USING FRACTAL THEORY IN DATABASES DESIGN

Raul ŞERBAN*

Abstract. *The objective of this paper is to present a new method for Databases Design by using fractal theory. This solution could be implemented only for price fluctuations analytic system. This article will propose a concept to improve the price fluctuations analysis.*

Keywords: *database, fractal, analytic system, Portfolio Theory.*

1. Introduction

Perhaps the most important thing in the economy is to be able to predict more or less accurately what will happen to the market after some time. The dominant theory that was used for this was the Portfolio Theory. According to it, the probability of various changes of the market can be shown using the standard bell curve:

Figure 1. Portfolio Theory graph [3].

Assuming that the Portfolio Theory is accurate, we can say that very small changes happen most often, while very big changes happen extremely rarely. In practice this is not true. While on the bell curve, one can observe the probability of rapid changes to approach zero, they can, be seen almost every month at the real market. Consider taking a year of

* Hyperion University, Bucharest, raulserban@gmail.com

market activity and graphing the price for every month. You will get a broken line with some rises and falls. Now, if you take one of the months and graph it in a more detailed way, with every week shown, you will get a very similar line with some rises and falls. If you make it more and more detailed by showing every day, every hour, and even every minute or second you will still get the same, only smaller, rises and falls. In order to correct the Portofolio Theory, Benoit Mandelbrot introduced a new fractal theory that can be used much more efficiently than the Portfolio Theory to analyze the market.

Benoit Mandelbrot came up with a method of creating fractals that fit the above description. He based in on simple generator iteration and created base-motif fractals that could model the market. In the February 1999 issue of Scientific American, he published some of his fractal "forgeries" next to real market lines, showing how remarkably similar they were.

The concept of fractals has been spread over all fields of sciences and represents "a rough or fragmented geometric shape that can be split into parts, each of which is (at least approximately) a reduced-size copy of the whole, a property called self-similarity.

This paper will present in next chapters a concept to design database in order to create a price fluctuations analytic system.

2. Design database for price fluctuations analytic system

In our days, in economics, we search for a balance between demand and supply, but it becomes evident that in reality such balances are hardly be realized for most of popular commodities in our daily life.

The important point is that demand is essentially a stochastic variable because human action can never be predicted perfectly; hence the balance of demand and supply should also be viewed in a probabilistic way. If demand and supply are balanced on average the probability of finding an arbitrarily chosen commodity on the shelves of a store should be 1/2, namely about half of the shelves should be empty. Contrary to this theoretical estimation shelves in any department store or supermarket is nearly always full of commodities. This clearly demonstrates that supply is much in excess in such stores. Excess supply generally holds for most of commodities especially foods in economically advanced countries.

In order to design a correct database for a price fluctuations analytic system we must identified the numerical market model. The model consists of speculative dealers who transact with others simply following the basic rule “buy at a lower price and sell at a higher price”. These two threshold prices are determined at each time step by each dealer taking into account the information of past market price changes. It is shown that even a smallest limit case of 3 dealers can show chaotic behaviours, implying that the transaction's nonlinear effect is very strong [1]. If we organize this data into a database structures like a table:

Id_price	Buying price	Selling price

we can obtain in time information that can be manipulating for creating analysis. The model of the table presented can be extended in order to create an analysis based on the formula:

$$C(T) = \frac{< \Delta r(T_0 + T)\Delta r(T_0) > - \Delta r(T_0) >^2}{< \Delta r(T_0)^2 > - < \Delta r(T_0) >^2} \,. \qquad (1)$$

In the paper FRACTAL PROPERTIES IN ECONOMICS written by HIDEKI TAKAYASU, MISAKO TAKAYASU, MITSUHIRO P. OKAZAKI, KOUHEI MARUMO, TOKIKO SHIMIZU is presented a typical example of yen-dollar rate changes in 3 different time scales. Intuitively this figure demonstrates a fractal property of exchange rates in the time axis measured by ticks, namely, Mandelbrot's classical finding also holds for this contemporary market price fluctuation. The statistics of this fluctuation is very close to random walk; actually it is easy to confirm that the power spectrum of this fluctuation clearly follows an inverse square law that is almost identical to a Brownian motion. The corresponding auto-correlation function for rate fluctuation per tick defined by the equation decay very quickly.

