# Social Network Analysis and Pension Fraud Prediction in Public Enterprises
## A Case of Developing Countries

Alocate Zvikaramba, Abdul Rashid Abban, Mujeeb Ur Rehman Rahimi

**Abstract —** The impact of social networks in business is real. Fraud robs economies of millions of dollars. This paper analyzes fraudulent activities in public enterprises particularly in pension administration that has witnessed a mismatch between capital gains and pension liabilities creating pension payment back-logs, anxiety amongst employees and retirees as pension payouts no longer meet basic needs. The phenomenon is linked to social network connections within and outside workplace. We use a fraud analytics process model in the preprocessing, analytics and post processing of data and employ association rules in data mining. Our findings show that there is an important association between social media activities, centrality of organizational networks and fraudulent activities.

**Key words: Social Network Analysis, Fraud Prediction, Public Enterprises, Pension**

— — — — — — — — — ◆ — — — — — — — — —

## 1 INTRODUCTION

THE increase in white collar crime in the world is a cause for concern. Robb [1] posits that 'white-collar crime is the soft underbelly of the modern economy, robbing the public of millions of dollars, undermining trust in commercial integrity and depressing the level of investment in new industries. Financial crime is as old as capitalism itself'. Many enterprises including state owned institutions have not found a solution to this business risk. According to a survey by PwC [2] government entities around the globe that suffered from economic crime reported that '57% of perpetrators were internal while only 37% were external'. Developing countries are worst hit by fraudulent activities within public enterprises further bleeding the already fragile economies. The problem is camouflaged in internal policies and procedures, employee behavior, organizational culture, technological effects of social media and networks.

Sandbrook and Gosling [3] argue that in developed countries like the UK, the pensions market is in flux and unpredictable with new providers emerging. In contrast, developing countries such as Pakistan, state owned institutions have their pension schemes tailor made for civil servants who contribute through compulsory pension savings. The problems faced are not unique but similar elsewhere like Shaikh [4] citing Arthur (2006) points that the history of state pensions is said to be a history of broken promises.

-----------------------------------------------------------

- *Alocate Zvikarambais a Certified E-Commerce Consultant, BTech E-Commerce and DISSM holder currently pursuing masters' degree program in Information Technology for Business Innovation at Ural Federal University, Russia, эмм151602. E-mail: alozvika@ecobua.com*
- *Abdul Rashid Abban holder of BA Economics and Mathematics and currently pursuing masters' degree program in Information Technology for Business Innovation at Ural Federal University, Russia, эмм151602.Email: abdulrashidabban@gmail.com*
- *Mujeeb Ur Rehman Rahimi holder of BBA Marketing and Finance and currently pursuing masters' degree program in Information Technology for Business Innovation at Ural Federal University, Russia, эмм151602 Email-mujeebrn@yahoo.com*

These include a mismatch between capital gains and pension liabilities as shown in figure 1, resulting in pension payment back-logs, anxiety amongst employees and retirees since most of their pension payouts do not meet their basic needs creating destitution.

Such circumstances have created conditions for corruption and fraudulent activities in the administration of pension funds. Hence, the growing calls to privatize the pension management. The severity of this challenge is not peculiar to Pakistan alone but a host of other developing countries.
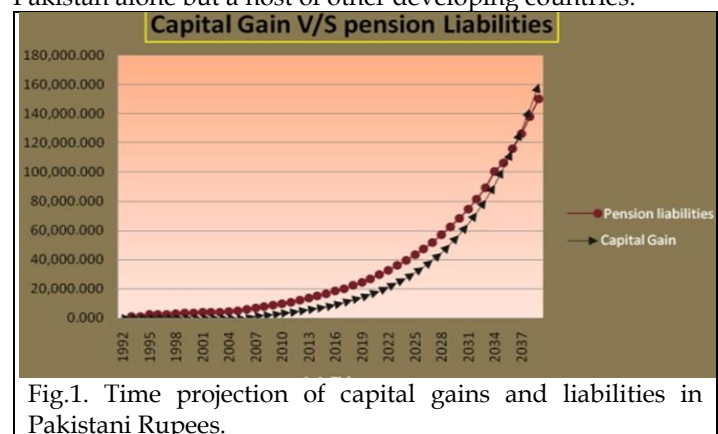


Fig.1. Time projection of capital gains and liabilities in Pakistani Rupees.

