Non-linear Point to Point Propagation Model Based on Complex Networks

Ya-wei ZHAO¹*, Pu WANG¹²³, Yu-tao WANG¹³ and Gang LUO¹³

¹Laboratory of Big Data Analysis Technology, Chinese Academy of Sciences, Beijing China
²Institute of Quantitative Economy and Technological Economy, Chinese Academy of Social Sciences, Beijing China
³AI Laboratory of Beijing Knowlegene Data Technology Co., Ltd., Beijing China

*Corresponding author

Keywords: Complex network, Non-linear point to point propagation, Edge weight, Propagation damping.

Abstract. Infectious disease model is used to describe the global level of a complex network propagate, but in the real world, people pay more attention to the point to point level of a complex network propagate, especially for the quantification of point to point propagation. In order to solve this problem, we proposed a non-linear point to point propagation model of complex network based on edge weight calculation. Characteristics of edge propagation weight between nodes is described and quantified from point to point perspective. At the same time, nodes have a damping effect on propagation. The propagation damping calculation is incorporated into characteristics of edge propagation weight and node which is functional. Referring to some ideas of the neuron model, series theory and information transmission theory, and the model is verified in a large-scale financial risk propagation experiment.

Introduction

Generally speaking, complex network in the real world are mostly heterogeneous, and the roles played by different nodes and edges in the process of network propagation are often very different. It has become an important issue that how to effectively measure the importance of nodes and edges, and the specific roles and significance played by nodes and edges in the process of point to point propagation. However, there is no general maturity model in the field of complex network point to point propagation, so the study of complex network point to point propagation has important theoretical and practical significance.

Taking into account the above point of views, this paper proposes a non-linear point to point propagation model of edge weight calculation based on complex network theory, which puts forward two improvements: (1) Using the neuron model, the nodes in the network are regarded as neurons, which are used to express the difference of nodes, enrich the information of nodes, and measure the importance of nodes in the process of propagation; (2) Using the non-linear distance to express different relationships between nodes, and putting forward the concept of propagation damping to measure the edges in the process of propagation.

Literature Overview

The currently propagation theory of complex network mainly focuses on the global level. Most scholars have studied the problem of propagation in the actual scenario [1-13]. To sum up, the current literature research mainly focuses on global level propagation, which can be classified as global propagation model, because of taking into account global, statistical and structural factors of complex network. Some documents take into account individual factors, such as “individual propagation ability [5]”, “individual weight adaptation mechanism [6]” and “individual current popularity [8]”, but these concepts are obviously insufficient in expressing individual differences. In addition, in the
above literature, no scholars put forward appropriate and accurate concepts and models for individual differences.

Therefore, the above-mentioned literature ignores the specific roles and significance of the differences of individuals and relationships in the process of point to point propagation, and can’t solve the problem of point to point propagation based on complex network. However, through the current literature research, no more mature and generalized point to point propagation model has been found. Therefore, the non-linear point to point propagation model proposed in this paper is precisely for solving the point to point propagation problem of complex network.

**Non-linear Point to Point Propagation Model**

In the real world, in order to outline the mechanism of non-linear point to point propagation more accurately, this paper builds non-linear point to point propagation network structure based on neuron theory, discrete-time system and series convergence theory.

**Assumptions**

**Assumption 1:**

This paper supposes that the propagation process of multiple sources can be decomposed into the superposition of the propagation of a single source. A single source propagates through the edges of complex network, furthermore, a single propagation source can be decomposed into the propagation from one node to another adjacent node in a complex network. Therefore, the single point to point propagation structure is the basic analysis unit of non-linear point to point propagation network model.

**Assumption 2:**

The real attenuation curve of the propagation process is a strictly monotone decreasing function curve. In practice, the attenuation trend is commonly seen. This paper only studies this curve. For example, because each bank has the ability and willingness to avoid resisting risks, when a risk event occurs at the node represented by a bank, the intensity of propagation in the process of propagation will decrease with the increase of the propagation path.

According assumption 1 and assumption 2, the following analysis focuses on the propagation process from point to point structure.

Analyzing the structure of the network from point to point perspective, there are four general propagation structures: (1) single-point to single-point; (2) single-point to multi-point; (3) multi-point to single-point; (4) multi-point to multi-point.

Among the four general structures, single point to multi-point can be seen as multiple single point to single point. But if the structure is multi-point to single point, there is a problem inevitable that multiple propagation intensities must be transformed into a comprehensive propagation intensity. Multi-point to multi-point can be decomposed into the first three cases.

Therefore, the following analysis focuses on the propagation process from single point to single point and multi-point to single point structure.

