Boosting Frameworks in Financial Applications: From Volatility Forecasting to Portfolio Strategy Optimization

Valeriy V. Gavrishchaka[1]

[1]Head of Quantitative Research, Alexandra Investment Management, New York, 10017

Abstract

Increasing availability of the multi-scale market data exposes limitations of the existing quantitative models such as low accuracy of the simplified analytical and statistical frameworks as well as insufficient interpretability and stability of the best machine learning algorithms. Boosting was recently proposed as a simple and efficient framework for intelligent combination of the clarity and stability of the analytical and parsimonious statistical models with accuracy of the adaptive data-driven models. Encouraging results of the boosting application to symbolic volatility forecasting have also been reported. However, accurate volatility modeling does not always warranty optimal decision making that leads to acceptable performance of the portfolio strategy. In this work, a boosting-based framework for a direct trading strategy and portfolio optimization is introduced. Due to inherent adaptive control of the parameter space dimensionality, this technique can work with very large pools of base strategies and financial instruments that are usually prohibitive for other portfolio optimization frameworks. Unlike existing approaches, this framework can be effectively used for the coupled optimization of the portfolio capital/asset allocation and dynamic trading strategies. Generated portfolios of trading strategies not only exhibit stable and robust performance but also remain interpretable. Encouraging preliminary results based on real market data are presented and discussed.

Keywords: Use “keywords” style here.

1. Introduction

Increasing availability of the multi-scale market data exposes limitations of the existing quantitative models such as low accuracy of the simplified analytical and simulation frameworks as well as insufficient interpretability and stability of the best data-driven algorithms [7]. Multivariate nature of many problems and data non-stationarity often lead to "dimensionality curse" [3] and incompleteness of the available training data. These complications are not tolerated by many mainstream machine learning and statistical algorithms.

Limitations of the existing quantitative models and frameworks are very clear in such complex and important areas as trading strategy discovery and portfolio optimization. These problems being high-dimensional by nature are especially sensitive to noise and nonstationarity of the market data. They are also technically challenging optimization problems. Emerging frameworks for the coupled optimization of the portfolio capital/asset allocation and trading strategies [2] are potentially even more challenging and will require novel approaches and algorithms. Interpretability of the obtained results is another important requirement (especially in financial applications) that is not properly addressed by the majority of the existing frameworks.

One of the machine learning approaches to compensate for deficiency of the individual models is to combine several models to form a committee [3]. Committee can be constructed using variety of ensemble learning algorithms [3,8,19]. Committee could effectively compensate for limitations of the individual models caused by both incomplete data and specifics of the algorithms (e.g., local minima in the NN error surface). Advantage of the ensemble learning approach is not only possibility of the accuracy and stability improvement but also its ability to combine models of different types and complexity [7]. This latter feature can significantly improve explanatory power of the combined model if building blocks are sufficiently simple and well-understood.

However, ensemble learning algorithms can be susceptible to the same problems and limitations as standard machine learning techniques. Therefore, optimal choice of both base models and ensemble learning algorithm with good generalization qualities and tolerance to data incompleteness and dimensionality is very important.

A very promising ensemble learning algorithm that combines many desirable features is adaptive boosting [14,15,17]. Boosting and its specific implementations such as AdaBoost [6] have been actively studied and successfully applied to many challenging classification problems [4,13]. One of the main features that sets boosting aside from other ensemble learning frameworks is that it is a large margin classifier [14] similar to support vector machine (SVM) [18]. This ensures superior generalization ability and better tolerance to incomplete data compared to other ensemble learning techniques.

In most cases discussed in the literature, boosting is used to combine data-driven models based on machine learning algorithms (classification trees, NNs, etc.). In our recent article [7] we have stressed out advantages of the boosting framework that allows intelligent combination of the existing analytical and other simplified parsimonious models specific to the field of interest. In this way one can combine clarity and stability typical for analytical and parsimonious parametric models with a good accuracy usually achieved only by the best adaptive data-driven algorithms. Moreover, since the underlying components are well-understood and accepted models, obtained ensemble/committee remains interpretable unlike typical data-driven models.

