doi:10.3969/j.issn.1673-9833.2015.02.013

基于灰色理论的网络安全态势预测方法

邓勇杰，文志诚，姜旭炜

（湖南工业大学 计算机与通信学院，湖南 株洲 412007）

摘 要：为了有效地预测网络安全态势，在态势因子和灰色理论的基础上，提出了将灰色GM(1, 1)和GM(1, *N*)模型相结合来预测网络安全态势的方法。首先筛选态势因子，再利用模型GM(1, 1)对态势因子的变化进行预测，得到*N*个态势因子变化函数，最后利用这些函数和模型GM(1, *N*)对网络安全态势进行预测。将灰色GM(1, 1)模型、神经网络模型和本文方法对网络安全态势进行预测，实验结果表明，本方法能够更准确地预测网络安全态势。

关键词：灰色理论；灰色模型；网络安全态势预测

中图分类号：TP393.08 **文献标志码**：A **文章编号**：1673-9833(2015)02-0069-05

Prediction Method of Network Security Situation Based on Grey Theory

Deng Yongjie，Wen Zhicheng，Jiang Xuwei

（School of Computer and Communication，Hunan University of Technology，Zhuzhou Hunan 412007，China）

Abstract：In order to forecast network security situation effectively, puts forward a forecast method which combines GM(1,1) and GM(1, *N*) model on the basis of situation factors and grey theory. First filters situation factors, then applies model GM(1,1) to forecast variation of the situation factors, and obtains *N* functions of situation factors variation, finally uses the functions and grey model GM(1, *N*) to forecast network security situation. The grey model GM(1,1), neural network model and the proposed method are used to forecasts network security situation, and the experimental results prove that the proposed method forecasts more accurately the network security situation.

Keywords：grey theory；grey model；network security situation prediction

0 引言

随着网络信息技术的发展，互联网成为了社会的基础信息设施。与此同时，网络攻击和破坏行为日益普遍，传统的网络安全防护设备如防火墙、入侵检测系统（intrusion detection systems，IDS）等存在着各自为战、功能单一等问题，不能全方位地对网络的安全状态做出整体的评价和估计[1]。如何及时准确、全面地掌握网络安全状况，针对网络安全的整体情况准确地做出预测[2]，发布预警和制定应对的策略来保障网络的健康运行成为影响社会经济发展的重要因素。

针对网络安全态势感知的研究就是在这种背景下产生的。网络安全态势感知最早由T. Bass[3]引入网络领域，并迅速成为网络安全领域的一个研究热点。M. R. Endsley于1995年提出的Endsley层次模型

收稿日期：2015-01-25

作者简介：邓勇杰（1986-），男，湖南邵阳人，湖南工业大学硕士生，主要研究方向为网络安全态势感知，
E-mail：1240858496@qq.com

通信作者：文志诚（1972-），男，湖南安东人，湖南工业大学副教授，博士，主要研究方向为软件工程，网络安全，
E-mail：7284353@qq.com

将态势感知模型分为3个部分：态势要素获取，态势评估，态势预测[4]。其中，态势预测作为态势感知的一个重要部分，它能够对网络系统整体运行的安全状况及未来趋势进行把握，实时地感知网络所面临的威胁，并为及时、准确的决策提供可靠依据，使由网络不安全带来的风险和损失降低到最低限度。

在网络安全态势预测之前，需先获取态势要素、态势评估来得出态势；然后通过分析态势序列的前N个时刻的态势值，来对未来的M个态势值进行预测[5]。具体分为2步：1）收集态势数据；2）利用数学模型和方法挖掘态势数据本身的变化规律，预测出未来某个阶段的网络安全态势。

传统的网络安全态势预测方法大部分采用单一形式，方法本身过于复杂，不适合复杂多变的网络环境。最大的缺陷就是仅针对网络安全态势本身的数据进行分析，利用数据挖掘技术挖掘安全态势数据序列的变化规律来达到预测目的，没有对网络安全态势因子进行深入研究，存在预测精度不高、参数确定困难等问题。这些缺陷迫切需要研究者们改变研究思路，从其它角度入手，更有效地去预测网络安全态势。为了准确预测网络安全态势，本文提出了一种基于灰色理论的网络安全态势预测方法。该方法首先从影响网络安全态势的因子入手，运用灰色模型中的数学方法处理网络安全态势因子的历史数据和当前数据序列，挖掘出态势因子的变化规律，然后分析态势因子与网络安全态势之间的联系，最后得到基于态势因子的网络安全态势变化函数，预测网络安全态势。