An academic inquiry through fraud analytics using descriptive, predictive and SNA may yield the desirable solutions. In this research we examine the connection between social networks and fraud in pension schemes of public enterprises. The structure of the paper is that we are going to look at definition of terms, literature review, methodology, presentation of results, discussion and conclusion.

## 2    DEFINITION OF TERMS:

### 2.1    Social Network
Xu [5] define a social network as a heterogeneous and multi relational dataset represented by a graph. Vertexes represent the objects (entities), edges represent the links (relationships or interaction), and both objects and links may have attributes.

### 2.2    Social Network Analysis (SNA)
SNA is the mapping and measuring of relationships and flows between people, groups, organizations, computers, URLs, and other connected information or knowledge entities [6]

### 2.3    Public Enterprise
A business organization wholly or partly owned by the state and controlled through a public authority [7] and the United Nations definition of a public enterprise is, 'an incorporated or large unincorporated enterprise in which public authorities hold a majority of the shares and or can exercise control over management decisions'.

### 2.4    Pension
Retirement benefit, series of periodic money payments made to a person who retires from employment because of age, disability, or the completion of an agreed span of service. The payments generally continue for the remainder of the natural life of the recipient, and sometimes to a widow or other survivor [7]

### 2.5    Fraud
is uncommon, well considered, imperceptibly concealed, time-evolving and often carefully organized crime which appears in many types of forms [8]. Fraud is also an interchangeable word deception, bribery, forgery, extortion, theft, conspiracy, embezzlement and misappropriation.

## 3    LITERATURE REVIEW

There is little literature on the study of social network analysis and the linkage to fraud in pension schemes of public enterprises per se. However, there is significant material on key terms used in this paper and we shall review some of this literature. According to Satell G [9], the story of networks started in 1736, when Leonhard Euler set out to conquer a famous math problem concerning the Seven Bridges of Königsberg, now Kaliningrad of Russia. To solve it, he created a new form of mathematics called graph theory, which concerned itself with links and nodes in a network. More scientific work contributions were made by people like Anatol Rapoport in the 1950s, later Erdős and Rényi and in the 1970s a sociologist named Mark Granovetter who argued that most

information is obtained through weak ties rather than close friends. Supporting this idea Watts and Strogatz (1990) further demonstrated that small clusters of people naturally organize themselves into far flung networks and by the late 1990s the field of network analysis become fully fledged science. In 2002, [6] teaming up with American Defense Department successfully applied knowledge to map terrorist networks.

**The value of SNA cannot be underestimated especially in information sciences and business**.
SNA is practically used extensively in a wide range of applications and disciplines. Some common network analysis applications include data aggregation and mining, network propagation modeling, network modeling and sampling, user attribute and behavior analysis, community-maintained resource support, location-based interaction analysis, social sharing and filtering, recommender systems development, and link prediction and entity resolution. In the private sector, businesses use social network analysis to support activities such as customer interaction and analysis, information system development analysis, marketing, and business intelligence needs. Some public sector uses include development of leader engagement strategies, analysis of individual and group engagement and media use, and community-based problem solving.

**Linking social media with business is topical and trendy**. Daniele Quercia [10] assert, 'there has been considerable success in leveraging large-scale social media data at the intersection of social sciences and computational sciences with myriad applications in socioeconomic measurement and prediction'. [10] further proves the importance of twitter, facebook in predicting market volatility, stock market fluctuations by citing the work of Antweiler and Frank (2004), Karabulut (2011), Bollen et al (2011) and the study of the correlation between emotional tweets and financial market indicators by Zhang etal (2010) which revealed that there is a positive correlation between percentage of tweets and volatility index. Further evidence is in the multidimensional mood analysis model that can track Twitter mood in six dimensions and its predictive power on price changes, work of Bollen et al (2011). Daniele [10] also mentions that the language of Twitter users is associated with the socioeconomic well-being those users experience in their physical communities'. To this end we set to investigate this claim too.

