Time Series Social Network Visualization Based on Dimension Reduction

Maha Al-Ghalibi
Computational Visualistics
University of Koblenz-Landau
Universitätsstr.1
56070 Koblenz,
Germany.
alghalibi@uni-koblenz.de

Adil Al-Azzawi
Electrical Engineering and
Computer Science, University
of Missouri-Columbia
USA (65211), MO, Columbia
aaadn5@mail.missouri.edu

ABSTRACT

Social networks are in general dynamically due to the involvement of many people on the web such as Facebook, Twitter, and Snapchat, etc. The meaningful visualization and analysis of social network is challenging due to its dynamic nature, the mobility of nodes in the network and extremely large size. In this paper, we consider the higher dimensionality issue of social networks regarding time series social network construction and visualization. To solve this issue, we develop a statically data-mining based approach for dimensionality reduction in social networks. Basically, we find that each sub-social network's model has different dimensions by nodes and links which are sampled originally from an m-dimensional metric space. Experimentally, we find that the m-dimensional features for each sub-network cause fail connections in time-series during the network reconstruction model for visualization. Therefore, we propose a new dimension reduction approach that is based on developing an SVD algorithm by relying on select significant sub features. Then we extract time features from the feature space of the original dataset to visualize the network in a deferent time interval. However, to monitor the network development and also the dimensionality reduction of features help us to speed up the computation time of the shortest path. The social circle Facebook dataset form Stanford is used with its corresponding attributes. The dataset includes node features (profile), circles, and ego networks. The obtained result shows better performances regarding the computation time and network visualization. Moreover, the experimental results show that the proposed system is much faster than the approach based on the whole feature space for closeness centrality computing.

Keywords

Network Visualization, SVD, Mutual Information, Dimensionality Reduction, Feature Selection.

1. INTRODUCTION

In recent years, the use of social networks has become a robust online communication platform supporting millions of users. User behavior in a large-scale data needs visualization and as well as data analysis methods to achieve higher performances [WCG+16]. This paper is focused on clustering to solve the challenging issues of visualization. A dynamic graph discretization and graph clustering used in presenting a hierarchical structure [GSZ+11]. Connected components of are characterized as a Depth First Search (DFS) tree that updates the dynamical changes in graphs. Clustering in dynamic social networks is applied to tolerate any changes that occur in the network. Dynamic changes in the social network are analyzed by monitoring the structural characteristics regarding patterns [Wu10]. This article uses the Top-k Weighted Clustering Coefficient supported maintenance of constructed a graph by performing vertex/edge insertion, vertex/edge deletion and monitoring top-K results [LCZ+17]. Insertion of a new edge or vertex involved without disturbing other vertices and edges. Also, the periodical graph is used that based on betweenness centrality to provide pseudo-polynomial time algorithm [FS15]. Here the edge weights are determined for predicting the shortest path for communication between nodes. For analyzing this type of network, a graphical representation involved in most of the research works. The taxonomy-based graph representation is introduced to analyze, i.e. Network topology, visualization, community detection, [ARK13] the work depends on centrality-based metrics that are computed as eigenvector, closeness centrality, betweenness centrality, page rank, hubs, and authority. Non-centrality-based metrics are reciprocity, transitivity, density, or similarity. Structural similarity-based dynamic network layout visualization is involved in identifying temporal changes in the network [XH16]. A single time slice method is used to determine node distances from which the node's moved distance estimated. The node's best mobile area identified, and then the adjustment is done among the nodes. The currently created network layout compared to the previous layout for predicting similarity to the final layout. Changes in the structure defined by the chang

ISBN 978-80-86943-42-8
https://doi.org/10.24132/CSRN.2018.2803.5

centrality metric that enables pairwise comparisons in the evolving network [FPA+12]. They have presented a set of novel metrics for the visual analysis of dynamic networks. Trying to enhance the perception of changes and the gaining of both, overview and detailed insights on the network growth. [CZ17] In this framework, the nodes are matched, and community is assigned. Then the edges of the nodes are updated. In [BJ15] a Time-Varying Graph model for Online Social Networks (TVG-OSN) is presented. An identity map is created at time intervals to analyze the link occurrences among nodes. TVG-OSN supports the determination of user characteristics using different link criteria. A temporal graph model plays a significant role in dynamic social networks. Therefore, the dynamic network environment has been involved with certain challenging issues, and many challenging issues are discussed by many authors that evaluated their performance metrics.

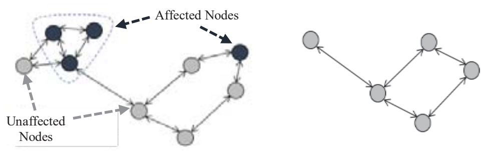
We developed visualization and sentiment analysis approach which first visualize the Facebook dataset in a different time series to show the changes in topology. For time series visualization, the proposed system allows the viewers to visualize the whole data of the Facebook social network in different categories depending on five different time series. We give a statistical visualization of the (total number of nodes, and connections) that has been significantly changed during the growth of the whole social network through the time. After time series visualization, we propose a novel approach for dimensionality reduction called Guided Dimensionality Reduction Approach (GDRA). GDRA solves the problem of dimension reduction by applying a mathematical model to avoid the biased feature selection. We suggest a threshold selection of the significant subfeatures that are ranked by Mutual Information. Moreover, we developed a new version of the Singular Value Description (SVD) by using a multi eigenvalue selection using SVD.

