

Criminal Statistics Analyzing Based on Deep Learning Methods

Xuepeng Huang

Information & Network Center
Hubei University of Police
P.R. China

Jianhui Lin*

Department of Information Technology
Hubei University of Police
P.R. China

*Corresponding author

Abstract—Social Security Administration analysis is studied based on intensive probe to present social fluctuation researches. A complex social system model is constructed by profile vector of complex system. According to the character of the complex social system, a simplified algorithm based on deep learning is proposed, which contains a system anomaly testing expression and a corresponding referenced anomaly testing expression according to the stochastically transfer delay of the complex system. Prototype system and experiment results show the precision and credibility of the social system anomaly index obtained from the algorithm.

Keywords—criminal statistical data; anomaly value; deep learning; transfer delay

I. INTRODUCTION

As the differences and changes of politics, economics and culture in a certain human society, social order tendency fluctuates correspondingly. How to distinguish the warning changes from routine changes is becoming event important for us. Quite simply, big data reflects the changing world we live in. The more things change, the more the changes are captured and recorded as data. Processing information like this illustrates why big data has become so important:

- Most data collected now is unstructured and requires different storage and processing than that found in traditional relational databases.
- Available computational power is sky-rocketing, meaning there are more opportunities to process big data.
- The Internet has democratized data, steadily increasing the data available while also producing more and more raw data.

Data in its raw form has no value. Data needs to be processed in order to be of valuable. Thus, crime tendency analysis can be obtained from the comprehensive processing toward many relevant factors which involve social status, population fluctuation, political reformation, crime quantity, crime type, crime structure, etc.[1] In historical researches, we deal with a novel algorithm which analyzes social order trend from criminal statistical data. [1,2]

Deep learning algorithms are contrasted with shallow learning algorithms by the number of parameterized transformations a signal encounters as it propagates from the

input layer to the output layer, where a parameterized transformation is a processing unit that has trainable parameters, such as weights and thresholds[3]. According to the character of the complex social system, it uses many layers of nonlinear processing units for feature extraction and transformation, the processing algorithms may be supervised or unsupervised and applications include pattern recognition and statistical classification, and it is based on the learning of multiple levels of features or representations of the data. This paper discusses the optimized algorithm based on deep learning.

II. COMPLEX SOCIAL SYSTEM DESCRIPTION WITH STATISTICAL DATA

A. Complex Social System Model

Assume a certain complex system S which contains n parameters $S = (s_1, s_2, \dots, s_n)$. Provided M can be described by a nonnegative integral within each fixed time span t , after a preset enough long time γ , S can be described by an N -dimensional random variable:

$$S = (\bar{s}_1, \bar{s}_2, \dots, \bar{s}_n)$$

Within it, $j = \text{int}(\gamma / t), s_i(j) = c_{ij}$.

Assuming matrix $S' = (\bar{s}'_1, \bar{s}'_2, \dots, \bar{s}'_n)$, its factor \bar{S}' contains q lines, then the relation between S and \bar{S}' can be described as follows :

$$\begin{cases} s'_1(x) = s_{last-i}(x) \\ s'_q(x) = s_{last-i-q+1}(x) \end{cases} \quad (1)$$

Then provided random variable \bar{S}_i observes Poisson distribution, we can generate N -dimensioned random variable:

$$Mean = (\bar{mean}_1, \bar{mean}_2, \dots, \bar{mean}_n)$$

which generate a new line when M' update every time :

$$mean_{i,j} = E(\bar{s}_j(x)) \quad (2)$$

Def. 1: Vector $profile = (mean'_{last,j}(x))$ is called profile vector of complex system C, if the following requirements are met:

$$mean'_{k,j}(x) = E(\bar{s}'_j(x)) \quad k=1$$

$$mean'_{k,j}(x) = \frac{mean_{k,j}(x) + \zeta \cdot mean'_{k,j-1}(x)}{1 + \zeta} \quad k > 1$$

ζ is a constant, and $0.9 \leq \zeta < 1$,

$Mean'$ is call profile set of complex system S.