1 相关工作

针对网络安全态势预测，国内外研究学者提出了多种多方位、多层面的预测方法和模型，主要有：灰色GM(1, 1)模型[6]、贝叶斯网络[7]、支持向量机（support vector machine，SVM）[8]、神经网络[9]以及各种复合模型。

灰色GM(1, 1)模型从随时间变化的态势数列中挖掘有关信息。优点是所需数据量少，弱化数列随机性，挖掘其潜在规律。缺点是精度不高，仅考虑单一情形，对态势本身的影响因子不予考虑，它只能做一个整体大概的预测。

人工智能领域的BP神经网络[10]具有对混沌、非线性数据分析处理能力，能完成极为复杂的模式抽取及趋势分析，非常适合解决预测的问题。不足之处是需要大量样本训练神经网络，模型中的关键参数获取困难。

贝叶斯网络是一种以概率推导事件先后关系的网络。它的态势预测要在已知态势基础上，分析当前态势与下一个态势的概率关系。这种完全以概率为推导的预测方法的预测结果随机性太大，预测准确度不高。

支持向量机具有收敛速度快、抗过拟合能力强等优点。但单独使用SVM做预测模型也存在SVM训练过程参数选取盲目性的问题。

各种复合方法是对已知预测方法的优化改进，以减少预测时间，提高准确度为目标。但复合方法本身就增加了工作量，改进的性能也受模型本身缺陷的限制。

灰色理论GM(1, N)模型是基于累加生成的数列建立一阶N个变量的微分方程模型。其适合于建立各变量的动态关联分析模型，解释性较其它模型有很大提高。

综上所述，针对传统网络安全态势预测方法存在的缺陷，本文提出将灰色GM(1, 1)与GM(1, N)模型相结合的方法来预测网络安全态势。

2 态势因子筛选

2.1 态势因子指标

在互联网实际应用中，网络安全态势的因子是从入侵检测日志、防火墙日志、病毒扫描系统、主机设备运行状态、节点流量监控设备、实时告警系统[11]上获得的多源异质观测数据。本文依据网络基础运行性、网络脆弱性、网络威胁性[12]，分析了以下态势因子：与基础运行相关的CPU使用率、内存占用量、流量状态、子网带宽占用率、数据丢包率等；与网络威胁性相关的病毒攻击频率、病毒数目、攻击危害度等；与网络脆弱性相关的网络安全设备数目、关键设备漏洞数目等。

2.2 因子筛选

影响网络安全态势的因子较多，本课题组只选取对网络安全态势有主要影响的因子。利用灰色理论中的灰色关联度分析法，分析因子与网络安全态势的关联度。

灰色关联系数为

$$\varepsilon_i(K)=\frac{\min\limits_i\min\limits_N\left|X_0^{(0)}(K)-X_i^{(0)}(K)\right|+\rho\max\limits_i\max\limits_N\left|X_0^{(0)}(K)-X_i^{(0)}(K)\right|}{\left|X_0^{(0)}(K)-X_i^{(0)}(K)\right|+\rho\max\limits_i\max\limits_N\left|X_0^{(0)}(K)-X_i^{(0)}(K)\right|}。$$

式中：$\rho\in[0,\infty)$为分辨系数；$X_0^{(0)},X_i^{(0)}$分别为随时刻 K 变化的参考数例、比较数例。

在网络安全态势感知中，$X_0^{(0)}(K)$为态势数例，$X_i^{(0)}(K)$为态势因子数例，它们在不同时刻 K 的取值不同。将其代入式（1）中，计算关联度，即

$$\rho_{0,i}=\frac{1}{L}\sum_{K=1}^{L}\varepsilon_i(K)。\qquad(1)$$

当关联度$\rho_{0,i}$大于阈值 R=0.50 时，因子被选用；否则，因子被剔除。

3　基于灰色理论的态势预测

灰色理论是由我国学者邓聚龙[13]于 1982 年提出的。之后，其相关研究工作迅速展开，现已经成为一个完备的理论体系。其中，灰色预测方法是根据过去及现在已知的或非确知的信息，建立一个从过去引申到将来的灰色 GM 模型，确定系统在未来的变化趋势，为规划决策提供依据。

3.1　灰色 GM（1, 1）模型

灰色预测模型中最基本的是 GM(1, 1)模型。其具体过程如下。设有 m 个随时间 K(K=1, 2, …, L)变化的数例为

$$X_i^{(0)}=\left(X_i^{(0)}(1),X_i^{(0)}(2),\cdots,X_i^{(0)}(L)\right),\ i=1,2,\cdots,m。$$

