一种基于Hurst指数的网络安全态势预测方法

申请号：201610871341.5
申请日：2016-09-30

申请(专利权)人　山东省计算中心(国家超级计算济南中心)
地址　250014 山东省济南市历下区科院路19号山东省计算中心
发明(设计)人　王继志　杨光　陈丽娟　杨英

主分类号　H04L12/24(2006.01)I
分类号　H04L12/24(2006.01)I　H04L29/06(2006.01)I
公开(公告)号　106411591A
公开(公告)日　2017-02-15
专利代理机构　济南诚智商标专利事务所有限公司 37105
代理人　李修杰

(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 106411591 A
(43)申请公布日 2017.02.15

(21)申请号 201610871341.5

(22)申请日 2016.09.30

(71)申请人 山东省计算中心(国家超级计算济南中心)
地址 250014 山东省济南市历下区科院路19号山东省计算中心

(72)发明人 王继志 杨光 陈丽娟 杨英

(74)专利代理机构 济南诚智商标专利事务所有限公司 37105
代理人 李修杰

(51)Int.Cl.
H04L 12/24(2006.01)
H04L 29/06(2006.01)

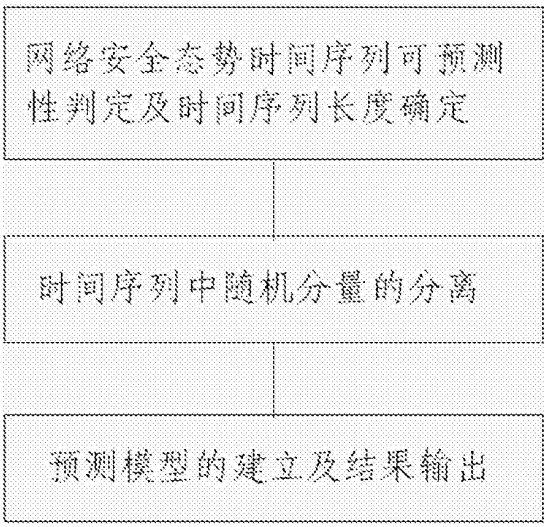
权利要求书2页 说明书6页 附图3页

(54)发明名称

一种基于Hurst指数的网络安全态势预测方法

(57)摘要

本发明公开了一种基于Hurst指数的网络安全态势预测方法,利用时间序列的自相似性指标Hurst指数来做为网络安全态势预测判定标准和优化目标,它包括以下三个过程:(1)网络安全态势时间序列可预测性判定及时间序列长度确定,(2)时间序列中随机分量的分离,(3)预测模型的建立及结果输出。本发明明确了根据实际的网络安全态势时间序列计算其是否具有可预测性,并通过计算得到最佳的用于预测的可变的时间序列长度,同时通过计算去除时间序列中无规律的随机分量,即对于预测无意义的噪声数据,在最大程度上保留时间序列中规律性的基础上,避免噪声数据的影响;并且通过预测模型计算得出预测结果和预测精度。



CN 106411591 A

1.一种基于Hurst指数的网络安全态势预测方法，其特征是，利用时间序列的自相似性指标Hurst指数来做为网络安全态势预测判定标准和优化目标，它包括以下三个过程：(1)网络安全态势时间序列可预测性判定及时间序列长度确定，(2)时间序列中随机分量的分离，(3)预测模型的建立及结果输出。

2.根据权利要求1所述的一种基于Hurst指数的网络安全态势预测方法，其特征是，所述网络安全态势时间序列可预测性判定及时间序列长度确定的过程包括以下步骤：

步骤101：设网络安全态势的时间序列为$x_1,x_2,\ldots,x_N$，设用于预测的时间序列长度用W表示，其中N和W均为正整数；

步骤102：如果$N>N_0$，则转入步骤103；否则终止计算；

步骤103：取W的初始值为P，即取时间序列$x_{N-P+1},x_{N-P+2},\ldots,x_{N-1},x_N$，共P个数值，计算其Hurst指数$H_1$，其中，$P<N_0$；

步骤104：令W的值加1，即W＝W+1，取时间序列$x_{N-W+1},x_{N-W+2},\ldots,x_{N-1},x_N$，共W个数值，计算其Hurst指数$H_2$；

步骤105：重复步骤104，直至W＝N为止，则共获得N-P+1个Hurst指数：$H_1,H_2,\ldots,H_{N-P+1}$；

步骤106：令$H_{max}=\max\{H_1,H_2,\ldots,H_{N-P+1}\}$，如果$H_{max}\leq 0.5$，则说明该网络安全态势的时间序列不可预测，终止计算；否则，设取得最大值H_{max}的相应的时间序列长度为k，则确定W＝k，即确定用时间序列$x_{N-k-1},x_{N-k-2},\ldots,x_{N-1},x_N$，共k个数值进行预测。

