A novel computation for predicting time series using fuzzy logical distance connectivity function and visibility graph theory

Ganesh Kumar Thakur¹, Bandana Priya²

¹Krishna Engineering College,
95, Loni Road, Mohan nagar, Ghaziabad, Uttar Pradesh 201007, India
²GL Bajaj Institute of Technology and Management,
Plot No.2, Knowledge Park III, Greater Noida, Uttar Pradesh 201306, India

(Received 18 June 2019; Revised 26 November 2019; Accepted 26 November 2019)

The visibility graph is a set of locations that lie in a line that can be interpreted as a graph-theoretical representation of a time series, while the fuzzy graph speaks about the connection between the lines by accurately demonstrating the level of the connection between the objects of a given set. Many graphs do not show proper previous values. Even knowing the previous values of times series, prediction of the future values will not be accurate. Therefore, to find the real values exactly, this paper introduces the Visibility graph by time series values \((x_t, y_t)\), \((x_u, y_u)\) along with the Fuzzy node values \(f_1, f_2, \ldots, f_n\). Considering the past nodes by Fuzzy logic, similarity does not give a more accurate prediction because the nodes similarity contains the past node values only. The fundamental target of this paper is to propose a calculation to predict a more exact strategy to measure information by finding the similarities of all fuzzy nodes \(f_1, f_2, \ldots, f_n\) with their distance function \(f_d(\alpha)\) and the connectivity function \(\alpha\). The results of the computational outcome \(Y(x+1)\) will demonstrate more accurate values of time series.

Keywords: time series, fuzzy logic, fuzzy distance, connectivity function, visibility graph.

2010 MSC: 03B52, 05C72

DOI: 10.23939/mmc2020.01.014

1. Introduction

Graph theory \(G\) is the investigation of graphs, which are arithmetical structures used to exhibit the pairwise connections between objects \((V, E)\). A graph might be undirected, implying that there is no refinement between the two vertices related to each edge, or its edges might be guided starting with one vertex then onto the next. When there is an irregularity in the delineation of the graph, it is typical that we need to design a Fuzzy Graph Model \(G: (\sigma, \mu)\). Fuzzy graphs [1] have numerous more applications in displaying continuous frameworks where the level of data intrinsic in the framework fluctuates with various levels of exactness. The continuous framework values in the Fuzzy graph may be in Time series.

Time series is an arrangement [2] of numerical esteems that portrays the measurable list of one certain marvel and is requested in chronological. In this paper, the time series values are represented as \((x_t, y_t)\), \((x_u, y_u)\). During our everyday life, numerous information exists as time series, for example, air temperature [3]. To enhance the exactness of estimations, early mathematicians created [3] models like Exponential Smoothing (ES) and Holt-ES. With the improvement of likelihood and measurements, various factual strategies have been advanced and they can refine forecasts in a large degree, for example, Auto-Regression (AR) and moving average (MA) [4]. In light of ES strategy and ARIMA models, a software bundle is produced, which can gauge time series automatically and proficiently [5]. However these current techniques have a decent execution, they are not sufficiently exact for all the mind-boggling circumstances, as a result of entangled information structure and expanding demand for accuracy. Therefore, our research will calculate the more accurate results for time series without the errors occurs during the transformation of the network.
To avoid the errors occurred by the random walks in a network, this paper introduced link prediction [6, 7] in a visibility graph so as to figure the conceivable future information. By and by, because of the way that vulnerabilities exist unavoidably, incorrect estimations happen from time to time. Inspired by GET, analysts have overseen vulnerabilities in crisis administration [7] human unwavering quality [8] and etymological basic leadership [8]. Another doable hypothesis to adapt to vulnerabilities is fuzzy rationale that is first proposed by Zadeh. Fuzzy rationale [9] has been broadly connected in both hypothetical research and designing field, for example, Dijkstra calculation design acknowledgment [10], stack recurrence control and flight network’s vulnerability [11].

