面向滥用检测的反馈预测机制[1]

张剑，龚俭
（东南大学计算机科学与工程系，南京 210096）

【摘要】:
在高速主干网络环境中的入侵检测系统应该满足两个要求：第一，需要尽早发现入侵企图；第二，要努力降低入侵检测的操作代价。两者的解决办法与入侵检测模型和测度密切相关。本文在一般的滥用检测系统中建立一个反馈预测机制，它不仅能预测用户当前行为是否入侵，而且能大幅度降低该入侵检测系统的操作代价，可适应在高速网络中的实时检测需要。实际测试结果表明反馈预测机制能比较精确地预测入侵，嵌入了反馈预测机制的滥用检测系统的数据处理能力有了较大的改善。

【关键字】入侵检测；滥用检测；异常检测；预测；

一．前言

计算机联网技术的发展改变了以单机为主的计算模式，但是，网络入侵的风险性和机会也相应地急剧增多。设计安全措施来防范未经授权访问系统的资源和数据，是当前网络安全领域的一个十分重要而迫切的问题。目前，要想完全避免安全事件的发生并不太现实，网络安全人员所能做到的只能是尽力发现和察觉入侵及入侵企图，以便采取有效的措施来堵塞漏洞和修复系统，这样的研究称为入侵检测，为此目的所研制的系统就称为入侵检测系统（intrusion detection system，简称 IDS）。

入侵检测技术根据检测方法的不同分为两大类，一种称为滥用检测方法[1]，该方法通过对采集的信息按已知的知识进行分析，发现正在发生和已经发生的入侵行为；另一种称为异常检测方法[2]，它通过采集和统计发现网络或系统中的异常行为，然后按照某种决策算子来判断它是否入侵。滥用检测方法的误报率较低，但它不能检测出未知的攻击方式，而且检测的操作代价较高，特别是在高速主干网络环境中，难以应付庞大的数据量；而异常检测方法能检测出未知的攻击方式，但其误报率较高。通常将两种方法结合起来能够扬长避短，从而取得比较好的检测效果。

本文所指的预测是细时间粒度的针对用户初始行为判断其是否攻击的技术，它能在掌握较少情报的情况下察觉入侵企图。该方法的优点是反应灵敏，能较早发现入侵，从而阻止其进一步深入；而且数据处理速度较快，能适应对庞大的数据量实时处理的需要。但预测系统存在误报较高的风险，因此它应当与入侵检测系统配合使用，一方面利用入侵检测的成果反馈到预测模型，使其有较高的准确性；一方面可以分担入侵检测系统的部分工作量，提高入侵检测系统的可用性。这种技术称为反馈预测机制 FPM（Feedback Prediction Mechanism）。

Wenke Lee[3]曾经做过类似的工作，他将数据挖掘技术应用到异常检测模型的建立中，该异常检测模型实际是入侵检测规则集。为了提高该模型的效率达到实时入侵检测和尽早发现入侵的目的，他将入侵检测测度按照计算代价的高低分为四类，计算代价不仅包含计算该测度所花费的时间，同时也包含能够计算该测度的等待时间。然后按照所需要用到的测度将入侵检测规则分为四类，即第一类规则只用到第一类测度，第二类规则用到第一、二类测度，以此类推。每当一个网络连接到达时，系统首先用第一类规则来评估该网络连接，如果匹配且置信度大于预定阀值，该规则就被激活，由于第一类测度最容易且能最早被计算，因此这时候入侵事件的检测效率是最高的；否则就用第二类规则来评估，并以此类推。Lee 的入侵检测系统实际也具有预测入侵的能力，但它需要从原始数据中用数据挖掘技术自动获取，原始数据尤其是高速主干网络的数据往往是海量的，因此定期更新入侵检测规则是制约该系统的瓶颈；而如果不定期更新，则随着时间的推移必然造成较高的误报率和漏报率。此外其入

