Hindawi Publishing Corporation
The Scientific World Journal
Volume 2014, Article ID 473504, 8 pages
http://dx.doi.org/10.1155/2014/473504

Research Article

Historical Feature Pattern Extraction Based Network Attack Situation Sensing Algorithm

Yong Zeng,[1,2] Dacheng Liu,[1] and Zhou Lei[1,3]

[1] *Department of Industrial Engineering, Tsinghua University, Beijing 100084, China*
[2] *Bengbu Automobile NCO Academy, Bengbu 233011, China*
[3] *Tianjin Port (Group), Ltd., Tianjin 300456, China*

Correspondence should be addressed to Dacheng Liu; liudacv@mail.tsinghua.edu.cn

Received 20 February 2014; Accepted 18 March 2014; Published 27 April 2014

Academic Editors: Y. Mao and Z. Zhou

The situation sequence contains a series of complicated and multivariate random trends, which are very sudden, uncertain, and difficult to recognize and describe its principle by traditional algorithms. To solve the above questions, estimating parameters of super long situation sequence is essential, but very difficult, so this paper proposes a situation prediction method based on historical feature pattern extraction (HFPE). First, HFPE algorithm seeks similar indications from the history situation sequence recorded and weighs the link intensity between occurred indication and subsequent effect. Then it calculates the probability that a certain effect reappears according to the current indication and makes a prediction after weighting. Meanwhile, HFPE method gives an evolution algorithm to derive the prediction deviation from the views of pattern and accuracy. This algorithm can continuously promote the adaptability of HFPE through gradual fine-tuning. The method preserves the rules in sequence at its best, does not need data preprocessing, and can track and adapt to the variation of situation sequence continuously.

1. Introduction

With attacks becoming more prevalent, the traditional static passive defense and whole system consolidation are hard to keep up with the changing rhythms, which have huge amounts of investment and affect the network performance. In this case, the dynamic, proactive, and targeted defending measures have been presented, most of which rely on attack situation forecast, that is, network attack situation sensing (NASS) [1, 2]. NASS aims at forecasting future evolution trend of network attack situation based on historical features and current attack indications, guiding dynamic defense, and allowing administrators to take corresponding measures in advanced, and effective manner to quickly respond to the complex and ever-changing attack threats [3, 4].

Rarely studying attack situation forecast, previous researches mostly using existing methods, such as autoregressive moving average model (ARMA), grey model (GM), and radial basis function neural network (RBFNN) [5–8]. ARMA identifies the dependence relationship and autocorrelation of situation sequences and establish mathematical prediction model [9]. It requests that situation sequences or their certain step difference satisfies the steady suppose, which is too strict to increase suitable scope. As one of GMs, GM(1,1) firstly weakens the randomness of situation sequences by using accumulation, secondly fits the born sequence through index curve, and then does regressive restitution after prediction, which can embody monotonously and slowly changing trend but hardly reflect some characteristics such as random rove and periodic fluctuation [10, 11]. Grey Verhulst is suitable to describe the situation sequences with swing development according to "S" or anti-"S" form [12], and the method dividing the changing line into several stages does not lack rationality, but the difficulty is how to predict the occurrence moment and lasting time of each stage [13]. RBFNN utilizes the nonlinear characteristic to describe the regulation contained in situation sequences [14]. However, evolving regulation of attack situation is infinite and changeable; a practical type neural network with small scale cannot solve well [15, 16].

Situation sequence contains massive complex and inconstant evolution trends, beyond the expression and prediction

capability of traditional methods only by some formulas, functions or via some training [17, 18]. Most traditional methods suffer from the confliction among training samples, rely on data preprocessing and artificial intervention heavily, do not support incremental training, and need to rebuild model once situation sequence changes [19–21]. Therefore, a situation prediction method based on historical feature pattern extraction (HFPE) is presented. The method measures the similarity between historical feature from the aspects of pattern and accuracy and utilizes multiple order difference operation to discriminate trends. It searches similar indications from recorded historical situation sequence, measures the link intensity of occurred indication upon subsequent effect, and infers the recurrence possibilities of some effects according to current indication. An evolution algorithm is introduced to measure prediction deviation and improve the adaptability of prediction algorithm continuously via gradual fine-tuning.

This paper proceeds as follows: Section 2 discusses algorithm principle for HFPE. Section 3 clarifies algorithm establishment and analysis. Section 4 presents the experiment results and Section 5 concludes the paper.