More importantly other paper shows the predictable values using visibility graph with fuzzy graphs with two-node similarity. But the values taken are included only with two nodes; the past values and the distance similarity are omitted. So, as a renovation, this paper proposed Fuzzy distance function \( f_d(\alpha) \) and connectivity function \( \alpha \) [12]. Considering the advantages of fuzzy logic along with their distance and connectivity value, the result of \( Y_{(x+1)} \) computed in this paper which predicts the future value more accurate.

Fig. 1. Representation of real visibility nodes \((x_1, y_1), \ldots (x_n, y_n), (x_{n+1}, y_{n+1})\) and the Fuzzy nodes \(f_1, f_2, \ldots, f_n\).

2. Related researches

Broumi S., et al. [13] have presented the idea of neutrosophic graphs with bipolar single valued in 2016, demonstrated that the most broadly utilized expansions of fuzzy graphs are specific instances of bipolar single-valued neutrosophic graphs. The investigation of certain sorts of bipolar single-valued neutrosophic graphs, for example, entire bipolar single-valued neutrosophic graphs, solid bipolar single-valued neutrosophic graphs, standard bipolar single-valued neutrosophic graphs. The idea of vitality of bipolar single-valued neutrosophic graphs. They learn about applications, particularly in rush hour gridlock light issue.

Anbalagan T., Maheswari S. U. [14] explained a Fuzzy Meta graph (FM) based securities exchange basic leadership, grouping and forecast are explained and now speculators of Indian securities exchange. The trial comes about, got by running on datasets taken from BSE India, demonstrates that their strategy performs superior to different techniques. Their examination is only a start and the long haul objective is to foresee the pattern of the value variety by including different compelling elements, for example, specialized investigation and central investigation. As a result, the framework can be additionally connected for the day by day exchanging reason.

Cai Q., et al. [15] have portrayed another fuzzy time series model joined with subterranean insect colony optimization (ACO) and auto-regression. To address the downside of the regular high-arrange fuzzy time series model, an AR high order fuzzy time series model is explained in this paper. While embracing the idea of the established AR model, the novel high request model makes a more viable utilization of the chronicled information and has been turned out to be more reasonable for useful utilization. The paper likewise explains another heuristic technique in light of ACO calculation to parcel the universe of talk, which had been the principal issue of fuzzy time series models.
Doucoure B., et al. [16] have depicted to develop a prediction method for renewable energy sources in order to achieve an intelligent management of a microgrid system and to promote the utilization of renewable energy in grid-connected and isolated systems. The depicted method is based on the multi-resolution analysis of the time-series by means of Wavelet decomposition and artificial neural networks. The analysis of predictability of each component of the input data using the Hurst coefficient is also explained. In this context, using the information of predictability, a prediction method based on wavelet networks using MRAWNN was implemented and evaluated by numerical simulations using real wind speed profiles. The analyzed prediction method includes the multi-resolution and predictability analysis of the components of the studied time-series with the aim of optimizing the prediction algorithm.

Sun B., et al. [17] have established a multivariate model and improves the accuracy of computation. By combining traditional fuzzy time series models and rough set method, used fuzzy c-mean algorithm to make the data into discrete. Furthermore, they deal with the rules in mature modules of the rough set and then refine the rules using data mining algorithms. This paper refines the rules obtained through the rough set. The paper uses Predictive algorithm to get two rules. It also might be used in empirical tests in other price predictions, such as stock prices, real futures.

Kuncheva Z., Montana G. [18] have depicted Multiplex networks, a special type of multilayer networks, are increasingly applied in many domains ranging from social media analytics to biology. They explained a community detection algorithm, LART (Locally Adaptive Random Transitions), for the detection of communities. The algorithm is based on a random walk on the multiplex, and the transition probabilities defining the random walk are allowed to depend on the local topological similarity between layers at any given node so as to facilitate the exploration of communities across layers. One advantage of LART is that it requires the definition of only one parameter $t$ which determines the length of the random walk.