If we extend the table in order to accumulate the information for analysis we obtain:

Id_price	Buying price	Selling price	Transaction Time

Also we can use the cash-flow model to build the structure of database and for a better analysis.

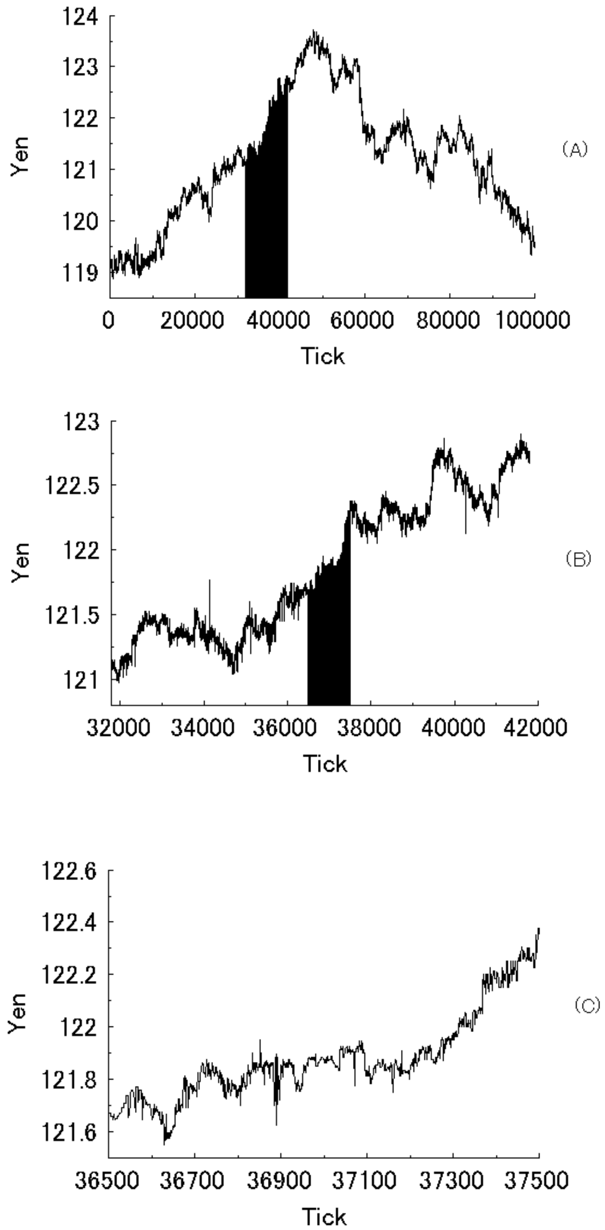

Figure 2. Fractal property in foreign currency exchange rate fluctuations. The dark part of (a) is magnified 10 times in (b), and the dark part in (b) is magnified in (c) [1].

The database model presented above is an example of organizing the data in order to obtain a price fluctuations analytic system based on fractal theory.