2. Algorithm Principle

2.1. Basic Definition. Looking from mathematical form, the continuous time-varied curve, $z = f(t)$, is commonly applied to describe the evolving process of attack situation. This curve is carried out by computer through sampling method, that is, to sample situation values with time interval τ, and then obtains discrete time sequences composed by (t_k, z_k), where z_k represents the situation value at moment t_k. To facilitate the research, a basic definition is made as follows: let $G(i,m)$ be the segmental subimage with m neighboring segments from moment t_i, let q_k be the segmental gradient, let $Q(i,m)$ be the gradient sequence, let $(q_i, q_{i+1}, \ldots, q_{i+m-1})$, $L(i,m)$ be the characteristic spectrum of $Q(i,m)$, let O be zero vector; then

$$q_k = \frac{z_{k+1} - z_k}{t_{k+1} - t_k},$$
$$L(i,m) = O, \quad Q(i,m) = O,$$
$$L(i,m) = \frac{Q(i,m)}{\|Q(i,m)\|}, \quad Q(i,m) \neq O. \tag{1}$$

For the kth component product of $L(i,m)$, l_{i+k}, the angle of inclination, θ_{i+k}, can be defined as

$$\theta_{i+k} = \arctan l_{i+k}, \quad -\frac{\pi}{2} < \theta_{i+k} < \frac{\pi}{2}. \tag{2}$$

The stretch rate from $Q(i,m)$ to $Q(j,m)$ can be calculated by $f_\sigma(i,j,m)$, which is defined as

$$f_\sigma(i,j,m) = 1, \quad Q(i,m) = O, \ Q(j,m) = O,$$
$$f_\sigma(i,j,m) = \frac{\|Q(j,m)\|}{\|Q(i,m)\|}, \quad Q(i,m) \neq O, \ Q(j,m) \neq O,$$
$$\text{not exist}, \quad \text{other conditions}. \tag{3}$$

$\gamma[i,m,\rho]$ is utilized to adjust the stretch rate, where ρ is the prediction steps.

The following three theorems through further analysis can be easily obtained: (1) $L(i,m)$ does not change with $G(i,m)$; (2) if $L(i,m) = L(j,m)$, then $Q(i,m)$ is linear correlative with $Q(j,m)$; (3) if $L(i,m) = L(j,m)$, then $G(i,m)$ can be the same with $G(j,m)$ through translating and magnifying.

2.2. Prediction Principle. Looking from probability theory and statistics, similar situation curve shapes are more probably derived from similar origin, mechanism, and impact, subsequently resulting in a similar subsequent effect. From the point of view of statistics, when the precedence relations of sequences in time appear frequently, it usually meant that the logical causal relationship exists in a certain degree.

It is supposed that $G(i,m)$ and $G(j,m)$ are known historical feature subpatterns, from the same pattern, $t_i < t_j$, and the further trend after $t > t_{j+m}$ is unknown and needed to be predicted. If $G(i,m)$ is similar with $G(j,m)$, then it can be deduced that the origin, mechanism, and impact in $[t_j, t_{j+m})$ are similar with those in $[t_i, t_{i+m})$, and the history after t_{i+m} may be repeated after t_{j+m} with some differences. According to this principle, the slope of the line segment behind can be forecasted by

$$\widehat{q}_{j+m+k} = f_\sigma(i,j,m) \times q_{i+m+k}. \tag{4}$$

ρ is utilized to control the predicting steps. When $k = 0, 1, 2, \ldots, \rho - 1$, the trend prediction curve can be recurred by τ and $\widehat{q}_{j+m+k}$.

2.3. Measurement System

2.3.1. Fitting Degree. Firstly calculate the angle cosine similarity between slope vectors, secondly introduce more order difference operators to obtain the trend difference of qualitative change and quantitative change, and then acquire the narrowing fitting degree by the difference of similarity degree and trend difference.

Let $\phi_\theta(i,j,m)$ represent the angle cosine similarity between $Q(i,m)$ and $Q(j,m)$; then

$$\phi_\theta(i,j,m) = 1, \quad Q(i,m) = O, \quad Q(j,m) = O,$$
$$\phi_\theta(i,j,m) = \frac{Q(i,m) \times Q^T(j,m)}{\|Q(i,m)\| \times \|Q(j,m)\|}, \quad Q(i,m) \neq O, \quad Q(j,m) \neq O$$
$$\phi_\theta(i,j,m) = 0, \quad Q(i,m) = O, \quad Q(j,m) \neq O \quad \text{or } Q(i,m) \neq O, \quad Q(j,m) = O. \tag{5}$$

The trend differences of qualitative change and quantitative change are denoted by $\phi_1(x)$ and $\phi_2(x)$, respectively, and the former of which stands for the pattern difference,

and the latter stands for the accuracy difference. The above two parameters can be derived by

$$\phi_1(x) = \begin{cases} +1, & x > 0 \\ 0, & x = 0 \\ -1, & x < 0, \end{cases} \qquad \phi_2(x) = \sin x. \tag{6}$$

Thus the composite trend, $\phi_\perp(x)$, can be defined by

$$\phi_\perp(x) = 0.2 \times \phi_1(x) + 0.8 \times \phi_2(x). \tag{7}$$

Let ∇ represent backward difference operator, and define

$$\nabla^0 \theta_k = \theta_k, \tag{8}$$

and then the α order differential recursive equation can be obtained by

$$\nabla^\alpha \theta_k = 0.5 \times \left(\nabla^{\alpha-1}\theta_k - \nabla^{\alpha-1}\theta_{k-1}\right), \tag{9}$$

in which α is a positive integer, and (9) meets

$$\frac{\pi}{2} < \nabla^\alpha \theta_k < \frac{\pi}{2}. \tag{10}$$

Let $\phi_\nabla(i, j, m)$ denote the trend difference between the feature patterns $L(i, m)$ and $L(j, m)$; then

$$\phi_\nabla(i, j, m) = \frac{2\sum_{\alpha=0}^{m-1}\sum_{k=\alpha}^{m-1}\left|\phi_\perp\left(\nabla^\alpha\theta_{i+k}\right) - \phi_\perp\left(\nabla^\alpha\theta_{j+k}\right)\right|}{m(m+1)}. \tag{11}$$