This approach helps our system to reconstruct the sub-network in the time series after removing the non-relevant feature space which causes the NAN connection during the visualization step. Moreover, it potentially reduces the time that has significantly consumed in the short path detection for the visualization of the experimental results. Our last contribution is the construction of a dynamic network model that manages the arrival of the new nodes and departure of nodes that were participating in the network as additional and optional view of our visualization system.

2. BACKGROUND THEORY

Social network analysis and visualization is based on the number of nodes and their connections to the whole entire network. Different nodes connections in a social network can affect the network visualization. Moreover, the fact that affects the decision of the other nodes on drive subject matters. Different influences regular connection node (unaffected nodes and affected nodes) based on their connection types such as direct connection, no connection, or either bi-connection during a specific time in the time series visualization task. This is challenging task in a social network visualization especially the time series task due to the difficulties of reconnect the other sub-network together based on their node connections [WU10].

Time series visualization can cause many issues especially for a complex social network such as the threshold in which the connection nodes can be under the influents decision at that period. Data mining approaches, such as clustering and dimensionality reduction techniques, can be used to model this problem in the social network visualization which is a way of asserting the affected nodes and unaffected nodes [XH16]. For such a visualization task, during time series, there is a decision about whether the influential participant's connection nodes that are either connected or not connected depends mainly on the original connection aspect. This connection is originally constructed based on the number of features that the whole network based on to draw the whole space. In this case, there is a feature vector that does fairly cause the affected influence connection based on the number of NAN this feature vector has. In this case, we assume that the dimension reduction of the feature space may help to solve this kind of complexity. Figure.1 shows an example of the affected and unaffected connections nodes in our situation where in this case, the unaffected node that originally connected to the affected rejoin should be reconstruct again to the unaffected rejoin (nodes).



(a) **(b)**

Figure1: Time series social network visualization issue (a) for the whole network, (b) specific time series

Data Mining Dimensionality Reduction and Feature Selection

The ageneral problem of data mining and machine learning approaches is to handle a high- dimensional data (feature space) due to the huge number of input variables. Dimensionality reduction and feature selection can be made in two ways, By using the feature selection techniques were in this way, keeping the most relevant data (variables) in the original data, or by using the dimensionality reduction techniques where in this case exploiting the redundancy of the input data and by selecting a subset of new data (new variables) [SVM14].

ISBN 978-80-86943-42-8

Singular-Value Decomposition (SVD)

Singular value decomposition (SVD) allows an exact representation of any matrix and makes it easy to eliminate the less important parts of that representation to produce an approximate representation with any desired number of dimensions. Let us assume that X is an $m \times n$ matrix and r is the rank of X. Recall that the rank of a matrix represents the largest number of rows (or equivalently columns) for which a non-zero linear combination of the rows is the all-zero vector 0. The Singular-Value Decomposition (SVD) is described in the Algorithm (1) below [SVM14].

Algorithm: Singular-Value Decomposition

Input: Generate data matrix $\boldsymbol{X}$

Output: New dimensions $\boldsymbol{C}$

1. **Repeat**
2. **Applying** SVD to the matrix X as $X = USV^T$
   $X \rightarrow$ is an $m \times n$ matrix ($m \rightarrow$ no. of vectors)
   $n \rightarrow$ is no. of attributes)
   $U \leftarrow XX^T$ matrthe ix of the eigenvectors
   $S \rightarrow$ is a matrix which is diagonal
   $V \leftarrow$ is a matrix of the eigenvectors.
3. **Construct** the covariance matrix from this decomposition by
   $XX^T\ XX^T \leftarrow (USV^T)(USV^T)^T$.
   $V \rightarrow$ orthogonal matrix ($V^T V = I$).
   $XX^T \leftarrow US2U^T$
4. **Compute** the square roots of the eigenvalues of $X\boldsymbol{X^T}$ which are the singular values of X
5. **Until** representing every transaction at vector $x(t)_i$
6. **Return** $U^T X$
7. **End**

Algorithm1: Singular-value decomposition (SVD)

Mutual Information

Mutual information is a statistical method that measures the relationship between two random variables that simultaneously sampled. It measures how much information has each variable about another. Intuitively, it asks how much each random variable tells us about another one [ZB16]. The formal definition of the mutual information of two random variables X and Y, whose joint distribution is defined by $P(X, Y)$.

Algorithm (2) shows how the MI is computed based on using two random variables [ZB16].

Algorithm: Mutual Information (MI)

Input: Feature Space $\boldsymbol{X}$

Output: Mutual Scoring $\boldsymbol{I}$

1. **Repeat**
2. **Compute** the MI for every two variables $\boldsymbol{X}$ and $\boldsymbol{Y}$ by $\boldsymbol{I(X,Y) = \sum_{x \in X} \sum_{y \in Y} P(x,y) log \frac{P(x,y)}{P(x)p(y)}}$
3. where
   $X \rightarrow$ is the first feature space
   $Y \rightarrow$ is the second feature space
   $P \rightarrow$ is the probability function
4. **Import** the two variables in the feature space
5. **Until** consumed all variable in the whole variables
6. **Return** the mutual scoring vector
7. **End**

Algorithm2: Mutual information (MI).

