计算机科学2002Vol. 29№. 2

免疫思想在计算机安全系统中的应用*)

The Application of the Immune Principle to Computer Secure Systems

闫 巧[1] 谢维信[2]

(西安电子科技大学133#信箱 西安710071)[1] (深圳大学 深圳518060)[2]

Abstract This paper overviews the development of immune principle applied in computer virus detection ,intrusion detection based host and intrusion detection based local area broadcast network. An inspiration that the features of natural immune system can be used to build a distributive intrusion detection system facing large- scale network is achieved.

Keywords Immune system intrusion detection,Negative selection

1. 引言

计算机安全问题在 Internet 时代变得空前重要起来。而将人类自然免疫系统的发现和研究成果应用到计算机安全系统是很自然的想法。人类的自然免疫系统的作用是保护人类免受病菌的侵害,而计算机安全系统所扮演的角色是防止计算机网络遭到病毒、蠕虫和入侵者的攻击,因此两个系统之间的类似之处是非常明显的。

人的自然免疫系统主要包括皮肤、生理条件、天生免疫系统和自适应免疫系统几部分。其中自适应免疫系统最复杂,它主要包括两类用于检测病原体的淋巴细胞即 B 淋巴细胞和 T 淋巴细胞。这两类淋巴细胞具体的作用可参考文[2],我们在这里只简单地介绍我们所要用到的几个概念:

1)自身:指体内的细胞或无害的体外细胞,我们用 S 表示自身集合。

2)非自身:指体外的有害细胞,即病原体,我们用 NS 表示非自身集合。

3)否定选择:所谓否定选择是自然免疫系统产生不会伤害自身的成熟的淋巴细胞的一种方法。具体是指在一定时间 T_N 内,若随机生成的淋巴细胞粘附住了体内自身的细胞,则这些淋巴细胞死亡,只有那些不会粘附任何体内自身细胞的淋巴细胞才能存活并离开产生它们的特定区域。我们用 D_0 表示随机生成的淋巴集合,用 D 表示经过了否定选择的淋巴集合。

4)自免疫:所谓自免疫是指淋巴细胞错误地将自身当作病原体而消灭掉的一种病理现象。

5)协同信号:是指为了减少自免疫的发生,淋巴细胞被激活时除了需要超过一定激活阈值以外还必须收到第二种信号,这种信号可由 T 淋巴细胞提供给 B 淋巴细胞或是由体内组织受损信号提供给 T 淋巴细胞。

2. 免疫思想在计算机安全中的应用

2.1 免疫思想在防计算机病毒中的应用

最早将自然免疫系统的一些思想引入信息安全的是 Forrest,她主要是借鉴免疫中否定选择的思想来保护静态数据免遭病毒的修改。她将信息安全问题看成一个更加普遍的问题即如何区分敌我或者说自身与非自身[3]。此处自身是指合法用户或被保护的数据、文件等,非自身是指非授权用户或被篡改了的数据。常见的病毒检测技术有三种:活动监视、特征扫描和文件认证。Forrest 提出的算法属于最后一种,是一种检测变化的方法。其算法具体说来包括三个阶段:

1)生成自身集合。将被保护数据串分成等长的短序列构成自身序列的集合。

2)生成一组检测器。类似免疫系统中否定检测的思想,随机生成一组随机串,将每个随机串与自身序列串进行匹配,若能够匹配则丢弃此随机串,若该随机串与任何自身序列的序列都不匹配,则将此随机串纳入检测器集合。具体生成方法如图1。