[1]本文受国家自然科学基金项目 90104031 资助

侵检测模型也具有异常检测方法的弱点，即误报率较高。

本文对一般的滥用检测系统结构进行改造，为其增加 FPM。由于 FPM 的核心技术是对滥用检测系统的入侵事件库进行数据挖掘，因此它不但误报率较低，而且也不容易出现海量数据处理的问题，从而提高了滥用检测系统的可用性和效率。在 CERNET 的高速主干网络环境中对带 FPM 的滥用检测系统进行了性能测试，结果表明 FPM 能较精确地预测入侵，嵌入了 FPM 的滥用检测系统的数据处理能力有了较大的改善。

本文以后的章节安排如下：第二节介绍 FPM 的机制及其理论依据，第三节讨论如何将 FPM 与滥用检测相结合，第四节是实验，最后是总结。

二．FPM 的基本原理

FPM 的目的有两个，一是使滥用检测系统具备入侵预测能力，二是提高滥用检测系统的数据处理能力。嵌入了 FPM 的滥用检测系统不但能在入侵完成前发现该行为，而且具备实时处理高速主干网络环境的庞大数据量的能力。下面分两节讨论 FFM 的预测原理和高速数据处理原理。

2.1 FPM 的预测能力

Wenke Lee[3]曾经将在网络入侵检测中用到的测度归为四类：第一类测度是能从网络连接的第一条报文中计算得到的，例如服务类型；第二类能在网络连接生命期的任一点计算得到，例如连接状态（SYN_WAIT，CONNECTED，FIN_WAIT 等）；第三类只能在连接最后才能计算得到，但只依赖当前连接的信息，例如当前连接的发送字节总数；第四类不但在连接的最后才能计算得到，而且还需要访问以前网络连接的信息，例如一些暂时、统计性的测度，这是计算代价最高的一种测度。

Lee 的测度分类方法总结起来就是第一、二类的测度是用户行为的初始特征、它们在用户行为结束前可以计算出来，而第三、四类的测度不但需在用户行为结束后才可计算出来，而且计算时间较长，即类别越高的测度的计算代价越高。

根据 Lee 对网络行为测度的分类，可知对某种入侵的检测仅需要计算第一、二类测度时，可以预测入侵行为；而检测需计算第三、四类测度时就不能做到预测（Prediction），因为这两种测度都必须在用户行为结束后才能计算出来。但是对于网络入侵检测系统而言，绝大多数的入侵行为的特征都包括第三、四类测度，而且这些测度包含的入侵信息量往往最大，如果忽略这些测度，可能会导致较高的误检率。因此提高预测的准确性是一个核心问题，下面介绍如何采用数据挖掘规则裁减和周期性更新来提高预测的准确性。

（1）建立入侵事件数据库，该库由入侵事件记录构成，每条记录包含时间戳、源地址/目标地址，源端口/目标端口、事件类型等属性，这些入侵事件记录都是系统的检测结果。

（2）运用数据挖掘技术对该库进行处理，目的是发现各属性与事件类型的关联规则。

（3）利用预定的置信度阀值对这些关联规则进行过滤，只保留置信度大于该阀值和对检测有意义的关联规则。由于这些规则仅需要匹配诸如源地址/目标地址等特征，因此它们比原来的规则的计算代价降低了。

（4）将保留的关联规则反馈回滥用检测系统的规则库，从而提高系统的数据处理能力。

（5）定期更新这些关联规则，以保持关联规则的准确性。

FPM 首先从一组标记为各种入侵方式的安全事件中提取出一组规则作为预测模型，每条规则根据安全事件第一、二类特征的不同的值将安全事件分类为各种入侵方式。由于这些规则可能导致较高的误报率，一个阀值 τ 被用来裁减置信度[2]低的规则，当规则的置信度小

[2]置信度的定义是：在一组训练数据 S 中，每条数据是一条安全事件，对于入侵方式 i，设 P 是预测为 i 的数据集，而 W 是 S 中标记为 i 的数据集，则置信度 $\tau = \frac{|P \cup W|}{|P|}$。