3. PROPOSED SYSTEM

The pipeline of our proposed system shown in Figure.2

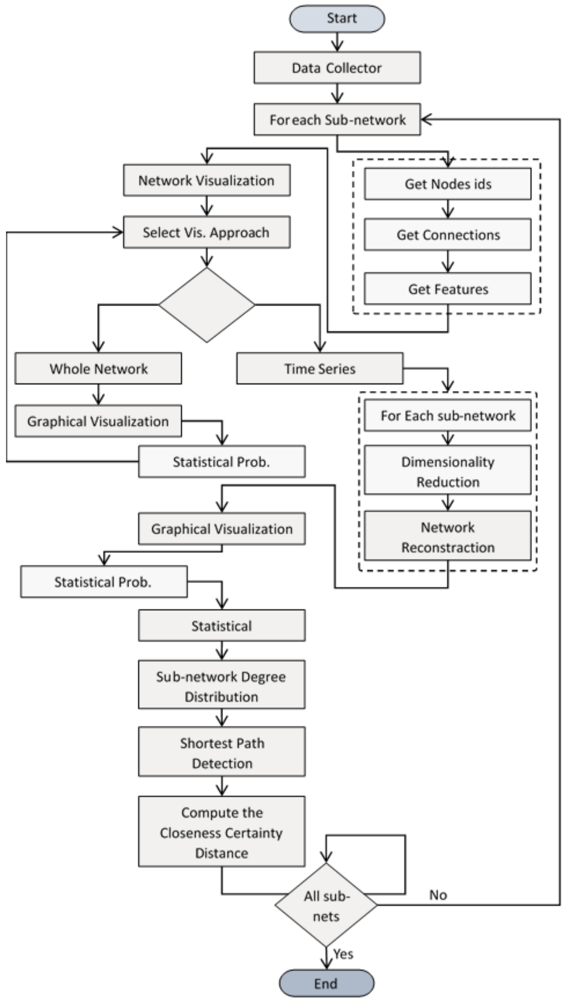


Figure2: Pipeline of Our proposed visualization approach

Any social network model has attributes such as a specific number of nodes (users) and link (connection). The network attributes allow the system to fit the best visualization model based on the dimensional metric space which based on the power-law distribution. Based on the logarithmic relational between the m dimensional scale and the number of users (nodes), we proposed a network visualization model, see Figure 2. It is based on the dimensionality

ISBN 978-80-86943-42-8

reduction approach for social network hypothesis confirming.

The main idea of our model is to reduce the original feature space which is performed in two steps. We using the mutual information (MI) which basically ranks the original features space from 1 to 0. Then, a sub-correlated features set is selected based on the pre-defined threshold. For the feature set selection, we introduce a new algorithm to perform dimensionality reduction (SVD). With this, We select most relevant and correlated sub feature space. The subset used within the model to reduce the unnecessary connection in the social network model according to the NAN connection that has already collected during the network constriction. This is a time-consumption process for whole network dimension, especially for detecting the distances measured by the shortest paths to compute closeness centrality.

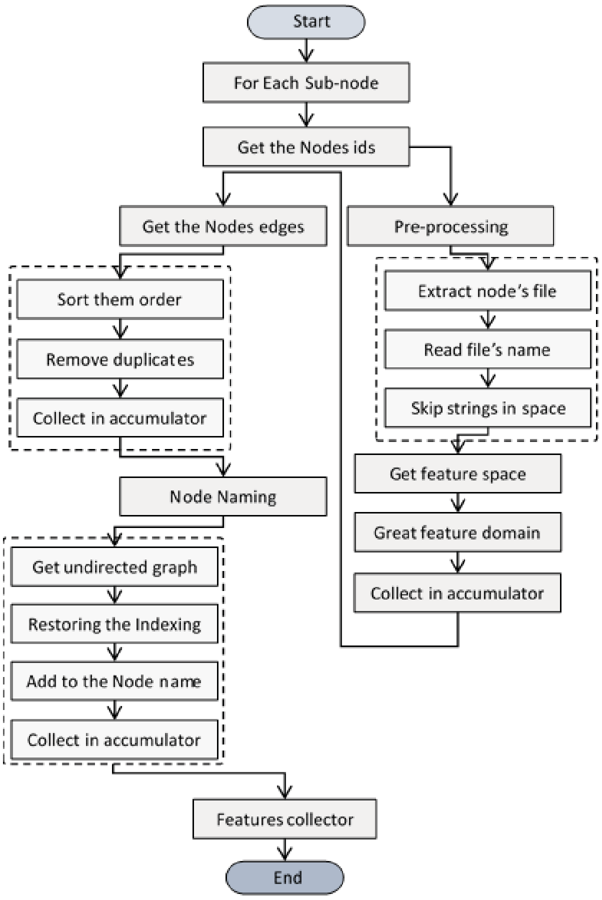


Figure3: Data collector and feature extractor

The visualization of time-series of social networks based on dimensionality reduction consists of two main approaches. The first approach is shown in Figure.3 it comprises the data collection and feature extraction for the whole network including ten subnetworks of the whole Facebook dataset.