Previously we have outlined several boosting-based modeling frameworks that can be used in financial applications [7]. One of them has been used to generate boosted collection of GARCH-type models for symbolic volatility forecasting of the stock market. Preliminary results indicated that boosted collection consistently outperforms (in-sample and out-of-sample) both single best model in the collection and classical GARCH(1,1) model. The same technique could be used to combine other models for symbolic time series forecasting.

Volatility model is an important component of many trading and risk management systems. However, accurate volatility forecasting alone does not always warranty optimal decision making that leads to acceptable performance of the portfolio strategy. Reliable approaches for the direct optimization of trading strategy and portfolio allocation are often required in practice.

Distinct feature of the trading strategy and portfolio optimization is that standard classification task required for classical adaptive boosting framework cannot be directly formulated. However, we will show that boosting can still be successfully used in this case. A weighted linear combination of the base strategies produced by the boosting algorithm has a direct and simple interpretation as a portfolio of complimentary trading strategies.

Another important feature of the proposed framework is that required CPU time increases linearly with a size of a base strategy pool and challenging optimization in a very high dimensional space is not required. Final portfolio size (not related to the pool size) can be easily controlled by a single regularization parameter to avoid overfitting. All this is in contrast to other portfolio strategy optimization frameworks where "dimensionality curse" is a major problem in many practical settings.

In the next section we give a short summary of the practical limitations of the existing frameworks for trading strategy and portfolio optimization. Short description of the standard algorithm for regularized adaptive boosting (AdaBoost) follows after that. Two other sections present proposed boosting framework for optimization and results of its application to real market data. Finally, short summary and discussion conclude the paper.

2. Limitations of the existing frameworks for trading strategy and portfolio optimization

Discovery and optimization of the dynamic trading strategies and efficient portfolio asset allocation and optimization are two intrinsically related fields. However, many practitioners and researchers consider these two problems as separate and are often biased in their opinions about comparative value of these two fields.

Development of dynamic trading strategies often rely on technical analysis [10,11,16] that assumes existence of occasional or regular intermittent periods of market inefficiency and predictable patterns. Even the most successful trading strategy is often characterized by insufficient and simplified risk management that may lead to occasional large losses.

On the other hand, portfolio selection and optimization frameworks [1,9,12] are focused on complex risk management techniques but consider only buy&hold strategies completely ignoring opportunities available from the dynamic trading. Existing risk management frameworks are useful to control portfolio exposure to particular markets and instruments. However, these frameworks are often useless in estimating risk of dynamic trading strategy, since they are based on assumptions of static position and purely stochastic time series of financial instruments excluding any possibility of predictable patterns.

Traditionally, dynamic trading strategies are constructed from the established market indicators that are able to detect trends, oscillations, and other market patterns [10,11]. Optimization procedures may involve parameter adjustments of these indicators for a particular company/index and/or market regime. For example, one can adjust the length of the time period in the moving average low-pass filters or "overbought/oversold" thresholds in typical oscillators [10]. More sophisticated optimization may involve simultaneous parameter optimization and simple rule combination into a complex strategy from a given pool of established rules/indicators [10,16].

The advantage of the trading strategies constructed from the well-known market indicators is their clarity and interpretability. This feature has a natural appeal for practitioners and decision makers. A common disadvantage is limitations in performance and stability of such strategies. It is usually difficult to optimize a single simple strategy to achieve an acceptable performance on a long time interval with many different intermittent regimes.

On the other hand, a simple strategy optimized on small data sample, covering only certain market regime, may become extremely risky when regime switching occurs. Strategies based on "AND" combination of simple rules can often be more stable compared to their individual components. However, the combined strategy of this type may be inactive during many valuable trading opportunities.

A different set of strategies is based on machine learning and statistical algorithms of the "black box" type (NN, SVM, etc.). They often rely on construction of adaptive time series predictors combined with simple entry/exit rules [16]. Due to their adaptive nature, these strategies may be less sensitive to market regime switching and in many cases can demonstrate excellent performance. However, as all "black box" algorithms, they often have limited interpretability and may experience all the problems associated with typical adaptive algorithms including occasional instabilities.

In contrast to trading strategy discovery, the problem of portfolio optimization focuses on comprehensive risk measures and usually considers only buy&hold strategy. Models for portfolio selection and optimization have evolved over the years from early mean-variance formats based on Markowitz's pioneering work [12] to more recent scenario-based stochastic optimization forms [2,9].

