基于完全无向图的贝叶斯分类器在入侵检测中的应用*)

焦从信　王崇骏　陈世福
(南京大学计算机软件新技术国家重点实验室　南京 210093)
(南京大学计算机科学与技术系　南京 210003)

摘　要　朴素贝叶斯分类器由于其强独立性假设，并不考虑属性之间的相互关系，而入侵检测的数据集不能很好地满足这一条件假设。为此，提出了一种基于有向完全图的贝叶斯分类器，将属性之间的关系加入到分类器的构造中，降低了朴素贝叶斯分类器的强独立性假设，并将其应用于入侵检测中。在 MIT 入侵检测数据集的实验表明，该算法能提高入侵检测的准确率，其效果很好。

关键词　贝叶斯算法，分类器，入侵检测模型

Complete Undirected Graph Augmented Bayes Classifier and its Application in Intrusion Detection System

JIAO Cong-xin　WANG Chong-jun　CHEN Shi-fu
(National Key Laboratory for Novel Software Technology, Nanjing University, Nanjing 210093, China)
(Department of Computer Science and Technology, Nanjing University, Nanjing 210093, China)

Abstract　The strong independence assumption made by the naïve Bayes classifier supposes that every attribute is independent from the rest of the attributes given the state of the class variable. This assumption rarely holds true in the intrusion detection datasets. This paper models a new algorithm based on the complete undirected Graph Augmented Bayes classifier (GAB) that takes into account all influences of attributes to reduce the naïve Bayes independence assumption. We conduct experiments by using MIT intrusion detection datasets. The experimental results show that the new algorithm results in a significant improvement in detection accuracy.

Keywords　Bayes, Classifier, Intrusion detection model

1　引言

入侵检测技术是近 20 年来出现的一种新型网络安全技术，它作为防火墙后的第二道安全闸门，能够检测出多种形式的入侵行为，是现代计算机网络安全体系的一个重要组成部分。入侵检测这一概念最早是由 Anderson 在 1980 年提出的，首次提出了利用审计数据发现入侵行为的思想[1]。其后，在 1987 年，SRI(Stanford Research Institute)的 Denning 提出了一种通用的入侵检测模型，成为入侵检测发展的基础[2]。1990 年，Heberlein 首次提出了网络入侵检测的概念并开发了第一个网络入侵检测系统 NSM(Network Security Monitor)[3]。在网络技术迅速发展、网络安全问题日益突出的环境下，传统的基于主机或基于网络的入侵检测系统已经难以满足对越来越复杂的网络攻击的检测任务。将具有自主性和智能性的 Agent 与机器学习等技术引入到入侵检测系统，已经成为入侵检测系统研究的主要方向之一。例如基于贝叶斯分类方法的、基于神经网络的和基于关联规则挖掘的等入侵检测技术。

朴素贝叶斯分类器是众多贝叶斯分类器中最简单的一种，属性的条件独立性假设是其区别与其它贝叶斯分类器的根本特征，即在给定类标的情况下其他属性的取值是相互独立的。这一假设大大简化了分类器的复杂性，虽然简单，但是在很多场合取得了较其它复杂分类器更优的性能，所以朴素贝叶斯分类器已经被引用到入侵检测系统中，如斯坦福研究院的 Valdes 和 Skinner 等用朴素贝叶斯分类器对网络流量进行分析，并设计了称为 eBayes 的入侵检测系统[4]；Barbara 等在其论文中也提出了使用朴素贝叶斯分类器对事件进行分类的入侵检测系统 ADAM[5]。然而，朴素贝叶斯的独立性假设在实际情况中不可能完全满足，事件属性之间会存在着一定的依赖关系。为了缓解朴素贝叶斯的条件独立性假设，贝叶斯网络被用来刻画属性之间的相互依赖。Kruegel 等人在其系统中使用贝叶斯网络刻画各个检测模型的置信度和相互之间的依赖关系，从而降低了系统的误报率[6]。