于 τ 时，该规则从预测模型中删除，最后预测模型中只保留了置信度大于 τ 的规则。

需要在预测模型的规则裁减和周期性更新间取得较好的平衡。如果规则数量太少，说明大部分规则都不能满足准确度的要求，这样就要计算第三、四类特征，实际上就丧失了预测的功能；而如果为了保持一定的准确度，而过分缩短更新周期，就会给系统带来较重的负担，实质上也是提高了预测的计算代价。

综上所述，FPM 实质是通过初始行为特征来预测行为是否入侵以及何种入侵的机制，同时采用预测模型规则裁减和周期性更新来提高预测的准确度。

2.2 FPM 的计算代价

滥用检测方法一般首先定义若干入侵检测的测度，然后搜集已知入侵方式的测度值，利用这些特征来匹配检测对象的当前行为，以判断系统是否正在遭受攻击以及遭受何种攻击。因此，这些特征组成了入侵行为的测度集。滥用检测系统在搜集入侵证据和检测时会遇到以下问题：

（1）某些特征的收集代价较高，甚至需要上下文相关知识。

（2）某些特征的匹配代价较高。例如在网络入侵检测中，某些特征的匹配需要搜索整个报文内容，而不是报文头，这样的代价要比单匹配报文头的某个字段大多了，这也是形成滥用检测系统瓶颈的主要原因。

（3）当某种入侵方式包括多个特征时，该入侵行为的检测代价就是各特征的计算代价之和。如果按照上面介绍的 Wenke Lee 的分类方法，记第一、二、三和四类特征的计算代价分别为 1、5、10、100，记该入侵方式的计算代价为 C，其一、二、三和四类特征的个数分别为 n_1，n_2，n_3，n_4，则

$$C_d = n_1 + 5n_2 + 10n_3 + 100n_4$$

如果计算代价比较高，就会严重影响系统的数据处理速度，特别在高速主干网络环境中，这会引起高的数据丢失甚至导致系统不可用。而利用 FPM 进行入侵预测，只需要计算第一、二类的特征，其计算代价为：

$$C_f = n_1 + 5n_2$$

设 $n_1=n_2=n_3=n_4$，则 $C_f / C_d \approx 1/20$。由此可见，预测机制对降低计算代价有显著的作用。特别如果在高速主干网络环境的入侵检测系统中嵌入了 FPM，由于 FPM 有较高的数据处理速度，因此可有效地提高入侵检测系统的数据处理量，降低由于入侵检测的计算代价过高带来的数据丢失。

Wenke Lee 曾提出一个基于多规则集的代价敏感入侵检测模型[4]，在该模型中根据计算代价的不同将规则分为四类，第一类规则只需要计算第一类特征，第二类规则需要计算第一、二类特征，如此类推。考虑到根据预测的需要，一、二类特征可以归成一类特征，而三、四类特征归为另一类特征，同时为了简化设计和实际的需要，本文提出两级结构的滥用检测模型，FPM 作为第一级检测模块只需计算第一、二类特征，且具有预测功能，而原有的滥用检测模块作为第二级需计算所有的特征。下面详细介绍该模型的实现。

三．FPM 与滥用检测系统的结合

提高滥用检测系统的数据处理能力和预测能力是 FPM 的两个主要目标，其方法是用计算和匹配代价低的特征来建立预测模型，将基于该模型的预测机制加入滥用检测系统中。预测模型可以是用数据挖掘技术建立的决策树或关联规则，也可以是多元回归模型或其他匹配规则，它被用来判断当前用户动作是哪种入侵行为。该模型具有以下的特点：

（1）预测模型中的变量均是在入侵完成前能计算得到的，即属于第一、二类特征。

（2）该模型的初始化是以滥用检测系统的成果为基础的。即在系统运行一段时间后，预测机制自动从系统的检测结果中反馈总结出预测模型，然后投入到实际运行中。因此该机制被称为反馈预测机制（FPM）。

