Green Referential Data Base for the Indian Stock Market

Krishna Kumar Singh
Research Scholar
Dept. of Computer Science,
GBPEC, Ghurdauri
Pauri Garhwal, Uttarakhand

Priti Dimri, Ph.D
Associate Professor and Head
Dept. of Computer Science
GBPEC, Ghurdauri
Pauri Garhwal, Uttarakhand

Soumitra Chakraborty
Associate Professor and Dean,
Dept. of Computer Science
Sriram Institute of Mangt. & Tech
Greater Noida, U.P

ABSTRACT

A traditional database methodology has been used in the Indian stock market. Forecasting of the market is not only based on the prices of the stocks but also on other integrated information like socio-economic factors, prices, politics etc. Changing data format and its behavior requires a new methodology to handle and integrate. Investments solely depend upon efficiency and accuracy of the data. Data required for the stock market decision making process is generated from every event and all events have its own impact on the market. Irrespective of the nature of the event and its format, market requires data integration of all kind. Because of acute scarcity of natural resources, processing of the stock market data requires green methodologies which contribute to save energy, power, time, space etc. Fractal behavior of the market shows repetition of the stock prices again and again. Large amount of space, time, power etc have been utilized to store and process these repetitive data. Referential data base is one of the answers to this problem. This paper proposes referential data base for the stock market prices without compromising efficiency and accuracy for market forecasting methods.

Keywords: Stock Market, Forecasting, Referential database, Green methodologies.

1. INTRODUCTION

Nature and behavior of data for stock market forecasting is changing with time very fast. Globalization has a great impact on stock market. Events in one part of the globe have direct impact on other part of stock market. Price variations in the stock market not only depend on the company's performance but also socio – economic dynamics of the economies. The data is piling at an alarming rate and until all the data is considered, prediction of the market is impossible. Discount rate of market is too high and traditional database is less relevant in this kind of scenario. There are many theories available to reduce data space and increase the processing speed of these data. This paper proposes a new methodology to reduce space of the data without compromising its forecasting capabilities. Use of Referential database is proposed to save memory for the stock market databases. Referential database is greener and efficient tool for the next generation technologies for predicting the stock market. Stock market data is being used by various players of the market. Although, exact calculation of green impact of referential database is known on case to case basis but we have tried to demonstrate its impact on small set of data from Indian stock market and on its basis the impact can be applied on the market as a whole.

2. LITERATURE REVIEW

With the rapid development and globalization of financial markets (especially emerging financial markets), financial information processing has become a hot research area due to its immense practical applications [13]. Technical trading is solely based on the price changes and its impact in the near future. Forecasting is an important activity in finance especially in stock market. Traditionally, forecasting has been done with in-depth knowledge in finance and the market. Advances in computational intelligence have created opportunities that were never there before [5]. Traditional database architecture distinguishes at least three layers: a conceptual layer describing the database schema in an abstract way, a physical layer implementing the schema and an external layer made out of views [9]. The rapid development of information technology has changed the dynamics of financial markets. Algorithmic trading enables investors to trade stocks through a computer program without the need for human interventions [8]. Because of its importance to the Financial Engineering field, there are many mathematical models for the Portfolio Optimization problem that can be solved by numerical methods. Because of these computational advantages, time-series analysis has always plays a significant role in the financial markets, where predictions (the decision making processes) are made based solely on the historical movements of the stock prices (time-series) [3]. There is dearth of literature on referential database in stock market database. Technology is crucial in forecasting methods. In a booming and developing economy like India, dealing with energy, space, time and power is highest priority. Choosing a particular technology to implement business strategy may have a significant impact a firm's stock performance [1]. E-markets can be thought of as trading environments over computer networks where millions of transactions typically occurred everyday [12]. Semantic access to databases has a long history and originated at early stages of database technology development. Unfortunately, they have not yet led to the creation of widely accepted industrial technologies [2]. Some of the researchers used a multi-stage optimized stock forecast model to grasp the changing trend of the stock market [7]. The market mechanism combined with software agents adjusts demand and supply through interaction among agents using prices as common information [6]. To gain faith in the market and increase in the participation of the individuals, an intelligent forecasting models selection system for refining portfolio structural estimation is needed. In the early stages of the market, volume of data is relatively low and economies are relatively closed. In the closed economies like India, markets were dependent on policies of the countries. Other country's policies have no or less impact on the other market. But now, economies are open to all and one's policies affect other's economies thus dynamics of data & information has been changing very fast. In this dynamic scenario, volume of data is increasing at a tremendous pace.