3.根据权利要求2所述的一种基于Hurst指数的网络安全态势预测方法，其特征是，所述时间序列中随机分量的分离过程包括以下步骤：

步骤201：将上述已选定时间序列长度为k的时间序列按顺序重新标号为$x_1,x_2,\ldots,x_W$，并将时间序列$x_1,x_2,\ldots,x_W$转换为M×K的矩阵E：

$$E=\begin{pmatrix} x_1 & x_2 & x_3 & \cdots & x_K \\ x_2 & x_3 & x_4 & \cdots & x_{K+1} \\ \vdots & \cdots & \cdots & \ddots & \vdots \\ x_M & x_{M+1} & x_{M+2} & \cdots & x_{M+K} \end{pmatrix} \qquad (1)$$

其中，$1<K<W$，$M=W-K$；

步骤202：将式(1)中矩阵E进行奇异值分解，令矩阵E的协方差矩阵$R=EE^T$，则有M个特征值和特征向量；令$\lambda_1,\lambda_2,\ldots,\lambda_M$是R的特征值，且$\lambda_1\geq\ldots\geq\lambda_M$；令$U_1,\ldots,U_M$是相应的特征向量，则$V_j=E^TU_j/\lambda_j^{1/2}$，$j=1,\ldots,M$；所以，$E_j=\lambda_j^{1/2}U_jV_j^T$，因此，$E=E_1+E_2+\ldots+E_M$；

步骤203：取i的初值为1，令$E^{(1)}=\Sigma_iE_i$；$E^{(2)}=E-E^{(1)}$，$E^{(1)}$代表时间序列中可预测的分量，$E^{(2)}$代表时间序列中不可预测的随机分量；

步骤204：按式(2)和式(3)将$E^{(1)}$，$E^{(2)}$重构为相应的时间序列$x_1^{(1)},x_2^{(1)},\ldots,x_W^{(1)}$和$x_1^{(2)},x_2^{(2)},\ldots,x_W^{(2)}$；

$$X(t)^{(1)}=\begin{cases}\dfrac{1}{t}\sum_{i=1}^{t}e^{*}_{i,t-i+2} & 1\le t<K^{*}\\ \dfrac{1}{K^{*}}\sum_{i=1}^{L^{*}}e^{*}_{i,t-i+2} & K^{*}\le t<M^{*}\\ \dfrac{1}{N-t}\sum_{i=t-M^{*}+2}^{N-M^{*}+1}e^{*}_{i,t-i+2} & M^{*}\le t<N\end{cases} \qquad (2)$$

$$X(t)^{(2)}=X(t)-X(t)^{(1)} \qquad (3)$$

式中，令$K^{*}=\min(K,M)$，$M^{*}=\max(K,M)$，e_{ij}是矩阵$E^{(1)}$的元素，则当K<M时，$e_{ij}^{*}=e_{ij}$，否则$e_{ij}^{*}=e_{ji}$；

步骤205：分别计算时间序列$x_1^{(1)},x_2^{(1)},\ldots,x_W^{(1)}$的Hurst指数$H^{(1)}$和$x_1^{(2)},x_2^{(2)},\ldots,x_W^{(2)}$的Hurst指数$H^{(2)}$；

步骤206：计算优化指标$e_i=|(1-H^{(1)})-(H^{(2)}-0.5)|$；

步骤207：令i的取值加1，即i＝i+1，重复步骤203至步骤206；并比较e_i和e_{i+1}，如果$e_i<e_{i+1}$，则终止计算，并取前i个分量作为$E^{(1)}$，及相应的时间序列$x_1^{(1)},x_2^{(1)},\ldots,x_W^{(1)}$进行网络安全态势预测，否则继续重复步骤3--6，直至i＝M为止。

4.根据权利要求3所述的一种基于Hurst指数的网络安全态势预测方法，其特征是，所述预测模型的建立及结果输出的过程包括以下步骤：

步骤301：选择参数p和q，其中，$1\le p\le[W/4]$，$1\le q\le[W/4]$；

步骤302：建立预测模型：

$$x_t'=a_1x_{t-1}+a_2x_{t-2}+\ldots+a_px_{t-p}-b_1u_{t-1}-\ldots-b_qu_{t-q} \qquad (4)$$

其中：$u_t=x_t-x_t'$，

将时间序列$x_1^{(1)},x_2^{(1)},\ldots,x_W^{(1)}$代入式(4)，并利用最小二乘法确定参数$a_1,\ldots,a_p$和$b_1,\ldots,b_q$；

步骤303：将$x_W^{(1)},x_{W-1}^{(1)},\ldots,x_{W-p+1}^{(1)}$和$u_W,u_{W-1},\ldots,u_{W-q}$代入式(4)，求得预测结果$x_{W+1}$。

5.根据权利要求4所述的一种基于Hurst指数的网络安全态势预测方法，其特征是，所述预测模型的建立及结果输出的过程还包括以下步骤：

步骤304：计算预测精度θ，预测精度θ的计算公式为：

$$\theta=\frac{1}{W}\sum_{i=1}^{W}\frac{|x_i^{(1)}-x_i|}{x_i} \qquad (5)$$

式中，θ为预测精度，$x_i^{(1)}$和x_i为时间序列。

6.根据权利要求2至5任意一项所述的一种基于Hurst指数的网络安全态势预测方法，其特征是，在网络安全态势时间序列可预测性判定及时间序列长度确定过程中，$6\le N_0\le 10$，$5\le P<N_0$，P和N_0为正整数。

