ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 6, Issue 1, January 2017

# Applying Data Visualization Methods on Australian Stock Investment Analytics

Mouataz Zreika, Maria Estela Varua, Ana Hol,Jie Hua

***Abstract*—Stock investment decisions are often made based on current events of the global economy and the analysis of historical data. Conversely, visual representations may assist investors gain a deeper understanding not only on the market overview structure but detailed information on specified targets including useful insights on stock market trends. The trend analysis is based on long-term data collection. This study adopts a hybrid method that combines both clustering algorithms and force-directed algorithms to overcome the scalability problem when visualizing large datasets. This methodology exemplifies the potential relationships and interaction among each individual stock, as well as determining the strength of connectivity, which in turn will provide investors adifferent angle of view on the stock relationships. Information derived from visualization will also assist investors to make better informed decisions with less human disturbance due to pure mathematic calculations. The results of the experiments reflect that the proposed method can produce visualized data aesthetically by providing clearer views on an entire structure and specific connections.***

*Index Terms*—**Data Visualization, Graph Drawing, Data Filtering, Clustering, Force-directed, Stock Investment Analysis.**

## I. INTRODUCTION

Stock investment decisions require time, knowledge and awareness including current events of the global economy and historical data. The stock market contains a huge amount of data that varies over time. Stock prices are influenced by various factors ranging from the performance of the company itself to the conditions of the economy in general [1]. Thus for investors to manage their portfolio well, they must analyze stock market data regularly in order to identify the potential connection between various companies as well as predicting the future movement of each stock based on available historical data. However, finding and analyzing useful information in such a complex data oriented market usually requires a high level of analytical skills and effort from non-expert investors. The proposed hybrid methodology is designed to assist investors make better informed decisions based on stock market trend analysis.

To reduce the complexity of the analysis of stock market raw data, the hybrid visualization method was developed to examine the historical price movements of publicly traded stocks including figuring out potential relationships amongst stocks which are chosen during the same time period. This method is aimed to cluster similar stock together and provide investors the information that will enable them to predict future trends on stocks. The components of the hybrid visualization proposed are *Graph Drawing*, *Clustering* and *Spring Forces*.
Visual representation is one of the most efficient ways to assist investors to have a clearer overview of the movement of the stock market, as well as providing a deeper understanding of individual stocks. The application of the graph drawing method can provide visualized data with specific attributes such as weight information which comes with graphical connections between each data element. Although Data Visualization technologies have been adopted in the financial sector for a long time, they are however normally limited to *Treemap* and *Parallel Coordinates* etc.
In the meantime, with the rapidly increasing size in stock market related raw data,drawing a large graph with clear representations of complex data and its network structures is becoming a big challenge to the graph drawing community. The key issue here is not only to provide users with a comprehensive display of large graphs on the screen, but also a user-friendly navigable visual structure for users browsing through the structure to find a particular detail of the data [19].
In the past, some attempts to overcome this problem have proceeded which includes *Clustering* the groups (clusters) related nodes into super-nodes. User sees a summary of the graph with the super-nodes (clusters) and super-edges between the super-nodes (clusters). [2, 3, 4]. E.g. *K-mean* clustering method, *MarkovClustering* method.*Clustering* is the task of grouping a set of objects in such a way that objects in the same group (called cluster) are more similar (in some sense or another) to each other than to those in other groups (clusters). Clustered graphs have been widely been incorporated in *Graph Drawing* to overcome the problem of drawing large (or

huge) graphs with thousands, or perhaps millions of nodes. [18] Cluster analysis itself is not one specific algorithm, but the general task to be solved. It can be achieved by various algorithms that differ significantly in their notion of what constitutes a cluster and how to efficiently find them. Popular notions of clusters include groups with small distances among the cluster members, dense areas of the data space, intervals or particular statistical distributions. *Clustering* can therefore be formulated as a multi-objective optimization problem [12].In practice; applying different clustering algorithms to the same clustered graphs might create very different final layouts.

*Force-directed* layout algorithms use a physical analogy to draw graphs. A graph is viewed as a system of bodies with forces acting between the bodies. The algorithm seeks a configuration of the bodies with locally minimal energy, that is, a position for each body, such that the sum of forces on each body is zero. And the method is easy to understand, the results is normally good [2, 13, 14, 15, 16, 17].However, *force-directed* methods can deal with only a limited number of nodes due to its slow convergence time.