The fitting degree function, $\phi(i, j, m)$, can be defined by

$$\phi(i, j, m) = \phi_\theta(i, j, m) - \phi_\nabla(i, j, m), \tag{12}$$

where the large value of $\phi(i, j, m)$ represents a fine fitting, and for $-1 < \phi_\theta(i, j, m) \leq 1$ and $0 < \phi_\theta(i, j, m) \leq 2$, it can be derived that

$$-3 < \phi_\theta(i, j, m) \leq 1. \tag{13}$$

The occurrence probability of $\phi_\theta(i, j, m) > 0$ may be 50% statistically, which is too big. Therefore, it is necessary to subtract the penalty term, $\phi_\nabla(i, j, m)$, and filter $\phi(i, j, m)$ by the threshold ε_ϕ $(0 < \varepsilon_\phi < 1)$ to narrow the fitting degree.

2.3.2. Universality Degree. Divide the attack situation subsequence into two parts, that is, occurred indication and subsequent effect; the values of the domination intensity of the former to the latter (or call link intensity between the two parts) may be high or low, some of which have a far-ranging representative, and some just have rare earth especially instance. If all the values are treated evenly, then the prediction accuracy will be affected seriously, so it is important to outstand inevitable link of the high intensity and weaken accidental link of the low intensity.

Let $\chi[k, m, \rho]$ be the universality value of $Q(k, m + \rho)$ in the historical feature pattern $G(0, n)$, where $\chi_{\max}$ can be derived by

$$\chi_{\max} = \max\left\{\chi[k, m, \rho] \mid 0 \leq k < n - m\right\}. \tag{14}$$

The value of $\chi_{\max}$ will be updated with the change of $\chi[k, m, \rho]$ and can be accessed directly without waiting to calculate.

The universality value can be shined upon to universality degree in $(0, n - m]$ by function $f_\chi(k, m, \rho)$, which is shown as follows:

$$f_\chi(k, m, \rho) = (n - m) \times \left(1 + 2\pi^{-1}\arctan\left(\chi[k, m, \rho] - \chi_{\max}\right)\right). \tag{15}$$

The larger value of universality degree reflects finer representativeness of $Q(k, m)$ and its extension and more exact patterns predicted by $Q(k, m + \rho)$. Otherwise, $Q(k, m + \rho)$ is just a special example, and the prediction effect is worse.

2.3.3. Contrast Degree. The predication results of situation are usually impacted by link intensity of several different weights. The function mechanism often changes; that is, sometimes they work with a community decision and sometimes with an individual domination. Therefore, it is necessary to trace and adjust between outstanding statistics effect and showing individual advantage.

It is supposed that $\widetilde{w}_1, \widetilde{w}_2, \ldots, \widetilde{w}_n$ are not normalized weights, which can be adjusted to $\widetilde{w}_1^\eta, \widetilde{w}_2^\eta, \ldots, \widetilde{w}_n^\eta$ by sensitization index η $(\eta > 0)$. Then the standardized weight w_k can be derived by

$$w_k = \frac{\widetilde{w}_k^\eta}{\sum_{k=1}^n \widetilde{w}_k^\eta}, \tag{16}$$

and comparison degree w_i/w_j can be obtained by

$$\frac{w_i}{w_j} = \frac{\widetilde{w}_i^\eta}{\widetilde{w}_j^\eta}. \tag{17}$$

If $\widetilde{w}_i \neq \widetilde{w}_j$, we can suppose $\widetilde{w}_i < \widetilde{w}_j$; then five generalized cutoff points can be acquired; that is, $0 < \widetilde{w}_i/\widetilde{w}_j < 1 < \widetilde{w}_j/\widetilde{w}_i < \infty$.

Equation (16) can be derived into the form of $\widetilde{w}_k^\eta = w_k \times \xi$ by

$$\frac{\widetilde{w}_k^{\eta\times x}}{\sum_{k=1}^n \widetilde{w}_k^{\eta\times x}} = \frac{(w_k \times \xi)^x}{\sum_{k=1}^n (w_k \times \xi)^x} = \frac{w_k^x}{\sum_{k=1}^n w_k^x}, \tag{18}$$

where η is utilized to adjust comparison degree; that is, $0 < \eta < 1$ outstands statistics effect, $\eta > 1$ shows individual advantage, and $\eta = 1$ maintains the present status.