**Network centrality measures were discussed in literature before** as Everett and Borgatti [11] note the work of Katz (1953) and additional work by Freeman (1979) who categorized centrality measures into degree, closeness, and betweenness with canonical measures for each category.

**Public enterprises are critical in today's economies particularly in developing countries as they can make or break a nation.** Leroy Jones [12] recognizes this and points that that public enterprises are socially accountable and have as one of their roles the strengthening of the community, as well as the state and the economy.

**Since fraud can disguise itself in corruption, networks are critical to study this phenomenon.** On networks and corruption, [13] cites a New York article that 'guanxi' is the Chinese word for 'connections' saying 'if you have guanxi in China, it is the opportunity to make big money' and in today's China, 'connections' have great economic value and smart people can capitalise on them to earn large 'incomes'. Robb [1] posits that 'white-collar crime is the soft underbelly of the modern economy, robbing the public of millions of dollars, undermining trust in commercial integrity and depressing the level of investment in new industries. Financial crime is as old as capitalism itself'. [8] links fraud with social network techniques. We further extend this work by analyzing a sample derived from public enterprises of pension administration in a developing country. Fraud is usually done by insiders [2].

**The essence of pension schemes in the contribution to the socio-economic well-being of a nation is worth noting.**
Wei Yang and Pradip Tapadar [14] observe that regulators and policy-makers are turning their attention to the pension sector, the other integral player in the financial markets. We note that this attention is not adequate as there are underfunding at retirement, an issue raised in [15]. Problems in the pension sector weigh down an economy. Frankfurter and Hill [16] reveal, 'unfunded liabilities (pension benefits payable net of pension assets) have become a sizable component of the capital structure of many firms. Pension funds (assets) also have grown rapidly due to liberalization of pension benefits, inflation, the higher funding requirements of the pension reform law, and the overall aging of the labour force'. Developing countries are plagued with corruption and fraud creating a mismatch between capital gains and pension liabilities as shown in fig.1. In developed countries like the UK, the pensions market is in flux and unpredictable with new providers emerging [3] Developed countries have not yet fully privatized this sector and challenges are bigger.

## 4 AIM

The aim of this research is to analyze social network and predict the possibilities of fraud activities in the pension scheme of developing countries



The intersection diagram

Fig.2. Intersection diagram
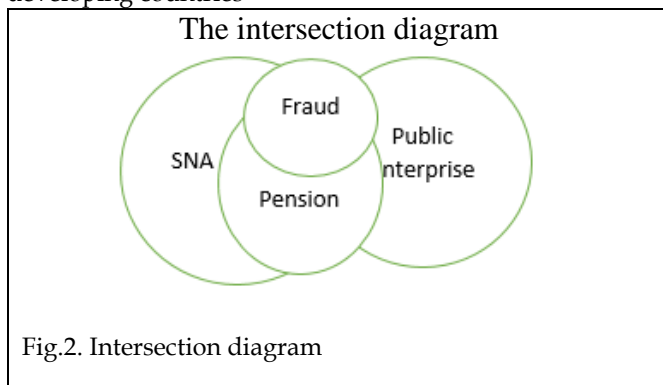
## 5 RESEARCH PROBLEM

There is perceived increase in fraudulent activities in public enterprises particularly of pension administration attributable to financial risks of loss due to mismatch between capital gains and pension liabilities resulting in pension payment back-logs, anxiety amongst employees and retirees since most of pension payouts do not meet retirees' basic needs thereby causing destitution. The cycle is linked to social network connections within and outside workplace.

## 6 RESEARCH QUESTIONS TO WARRANT RED FLAGS 🚩

1.Are there workers living beyond their means?
2.What is the link between personal savings and salary of employees?
3.The capacity cum possibility of bending rules according to nature of office:

a. Employing ghost employees?
b. Fast tracking retirement?
c. undervaluing lump sum entitlement?
d. Over valuing lump sum entitlement?
e. Forgery of documents
f. Bribery

4.Is there a connection between active nodes in the social network with fraudulent activities.
5. What algorithms can be used to associate possible fraudulent activities with offices?
6. Can disciplinary records at work be linked to abnormal savings?