The second approach is the visualization of time-series of the network. Figure.4 shows the steps of our approach to select the time series of social networks as well as to visualize whole networks.

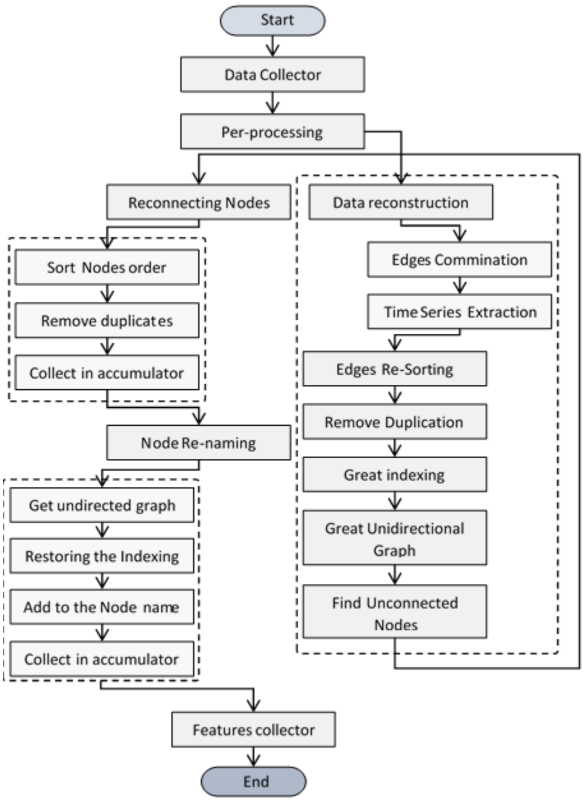


Figure4: Time series network visualization

Sub Feature Selection based Non-Negatively MI.

Our main contribution is a new dimensionality reduction approach. Our approach comprises two steps. The first one is the sub-feature selection based on MI, and the second step the SVD approach is applied. The SVD is based on the multi-Eigen value selection which is a new method for dimensionality reduction in data mining and machine learning.The principle idea is that we want to find a non-biased threshold for the subfeature selection. For this purpose, We developed a mathematical model that determines a significant threshold for feature selection based on the MI scoring. For this, We determine the uncertainty of MI scoring-based feature selection. Therefor we select the positive scoring produced by the MI. Afterward, We normalize the whole feature space. To ensure the non-negativity selection for the MI. We assume that the MI measures the inheritance dependence expressed in the joint distribution x and y which is related to the joint distribution of x and y under the assumption of independence. In this case, MI measures the dependency $I(x, y) = 0$ if x and y is independent random variables. That means in the case x and y are independent, the probability of x is represent by p [Pea01] [ZB16] as in equation 1.

$$(x, y) = p(x) \times p(y) \quad (1)$$

Applying the log to the probability results in:

$$log\left(\frac{p(x)}{P(x)P(y)}\right) = log(I) = 0 \quad (2)$$

 ISBN 978-80-86943-42-8

That means the MI scoring has non-negative values.

$$I(x,y) \geq 0 \quad (3)$$

Moreover, the MI scoring is asymmetric as shown in Equation 4:

$$I(x,y) = I(y,x) \quad (4)$$

After we ensure that MI is a non-negative value, we need to find the relation between the conditional and the joint entropy. MI can equivalently express by:

$$\begin{aligned} I(x,y) = H(x) - H(y) &\approx H(x) - H\left(\frac{y}{x}\right) \\ &\approx H(x) + H(y) - H(x,y) \\ &\approx H(x,y) - H\left(\frac{x}{y}\right) \\ &- H\left(\frac{y}{x}\right) \end{aligned} \quad (5)$$

where $H(x)$ and $H(y)$ are original entropies, and $H\left(\frac{x}{y}\right)$, $H\left(\frac{y}{x}\right)$ are conditional entropies.

Because of $I(x,y)$ is a non-negative value, $H(x) \geq H(x/y)$, because of

$$I(x;y) = H(y) - H\left(\frac{y}{x}\right). \quad (6)$$

The proof of this is illustrated in the following:

$$\begin{aligned} I(x,y) &= \sum p(x,y)\, log\left[\frac{p(x,y)}{p(x)p(y)}\right] \\ &= \sum p(x,y)\, log\left(\frac{p(x,y)}{p(x)}\right) \\ &- \sum p(x,y) log p(y) \\ &= \sum p(x) p\left(\frac{y}{x}\right) \log p\left(\frac{y}{x}\right) \\ &- \sum p(x,y) \log p(y) \\ &= \sum p(x) \left(\sum p\left(\frac{y}{x}\right) \log p\left(\frac{y}{x}\right)\right) \\ &- \sum \log p(y) \left(\sum p(y,x)\right) \\ &= -\sum p(x) H\left(\frac{Y}{X=x}\right) - \sum \log p(y)\, p(x) \\ &= -H\left(\frac{Y}{X}\right) + H(y) = H - H\left(\frac{Y}{X}\right) \end{aligned} \quad (7)$$

If the entropy $H(y)$ measures the uncertainty of the random variable, $H\left(\frac{Y}{X}\right)$ measures what X dos not say about Y. This means that the amount of reaming uncertainty of Y after X is known. Then, the amount of MI as the amount of information (that reduction in uncertainty) knowing either variables provide information about the other.