In this study, a new approach is proposed which combines *clustering* method and the traditional *force-directed* algorithm, to represent a clear overview of the whole structure on relevant stocks in reasonable convergence time by dividing a long convergence. The proposed method is applied to drawing weighted graphs. The early outcome of our approach indicates improvement in computation time and better graph aesthetics that provides a clearer view of the properties associated with the weighted graph in terms of its connectivity and edge weights. The preliminarily experimental results also show that the combination of the *clustered graph drawing* method and the *force-directed* layout algorithm could be used in *large graph drawing*.

## II. METHODS

The proposed hybrid method that combines the *Clustering* algorithms and *Force-directed* algorithm was tested in experiments. In our experiments, we use the clustering method based on edge weight to group vertices for pre-handling and then applied forces within each cluster. Details are described in the following subsections.

### *A. Chinese Whispers Clustering Method (CW)*

*CW* is an effective algorithm to partition the nodes of undirected graphs. It is motivated by the eponymous children's game, where children whisper words to each other. While the game's goal is to arrive at some funny derivative of the original message by passing it through several noisy channels, the *CW* algorithm aims at finding groups of nodes that broadcast the same message to their neighbors. It can be viewed as a simulation of an agent-based social network; for an overview of this field.

Intuitively, the algorithm works as follows in a bottom-up fashion:

1) First, all nodes get different classes;
2) Then the nodes are processed for a small number of iterations and inherit the strongest class in the local neighborhood.

This is the class whose sum of edge weights to the current node is maximal. In case of multiple strongest classes, one is chosen randomly. Regions of the same class stabilize during the iteration and grow until they reach the border of a stable region of another class. Note that classes are updated immediately: a node can obtain classes from the neighborhood that were introduced there in the same iteration.

**The *CW* algorithm can be described as follows:**

Chinese Whispers Clustering Method

```
initialize:
 forall v_i in V: class(v_i)=i;

while changes:
 forall v in V, randomized order:
      class(v)=highest ranked class
          in neighbourhood of v;
```

### *B. Markov Cluster Algorithm (MCL)*

Markov Cluster Algorithm (MCL) is based on the Markov Chain method which calculates the random walkers' chance between every pair of nodes in the graph, and then the nodes could be grouped according to the connection possibilities among them. The MCL adds the inflation operator for both strengthening and weakening of current [3] (Strengthens strong currents, and weakens already weak currents). The details of MCL are as below:

1) Expand by taking the eth power of the matrix;
2) Inflate by taking inflation of the resulting matrix with parameter *r*;
3) Repeat steps 1) and 2) until a steady state is reached.

*MCL* normally takes long running time due to its time complexity.

### *C. Decreasing Progressively Cluster on Weighted Graph (DPCW)*

The *DPCW* clustering is based on the connectivity of vertices and weight on each edge in the graph. The basic idea is that if a vertex $v_i$is assigned in a cluster $c_j$, then we intend to include all its connected vertices with the most weights in the graph in this cluster.

Suppose that $W = (w_0, w_2, \ldots, w_k)$ is the set of weights on every edge, $w_k$ is the maximum weight and $w_0$ is the minimum weight in *W*.

Assume that $G = (V, E)$ is a *connected undirected weighted graph*, where *V* is the set of vertices and *E* is the set of edges among *V*. A cluster graph $C = (G', T)$ consists of graph $G' = (V', E')$ and a rooted tree *T*, where *G'* is a sub-graph of *G*. The *DPCW* algorithm can be described below:

a) If $(v_m, v_n) \in V$, where $e_i = (v_m, v_n)$ and its weight $w_i = w_k$, then we add two vertices $v_m$ and $v_n$into the same cluster $c_k^1$;

b) If $(v_{m1}, v_{n1}) \in V$, where $e_{i1} = (v_{m1}, v_{n1})$ and its weight $w_{i1} = w_k$.
   1) If ($m = m_1$and $n \neq n_1$), then we add vertex $v_{n1}$into the cluster $c_k^1$;
   2) If ($m \neq m_1$and $n = n_1$), then we add vertex $v_{m1}$into the cluster $c_k^1$;
   3) If ($m = n_1$and $n \neq m_1$), then we add vertex $v_{m1}$into the cluster $c_k^1$;
   4) If ($m \neq n_1$and $n = m_1$), then we add vertex $v_{m1}$into the cluster $c_k^1$;
   5) If ($m \neq m_1$and $n \neq n_1$ and $m \neq n_1$and $n \neq m_1$), then we add two vertices $v_{m1}$ and $v_{n1}$into the same cluster $c_k^2$;

c) Repeat step (b) until every vertex which satisfies the conditions described in (b) are included in clusters, and the cluster $c_k = \{ c_k^1, c_k^2, \ldots, c_k^{xk} \}$;