3. Algorithm Establishment and Analysis

3.1. Prediction Algorithm. The prediction algorithm flow is given in Figure 1. According to historical feature pattern

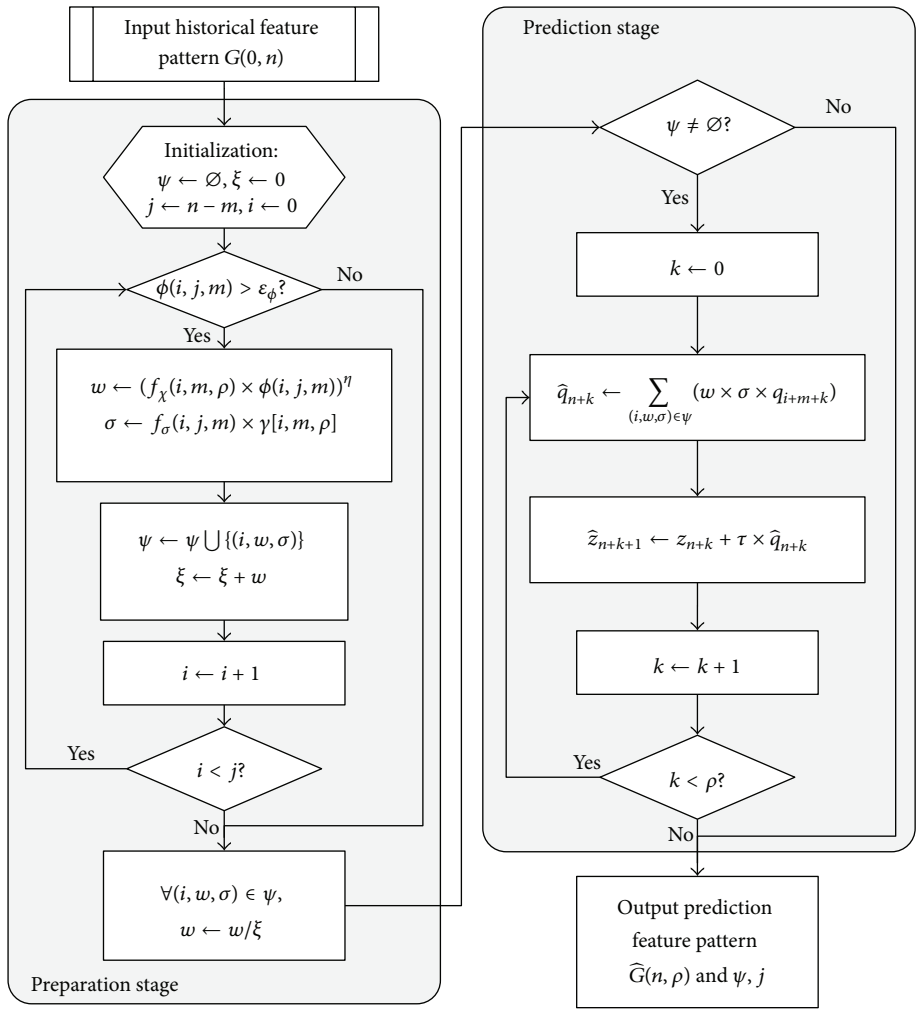

Figure 1: Prediction algorithm flow chart.

$G(0, n-1)$ and occurred indication $G(n-m, m)$, the algorithm can predict subsequent effect $\widehat{G}(n, \rho)$ through two main stages, that is, preparation stage and prediction stage, which are marked in the chart.

As shown in the figure, the preparation part circularly promotes the sliding window $G(i, m)$, selects poor values of fitting degree $Q(i, m)$ to reject, and sensitizes the product of universality degree $f_\chi(i, m, \rho)$ and fitting degree $\phi(i, j, m)$, which is assigned to w. The prediction part first checks whether historical feature pattern set ψ has record. What calls for special attentions is that the value of $Q(i, m)$ in the sliding window or the fitting degree value of it with $Q(j, m)$ in the occurred indication cannot be too small, because the smaller the above value, the poorer the contribution to prediction value $\widehat{q}_{n+k}$.

3.2. Evolution Algorithm. Evolution algorithm is introduced to measure predicting deviation from the views of pattern and accuracy, which can be fine-tuned to raise the adaptability of prediction algorithm.

The accuracy of adjusting η to $\eta \times x$ can be derived by

$$f_k(x) = \frac{\sum_{(i,w,\sigma)\in\psi} \left(w^x \times \phi\left(i + m, n, \rho\right)\right)}{\sum_{(i,w,\sigma)\in\psi} w^x}, \tag{19}$$

which is based on current weight set and (18) and meets $-3 < f_k(x) \leq 1$. And the definition can be popularized to $f_\Lambda(x_1, x_2, \ldots)$, only when $f_k(x_i)$ is the largest value first met in $\{f_k(x_1), f_k(x_2), \ldots\}$, which is in ascending order by k of x_k; it can be obtained that

$$f_\Lambda(x_1, x_2, \ldots) = x_i. \tag{20}$$

As shown in Figure 2, the evolution algorithm carries on the variables and results of the prediction algorithm and works to promote adaptability after acquiring measured value. $\Delta\varepsilon_\phi$ is an adjustment variable for ε_ϕ and meets $-n^{-1} \leq \Delta\varepsilon_\phi \leq n^{-1}$. If n rises or the distance between $|\psi|$ and $\ln n$ drops, then the adjustment amplitude becomes lower, else becomes higher. If $|\psi| < \ln n$, then decrease the threshold to soften the terms, else increase the threshold. $\Delta\chi$ is calculated

