Improving portfolio optimization using weighted link prediction in dynamic stock networks

Douglas Castilho[1,2,3], João Gama[2],
Leandro R. Mundim[3], and André C. P. L. F. de Carvalho[3]

[1] Institute of Mathematical and Computer Sciences (ICMC)
University of São Paulo (USP), São Carlos, Brazil
douglas.castilho@usp.br, {mundim,andre}@icmc.usp.br
[2] Institute for Systems and Computer Engineering, Technology and Science
University of Porto (UP), Porto, Portugal
jgama@fep.up.pt
[3] Laboratory of Technology and Innovation (LATIN)
Federal Institute of South of Minas Gerais (IFSULDEMINAS)
Poços de Caldas, Brazil

Abstract. Portfolio optimization in stock markets has been investigated by many researchers. It looks for a subset of assets able to maintain a good trade-off control between risk and return. Several algorithms have been proposed to portfolio management. These algorithms use known return and correlation data to build subset of recommended assets. Dynamic stock correlation networks, whose vertices represent stocks and edges represent the correlation between them, can also be used as input by these algorithms. This study proposes the definition of constants of the classical mean-variance analysis using machine learning and weighted link prediction in stock networks (method named as MLink). To assess the performance of MLink, experiments were performed using real data from the Brazilian Stock Exchange. In these experiments, MLink was compared with mean-variance analysis (MVA), a popular method to portfolio optimization. According to the experimental results, using weighted link prediction in stock networks as input considerably increases the performance in portfolio optimization task, resulting in a gross capital increase of 41% in 84 days.

Keywords: Stock Market · Dynamic Stock Networks · Machine Learning · Portfolio Optimization.

1 Introduction

Portfolio optimization in stock markets is the process of selecting a subset of assets that maintain an expected trade-off control between risk and return[16]. The portfolio selection process consists of finding, in a large collection of stocks, the participation (i.e. individual proportion) of each stock that minimizes the portfolios risk at a given portfolio return, or maximizes the portfolios return at a given risk [8]. This topic has been investigated by many researchers from many

ICCS Camera Ready Version 2019
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-22744-9_27

areas, such as optimization, machine learning (ML) and economics. Usually, portfolio selection algorithms use return and risk measures from a set of assets to make decision. This trade-off between risk and return is used to suggest a subset of assets that will be in the portfolio. Commonly used measures include price mean-return and price variance.

The correlation between asset prices is also important for portfolio management. It has been used in several works to create a network perspective that characterizes the complex structures in the stock market, also known as stock networks or financial networks [15] [18][5]. The correlation is so important that some authors suggest the use of topological information derived from financial networks for portfolio management [24] [32] [13]. As an example, we can create portfolios by applying clustering algorithms or centrality measures in stock networks. However, despite its importance, we did not find works exploring the prediction of weighted links in financial networks to improve the performance of portfolio optimization algorithms. However, we found works using price and return forecasting to improve portfolio management results [19].

This study proposes a new approach to define the constants of the classic mean-variance analysis (MVA) from [16]. The new measures for expected return and asset correlation are calculated using by a new method. The proposed method, ML Weighted Link Prediction Analysis (MLink), provides information to the mathematical model of Markowitz, which is often used as basis for portfolio optimization. MLink was applied to dynamic temporal stock networks to induce a predictive model for return price forecast and weighted link prediction. We executed several experiments to assess the performance of our method. When it was compared with MVA, experimental results shows that MLink has 56% over MVA. In addition, MLink has better results than other variant methods. These variants use ARIMA, Median and Mean to return forecast combined with Weighted Link Prediction, named as ARIMA-MLink, Median-MLink and Mean-MLink, respectively.

2 Problem Definition and Related Works

The analysis of the behavior and interaction between assets in stock markets has been widely studied in the literature [30] [3] [4]. To formally describe it, consider i and j two distinct assets belonging to a specific set A. Let S_i and S_j be time series related to i and j, respectively. A weighted stock network can be represented by a graph $G = (V, E)$, where V is the set of assets belonging to A, $V \subset A$, and E contains all possible pairs of assets $i, j \mid i, j \in V$. The relationship between i and j is measured using correlation metrics, assigning a weight w to each edge in E. A set of time ordered sequence of weighted stock networks graphs represents Dynamic Temporal Stock Networks [10].

In this work, we want to to answer the following question: given a set of graphs $G_1, G_2, ..., G_N$ related to a temporal sequence of weighted stock networks in time N and G_{N+1} a graph whose link weights were predicted using ML algorithms, can

ICCS Camera Ready Version 2019
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-22744-9_27

the combination of weighted link prediction by ML with mathematical models of portfolio optimization improve the trade-off between risk and return?