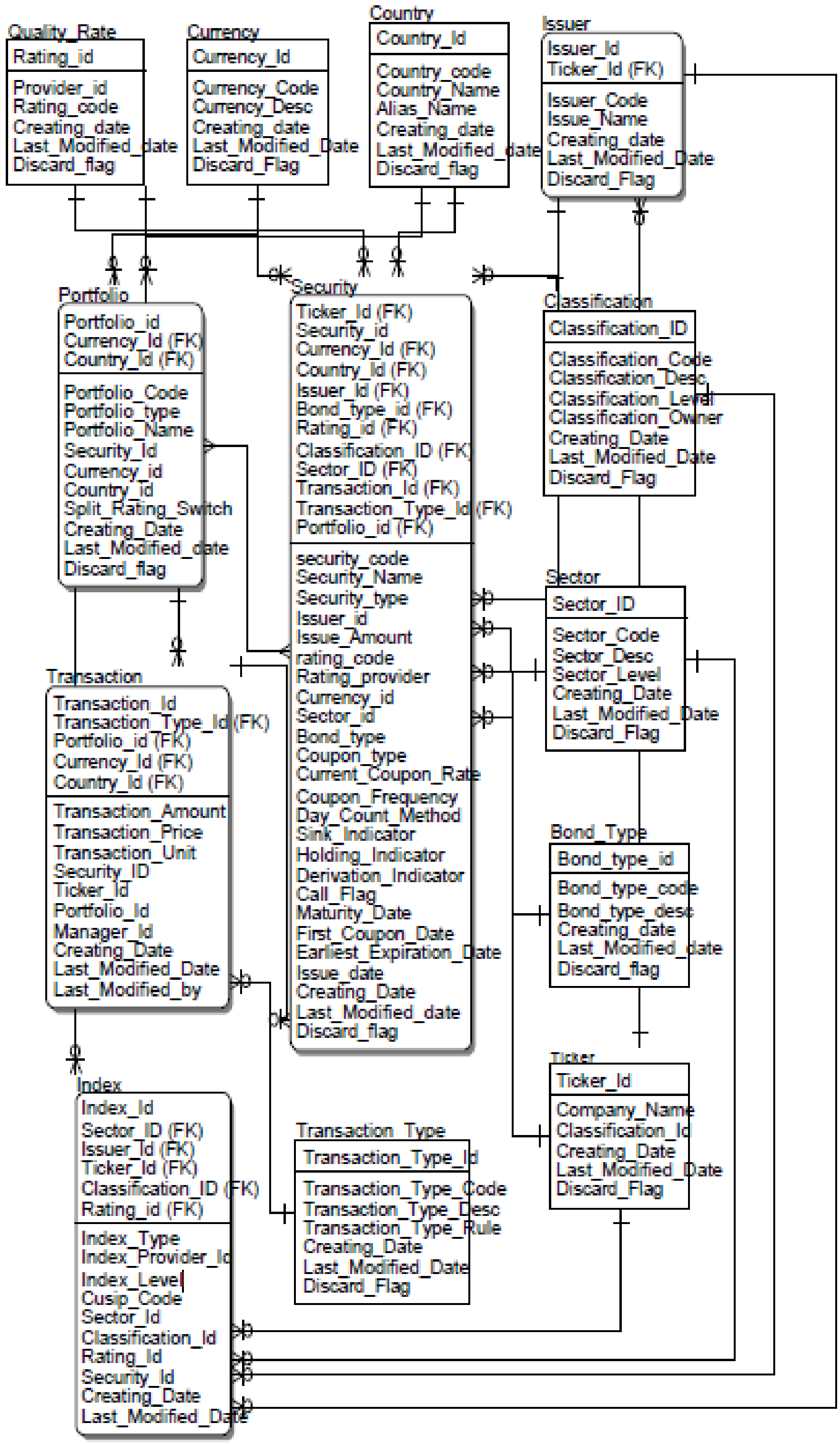

Figure 3. Cash-flow model for database design [2].

3. B-tree database indexes

In order to implement the model of database presented in previous chapter we must use b-tree. The B-tree is a generalization of a binary search tree in that a node can have more than two children.

In a tree, records are stored in locations called leaves. The starting point is called the root. The maximum number of children per node is called the **order** of the tree. The maximum number of access operations required to reach the desired leaf (data stored on the leaf) is called the **depth** (level). Some databases indexes are balanced b-trees; the order is the same at every node and the depth is the same for every leaf [4].

real tree (in nature)	b-tree
grows up	grows down
main trunk	root
branch	node
leaf	leaf

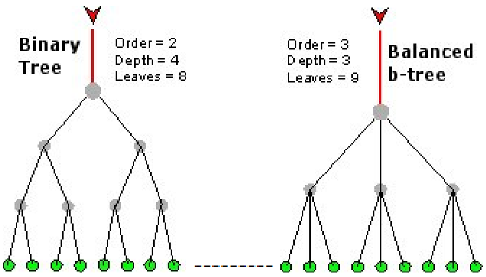

Figure 4. b-tree model [4].

The bigger the order, the more leaves and nodes you can put at a certain depth. This means that there are fewer levels to traverse to get to the leaf (which contains the data you want). In the example above and all balanced b-trees, the number of hops to a leaf = depth.

To obtain a fast response from database where most indexes are too large to fit into memory, this means that they are going to be stored on disk. Since I/O is usually the most expensive thing we can do it in a computer system, these indexes need to be stored in by an efficient I/O way. A b-tree is a good way to do this. If we make the nodes the size of a physical I/O block, it would take one I/O to move to a lower depth in the tree.

