

# An Investigation into the Recurring Patterns of Forex Time Series Data

Yoke Leng Yong

Department of Computing & Information Systems  
Sunway University  
Selangor, Malaysia  
14071856@imail.sunway.edu.my

Yunli Lee, David Ngo

Department of Computing & Information Systems  
Sunway University  
Selangor, Malaysia  
{yunli, dngo}@sunway.edu.my

**Abstract**—Countless theories have been developed by both researchers and financial analyst in an attempt to explain the fluctuation of forex price. By obtaining an intimate understanding of the forex market, traders will hopefully be able to forecast and react to forex price oscillations on-the-fly towards making a profitable investment. In this paper, an investigation into the underlying theory that there exists repeating patterns within the time series data which forms the basis of technical analysis is conducted. The assumption that certain patterns do develop over time and the forex market does not fluctuate in a random manner is used to establish the fact that history repeats itself in forex trading. The patterns and repetitions unveiled within the forex historical data would be an important element for forex forecasting.

**Keywords**—forex forecasting; SAX; time series data

## I. INTRODUCTION

Ever since the inception of the Gold Standard Exchange in 1876, the evolution of the currency market [1] alongside technology advancement has propelled Forex into one of the biggest financial market surpassing even the stock exchange. This is evident as proven by the massive amount of daily trading which amounts to 5.3 million as recorded by the Triennial Central Bank Survey [2]. Coupled with the advancement on technology over the years, it has only served in providing the influx of Forex traders with ease of access via the introduction of automated trading platforms. The trading platform not only serve as means for traders to monitor and trade Forex data in real time, it has also evolved into something much more sophisticated with the incorporation of additional tools such as basic trend analysis and expert advisors.

Trend analysis and expert system is where researchers contribute by introducing novel methods in the analysis and forecasting of the Forex market. Forex market prediction is particularly complicated as there are many external factors and assumptions affecting the fluctuation of the Forex prices as well as the size of the market. Not to mention that traders often rely on heuristics and experience in data analysis. While traders incorporate all sorts of statistical based methods and trading rules in forecasting, researchers approaching it from the computational finance aspect will be tempted to incorporate elements of Artificial Intelligence (AI), especially soft computing methods such as: Artificial Neural Network (ANN), Fuzzy Systems, Evolutionary Computing and Swarm

Intelligence which is generally more tolerant of imprecision and uncertainty.

Herein this paper, not one but two of the most important question/assumption concerning the forex time series data will be investigated, namely: 1) history repeats itself in the Forex market and 2) there are unknown underlying patterns that develop over time that are useful in forex forecasting. It is well known that traditionalist believe in the random nature of the forex market which is tough to forecast. However, recent years have seen the emergence of a new breed of traders who believes in the existence of trends and patterns in forex trading. In this paper, the repetitions and formation of patterns either known/unknown within the forex time series data are investigated.

## II. FOREX FORECASTING

### A. Overview

Taking into consideration that the current Forex market forecasting research encompasses a variety of methods from different academic disciplines and its evolution through the years, a brief overview of forex data analysis and forecasting methods used over the years as depicted in Fig 1 is essential. Historically, proponents of the Random Walk Hypothesis [3] and Efficient Market Hypothesis [4] promoted the idea that the Forex market fluctuations are unpredictable. However, traders who believe in the three remaining analysis methods namely: fundamental analysis, sentiment analysis and technical analysis have their own interpretation of what actually affects the Forex market. In truth, it is extremely difficult to gauge the effects of each individual element within the analysis methods studied when even the source of information itself differs greatly as news delivery platform evolve along with the ability of the traders to keep up with the latest updates.

The forex historical data proves to be a crucial element in analysis methods widely utilised by technical analyst as well as researchers in the area of computational finance. Although technical analysis has traditionally been used by traders, researchers have begun adopting the use of historical data for various analysis using soft/hard computing methods introduced in Artificial Intelligence [5, 6]. Forex historical data required obtained from sources such as OANDA.com, HistData.com, and TradeStation [7-10] have been integrated into the innumerable application of technical analysis and AI methods.

Subsection B below provides more in depth discussion on the technical analysis methods as well as methods currently used for forex data analysis and forecasting.