2.1 Portfolio Optimization using Stock Network and Forecast

Several works apply ML or optimization algorithms for portfolio management [16] [17]. Despite its relevance, few papers investigate the use of stock network structure to portfolio management. In [24], the authors explore centrality in complex networks to improve portfolio selection process via targeting a group of stocks belonging to certain region of the stock market network. The work [32], like this paper, analyze time evolving stock markets by using temporal network representation. The authors also propose a portfolio selection tool using temporal centrality in stock networks. A portfolio optimization based on network topology, using cross-correlation of the daily price returns for the American and Chinese stock markets to create networks, is proposed in [13]. Taking into account the importance of correlation matrices, and the possible presence of noise values in these matrices, [22] introduces an approach that allows a systematic investigation of the effect of the different sources of noise in financial correlations in the portfolio and risk management context. We also analyzed different aspects related to noise presence, like the size of time series.

Other works use price and return forecasting to improve portfolio management [19]. In [8], a neural network is used to predict future stock returns. The prediction errors are used as a risk measure. [1] also shows that, for asset allocation decisions, the use of models able to predict return is better than using historical averages. This motivated the proposal of the new method MLink that uses prediction of weighted link formation to improve results of portfolio optimization models. The problem of predicting links in weighted networks is an extension of the problem of link prediction, whose main objective is the detection of hidden links or links that will be formed [28]. In weighted link prediction problem, it is necessary to predict both link and edge weight [14].

3 Methodology

The MLink framework has two steps: (i) ML to predict stock returns and weighted links in dynamic stock networks; (ii) mathematical portfolio optimization model using stock returns and predicted weighted stock networks as input. Figure 1 illustrates this framework.

This section introduces the data set (Section 3.1), the weighted link prediction method (Section 3.2), return forecast (Section 3.3) and the mathematical model (Section 3.4).

3.1 Market Data Set

The real data set used in this work was collected from the Brazilian Stock Exchange (BM&F Bovespa) between January 2018 e October 2018. We performed

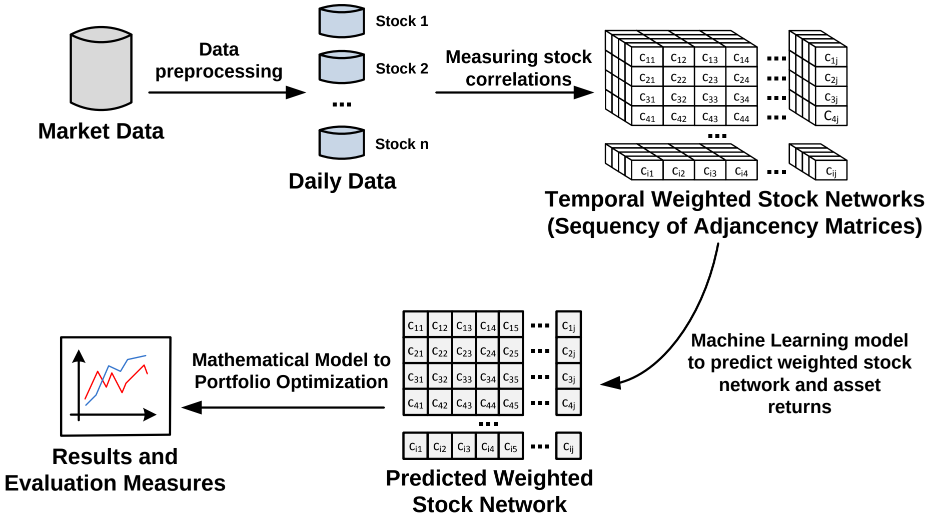

Market Data
Data preprocessing
Stock 1
Stock 2
...
Stock n
Daily Data
Measuring stock correlations
c11 c12 c13 c14 ... c1j
c21 c22 c23 c24 ... c2j
c31 c32 c33 c34 ... c3j
c41 c42 c43 c44 ... c4j
ci1 ci2 ci3 ci4 ... cij
Temporal Weighted Stock Networks (Sequency of Adjancency Matrices)
Machine Learning model to predict weighted stock network and asset returns
c11 c12 c13 c14 c15 ... c1j
c21 c22 c23 c24 c25 ... c2j
c31 c32 c33 c34 c35 ... c3j
c41 c42 c43 c44 c45 ... c4j
ci1 ci2 ci3 ci4 ci5 ... cij
Predicted Weighted Stock Network
Mathematical Model to Portfolio Optimization
Results and Evaluation Measures

Fig. 1. Framework of the MLink

experiments using data regarding Bovespa Index assets (*Ibovespa*) [4]. According to its dynamical structure, this theoretical portfolio has 65 assets currently, but we used 56 assets that remained in Ibovespa list during the entire period. The list of assets used in the experiments can be seen at link below [5]. In the first step, our objective is to identifying link weight between all pairs of assets in a complete weighted network. As traditional portfolio management algorithms, we are interested in reorganize the portfolio every day. Thus, the data were processed to obtain daily price time series. Note that our approach can be easily adapted for weekly, monthly or intraday strategies, according to the main purpose of the investor. It is important to emphasize that these data include an election period with high stock price variations.

