PRESENTED AT THE
FINANCIAL SYSTEMS CENTER LAUNCH EVENT
POSTER SESSION
APRIL17, 2012

STEVENS INSTITUTE of TECHNOLOGY
THE INNOVATION UNIVERSITY

FINANCIAL SYSTEMS CENTER

Detection and Analysis of Rare Events in High-Frequency Financial Data

Dragos Bozdog[1], Ionut Florescu[1], Khaldoun Khashanah[2], Jim Wang[2]
[1]Department of Mathematical Sciences, Stevens Institute of Technology
[2]Program in Financial Engineering, Stevens Institute of Technology

Abstract

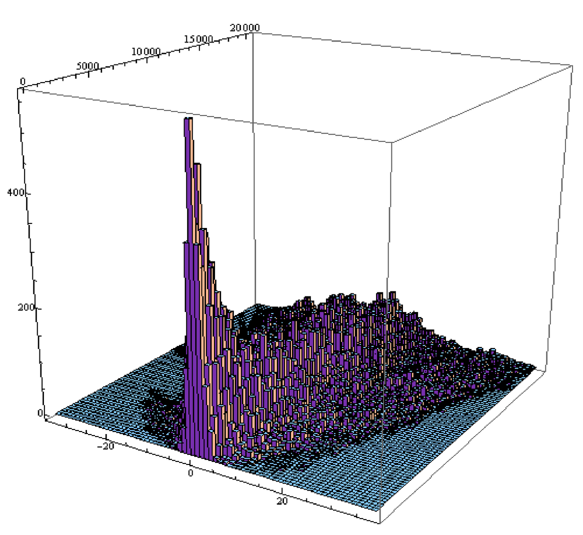
We present a methodology to detect unusual trading activity defined as high price movement with relatively little volume traded. The analysis is applied to high-frequency transactions of thousands of equities and the probability of price recovery in the proximity of these rare events is calculated. Similar results are obtained when analyzing commodities with different expiration dates. The propagation of rare events in the commodity structure and the liquidity problems are addressed.

Specific objectives of the study

- Develop a method to detect large price movements corresponding to small volume of shares traded.
- Analyze the evolution of price after these unusual events and study the probability of price recovery.
- Estimate the expected return if a trade is placed at the detected event.
- Analyze rare events propagation in futures with several expiration dates.
- Liquidity considerations using rare events & aggressor indicator

Methodology

In this analysis we use tick-by-tick data of 5,369 equities (TAQ), US and EUREX futures.



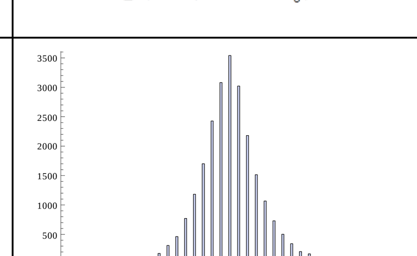
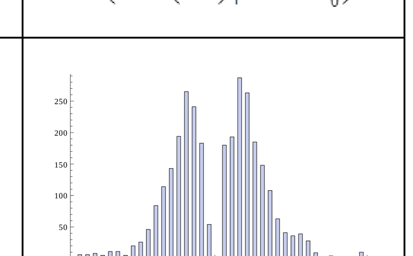
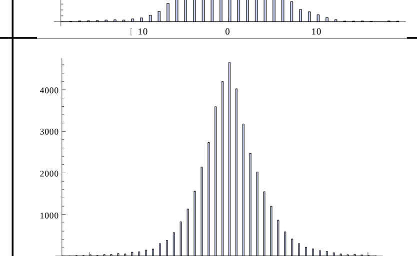
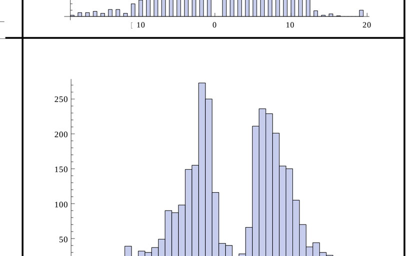
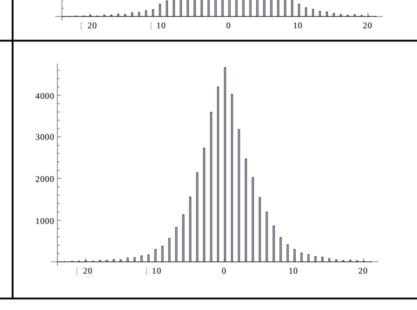
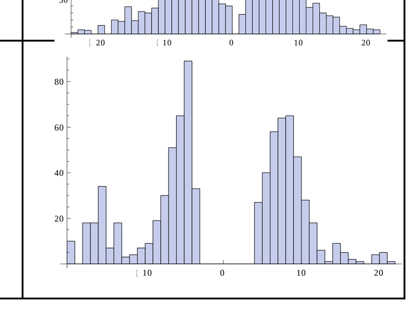
The 3D image on left represents the joint distribution of volume and price movement. The total number of pairs used for this distribution is 159,583. This is an extremely large number of data points to be calculated and analyzed for every stock and for every day.