Future forecasting of the market requires new and advance tools & methodologies which deal with these challenges. Due to this large amount of data and information available and the pace at which the technologies are being developed, companies must innovate different database technologies in order to keep up with the latest ideas [4]. Indian Markets have fractal properties and volume will be a crucial parameter to implement fractal market hypothesis in the market. Some of the researchers follow algorithmic path to predict the market and recent explosion of interest in streaming data and online algorithms, clustering of time-series subsequences, extracted via a sliding window, has received much attention [10]. There is great need for more flexible structural similarity measures between time series sequences, which are based on the extraction of important periodic features. Non-parametric methods for accurate periodicity detection and new periodic distance measures for time-series sequences are being considered [11]. Green practices in the stock market are very less in the countries like India. Hindrances like low education rate and poor administration have played a crucial role in the implementation of greener technologies. Use of integer value data with integration of other data set will save energy, time, cost, space etc with the same efficiency in forecasting of the market.

3. REFERENCTIAL DATABSE FOR THE STOCK MARKET DATA.

In the forecasting tools of stock market, fractional value is not very significant. Integer value dataset with not duplicate value have great impact on green trend analysis with same pace and accuracy. Stock market prices are repetitive in nature and same value is stored many times. Referential database has been created with the algorithm given below –

1. Start:

2. Initialize database:

3. Initialize col(open), col(high), col(low) and col(close)

\\ in step4 all values in the matrix is converted into integer except date col.

4. NewVal = int(oldVal)

   \\ Step 5 to 8 is for comparing values and replacing duplicate values with address of previous (same) value.

5. For (start = first data of matrix; Start = eof(matrix); Start ++)

6. Compare Start(val) with all values(Matrix)

7. If found same value

8. Then replace duplicate value with previous value's address and search for the next value of the matrix.

9. Save all duplicate values with address of previous value in the database matrix.

10. Exit.

In the above algorithm, we have tried to search duplicate value (if any exist) and replace the duplicate value from the address of the previous value. Let us try to understand with the help of an example given below. We took prices of Reliance Industries Limited from national stock exchange

(NSE) from Jan/2000 to Aug/2013. A matrix has been formed with Date, Open price if the day, highest value of the day, lowest prices of the day and closing price of the day. In table 01, traditional methodology is used. Value stored in the matrix is repeated again and again with difference of fractional value.

Date	Open	High	Low	Close
03-01-2000	120	126.18	120	126.18
04-01-2000	127.55	136.25	125.5	136.25
05-01-2000	136	143.4	126.85	142.75
06-01-2000	143.95	149.75	143.95	146.48
07-01-2000	146.8	158.18	146.1	156.75
-	-	-	-	-
06-03-2000	140	143.9	124.25	130.45
-	-	-	-	-
01-06-2006	476.5	484.5	462.5	462.7
-	-	-	-	-
09-07-2013	869.1	881.8	867.05	874.8
-	-	-	-	-
01-08-2013	875.15	882.85	844.85	850
-	-	-	-	-
30-08-2013	843.8	860	828.3	853.85

Table 01: Prices value matrix of Reliance Industries Limited (RIL) from Jan/2000 to Aug/2013 (Before referential value conversion) at NSE, India [14].

Date	Open	High	Low	Close
03-01-2000	120	126	Address of 120	Address of 126
04-01-2000	127	136	125	Address of 136
05-01-2000	Address of 136	143	Address of 126	142
06-01-2000	143	149	143	146
07-01-2000	Address of 146	158	Address of 146	156
-	-	-	-	-
06-03-2000	140	Address of 143	124	130
-	-	-	-	-
01-06-2006	476	484	462	462
-	-	-	-	-
09-07-2013	Address of 869	881	867	874
-	-	-	-	-
01-08-2013	875	882	Address of 844	850
-	-	-	-	-
30-08-2013	843	Address of 860	828	853

Table 02: Price value matrix of Reliance Industries Limited (RIL) from Jan/2000 to Aug/2013 (After referential value conversion) at NSE, India [14].

In table 02, matrix contains original prices of the stock as well as the address of the previous value stored (if same).

1. To remove this discrepancy from table 01, following steps has been followed:

2. All the prices from table 01 have been converted into integer value.

3. Same value data in the matrix I then searched.

4. If data is stored in more than one place then the data in the first place is left as it is but its address is saved in place of the other duplicate entries.

5. Exit.

Number of rows in the both matrix are the same i.e. 3242.

## 4. COMPARISON OF REFERENTIAL AND TRADITIONAL DATABASE.

Scope of traditional database is very vast and applicable every where. But stock market data requires more customized database for forecasting and decision making. Very large number of data has been repeated in the stock market. Referential database doesn't repeat data in the database. Instead of data repetition in the database, it replaces address of the same data previously saved. In the referential database, it saves space without removal of data and incorporating features of both types of databases traditional as well as referential. Referential database generates same results at the time of generation of the chart for the decision makers. Referential database is suggesting methodology for reducing the quantity of data from tradition database without compromising its behavior. It provides all required features for market players. Memory required for this referential database is 39 percent less than that is required for the traditional database.