Stephen M., Gu C., Yang H. [19] have depicted a method to convert a mono-variant time series to a weighted network of networks while preserving the local network behavior by using the visibility graph. By this, one can distinguish transformative practices of the framework. Beginning from the stage space reproduction method, the arrangement fragments with a predefined length are changed over to visibility graphs as being portrayals of the states comparing to the time interims. Connecting progressively happening states prompts a state exchange system of recognizable states. The center points, themes, and nontrivial designs (solid connections, circles, and so forth.) between the hubs disclose to us the exchange probabilities which are useful in prompt expectation. The happening places of the themes demonstrate the long haul practices that are helpful in naturally visible expectations.

3. Terminologies

3.1. Visibility graph

Visibility graph $G(V, E)$ is a graph of intervisible locations, [20] typically for a set of points and obstacles in the Euclidean plane. Each node in the graph represents a point location, and each edge represents a visible connection between them. That is, if the line segment connecting two locations does not pass through any obstacle, an edge is drawn between them in the graph.

3.2. Fuzzy graph

A fuzzy graph is a pair $G: (\sigma, \mu)$ where $\sigma$ is a fuzzy subset of $S$, $\mu$ is asymmetric fuzzy relation on $\sigma$. The elements of $S$ are called the nodes or vertices of $G$ and the pair of vertices as edges in $G$. The underlying crisp graph of the fuzzy graph $G: (\sigma, \mu)$ is denoted as $G^*: (S, E)$ where $E \subseteq S \times S$. The crisp graph $(S, E)$ is a special case of the fuzzy graph $G$ with each vertex and edge of $(S, E)$ having degree of membership 1.
3.3. Local random walk

A random walk is most conveniently represented by its transition matrix $P$. $P$ is a square matrix denoting the probability of transitioning from any vertex in the graph to any other vertex. Formally, $p_{uv} = p_r$ (going from $u$ to $v$, given that we are at $u$). Thus for a random walk,

$$p_{uv} = \frac{1}{d_u} \quad \text{if} \quad (u, v) \in E,$$

and 0 otherwise (where $d_u$ is the degree of $u$).

3.4. A contrivance of fuzzy logical time series prediction in visibility graph

Visibility graph $G(V, E)$ analysis could be applied to almost all types of time series $(x_1, y_2), \ldots, (x_n, y_n), (x_{n+1}, y_{n+1})$ which is a fast computational method that converts a time series into a graph. For example a data structure can be designed in the form of tree which in turn utilized vertices and edges. Similarly modeling of network topologies can be done using graph concepts. The representation of series with graph are, the periodic series convert into regular graphs, random series convert into random graphs, and fractal series convert into scale-free graphs. Thereby, we can investigate the characteristics of time series by means of new approach. In many cases, the value of time series invisibility graph should not in a regular manner. At that scenario, the prediction of the time series will be more complicated. So as to rectify the problems, we introduced the Fuzzy Graphs with distance and connectivity functions. Fuzzy graphs contain a set of values which was the subset of visibility values. While arranging this graph into network errors are occurred due to various causes. Thus to avoid the errors occurs during random walk, the nodes in the graph are connected by a link prediction. So that we can obtain a nearly exact real value. By obtaining the real and fuzzy values, here we proposed the computational value by introducing the Fuzzy connection $f_d(\alpha)$ and distance Function $\alpha$ which predict more accurate future values as shown in the results.

3.5. Formulation of visibility graph theory

Visibility graph has established itself as a powerful tool for analyzing time series [21]. For a given multivariate time series, we first perform a coarse-grained operation to define temporal scales and then reconstruct the multivariate phase-space for each scale to infer multi-scale complex networks [22]. In addition, develop a novel clustering coefficient entropy to assess the derived multi-scale complex networks [23] aiming to characterize the coupled dynamical characteristics underlying multivariate time series.