The common thread in all these models is the minimization of some measure of risk (uncertainty), while simultaneously maximizing expected return. In most frameworks the risk metric is a function of the entire range of possible portfolio returns. For example, overall portfolio variance is used in a mean-variance framework while concave utility functions are applied across the set of all possible outcomes in stochastic programming frameworks.

However, complete ignorance of dynamic trading strategies could often lead to significantly sub-optimal portfolios from both down-side risk control and expected return perspectives. For example, certain oscillating market regimes can be effectively exploited by dynamic oscillator-type strategies [10,11], while return of the buy&hold strategy will be minimal in such cases (even with dynamic rebalancing). Advantages of the dynamic trading could be especially evident for short-to-medium time horizons (from several months to 1-2 years).

It is usually not possible to achieve a desirable market-neutral performance using portfolio based solely on buy&hold strategy. Coupled portfolio and dynamic trading strategy optimization should provide a much more powerful framework to construct market neutral portfolio strategies with acceptable returns.

Cointegration-based frameworks for selection and optimization of the market-neutral portfolios have been proposed recently [2]. Concept of cointegration and related statistical tests have been introduced by Engle and Granger [5] as tools for finding stable dependencies between different econometric time series. Cointegration relationship is able to explain the long run behavior of cointegrated time series, while correlation usually lacks stability, being only a short-run co-dependency measure.

Cointegration-based frameworks have many advantages over classical portfolio selection procedures. However, their application is limited to buy&hold and pairs trading strategies [2]. Also, even though time series co-dependency found from cointegration is more stable than correlation, a sudden switch from such dynamics is observed quite often.

More universal frameworks capable to discover and optimize portfolio strategy that combine market neutrality with stable down-side risk protection and unlimited growth are required. As discussed in the following sections, we propose such a framework based on boosting. It can be applied to any collection of base trading strategies and financial instruments. Our framework offers an open architecture for incorporation of any particular preferences and apriory beliefs.

3. Regularized adaptive boosting for classification

The basic idea of boosting and other ensemble learning algorithms is to combine relatively simple base hypotheses/models for the final prediction (e.g., classification) [14,19]. In most algorithms base hypotheses are linearly combined. Intuitively, combining multiple models only helps when these models are significantly different from one another and each one treats a reasonable portion of the data correctly. Ideally, the models should complement one another, each being an expert in a part of the domain where performance of other models is not satisfactory.

The boosting method for combining multiple models exploits this insight by explicitly seeking models that complement one another [14,15,17]. Boosting is iterative procedure where each new model is influenced by the performance of those built previously. Boosting encourages new models to become experts for instances handled incorrectly by earlier ones. Base models are weighted by their performance and weighted linear combination of these models is used as a final expert.

The initial motivation for boosting was a procedure that combines the outputs of many "weak" classifiers to produce a powerful committee [14,15,17]. The purpose of boosting is to sequentially apply the weak classification algorithm to repeatedly modified versions of data, thereby producing sequence of weak classifiers. Data modification is achieved by increasing weights for incorrectly classified samples.

Empirical comparative studies of different ensemble methods often indicate superiority of boosting over other techniques [e.g., 4,13]. Using PAC (probably approximately correct) theory, it was shown that if base learner is just slightly better then random guessing, AdaBoost (a specific boosting implementation) is able to find combined hypothesis with arbitrary high accuracy [17]. Later it was found that boosting is a large margin classifier and that SVM and AdaBoost are intimately related [14]. Among other things this ensures robust generalization ability of boosting and partially explains its superiority over other ensemble techniques.

Many traditional statistical and machine learning techniques often have the problem of "dimensionality curse" [3]. The upper bound of the boosting generalization error depends on the margin and not on the dimensionality of the input space [14]. Hence it is easy to handle high-dimensional data and learn efficiently if the data can be separated with large margin.

