和告警相关准确率。告警相关识别率用来表示告警相关的识别比例，也就是用已正确识别的入侵过程数除以总的入侵过程数；告警相关准确率表示告警相关的准确程度，也就是用已正确检测出来的入侵过程数除以实际检测到的入侵过程数。

本实验是在湖南计算机股份有限公司(股票代码0748)网络安全测试实验室分两次进行。第一次实验所用的数据是从客户处收集到的实际告警信息，我们将这些告警信息重组后统一发送给告警相关处理部件进行处理。这次实验使用了三个月的告警信息分别进行，根据实验结果统计出带分支的入侵过程和含有失败入侵步骤的入侵过程的比例，如图2所示。

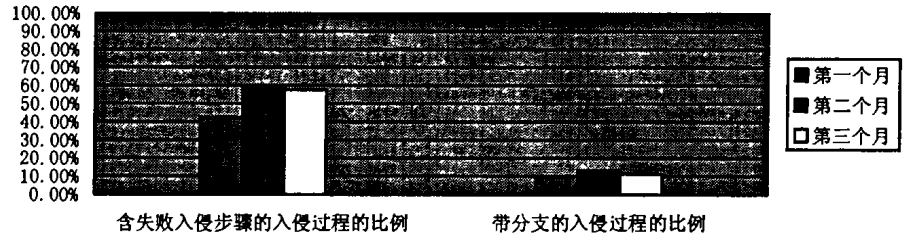

图2 第一次实验统计结果

从上图可以看也，在所有的入侵过程中，含失败入侵步骤的入侵过程在40%以上，而带分支的入侵过程数相对比较少，大约占10%。实验结果表明，使用这种改进的方法可以更加完整地识别入侵过程。

第二次实验采用模拟攻击的方法，在各检测点使用入侵检测系统进行检测，以便产生初级告警信息。在这次测试中，有8个检测点，均采用湖南计算机股份有限公司的分布式入侵检测设备。我们利用本文提出的方法与 Peng Ning 等提出的基于因果关系的方法分别进行，实验结果如图3所示。

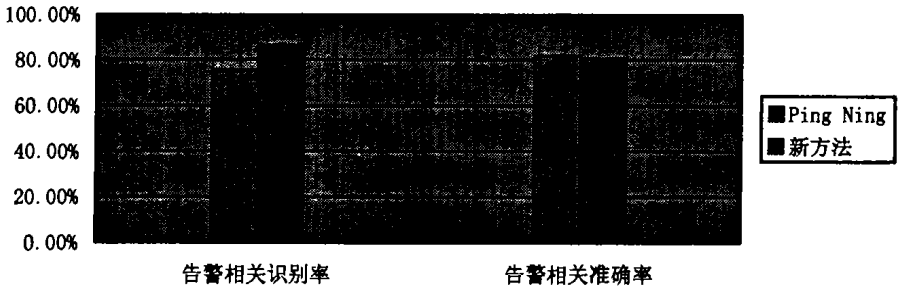

图3 第二次的两种方法对比的实验结果

从上图可以看出，新方法的识别率有一定提高，而在准确率上基本保持不变。虽然 Ping Ning 的方法识别率也比较高，但其所识别的入侵过程大多是不完整的。

结束语 在入侵检测系统中加入告警相关功能后，将入侵检测系统的功能由只能识别单个的入侵事件转变为能识别黑客的入侵过程，可以大大减轻系统安全管理员的工作，提高

入侵检测系统的可用性。下步工作将实现一个相关图的形成部件，以便更好地表示入侵过程。

参 考 文 献

1 李娜，秦拯，张大方，陈蜀宁. 基于 Markov Chain 的协议异常检测模型. 计算机科学，2004，31(10)

2 Yang Jin-Min，Zhang Da-Fang，Qin Zheng. et al. WINDAR：A Multithreaded Rollback-Recovery Toolkit on Windows. In：IEEE 10th Proc. of Pacific Rim Dependable Computing (PRDC10)，2004

3 Qin Zheng，Wu Zhong-fu，Liao Xiaofeng，et al. A Network Intrusion Detection Architecture Based on Intelligent Agents. In：Proc. of the Intl. Conf. on NIT，2001

4 Ning P，Reeves D，Cui Yun. Correlating alerts using prerequisites of intrusions：[Technical Report TR-2001-13]. North Carolina State University，Department of Computer Science，Dec. 2001

5 Debar H，Wespi A. Aggregation and correlation of intrusion-detection alerts. In Recent Advances in Intrusion Detection，number 2212 in Lecture Notes in Computer Science，2001. 85～103

6 Valdes A，Skinner K. Probabilistic alert correlation. In：Proc. of the 4th Intl. Symposium on Recent Advances in Intrusion Detection (RAID 2001)，2001. 54～68

7 Ning P，Cui Y. An intrusion alert correlator based on prerequisites of intrusions：[Technical Report TR-2002-01]. North Carolina State University，Department of Computer Science，2002

8 Bace R G. Intrusion Detection. Macmillan Technology Publishing，2000

9 Burroughs D J，Wilson L F，Cybenko G V. Analysis of Distributed Intrusion Detection Systems Using Bayesian Methods. http://www. ists. dartmouth. edu/IRIA/published，2002

10 秦拯，龚发根，张大方. 分布式入侵检测系统中告警相关的研究与实现. 计算机科学与工程，已录用

(上接第117页)