一种基于Hurst指数的网络安全态势预测方法

技术领域

[0001] 本发明涉及一种基于Hurst指数的网络安全态势预测方法，属于网络信息安全技术领域。

背景技术

[0002] 随着互联网的应用普及，网络规模越来越大，也越来越复杂。相应的，网络攻击也向着分布式、规模化、复杂化的方向发展。对于网络管理人员来说，迫切需要能对网络安全整体态势进行展示的安全产品。

[0003] 所谓网络安全态势是指能够表达网络攻击行为和网络防御措施等由各种因素所构成的网络安全状态，并以数值量化的形式展示出来。由各个时刻的网络安全态势值，就构成了一个表达网络安全状态的时间序列，再基于网络安全态势的时间序列，就可以对未来时刻的网络安全态势值进行预测，以帮助网络管理人员了解网络安全态势的演化趋势，采取相应的安全措施。

[0004] 目前，已有一些涉及网络安全态势预测方法的专利，如“一种网络安全态势预测方法”(申请号：2011101052724)、“一种网络安全态势预测方法及系统”(申请号：2013105443158)、“一种误报自适应的网络安全态势预测方法”(申请号：2014107050406)和“一种基于证据理论的网络安全态势预测方法”(申请号：2015101398133)。这些专利基于不同的理论，从不同的角度提出网络安全态势的预测方法。

[0005] 然而，上述网络安全态势的预测方法都忽略了一个前提条件，即并非所有的时间序列都是可以用来预测的，如随机产生的时间序列，对其进行预测是没有意义的。因此，需要首先判定用于安全态势预测的时间序列是否具有可预测性。在此基础上，还要解决用于预测的时间序列长度问题。因为随着时间的积累，网络安全态势的时间序列会越来越长，很久之前的网络安全态势数值，从直观上来说，可能不会影响对未来时刻网络安全态势值的预测。因此，有必要确定一个合适的用于预测的时间序列长度，然而目前的方法都是主观选择一个定长的时间序列的长度，如最近一周内的网络安全态势值用于预测、最近一个月内的网络安全态势值用于预测等。同时，由于网络攻击行为具有一定的随机性，这导致网络安全态势的时间序列中也含有随机性，也就是说，网络安全态势并非完全可以预测的，对于其中的随机分量是无法进行预测的。然而现有的方法并没有考虑这一问题，试图对含有随机分量的网络安全态势值进行精确的预测。由于对随机分量的预测是无意义的，更合适的做法，是将网络安全态势时间序列中的随机分量分离出来，只对其中有规律的、可预测的分量进行预测，并确定预测精度。

发明内容

[0006] 针对上述不足，本发明提供了一种基于Hurst指数的网络安全态势预测方法，其能够根据网络安全态势的历史数据预测未来一段时间内的网络安全态势值。

[0007] 本发明解决其技术问题采取的技术方案是：一种基于Hurst指数的网络安全态势

预测方法，其特征是，利用时间序列的自相似性指标Hurst指数来做为网络安全态势预测判定标准和优化目标，它包括以下三个过程：(1)网络安全态势时间序列可预测性判定及时间序列长度确定，(2)时间序列中随机分量的分离，(3)预测模型的建立及结果输出。

[0008] 进一步地，所述网络安全态势时间序列可预测性判定及时间序列长度确定的过程包括以下步骤：

[0009] 步骤101：设网络安全态势的时间序列为$x_1, x_2, \ldots, x_N$，设用于预测的时间序列长度用W表示，其中N和W均为正整数；

[0010] 步骤102：如果$N > N_0$，则转入步骤103；否则表示网络安全态势的历史数据过少，不满足计算要求，终止计算；

[0011] 步骤103：取W的初始值为P，即取时间序列$x_{N-P+1}, x_{N-P+2}, \ldots, x_{N-1}, x_N$，共P个数值，计算其Hurst指数$H_1$，其中，$P < N_0$；

[0012] 步骤104：令W的值加1，即$W = W+1$，取时间序列$x_{N-W+1}, x_{N-W+2}, \ldots, x_{N-1}, x_N$，共W个数值，计算其Hurst指数$H_2$；

[0013] 步骤105：重复步骤104，直至$W = N$为止，则共获得N-P+1个Hurst指数：$H_1, H_2, \ldots, H_{N-P+1}$；

[0014] 步骤106：令$H_{max} = \max\{H_1, H_2, \ldots, H_{N-P+1}\}$，即取这N-P+1个Hurst指数中的最大值，如果$H_{max} \leq 0.5$，则说明该网络安全态势的时间序列不可预测，终止计算；否则，设取得最大值H_{max}的相应的时间序列长度为k，则确定$W = k$，即确定用时间序列$x_{N-k-1}, x_{N-k-2}, \ldots, x_{N-1}, x_N$，共k个数值进行预测。