For our purposes it is sufficient to describe boosting algorithm only for two-class classification problem, where classifier outputs either +1 or -1. Regularized AdaBoost for two-class classification consists of the following steps [14]:

$$w_n^{(1)} = 1/N \tag{1}$$

$$\gamma_t = \sum_{n=1}^{N} \left(w_n^{(t)} y_n h_t(x_n) \right) \tag{2}$$

$$\alpha_t = \frac{1}{2} \ln\left(1 + \frac{\gamma_t}{1-\gamma_t}\right) - \frac{1}{2} \ln\left(1 + \frac{C_t}{1-C_t}\right) \tag{3}$$

$$w_n^{(t+1)} = w_n^{(t)} \exp\left(-\alpha_t y_n h_t(x_n)\right) / Z_t \tag{4}$$

$$f(x) = \sum_{t=1}^{T} \alpha_t h_t(x) / \sum_{t=1}^{T} \alpha_t \tag{5}$$

Here N is a number of training data points, x_n is a model/classifier input set of the n-th data point and y_n is the corresponding class label (i.e., -1 or +1), T is number of boosting iterations, $w_n^{(t)}$ is a weight of the n-th data point at t-th iteration, Z_t is weight normalization constant at t-th iteration, $h_t(x_n)$->[-1; +1] is the best base hypothesis/model at t-th iteration, C is a regularization (soft margin) parameter, and $f(x)$ is a final weighted linear combination of the base hypothesis.

Boosting starts with equal and normalized weights for all training data (step (1)). Steps (2)-(4) are repeated at each iteration until stop criteria $\gamma_t < C$ or $\gamma_t = 1$ occurs. It is clear from step (4) that at each boosting iteration, data points misclassified by the current best hypothesis/model (i.e., $y_n\ h_t(x_n) < 0$) are penalized by the weight increase for the next iteration. Step (5) represents the final combined (boosted) model that is ready to use. The model classifies unknown sample as class +1 when $f(x) > 0$ and as -1 otherwise. Details of the more general versions of the regularized AdaBoost and its extensions are given in [14].

Original AdaBoost algorithm [6] can be recovered from (1)-(5) for $C = 0$. Similar to SVM, regularization parameter C represents so-called soft margin. Soft margin is required to accommodate the cases where it is not possible to find a boundary that fully separates data points from different classes. Regularization (soft margin) is especially important for financial applications where large noise to signal ratio is a typical case.

4. Boosting framework for portfolio strategy optimization

Boosting frameworks can be used in financial applications in a variety of conceptually different ways [7]. The most appealing feature of boosting, especially for financial applications, is that it combines the power of the advanced learning algorithm with the ability to use existing models specific to the field of interest. Among other things, this allows to keep the final model interpretable and to integrate a priory problem-specific knowledge in a natural way.

In the case of well-defined classification problems application of the standard boosting procedure (1)-(5) is straightforward. Recently, a successful application of boosting to symbolic volatility forecasting (a typical classification problem) has been demonstrated [7]. The same technique could be used to combine other models for symbolic time series forecasting.

Similarly, it would be desirable to have a boosting-like framework for the discovery and optimization of the market-neutral portfolio strategy based on existing and well-understood trading strategies. Unfortunately, straightforward application of the standard boosting framework for trading strategy and portfolio optimization is not possible, since it is a direct optimization and not classification problem. However, we argue that for a large class of objective functions, boosting for classification (1)-(5) can be efficiently used as a basis for the framework that can be labeled as "boosting for optimization".

One of the natural and robust objectives for trading strategy optimization is to require returns (r) generated by the strategy on a chosen time horizon (τ) to be above certain threshold (r_c). By calculating strategy returns on a series of intervals of length τ shifted with a step $\Delta\tau$ and encoding them as +1 (for $r >= r_c$) and -1 (for $r < r_c$), one obtains symbolically encoded time series (distribution) of strategy returns.

Contrary to symbolic time series forecasting discussed in [7], here the purpose is not to correctly classify (between +1 and -1), but rather to increase the number of +1 samples (i.e., the number of cases with supercritical returns). This can still be considered as classification problem with potentially uneven sample number between two classes. Namely, we would like maximum number of samples to belong to a single class ($r >= r_c$) and incorporate this into boosting operation (1)-(5) (i.e., -1 output is considered as misclassification). In such setting, boosting (1)-(5) provides a framework for optimization, where maximization objection function is a "hit rate", i.e., number of +1 samples divided by the total number of samples.

Objective function described above is not only suited for boosting framework but is also appealing for practical applications. Indeed, it does not rely on any assumptions about return distribution as many standard measures do (i.e., Sharpe ratio). Instead, it naturally incorporates required investment horizon (τ) and focuses on maximization of the lower end of the profit and loss (P&L) distribution. Our empirical results suggest that using this objective function one can effectively discover realistic market-neutral strategies with downside protection and no significant upside limitations.