3.2 Weighted Link Prediction in Dynamic Stock Networks

In this work, we use a modified version of the method proposed in [15] to create weighted dynamic financial networks. In this method, nodes of the graph represent assets and edges represent the relationship between them. This relationship is based on price time series correlation. Let S_i and S_j both price time series with length L regarding two distinct assets i and j. We can transform these non stationary price time series into a stationary return time series using the following equation:

$$Y_i = ln(P_i(t)) - ln(P_i(t-1))$$

[4] www.bmfbovespa.com.br

[5] iccs2019.douglascastilho.com

where, $P_i(t) \in S_i$ is the closing price of asset i at day t. In terms of market definitions, this transformation represents the logarithmic return price. The average of this type of time series tends to be close to zero. Next, we applied the Pearson [23] correlation coefficient between all possible pairs of stocks logarithmic return time series Y_i and Y_j:

$$\rho_{ij} = \frac{cov(Y_i, Y_j)}{\sqrt{var(Y_i) \cdot var(Y_j)}}$$

The number of points L to measure the correlation between all possible pairs of stocks is another aspect to be considered. We performed experiments using $L = \{10, 15, 20, 25, 30\}$, as suggested by [31]. The correlation between assets is assigned to edges weights. This information is used as input to mathematical portfolio optimization algorithm. We used correlation ρ_{ij} between all possible pairs of stocks present in the Bovespa Index to create an adjacency matrix $\mathbf{C}$. By definition, the elements ρ_{ij} are in the range of -1 to 1, where -1 corresponds to perfect anti-correlation, 1 corresponds to perfect correlation and 0 corresponds to absence of correlation. Consider that $\mathbf{C}$ represents a complete undirected stock network. Optimization models generally use similar correlation matrix as input to create a subset of assets with reasonable trade-off between return and risk. In this work, instead using the known correlation matrix at day t, we propose to use a step forward correlation matrix to improve the results of optimization portfolio. Our first main problem is to create machine learning (ML) algorithms able to induce models that can predict the weight of all edges in a future dynamic stock network Δ_{t+1}. For such,

we created a method using ML to predict these values. This method uses three different sources of features to build models able to predict correlation values, represented by edge weights. In addition, since our networks are undirected, the number of weighted edges for each graph $G = (V, E)$ is given by $|E| = |V| * (|V| - 1)/2$. Thus, we have $|E| = 56 * (55)/2 = 1540$ number of values to predict for each day.

In this work, we address the weighted link prediction problem as a regression task. For such, we propose the use of supervised ML algorithms. We use three sources of features to predict weighted link formation between assets in dynamic stock networks: *(i)* complex network derived features; *(ii)* domain derived features; *(iii)* return time series forecasts. Each example in the data set is labeled with correlation value between a pair of stocks i and j for a future given period. Next, we present the set of input features used to train the ML algorithms.

Network Derived Features are computed at each iteration using complex network statistical measures. The measures can be divided according to the level of analysis to be performed: at the node-level, where nodes represent assets, and at the link-level [21]. To create each example for the supervised learning data set, metrics related with node i and j are inserted for both nodes. Metrics related with edges are inserted calculating the measures between nodes i and j[20]. Consider $|i|$ as node degree or number of edges.

- **Node-Level Derived Features**: related to the position of the node within the overall structure of the complex network [21]. The following node metrics presented in Table 1 were used as stock i derived features.
- **Link-Level Derived Features**: related to both contents and patterns of edges in complex networks [21]. The following link metrics presented in Table 2 were calculated between i and j stocks.