[0015] 进一步地，所述时间序列中随机分量的分离过程包括以下步骤：

[0016] 步骤201：将上述已选定时间序列长度为k的时间序列按顺序重新标号为$x_1, x_2, \ldots, x_W$，并将时间序列$x_1, x_2, \ldots, x_W$转换为$M \times K$的矩阵E：

[0017] $$E=\begin{pmatrix} x_1 & x_2 & x_3 & \cdots & x_K \\ x_2 & x_3 & x_4 & \cdots & x_{K+1} \\ \vdots & \cdots & \cdots & \ddots & \vdots \\ x_M & x_{M+1} & x_{M+2} & \cdots & x_{M+K} \end{pmatrix} \quad (1)$$

[0018] 其中，$1 < K < W$，$M = W-K$；

[0019] 步骤202：将式(1)中矩阵E进行奇异值分解，令矩阵E的协方差矩阵$R = EE^T$，则有M个特征值和特征向量；令$\lambda_1, \lambda_2, \ldots, \lambda_M$是R的特征值，且$\lambda_1 \geq \ldots \geq \lambda_M$；令$U_1, \ldots, U_M$是相应的特征向量，则$V_j = E^T U_j / \lambda_j^{1/2}$，$j = 1, \ldots, M$；所以，$E_j = \lambda_j^{1/2} U_j V_j^T$，因此，$E = E_1 + E_2 + \ldots + E_M$；

[0020] 步骤203：取i的初值为1，令$E^{(1)} = \Sigma_i E_i$；$E^{(2)} = E - E^{(1)}$，$E^{(1)}$代表时间序列中可预测的分量，$E^{(2)}$代表时间序列中不可预测的随机分量；

[0021] 步骤204：按式(2)和式(3)将$E^{(1)}$，$E^{(2)}$重构为相应的时间序列$x_1^{(1)}, x_2^{(1)}, \ldots, x_W^{(1)}$和$x_1^{(2)}, x_2^{(2)}, \ldots, x_W^{(2)}$；

[0022] $$X(t)^{(1)}=\begin{cases}\frac{1}{t}\sum_{i=1}^{t}e^{*}_{i,t-i+2} & 1\le t<K^{*}\\ \frac{1}{K^{*}}\sum_{i=1}^{L^{*}}e^{*}_{i,t-i+2} & K^{*}\le t<M^{*}\\ \frac{1}{N-t}\sum_{i=t-M^{*}+2}^{N-M^{*}+1}e^{*}_{i,t-i+2} & M^{*}\le t<N\end{cases} \quad (2)$$

[0023] $X(t)^{(2)}=X(t)-X(t)^{(1)}$ (3)

[0024] 式中，令$K^{*}=\min(K,M)$，$M^{*}=\max(K,M)$，e_{ij}是矩阵$E^{(1)}$的元素，则当K<M时，$e_{ij}^{*}=e_{ij}$，否则$e_{ij}^{*}=e_{ji}$；

[0025] 步骤205：分别计算时间序列$x_1^{(1)}$，$x_2^{(1)}$，...，$x_W^{(1)}$的Hurst指数$H^{(1)}$和$x_1^{(2)}$，$x_2^{(2)}$，...，$x_W^{(2)}$的Hurst指数$H^{(2)}$；

[0026] 步骤206：计算优化指标$e_i=|(1-H^{(1)})-(H^{(2)}-0.5)|$；

[0027] 步骤207：令i的取值加1，即i＝i+1，重复步骤203至步骤206；并比较e_i和e_{i+1}，如果$e_i<e_{i+1}$，则终止计算，并取前i个分量作为$E^{(1)}$，及相应的时间序列$x_1^{(1)}$，$x_2^{(1)}$，...，$x_W^{(1)}$进行网络安全态势预测，否则继续重复步骤3--6，直至i＝M为止。

[0028] 进一步地，所述预测模型的建立及结果输出的过程包括以下步骤：

[0029] 步骤301：选择参数p和q，其中，$1\le p\le[W/4]$，$1\le q\le[W/4]$；

[0030] 步骤302：建立预测模型：

[0031] $x_t'=a_1x_{t-1}+a_2x_{t-2}+...+a_px_{t-p}-b_1u_{t-1}-...-b_qu_{t-q}$ (4)

[0032] 其中：$u_t=x_t-x_t'$，

[0033] 将时间序列$x_1^{(1)}$，$x_2^{(1)}$，...，$x_W^{(1)}$代入式(4)，并利用最小二乘法确定参数a_1，...，a_p和b_1，...，b_q；

[0034] 步骤303：将$x_W^{(1)}$，$x_{W-1}^{(1)}$，...，$x_{W-p+1}^{(1)}$和u_W，u_{W-1}，...，u_{W-q}代入式(4)，求得预测结果x_{W+1}。