## 7 METHODOLOGY

In this research, a fraud analytics process model is used in the preprocessing, analytics and post processing of data. In tandem with the model, we identified a business problem first, identified data sources, cleaned and transformed the data before it was analyzed using various tools like Microsoft Azure Machine learning, R Programming, Excel and MySQL. A survey from a sample of 100 public employees was done. The data was gathered through questionnaires, interviews and observations and it is contained in files attached.

## 8     PENSION POLICY INFORMATION SYSTEM



Fig.3.Policy Rules on Retirement.
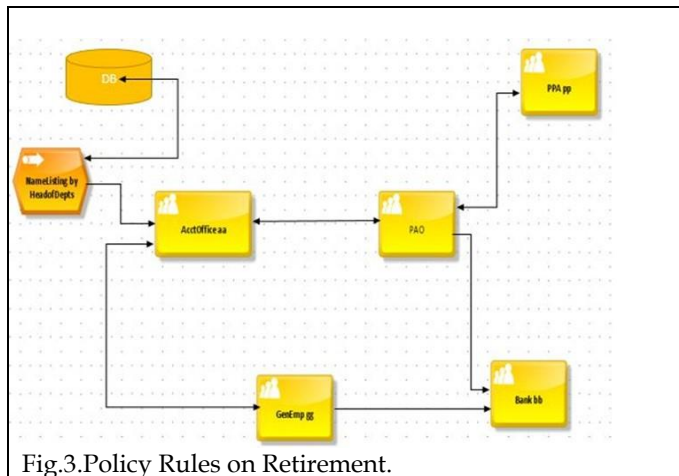
The process of retirement is initiated by the Head of Departments Office through his interaction with the database. Papers are sent to the Accounts office for further processing, then to the Pay Accounting Office(PAO) and Principal Pension Authority(PPA) for Approval then back to Pay Accounting Office and finally to Bank for payment. Concerned Employees get notification from the accounting office and collect money from the bank.

## 9     FRAUD-DETECTION RULES FED INTO FRAUD DETECTION ENGINE

**Business Policy Rule 1**
IF multiple advance payment of one retiree claim, THEN suspicious case

**Business Policy Rule 2**
IF payment is made before approval, THEN suspicious case

**Business Policy Rule 3**
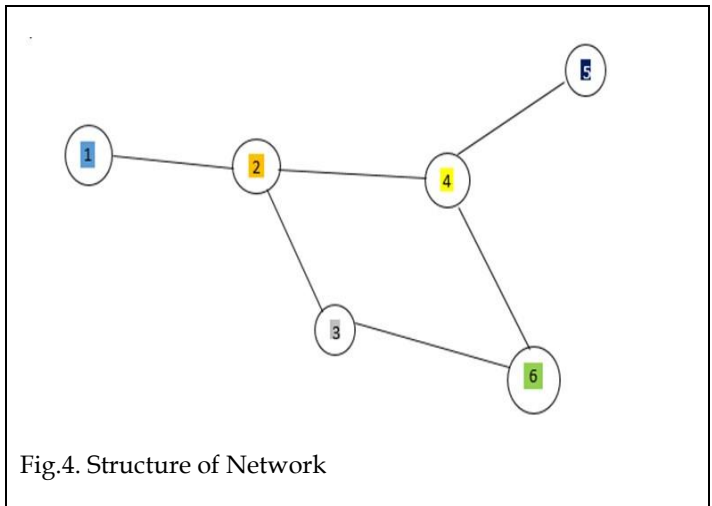IF payment is made AND approved by same person, THEN suspicious case

**Business Policy Rule 4**
IF payment is made AND approved but changed afterwards, THEN suspicious case

**Business Policy Rule 5**
IF payment is rejected AND processed afterwards (without reason), THEN suspicious case

## PENSION OFFICES ORGANISATIONAL NETWORK



Fig.4. Structure of Network

### Discussion 1

Examining the degree of centrality, from the network in fig.4, node 2 has the most direct connections in the network, making the most active node in the network. Node 2 also has the highest degree of betweenness followed by node 4. Nodes 1 and 5 are the peripheral nodes, they are connected to networks that are not currently mapped. Otte and Rousseau [17] say, every centrality measure can be used to derive a centrality measure, C, for the whole network. Done as follows:

$$C_{network} = \frac{\Sigma_j \left( C_{max} - C_j \right)}{\text{max value possible}}$$

max value possible. This formula can be applied for determining degree, closeness and betweenness centrality. The summation goes over all nodes of the network; Cmax is the largest value obtained in the network under study, and 'max value possible' refers to the maximum value possible for the numerator, given the total number of nodes.
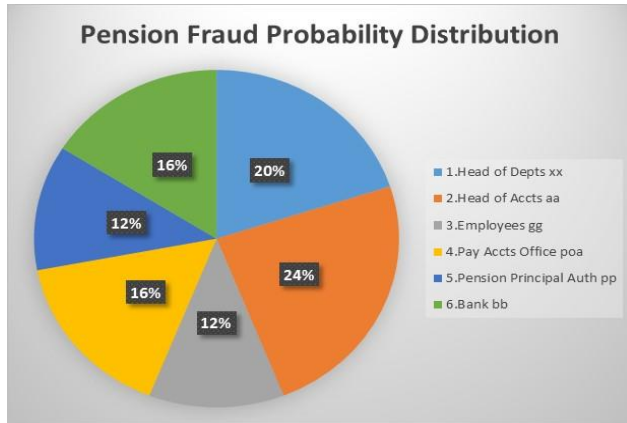
Fig. 5. Pension Fraud Probability Distribution According to Offices. Based on the capacity and probability to bend policy rules. Bayesian Rule Applied.

## Discussion 2

The probability diagram indicates that 56.41% of these employees exposed to fraudulent activities are working in the pension scheme offices. 70.97% have been exposed to fraudulent activities. This result suggests there is huge possibility of fraudulent activities in this sector.
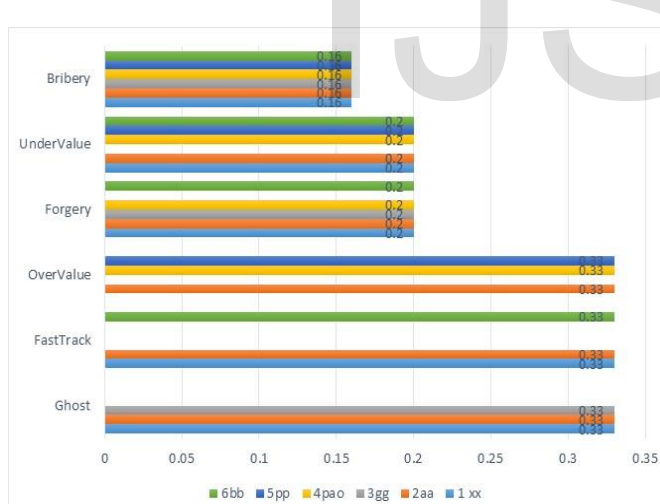


Fig.6. Distribution of Fraud Types According to Offices Bayesian Rule Applied.

## Discussion 3

On figure 6 above , fraud type tends to be associated with offices. Some offices have a greatest share of a certain type of fraud than others, possibly due to the nature of duties of their holders. It makes sense to link this with the offices location in the network diagram discussed earlier on.Bribery is fairly

common among all offices whilst employement of ghost workers is fraud 'reserved' for Head of Departments and Accounts Department. From the business rules, Head of Department is responsible for making a list for retirees, there is a possibility to recruit ghost employees in connivance with Accounting Department.
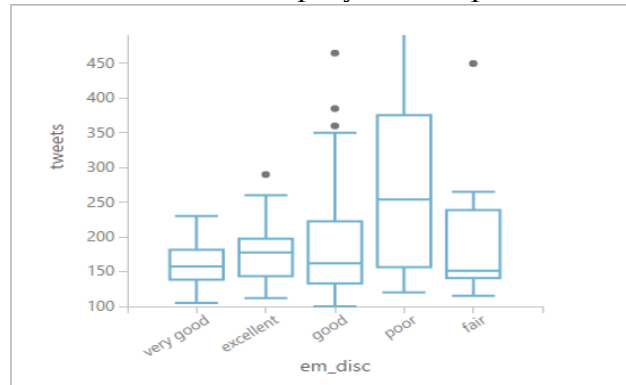


Fig.7. It can be learned that as discipline drops tweets tend to increase.