Def. 2: N-dimensioned random variable D is called anomaly index if following requirement are met:

$$d_i = (0), \text{ if } i = 1$$

$$d_i = mean_i - mean'_{i-1}, \text{ if } i > 1$$

B. Abnormal Detection

Since every column of D is difference value between the mean value of a Poisson distribution variable and the weighted value of which, obviously d_i observes a Normal distribution, we can draw a conclusion that:

$$\bar{d} \sim N(0,1),$$

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$$

And they are independent of each other, then

$$\bar{d} / \sqrt{\frac{(n-1)S^2}{\sigma^2(n-1)}} = \frac{\bar{d}}{S/\sqrt{n}} \sim t(n-1)$$

Now making a hypothesis:

$$H_0: \mu_d = 0$$

Note sample mean value and sample variance value of d_1, d_2, \dots, d_n as \bar{d}, S^2 separately, then:

$$\bar{d} = \frac{1}{n} \sum_{i=1}^n d_i$$

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (d_i - \bar{d})^2$$

Now setting a small positive number $\alpha > 0$, we suppose Eq.3 as anomaly detecting formula of complex system C and Eq.4 as referenced anomaly detecting formula.

$$|t| = \left| \frac{\bar{d}}{S/\sqrt{n}} \right| \geq t_{\alpha/2}(n-1) \quad (3)$$

$$p = P(t \geq t_{\alpha/2}(n-1) | \mu_d = 0) \quad (4)$$

III. ANALYZE CRIMINAL STATISTICAL DATA WITH DEEP LEARNING METHODS

A. Defetion of Anomaly Value.

For $mean'_i$, on consumption that the effective cardinality is p, then its sample mean is:

$$\bar{f}_i = \sum_{j=1}^p mean'_{ij} / p$$

Thus we can call the $n \times n$ Positive definite matrix W as its sample covariance matrix.

$$W = \sum_{j=1}^n (mean'_{ij} - f_i)'(mean'_{ij} - f_i) / (n-1)$$

If there is a vector x with known mean μ , then we call:

$$t = p(\bar{f} - \mu)W^{-1}(\bar{f} - \mu)^T \sim T^2(n, p-1)$$

as the abnormal testing value of the complex system S.

In order to present the affection that transfer phenomenon impact on the statistical data, covariance matrix has to be optimized as follows.

An $n \times n$ matrix $Cov_{n \times n}$ is defined to substitute for original variance matrix.

Then

$$Cov'(x, y) = \frac{1}{n} \sum_{i=1}^n m'_i(x)m'_i(y) - mean_{last-i}(x) \cdot mean_{last-i}(y)$$

$$= \frac{1}{n} \sum_{i=1}^n m'_i(x)m'_i(y) -$$

$$\left(\frac{mean_{max-k}(x) + \zeta \cdot mean'_{max-k-1}(x)}{1 + \zeta} \right) \cdot \left(\frac{mean_{max-k}(y) + \zeta \cdot mean'_{max-k-1}(y)}{1 + \zeta} \right)$$

B. Transfer Delay Analyzing Based on Deep Learning

Since the transfer delay of complex network possesses stochastic characteristic, if the controlled system is a non-linear system, the complex social system will be a non-linear system with stochastic delay^[4]. If there are r operation points in the non-linear system, we can build r local linear subsystem models with stochastic time-varying delay at each point based on the theories of T-S fuzzy theory. Additionally, the transfer delay of network is stochastic; we can divide each of them into l linear subsystems again according to the length of time-delay^[5].