Our Approach for the Singular Value Description (SVD)

Our dimensionality reduction/feature selection approach depends on the regular SVD approach. Our method is based on ranking the eigenvalues first, followed by an accelerated singular value decomposition. For this purpose, multiple eigenvalues have been selected based on ranking step [Pea01]. Our function produces a diagonal matrix S of the dimension as the rank of X (whole data dimensions) with non-negative diagonal elements in decreasing order, and unitary matrices U and V with U represents the eigenvalues, and V represents the eigen vectors.

Algorithm: Our Singular Value Decomposition (SVD)

Input: Generate data matrix $\boldsymbol{X}$
Output: New dimensions $\boldsymbol{C}$
Determine the max matrix size
Determine the eigenvector ratio = 0.1
Compute the $\boldsymbol{S}$ matrix (compute data $\boldsymbol{X^t}.\boldsymbol{X}$)
  Check the optimally reduced dimension
    Compute the diagonal data matrix
    Find the max value of data $\boldsymbol{X}$
    Find the data size
    Check the reduced dimension
      Compute the diagonal eigenvalues
      Sort the value of the eigenvalue
      Get the index of the eigenvalue
      Get $\boldsymbol{U}$ eigenvalue matrix
      Select the max eigenvalue
    Get the index of the eigenvalue
    Replace the eigenvalue with the index
  Select the Eigenvalue
  Compute the unitary matrices $\boldsymbol{U}$
  Produce diagonal matrices of the dimension as the rank $\boldsymbol{X}$ and with non-negative diagonal elements in decreasing order $\boldsymbol{S}$
Find minimum half of the eigenvalue
Compute the unitary matrices $\boldsymbol{V}$ eigenvector

Algorithm 3: Our approach to compute the SVD

4. EXPERIMENTAL RESULTS

In this section, the obtained results of the proposed visual dynamic network evaluated. Our methods are implemented by using Matlab 2017a. For analyzing the proposed dimension reduction, topological observations, mainly a Facebook dataset used. Facebook is one of the most popular social networks that is widely used by millions of people. The used Facebook dataset is given from SNAP consists of 'circles' (or 'friend's lists') from Facebook. The dataset holds node features (profiles), circles, and ego networks. In our experiments, we evaluate our system according to different criteria such as visualizing the whole social network once with each social sub-network as well as the visualization of time series within social networks. Finally, we determine the time-consuming for computing closeness centrality in two cases. First, by using our dimension reduction approach on social network visualization secondly, without dimension reduction.

Whole Network Combination and Visualization

In this section, we visualize the whole social network by one single plot, see Figure.5 the whole network has many sub-networks (ten sub-networks).

 ISBN 978-80-86943-42-8

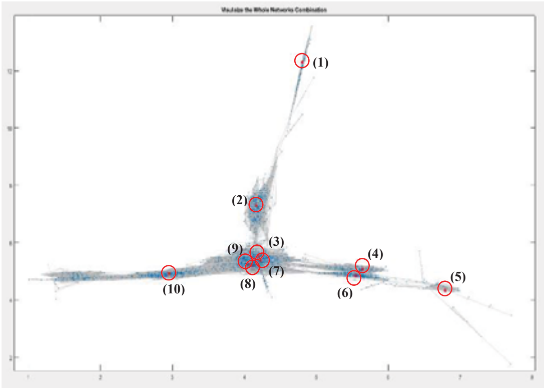


Figure5: Whole social network visualization

In this case, we combine all ten sub-social networks or (ego networks) into one visualization of the social network. The first plot shows that there are more than 4,000 nodes and more than 84,000 edges.