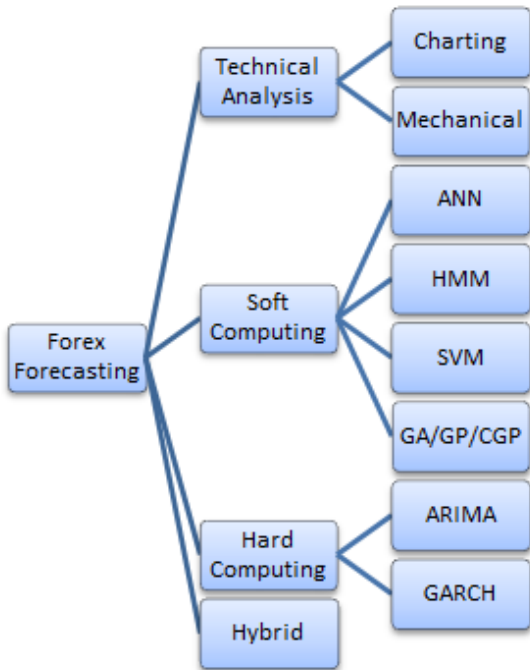


Fig. 1. Forex forecasting methods

### B. Technical Analysis

Technical analysis can be seen to encompass two main methods: 1) mechanical and 2) charting based analysis. Charting relies upon predefined chart pattern such as available in [11] and used in research for forex forecasting where identification of chart patterns are performed for forecasting purposes [12, 13]. Recorded in the aptly named Encyclopedia of Chart Patterns, a total of 53 chart patterns have been identified [11] that are useful in market analysis. However, charting is not without its shortcoming as it is extremely subjective and depends heavily on the interpretation of the trader performing the analysis to identify the pattern. Moreover, it does not offer flexibility in dealing with unknown and hidden patterns within the time series data.

The incorporation of AI methods have vastly expanded the options available in the analysis and forecasting of Forex data. While the development of soft computing have seen the introduction of machine learning algorithm such as SVM, HMM, ANN, Genetic Algorithm for analysis and forecasting, hard computing focuses on rigid statistical based methods to forecast the data. Methods such as ARIMA and GARCH are among some of the techniques used. Hybrid methods combining both soft and hard computing methods have also been introduced as discussed by Abraham [5].

### C. SAX Time Series Analysis

SAX algorithm has been widely used for time series analysis [14]. The SAX representation proves to be an excellent candidate as feature representation for forex data analysis as it is a simple yet powerful feature representation for such a huge set of data. The simple symbolic representation used initially for time series analysis has been adopted for stock investment strategies optimization [15-18]; focusing primarily on the analysis of stock market data. The processing algorithm involves the use of SaX features along with genetic algorithm optimization kernel which determines the buy/sell decisions.

In order to better understand the feature representation offered by SAX, it is important to look at the algorithm implementation in detail. As outlined by previous researches [14-18], the three major components of SAX algorithm involves: 1) Normalisation of the time series data, 2) Piecewise Aggregate Approximation (PAA) computation and 3) converting the PAA to a symbolic representation [14]. Performing the z-score normalization on the time series data ( $z_i$ ) by subtracting each of the time series element ( $c_i$ ) by the mean value ( $\mu$ ) and dividing it by the standard deviation ( $\sigma$ ) as denoted in (1) is important to obtain meaningful comparison of the time series data as the data might have different offsets and amplitudes [14].

$$z_i = (c_i - \mu) / \sigma \quad (1)$$

Subsequent to normalization, the PAA could then be applied to the data as proposed by Lin et al [14]. Finally, the SAX string representation could be obtained once the discretization of the time series is performed whereby breakpoints are allocated as denoted in Fig 2 to produce symbols with equiprobability [14].

$\beta_i \backslash a$	3	4	5	6	7	8	9	10
$\beta_1$	-0.43	-0.67	-0.84	-0.97	-1.07	-1.15	-1.22	-1.28
$\beta_2$	0.43	0	-0.25	-0.43	-0.57	-0.67	-0.76	-0.84
$\beta_3$		0.67	0.25	0	-0.18	-0.32	-0.43	-0.52
$\beta_4$			0.84	0.43	0.18	0	-0.14	-0.25
$\beta_5$				0.97	0.57	0.32	0.14	0
$\beta_6$					1.07	0.67	0.43	0.25
$\beta_7$						1.15	0.76	0.52
$\beta_8$							1.22	0.84
$\beta_9$								1.28

Fig. 2. SAX breakpoint lookup table from [14]

### III. PROPOSED SYSTEM

There is no doubt that extensive work has been done on the identification of forex time series patterns. However, the proposed method ventures into uncharted territory with a novel proposition of dealing with unknown patterns. The proposed implementation involves the use of SAX representation as features in the identification and matching of unknown patterns within the forex time series. For the proposed system

developed, a detailed discussion of the dataset used is available in subsection A below while the algorithm implementation is further deliberated in subsection B.