Symbolic encoding and corresponding objective function can be based on any complex condition that combines different measures of profit maximization and risk minimization specified by the utility function of interest. This can be easily achieved with combination of several simple conditions.

In the case of trading strategy optimization, the final usage of boosting output is very natural and different from the classical case of boosting for classification. Instead of using weighted linear combination (5) of the base hypothesis as a final model for classification, one uses boosting weights to construct portfolio of strategies. The initial capital is distributed among different base strategies in amounts according to the weights obtained from boosting.

Boosting can be used to discover optimal combination of different dynamic trading strategies for a single financial instrument as well as simultaneous combination of trading strategies and different instruments. In both cases, boosting steps (1)-(5) are applied to a pool of base strategies $\{S_i(\boldsymbol{p}_i)\}$, where $\boldsymbol{p}_i$ is a vector of adjustable parameters for strategy S_i. However, in the first case all base strategies are applied to a time series of a single instrument, while set $\{S_i\}$x$\{I_j\}$ of all possible pairs of strategies (S_i) and instruments (I_j) should be used in the latter case.

For a pure technical trading with equities or indexes, pool of base strategies can include various types of low-pass filters for trend detection (moving averages, wavelet filters, etc.), oscillators and cycle-based strategies (stochastic, relative strength index, wavelet filter banks, etc.), and many others. Each entry/exit strategy should be paired with risk management exits such as different types of parametrized trailing stops [10].

Optimal portfolio of dynamic trading strategies for a single instrument can be considered as a new "quasi-instrument" with a desirable feature of down-side protection. Different quasi-instruments can be combined in a portfolio for further reduction of volatility and improvement of down-side protection. Unlike other types of portfolios that directly based on correlations and other measures of instrument co-dependency, this portfolio will often remain stable when historical co-dependency patterns change.

The size of the final portfolio of trading strategies can be controlled directly by limiting number of boosting iterations suggested by problem-specific heuristic for a given regularization parameter C or by changing C to activate stop condition at the required iteration number.

The proposed boosting framework can work with large pools of base trading strategies and financial instruments that are often prohibitive for other portfolio optimization frameworks. The size of the boosted portfolio is not related to the size of the pool of base strategies, i.e., there is no direct threat of overfitting. Besides that, base strategies are optimized one at a time, i.e., there are no principal or technical problems of optimization in a very high dimensional space that are typical for other frameworks.

5. Application example

In this section we demonstrate application of boosting for optimization using ~10 years of the

Nikkei500 daily data. This index offers combination of challenging complexity and absence of the direct trading vehicle to avoid any proprietary information disclosure. Similar results have been obtained for many traded indexes and stocks.

To demonstrate power of the boosting itself, a base strategy pool is restricted to a single simple strategy that is exponential moving average EMA(n,a) for entry combined with adaptive trailing stop ATS(m,α) for exit. Entry into long (short) position occurs when $EMA_t > EMA_{t-1}$ ($EMA_t < EMA_{t-1}$). Here n is a number of points to be averaged and a is a smoothing constant. Long (short) position is closed when intraday price fells below $c_{max}[1-\alpha\sigma(m)]$ (jumps above $c_{min}[1+\alpha\sigma(m)]$). Here c_{max}/c_{min} is a local max/min of the closing prices and $\sigma(m)$ is a daily return standard deviation computed using m last points. Index is traded as a regular share with transaction cost 1c/share. Profitability of the obtained strategy survives much larger transaction costs.

Boosting (1)-(5) with C=0.1 and T=7 is applied to the described base strategy. Binary hit rate objective function with r_c=15% is used (τ=126 days, $\Delta\tau$=40 days). Training data period is 1995/07 - 2003/07. The result of boosting is portfolio of 7 complimentary multi-scale strategies [EMA(n,a), ATS(m,α)], where n_{min}=22, n_{max}=170, a~0.97 (i.e., EMA(n,a) ~ MA(n)), m_{min}=7, m_{max}=37, α_{min}=0.5, and α_{max}=1.