首先对数列进行一次累加生成，得

$$X_i^{(1)}=\left(X_i^{(0)}(1),\sum_{j=1}^{2}X_i^{(0)}(j),\cdots,\sum_{j=1}^{L}X_i^{(0)}(j)\right)=$$

$$\left(X_i^{(1)}(1),X_i^{(1)}(2),\cdots,X_i^{(1)}(L)\right)。$$

如将数列$X_i^{(1)}$ 的时刻 K 看成连续的时间变量 t，则可将数列$X_i^{(1)}$ 看成时间 t 的连续函数$X_i^{(1)}(t)$，建立一阶微分拟合方程

$$\frac{\mathrm{d}X_i^{(1)}}{\mathrm{d}t}+aX_i^{(1)}=b,\qquad(2)$$

式中发展系数 a 和常数项 b 通过最小二乘法拟合得到。令$\boldsymbol{\alpha}$=(a, b)t，按最小二乘法得$\boldsymbol{\alpha}=(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B})^{-1}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{Y}^{1}$，其中 $\boldsymbol{B}$，$\boldsymbol{Y}^{1}$ 可以从初始参数得到。求出式（2）的解后，再对其累减还原得到的随时间 K 变化的函数

$$X_i^{(0)}(K+1)=X_i^{(1)}(K+1)-X_i^{(1)}(K)=$$

$$\left(X_i^{(0)}(1)-\frac{b}{a}\right)\left(\mathrm{e}^{-bK}-\mathrm{e}^{-b\cdot1}\right)。\qquad(3)$$

利用灰色 GM(1, 1)模型对筛选好的态势因子数例进行预测，可得到 m 个态势因子变化函数。

3.2　灰色GM(1, N)模型

灰色模型中预测精度较高的是GM(1, N)模型。同样按照 GM(1, 1)模型对数例进行累加生成处理，可以建立含 N 个因子数例的白化式微分方程，即

$$\frac{\mathrm{d}X_1^{(1)}}{\mathrm{d}t}+aX_1^{(1)}=b_1X_2^{(1)}+b_2X_3^{(1)}+\cdots+b_{N-1}X_N^{(1)}。\qquad(4)$$

式（4）是一个含 N 个变量的微分方程模型 GM(1, N)。方程的参数例为$\boldsymbol{\alpha}$=(a, b_1, b_2, …,b_{N-1})，设$\boldsymbol{Y}^N$=($X_i^{(0)}$(2), $X_i^{(0)}(3),\cdots,X_i^{(0)}(K)$)$^{\mathrm{T}}$，将式（4）按差分法离散：若要确定系数$\boldsymbol{\alpha}$，必须先求出 $\boldsymbol{B}$，令残差 $\boldsymbol{e}=\boldsymbol{Y}^N-\boldsymbol{B}\boldsymbol{\alpha}$，按右端对左端进行预报，则要求残差 $\boldsymbol{e}$ 的平方和 $\|\boldsymbol{e}\|^2=\boldsymbol{e}^{\mathrm{T}}\boldsymbol{e}=\|\boldsymbol{Y}^N-\boldsymbol{B}\boldsymbol{\alpha}\|^2$ 最小（用最小二乘法得到），其中$\boldsymbol{\alpha}$必须满足$\boldsymbol{\alpha}=(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B})^{-1}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{Y}^N$，因此，利用 2 点滑动平均思想得到矩阵 $\boldsymbol{B}$ 后，即可求得$\boldsymbol{\alpha}$。式（4）的解为

$$X_1^{(1)}(t)=\mathrm{e}^{-at}\left[\sum_{i=2}^{N}\int b_{i-1}X_i^{(1)}(t)\mathrm{e}^{at}\mathrm{d}t+X_1^{(1)}(0)-\right.$$

$$\left.\sum_{i=2}^{N}\int b_{i-1}X_i^{(0)}(t)\mathrm{d}t\right]。\qquad(5)$$

当$X_1^{(1)}(t)$, $X_i^{(0)}(t)$变化幅度较小，为灰常量时，令 t=K+1，对式（5）累减还原，得

$$X_1^{(0)}(K+1)=X_1^{(1)}(K+1)-X_1^{(1)}(K)。$$

将由 GM(1,1)模型求得的 N 个态势因子变化函数代入方程，即可得到含 N 个态势因子的网络安全态势变化函数。

3.3　精度检验

衡量一个模型或方法好坏的标准就是对预测结果进行精度检验，只有满足精度要求的模型或方法才能被选择和使用。检验方法有残差、相对误差检验等。

根据实际态势值序列 $X_0^{(0)}(K)$和模型得到的态势值预测序列 $Y_0^{(0)}(K)$，求残差序列，即

$$\varepsilon(K)=X_0^{(0)}(K)-Y_0^{(0)}(K)。$$

平均残差为

$$\bar{\varepsilon}=\frac{1}{L}\left|\sum_{K=1}^{L}\varepsilon(K)\right|。\qquad(6)$$