Table 1. Node-Level Derived Features

Name	Description
Weighted degree	$deg_w(i) = \sum_{j \in \|i\|} w_{<i,j>}$
Average neighbor weighted degree	$avg_w(i) = \frac{\sum_{j \in \|i\|} \|j\| * w_{<i,j>}}{\|i\|}$
Propensity of i to increase its degree	$\gamma(i) = \frac{\|i\|}{deg_w(i)}$
Node betweenness	$b_v = \sum_{i,j \in V \setminus v} \frac{\sigma_{ij}(v)}{\sigma_{ij}}$, where $\sigma_{ij}(v)$ is the number of shortest weighted paths between i and j passing through v
Node eigenvector	$x_i \frac{1}{\lambda} \sum_{j=1}^{n} d_{ij} x_j$, where d_{ij} represents an entry of the adjacency matrix C and λ denotes the largest eigenvalue

Table 2. Link-Level Derived Features

Name	Description
Correlation value	value of C_ij
Edge betweenness	$b_e = \sum_{i,j \in V} \frac{\sigma_{ij}(e)}{\sigma_{ij}}$, where $\sigma_{ij}(e)$ is the number of shortest weighted paths between i and j passing through edge e
Same Louvain community [2]	value 1 if i and j belongs to the same Louvain community, 0 otherwise
Same Girvan-Newman community [9]	value 1 if i and j belongs to the same Girvan-Newman community, 0 otherwise
Preferential attachment weighted	$PA(i,j) = deg_w(i) * deg_w(j)$

Domain Features are computed at each day using a set of Technical Analysis Indicators (TAI). An indicator can be defined as a series of data points derived from assets price information applying a mathematical formula [26]. These metrics were calculated using the same set of daily price time series used to create the networks. The relationship between these domain features and the graph networks can be seen as characteristics to describe nodes in a complex network. The domain features used are: Relative Strength Index (RSI), Simple Moving Average (SMA), Exponential Moving Average (EMA), Moving Average Convergence/Divergence (MACD), Average Directional Movement Index (ADX), Aroon Indicator (Aroon), Bollinger Bands (BB), Commodity Channel Index (CCI), Chande Momentum Oscillator (CMO), Rate of Change (ROC) and Average True Range (ATR). More information regarding how to calculate these features can be found in [27] and [26].

To create each example for the ML algorithm training data set, the subset of network derived features is concatenated to the subset of domain derived features. Thus, TAI related to assets i and j are inserted for both nodes.

3.3 Logarithmic Return Forecast Using Machine Learning

Another issue to be considered in our MLink framework is the stock logarithmic return. According to the mathematical optimization model, which will be presented in the following section, it is necessary a measure of return and risk of each asset. Usually, the mean of the returns and the standard deviation are used in optimization algorithms to estimating return and future risk, respectively. For this, we used a ML algorithm (MLP) to forecast the logarithmic return of all assets. For comparative analysis, we also used three statistical methods to estimate assets return. These methods are:

- MLP - Multilayer Perceptron neural network - has powerful approximation capabilities and its self-adaptive data driven modelling approach allow them great flexibility in modelling time series data [12]. The MLP network used has one hidden layer with 5 neurons and is trained using the resilient backpropagation algorithm, a fast weighted update mechanism to feedforward artificial neural networks [25].
- ARIMA - Autoregressive Integrated Moving Average - models are fitted to the time series data to predict future points in the series [6].
- Mean - Simple mean price return using time series data.
- Median - Simple median price return using time series data.

Based on each method result, we then calculated the risk measure, which is given by standard deviation of logarithmic return time series. These return forecasts are used as input to mathematical portfolio optimization model and also as input feature to Mlink to predict weighted link in the stock network. To create each example for the supervised learning data set regarding weighted link prediction in MLink, these forecasted return values of each pair of assets i and j are concatenated to input set of features.

3.4 Mathematical Model to Portfolio Optimization

In [16] Markowitz proposed the first mean-variance model that served as the basis for the Modern Portfolio Theory in financial management. In this theory, an investor wishes to distribute an initial wealth in a set of investments in order to minimize the risk and maximize the return. Naturally, these two objectives are conflicting because if there is a minimum risk investment and maximum return the decision is trivial. Usually the higher the risk the higher the expected return.

For this, [16] proposed a bi-objective quadratic programming model for find the Markowitz efficient front. In mathematical terms: given n assets with return vector $\mu \in \mathbb{R}^n$, estimated covariance matrix $\sigma \in \mathbb{R}^{n \times n}$, and the invested fraction of each asset in optimal portfolio is $x \in \mathbb{R}^n$. To computing the Markowitz efficient front is given by maximizing expected return for a given level of the risk (mean-variance model 1) or minimizing the risk for a given level of the expected return (mean-variance model 2).

$$\begin{aligned} \text{Maximize} \quad & \mathcal{E} = \sum_{i=1}^{n} x_i \mu_i \\ \text{Subject to:} \quad & \sum_{i=1}^{n} \sum_{j=1}^{n} x_i x_j \sigma_{ij} \leq v^2, \\ & \sum_{i=1}^{n} x_i = 1, \\ & x_i \geq 0, \ \forall\, i = 1, ..., n. \end{aligned} \tag{1}$$