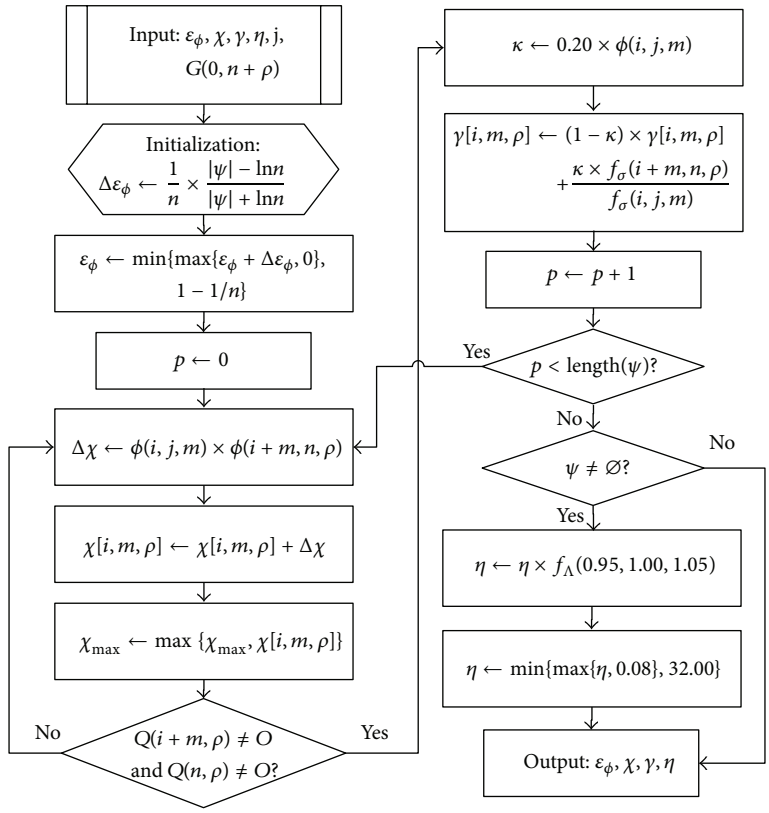

Figure 2: Evolution algorithm flow chart.

through fitting degree, $\phi(i, j, m)$, selected by the prediction algorithm, and meeting $0 < \phi(i, j, m) \leq 1$. When the fitting of $Q(i, m)$ and $Q(j, m)$ is poor, the adjustment to $\chi[i, m, \rho]$ needs to be cautious; that is, if $\phi(i + m, n, \rho) > 0$, then it needs to raise the value of $\chi[i, m, \rho]$, and if the extension value $Q(i + m, \rho)$ approximates $Q(n, \rho)$, then the prediction according to $Q(i + m, \rho)$ is accurate, and the value raising can be large. To determine η, select the best one among value lowering by 5%, current value, and value rising by 5%, and restrict it by a reasonable range to prevent passivating or sharpening.

3.3. Example Analysis. Figure 3 gives an example of predicting $\widehat{G}(13, 2)$ according to the historical feature pattern $G(0, 13)$, in which the value of (n, m, ρ) is $(13, 3, 2)$, $\varepsilon_\phi = 0.2$, $\chi[i, m, \rho] = 0.0$, $\chi_{\max} = 0.0$, $\gamma[i, m, \rho] = 1.0$, and $\eta = 1.0$.

According to the prediction algorithm, it can be known through comparing all the values of $Q(i, 3)$ with $Q(10, 3)$ that when $i = 0$ and $\phi(0, 10, 3) = 1.00$, so $Q(0, 3)$ is selected, and when $i = 5$ and $\phi(5, 10, 3) = 0.42$, so $Q(5, 3)$is also selected, and other values are excluded for they meet $\phi(i, 10, 3) \leq \varepsilon_\phi$, which are partly listed in Table 1. When $i = 5$, the slope becomes larger at $t = 7$ and smaller at $t = 12$, which are reflected through $\nabla^1\theta_7 > 0$ and $\nabla^1\theta_{12} < 0$ derived by (9), and the relative penalty value is recorded by $\phi_\nabla(5, 10, 3)$. And through normalization process, the elements of set ψ are $(0, 0.705, 0.67)$ and $(5, 0.295, 0.98)$. Thus, the prediction value $\widehat{q}_{13}$ is equal to $0.705 \times 0.67 \times (-1.29) + 0.295 \times 0.98 \times 0.50$,

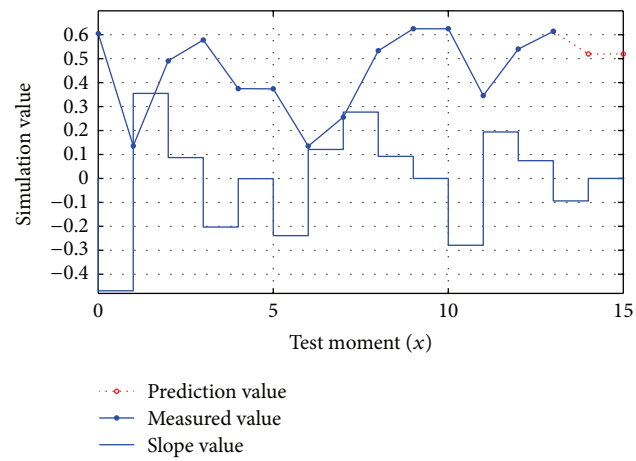

Figure 3: Trend prediction illustration.