Distributions of annualized returns of different strategies for horizon of τ=126 days are plotted in figure 1, where period from 1995/07 to 2004/11 is covered. Complexity of the index time series is evident from the distribution of the buy&hold strategy (dotted line), which has almost equal number of negative and positive returns. Performance of the single optimized [EMA,ATS] strategy (dashed-dotted line) is better. However, it still has ~ 10% of periods with negative returns.

Only boosted portfolio (dashed line) demonstrates very stable performance with significant positive returns at all periods. It is clear that this stability is achieved at the expense of some attenuation of the most profitable periods which is a fair price for the down-side protection.

In this example we also have a desired byproduct of boosting operation that is often encountered. It turns out that considered boosted portfolio of the long/short strategies uses maximum 50% of the available capital at all times, i.e., there is always some compensation of the long and short signals generated by different strategies from the portfolio. It means that returns of the boosted portfolio should be scaled up by a factor of 2 (solid line in figure 1).

Figure 1 covers both in-sample (training) periods and up to 2 years of out-of-sample data. The final goal

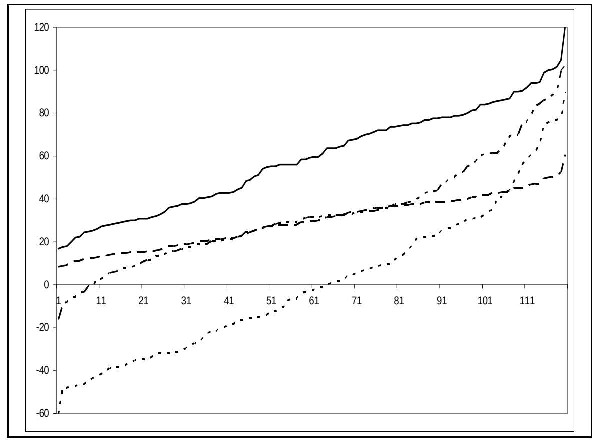

Fig. 1: Distribution of the annualized returns (%): Boosted portfolio of strategies (solid line) and buy&hold strategy (dotted line).

is robust out-of-sample performance that is often the case with boosted portfolios. Out-of-sample P&L times series (2003/11-2004/11) of the scaled boosted portfolio is plotted in figure 2 (solid line). Trading starts 6 month after the last training point. It demonstrates excellent performance in a rather difficult market regime as evident from the poor performance of the buy&hold strategy (dotted line).

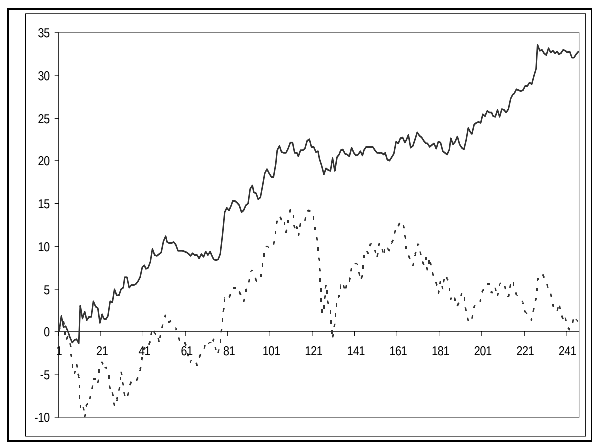

Fig. 2: Out-of-sample P&L(%) time series: Boosted portfolio of strategies (solid line) and buy&hold strategy (dotted line).

Finally, we illustrate mechanism of stability and robustness of boosted portfolios. As obvious from the very boosting operation, obtained single strategies are locally uncorrelated and capable to support each other through many different market regimes. Annualized returns of two strategies from the current portfolio are plotted in figure 3 chronologically. One strategy (solid line) operates on large scales (n=170), while the other (dotted line) works on smaller scales (n=22). It is clear from figure 3 that these two strategies are complimentary to each other and locally uncorrelated.