（3）最近的检测结果必须以一定周期反馈到预测模型中，以提高该模型的准确性。

根据上述思想，本文以基于网络的滥用检测系统为改造对象，提出一个带 FPM 的滥用检测系统结构框架。如图 1 所示：

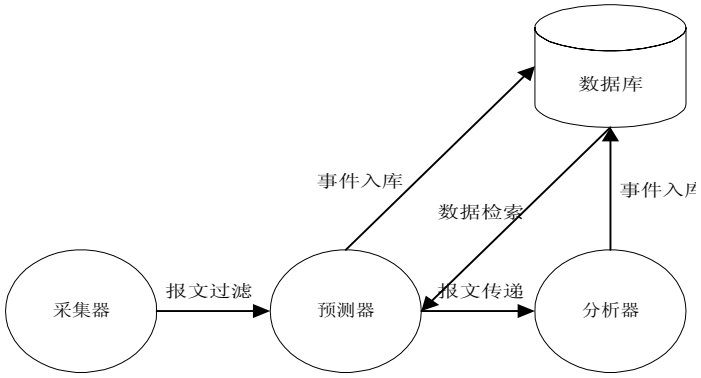

图1 带FPM的滥用检测系统结构

该系统结构包括采集器、预测器、分析器和数据库组成，采集器从事简单的报文过滤工作，并对过滤的报文按照预测器的要求进行适当的格式化处理，然后送到预测器。为了尽可能提高数据处理速度，减少数据丢失，采集器上的规则只可能是对报头进行简单的过滤。预测器检查到的入侵事件被写入数据库；每隔一段时间，预测器从数据库取出最近的入侵事件测度数据更新预测模型。预测器将不能识别的报文传递给分析器，由分析器根据滥用检测模型来识别入侵，并将检测到的入侵事件写入数据库中。

FPM 用预测器实现，它根据预测模型判断当前报文是哪种入侵行为，预测模型中的变量都可以从报文头的字段直接获取，例如发起攻击的源地址，目标地址和端口等，因此预测器理论上会为分析器分担大量甚至多数的数据处理任务。预测器的特点是处理报文的速度极快，能够承受从采集器下来的庞大数据量。作为 FPM 的响应政策的一部分判断为入侵的行为被写入数据库中，而不能为预测器识别的报文将传递到分析器中。预测器按一定的周期更新预测模型，更新的方法是从数据库中取出最近的数据建立临时的预测模型，然后临时模型与旧模型按一定的算法合成新的预测模型。

分析器是传统的滥用检测模块，它的任务与原来的一样，根据滥用检测模型对输入报文进行检查，这部分的检测代价比较高，可能包括检索报文内容之类的和上下文相关匹配的工作。当它检测到入侵，作为响应政策的一部分也将其记录到数据库中，其格式与预测器的一样，一般都是第一、二类的特征，因此分析器的数据丢失会大大降低。而在试验中也验证了我们的想法。

四．试验及其结论

4.1 试验系统结构及其运行环境

本试验的目的是测试 FPM 的准确性和数据处理能力（即吞吐量），本试验系统使用了 CERNET 主干信道所提供的实际数据，其中采集器和监测器的运行环境为 Intel ISP4400 server，Redhat 6.2；而预测器和分析器的运行环境为 Intel ISP2150 server，Redhat 6.2；后处理器为 Sun Sparc20，Solaries 7。其系统框架如图 2 所示：