## 5. FORECASTING WITH THE HELP OF REFERENTIAL DATABASE.

For forecasting and decision making, decision support software's must generate reports with the changing pace of data. Reducing the space of data will help to increase the speed of the decision making process. Referential database will reduce space & time required with increase in speed. Fig 01 (01) & Fig 02 (01) is drawn from table 01 and Fig 01 (02) & Fig 02 (02) is drawn from table 02. Although data set is different, Fig 01(01) and Fig 01 (02) are the same in nature. Same inference has been drawn form both the figures. Similarly in Fig 02 (01) & Fig 02 (02), Chart with high, low and closing prices is drawn with table 01 & table 02 and it also demonstrated the same result. Chartists and players of market require only charts without being concerned of the back ground data processing. If charts drawn from referential dataset are similar to that drawn from the traditional database then there is no adverse impact on forecasting methodologies. Referential dataset can be equally efficient and valuable for the traders and chartists.
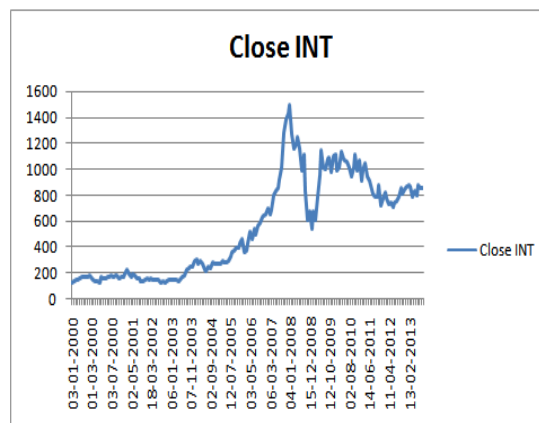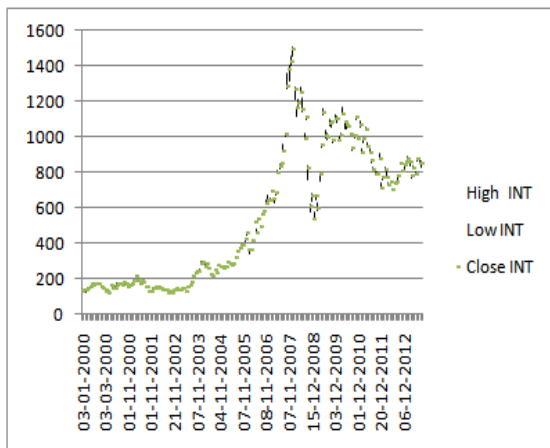

Fig. 01 (01): Line chart with Closing prices of Reliance Industries Limited with data from Jan/2000 to Aug/2013 (From table 01).
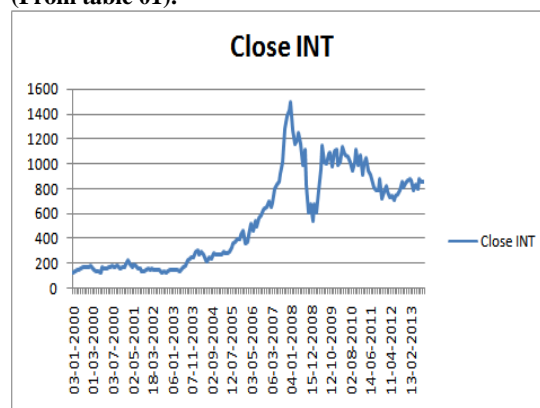
Fig. 01 (02): Line chart with closing prices of Reliance Industries Limited after referential value conversion from Jan/2000 to Aug/2013 (From table 02).
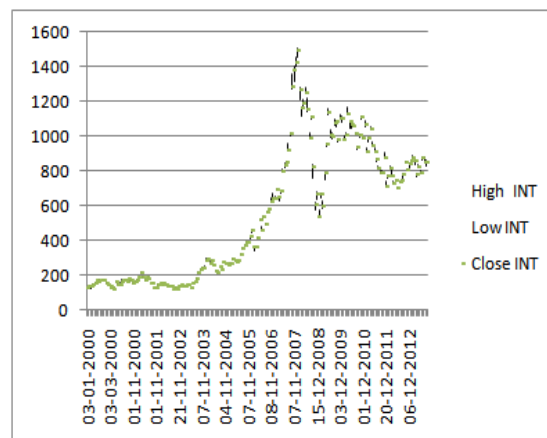
Fig. 02 (01): Chart with high, low and closing prices of Reliance Industries Limited with traditional database with data from Jan/2000 to Aug/2013 (From table 01).