本文在深入研究各种贝叶斯分类方法和入侵检测技术的基础上，对朴素贝叶斯分类方法进行了改进，提出了一种基于完全无向图的贝叶斯分类器(Graph Augmented Bayes, GAB)。GAB 考虑到了每个属性对的依赖关系，从而在一定程度上缓解了朴素贝叶斯分类器的条件独立性假设。由于刻画入侵事件的属性并不能够满足朴素贝叶斯的条件独立性假设，即这些属性之间是有条件依赖的，所以将 GAB 应用到入侵检测中时，可以提高检测的准确率。为了评估 GAB 的性能，本文给出了 GAB 及其它一些贝叶斯分类器在 MIT 入侵检测数据集上的实验结果。这些试验结果表示 GAB 在大样本类标上的分类性能与其它分类器相当的情况下，在小样本

*)本文得到国家自然科学基金(编号：60503021)，江苏省自然科学基金(编号：BK2005075)，江苏省高新技术计划(编号：BG2006027，BG2007038)的资助。**焦从信**　博士研究生，研究领域是机器学习与网络安全；**王崇骏**　博士，副教授，研究领域是机器学习与数据挖掘；**陈世福**　教授，博导，研究领域是人工智能与知识工程。

类标上的分类性能较其它分类器有明显的优势。

2 贝叶斯分类模型

分类的基本任务是能够正确地给出无类标实例的类标。通常，首先要对给定的已知类标的训练集进行分析，构造出一个分类器，然后使用此分类器对任何未知类别的实例进行分类。朴素贝叶斯分类器(Naive Bayes)的分类过程也是如此：首先从训练集样本中估算出每个类标的先验概率及每个类标下所有属性取值组合的似然度，然后使用贝叶斯公式(1)计算出所有类标的后验概率，并根据最小风险理论取具有最大后验概率的类标为分类结果。

$$P(C_j|A)=\frac{P(A|C_j)P(C_j)}{P(A)} \tag{1}$$

其中 A 是刻画事件的所有属性的集合，$A=\{A_1,A_2,\cdots,A_n\}$，A_i 是其中的一个属性。C_j 是事件的一个类标，$P(C_j)$为类标的先验概率，$P(A|C_j)$为似然度，$P(C_j|A)$为类标的后验概率，$P(A)$是所有属性取值组合的概率。因对于所有的类标 $P(A)$总是相同的，所以在实际分类时可以不予考虑。由于朴素贝叶斯中假设给定类标不同属性之间的取值是独立的，所以在估计似然度时相对简单，如式(2)所示。

$$P(A|C_j)=\prod_{i=1}^{n}P(A_i|C_j) \tag{2}$$

根据最小风险理论，朴素贝叶斯中所用的判别函数为式(3)：

$$v_{NB}=\underset{c}{\operatorname{argmax}}P(C)\prod_{i=1}^{n}P(A_i|C) \tag{3}$$

与其他更复杂的分类器相比，朴素贝叶斯分类器在许多数据集上表现出了出色的性能[7]。作为分类器，朴素贝叶斯方法具有实现简单、训练与分类速度快、空间复杂度低等特点[8]。从式(3)可以看出，不需要搜索，只要直接从样本中估计出先验概率和似然度。

很明显，朴素贝叶斯分类器的条件独立性假设在现实世界中是很难成立的，目前已有研究者提出了许多方法试图通过放宽朴素贝叶斯的条件独立性假设来改善朴素贝叶斯分类器的性能和适用范围。主要的思路是：采用一定方式来表示属性的依赖关系。贝叶斯网络[9]是其中的一种，它通过有向边来表示属性之间的依赖关系，如 Friedam 和 Goldsmidt 提出的 TAN(Tree Augmented Naive Bayes)方法[10]。在 TAN 中，允许每个属性节点拥有一个非类标属性作为其父节点，如图 1 所示。研究表明，在 UCI 机器学习数据集上，TAN 分类精度性能较朴素贝叶斯有很显著的改进。除了 TAN 以外，还有许多其它改进朴素贝叶斯方法的技术，例如，NB-Tree[11]、SBC(Selective Bayesian Classifier)[12]，以及近年来的 AODE[13]、L^2DLNB[16]等。