Table 1: Similarity metric and punishment value.

i	0	2	3	5	6	8
ϕ_Θ	1.00	−0.85	0.42	0.72	0.32	−0.32
ϕ_∇	0.00	1.01	0.34	0.30	0.30	0.68

and so forth, and ρ step trend can be predicted. It can be seen that this scheme has the ability to identify multiple long-range correlation contained in the same situation sequence.

Table 2: Effects of prediction and evolution.

	1	3	5	7	9
ω_0	0.71	0.95	0.98	0.99	1.00
ω_5	0.29	0.05	0.02	0.01	0.00
$\widehat{q}_{13}$	−0.46	−0.78	−0.83	−0.85	−0.86
$\widehat{q}_{14}$	0.00	0.00	0.00	0.00	0.00
ε_ϕ	0.19	0.17	0.15	0.13	0.11
$f_\chi(0,3,2)$	10.00	10.00	10.00	10.00	10.00
$f_\chi(5,3,2)$	3.27	1.18	0.72	0.51	0.40
$\gamma[0,3,2]$	1.00	1.00	1.00	1.00	1.00
$\gamma[5,3,2]$	1.06	1.17	1.27	1.35	1.41
η	1.05	1.16	1.28	1.41	1.55

This part is analyzed according to evolution algorithm. Assuming that $q_{13} = -0.86$ and $q_{14} = 0.00$, so $|\psi| = 2$, which is smaller than ln 13; thus, the value of ε_ϕ needs to be lower, and once $|\psi| > 2$, then raise the value of ε_ϕ. It can be known that the changing value of universality degree $\chi[0,3,2]$ is 1.00×1.00, and raising this degree can strengthen the role of vector $Q(0,5)$. The changing value of $\chi[5,3,2]$ is $0.42 \times (-1.85)$, and lowering this value can weaken the interference of vector $Q(5,5)$. To bridge the gap between epitaxial scale and measured scale, $\gamma[0,3,2]$ is adjusted to $(1-0.20)\times 1.00 + 0.20 \times 0.67/0.67$, and $\gamma[5,3,2]$ is adjusted to $(1 - 0.08) \times 1.00+0.08\times 1.72/0.98$. And for $(f_\kappa(0.95), f_\kappa(1.00), f_\kappa(1.05)) = (0.13, 0.16, 0.19)$, the value of η needs to rise. Table 2 shows that $f_\chi(5,3,2)$ becomes smaller, and η becomes larger with continued evolution, which results in rapid rise of w_0/w_5, and approach between prediction value and measured value.

With the passage of time, m and ρ keep unchanged, n grows linearly, and the algorithm can delete stale data, save recent data, and correct fitting threshold and universality degree. The above process can be complicated not only by autonomous evolution, but also by artificially modified parameters.

4. Experiment Results Analysis

The traditional indexes utilized to measure the prediction accuracy include mean absolute error (MAE), standard deviation error (SDE), and mean absolute percentage error (MAPE) [21] derived by

$$\text{MAPE} = \frac{1}{\rho}\sum_{k=n+1}^{n+\rho}\left|\frac{\widehat{z}_k - z_k}{z_k}\right|. \tag{21}$$

This section selects MAPE to obtain the relative error between prediction pattern and measured pattern, which is denoted by E_r. The standard deviation of ρ relative error components is denoted by E_{std}.

4.1. Experiment 1. Figure 4 is a critical subsequence selected from actual network attack situation records, which includes various features such as ascent trend, saturation trend, decline trend, periodic fluctuation, and stochastic disturbance.

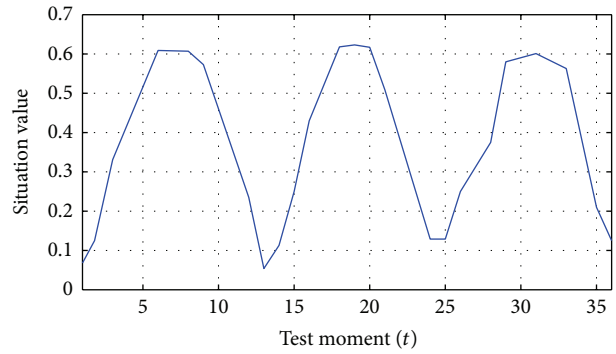

Figure 4: Historical records of network attack situation.

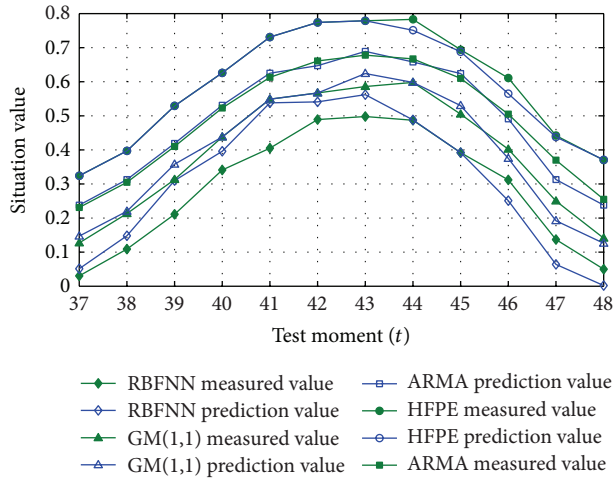

Figure 5: Prediction results of network attack situation.