Fig. 02 (02): Chart with high, low and closing prices of Reliance Industries Limited after referential value conversion with data from Jan/2000 to Aug/2013 (From table 02).

6. GREEN EVALUATION OF REFERENTIAL DATABASE

Matrix table 01, contains 3410 (Number of Rows) * 5 (Number of Columns) = 17050 data. Table 02, contains 12276 data (total number of data without duplication) and 4774 data (total number of data with duplicate). Total address replaced in place of duplicate data is 4774. Lets, x is the amount of address saved by storing one address in the place of data then total space saved is 4774 * x. More than 01 Billion data is generated per day in Indian stock market. So, Total space saved is more than 28 million * x per day. In other words, 28 percent of address has been store in the matrix. Let y amount of total memory requires storing binary number and z amount of total memory require to save the same digit in hexadecimal then total memory space saved by the method is $y - z$. Processing time and other parameters like energy and cost is also calculated on case to case basis.

7. CONCLUSION AND FUTURE WORK

The database model proposed will be helpful for the players of the stock market who provide services to their clients without being concerned with the kind of data or its format being used. Forecasting Analysts generally focus on the charts and are concerned only on the output of the charts.

By using the model proposed above we can say that:

1. Referential database can save a substantial amount of memory space.

2. Referential database has adaptability capabilities which help in integrating the existing database easily with new technologies like Cloud Computing.

3. This database will be more efficient and accurate in forecasting than the traditional database.

4. Referential database is a milestone towards green stock market database.

There is tremendous scope to work in the develop architecture and enhance referential database. It will be a milestone in the field of green computing for the financial market data.

8. REFERENCES

[1] Aurore J. Kamssu, Brian J. Reithel, Jennifer L. Ziegelmayer, "Information Technology and Financial Performance: The Impact of being an Internet-Dependent Firm on Stock Returns", Information Systems Frontiers 5:3, pp. 279–288, 2003.

[2] Chien-Jen Huang, Peng-Wen Chen, Wen-Tsao Pan : "Using multi-stage data mining technique to build forecast model for Taiwan stocks", Springer-Verlag London Limited , DOI 10.1007 / s00521 − 011 − 0628 − 0, 2011.

[3] Di WU, Gabriel Pui Cheong FUNG, Jeffrey Xu YU,QiPAN, "Stock Prediction: An event driven approach based on bursty keywords", Front. Comput. Sci. China, vol. 3(2), pp. 145–157, 2009.

[4] Eamonn Keogh,Jessica Lin, "Clustering of time-series subsequences is meaningless: implications for previous and future research", Knowledge and Information Systems, vol. 8, pp. 154–177, 2005.

[5] Edward TSANG, "Forecasting − where computational intelligence meets the stock market", Front. Comput. Sci. China vol. 3(1), pp. 53–63, 2009.

[6] Hsing-Wen Wang,"Intelligent forecasting models-selection system for the portfolio internal structure change", Springer-Verlag, DOI 10.1007/s00500-007-0177-8, 2007.

[7] Hui Ma, Klaus-Dieter Schewe, Bernhard Thalheim, Qing Wang, "A theory of data-intensive software services", SOCA, vol. 3, pp. 263–283, 2009.

[8] Ji-Yong Seo & Sangmi Chai, "The role of algorithmic trading systems on stock market Efficiency", Inf Syst Front. Vol. 15, pp. 873–888, 2013.

[9] Masayuki ISHINISHI, Yuhsuke KOYAMA and Hiroshi DEGUCHI, Hajime KITA, "A Futures Market-based Model for Dynamic Network Resource Allocation", New Generation Computing, Vol. 23, pp. 43-56 Ohmsha, Ltd. and Springer, 2005.

[10] MICHAIL VLACHOS, PHILIP S. YU, VITTORIO CASTELLI: "Structural Periodic Measures for Time-Series Data", Data Mining and Knowledge Discovery, Vol. 12, pp. 1–28, 2006.

[11] PATRICK THOMAS, "A relationship between technology indicators and stock market performance", Jointly published by Kluwer Academic Publishers, Dordrecht Scientometrics, and Akadémiai Kiadó, Budapest, Vol. 51, No. 1, pp. 319–333, 2001.

[12] Piyanath Mangkorntong & Fethi A. Rabhi, "A Domain-Driven Approach for Detecting Event Patterns in E-Markets", World Wide Web, vol. 12, pp. 69–86, 2009.

[13] Shuo BAI, Shouyang WANG, Lean YU, Aoying ZHOU, "Financial information processing and development of emerging financial markets", Front. Comput. Sci. China, vol. 4(2), pp. 185–186, 2010.

[14] www.nseindia.com