d) Find the smaller weight $w_{k-1} \in W$, where $w_{k-1} < w_k$and $w_{k-1} > \{ w_0, w_2, \ldots, w_{k-2} \}$, set $w_k = w_{k-1}$, Repeat step (b) and (c) until every vertex which satisfies the conditions described in (b) are included in clusters, and the cluster $c_{k-1} = \{ c_{k-1}^1, \ldots, c_{k-1}^{x(k-1)} \}$;

e) Repeat step (d) until $w_i = w_0$ and every weight in *W* has been handled;

f) The final clusters $C = \{ c_k^1, \ldots c_k^x, \ldots, c_{k-1}^1, \ldots c_{k-1}^{x(k-1)}, \ldots, c_0^1, \ldots c_0^{x0} \}$.

### *D. A Classical Force-Directed Algorithm*

The *force-directed* algorithm aims to position nodes with as few crossing edges as possible by assigning forces among the set of nodes and edges for drawing graphs in an aesthetically pleasing way. The spring forces are used to keep all elements in reasonable distances in such a way that it not too close and not too far. The *force-directed* algorithms achieve this by assigning forces amongst the set of edges and the set of nodes. The entire graph is then simulated as if it were a physical system. In the *force-directed* algorithm, we need to calculate all the forces work on every element, and then place them to a suitable position to avoid edge crossings. The three steps for each iterative calculation are:

1) Calculate the effect of attractive forces $f_a(d) = d^2/k$ between adjacent vertices;
2) Calculate the effect of repulsive forces $f_r(d) = -k^2/d$ between all pairs of vertices;
3) Finally stop the iteration if $f_a$ and $f_r$ tend to not be changed.

Where *d* above represents the distance between two vertices while *k* is the optimal distance between vertices.

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 6, Issue 1, January 2017

### *E. Relevant Rate/Weight Computing Algorithm*

Based on the finalized layout using the combined methods of *clustering* and *force-directed* algorithm, our solution provides investors an overview of all the stocks, as well as the graphical representation of the relationships. The relevant aspects involved are summarized as follows:

#### *1) Trend Computing*

1) Individual rate
   Suppose the costs of an independent stock in two continuous business days are $c_1$ and $c_2$, the rate $r_1 = (c_2 - c_1) / c_2$, then the rate array of stock $k$ is $r_k = \{r_1, r_2, \ldots, r_k\}$;
2) Relevant rate
   Suppose the rates of two independent stocks in two same continuous business days are $r_{mi}$ and $r_{ni}$, the rate $r_{mni} = (r_{ni} - c_{mi}) / c_{ni}$, then the rate comparison array of stock $m$ and $n$ is $r_{mn} = \{r_{mn}^{1}, r_{mn}^{2}, \ldots, r_{mn}^{k}\}$, all the different time periods are dropped, only those rates changes happened within the same time period on both stocks are taken into account.

#### *2) Weight Computing*

Edge thickness is applied for displaying how close the connection is between stocks based on rates computing from above, the transmission from Rate to Weight is shown as:

1) If the absolute value of the original relevant rate $r_{mn}$ is bigger than 1, then the weight $w_{mn}$ = -1, which means no connection;
2) If the absolute value of the original relevant rate $r_{mn}$ is smaller than 1, then the weight $w_{mn}$ = Int((1 - Abs($r_{mn}$)) * 10);

Based on the weight calculation results above, edge thickness is computed as 0.5 * $w_{mn}$

## III. OUR APPROACH

This section describes the various steps of our hybrid approach. Graphically, the approach proposed in this study is depicted in Figure. 1. And the proposed approach is summarized in the following steps:

a) Stock raw data collecting from the ASX;
b) Data cleansing on raw data collected, including *Data Splitting*, *Data Comparing*, *Data Filtering* and *Graph Creating*;
c) Experiments on imported graphs, *which involves Graphs Clustering and Force-directed* methods;
d) Stocks relationship analytics based on finalized graph and relevant stock information;
e) Experimental outcomes;