In the i th ($i = 1, 2, \dots, r$) linear subsystem with time-varying delay, the j th rules of T-S fuzzy model is of the following form:

Model rule j : ($j = 1, 2, \dots, l$)

IF

$$(j-1)T < \tau_j \leq jT,$$

THEN

$$x(k+1) = A_i x(k) + B_i u(k-j), \quad y(k) = C_i x(k).$$

Using parallel distributed compensation:

$$u(k) = F_{ij} x(k).$$

Choosing F_{ij} properly can make the i th system stability, then the model of the i th linear subsystem with time-varying delay is:

$$\begin{aligned} x(k+1) &= A_i x(k) + B_i \sum_{j=1}^l \mu_j(\tau_j) u(k-j), y(k) \\ &= C_i x(k), \quad u(k) = \sum_{j=1}^l \mu_j(\tau_j) F_{ij} x(k) \end{aligned}$$

where $\mu_j(\tau_j)$ present the probability of τ_j occurs in the system, that is: $\mu_j(\tau_j) = p(\tau_j)$ and

$$\begin{aligned} \sum_{j=1}^l \mu_j(\tau_j) &= 1 \\ 0 \leq \mu_j(\tau_j) &\leq 1 \quad \forall j = 1, 2, 3, \dots, l \end{aligned}$$

IF

$$z_1(k) \text{ is } M_{i1}, z_2(k) \text{ is } M_{i2}, \dots, z_p(k) \text{ is } M_{ip}$$

THEN

$$\begin{aligned} x(k+1) &= A_i x(k) + B_i \sum_{j=1}^l \mu_j(\tau_j) u(k-j), \quad y(k) \\ &= C_i x(k), \quad u(k) = \sum_{j=1}^l \mu_j(\tau_j) F_{ij} x(k) \end{aligned}$$

And we can obtain the two-layers fuzzy model of social system with stochastic transfer delay as follows:

$$\begin{aligned} x(k+1) &= \sum_{i=1}^r \mu_i(z(k)) \{A_i x(k) + B_i \sum_{j=1}^l \mu_j(\tau_j) u(k-j)\} \\ y(k) &= \sum_{i=1}^r \mu_i(z(k)) C_i x(k) \\ u(k) &= \sum_{i=1}^r \mu_i(z(k)) \sum_{j=1}^l \mu_j(\tau_j) F_{ij} x(k) \end{aligned}$$

where $\mu_i(z(k))$ is the membership of $z_i(k)$, which meets the condition:

$$\begin{aligned} \sum_{i=1}^r \mu_i(z(k)) &= 1 \\ 0 \leq \mu_i(z(k)) &\leq 1 \quad \forall i = 1, 2, \dots, r \end{aligned}$$

Rewriting it we can obtain:

$$\begin{aligned} x(k+1) &= \sum_{i=1}^r \mu_i(z(k)) \{A_i x(k) + B_i \sum_{j=1}^l \mu_j(\tau_j) u(k-j)\} \\ y(k) &= \sum_{i=1}^r \mu_i(z(k)) C_i x(k) \\ u(k) &= \sum_{i=1}^r \sum_{j=1}^l \mu_i(z(k)) \mu_j(\tau_j) F_{ij} x(k) \end{aligned}$$

Therefore, the computation can be simplified for convenience as follows:

$$\text{Cov}(x, y) \approx \frac{1}{n} \text{mid}(x, y) - \text{mean}(x) \cdot \text{mean}(y)$$

And the anomaly value of system can be calculated:

$$t = n(\text{mean} - \text{profile}) \text{Cov}^{-1}(\text{mean} - \text{profile})^T \quad (5)$$

C. System Detection Algorithm Process

In general, the Hypothesis test Algorithm we supposed can be described as follows:

Composite Statistic Profiling-Vector Hypothesis test Algorithm

Input: attenuation period ζ , parameter statistical value $measure_i$, detection segment ds, statistical segment ts

Output: anomaly index of complex system C

Method:

while .t.

line=countline(M);//lines of M

insertline(M,measure);// insert a new line into M

for each measure

if line<ts

 window_start=0;

 window_end=line;

else

 window_start+=1;

 window_end+=1;//slide windows

endif

mean_value=getmean(M,window_start,window_end);