In mean-variance model 1 the expected value of the portfolio ($\mathcal{E}$) is maximized, subject to a minimum variation (v^2), the sum from fraction of the portfolio is equal 1, and no investment can be negative.

$$\begin{aligned} \text{Minimize} \quad & v^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i x_j \sigma_{ij} \\ \text{Subject to:} \quad & \sum_{i=1}^{n} x_i \mu_i \geq \mathcal{E}, \\ & \sum_{i=1}^{n} x_i = 1, \\ & x_i \geq 0, \ \forall\, i = 1, ..., n. \end{aligned} \tag{2}$$

In mean-variance model 2 the objective is minimize the standard deviation (v^2) subject to a level of return ($\mathcal{E}$). Note that, if we vary the of standard deviation (v^2) in mean-variance model 1 or the desired level return ($\mathcal{E}$) in mean-variance model 2 we can build the Markowitz efficient front, i.e. the trade-off between risk and expected return.

[11] examines an alternative formulation for the problem using a measures of absolute and relative risk aversion. Consider u is a von Neumann-Morgenstern utility function. The absolute risk aversion is defined by $R_a = \frac{u''(w)}{u'(w)}$, where w is the valuation of the portfolio. The formulation result is presented below:

$$\begin{aligned} \text{Maximize} \quad & F = \sum_{i=1}^{n} x_i \mu_i - \frac{R_a}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} x_i x_j \sigma_{ij} \\ \text{Subject to:} \quad & \sum_{i=1}^{n} x_i = 1 \\ & x_i \geq 0, \ \forall \ i = 1, ..., n. \end{aligned} \tag{3}$$

For the model 3 the objective is maximized the expected return of portfolio less $\frac{R_a}{2}$ time standard deviation of portfolio. According to [11] the utility functions (R_a) of negative exponential energy generate very risk-averse portfolios. Thus, the efficient boundary can be obtained for values of $R_a > 0$. The empirical results indicate that: risky portfolios have values of $R_a \leq 2$; moderate risk portfolios have $2 \leq R_a \leq 4$; risk-averse portfolio have $R_a \geq 4$.

4 Experiments

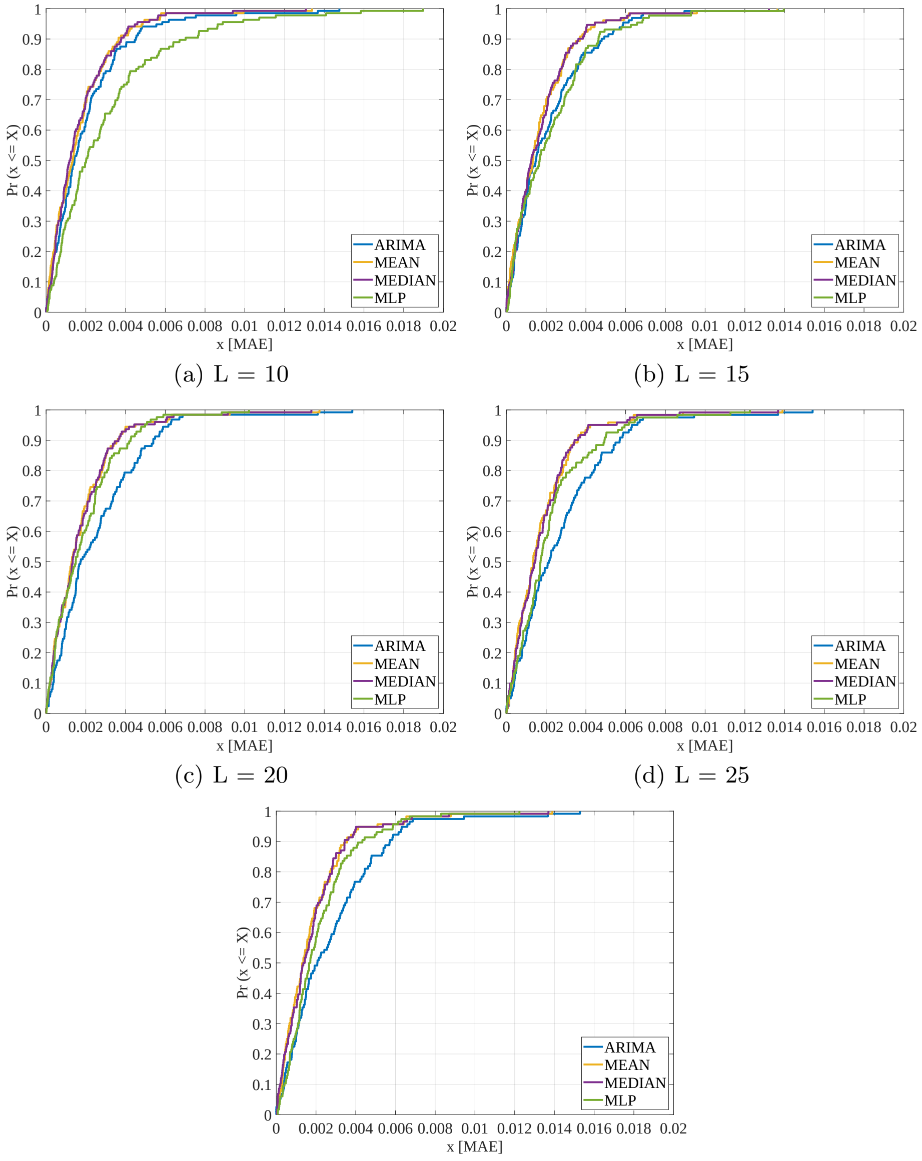
In this section we present results separately. First, we present results regarding to price return forecast using different lengths of L. Second, we present weighted link prediction results related to stock networks. We used the same time series length L to perform both return forecast and weighted link prediction experiments. Finally, we present financial results comparing MLink with Ibovespa, MVA and the three proposed variants ARIMA-MLink, Median-MLink and Mean-MLink.