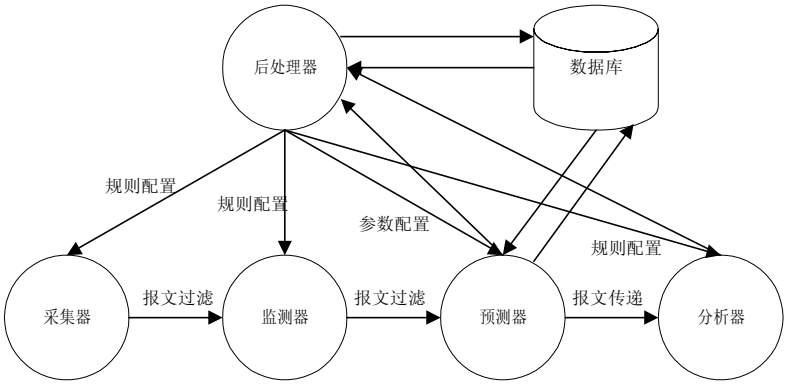

图2 带FPM的滥用检测系统结构2

本系统实际是图 1 所示系统的细化，增加的后处理器的功能有三个：

（1）配置采集器和检测器的报文过滤规则，预测器的相关参数和分析器的匹配规则。

（2）接受预测器和分析器的检测结果并负责将其格式化后入库。

（3）向系统管理员报告最近发生的入侵事件。

采集器和监测器都是报文过滤机制，两者的过滤条件都取自报文头。采集器的过滤规则较简单，它是基于<地址，端口>的二元组过滤机制。而监测器的过滤条件可以涉及报文头的所有字段，通过多维的高速分类算法来实现。

所有检测到的入侵事件都以统一的格式标识并入库，该格式包括事件类型、源地址/端口、目的地址/端口、开始时间，结束时间等。

分析器实际是一个专家系统，其检测模型是一组规则，每条规则由入侵行为名称、编号、特征描述和响应方式组成。

预测模型采用分类模型中的 Discriminent 过程，该过程适合处理因变量有两个以上分类值的情况，自变量是源网络、目的网络、目的端口，因变量是事件类型编号。该模型的目的是利用源网络、目的网络和目的端口来预测入侵事件类型。设初始的预测模型更新周期是 T，每当更新时间到达，利用数据库中最近 T 时间段内的数据计算临时预测模型 F_t，设原预测模型为 F_o，则更新后的预测模型为 $F_n = \omega_t * F_t + \omega_o * F_o$，其中 $\omega_t > \omega_o$。ω_t 和 ω_o 分别是临时预测模型和原预测模型所占的比重。FPM 正是由预测器和数据库组成。

4.2 试验 1：测试 FPM 的准确性

本试验包括两部分，第一部分测试预测模型在不更新的情况下的准确性，第二部分测试该模型在不同的更新周期下的准确性。试验的方法是关闭预测器，让系统在没有预测器的情况下运行一天。然后利用数据库中这一天的数据测试预测器。数据量大约是 60 万条记录。

第一部分测试时，设预测模型的更新周期为 1 小时，用当天第一小时的入侵事件（2 万多条）训练分类模型。用该模型来预测剩余 23 个小时的入侵事件，误测率由事件编号判断错误的数量比例代表，由此形成的时间－误测率曲线如图 3 所示：

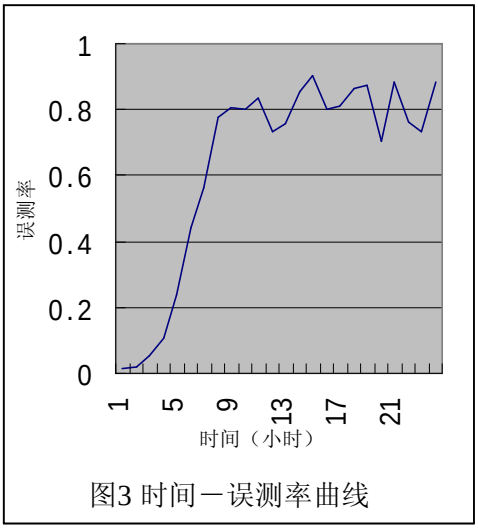

图3 时间－误测率曲线

由图 3 可看出，随着时间的推移，预测模型的误测率迅速升高，这是因为预测模型对入侵的预测是建立在不完整的特征描述基础上，它只能在短期内比较精确，而从长期来看非常不稳定的。因此，必须定期更新预测模型，以保持其准确性。