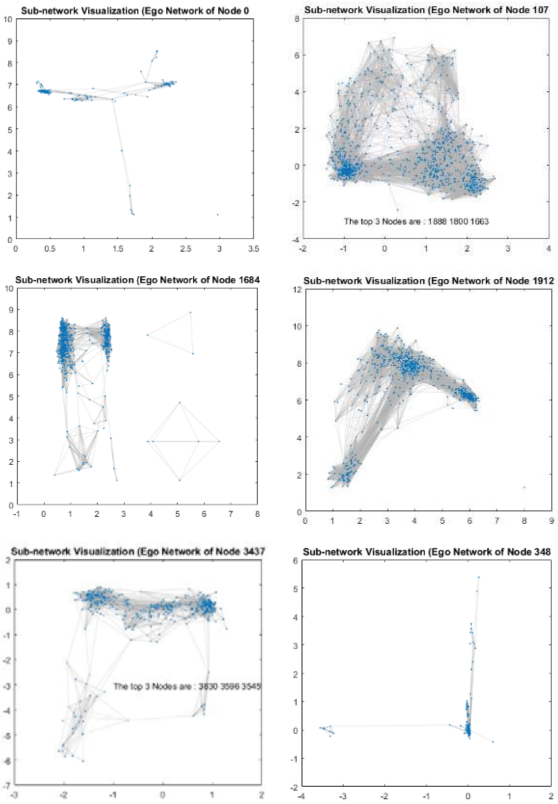


Figure 6: Individual sub-network

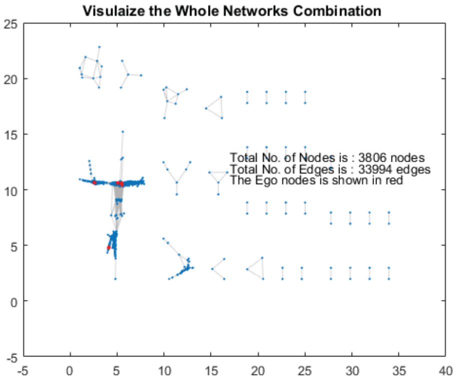
The ten sub-social networks or ego-node that has been marked by red dots

Secondly, every sub-social network (ego-nodes) has visualized individually. Figure.6 shows examples of six sub-networks out of ten. Each sub-social network is visualized.

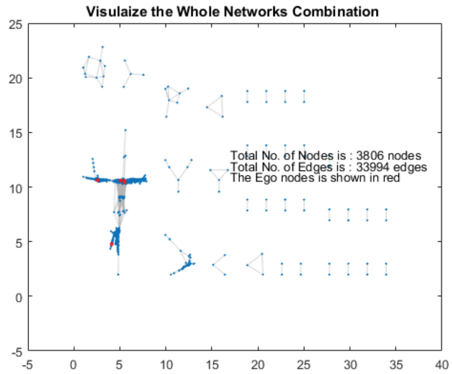
Time Series Social Network Visualization

One of Our main contributions in this paper are the an approach to visualize time series. Our aim to visualize effeteness of the time series overall social network configurations based on the number of connections and the total number of nodes at that time. For this purpose, we pre-processed the Facebook dataset by adding the acquisition time to the main CVS file. That means we collected four updates on the same network according to five-time periods 2008, 2009, 2010, 2011, and 2012. The outcome of our visualization system is a visualization of time series for social networks.

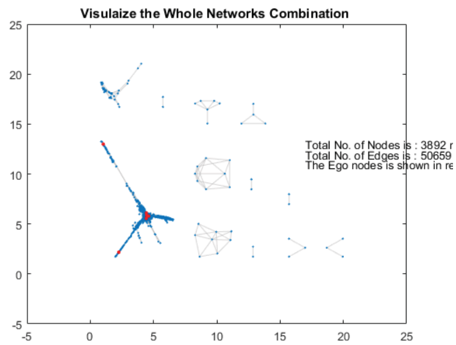
In this case, we performed a variety of Experiments to visualize every single time and the effeteness of this selection on the whole network. Figure.7 shows the effects of the time-series according to the total number of nodes in each selection as well as the total number of edges (connection) in each time selection.



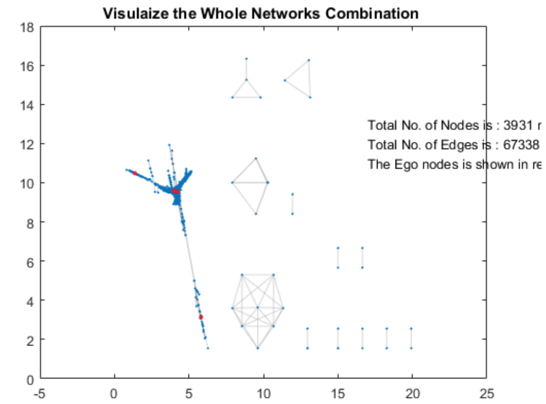
One-year Visualization 2008



Two-years Visualization 2008-2010



Three -years Visualization 2008-2009-2010



Four-year Visualization 2008-2009-2010-2011

Figure7: Time-series network visualization

Time-series network evolution

Figure.8 shows the main computation (total number of nodes and number of connections) that have changed in the meantime.

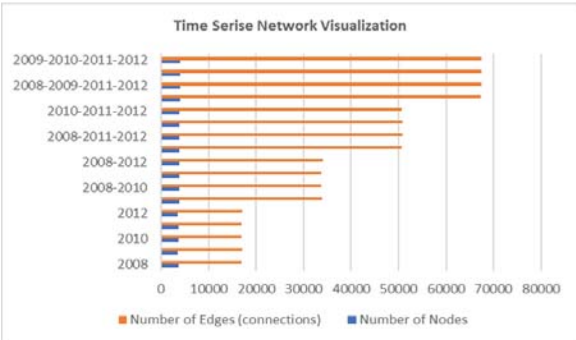


Figure8. A statistical measurement of time-series visualization

It has been shown here that the absolute visualization majority for each time-series that we selected in every single experiment. We can notice that the total number of nodes, as well as the total number of connections, are increased especially when more than one period was selected. For instance, the number of nodes in 2008 was 3583. Then it has been increased to 3940 when we select time series 2008-2012. Also, the total number of connections increased as from 16916 in 2008, to 67400 in the same period (2008 to 2012). This show that the number of connections and nodes is monotonically increased over the five years until 2012.