4.1 Machine Learning for Return Forecast

In this section we present a set of experimental results using different methods to asset return forecast. We used different time series lengths $L = 10, 15, 20, 25, 30$ as input to time series forecast algorithms. Figure 2 shows a comparative result of Mean Absolute Error (MAE) evaluation metric [29]. For each length L, we plotted the Cumulative Distribution Function (CDF) related to each return forecast model. Thus, we can see the behavior of each model when different time series lengths are used.

An important result presented in Figure 2 is that Mean and Median have great MAE values compared with ML model (MLP). ARIMA has better results in small L values, but in $L = 30$ present the worst result. Besides that, MLP has good results for large values of L.

ICCS Camera Ready Version 2019
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-22744-9_27



(a) L = 10

(b) L = 15

(c) L = 20

(d) L = 25

(e) L = 30

Fig. 2. Return forecast results for each size of L

ICCS Camera Ready Version 2019
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-22744-9_27

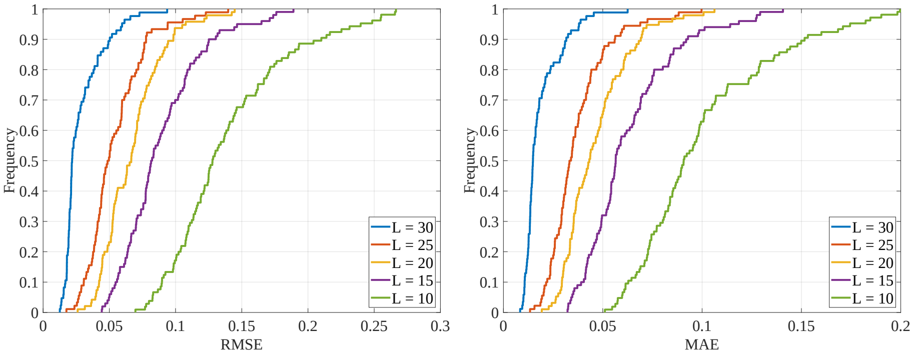
4.2 Machine Learning for Weighted Link Prediction

This section presents results related to weighted link prediction in dynamic stock networks. To perform these experiments, the training and test set was built using sliding window [26]. For such, we used 30 daily graph snapshots in the training set. The test set corresponds to the next trading daily data. The sliding window moves one day ahead to create new training and test set.

We used XGboost [7] as main ML model to weighted link prediction. It is a fast, a highly effective and widely used machine learning method. We did not perform an exhaustive search for model parameters because this is not our main objective. Our intention is to show how predictive a machine learning model can be using the set of features that we proposed. The set of model parameters are:

- booster = ”gbtree”;
- objective = ”reg:linear”,
- eta = 0.05,
- max_depth = 2,
- min_child_weight = 100,

Figure 3 presents comparative results using CDF for both MAE and Root Mean Square Error (RMSE) evaluation metrics.



(a) Root Mean Square Error

(b) Mean Absolute Error

Fig. 3. Weight link prediction using different sizes of L

Figure 3 presents significant weighted link prediction results using the proposed method. This comparison between the behaviors of the machine learning model for the different values of L allows us to visualize that, for greater L values, the model can better predict the values of the edges weights (correlations). A possible explanation is that: the greater the size of L, more stable the financial network tends to be, facilitating the prediction of edges weights. With smaller values of L, the network tends to be more unstable. Note that the value of L influences both return forecast and weighted link prediction. Considering that edge weights are between −1 and 1, MAE results for $L = 30$ are almost 95% under 0.05, which is very expressive in terms of weighted link prediction. At each day, the model predicts 1540 edge weights (correlations).