From the view of the experimental prediction results, the relative errors of HFPE, ARMA, GM(1,1), and RBFNN are 3.28%, 5.89%, 7.18%, and 16.11%, respectively. As shown in Figure 5, in the experiment, ARMA, GM(1,1), and RBFNN need to be artificially identified and protected against cyclical situation fluctuations. The difference transformation utilizes 12 as the distance and is restored after prediction to prevent poor prediction effect; otherwise, the relative errors of GM(1,1) and RBFNN may reach 59.67% and 73.99%, respectively. However, the above method is special, cannot be spread for that data preprocessing of these algorithms does not exist in universal law. On the contrary, HFPE can maintain adaptation to complicated and changeable trends but does not need data preprocessing or artificial cognition.

4.2. Experiment 2. This experiment is to randomly choose subsequences with similar parts, repeat 20 times, and then calculate the average value.

From the view of the experimental prediction results, the relative errors of HFPE, ARMA, GM(1,1), and RBFNN are 8.09%, 20.89%, 44.89%, and 34.75%, respectively. If the situation sequence selected does not exist in any principle, then the relative errors will be 3.96%, 21.72%, 37.47%, and 53.54%, respectively. Figure 6 shows one group of data, in

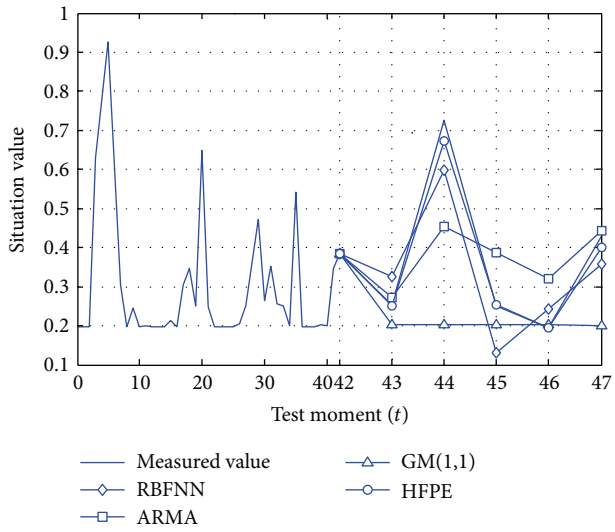

Figure 6: Prediction results of network attack situation.

which $t = 42$ is a boundary for historical feature pattern and prediction feature pattern.

If put all groups of the historical feature patterns into a new long sequence, and repeat above prediction, then the performance of ARMA and GM(1,1) drops rapidly, and that of HFPE does not change much for that longer sequence containing more correlation is benefit to prediction.

4.3. Experiment 3. To compare differences among four algorithms, random data are utilized to simulate situation sequences. First, extract random data with ξ bits from the entropy pool of Windows 7 system. Then randomly gather subsequence with 16 bits, the former 8 bits of which are occurred indication and the latter 8 bits are subsequent effect. Thirdly, splice occurred indication behind the random sequence to form a historical feature pattern and treat the subsequent effect as a prediction feature pattern. Let us make 100 groups of experiments to test each algorithm's capacity in resisting random interference and in identifying the correlation with far distance. The average results are listed in Table 3.

It can be found from the table data that HFPE has the best performance among the four algorithms. When the scale of experiment is large, this conclusion can be repeated well. And ARMA and RBFNN cannot deal with the random sequences with long bits, while HFPE can perform smoothly.

5. Conclusion

This paper proposes a prediction method based on historical feature pattern, that is, HFPE. The main principle of this algorithm is shown as follows. Fitting degree is introduced to measure the similarity among subsequences from the views of pattern and accuracy. Universality degree is utilized to test the representation of subsequence and its epitaxy. Contrast of the weight system is adjusted by sensitized index, which gives prominence to statistical effect in passivation area and highlights individual strengths in sharpening area. Prediction algorithm and evolution algorithm are proposed to predict situation results according to historical feature patterns. HFPE algorithm maximally reserves the rules in the situation sequences, does not need data preprocessing, and can adapt to the situation changes automatically. The experiment results prove the performance of HFPE.

Table 3: Prediction effects of network attack situation.

	ξ	ε_ϕ	E_r	E_{std}
HFPE	1.2×10^2	0.930	2.49	0.09
	1.2×10^3	0.980	1.75	0.09
	1.0×10^6	0.998	1.88	0.09
ARMA	1.2×10^2	—	32.96	0.15
GM(1,1)	1.2×10^2	—	49.32	0.20
	1.2×10^3	—	51.06	0.18
RBFNN	1.2×10^2	—	27.87	0.23

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

This work was supported by China Postdoctoral Science Foundation (no. 2013M540107).