To further simplify we retain only the maximum price movement with respect to change in volume for each window with $\Delta V < V_0$. Consequently, all the analysis that follows is based on the following distribution:

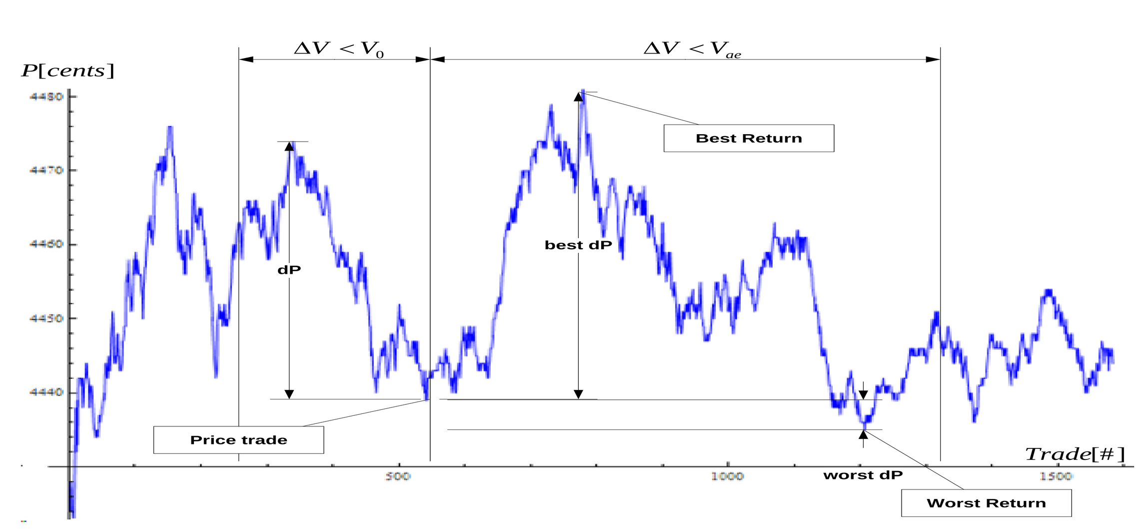
$$h(Max(\Delta P) \mid \Delta V < V_0)$$

For example for IFF stock, we obtain 1,570 observations for $V_0 = 3{,}000$ shares, 1,562 observations for $V_0 = 5{,}000$ shares, 1,544 observations for $V_0 = 10{,}000$ shares.

The proposed sampling technique generates new distributions that share similar behavior of the tails and provides a computationally feasible approach to this analysis.

$< V_0$	$g(\Delta P \mid \Delta V < V_0)$	$h(Max(\Delta P) \mid \Delta V < V_0)$
$< 3{,}000$		
$< 5{,}000$		
$< 10{,}000$		