定义相对误差序列为

$$\Delta_K=\left|\frac{\mu(K)}{X_0^{(0)}(K)}\right|(K=1,2,\cdots,L),$$

平均相对误差为

$$\bar{\Delta}=\frac{1}{L}\sum_{K=1}^{L}\Delta_K。\qquad(7)$$

4　实验分析

为了验证本文提出的基于灰色理论的网络安全态势预测方法的合理性，本课题组构建了一个特定的网络，采用 MATLAB 7.0 平台进行仿真实验[14−15]。实验数据主要是通过观测网络中的 IDS，Netflow，

Firewall日志等获得。实验前，对所有数据进行了无量纲归一化处理，处理后的数据的取值范围为[0, 1]。再将各类恶意攻击流量注入到本文特定的正常网络中，从各网络节点汇聚获得异常数据。

4.1 因子筛选结果

通过灰色关联度分析计算，筛选出来以下8个影响网络安全态势的主要指标：与基础运行相关的CPU使用率、内存占用量、流量状态；与网络威胁性相关的病毒攻击频率、病毒数目、告警数目；与网络脆弱性相关的网络安全设备数目、关键设备漏洞数目。

4.2 预测过程

为了验证本方法的优越性，本文分别将灰色GM(1, 1)模型、神经网络模型和本模型对网络安全态势进行预测，并比对其预测结果。

神经网络模型。神经网络模型采用RBF(radical basis function)神经网络，参数选取如下：1）神经网络输入向量$\boldsymbol{X}$的维数N=9，即以9天为周期单位对网络进行预测；2）神经网络输出向量$\boldsymbol{Y}$的维数M=1，即根据神经网络历史9天的输入向量，预测未来1天态势值；3）训练样本数量K=100。表1为9天的网络安全态势。

表1 3种模型的预测结果表

Table 1 Predicted results of three models

时间/天	实际值	预测值		
		灰色GM(1, 1) 模型	神经网络模型	本模型
1	0.32	0.39	0.21	0.30
2	0.44	0.44	0.31	0.37
3	0.21	0.44	0.34	0.27
4	0.30	0.45	0.34	0.27
5	0.42	0.47	0.56	0.37
6	0.63	0.51	0.72	0.58
7	0.50	0.53	0.60	0.49
8	0.71	0.60	0.60	0.66
9	0.80	0.64	0.70	0.72

将表1中3种模型的预测值进行残差、相对误差检验，检验其准确性。精度检验结果见表2。

表2 精度检验结果表

Table 2 Accuracy test results

方法	平均残差	平均相对误差
GM(1, 1)	0.127	0.282
神经网络	0.103	0.250
本方法	0.050	0.115

4.3 结果分析

从3种模型的网络安全态势预测结果可以看出，当恶意攻击流量注入正常的网络流量中后，预测结果均显示网络安全态势的总体趋势在逐渐上升，网络安全状况有逐渐变差的趋势，需要即时采取相应安全措施。

3种模型的预测结果误差和适用范围各有不同。灰色GM(1, 1)模型预测结果的平均残差和平均相对误差较大，主要反映了安全态势的总体平滑趋势。虽然神经网络模型的预测结果误差相对于灰色GM(1, 1)模型要小些，但是其算法收敛速度慢，尤其在网络结构复杂、时间复杂度过高的情况下，会失去其实际价值。本方法综合考虑了影响安全态势的因子，预测结果的平均残差和平均相对误差较灰色GM(1, 1)模型和神经网络模型要小很多，准确地反映出网络安全态势变化趋势，更能适应于复杂多变的网络环境。

5 结语

面对日益复杂的网络安全形势，如何及时准确地为网络管理员提供网络安全态势的信息成为一个研究热点。传统的态势预测方法只是片面单一地从网络安全态势数据本身进行分析，存在误差较大、或者训练数据样本过大、参数确定困难、适应范围有限等问题。本文针对影响网络安全态势的因子，将灰色理论运用在网络安全态势预测中，该方法较好地克服了传统方法的不足，能更准确地预测网络安全态势，为保障网络的健康运行提供了一种可行的办法。

参考文献：

[1] 国家计算机网络应急技术处理协调中心. 2011年中国互联网网络安全态势报告[J]. 信息网络安全，2012，10(4)：98-100.
CNCERT/CC. 2011 Chinese Internet Network Security Situation Report[J]. Netinfo Security，2012，10(4)：98-100.