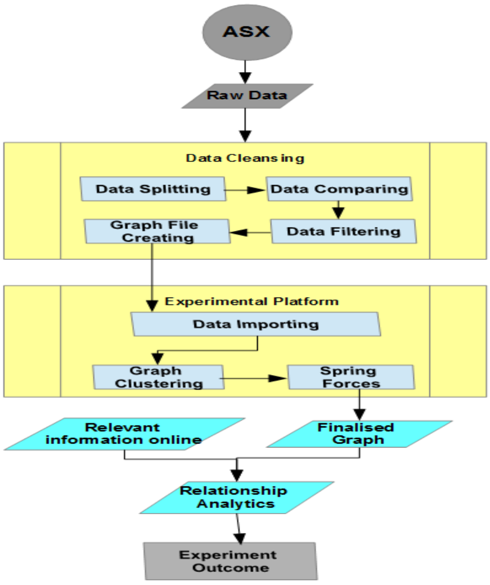



**Fig 1. Workflow of our approach**

### *A. Raw Data Description*

All raw data were collected from the Australian Securities Exchange (ASX), which includes 4869 stocks in Australia, ranges from 01/1997 to 04/2016, around 7.8 million data entries. A sample format is shown as:

**Data format sample of raw data**

| *Stock code* | Date | Open value | High Value | Low Value | Close Value | Volume |
|---|---|---|---|---|---|---|
| *BHP* | 19970102 | 17.779 | 17.88 | 17.64 | 17.82 | 1378841 |

**Table 1.Raw data format example**

This indicates relevant information of the specific stock of Stock Code, Date, Open Value, High Value, Low Value, Close Value, and Volume. The experiments are based on value comparing between every two stocks in the same time period.

### *B. Experiment Data Description*

For experiment purpose, data needs to be formatted as follows:

Graph File Format

```
<graph id="G" edgedefault="undirected">
<node id="n12" name="AAA"/>
<node id="n27" name="AAP"/>
<node id="n29" name="AAR"/>
.......
<edge id="e183" source="n12" target="n334" />
<edge id="e186" source="n12" target="n608" />
.......
</graph>
</graphml>
```

The format above is ready for import in our experiment platform, which defines vertices and edges.

### *C. Data Cleansing*

Prior to the experiment, the raw data is massive and unformatted and contains individual information on each stock. Relevant processes need to be done before importing to the experimental platform.
The entire processes are shown as:

***A. Data Splitting;***

Splitting mixed data of the entire stock on specific dates into individual data files following the time series. Hence, in our experiment, 4869 files have been created, corresponding to 4869 Australian stocks.

***B. Data Comparing;***

Based on those individual files, only compare values on the same time period of each two stocks, following the trend computing method in subchapter 1.4.1.

***C. Data Filtering;***

Due to the stock data feature, some stocks were not on market on certain dates, which means some of those stocks did not have values on specified dates; and some stocks may not exist nowadays, and those factors may affect the accuracy of the final analysing outcomes. Therefore, we did the *data filtering* onthe data processed from step B. the filtering rules are:

- Only apply data range from 01/2000 to 04/2016; and
- The total absence days is less than 10; and
- The weight is more than 5 following the rules from subchapter 1.4.2.

After the *data filtering* step, we have got 380 stocks with data which satisfy the selection, and those relevant files have all been taken into account in our experiments.

**D. *Graph Creating*;**

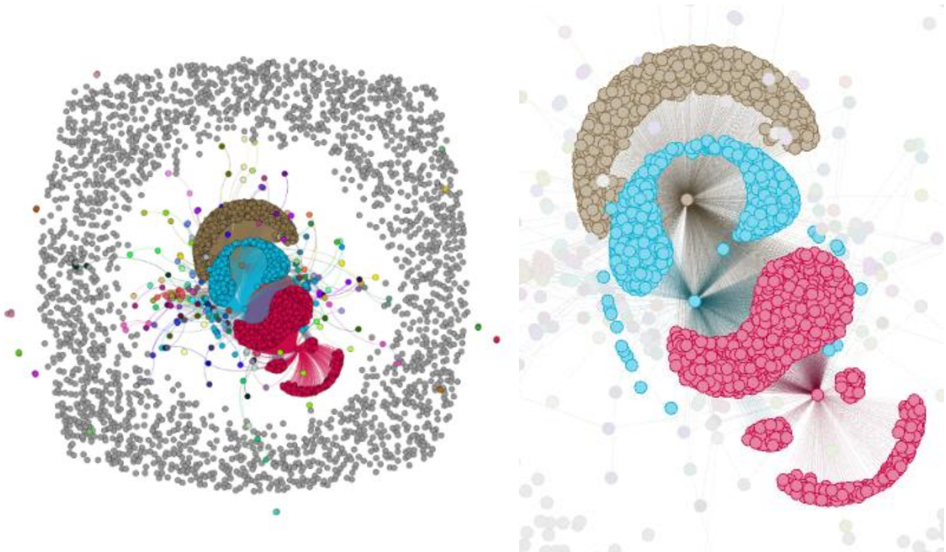
For the testing on the experimental platform, graph files reach the definition in subchapter 2.2 need to be created, relevant graph files have been built which define vertices (each stock) and edges (stock connection);