## Discussion 4

From our findings, employees who are ill disciplined according to disciplinary records,had more tweets and twitter followers Fig.7. shows. Similiary, Fig.8 shows employees who had savings greater than salaries meaning that they are living beyond their means warranting a red flag.
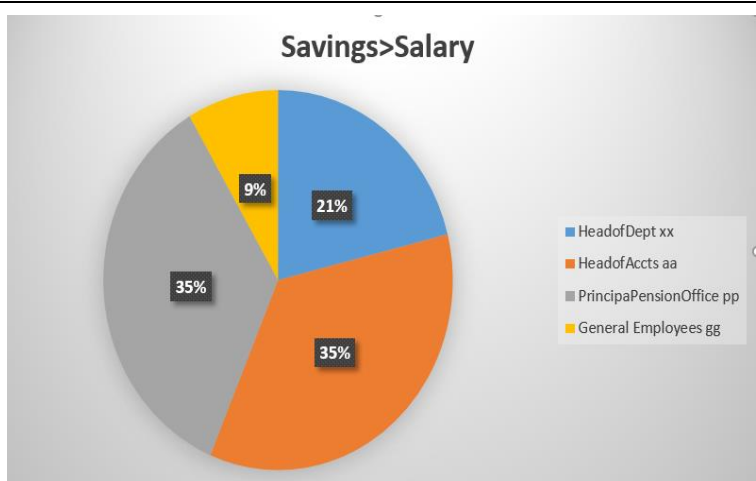


Fig.8. Employee with savings greater than salary, majority working in the Head of Accounts Department and Principal Pension Office.

## 9. 1 Application of the Apriori Algorithm to problem of association

 Coenen and  Leng[18]postulate that Association rule mining (ARM) is now a well-established branch of knowledge discovery in databases (KDD). The Apriori property states that if an itemset is considered frequent, then any subset of the frequent itemset must also be frequent. For example, if an itemset {milk, bread} appears in 65% of all retail transactions of a large retail corporation, the itemset{milk} as well as the itemset {bread} will appear at least 65% of the transactions [20].

   Given an itemset Q, the 'support' of Q is the % of transactions containing Q. To meet support criteria of 0.5 we need to show up the sets of transactions that show up at least 50% of the time.

| Itemset | Frequency | Support |
|---------|-----------|---------|
| xx | 5 | 0.20 |
| aa | 6 | 0.24 |
| gg | 3 | 0.12 |
| pao | 4 | 0.16 |
| pp | 3 | 0.12 |
| bb | 4 | 0.16 |

Confidence $(X \longrightarrow Y)$= (Support(X ∩ Y))/Support (X)

A higher confidence indicates that the rule is more interesting and more trustworthy based on the sample dataset.

Confidence $(gg \longrightarrow aa)$= (Support(gg ∩ aa))/Support (gg)
= 0.12*0.24/0.12
=24%

### Discussion 5

It is interesting to note that fraud recorded a highest support of 24% in Accounting Office(aa) and minimum of 12% in the general offices. This could signify that employees working in the accounts department of Pension Administration are more exposed to the risk of fraud than those working in other offices. This is also true with the degree of centrality of network fig.4 Head of Accountants has the most direct connections in the network, making this node the most active node in the network.

## 10    FINDINGS

   After our analysis, we realized that 39% of the sampled employees have been exposed to fraudulent activities. 56.41% of these employees exposed to fraudulent activities are working in the pension scheme offices. Moreover, considering only the 31 employees in our sample that were working with the pension scheme offices, we derived from our analysis that 70.97% have been exposed to fraudulent activities. This result suggests there is huge possibility of fraudulent activities in this sector. In addition, our analysis showed that 41.94% of the sampled employees working in the pension scheme offices save more money than their salary and this is also a major suspicion that these people might be engaging in fraudulent activities. These employees working at the pension scheme offices have great connections with lots of people and they again have huge followers on twitter. It is again clear to deduct that these people who have been previously exposed to fraud and have such wide connections are likely to be tempted again for favours from people they are connected to.