Visualization of the quantities used in the study

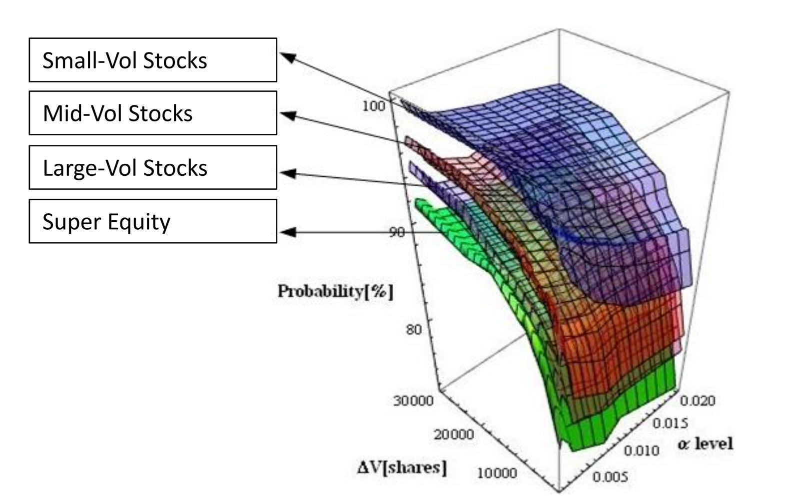


Comparative Study of Equity Groups

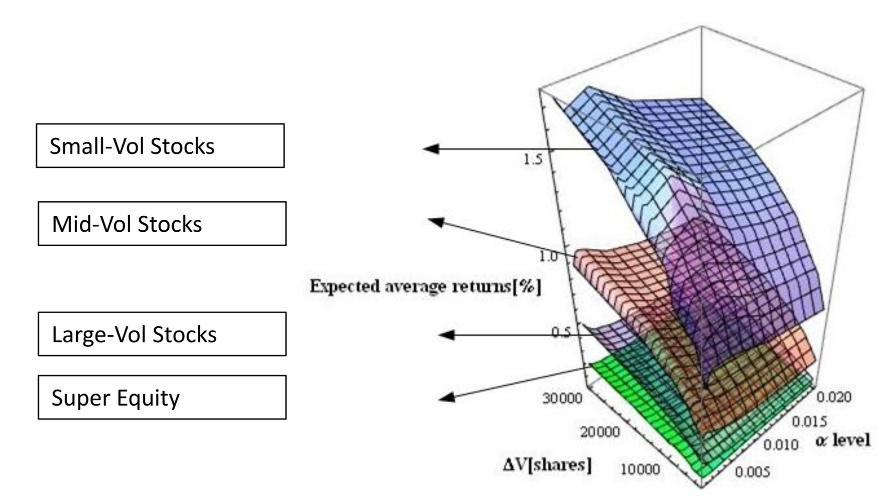
We analyze the change in price from the volume perspective. We classify stocks into classes based on the average daily traded volume. We refer to this classification as the ***multi-scale volume classification*** of the 5,369 stocks considered in this study.

Table 1: Asset Classification

	Class	Average daily volume (shares)	Number equities
1		$ADV \le 30{,}000$	1,305
2	Small-Vol Stocks	$30{,}000 < ADV \le 100{,}000$	1,088
3	Mid-Vol Stocks	$100{,}000 < ADV \le 1{,}000{,}000$	2,117
4	Large-Vol Stocks	$1{,}000{,}000 < ADV \le 10{,}000{,}000$	799
5	Super Equity	$10{,}000{,}000 < ADV$	60