4.3 Portfolio Optimization Experiments

This section presents a comparative experiment using $L = 30$, which is the best return forecast and weighted link prediction results. We executed the Markowitz model using results of weighted link prediction and all methods applied to return forecast. Figure 4 shows a financial simulated gross return for each approach. Each execution uses 84 trading days. For a utility functions ($R_a = 2$) a threshold value between a moderate risk and risky.

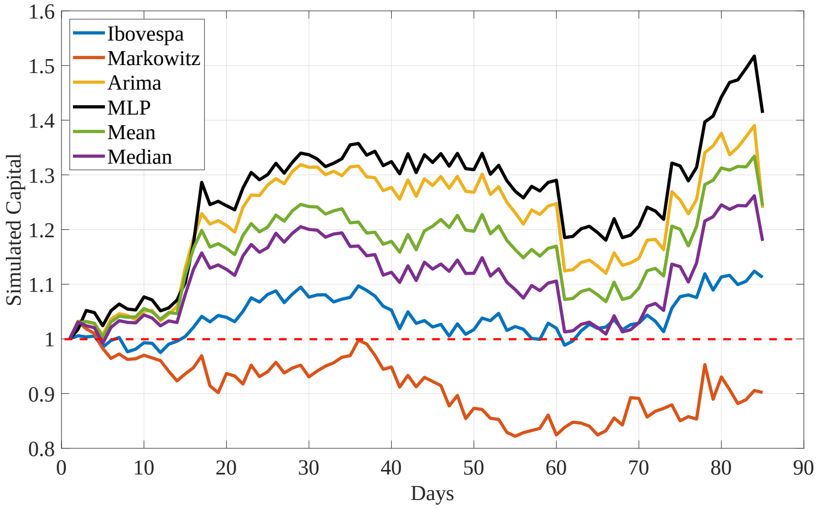


Fig. 4. Comparing accumulated return for each approach

Figure 4 shows that our MLink method outperforms MVA in over 56% comparing accumulated gross return. ARIMA-MLink, Median-MLink and Mean-MLink also has better results than MVA and Bovespa Index. This is an impressive result in terms of financial return.

5 Conclusions and Future Works

This study proposed determining the constants of the MVA from [16] using machine learning and a new weighted link prediction in stock networks defined in this paper as MLink. Portfolio optimization models use data from the past to create portfolios with a good trade-off between return and risk. Using the correlation between asset price series in the stock networks, when vertices represent

ICCS Camera Ready Version 2019
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-22744-9_27

assets and edges represent the correlation between them, the proposed method can predict the weights of all edges in dynamic stock networks. Like the MVA, the proposed method also use return forecasts. The experimental results show that using both return forecast and weighted link prediction data, the proposed method performance is superior to the performance obtained by the MVA. The experiments show that the MLink capital increases almost 41.34% in 84 days, a difference of 56.68% for the MVA (which reduced the capital in 9.79%).

These findings open a range of possibilities for future works. Several analyzes will be carried out, such as the use of different optimization models and run experiments using known market data sets for portfolio optimization. In addition, we can use other ways of predicting edge weights, such as tensor or deep learning, trying to improve weighted link prediction results. In addition, we can make available the created data set for other researchers evaluate their optimization models.

Acknowledgments

The authors would like thank CAPES, FAPESP (2013/07375-0), Intel - FAFQ, National Council for Scientific and Technological Development - CNPq - Brazil (202006/2018-2) and Federal Institute of Education, Science and Technology of South of Minas Gerais - IFSULDEMINAS - for their support.

References

1. Bessler, W., Wolff, D.: Portfolio optimization with return prediction models evidence for industry portfolios. In: World Finance and Banking Symposium (2015)
2. Blondel, V.D., Guillaume, J.L., Lambiotte, R., Lefebvre, E.: Fast unfolding of communities in large networks. Journal of statistical mechanics: theory and experiment **2008**(10), P10008 (2008)
3. Bonanno, G., Caldarelli, G., Lillo, F., Mantegna, R.N.: Topology of correlation-based minimal spanning trees in real and model markets. Physical Review E **68**(4), 046130 (2003)
4. Bonanno, G., Caldarelli, G., Lillo, F., Micciche, S., Vandewalle, N., Mantegna, R.N.: Networks of equities in financial markets. The European Physical Journal B-Condensed Matter and Complex Systems **38**(2), 363–371 (2004)
5. Bonanno, G., Lillo, F., Mantegna, R.N.: High-frequency cross-correlation in a set of stocks. Quantitative Finance **1**(1), 96–104 (2001)
6. Brockwell, P.J., Davis, R.A., Calder, M.V.: Introduction to time series and forecasting, vol. 2. Springer (2002)
7. Chen, T., Guestrin, C.: Xgboost: A scalable tree boosting system. In: Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining. pp. 785–794. ACM (2016)
8. Freitas, F.D., De Souza, A.F., de Almeida, A.R.: Prediction-based portfolio optimization model using neural networks. Neurocomputing **72**(10-12) (2009)
9. Girvan, M., Newman, M.E.: Community structure in social and biological networks. Proceedings of the national academy of sciences **99**(12), 7821–7826 (2002)
10. Holme, P., Saramäki, J.: Temporal networks. Physics reports **519**(3), 97–125 (2012)