[0035] 进一步地，所述预测模型的建立及结果输出的过程还包括以下步骤：

[0036] 步骤304：计算预测精度θ，预测精度θ的计算公式为：

[0037] $$\theta=\frac{1}{W}\sum_{i=1}^{W}\frac{|x_i^{(1)}-x_i|}{x_i} \quad (5)$$

[0038] 式中，θ为预测精度，$x_i^{(1)}$和x_i为时间序列。

[0039] 优选地，在网络安全态势时间序列可预测性判定及时间序列长度确定过程中，$6\le N_0\le10$，$5\le P<N_0$，P和N_0为正整数。

[0040] 本发明的有益效果是：

[0041] 本发明首先基于Hurst指数对网络安全态势时间序列的可预测性进行判定，在时间序列具有可预测性的前提下得到用于预测的时间序列长度，然后基于Hurst指数分离出其中不可预测的随机分量，只保留可预测的分量，最后根据预测模型给出最终的预测结果及预测精度。

[0042] 与现有技术相比，本发明明确了根据实际的网络安全态势时间序列计算其是否具有可预测性，而非假定其具有可预测性，并通过计算得到最佳的用于预测的可变的时间序列长度，而非凭主观经验选择一个固定的时间序列长度；同时通过计算去除时间序列中无

规律的随机分量，即对于预测无意义的噪声数据，在最大程度上保留时间序列中规律性的基础上，避免噪声数据的影响；并且通过预测模型计算得出预测结果和预测精度。

附图说明

[0043] 下面结合说明书附图对本发明进行说明。

[0044] 图1为本发明的方法流程图；

[0045] 图2为本发明所述网络安全态势时间序列可预测性判定及时间序列长度确定过程的方法流程图；

[0046] 图3为本发明所述时间序列中随机分量的分离过程的方法流程图；

[0047] 图4为本发明所述预测模型的建立及结果输出过程的方法流程图。

具体实施方式

[0048] 为能清楚说明本方案的技术特点，下面通过具体实施方式，并结合其附图，对本发明进行详细阐述。下文的公开提供了许多不同的实施例或例子用来实现本发明的不同结构。为了简化本发明的公开，下文中对特定例子的部件和设置进行描述。此外，本发明可以在不同例子中重复参考数字和/或字母。这种重复是为了简化和清楚的目的，其本身不指示所讨论各种实施例和/或设置之间的关系。应当注意，在附图中所图示的部件不一定按比例绘制。本发明省略了对公知组件和处理技术及工艺的描述以避免不必要地限制本发明。

[0049] 为了克服目前技术的不足之处，本发明提供了一种新的网络安全态势预测方法，该方法的主要思路是利用时间序列的自相似性指标Hurst指数来做为判定标准，若网络安全态势的时间序列具有自相似性，则是可以预测的；否则该时间序列是随机的，是不可预测的。同时，利用Hurst指数来作为优化目标，分别确定用于预测的时间序列的长度、时间序列中分离随机分量的标准，以及估计最终预测的精度。

[0050] 如图1所示，本发明的一种基于Hurst指数的网络安全态势预测方法，利用时间序列的自相似性指标Hurst指数来做为网络安全态势预测判定标准和优化目标，它包括以下三个过程：(1)网络安全态势时间序列可预测性判定及时间序列长度确定，(2)时间序列中随机分量的分离，(3)预测模型的建立及结果输出。

[0051] 如图2所示，以N_0取值为6，P取值为5为例，本发明所述网络安全态势时间序列可预测性判定及时间序列长度确定的过程包括以下步骤：

[0052] 步骤101：设网络安全态势的时间序列为$x_1, x_2, \ldots, x_N$，共N个，本发明预测方法的目的就是求解未来下一个时刻的网络安全态势值x_{N+1}；设用于预测的时间序列长度用W表示，其中N和W均为正整数；

[0053] 步骤102：如果$N>6$，则转入步骤103；否则表示网络安全态势的历史数据过少，不满足计算要求，终止计算；

[0054] 步骤103：取W的初始值为5，即取时间序列$x_{N-4}, x_{N-3}, x_{N-2}, x_{N-1}, x_N$，共5个数值，计算其Hurst指数$H_1$；

[0055] 步骤104：令W的值加1，即$W=W+1$，取时间序列$x_{N-W+1}, x_{N-W+2}, \ldots, x_{N-1}, x_N$，共W个数值，计算其Hurst指数$H_2$；

[0056] 步骤105：重复步骤104，直至$W=N$为止，则共获得N-4个Hurst指数：$H_1, H_2, \ldots, H_{N-4}$；

[0057] 步骤106：令$H_{max}=\max\{H_1,H_2,\ldots,H_{N-P+1}\}$，即取这N-P+1个Hurst指数中的最大值，如果$H_{max}\leq 0.5$，则说明该网络安全态势的时间序列不可预测，终止计算；否则，设取得最大值H_{max}的相应的时间序列长度为k，则确定W＝k，即确定用时间序列$x_{N-k-1},x_{N-k-2},\ldots,x_{N-1},x_N$，共k个数值进行预测。

[0058] 如图3所示，本发明所述时间序列中随机分量的分离过程包括以下步骤：

[0059] 步骤201：为描述方便，将上述已选定时间序列长度为k的时间序列按顺序重新标号为$x_1,x_2,\ldots,x_W$，并将时间序列$x_1,x_2,\ldots,x_W$转换为M×K的矩阵E：