### A. Dataset

From a wide selection of forex historical data accessible from HistData website [19], the EUR/USD exchange rate have been chosen as it is recognised as one of the five major currencies in the forex market. The data downloaded is subsequently pre-processed and the close price captured used for computation. The illustration of the entire dataset is as depicted in Fig 3 and detailed breakdown of the dataset used are available in Table I.

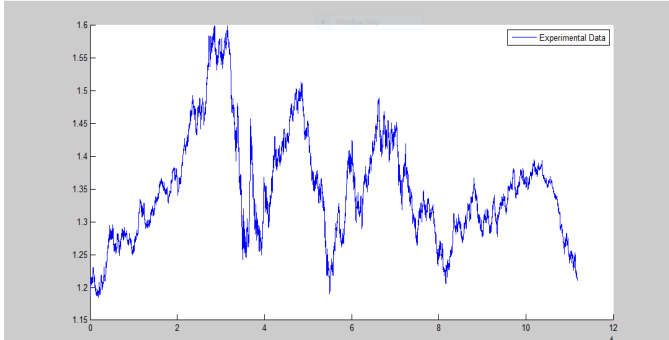


Fig. 3. Forex forecasting methods

TABLE I. DATASET INFORMATION

Experimental Data	Details
Currency dataset	EUR/USD
Dataset information	Date/Time, Open, High, Low and Close Price, Volume
Total Time Frame	2 Jan 2006 – 31 Dec 2014
Pre-processed dataset	30 minutes interval
Time series segmentation	Bi-weekly segmentation

### B. Algorithm Implementation

The proposed algorithm involves the use of SAX symbolic representation in detecting unknown patterns within the forex time series and seeking its repetitions, if it exists. The algorithm consists of the following major steps: pre-processing, SAX symbolic representation and lastly, finding the minimum distance between the SAX feature representation of the time series obtained. The pre-processing steps involve segmenting the time series data into predefined time segment to acquire the SAX representation whereby for the purpose of this study is set on a bi-weekly basis. Once the SAX symbolic representations of all the segments are obtained, the distances comparing each of the strings are computed. From the distance measures obtained, the repeating patterns within the forex time series are easily identified. A detailed representation of the algorithm implementation is as depicted in the pseudocode as described in Fig 4.

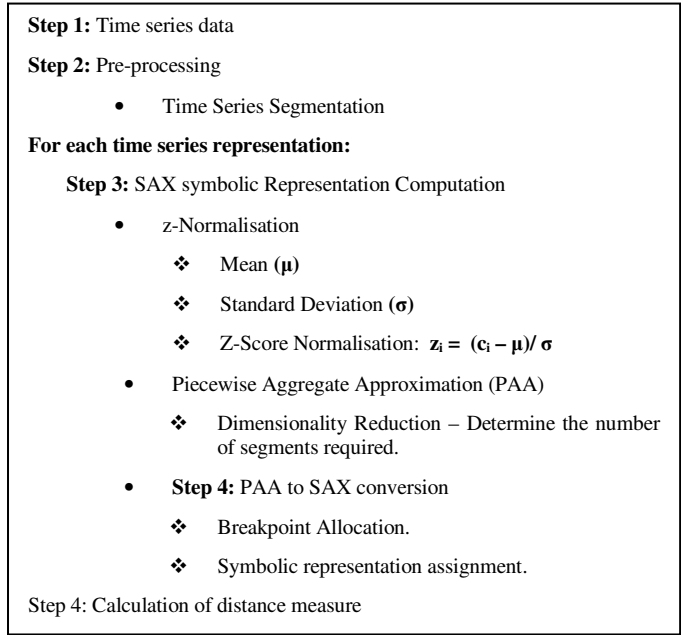


Fig. 4. Proposed algorithm

## IV. RESULTS

From the algorithm as outlined in Section III, 238 SAX strings are extracted and the tabulation of distance measures for all the SAX strings are as seen in Fig 5. As the distance measure does not provide enough information, the visualisation of the close price for matching strings with the closest distance confirms the similarities of patterns extracted such as shown in Fig 6. Within the graphs, the y-axis represents the price fluctuation of the EUR/USD exchange with the scale set at a hundredth of the exchange unit (0.01). It can clearly be seen that that the forex close price which produces short distance measure depicted by Fig 6 have similar patterns and trends while those bigger distance measure have either dissimilar patterns or huge price gaps exist between both the patterns such as seen in Fig 7.