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 6, Issue 1, January 2017

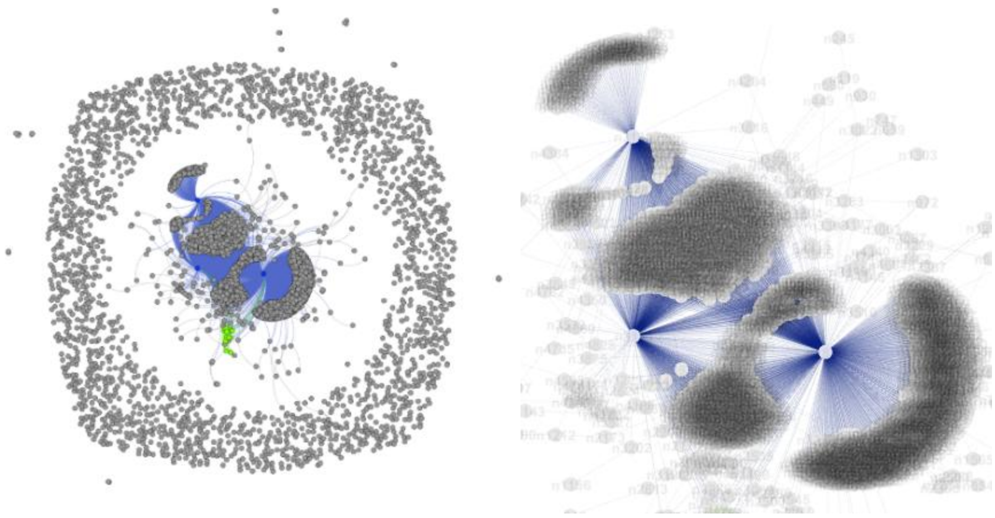
### *D. Experiment Approach*

Based on the graph files created, we apply *graph clustering* methods (*CW* and *MCL*) on the original graph *G= (V, E)*, and then apply *Spring Forces* on clustered graph *C(G) = (G', T)* with all its vertices in the layout until the convergence process is completed and reaches the energy balance. For any individual stock *x* details, apply forces on vertices connected to the specific vertex *x*, edge length needs to be adjusted based on weight on each edge, as well as the edge thickness. This approach could be applied on large data set, providing reasonable time complexity, comes with aesthetical visualized results. According to the *edge degree* and *edge weight* we could see the overview of the whole connections of the stock network; and detailed relationship of specific stocks via zoom-in/out. Examples can be found in the Figure 2.

## IV. EXPERIMENTAL EVALUATION



**(a) Open Value (*CW*)** **(b) Centre of Open Value graph (*CW*)**



**(c) Open Value (*DPCW*)** **(d) Centre of Open Value graph (*DPCW*)**

**Fig 2. Examples of visualized relationships after applied graph clustering methods and forces. (Without data filtering)**

In our experiment, we created artificially connected / undirected graphs based on around 7,800,000 data entries (before filtered) collected from the ASX to test the proposed method. The data was used for evaluation and different clustering methods have been adopted to represent connections among every indusial stock. We then applied the different *graph clustering* methods and *forces* on each graph and then compared the final layouts. See Figure. 2 regarding the initial layout including all vertices and edges. (We only adopted *CW* and *DPCW* clustering

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 6, Issue 1, January 2017
algorithm on this example due to the high computing time of *MCL*) From Figure. 2, 4,861 vertices and more than 40,000 edges have been involved.

Based on factors such as

- Connections (edges);
- Groups (clusters);
- Colors (ranking based on clusters);

The grey vertices ((Figure 2. (a), (c))) distributed peripheral come with less edge degree, and those nodes (Figure 2. (b), (d)) in the centre of the graph have much more connections, which should be emphasised in potential relationship analytics, since their trends may affect more stocks in market.

Stakeholders may be offered capability of analyzing connections among stocks based on long-term analytics, and the new view comes with less human disturbance, only based on the changing rates calculation, and this may assist stakeholders to consider more factor to handle investments other than advices from stock experts. A detailed case study is shown in next subchapter.

## V. CASE STUDY

A case is provided here to explain the details of our proposed methodology. For this case study, the raw data was collected from Australian Stock Exchange (ASX) for the period 1997 to 2016 (some stocks may be less) of 4869 companies. The detailed data aspects of each stock include:

- Open Value,
- Close Value,
- High Value,
- Low Value,
- Adjusted Close Value,
- Volume and Average Value.