[0060] $$E=\begin{pmatrix} x_1 & x_2 & x_3 & \cdots & x_K \\ x_2 & x_3 & x_4 & \cdots & x_{K+1} \\ \vdots & \cdots & \cdots & \ddots & \vdots \\ x_M & x_{M+1} & x_{M+2} & \cdots & x_{M+K} \end{pmatrix} \quad (1)$$

[0061] 其中，$1<K<W$，M＝W-K；

[0062] 步骤202：将式(1)中矩阵E进行奇异值分解，令矩阵E的协方差矩阵$R=EE^T$，则有M个特征值和特征向量；令$\lambda_1,\lambda_2,\ldots,\lambda_M$是R的特征值，且$\lambda_1\geq\ldots\geq\lambda_M$；令$U_1,\ldots,U_M$是相应的特征向量，则$V_j=E^TU_j/\lambda_j^{1/2}$，j＝1，...，M；所以，$E_j=\lambda_j^{1/2}U_jV_j^T$，因此，$E=E_1+E_2+\ldots+E_M$；

[0063] 步骤203：取i的初值为1，令$E^{(1)}=\Sigma_iE_i$；$E^{(2)}=E-E^{(1)}$，$E^{(1)}$代表时间序列中可预测的分量，$E^{(2)}$代表时间序列中不可预测的随机分量；

[0064] 步骤204：按照式(2)和式(3)将$E^{(1)}$，$E^{(2)}$重构为相应的时间序列$x_1^{(1)},x_2^{(1)},\ldots,x_W^{(1)}$和$x_1^{(2)},x_2^{(2)},\ldots,x_W^{(2)}$；

[0065] $$X(t)^{(1)}=\begin{cases} \dfrac{1}{t}\displaystyle\sum_{i=1}^{t}e^*_{i,t-i+2} & 1\leq t<K^* \\ \dfrac{1}{K^*}\displaystyle\sum_{i=1}^{L^*}e^*_{i,t-i+2} & K^*\leq t<M^* \\ \dfrac{1}{N-t}\displaystyle\sum_{i=t-M^*+2}^{N-M^*+1}e^*_{i,t-i+2} & M^*\leq t<N \end{cases} \quad (2)$$

[0066] $X(t)^{(2)}=X(t)-X(t)^{(1)}$ (3)

[0067] 式中，令$K^*=\min(K,M)$，$M^*=\max(K,M)$，e_{ij}是矩阵$E^{(1)}$的元素，则当K<M时，$e_{ij}^*=e_{ij}$，否则$e_{ij}^*=e_{ji}$；

[0068] 步骤205：分别计算时间序列$x_1^{(1)},x_2^{(1)},\ldots,x_W^{(1)}$的Hurst指数$H^{(1)}$和$x_1^{(2)},x_2^{(2)},\ldots,x_W^{(2)}$的Hurst指数$H^{(2)}$；

[0069] 步骤206：计算优化指标$e_i=|(1-H^{(1)})-(H^{(2)}-0.5)|$；

[0070] 步骤207：令i的取值加1，即i＝i+1，重复步骤203至步骤206；并比较e_i和e_{i+1}，如果$e_i<e_{i+1}$，则终止计算，并取前i个分量作为$E^{(1)}$，及相应的时间序列$x_1^{(1)},x_2^{(1)},\ldots,x_W^{(1)}$进行网络安全态势预测，否则继续重复步骤3--6，直至i＝M为止。

[0071] 如图4所示，本发明所述预测模型的建立及结果输出的过程包括以下步骤：

[0072] 步骤301：选择参数p和q，其中，$1\leq p\leq[W/4]$，$1\leq q\leq[W/4]$；

[0073] 步骤302：建立预测模型：

[0074] $x_t'=a_1x_{t-1}+a_2x_{t-2}+\ldots+a_px_{t-p}-b_1u_{t-1}-\ldots-b_qu_{t-q}$ (4)

[0075] 其中：$u_t = x_t - x_t'$，t为正整数；

[0076] 将时间序列$x_1^{(1)}, x_2^{(1)}, \ldots, x_W^{(1)}$代入式(4)，并利用最小二乘法确定参数$a_1, \ldots, a_p$和$b_1, \ldots, b_q$；

[0077] 步骤303：将$x_W^{(1)}, x_{W-1}^{(1)}, \ldots, x_{W-p+1}^{(1)}$和$u_W, u_{w-1}, \ldots, u_{w-q}$代入式(4)，求得预测结果$x_{W+1}$；

[0078] 步骤304：计算预测精度θ，预测精度θ的计算公式为：

[0079] $$\theta = \frac{1}{W}\sum_{i=1}^{W}\frac{|x_i^{(1)} - x_i|}{x_i} \qquad (5)$$

[0080] 式中，θ为预测精度，$x_i^{(1)}$和x_i为时间序列。

[0081] 上述实施例中，在网络安全态势时间序列可预测性判定及时间序列长度确定过程中，以N_0取值为6，P取值为5为例。为了计算准确快速，N_0的取值范围优选为[6,10]，P的取值范围优选为[6,N_0]，但本申请并不将其保护范围仅仅限定于此，对N_0和P的取值进行适当调整也可实现网络安全态势的预测，同样也应视为本申请的保护范围。