[2] 王 庚，张景辉，吴 娜. 网络安全态势预测方法的应用研究[J]. 计算机仿真，2012，29(2)：98-101.
Wang Geng，Zhang Jinghui，Wu Na. Application Research on Network Security Situation Prediction Method[J]. Computer Simulation，2012，29(2)：98-101.

[3] Bass T. Intrusion Detection Systems and Multisensor Data Fusion：Creating Cyberspace Situational Awareness[J]. Communications of the ACM，2000，43(4)：99-105.

[4] Snyder L. Formal Models of Capability-Based Protection Systems[J]. IEEE Transactions on Computers，1981，30(3)：172-181.

[5] 刘 鹏，孟 炎，吴艳艳. 大规模网络安全态势感知及预测[J]. 计算机安全，2013 (2)：28-35.
Liu Peng，Meng Yan，Wu Yanyan. Large-Scale Network Security Situation Awareness and Forecast[J]. Journal of Computer Security，2013(2)：28-35.

[6] 邓聚龙. 灰理论基础[M]. 武汉：华中科技大学出版社，2003：5-10.
Deng Julong. Gray Theory[M]. Wuhan：Huazhong University of Science & Technology Press Co., Ltd.，2003：5-10.

[7] Cooper G F. The Computational Complexity of Probabilistic Inference Using Bayesian Belief Networks[J]. Artificial Intelligence，1990，42 (2)：393-405.

[8] 汪材印. 灰色关联分析和支持向量机相融合的网络安全态势评估[J]. 计算机应用研究，2013，30(6)：1859-1862.
Wang Caiyin. Assessment of Network Security Situation Based on Grey Relational Analysis and Support Vector Machine[J]. Application Research of Computer，2013，30 (6)：1859-1862.

[9] 谢丽霞，王亚超，于巾博. 基于神经网络的网络安全态势感知[J]. 清华大学学报：自然科学版，2013，53(12)：1750-1760.
Xie Lixia，Wang Yachao，Yu Jinbo. Network Security Situation Awareness Based on Neural Network[J]. Journal of Tsinghua University：Science and Technology，2013，53(12)：1750-1760.

[10] 唐成华，余顺争. 一种基于似然BP的网络安全态势预测方法[J]. 计算机科学，2009，36(11)：97-101.
Tang Chenghua，Yu Shunzheng. Method of Network Security Situation Prediction Based on Likelihood BP[J]. Computer Science，2009，36(11)：97-101.

[11] 黄同庆，庄 毅. 一种实时网络安全态势预测方法[J]. 小型微型计算机系统，2014，35(2)：303-306.
Huang Tongqing，Zhuang Yi. An Approach to Real-Time Network Security Situation Prediction[J]. Journal of Chinese Computer Systems，2014，35(2)：303-306.

[12] 贾 焰，王晓伟，韩伟红，等. YHSSAS：面向大规模网络的安全态势感知系统[J]. 计算机科学，2011，38(2)：4-8.
Jia Yan，Wang Xiaowei，Han Weihong，et al. YHSSAS：Large-Scale Network Oriented Security Situational Awareness System[J]. Computer Science，2011，38(2)：4-8.

[13] 邓聚龙. 灰色预测与决策[M]. 武汉：华中科技大学出版社，1986：12-25.
Deng Julong. Gray Forecast and Decision Making[M]. Wuhan：Huazhong University of Science & Technology Press Co., Ltd.，1986：12-25.

[14] 叶健健，文志诚，吴欣欣. 基于贝叶斯方法的网络安全态势感知模型[J]. 湖南工业大学学报，2014，28(3)：65-70.
Ye Jianjian，Wen Zhicheng，Wu Xinxin. The Network Security Situational Awareness Model Based on Bayesian Method[J]. Journal of Hunan University of Technology，2014，28(3)：65-70.

[15] 蒲天银. 基于灰色理论的网络安全态势感知模型[D]. 长沙：湖南大学，2009.
Pu Tianyin. Probe on the Network Security Situational Awareness Model Based on the Gray Theory[D]. Changsha：Hunan University，2009.

（责任编辑：邓　彬）