**Definition 1.** For any two values in the time series $(x_t, y_t), (x_u, y_u)$ will have Visibility and consequently connected with two nodes of the associated graph and if any other data $(x_v, y_v)$ placed between fulfills by

$$(y_v < y_u) + (y_t - y_u) \left( \frac{x_u - x_v}{x_u - x_t} \right).$$

**Theorem 1.** If $\{x_i\}, i = 1, \ldots, N$ be a time series of $N$ node. The two node $i$ and $j$ in the graph are connected. Then the two nodes $i$ and $j$ satisfies

$$y_i, y_j > y_n, \quad \forall n | i < n < j$$

**Proof.** Let the time series $(x_t, y_t)$ can represent as $x_t = \{x_i\}, i = 1, \ldots, N$ the adjacency matrix is given by

$$A = (a_{ij})$$

Assume that the value of $a_{ij} = -1$, if the value of $x_i$ and $x_j$ are connected horizontal. The elements in the matrix $a_{ij}$ is given as $P$,

$$p(a_{ij}) = \frac{e_{ij}}{k_i} \quad \left\{\begin{array}{ll}
1 \leq i \leq N; \\
1 \leq j \leq N;
\end{array}\right.$$
where $i$ represents the departure node and $j$ represents the arrival node. Thus the node satisfies by the element
\[ p(a_{ij}) \text{ as } x_i, x_j > x_n, \quad \forall n | i < n < j \] (6)
Hence equation (3) verified.

### 3.6. Link prediction method

By positioning the values in the graph of equation (2) as a network some errors are occurs due to various causes. The major causes are random walk. So the random walk also calculated as the superposed random walk by Link prediction method [24]. Link prediction positioning measures assess the normal neighborhood of a couple of node in a system, with a specific end goal to evaluate the probability of another link. Then again, a similar zero rank esteem is given to node sets with no basic neighborhood [25], which generally are an expansive number of possibly new links, in this way bringing about exceptionally low-quality general link positioning regarding normal alter separation to the ideal rank.

**Definition 2.** Consider a random walk on the space $U$ which starts from the node $i$ and iteratively move to the neighboring node $j$, the chosen uniformity from the random $\Gamma(x)$. Then the Hitting time $HT_{i,j} = -H_{ij}$ and the Commute time $HT_{i,j} = -C_{ij}$ where
\[ C_{ij} = H_{ij} + H_{ji} \] (7)

**Definition 3.** The link prediction for the two nodes is computed using the values of element in $p(a_{ij})$. The probability transfer matrix $p(a_{ij})$ is used to represent the probability that a random walker departs from one node and arrives at another node.

**Lemma 2.** In initial state, Random walk located at the node $i$ is represented by the $N \times 1$ row vector and random walk done by the node $i$ is given as $\lambda_i(0)$ and the node $j$ is $\lambda_j(0)$ at its initial state. The similarity of the nodes can be expressed as superposing Random walk
\[ S^*_{ij}(t) = \sum_{l=1}^{t} S^\theta_{ij}(t)(l). \] (8)

**Proof.** Let, $N \times 1$ is the row vector of the matrix $a_{ij}$ at the node of $i$ and $j$, $E$ be the total number of edges, $K$ be the degree of the two nodes the initial walk of the node $i$ is given as $\lambda_i(0)$ and the node $j$ is $\lambda_j(0)$. After $t$ time, the probability of the random walk reaches the is given as,
\[ \lambda_i(t) = p^T(a_{ij}) \lambda_i(t - 1), \] (9)
\[ \lambda_j(t) = p^T(a_{ij}) \lambda_j(t - 1). \] (10)
By the equation (3), (9) the Similarity of the node by the random walk can be expressed by,
\[ S^\theta_{ij}(t) = \frac{K_i}{2 |E|} * \lambda_i(t) + \frac{K_j}{2 |E|} * \lambda_j(t). \] (11)
The superposed Random walk with the 8 length value $l$ can be calculated by the equation (10) as
\[ S^*_{ij}(t) = \sum_{l=1}^{t} S^\theta_{ij}(t)(l). \] (12)
Thus, verified in equation (12).
decision making and have broader applications. Experimenting with random walk also shed light on
the transition from classical to fuzzy probability.