[0082] 本发明首先对网络安全态势时间序列的可预测性进行判定，然后分离出其中不可预测的随机分量，最后给出最终的预测结果，它明确了根据实际的网络安全态势时间序列计算其是否具有可预测性，而非假定其具有可预测性，并通过计算得到最佳的用于预测的可变的时间序列长度，而非凭主观经验选择一个固定的时间序列长度；同时通过计算去除时间序列中无规律的随机分量，即对于预测无意义的噪声数据，在最大程度上保留时间序列中规律性的基础上，避免噪声数据的影响；并且通过预测模型计算得出预测结果和预测精度。

[0083] 以上所述只是本发明的优选实施方式，对于本技术领域的普通技术人员来说，在不脱离本发明原理的前提下，还可以做出若干改进和润饰，这些改进和润饰也被视为本发明的保护范围。

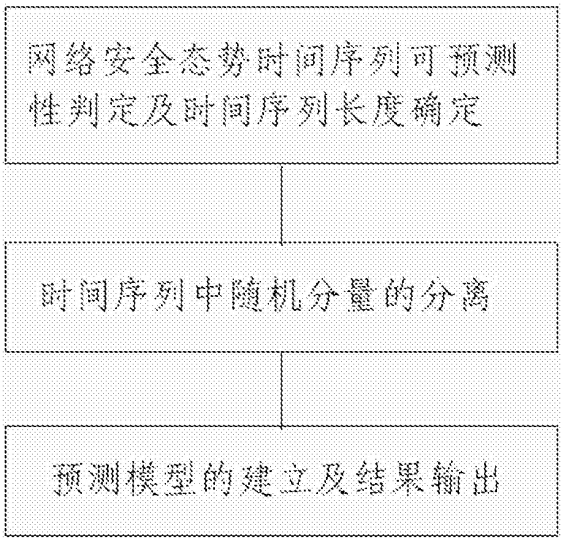


图1

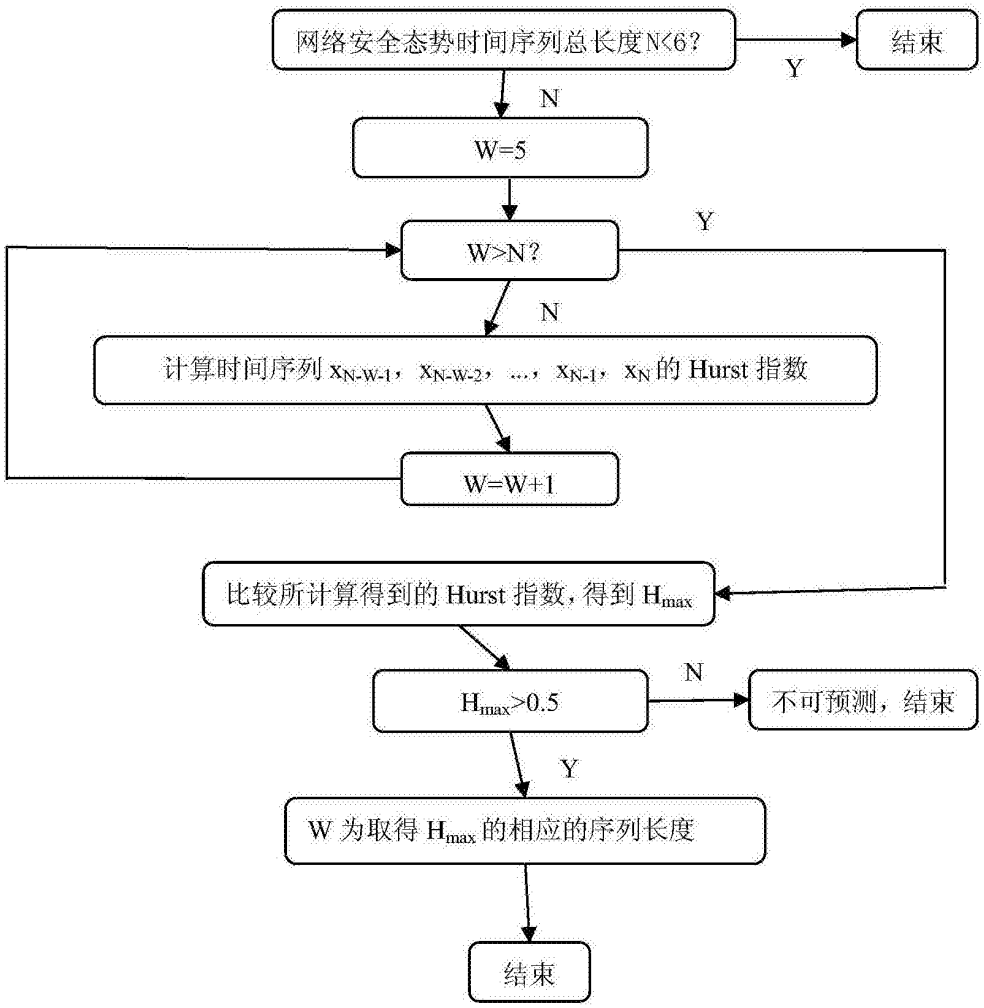

网络安全态势时间序列总长度N<6?
结束
Y
N
W=5
W>N?
Y
N
计算时间序列 x_{N-W-1}，x_{N-W-2}，...，x_{N-1}，x_N 的 Hurst 指数
W=W+1
比较所计算得到的 Hurst 指数，得到 H_{max}
H_{max}>0.5
N
不可预测，结束
Y
W 为取得 H_{max} 的相应的序列长度
结束