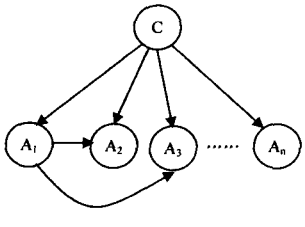

图 1 TAN 分类器的属性信度图

3 基于无向完全图的贝叶斯分类器

朴素贝叶斯分类器由于其条件独立性假设的原因，并不考虑属性之间的相互联系，而入侵检测的数据集不能很好地满足这一假设。因此，很自然地，如果把实例属性之间的依赖关系考虑进分类器的构造中的话，分类正确性应该会有一定的提高。基于这个出发点，本文提出了一种基于无向完全图的贝叶斯分类器，并将其引用到入侵检测的模型中，以提高入侵检测的准确性。

3.1 GAB算法描述

3.1.1 GAB图模型

GAB图模型是一般图模型的一种，也包括两部分：图的基本结构和表示事件属性顶点的概率分布描述。令 $A=\{A_0,A_1,A_2,\cdots,A_n\}$是离散随机变量的有限集，其中 $A_1,A_2,\cdots,A_n$ 是样本的属性变量，A_i 取值为 a_i。

GAB的基本结构是一个无向完全图。其中，每个顶点对应属性集 A 中的一个属性，每一对顶点之间都有一条边来显示地表示属性之间的联系，并使用给定类标的两属性的联合条件概率分布来定量这种依赖关系。因此，给定属性集 A 上的 GAB 图的结构可以表示为 $BG=(V,E)$。这里，V 对应属性顶点的集合 $V=\{A_0,A_1,A_2,\cdots,A_n\}$，$E$ 则是连接 V 中不同顶点的边的集合，$E=\{\{A_i,A_j\}|A_i,A_j\in V\wedge i\neq j\}$。图 2 描述了一个有 4 个顶点的贝叶斯图。

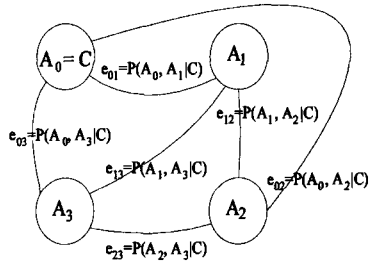

图 2 4 个顶点的贝叶斯图

GAB图模型描述的一个概率函数如下：

$$f_{BG}(A_1,A_2,\cdots,A_n,C)=P(C)\prod_{i=0}^{n-1}\prod_{j=i+1}^{n}P(A_i,A_j|C) \tag{4}$$

3.1.2 GAB分类器

令未知分类样本实例 $x=(a_1,a_2,\cdots,a_n)$，属性值 a_i 对应属性集$\{A_1,A_2,\cdots,A_n\}$中的属性 A_i。

根据上节中的贝叶斯图给出的概率分布公式(3)与公式(2)，有

$$v_{GAB}=\underset{c}{\operatorname{argmax}}(P(C)(\prod_{i=0}^{n-1}\prod_{j=i+1}^{n}P(A_i,A_j|C))) \tag{5}$$

把样本的属性值 a_i 代入，结合从训练集上获得的相应的先验概率，就可以得到该样本后验概率最大的类标属性值 c。

为了方便后面分类器的学习，由 $A_0=C$，我们可以重写式(6)为

$$\begin{aligned}v_{GAB}&=\underset{c}{\operatorname{argmax}}(P(C)(\prod_{i=1}^{n}P(A_i,A_0|C)(\prod_{i=1}^{n-1}\prod_{j=i+1}^{n}P(A_i,A_j|C)))\\&=\underset{c}{\operatorname{argmax}}(P(C)(\prod_{i=1}^{n}P(A_i|C)(\prod_{i=1}^{n-1}\prod_{j=i+1}^{n}P(A_i,A_j|C)))\end{aligned} \tag{6}$$