Roughly 7.8 million data entries are applied in our experiments. After filtering raw data, 380 stocks have been taken into account to next stage for analytics. Based on the long-term stock value changing analytics, potential relationships between different stocks (companies) are provided depending on the final layouts of graphs. Bigger changing rate leads to weaker connection, and stocks tend to affect each other are put into the same group (cluster) with connections (edges).

As shown in the Figure 3, details of stocks are as (only partial stocks have been taken into account):

- NCM: Newcrest Mining
- CPU: Computershare Limited
- TAM: Tanami Gold
- BHP: BHP Billiton Ltd
- ANZ: ANZ Banking Group

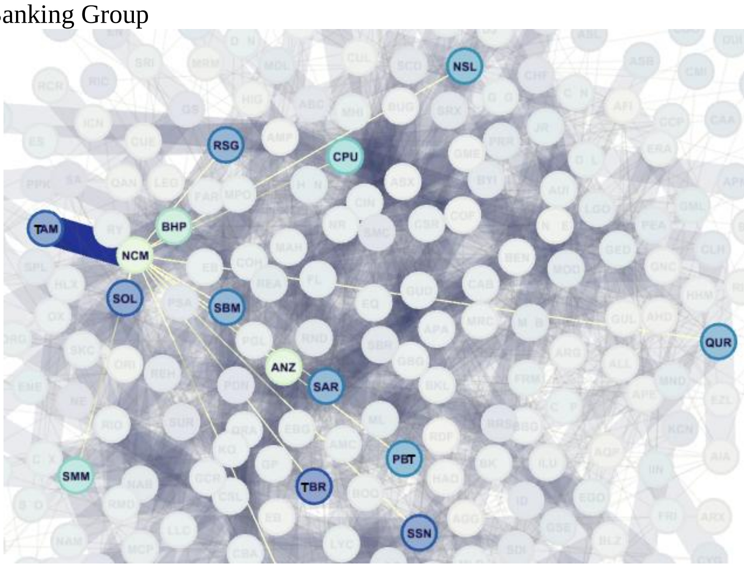



**Fig 3. Example of final layout of NCM (based on Open Value comparison and *CW* clustering).**

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 6, Issue 1, January 2017
The Figure 3above provides investors with visualized representation of connections of NCM. This will aid investors make informed decisions regarding their portfolio. The potential relationships among those stocks above are:

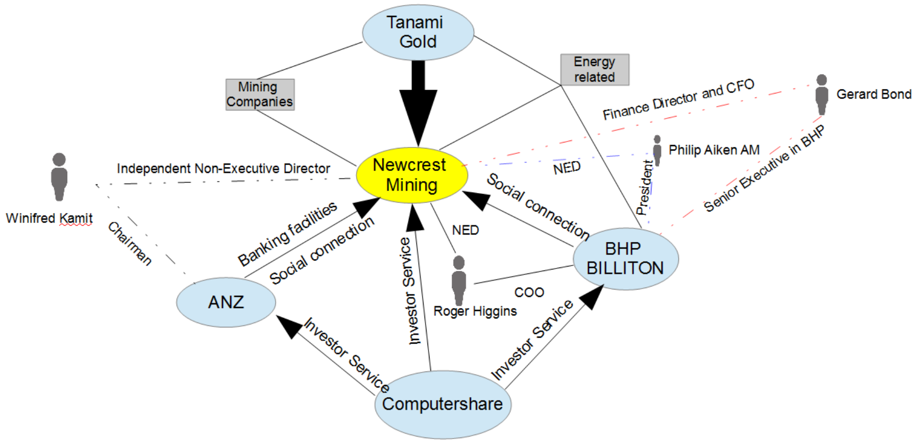



**Fig 4. Example of stock relationships analytics of NCM (based on information googled online).**

All those information were obtain from searching engine, details are as follows:

- NCM: Newcrest Mining

Newcrest is one of the world's largest gold mining companies and operates mines in four countries. (http://www.newcrest.com.au/)

Gerard Bond, who is the Finance Director and Chief Financial Officer at NCM and was the Head of Group Human Resources at BHP Billiton; Philip Aiken AM, who is the Independent Non-Executive Director at NCM and was Group President Energy BHP Billiton; Roger Higgins. Who is the Independent Non-Executive Director, and was the Chief Operating Officer with BHP Billiton. (http://www.newcrest.com.au/about-us/executive)