图2

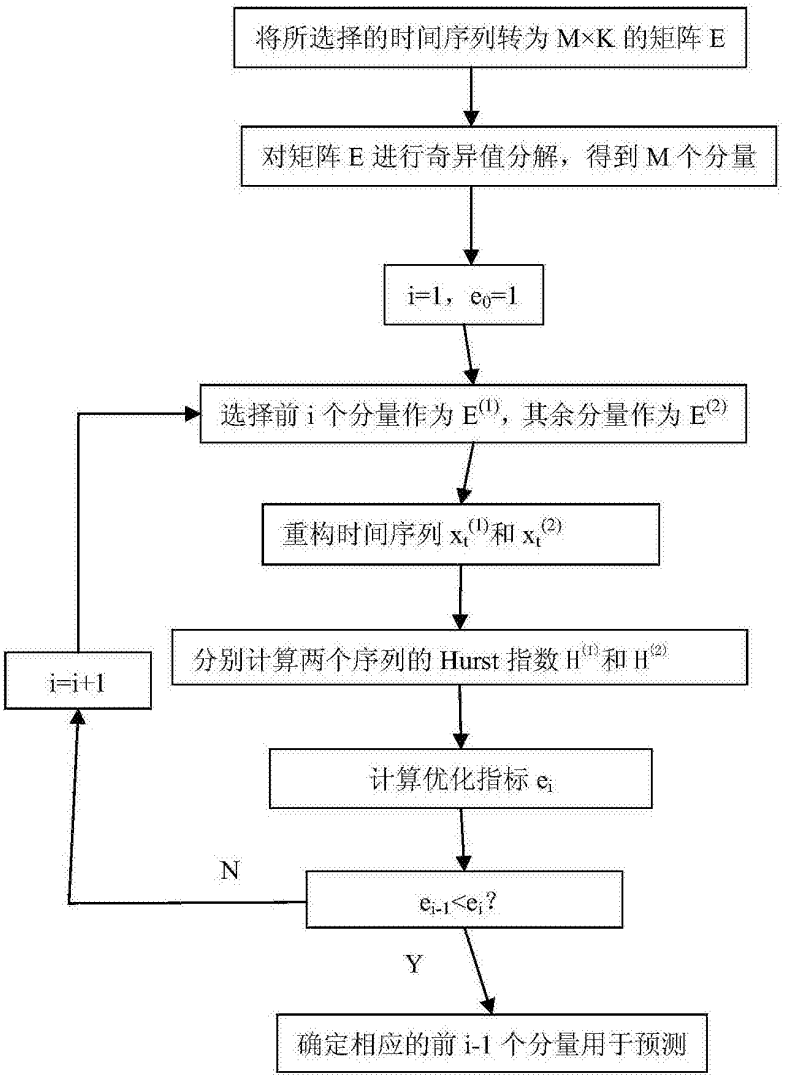

将所选择的时间序列转为 M×K 的矩阵 E
对矩阵 E 进行奇异值分解，得到 M 个分量
i=1，e_0=1
选择前 i 个分量作为 $E^{(1)}$，其余分量作为 $E^{(2)}$
重构时间序列 $x_t^{(1)}$和 $x_t^{(2)}$
分别计算两个序列的 Hurst 指数 $H^{(1)}$和 $H^{(2)}$
计算优化指标 e_i
e_{i-1}<e_i?
N
Y
i=i+1
确定相应的前 i-1 个分量用于预测

图3

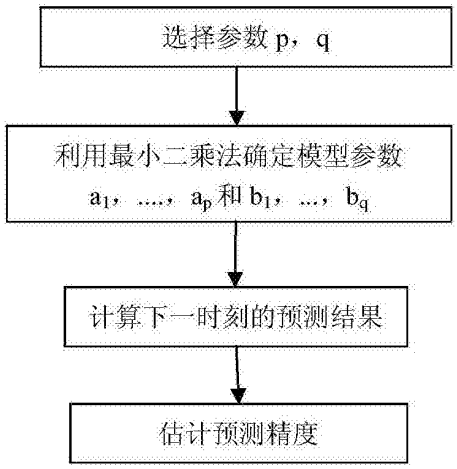

选择参数 p，q
利用最小二乘法确定模型参数
a_1，....，a_p 和 b_1，...，b_q
计算下一时刻的预测结果
估计预测精度

图4