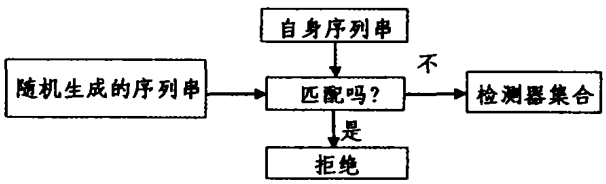

图1 生成检测器

3)监控被保护的数据。将被保护数据与检测器相比较,若检测器被激活,则说明有变化发生。具体算法如图2。

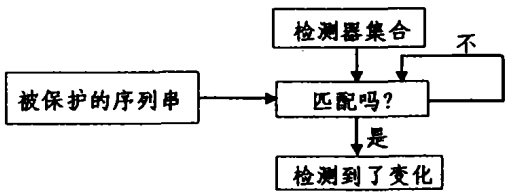

图2

该算法采用 R 连续字符串匹配的不完全匹配的方法即若两个串在至少连续 R 个位置相匹配被认为是匹配。该方法其检测算法的每个拷贝都是唯一的、检测时概率可调、提供的保护是双向的,即同一机制既保护被保护数据又保护检测器自身等多个优点。

2.2 免疫思想在基于主机的入侵检测中的应用

*)国家863应急项目 信息安全技术 项目号301-6-6.闫 巧 博士生.谢维信 博士生导师,深圳大学校长。

后来 Forrest 对区分自身与非自身的思想进一步进行延伸，建立了对 UNIX 特权进程自身的定义。他们通过大量实验说明运行进程的系统调用的短序列有稳定的正常行为特征。当异常事件发生时，该特征会发生变化。

在他们所提出的时延嵌入序列(TIDE)的方法中，通过列举出现在训练数据中所有唯一的、预先给定长度为 K 的连续序列来构造程序正常行为轮廓数据库。当选择序列长度为 K 时，他们将长度为 K 的窗口通过每个正常轨迹，一次滑动一个系统调用，向正常轮廓库中添加唯一的序列。建立这样的数据库轮廓只需要一次遍历数据。为了节省存储空间和加速比较，将序列树状存储。检测时，来自检测轨迹的序列与正常数据库轮廓中的序列相比较，在数据库中找不到一样的序列叫做不匹配。任何一次不匹配都说明该序列是没包括在正常训练数据库轮廓中的序列，它可能是异常行为。通过计数不匹配的次数，并求出不匹配次数占数据库中总序列的百分比，再将这个百分比与预先给定的阈值相比较，就可以判断程序每次执行是正常还是异常。该项技术对 UNIX 的程序 SENDMAIL、LPR、FTPD 等都很有效[1]。

该方法采用短序列的完全匹配方法，具有自然免疫系统的多样性的特点。每一个站点都有基于本地软件的独特的正常轮廓，因此利用同一弱点的一个入侵很难在多个站点同时得逞。

2.3 在基于网络的入侵检测中的应用

Hofmeyr 等人将免疫的原理和思想推广到网络入侵中去。在一个局域广播网内，他将自身定义成为计算机间正常的连接(包括局域网内部计算机的连接以及外部计算机与局域网内部计算机的连接)。一次连接用相连的计算机的源 IP 地址、目的 IP 地址和服务共49位数据串来表示，如图3所示[4]。而将非自身定义为局域网上不常见的连接。

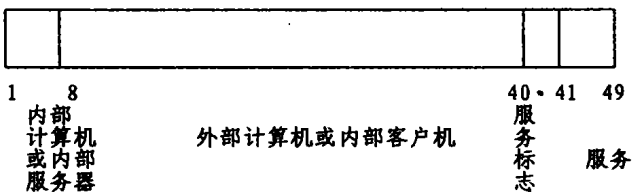

图3 连接表示

每个检测器有随机生成的49位的数据串和检测器状态(随机的、成熟的、激活的和记忆)。两个串匹配用 r 个连续位置相同的不完全匹配。首先在局域网的每台计算机上随机异步地生成随机检测器集合 D_0。由于该局域网是广播网，因此检测器可以监听到网上所有连接。我们将一段时间 T_S 内局域网上的所有连接定义为自身集合 S(假设这段时间内异常连接发生的概率极低)。而后在否定选择时间 T_N 内让随机生成的检测器集合经历否定选择，即若 $d_0 \in D$ 与任何一个 $s \in S$ 相匹配，则 d_0 死亡，经过了 T_N 时间后，称经过了否定选择的检测器 d_0 为成熟的检测器 d。所有经过否定选择的 d_0 全体构成成熟的检测器集合 D。D 中的所有的检测器 d 与自身集合 S 中的检测器 s 都不匹配。成熟的检测器 d 可以进行独立的检测，若一个在激活时间 T_A 内 d 匹配了超过检测阈值 TH_D 个连接包，则 d 变为激活检测器 d_a，否则该成熟检测器 d 死亡。激活检测器 d_a 若在等待协同信号时间 T_C 内收到协同激活信号(文[4]是由管理员发出的肯定是入侵的邮件)，则激活检测器 d_a 变成记忆检测器 d_m，否则认为本次激活是虚警，所以激活检测器 d_a 死亡。记忆检测器 d_m 有比 TH_D 更低的激活阈值 TH_M。每台主机上都同时有大量不同状态的检测器存在，其中 d 可用来检测新的未知的异常连接，而对已知的异常 d_m 更为敏感。检测器的生命周期如图4所示。