DOI: 10.1007/978-3-030-22744-9_27

11. Kallberg, J.G., Ziemba, W.T.: Comparison of alternate utility functions in portfolio selection problems. Management Science **29**, 1257–1276 (1983)
12. Kourentzes, N., Barrow, D., Crone, S.: Neural network ensemble operators for time series forecasting. Expert Systems with Applications **41**(9), 4235–4244 (7 2014)
13. Li, Y., Jiang, X.F., Tian, Y., Li, S.P., Zheng, B.: Portfolio optimization based on network topology. Physica A: Statistical Mechanics and its Applications **515**, 671–681 (2019)
14. Lü, L., Zhou, T.: Link prediction in weighted networks: The role of weak ties. EPL (Europhysics Letters) **89**(1), 18001 (2010)
15. Mantegna, R.N.: Hierarchical structure in financial markets. The European Physical Journal B-Condensed Matter and Complex Systems **11**(1), 193–197 (1999)
16. Markowitz, H.: Portfolio selection. The Journal of Finance **7**(1), 77–91 (1952)
17. Markowitz, H.: Portfolio selection: Efficient diversfication of investments. New York: Wiley **7** (1959)
18. Marti, G., Nielsen, F., Bińkowski, M., Donnat, P.: A review of two decades of correlations, hierarchies, networks and clustering in financial markets. arXiv preprint arXiv:1703.00485 (2017)
19. Mishra, S.K., Panda, G., Majhi, B.: Prediction based mean-variance model for constrained portfolio assets selection using multiobjective evolutionary algorithms. Swarm and Evolutionary Computation **28**, 117–130 (2016)
20. NARASIMHAN, S.: Link Prediction in Dynamic Networks. Ph.D. thesis (2015)
21. Oliveira, M., Gama, J.: An overview of social network analysis. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery **2**(2), 99–115 (2012)
22. Pafka, S., Kondor, I.: Estimated correlation matrices and portfolio optimization. Physica A: Statistical Mechanics and Its Applications **343**, 623–634 (2004)
23. Pearson, K.: Note on regression and inheritance in the case of two parents. Proceedings of the Royal Society of London **58**, 240–242 (1895)
24. Peralta, G., Zareei, A.: A network approach to portfolio selection. Journal of Empirical Finance **38**, 157–180 (2016)
25. Riedmiller, M., Braun, H.: A direct adaptive method for faster backpropagation learning: The rprop algorithm. In: Neural Networks, 1993., IEEE International Conference on. pp. 586–591. IEEE (1993)
26. Silva, E., Brandao, H., Castilho, D., Pereira, A.C.: A binary ensemble classifier for high-frequency trading. In: Neural Networks (IJCNN), 2015 International Joint Conference on. IEEE (2015)
27. Silva, E., Castilho, D., Pereira, A., Brandao, H.: A neural network based approach to support the market making strategies in high-frequency trading. In: Neural Networks (IJCNN), 2014 International Joint Conference on. IEEE (2014)
28. Wang, P., Xu, B., Wu, Y., Zhou, X.: Link prediction in social networks: the state-of-the-art. Science China Information Sciences **58**(1), 1–38 (2015)
29. Willmott, C.J., Matsuura, K.: Advantages of the mean absolute error (mae) over the root mean square error (rmse) in assessing average model performance. Climate research **30**(1), 79–82 (2005)
30. Yang, C., Chen, Y., Niu, L., Li, Q.: Cointegration analysis and influence ranka network approach to global stock markets. Physica A: Statistical Mechanics and its Applications **400**, 168–185 (2014)
31. Yang, Y., Yang, H.: Complex network-based time series analysis. Physica A: Statistical Mechanics and its Applications **387**(5-6), 1381–1386 (2008)
32. Zhao, L., Wang, G.j., Wang, M., Bao, W., Li, W., Stanley, H.E.: Stock market as temporal network. Physica A **506**, 1104–1112 (2018)