3.7. Modernism of fuzzy logical distance and connectivity functions

Fuzzy logic is determined on the ideas of fuzzy sets [3, 5, 6, 15, 26] and emblematic logic. In fuzzy
logic and its applications to exchanging frameworks and some properties of fuzzy logic, the discourse
of fuzzy logic was constrained to propositional math. We may see fuzzy logic as an uncommon sort
of many-esteemed logic (2, 1, 2). In fuzzy logic, the truth-estimation [26] of an equation, rather than
accepting two esteems, (0 and 1), can expect any an incentive in the interim [0, 1] and is utilized to
demonstrate the level of truth spoke to by the technique. In Fuzzy transform of a function from
$X$ is defined with respect to a Fuzzy partition of its domain i.e. the interval $(i, j)$. Fuzzy set on $(i, j)$ are
identified with their member function. They are mapping from $(i, j)$ into $(0, 1)$. If $I$ is a Fuzzy set in
$(i, j)$ then we get $I \subseteq (i, j)$.

Definition 4. Let $U$ be a space of points involves in the Fuzzy set by $x$ i.e. $X = \{x\}$. A Fuzzy set
$X$ is characterized by the member function $\mu_X$ which maps from 0 to 1. The Fuzzy set $A$ can be given
as,

$$A = \{x, \mu_X(x) \mid x \in U\}.$$  \hfill (13)

Definition 5. A Fuzzy graph $G$ is defined by

$$G = (x, y), \quad x = \{x_1\}, \quad F = (f_{ij}), \quad 0 \leq f_{ij} \leq 1.$$  \hfill (14)

The fuzziness of the node $i$ and $j$ is denoted as $f_i$ and $f_j$. The similarity relation between the node
is calculated from the symmetric relation matrix $S$ by

$$S = \begin{bmatrix} S_{ij}, 2 \frac{S_{ij}}{S} \end{bmatrix} = \frac{1}{f_{ij}} + \frac{1}{f_{ji}}.$$  \hfill (15)

Lemma 3. If $f_d(\alpha)$ be the fuzzy distance function and $\alpha$ be the fuzzy connectivity function, then
the value of final prediction should be given by the relation of fuzzy real number $\hat{X}$ and superposed
link prediction as

$$Y_{x+1} = S^*_ij(t) \ast y_{x-1} \ast |f(\lambda)| \ast y_x \ast \hat{X}.$$  \hfill (16)

Proof. Let the fuzzy set $I = \{(x, y), \mu_I(x, y) \mid (x, y \in X \times Y)\}$. If the number of rules in fuzzy logic
is combined, the centre of length of fuzzy set $K^{-1}$, height of the set $\omega$ then the real value should be

$$\hat{X} = \sum_{m=1}^{N} K_{m}^{-1} \omega_m \sum_{l=1}^{N} \omega_l.$$  \hfill (17)

The nodal distance of the fuzzy set be

$$f(\lambda) = (f_d(\lambda), f_e(\lambda)),$$ \hfill (18)

where $d$ represents the distance from visibility node $e$ represents the connectivity function of the
visibility node. The value of the fuzzy connectivity function is

$$f_d(\lambda) = \frac{d_i - d(\lambda)}{d_i - d_j},$$ \hfill (19)

the value of fuzzy distance function is

$$f_e(\lambda) = \frac{e(\lambda) - e_i}{e_i - d_j}.$$ \hfill (20)
By the above-said equations, the final prediction can be computed by the equation (11), (15) as

\[ Y_{(x+1)} = S_{ij}^*(t) * y_{x-1} + |f(\lambda)| y_x * \hat{X} \]  

(21)

Thus the final value \( Y_{(x+1)} \) of time series will be predicted by the fuzzy distance and connectivity function.

4. Conclusion

Time series using Fuzzy logic is a fast-growing area of research and as such provides much scope for future works. One of them is the Combining Approach, i.e. to combine a number of different and dissimilar methods to improve forecast accuracy. In this paper, we combine Visibility graph theory and Fuzzy logic with link prediction along with fuzzy connectivity function \( f_d(\lambda) \) and fuzzy distance function \( f_e(\lambda) \). As a result, the overall equation portrays the precision of time series values more accurate than any other method.
A novel computation for predicting time series using fuzzy logical distance . . .