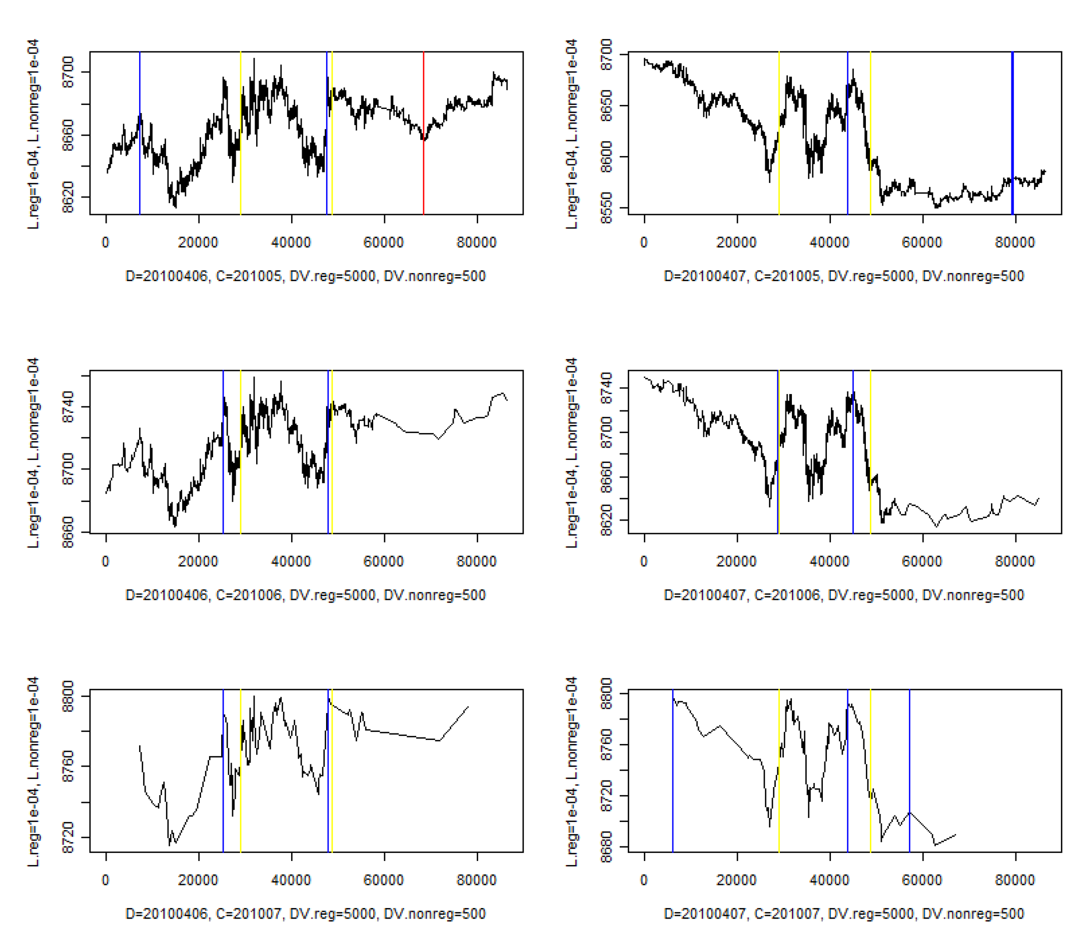
Probability of elastic behavior for each asset class



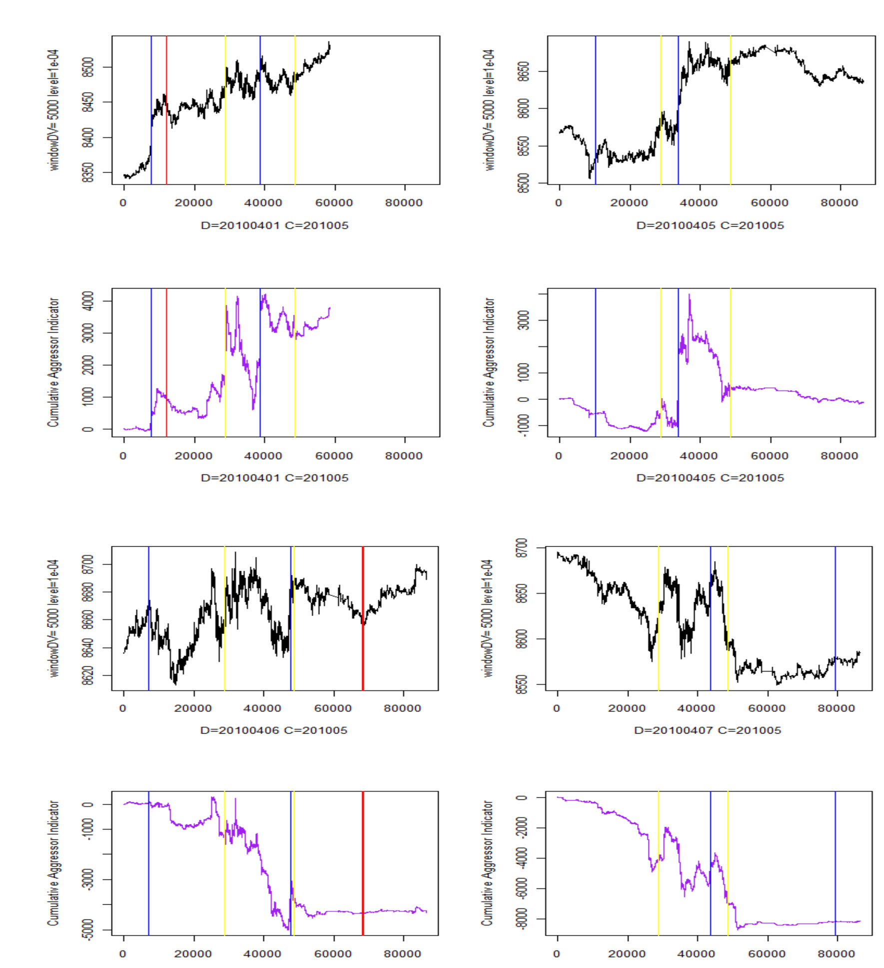
Expected returns for a simple trading rule

Application to Futures

The Rare Events Detection is implemented for several expirations dates of the same underlying futures contract. The propagation of the rare events in the future structure is influenced by the trading activity and results in asynchronous behavior.