**Non-linear Point to Point Propagation Network Structure**

**Single Point to Single Point Structure.** Because most of the real world propagation is not infinitely spread, the length and range are limited. In order to simulate the actual situation, this paper cites the neuron structure and uses the rectified linear unit (ReLU) as the activation function when the propagation intensity is less than the threshold value and cannot propagate. Otherwise, the propagation is expected to continue.

To illustrate the single point to single point structure, it is assumed that a propagation process \( p \) is propagated from node \( i \) to \( j \), which is one of its neighbors. The structure is as follows:
Symbolic and Formal Descriptions

Description 1:
Assuming that the force of a certain propagation process $p$ acting on the node $i$ is $p_i$, which made node $i$ happen to a $\delta_i$ change, then the force $p_i$ is called propagation intensity of propagation process $p$ on node $i$.

In the process of propagation, the intensity of propagation is often reflected in the strength of the information or signal. Therefore, it is converted into a numerical description. Generally, the larger the value, the stronger the intensity of propagation, the smaller the value, the weaker the intensity of propagation.

Description 2:
The set of node attributes of a node $i$ in a network is used for representing the individual differences of nodes in the non-linear point to point network structure. The set is named $\mathbf{v}_i$.

Description 3:
Assuming that the propagating edge weight $r_{ij}$ is defined as the weight of the whole relationship between two adjacent nodes $i, j$, which is mainly determined by the set of relational attribute of two adjacent nodes. The propagation edge weight is an important parameter of non-linear point to point propagation model because propagating is transmitted by edge.

Description 4:
In complex network $G$, the structure damping of two adjacent nodes $i, j$ is $c_{ij}$ which is a function with the parameters of $\mathbf{sys}_i, \mathbf{sys}_j$. $c_{ij}$ expresses the network structural factors causing propagation attenuation in the process of propagation. The parameters $\mathbf{sys}_i, \mathbf{sys}_j$ are the systemic importance of node $i, j$ in $G$.

Description 5:
In complex network $G$, the self-damping of a node is $g_i$ which expresses the propagation attenuation in the process caused by self-factors through self-damp function $G_i(\mathbf{v}_i)$.

Since normal form distribution can be widely applied to the study of human social phenomena and economic phenomena [14]. This paper assumes that the function form of $G_i(\mathbf{v}_i)$ is a normal distribution function.

Description 6:
In complex network $G$, the propagation process tends to be attenuated. Assuming that a propagation process $p$ propagates from node $i$ to adjacent node $j$, the propagation intensities of node $i, j$ are $p_i, p_j$, there is an attenuation of $p$, the attenuation amount is $\Delta p$, and then the factors leading to attenuation are defined as propagation damping. The propagation damping between two adjacent nodes is related to the structural damping between the two nodes, the self-damping of the node to be propagated and the propagation boundary weight. The function forms are as follows:

\[
d_{ij} = D_{ij}(c_{ij}, g_j) \tag{1}
\]

\[
p_j = (p_i - \Delta p) \times r_{ij} = (p_i - p_i \times d_{ij}) \times r_{ij} = p_i \times (1 - d_{ij}) \times r_{ij} \tag{2}
\]

$d_{ij}$ represent the propagation damping and $D_{ij}(c_{ij}, g_j)$ is the propagation damping function. According to formula (1), the attenuation value of each propagation step and the propagation intensity of the propagated nodes can be calculated.

Process Description
The specific propagation process is as follows, $p_i$ is the propagation intensity of the propagation $p$ at the node $i$, assuming that $p_i$ is greater than the propagation threshold, passing through the
activation function, and $p_i$ attenuate because of the propagation damping, $p_j$ is the propagation intensity when it reaches the node $j$. Repeating the above process, he propagation attenuation process including the propagation intensity at each node being propagated and the path is calculated.

The propagation of edge weights between nodes in a complex network, its own differences and structural differences lead to different attenuation processes and propagation strengths of each node undergoing propagation.

The above is the single point to single point structure. In the experimental part of this paper, it is proved that the structure proposed in this paper can describe the model well.

**Multiple Point to Single Point Structure.** In practical application scenarios, when one or more propagations are transmitted to the same node through multiple nodes, a multi-point to single-point structure is formed. At this time, there are multiple propagation intensities at the node, the overall propagation strength is generally convergent and bounded. It is assumed that $p_1, p_2, p_3, \ldots, p_n$ are the propagation intensities at a node and the overall propagation intensity is generally greater than $p_{\text{max}}$ which is the maximum value of the intensities. Because the simple addition operation cannot satisfy the overall convergence and boundary, so this paper uses the linear difference equation with constant coefficients method in discrete time systems to calculate the overall propagation intensity.