- ANZ

Lady Winifred Kamit who is the Independent Non-Executive Director in NCM and Chairman of ANZ Banking Group (PNG) Limited. (http://www.newcrest.com.au/about-us/board-of-directors)

The arrival of the ANZ banktainer has definitely sparked the interest of the Lihir community; Importantly for Newcrest, the additional banking facilities are proving to be a great convenience to the Lihirian community and its arrival will no doubt enable more people to be introduced to and have access to banking services. Newcrest congratulates ANZ for establishing the bank trainer. (http://media.corporate-ir.net/media_files/IROL/96/96910/reports-20121115/corporate-responsibility/bridging-urban-and-rural-divides/)

- CPU: Computershare Limited

It is an Australian stock transfer company that provides corporate trust, stock transfer and employee share plan services in a number of different countries. (http://www.theinfolist.com/php/SummaryGet.php?FindGo=Computershare)

- BHP: BHP Billiton limited

BHP Billiton, Newcrest make rival bids for copper, gold junior Solgold. (http://www.afr.com/business/mining/bhp-billiton-newcrest-make-rival-bids-for-copper-gold-junior-solgold-20161010-grz8cc)

Computershare Investor Services provides shareholders with a secure online facility at BHP. (http://www.bhpbilliton.com/investors/shareholderinfo/onlineservices)

The Figure 4 is based on those evidences collected from the Internet as above, which support our outcomes on the connections of the NCM from the side.

## VI. CONTRIBUTIONS

Our research's contributions are:

- The results of experiments show that the proposed method is able to produced visualized data aesthetically which is easy-to-understand to stakeholders, especially to none-experts;
- It bridges the gap to what has evolved before via the assistance of the power on direct vision to some degree;
- The graphical representations of data made possible by visualisation can communicate trends and outliers much faster than tables containing numbers and text. Users can easily identify issues and details via visualized data;
- The outcome offers not only the entire network of the stock market structure, but also detailed connections of specific stocks;
- It provides investors a different angle of view on the stock relationships for reference. Information derived from visualization will also assist them make an informed decision with less human disturbance due to pure mathematics calculation;
- The results of experiments show that the proposed method is able to produced visualized data aesthetically by providing clearer views on entire structure and specific connections.
- The methodology has been extended onto other fields such as river/rainfallgovernment project, to visualize the data handled from *machine learning* method, to provide executives judgements on decision making on priority of dam building in NSW, to reduce the risk on drought and flood in relevant area. And in the future, we are applying it on BIM (building information modelling) project.

Revised the traditional *force-directed* algorithm on termination condition, the combination of *graph clustering* methods and *graph drawing* algorithm has improved the time complexity.

## VII. CONCLUSION AND FUTURE WORKS

In this paper we have presented an approach for potentially visualizing the relationships between stocks and presenting trends for individual stocks by combining *graph clustering* and *force-directed* algorithms based on long-term historical data analysis. The experimental results of the proposed method demonstrate its effectiveness in terms of providing investors reasonable visualized information to assist in their decision regarding stock investment; however we have identified several issues or limitations that we need to address in our future works. These issues are listed below:

- Stock market analysis is complex, and affected by multiple factors, this methodology could only provide an additional view from another angle to assist investors to have a better understanding of the potential relationships between each stock.
- Different value comparison may cause different final connection structure, the accuracy of the value comparison is still unknown.
- The methodology is only applicable to stocks with a long history of available data (long-term) but cannot be adopted to those with limited data availability (short-term).
- The methodology's outcome is rely on detailed*graph clustering* algorithms, they may cause the wrong group divisions, relevant algorithms need to be specified based on users' detailed needs.
- This methodology can only assist stakeholder with different angle of views on investments, can't replace advices from financial experts.

In our future works, we will apply the proposed revised method to a wider and larger set of data and applications. The future work will be addressing the limitations identified by:

- Carrying out further experiments that use more data from other fields. (We have adopted this methodology on a river/rainfall data analytics government research project, it is still in process)
- Revising the *graph clustering* algorithm to emphasis the weight computing on each connection.
- Undertaking a study to formally evaluate the effectiveness of the proposed method, working with experts in stock market to find more factors affects the finalized methodology.

The current research only focus on relationship analytics, we could adopt *SPC* (statistical process control) concept onto data analytics, which could provide early-warning feature on investments, then stakeholders are capable of adjusting investments in advance to some degree.

## REFERENCES

[1] Cindy Wang. 2014. Investing 101: How to Analyze Stock Market Trends. [ONLINE] Available at: http://blog.sprinklebit.com/investing-101-how-to-analyze-stock-market-trends/. [Accessed 13 July 15].