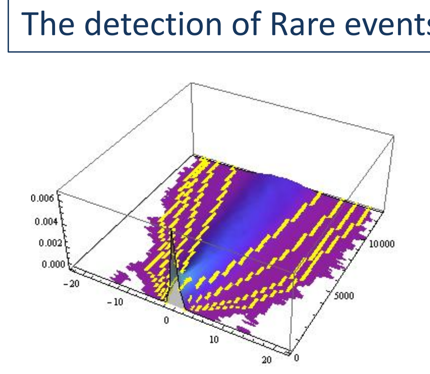
Three futures (different expirations) on the same commodity.
Blue line indicates unusual increase in price relative to the volume, Red unusual decrease. Open Pit trading hours are between the yellow lines.



The most liquid future (black line), Aggressor Indicator (purple line), for 4 different days. These two are strongly correlated, but the later is not visible to traders. Rare events are capable to detect sharp liquidity drops (bid or ask) by only analyzing the price/volume movement.

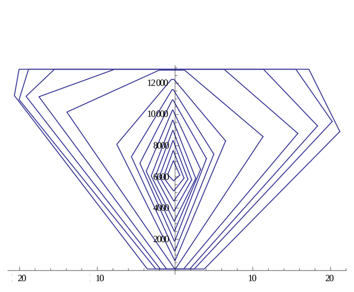
Comparison of Rare Events Detection with other, computationally expensive Algorithms

The detection of Rare events is compared with p-level points generation and zonoid trimmed regions method.



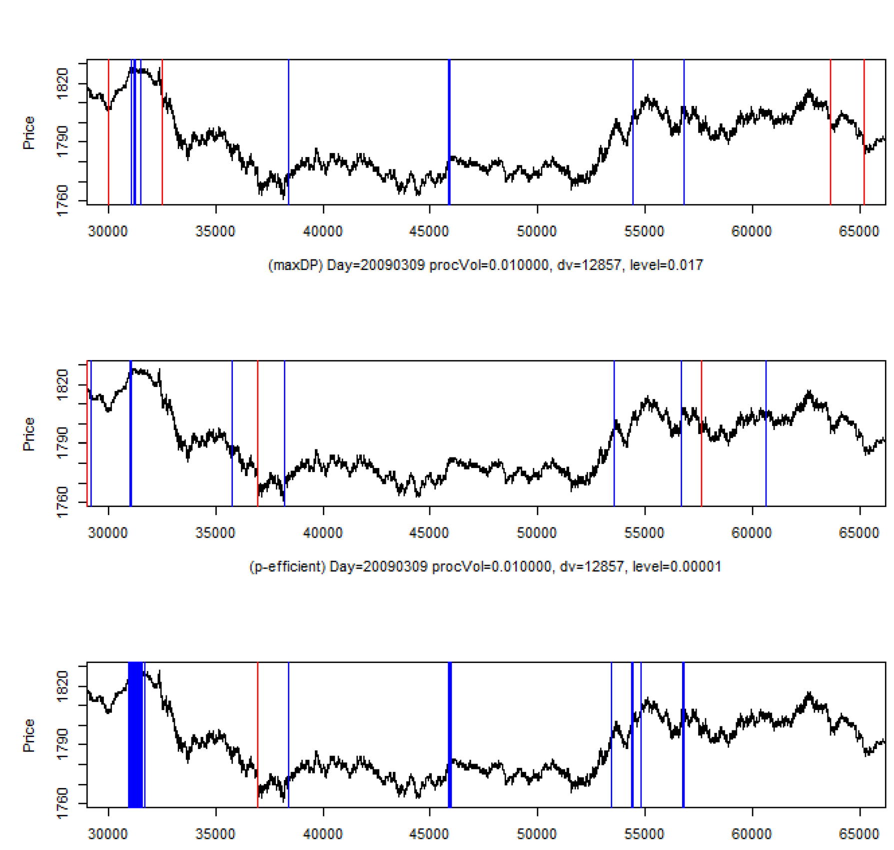
p-level points
(all points more extreme than p-efficient points)