The specific calculation process as follows:

The first step: Because the smaller the propagation intensity relative to the maximum value, the smaller the contribution and influence on the overall propagation intensity, and vice versa. Therefore, the propagation intensity sequence is sorted from small to large to form a skew sequence $p = \{p^{(1)}, p^{(2)}, p^{(3)}, \ldots, p^{(n)}\}$ (see Figure 2).

![Figure 2. Propagation intensity sequence.](image)

The second step: Suppose $y(n)$ is the final state of the overall propagation intensity, $y(n)$ depends on the last value $p^{(n)}$ which is in the sequence of propagation intensity and the previous state value $y(n-1)$, and pushes sequentially, so the linear difference equation with constant coefficients is constructed as follows:

\[
\begin{align*}
y(n) &= ay(n-1) + p^{(n)} \\
y(n-1) &= ay(n-2) + p^{(n-1)} \\
&\quad \vdots \\
y(2) &= ay(1) + p^{(2)} \\
y(1) &= ay(0) + p^{(1)} \\
y(0) &= 0
\end{align*}
\]

The constant coefficient $\alpha$ is generally greater than 0, and the iteration method is used to solve $y(n)$:

\[
y(n) = p^{(n)} + \alpha p^{(n-1)} + \alpha^2 p^{(n-2)} + \cdots + \alpha^{n-1} p^{(1)}
\]

(3)

According to the series convergence criterion, when $\alpha < 1$, the series converges, and then, $y(n)$ converges and is bounded.

In summary, this paper uses the difference equation method to obtain the propagation strength calculation model of the multi-point to single-point structure. In practical applications, the values can be adjusted according to specific data and target tasks.
Experiment

The non-linear point to point propagation model proposed in this paper is used to solve the tasks and problems in the actual scenario, and verify the correctness and effectiveness of the model.

The spread of credit risk in enterprise networks is a problem that all banks pay more attention to and has a strong practical significance. Therefore, this paper uses the model to solve the problem of early risk warning of corporate network credit risk propagation.

There are two main indicators for the accuracy and effectiveness of the model. One is the precision rate, and the second is the recall rate.

Experiment Process

The data comes from a large bank (including corporate attributes, corporate relationship attributes, corporate loan data), and the time range is from January 2017 to June 2018. Taking the five-category classification of corporate loans as the observation target, this paper regards the five-level classification as the risk of the enterprise. The predicted time window is 6 months, that is, the events that are pushed forward for 6 months at the predicted time point are valid events, and the time range of the predicted results is the next 6 months.

The precision and recall rate of the five-level classification forecast for corporate loans reflect the accuracy and effectiveness of the model.

(1) Data preparation
The relationship between the enterprise is abstracted into a complex network. The bank risk management experts provide 34 types of common risk events about enterprises and score them by numerical value as the propagation source.

(2) Modeling
Determine the actual parameters of the model based on the actual data provided by the bank.

(3) Calculation
This article uses a computer program to implement a corresponding algorithm to calculate and statistically analyze the results.

Analysis of Experimental Results

Fig. 3 is a comparison between the results of non-linear point to point propagation risk early warning and traditional risk early warning. The traditional risk early warning methods include expert analysis method, financial ratio analysis method and the logistic regression method. Traditional risk early warning results are provided by banks.

Figure 3. Comparison of precision rate of traditional methods and non-linear point to point propagation on risk early warning.
In a given period of time: January 2017 to June 2018, the accuracy rate of traditional methods is stable at about 7%, while the precision rate of non-linear point to point propagation risk early warning is significantly higher than that of traditional risk early warning, which is stable between 23% and 33%. The increase is about two to three times, which shows that the non-linear point to point propagation model of complex networks, is remarkably accurate and effective, so that the overall idea of the model is correct.

Conclusion
This paper proposes a non-linear point to point propagation model based on complex network, which is a useful attempt to quantitatively analyze and predict local and individual dynamic behaviors of nodes in a complex network. Through the application of the model and experiments, the accuracy and effectiveness of the model are verified, which provides a practical solution to the problem of point to point propagation of complex networks. The model can be widely used in public opinion propagation, traffic flow analysis, infectious disease analysis, bank risk event propagation analysis and so on.

Acknowledgement
This research was financially supported by : (1) Natural Science Foundation of China (Grant No.61872331); (2) The 13th Five-Year Plan Information Project of the Chinese Academy of Sciences Foundation (Grant No. XXH13504-05).

References