MsLB-Triang 较好地改进了目前广泛采用的 Min. Weight Heuristic 算法，是一种高效率的三角化方法。

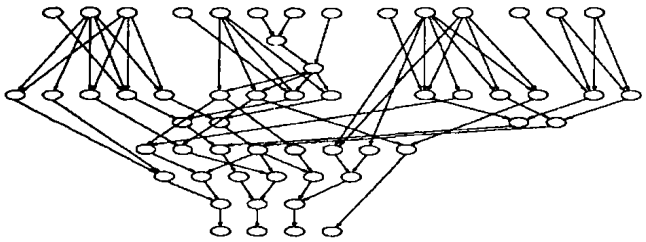

图5 Medianus-II 信度网

MsLB-Triang 算法已经用于作者开发的信度网推理程序中。在本研究中，算法采用了比较简单的实现方式，其复杂性还有进一步降低的可能。今后，我们将进一步对此算法的复杂度进行深入研究，争取提高算法的效率，降低其复杂度。

参 考 文 献

1 Special issue on probability forecasting. International Journal of Forecasting，1995(11)

2 Special issue on real_world applications of uncertain reasoning. Communications of the ACM，1995，38

3 Andreassen S，Jensen F V，Andersen S K，et al. MUNIN — an expert EMG assistant. In：J E Desmedt，ed. Computer-Aided Electromyography and Expert Systems，chapter 21. Elsevier Science Publishers，Amsterdam，1989

4 Becker A，Geiger D. Optimization of Pearl's method of conditioning and greedy-like approximation algorithms for the vertex feedback set problem. Artifciial Intelligence，1996，82：1～22

5 Berry A，Bordat J P，Heggernes P，et al. A wide-range algorithm for minimal triangulation from an arbitrary ordering：[Reports in Informatics 243]. University of Bergen，Norway，2003

6 Berry A，Blair J，Heggernes P. Maximum Cardinality Search for Computing Minimal Triangulations. In：Proc. WG 2002 - 28th Workshop on Graph Theoretical Concepts in Computer Science，Cesky Krumlov，Czech Rebublic. Springer Verlag，Lecture Notes in Computer Science 2573，June 2002. 1～12

7 Cano A，Moral S. Heuristic Algorithms for the Triangulation of Graphs. In：Proc. of the Fifth Intl. Conf. on Information Processing and Management of Uncertainty in Knowledge-Based Systems，IPMU94，Paris，1994. 166～171

8 Cooper G F. The computational complexity of probabilistic inference using Bayesian belief networks. Artificial Intelligence，1990，42(2-3)：393～405

9 Dirac G A. On rigid circuit graphs. Anh. Math. Sem. Univ. Hamburg，1961，25：71～76

10 Gamez J，Puerta J. Searching for the best elimination sequence in Bayesian networks by using Ant Colony based optimization：[Technical Report # DIAB-01-04-13]. Department of Computer Science，University of Castilla-La Mancha，Jan. 2000

11 Huang C，Darwiche A. Inference in belief networks：A procedural guide. Intl. J of Approximate Reasoning，1996，15(3)：225～263

12 Jensen F V，Jensen F. Optimal junction trees. In Uncertainty in Artificial Intelligance. In：Proc. of the Tenth Conf. San Mateo，CA，Morgan Kaufman，July 1994. 360～366

13 Jensen F V，Lauritzen S L，Olesen K G. Bayesian updating in causal probabilistic networks by local computations. Computational Statistics Quarterly，1990，4：269～282

14 Kjærulff U. Triangulation of Graphs - Algorithms Giving Small Total State Space：[Technical Report R 90-09]. Department of Mathematics and Computer Science，Aalborg University，Denmark，1990

15 Lauritzen S L，Spiegelhalter D J. Local computations with probabilities on graphical structures and their application to expert systems. Journal of the Royal Statistical Society，Series B (Methodological)，1988，50(2)：157，224

16 Lekkerkerker C G，Boland J C. Representation of a finite graph by a set of intervals on the real line. Fund. Math.，1962，51：45～64

17 Murphy K. The Bayes Net Toolbox for Matlab. Computing Science and Statistics. Available at：http://citeseer. ist. psu. edu/murphy01bayes. html. 2001，33

18 Pearl J. Probabilistic Reasoning in Intelligent Systems：Networks of Plausible Inference. In：Proc. National Conf. on AI，Pittsburgh. Morgan Kaufmann，San Mateo，CA，1988. 133～136

19 Pearl J. Probabilistic Reasoning in Intelligent Systems. Morgan Kaufmann，San Mateo，California，1988

20 Rose D J. A graph-theoretic study of the numerical solution of sparse positive definite systems of linear equations. In：R C Read，ed. Graph Theory and Computing，Academic Press，1972. 183～217

21 Rose D J，Tarjan R E，Lueker G S. Algorithmic aspects of vertex elimination on graphs. SIAM J. Comput.，1976，5：266～283

22 Shachter R D. Evidence absorption and propagation through evidence reversals. In：Henrion. M，Shachter R d，Kanal L N，Lemmer J F，eds. Uncertainty in Artificial Intelligence，Elsevier Science Publishers，Amsterdam，The Netherlands，1990，5：173～190

23 Shachter R D，Andersen S K，Szolovits P. The equivalence of exactr methods for probabilistic inference on belief networks：[Technical report]. Department of Engineering-Economic Systems，Stanford University，Stanford，California，1991

24 Tarjan R E，Yannakakis M. Simple linear-time algorithms to test chordal-ity of graphs，test acyclicity of hypergraphs，and selectively reduce acyclic hyper-graphs. SIAM J. Comput.，1984，13：566～579