$$D_p(x_1,\dots,x_n) = \{x_i : P(x_i \le \xi_i, i = 1,\dots,n) \le p\} \bigcup \{x_i : P(x_i \ge \upsilon_i, i = 1,\dots,n) \le 1-p\}$$



zonoid trimmed regions

$$D_\alpha(x_1,\dots,x_n) = \left\{ \frac{1}{n\alpha}\sum_{i=1}^{n}\lambda_i x_i : \sum_{i=1}^{n}\lambda_i = n\alpha, 0 \le \lambda_i \le 1 \text{ for all } i \right\}$$



Publications

1. Bozdog, D., Florescu, I., Khashanah, K. and Wang, J. (December 2011). A study of persistence of price movement using High Frequency Financial Data. Chapter in *Handbook of Modeling High-Frequency Data in Finance*. Wiley. [ISBN: 978-0-470-87688-6]
2. Bozdog, D., Florescu, I., Khashanah, K. and Wang, J. (July 2011). Rare Events Analysis of High-Frequency Equity Data. *Wilmott Journal*, Vol. 54, pg. 74-81. [DOI: 10.1002/wilm.10016]

References

1. Alfonsi, A., A. Schied and A. Schultz (2007). Optimal execution strategies in limit order books with general shape functions. http://www.citebase.org/abstract?idoai:arXiv.org:0708.1756.
2. Beaver W.H. (1968). The information content of annual earnings announcements. *Empirical research in Accounting: Selected Studies*; suppl. to *Journal of Accounting Research*, vol. 6, pp 67-92
3. Bollerslev, T. and Jubinski, D. (1999). Equity trading volume and volatility: latent information arrivals and common long-run dependencies. *Journal of Business & Economic Statistics* 17, pp. 9–21.
4. Gallant, A.R., Rossi, P.E. and Tauchen, G.E, (1992), Stock prices and volume. *The Review of Financial Studies* 5, pp. 199–242.
5. Karpoff, J. (1987). The relation between price change and trading volume: A survey, *Journal of Financial and Quantitative Analysis*, 22,March, pp 109-126.
6. Lo, A. W., H. Mamaysky, and J. Wang (2000). Foundation of technical analysis: Computational algorithms, statistical inference, and empirical implementation. *The Journal of Finance* 55 (4), pp 1705-1765.
7. Lo, A.W, and J. Wang. (2002) Trading volume: Implications of an intertemporal capital asset price model. *Advances in Economic Theory: Eighth World Congress*, pp 1-23
8. Osborne M.F.M. (1959) Brownian motion in the stock market, *Operations Research*, 7(2), pp 145-173.
9. Sun W. (2003) Relationship between Trading Volume and Security Prices and Returns, *MIT LIDS Technical Report* 2638, February 2003Area Exam.
10. Tsay, A. S. and C. Ting (2006, January). Intraday stock prices, volume, and duration: a nonparametric conditional density analysis. *Empirical Economics* 30 (4), pp 253-268.
11. Zhang, M. Y., Russell, J. R., & Tsay, R. S. (2008). Determinants of bid and ask quotes and implications for the cost of trading. *Journal of Empirical Finance* 15(4), pp 656-678.
12. D. Dentcheva, A. Prekopa, A. Ruszczynski, (2000). Concavity and efficient points for discrete distributions in stochastic programming, *Mathematical Programming*, 89 55—79.
13. R. Dyckerhoff (2000). Computing zonoid trimmed regions of bivariate data sets, Chapter in *COMPSTAT - Proceedings in computational statistics* : 14th symposium held in Utrecht, 1st Ed, Physica-Verlag HD.

Contact

Dragos Bozdog
Department of Mathematical Sciences, Stevens Institute of Technology, Castle Point on Hudson, Hoboken, NJ 07030, Phone: (201) 216-5447, Fax: (201) 216-8321
Email: dbozdog@stevens.edu , Web: http://personal.stevens.edu/~dbozdog/