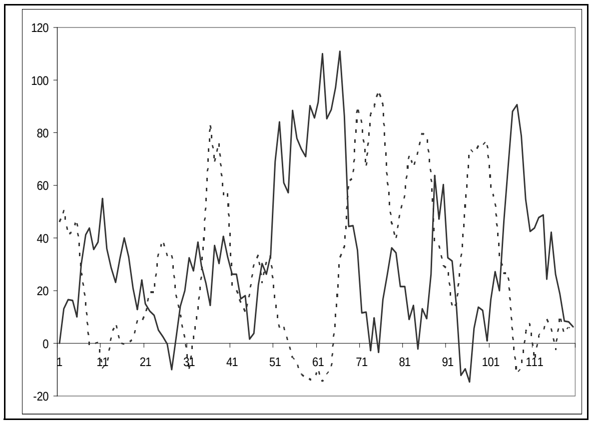

Fig. 3: Shifted annualized returns (%) of 2 individual strategies from the boosted portfolio (in chronological order): Large-scale (solid line) and small-scale (dotted line).

6. Discussion and Conclusions

Limitations of the existing frameworks for portfolio and trading strategy optimization have been identified. A generic boosting-based framework for the discovery and optimization of the portfolio strategy that can effectively address many key issues has been proposed.

Due to inherent adaptive control of the parameter space dimensionality, this framework can work with large pools of base trading strategies and financial instruments that are usually prohibitive for other portfolio optimization frameworks. Given that initial pool consists of simple standard strategies, generated portfolio strategy remains clear and interpretable.

Details of the novel framework application have been demonstrated using real market data (Nikkei500). It has been shown that boosting can effectively discover complimentary trading strategies to form market-neutral portfolio that can deliver very stable returns for short-to-medium investment horizons. This goal is usually not attainable by either individual dynamic strategies or by static buy&hold strategy.

Proposed framework can also be used for the discovery and optimization of the multi-agent strategies in other fields: operation research, medicine, computational biology, etc. Further details will be published elsewhere.

7. References

[1] D. Bai, T. Carpenter, J. Mulvey, "Making a case for robust optimization", *Management Science*, 43, 895, 1997

[2] A.N. Burgess, "Using cointegration to hedge and trade international equities", in "*Applied quantitative methods for trading and investment*", edited by C. Dunis, J. Laws, and P. Naim, *Wiley*, 2003.

[3] C.M. Bishop, "Neural Networks for Pattern Recognition", *Clarendon Press*, *Oxford*, 1995

[4] H. Drucker, et al., "Boosting and other ensemble methods", *Neural Computation*, 6, 1994

[5] R.F. Engle, and C.W.J. Granger, "Cointegration and error-correction: Representation, Estimation and Testing", *Econometrica*, 55, 251, 1987.

[6] Y. Freund, and R.E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting", *Journal of Computer and System Sciences*, 55, 119, 1997

[7] V.V. Gavrishchaka, "Boosting-based frameworks in financial modeling: Application to symbolic volatility forecasting", *Advances in Econometrics*, 20, 2005 (in print)

[8] T. Hastie, R. Tibshirani, and J. Friedman, "The Elements of Statistical Learning: Data Mining, Inference, and Prediction", *Springer*, 2001

[9] R.S. Hiller, and J. Eckstein, "Stochastic dedication: Designing fixed income portfolios using massively parallel benders decomposition", *Management Science*, 39, 1422, 1993

[10] J.O. Katz, and D.L. McCormick, "The encyclopedia of trading strategies", *McGraw-Hill*, 2000

[11] P.J. Kaufman, "Trading systems and methods", *Wiley*, *New York*, 1998

[12] H.M. Markowitz, "Portfolio selection: Efficient diversification of investments", *Wiley*, *New York*, 1959

[13] D. Opitz, and R. Maclin, "Popular ensemble methods: An empirical study", *Journal of Artificial Intelligence Research*, 11, 169, 1999

[14] G. Ratsch, "Robust boosting via convex optimization: Theory and applications", *PhD thesis*, *Potsdam University*, 2001

[15] R.E. Schapire, "The design and analysis of efficient learning algorithms", *PhD thesis*, *MIT Press*, 1992

[16] F. Tay, L. Shen, and L. Cao, "Ordinary shares, exotic methods: Financial forecasting using data mining techniques", *World Scientific*, 2003

[17] L.G. Valiant, "A theory of the learnable", *Communications of the ACM*, 27, 1134, 1984

[18] V. Vapnik, "Statistical Learning Theory", *Wiley*, *New York*, 1998

[19] I.H. Witten, and E. Frank, "Data Mining", *Morgan Kaufmann Publishers*, 2000