3.2 GAB分类器的训练和分类

GAB分类器的学习过程是简单而且直观的。

设给定一个有限的训练集 $V=\{A_0,A_1,A_2,\cdots,A_n\}$，有实例 $v=\{a_1,a_2,\cdots,a_n,c\}$，其中 a_i 对应属性 A_i 的属性值，c 是样本的类标属性值，对应 A_0。所以，在计算未知类标属性 c

极大后验概率时，将这些属性值代入式(6)得到

$$v_{GAB}=\underset{c}{\operatorname{argmax}}(P(c)(\prod_{i=1}^{n}P(a_i|c)(\prod_{i=1}^{n-1}\prod_{j=i+1}^{n}P(a_i,a_j|c))) \tag{7}$$

在这里，我们给出一个简单的训练和分类的实现。

对式(7)中的 $P(c)$，$P(a_i|c)$ 和 $P(a_i,a_j|c)$，可根据给定的样本实例 v，代入实例的属性值，然后根据训练集中类标以及其他属性出现的频率直接计算对应的概率。为了防止概率为零的情况出现，计算这些概率时我们采用了 Laplace 修正：

$$P(c)=\frac{N_k+1/N_c}{N_t+1}$$
$$P(a_i|c)=\frac{N_{ic}+1/n_i}{N_k+1} \tag{8}$$
$$P(a_i,a_j|c)=\frac{N_{ijc}+1/(n_i*n_j)}{N_k+1}$$

其中，N_t 是训练集的总样本个数，N_k 是类标属性值为 c 的样本个数，N_c 是类的个数，N_{ic} 是类 c 中属性 A_i 的属性值为 a_i 的样本个数，N_{ijc} 是类 c 中满足属性 $A_i=a_i$ 且属性 $A_j=a_j$ 的样本个数，n_i 是离散属性 A_i 的属性个数。

3.3 GAB 的计算复杂度

根据式(7)和(8)可以看出，GAB 训练和分类的过程与朴素贝叶斯十分相似。

GAB 计算的时间复杂度是 $O(n^2)$，其中计算 $P(c)$ 和 $P(a_i|c)$ 的复杂度都是 $O(n)$，加上计算 $P(a_i,a_j|c)$ 的复杂度 $O(n^2)$，所以 GAB 训练和分类一个实例的计算复杂度是 $O(n^2)$，其中 n 是该实例不同属性的个数。

4 基于 GAB 的入侵检测模型

基于 GAB 的入侵检测模型可分为两个部分：训练过程和检测过程，如图 3 所示。

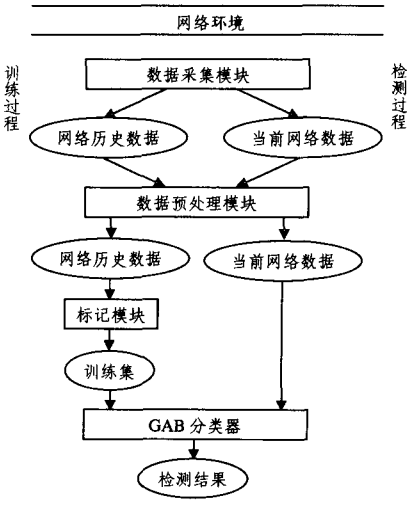

图 3 基于 GAB 的入侵检测模型

训练过程中，数据采集模块对历史网络数据包或主机系统日志等原始数据进行采集。通过预处理模块，将原始数据转化为可分类的样本实例。标记模块是在训练过程中对经过预处理的历史记录进行标记，区分正常行为和攻击记录，从而生成构造 GAB 分类器所需的训练集。

检测过程中，数据采集模块同样从系统环境中采集原始数据，经预处理后，转化为当前的待检测记录。该记录被送到 GAB 分类器后，由分类器进行分类检测，从而判断是否有入侵行为发生。