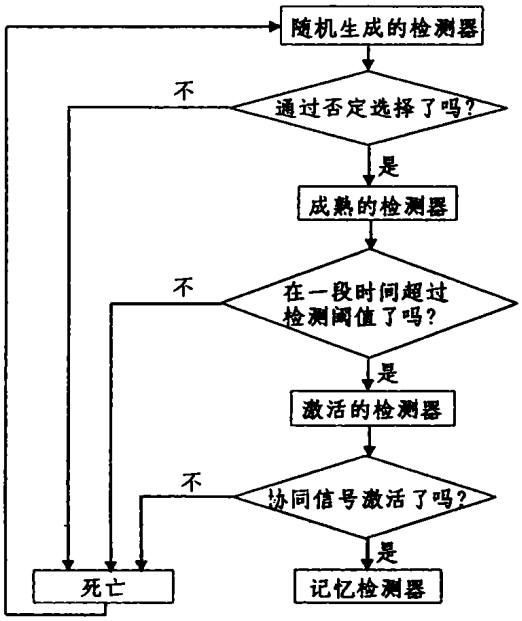

图4 检测器生命周期

3. 启发与研究方向

如前所述免疫思想已经成功地应用在保护静态数据、基于主机的入侵检测和基于局域广播网络的入侵检测之中。我们从中可看到将自然免疫系统用于计算机安全问题的几个关键问题是：1)如何定义自身与非自身。2)如何定义匹配，已经采用的匹配方法有完全匹配和 R 连续位置匹配的不完全匹配。3)怎样将自然免疫的一些机理用于我们的具体研究问题上去。比如否定选择、协同信号等。

自然免疫系统还具有许多特点，这些特点包括多层次、分布性、多样性、容错性、动态防护、自适应性、联想记忆、自我防护等等[5]，我们可进一步将这些特点扩展到面向大规模网络的分布入侵检测之中。我们可将每一台计算机看成一个细胞，每一个子网看成一个个人，而将整个 Internet 网络看成人群，可在每个子网建立独特的入侵检测系统，子网内部的入侵检测系统具有自然免疫系统所具有的分布性、多样性、自适应性等特点，可实时检测所有入侵，并对检测到的入侵采取相应的防范措施。而子网之间可相互通信，当有一个子网受到入侵后可发出信息，告知所有子网采取相应的预警措施。目前基于这种设想的大规模网络入侵检测系统我们正在研究之中。

参 考 文 献

1 Forrest S, et al. A Sence of Self for Unix Processes . [A] IEEE Symposium on Security and Privacy[C]. Oakland: California IEEE Computer Society, 1996. 120～128

2 Forrest S, Hofmeyr S A, Somayaji A. Computer Immunology, Communications of the ACM, 1997, 40(10): 88～96

3 Forrest S, et al. Self-nonself discrimination in a computer. In: Proc. of the 1994 IEEE Symposium on Research in Security and Privacy, Los Alamos, CA, IEEE Computer Society Press, 1994

4 Hofmeyr S A. A Immunological Model of Distributed Detection and its Application to Computer Security: [PhD thesis]. Department of Computer Sciences, University of New Mexico, Albuquerque, NM, April 1999

5 Somayaji A et al. Principles of a Computer Immune System. In: Proc. of the Second New Security Paradigms Workshop, 1997