第二部分测试在不同的更新周期下模型的准确性，取更新周期为 30、60、90 分钟等 10 个点，绘制而成的“更新周期－误测率”曲线如图 4 所示：

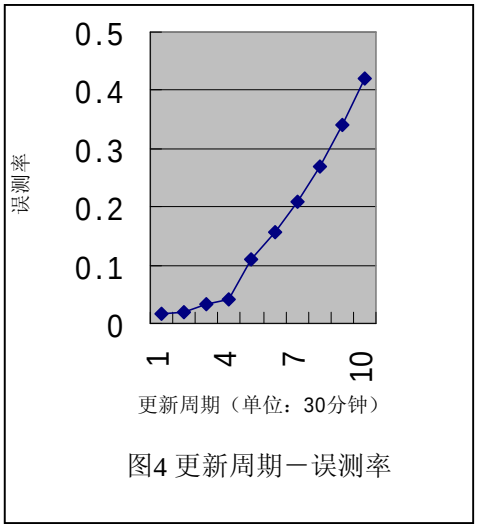

图4 更新周期－误测率

从理论上讲，更新周期越短，预测模型就越准确，实际上整个趋势也是如此，但在一定的更新周期范围内误测率相差无几。考虑到预测模型的频繁更新会给预测器带来较重的负担，因此需要在可容忍的误测率范围内选择尽可能长的更新周期。

4.3 试验 2：测试 FPM 的数据处理能力

本试验的方法是启动整个系统（包括预测器），根据预测器的输入量和输出量，计算丢包率衡量 FPM 的吞吐量。实际从监测器出来的流量大约为 100M/s，即预测器的输入量为 100M/s，输出量为 90M/s，即丢包率约为 10%，而流向分析器的流量仅为 20M/s，分析器的输出量也大约为 20M/s，而在缺少 FPM 的情况下，分析器的丢包率高达 70%，实际的吞吐量只有约 30M/s。这说明

（1）预测器为分析器分担了大部分的数据量，使分析器的丢包率大大降低。

（2）预测器的吞吐量约为 90M/s，大大高于分析器的 30M/s。

五．结论

对入侵行为的预测根据时间粒度可以分为长期和短期两种，本文所研究的内容属于后者，其手段是通过建立入侵行为的预测/趋势模型，从而根据当前用户行为的初始特征判断它是否入侵行为，从而能做到预测，同时数据处理速度也大大加快。而一般的滥用检测系统对入侵的判断需要复杂而完整的证据，既不能保证检测速度，也不能保证对入侵的预测。本文介绍了一种预测机制——反馈预测机制，FPM 作为入侵检测的先锋分担其部分工作量，还能从入侵检测系统的成果中挖掘出特殊的知识反馈给预测器，以提高预测的准确性。因此，将 FPM 加入入侵检测系统中，将能提高入侵检测系统的数据处理能力和实现预测。

在试验中可以看出，随着预测模型更新周期的延长，入侵预测的准确度迅速下降，这说明了 FPM 不能做到长期预测，更新周期的确定应综合考虑到效率和准确度两方面因素。预测模型测度的选择是本方法的另一关键，例如在试验中，如果选源/宿地址作为自变量，则预测模型的准确度太低，因为这些变量的值分布比较分散，规律性不强。

【参考文献】

1．Steven R. Snapp and Stephen E. Smaba；*Signature Analysis Model Definition and Formalism*； In Proceedings of the Fourth Workshop on Computer Security Incident Handling,Denver,Colorado,August 1992

2．W Lee，and S J Stolfo. *Data mining approaches for intrusion detection.* In Proceedings of the 7th USENIX Security Symposium，1998，21(3)：181～199.

3. W. Lee, S. Stolfo, P. Chan, E. Eskin, W. Fan, M. Miller, S. Hershkop, and J. Zhang. *Real time data mining-based intrusion detection.* Proc. Second DARPA Information Survivability Conference and Exposition, pp. I85-100, 2001.

4. Wenke Lee, Wei Fan, Matthew Miller, Sal Stolfo and Erez Zadok；*Toward Cost-Sensitive Modeling for Intrusion Detection and Response*；Journal of Computer Security, 2002, 10, 1: 318-336.

5．[Cohen, 1994] W. W. Cohen. *Fast effective rule induction.* In Machine Learing:the 12th International Conference, Lake Taho, CA, 1995. Morgan Kaufmann.