5 实验结果与分析

5.1 实验环境与说明

本实验采用的平台是 Weka[14]，实验中采用的数据集来自 KDDCUP99[15]，这些数据来自美国 MIT Lincoln 实验室通过模拟一个典型的美国空军网站并将 9 个星期所收集的原始数据加工而成的 500 万条样本数据。每个样本由 41 个特征属性和类标属性构成，其中 41 个特征属性中有 34 个连续属性、7 个离散属性，而类标属性则分为 5 大类。

1)Normal：正常行为。

2)Denial of Service Attacks(DOS)：通过大量服务请求使目标主机陷于瘫痪。

3)User to Root Attacks(U2R)：攻击者试图以一个普通用户的身份获得对系统访问的根用户权限。

4)Remote to User Attacks(R2L)：攻击者试图通过利用系统缺陷控制远程主机。

5)Probe：攻击者试图在网络上收集主机和可用服务的信息为攻击而准备。

对实验过程中的一些问题，说明如下。

1)由于 MIT 提供数据集的总量十分庞大，共 494019 条样本(其中 U2R 只有 52 个)，我们选取了全部数据集的 10% 作为实验的训练集。由于 U2R 是稀缺样本，我们把其他样本随机分为 10 份再与 U2R 样本合并作为一组训练集。这样，每组约有 49450 个样本。MIT 提供的测试集一共有 311029 条样本记录，同样随机选取 10%，共 31103 条记录作为实验的测试集。需要说明的是，MIT 提供的训练集和测试集虽然把记录都分成上述的 5 个大类，但测试集详细的记录类型从训练集的 23 种增加到了 38 种。这样，虽然 5 大类没有变化，但是通过对样本的观察可以发现，同大类中有许多差异很大的新样本出现，这就模拟了网络入侵的真实场景，特别是对异常入侵的检测。

2)对连续属性的离散化，采用 Weka 中的 Filter，每个属性离散化为 50 个属性值。

3)本实验分以下 3 种情况对 Naive Bayes，TAN[36] 和 GAB3 种算法进行了比较：

- 首先在训练集上进行 10 倍交叉验证。
- 将训练集和测试集使用同样的方法进行离散化，然后用训练集构造 3 种分类器，并使用测试集对这些分类器进行性能评估。
- 将训练集与测试集合并离散化后进行 10 倍交叉验证。

5.2 实验结果与分析

(1) 实验一

我们首先在训练集上进行 10 倍交叉验证，实验结果的比较如表 1 所示。通过对这 3 种算法的比较和分析，由于 GAB 把一个属性与其他所有属性的联合条件概率都引入到分类器的构造中，充分利用了属性之间的依赖关系，所以在本数据集的小样本分类上取得较朴素贝叶斯以及 TAN 算法较好的分类性能。

表1 三种算法在训练集上的10倍交叉验证分类结果

	Naive Bayes	TAN	GAB
Total	98.37%	99.14%	99.21%
Normal	99.5%	99.9%	99.6%
U2R	23.1%	51.9%	71.2%
DOS	98.5%	99.2%	99.2%
R2L	27.4%	44.2%	84.1%
Probe	92.9%	94.2%	95.9%

(2) 实验二

3种算法在测试集上分类结果的比较如表2所示。GAB算法虽然在总正确率和Normal的分类上略逊于TAN算法，但是和在训练集上的10倍交叉验证结果一样，GAB算法在U2R和R2L的分类上较TAN有明显的优势。和朴素贝叶斯分类器比较，除了Probe分类以外，GAB在其他分类正确率上都超过了朴素贝叶斯。

表2 三种算法在测试集上的分类结果

	Naive Bayes	TAN	GAB
Total	89.59%	91.87%	91.78%
Normal	98.3%	99.4%	98.4%
U2R	42.9%	14.3%	57.1%
DOS	93.4%	96.5%	96.5%
R2L	2.4%	2.7%	6.9%
Probe	91.8%	73.4%	77.2%