Degree centrality visualize for single sub-network

As a basic visualization evaluation technique based on the link analysis, this kind of visualization for each sub-network is the most basic analysis that can perform on a social network. With this experiential evaluation, we want to figure out how are the best connected to each network. For this purpose, we change the color of the connection based on the connection number in that sub-network which also represents the connection degree. However, for each sub-network, there is a metric of the connection degree that is known as a connection degree centrality. Figure.9 shows some example of six sub-networks out of ten that depict the centrality degree of each sub-social network.

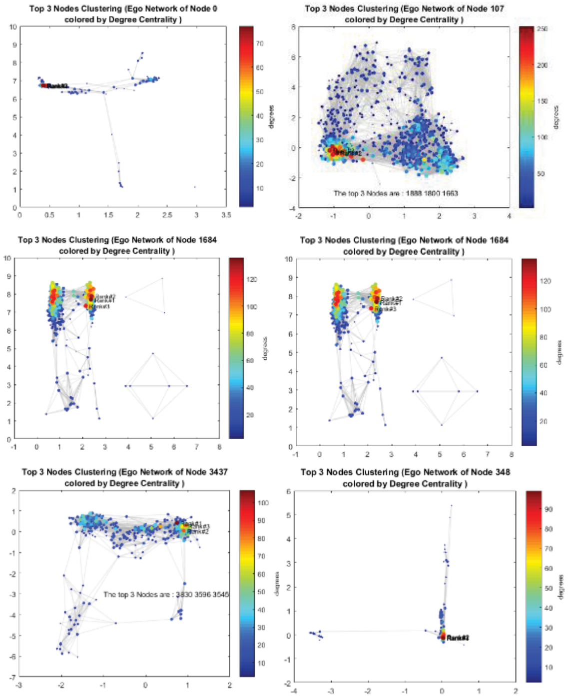


Figure9: Centrality degree visualization for each sub-network

The top three nodes that have been marked by the degree of connection centrality highlighted in the same plot. The nodes are closely connected to each other in the visualized network.

Sub-Social Network Degree Distributed Visualization

Another outcome of our visualization system, we visualize the degrees distribution between each sub-network and whole combined network. For this purpose, the histogram of the degree for each sub-network has been computed and compared with the original network (combined one). People who are active on the social network Facebook have a stronger edge distribution than others. Moreover, a few people have a large number of degrees in a social network such as a Facebook that we want to visualize in our exponential results. The majority, in this case, exhibits small number of degrees and a large difference in the edges distribution. This gives us an indication that is exponential. Figure.10 shows the individual sub-network no. (2) according to the degree centrality visualization and Figure.11 shows the histogram visualization of the degree distribution of the same sub-network no. (2).

ISBN 978-80-86943-42-8

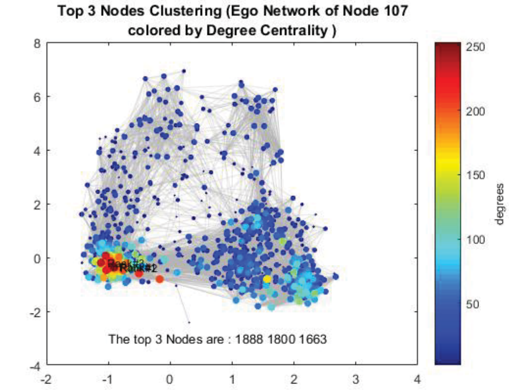


Figure10: Centrality degree visualization of sub-network no. (2)

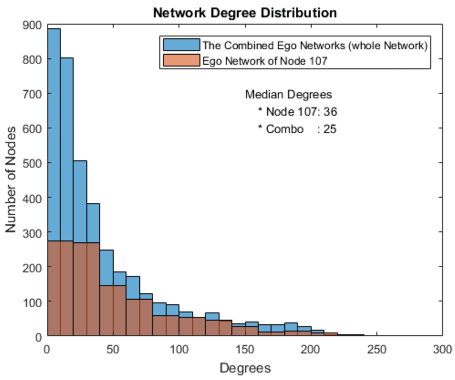
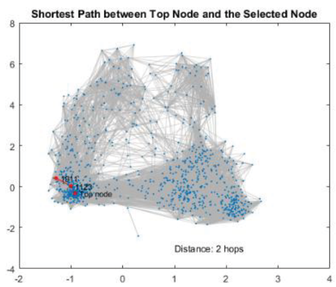


Figure11: Newtown degree distribution of sub-network no. (2)

We can notice that the median degree in the sub-network is less than the degree in the whole combine network. That is because the degree of the node connection in the subnetwork has been only computed based on the node in this individual sub-network. However, in this case, we do not visualize the other nodes that not connected to the other sub-network.

Shortest Path Visualization

The other visualization metric is the degree of the metric evaluation nodes which means more nodes have a better connection. That means the short path indicates how many hops exist in the metric. Figure.11 shows the short path between the top nodes that has been detected in the subnetwork no. (2) and the selected node number 1911, the selected node number describes how many minimum numbers of the node we need to reach our destination although Figure 12 shows the zoomed plot of the short path between the top node and the node number 1911.