References

[1] J. Wang, Z.-G. Qin, and L. Ye, "Research on prediction technique of network situation awareness," in *Proceedings of the IEEE International Conference on Cybernetics and Intelligent Systems (CIS '08)*, pp. 570–574, Chengdu, China, September 2008.

[2] H. Zhu, R. Lu, X. Shen, and X. Lin, "Security in service-oriented vehicular networks," *IEEE Wireless Communications*, vol. 16, no. 4, pp. 16–22, 2009.

[3] G. Jakobson, J. Buford, and L. Lewis, "Situation management," *IEEE Communications Magazine*, vol. 48, no. 3, pp. 110–111, 2010.

[4] F. Lan, W. Chunlei, and M. Guoqing, "A framework for network security situation awareness based on knowledge discovery," in *Proceedings of the 2nd International Conference on Computer Engineering and Technology (ICCET '10)*, vol. 1, pp. 226–231, Chengdu, China, April 2010.

[5] W. Lou and K. Ren, "Security, privacy, and accountability in wireless access networks," *IEEE Wireless Communications*, vol. 16, no. 4, pp. 80–87, 2009.

[6] A. Feng, M. Knieser, M. Rizkalla, B. King, P. Salama, and F. Bowen, "Embedded system for sensor communication and security," *IET Information Security*, vol. 6, no. 2, pp. 111–121, 2012.

[7] B. Magoutas, G. Mentzas, and D. Apostolou, "Proactive situation management in the future internet: the case of the smart power grid," in *Proceedings of the 22nd International Workshop on Database and Expert Systems Applications (DEXA '11)*, pp. 267–271, Toulouse, Greece, September 2011.

[8] C. G. Keller, T. Dang, H. Fritz, A. Joos, C. Rabe, and D. M. Gavrila, "Active pedestrian safety by automatic braking and evasive steering," *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 4, pp. 1292–1304, 2011.

[9] X. Chen, H. Gao, and Y. Fu, "Situation analysis and prediction of web public sentiment," in *Proceedings of the International Symposium on Information Science and Engineering (ISISE '08)*, vol. 2, pp. 707–710, Shanghai, China, December 2008.

[10] N. Zhang, N. Cheng, N. Lu, H. Zhou, J. W. Mark, and X. Shen, "Risk-aware cooperative spectrum access for multi-channel cognitive radio networks," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 3, pp. 516–527, 2014.

[11] R. K. Iyer, Z. Kalbarczyk, K. Pattabiraman et al., "Toward application-aware security and reliability," *IEEE Security & Privacy*, vol. 5, no. 1, pp. 57–62, 2007.

[12] K. Janac, "Control of large power systems based on situation recognition and high speed simulation," *IEEE Transactions on Power Apparatus and Systems*, vol. 98, no. 3, pp. 710–715, 1979.

[13] R. K. Iyer, Z. Kalbarczyk, K. Pattabiraman et al., "Toward application-aware security and reliability," *IEEE Security & Privacy*, vol. 5, no. 1, pp. 57–62, 2007.

[14] A. Kitadai, M. Nakagawa, H. Baba, and A. Watanabe, "Similarity evaluation and shape feature extraction for character pattern retrieval to support reading historical documents," in *Proceedings of the 10th IAPR International Workshop on Document Analysis Systems (DAS '12)*, pp. 359–336, Gold Cost, Australia, March 2012.

[15] J. R. Almeida, J. B. Camargo, and P. S. Cugnasca, "Safety and security in critical applications and in information systems ?? A comparative study," *IEEE Latin America Transactions*, vol. 11, no. 4, pp. 1127–1133, 2013.

[16] M. Panteli, P. A. Crossley, D. S. Kirschen, and D. J. Sobajic, "Assessing the impact of insufficient situation awareness on power system operation," *IEEE Transactions on Power Systems*, vol. 26, no. 3, pp. 2967–2977, 2013.

[17] M. M. Masud, Q. Chen, L. Khan et al., "Classification and adaptive novel class detection of feature-evolving data streams," *IEEE Transactions on Knowledge and Data Engineering*, vol. 25, no. 7, pp. 1484–1497, 2013.

[18] L. Jibao, W. Huiqiang, L. Xiaowu, and L. Ying, "A quantitative prediction method of network security situation based on wavelet neural network," in *Proceedings of the 1st International Symposium on Data, Privacy, and E-Commerce (ISDPE '07)*, pp. 197–202, Chengdu, China, November 2007.

[19] W. He, G. Hu, and H. Xiang, "Apply anomaly grey forecasting algorithm to cyberspace situation prediction," in *Proceedings of the IEEE International Conference on Cybernetics and Intelligent Systems (CIS '08)*, pp. 503–505, Chengdu, China, September 2008.

[20] R.-F. Wu and G.-L. Chen, "Research of network security situation prediction based on multidimensional cloud model," in *Proceedings of the 6th International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS '12)*, pp. 409–414, Palermo, Italy, 2012.

[21] X. Cheng and S. Lang, "Research on network security situation assessment and prediction," in *Proceedings of the 4th International Conference on Computational and Information Sciences (ICCIS '12)*, pp. 864–867, Chongqing, China, August 2012.