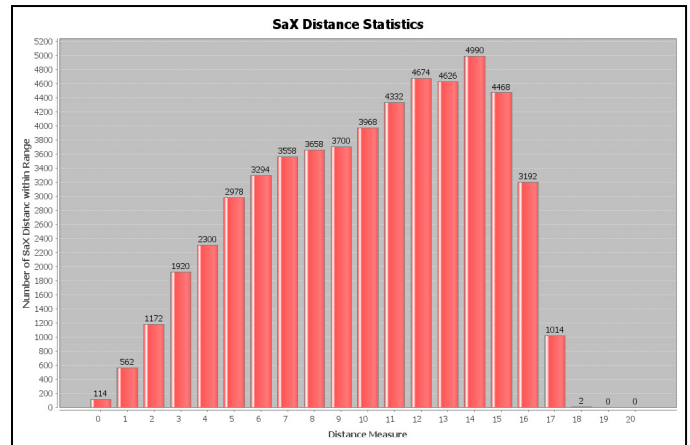


Fig. 5. Distance measure statistics

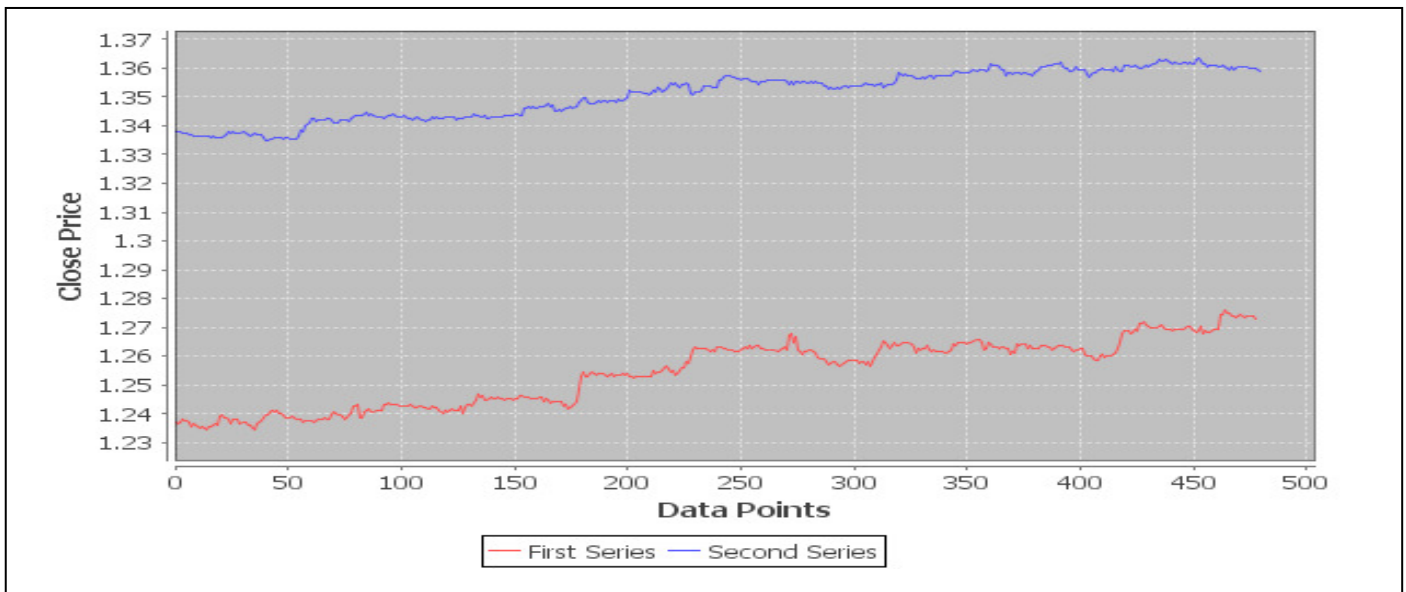


Fig. 6. Data with distance measure 0

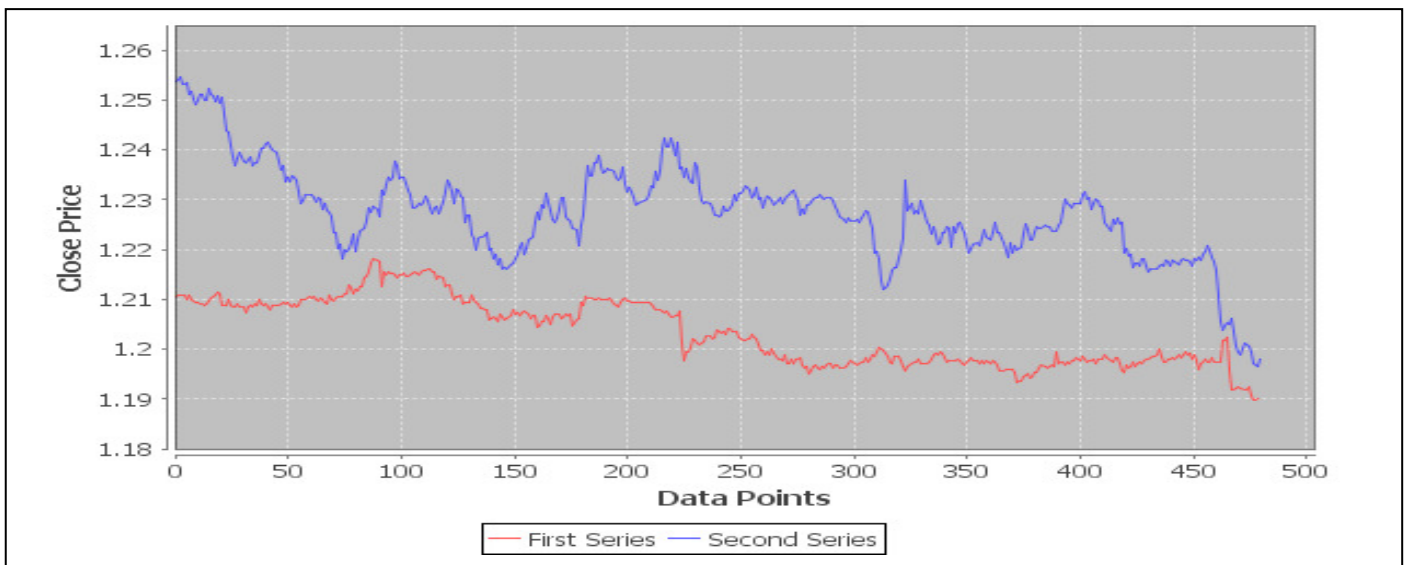


Fig. 7. Data with distance measure 6

## V. CONCLUSION

In conclusion, the proposed method is able to extract and identify unknown patterns which reoccur periodically within the forex time series data. The finding is extremely pertinent in establishing the fact that history does in fact repeat itself in the forex market albeit with slight difference, taking into account the time warping element of the time series data and the repetition occurs in a random manner. This contributes positively to the forecasting of forex price with the ability to handle unknown patterns as the ability to identify repeating patterns are well known in contributing to the forecasting of forex price.

## REFERENCES

- [1] C. Gallo, "The Forex Market in Practice : A Computing Approach for Automated Trading Strategies," *International Journal of Economics and Management Sciences*, vol 3, no.169, pp. 1-9, 2014.
- [2] Bank for International Settlement, "Triennial Central Bank Survey, Foreign Exchange turnover in April 2013: Preliminary Global Results.," 2013.
- [3] B.G. Malkiel, *A Random Walk Down Wall Street*, Norton, 1973.