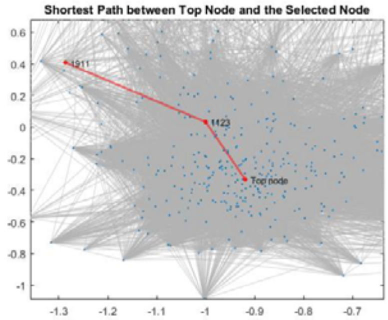


Figure12: Shortest path visualization

Closeness Centrality Visualization

The distances measured by shortest paths been shown in the previous section to be used to compute closeness centrality this is time-consumption issue in the whole network dimension, especially for computing closeness centrality. In this section, we measure the time to compute the closeness centrality for each subnetwork using our dimensionality reduction approach. Table.1.

Sub-Network	Sub-Network Dimensionality	
	Original Feature Space	Reduced Feature Space
1	225	66
2	577	161
3	320	107
4	481	151
5	263	71
6	162	50
7	43	15
8	106	41
9	64	22
10	49	20

Table 1. Dimensionality Reduction for Each Sub-Network

Then, we compare our time with the original dimension for each subnetwork without dimension reduction. Figure.12 shows the results for the time consuming (in minutes) between the approach after dimensionality has reduced for each sub-network and the original feature space. We can notice that the average time consuming for closeness centrality using the original feature space is 52 minutes where our dimension reduction approach consumed 4 minutes. This shows that our proposed approach is significantly faster than the original approach (using the whole feature space for computing closeness centrality).

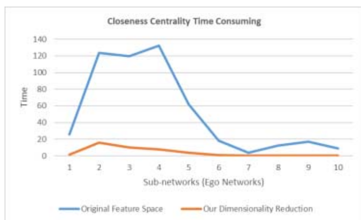


Figure13: Closeness centrality time consumption

ISBN 978-80-86943-42-8

5. CONCLUSION

This paper solves two significant problems of dynamic networks. The problem of dimension reduction and topological observation is resolved using MI and Feature selection based on developing an SVD algorithm. Our results show that the proposed approach for the time-series-network visualization based on the dimensionality reduction is significantly faster than the original dimension for distances measured by shortest paths detection based on the closeness centrality computing

6. ACKNOWLEDGMENTS

This work was partially supported by Ministry of Higher Education and Scientific Research (MHESR), Iraq, University of Koblenz Landau, Germany, HCED (Higher Committee for Education Development in Iraq), and University of Diyala-College of Science, Iraq.

7. REFERENCES

[ARK13] Andry Alamsyah, Budi Rahardjo, and Kuspriyanto. Social Network Analysis Taxonomy Based on Graph Representation. The 5th Indonesian International Conference on Innovation, Entrepreneurship, and Small Business (IICIES), (June):341–349, 2013.

[BJ15] B Birregah and O Jaafar. TVG-OSN: A Time-Varying Graph Model for Online Social Network Dynamics Analysis. Proceedings - 2nd European Network Intelligence Conference, ENIC 2015, pages 17–24, 2015.

[CZ17] Hamideh Sadat Cheraghchi and Ali Zakerolhosseini. Toward a novel art inspired incremental community mining algorithm in the dynamic social network. Applied Intelligence, 46(2):409– 426, 2017.

[FPA+12] Paolo Federico, Ju¨rgen Pfeffer, Wolfgang Aigner, Silvia Miksch, and Lukas Zenk. Visual Analysis of Dynamic Networks Using Change Centrality. Asonam, pages 179–183, 2012.

[FS15] Norie Fu and Vorapong Suppakitpaisarn. Clustering 1-Dimensional periodic network using betweenness centrality. Lecture Notes in Computer Science, 9197:128–139, 2015.

[GSZ+11] Frederic Gilbert, Paolo Simonetto, Faraz Zaidi, Fabien Jourdan, and Romain Bourqui. Communities and hierarchical structures in dynamic social networks: analysis and visualization. Social Network Analysis and Mining, 1(2):83– 95, 2011.

[LCZ+17] Xuefei Li, Lijun Chang, Kai Zheng, Zi Huang, and Xiaofang Zhou. Ranking weighted clustering coefficient in large dynamic graphs. World Wide Web, 20(5):855–883, 2017.

[Pea01] Karl Pearson. LIII. <i>On lines and planes of closest fit to systems of points in space</i>. Philosophical Magazine Series 6, 2(11):559–572, 1901.

[SVM14] C O S Sorzano, J Vargas, and a Pascual Montano. A survey of dimensionality reduction techniques. arXiv preprint arXiv:1403.2877, pages 1–35, 2014.

[WCG+16] Yingcai Wu, Nan Cao, David Gotz, YapPeng Tan, and Daniel A. Keim. A Survey on Visual Analytics of Social Media Data. IEEE Transactions on Multimedia, 18(11):2135–2148, 2016.

[Wu10] Yu Wu. Pattern Analysis in Dynamic Social Networks. 117(May):43–48, 2010.

[XH16] Wang Xiangang and Song Hanchen. A Dynamic Network Layout Visualization Method Based on Structural Similarity. 2016 International Conference on Cyberworlds (CW), pages 119–126, 2016.

[ZB16] Charles Y. Zheng and Yuval Benjamini. Estimating mutual information in high dimensions via classification error. (Nips), 2016.