需要说明的是，3种方法在R2L上的分类效果都不佳。通过对测试集的分析以及与训练集的对比可以发现，测试集中新增加的攻击类型实例主要集中在R2L上，这对分类器的性能有较大的影响，说明方法对异常入侵的检测性能还有待提高，这一点也可以通过增加新的样本实例来进行改进。

(3)实验三

我们将训练集和测试集合并再进行10倍交叉验证，3者在R2L上的分类性能有了明显的提高，如表3所示。

表3 三种算法在合并训练集与测试集后的数据集上的分类结果

	Naive Bayes	TAN	GAB
Total	94.82%	97.45%	96.19%
Normal	85.8%	92.3%	88.2%
U2R	23.7%	44.1%	83.1%
DOS	97.3%	99.1%	98.3%
R2L	93.9%	88.4%	94.3%
Probe	87.3%	92%	96.4%

结束语 本文提出了一种基于完全无向图的贝叶斯分类方法(GAB)，并使用MIT入侵检测数据集(KDDCUP99)对GAB的性能进行评估。与朴素贝叶斯和TAN相比较，GAB由于在构造分类器时考虑到了每个属性对之间的可能联系，从而在一定程度上缓解了朴素贝叶斯的条件独立性假设，并达到了提高算法在入侵检测数据集上的分类准确性的目的。实验结果表示GAB比其它的一些贝叶斯分类器确实有一定的性能优势，特别在U2R和R2L这两个小样本的分类效果上有明显的提高。

参考文献

[1] Anderson. Computer Security Threat Monitoring and Surveillance. Technical report. 1980

[2] Denning D E. An Intrusion-Detection Model. IEEE Transaction on Software Engineer, 1987

[3] Heberlein L T, Dias G, Levitt K, et al. A Network Security Monitor//Proceedings of 1990 Symposium on Research in Security and Privacy. 1990

[4] Valdes A, Skinner K. Adaptive, Model-based Monitoring for Cyber Attack Detection//Proceedings of RAID 2000. 2000

[5] Barbara D, Wu N, Jajodia S. Detecting Novel Network Intrusions Using Bayes Estimators//Proceedings of the First SIAM International Conference on Data Mining. 2001

[6] Kruegel C, Mutz D, Roberston W, et al. Bayesian Event Classification for Intrusion Detection//Proceedings of the 19th Annual Computer Security Applications Conference. 2003

[7] Dougherty J, Kohavi R, Sahami M. Supervised and unsupervised discritization of continuous features//Proceedings of the 12th Int. Conference on Machine Learning. 1995

[8] Domingos P, Pazzani M. On the optimality of the simple Bayesian classifier under zero-one loss. Machine Learning, 1997

[9] Pearl J. Probabilistic Reasoning in Intelligent Systems. Morgan Kaufmann, 1988

[10] Friedman N, Goldszmidt M. Building classifiers using Bayesian networks//Proceedings of National Conference on Artificial Intelligence. 1996

[11] Kohavi R. Scaling upthe Accuracy of Naive Bayes Classifiers: A Decision-tree Hybrid//Proceedings of the Second International Conference on Knowledge Discovery and Data Mining. 1996

[12] Langley P, Sage S. Induction of Selective Bayesian Classifiers//Proceedings of the Tenth Conference on Uncertainly in Artificial Intelligence. 1994

[13] Webb G I, Boughton J, Wang Z. Not so Naive Bayes: Aggregating One-dependence Estimators. Machine Learning, 2005

[14] Weka. http://www.cs.waikato.ac.nz/ml/weka/s

[15] KDDCup1999 Data. http://kdd.ics.nci.edu/databases/kddcup99/kddcup99.html, 1999

[16] 孙江文，王崇骏，王珺，等. L^2DLNB：懒惰学习双层朴素贝叶斯分类器. 计算机科学，2007，34(1)：136-139