// computing mean value

endfor

insertline(mean,mean_value);// insert a new line in mean

if first_line_of_mean

 insertline(mean1,mean_value);

else

$mean1 = (mean + \zeta \cdot mean1) / (1 + \zeta)$;

endif

for each end_of_ts

 d=mean-mean1 ;

 insertline(D,d);

end for

s=variance(d) ; count statistical standard variance

d1=average(d) ;

t=abs(d1).sqrt(n).s

insertline(T,t); //record the testing anomaly value

if t>threshold_value

 if test_fail // refuse 0 assumption

 send_abnormal(serious_alert); // serious alert

 else

 send_abnormal(warning); //warning

 endif

else

 send_normal();

endif

return t;

endwhile

As far as the process of anomaly value is concerned, there are two opinions: (1) to eliminate the effect system property vector upon system property value to make system property stable; (2) since the system anomaly value implies the fluctuation of social status, it should be remained to affect and update social order vector in time. Considering the fluctuation characteristics of social order status against relatively stable status, its value may not be constant but changed slowly. This kind of movement derived from the anomaly status of complex system, as a result the anomaly status should remain. Of course, anomaly rises from statistical error have to be eliminated.

IV. EXPERIMENTS AND CONCLUSIONS

In experiment environment, we collect the statistical data of criminal case occurrence in a certain city in 2008. Considering that the algorithm bases upon historical data, we abstract only last 20 weeks data as our detection data and others as training

data. Setting $\zeta = e^{\frac{d(\ln 0.5)}{s}} = 0.9330$, experiments results are showed as Fig. 1. Setting alert threshold as 0.75 and system can receive 4 alerts. Comparing with initial data we collect, urban criminal statistics data changes sharply when alert signal received. Consequently we can draw a conclusion that analysis prediction results are scientific and accurate.

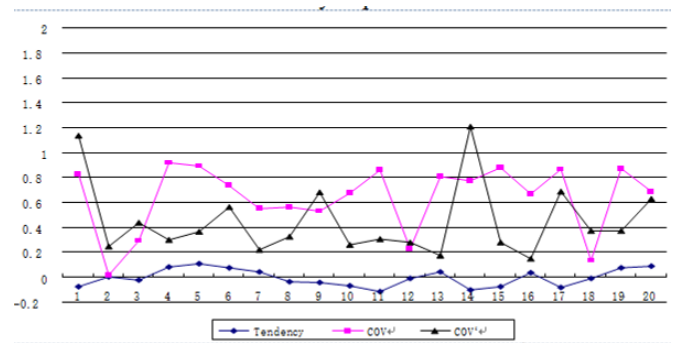


FIGURE 1. EXPERIMENT RESULTS AND COMPARISON

ACKNOWLEDGEMENT

It is a project supported by Social Security Administration of Hubei Province Social Science key research bases(2014-24), the humanities and social science research projects (2012G121) , the humanities and social science research

projects (2013G351) , Science research project of Hubei Province Education Departmentand (B2013043) ,the universities youth science and technology innovation team project of Hubei Province (T201222).

REFERENCES

- [1] J. H. Lin, T. S. Huang: Analysis and realization of a crime prediction system based on statistics algorithm. Journal of Information and Computational Science. Vol. 4(2007), p.1045
- [2] J. H. Lin, T. S. Huang: Crime Information Analysis and Complex Social System Tendency Researches. Journal Computer Engineering and Application. Vol. 32(2010), p.47
- [3] P. E. Utgoff, D. J. Straczzi: Many-layered learning. Neural Computation. Vol.14(2002)7,p. 2497
- [4] S. C. Tong, H. X. Li: Observer-based robust fuzzy control of nonlinear systems with parametric uncertainties. Fuzzy Sets and Systems. Vol. 131(2002), p.165-184
- [5] X.J. Ma, Z. Q. Sun: Analysis and design of fuzzy reduced-dimensional observer and fuzzy functional observer. Fuzzy Sets and Systems. Vol. 120 (2001), p. 35.