[2] Omote, H. & Sugiyama, K. 2007, Force-Directed Drawing Method for Intersecting Clustered Graphs, APVIS 2007, 6th International Asia-Pacific Symposium on Visualization 2007, 5-7 February 2007, Sydney, Australia, pp.85-92

[3] Huang, X. & Lai, W., 2005, clustering graphs for visualization via node similarities. J. Vis. Lang. Comput. 17, 3 (June 2006), pp. 225-53. Elsevier Ltd.

[4] Huang, M., Nguyen, Q. V., A space efficient clustered visualization of large graphs Image and Graphics, 2007. ICIG 2007. Fourth International Conference on, pp. 920-27

[5] Huang, M.L., & Nguyen, Q.V. 2007, Navigating Large Clustered Graphs with Triple-Layer Display, 11th International Conference Information Visualization, 2-6 July 2007, Zürich, Switzerland, pp. 684-92.

[6] Sarkar, M. & Brown, M.H. 1994, 'Graphical fisheye views', Communications of the ACM, Volume 37 Issue 12, pp. 73–83.

[7] Huang, M.L., Eades, P. & Wang, J. 1998, On-line animated visualization of huge graphs using a modified spring algorithm, Journal of Visual Language and Computing, no. vl980093, pp. 623-45.

[8] Huang, M.L., Eades, P. & Cohen, R.F. 1998, Webofdav – navigating and visualizing the web on-line with animated context swapping, Computer Networks and ISDN Systems, 30(1), pp. 638-42.

[9] Nguyen, Q.V. & Huang, M.L. 2005: EncCon: an approach to constructing interactive visualization of large hierarchical data. Information Visualization 4(1): 1-21.

[10] Qiu, M., Zhang, K., Huang, M., 2006, Usability in mobile interface browsing, Web Intelligence and Agent Systems, Vol.4 (1), pp. 43-59, IOS Press.

[11] Nguyen, Q. V. and Huang, M. L., A focus+ context visualization technique using semi-transparency, The Fourth International Conference on Computer and Information Technology, 2004. CIT'04, pp. 101-08.

[12] Brandes, U., Delling, D., Gaertler, M., Gorke, R., Hoefer, M., Nikoloski, Z., Wagner, D., On Modularity Clustering, Knowledge and Data Engineering, IEEE Transactions on, pp. 172 - 88 Volume: 20, Issue: 2, Feb. 2008.

[13] Battista, G.D., Eades, P., Tamassia, R. & Tollis, I.G. 1999, Graph drawing algorithms for the visualization of graphs, Prentice-Hall, New Jersey, USA.

[14] Eades, P., A heuristic for graph drawing. Congress Numerantium, 42:149-160, 1984.

[15] Lin, C.C., Yen, H.C. 2005, A New Force-Directed Graph Drawing Method Based on Edge-Edge Repulsion, Ninth International Conference on Information Visualization, 6-8 July 2005, London, UK, pp.329-24

[16] Battista, G.D., Eades, P., Tamassia, R. & Tollis, I.G. 1999, Graph drawing algorithms for the visualization of graphs, Prentice-Hall, New Jersey, USA.

[17] Lin, C.C. & Yen, H.C. 2008, A new force-directed graph drawing based on edge-edge repulsion, 9th International Conference on Information Visualization IV2008, 6-8July, London, England, pp. 329-34.

[18] Hua, J. & Huang M.L. (2013). Improving the Quality of Clustered Graph Drawing through a Dummy Element Approach. In Computer Graphics, Imaging and Visualization (CGIV), 2013 10th International Conference. Macau, 6-8 Aug. pp. 88-92.

[19] Hua, J., Huang, M.L & Nguyen, Q.V, (2014). Drawing Large Weighted Graphs Using Clustered Force-Directed Algorithm. In Information Visualization (IV), 2014 18th International Conference on. Paris, 16-18 July 2014. Paris: IEEE. pp. 13-17.

[20] Zreika, M, 2016. Applying Hybrid Graph Drawing and Clustering Methods on Stock Investment Analysis. International Journal of Mathematical and Computational Sciences, 3, pp. 689-98.

[21] Hua, J., Liang, J. & Huang, M.L. Comparison of Clustered Force-directed Graph Drawing for Large Dataset Visual Analytics. Poster presented at: Workshop on Data Mining Meets Visual Analytics at Big Data Era (DAVA'16), Indianapolis, US. Oct. 28, 2016