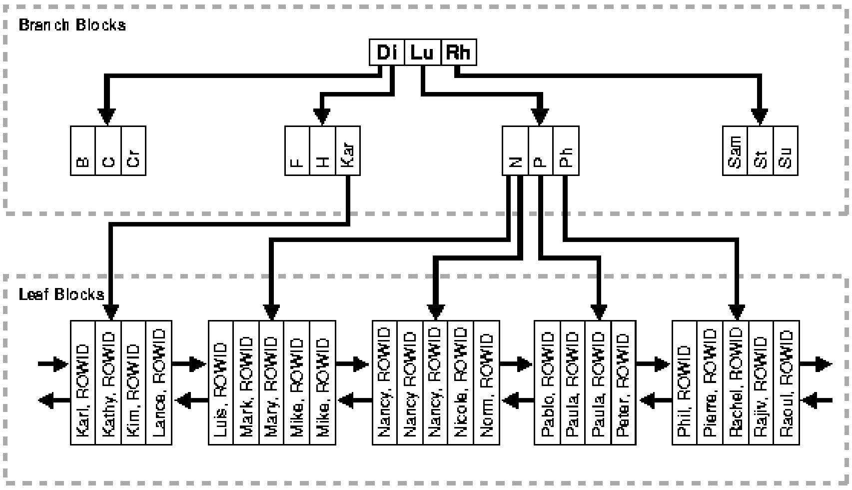

Figure 5. b-tree model organization in database design [4].

4. Conclusions

Economic systems tend to be around the critical point in general. Open markets are typical examples. When there are more buyers than sellers, the market price goes up causing decrease in the number of buyers and increase in sellers. Such well-known pulling back mechanism obviously keeps the system around the critical point. The important fact is that this criticality does not stabilize the absolute value of the market price; it only stabilizes the statistics of market price to follow the critical fluctuations. As the dealers in open markets tend to care about only the relative market price, whether it goes up or down, the absolute value is nearly meaningless for the determination of the market price. Therefore, the market prices generally wander almost randomly having statistics characterized by the critical point. From the viewpoint of fluctuations in demand and supply, company's statistics may also be considered in a similar way. When there appears a new demand in a new field of industry, companies competitively try to supply it. The winner will get a big amount of income and the company will grow. The followers will share the rest of demand and the total supply will increase. When the sum of supply becomes close to the demand, some companies cannot gain their incomes and the growth of supply is weakened. As a result demand and supply may nearly balance on average and the whole system may show critical characteristics such as fractal distributions. Therefore we must create a real solution for database design in order to use fractal theory.

REFERENCES

[1] Hideki Takayasu, Misako Takayasu, Mitsuhiro P. Okazaki, Kouhei Marumo, Tokiko Shimizu, Fractal properties in economics.

[2] Hong Zeng, *Fixed Income Database Design & Architecture*, Worcester Polytechnic Institute, 2005.

[3] http://library.thinkquest.org/26242/full/ap/ap7.html

[4] http://mattfleming.com/node/192

[5] B. B. Mandelbrot, The variation of certain speculative prices, J. of Bussiness (Chicago) **36**(1963) pp. 394-419.

[6] J. Kertesz and I. Kondor (Eds.), *Econophysics: an emerging science* (Kluwer Academic Publisher, Dordrecht, to appear).

[7] R. Mantegna and E. H. Stanley, *Introduction to Econophysics* (Cambridge Univ. Press, London, to appear).

[8] H. Takayasu, A.-h. Sato and M. Takayasu, *Power law behaviours of dynamic numerical models of stock market prices*, (to appear in ref. 2)

[9] H. Takayasu and M. Takayasu, *Critical fluctuations of demand and supply*, Physica **A** 269 (1999) pp. 24-29.

[10] D. Sornette, D. Stauffer and H. Takayasu, Market fluctuations II: multiplicative and percolation models, size effects and predictions, (Procedings book of conference in Rauischholzhausen 1999, to appear (Eds. A. Bunde) (http://xxx.lanl.-gov/abs/condmat /9909439)).

[11] H. Takayasu, H. Miura, T. Hirabayashi and K. Hamada, *Statistical properties of deterministic threshold elements – the case of market price*, Physica **A 184**(1992) pp. 127-134.

[12] R. N. Mantegna and H. E. *Stanley, Scaling behaviour in the dynamics of an economic index*, Nature376 (1995), pp. 46-49.

[13] Oracle 9i Database Administration, Published by Oracle Corp.

[14] Len Boss. From Carnegie Mellon, Attribute Driven Design Method.