## 11    CONCLUSION

From our survey it can be concluded that workers in the accounts and finance departments have a higher exposure to risks of fraud than those in general offices of pension schemes. Similarly, offices that have direct connections in the network, recorded more activities on the social media and had significant exposure to risk of fraud. The same employees who had records of indiscipline are well connected and live beyond their means. In summary, our analysis has showed that there are lot of possibilities that can trigger fraud activities in the pension scheme of most developing countries, looking at the past records, behaviors, salaries, savings and connections of people working under this scheme.

## 12    RECOMMENDATIONS

- More checks and balances are required in the Accounts and Head of Department Offices.
- Workers living beyond their means should be closely monitored.
- More awareness campaigns be done to educate employees about the pension scheme and their entitlement.
- Further research needs to be done in this field particularly social media to establish the connection with fraud and the pension scheme to further support our findings.

## 13    References

[1]    George Robb White-Collar Crime in Modern England Financial Fraud and Business Morality, 1845–1929 Publisher: Cambridge University Press Online Publication December 2009.

[2] PricewaterhouseCoopers(PwC), Fighting fraud in the public sector, Australia, 2011.

[3] Will Sandbrook and Tim Gosling, Pension Reform in the United Kingdom: The Unfolding NEST Story Rotman International Journal of Pension Management ,2014.

[4]  Khalid Mehmood Shaikh Pension Privatization in Paki stan, Computer Cell (Resource), Finance Department, Government of Sindh, 2010.

 [5] Xu G, Zhang Y, Li L. Web Mining and Social Networking Techniques and applications, Springer 2011

[6] Valdis Krebs,2002 http://www.orgnet.com/VKbio.html

[7] http://global.britannica.com/topic/public-enterprise
What is a public enterprise?
https://www.capdm.co.uk//demos/gi/gi/GI0802.html

[8] Bart Baesens, Veronique Van Vlasselaer and Woulter
Verbeke, Fraud Analytics Using Descriptive, Predictive
and Social Network Techniques: A Guide to Data Science
for Fraud Detection.SAS Institute North Carolina USA,
2015.

[9] Satell Greg http://www.digitaltonto.com/2010/the-story-of-networks/

[10] Daniele Quercia Hyperlocal Happiness from Tweets pp.
96-110 Twitter: A Digital Socioscope
Publisher: Cambridge University Press Online
Publication May 2015.

[11] Martin G. Everett and Stephen P. Borgatti
Extending Centrality pp. 57-76 Models and Methods in
Social Network Analysis ,Cambridge University Press
,2012

[12] By Leroy P. Jones, Social Accountability of Public Enter
prises: Law and Community Controls in The New
Development Strategies pp. 77-100. Cambridge University
Press, August 2010.

[13] Vito Tanzi ,Corruption: arm's-length relationships and
Smarkets pp. 161-180 The economics of organised crime
Cambridge Online ,2010.

[14] Wei Yang and Pradip Tapadar, Role of the Pension
Protection Fund in Financial risk management of UK
defined benefit pension sector: a multi-period economic
capital study. Annals of Actuarial Science, 9, pp 134-166
doi:10.1017/S1748499514000256, 2015.

[15] Torsten Kleinow Pension Fund Management and
Conditional Indexation. ASTIN Bulletin, 41, pp 61-86.
doi:10.2143/AST.41.1.2084386.,2011.

[16] George M. Frankfurter and Joanne M. Hill, A Normative
Approach to Pension Fund Management. Journal of
Financial and Quantitative Analysis, 16, pp 533-555.
doi:10.2307/2330373, 1981.

[17] Otte E., Rousseau R. Social network analysis: A powerful
strategy, also for the Journal of Information Science, 28
(6) , pp. 441-453 ,2002.

[18] Frans Coenen and Paul Leng, Partitioning strategies for
distributed association rule mining. The Knowledge
Engineering. Review, 21, pp 25-47, 2006.

[19] By S. Y. Kung Unsupervised learning models for cluster
analysis pp. 139-140 Kernel Methods and Machine
Learning Online Publication 2014

[20] EMC Education Services Data Science and Big Data
Analytics, 2015, EMC Corporation.