- [4] E.F. Fama, "Random Walks in Stock Market Prices," *Financial Analysts Journal*, vol. 21, no. 5, pp. 75 – 80, 1965.
- [5] A. Abraham, "Analysis of Hybrid Soft and Hard Computing Techniques for Forex Monitoring Systems," In *Proceedings of the 2002 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE'02)*, vol. 2, pp. 1616 – 1622, 2002.
- [6] A. Mochón, D. Quintana, Y. Sáez, and P. Isasi, "Soft Computing Techniques Applied to Finance," *Applied Intelligence*, vol. 29, no. 2, pp. 111-115, 2008.

- [7] N.G. Pavlidis, D.K. Tasoulis, and M.N. Vrahatis. "Financial forecasting through unsupervised clustering and evolutionary trained neural networks," In *The 2003 Congress on Evolutionary Computation, CEC'03*, vol. 4, 2003.
- [8] R.F.B. De Brito and A.L.I. Oliveira, "A Foreign Exchange Market Trading System by Combining GHSOM and SVR," In *Proceedings of the International Joint Conference on Neural Networks (IJCNN)*, pp. 1 – 7, 2012.
- [9] Y.L. Lee, L.C.O. Tiong, and D.C.L. Ngo. "Hidden Markov Models for Forex Trends Prediction," In *2014 International Conference on Information Science and Applications (ICISA)*, 2014.
- [10] A. Paukste and A. Raudys. "Intraday forex bid/ask spread patterns-Analysis and forecasting," *2013 IEEE Conference on Computational Intelligence for Financial Engineering & Economics (CIFER)*, 2013.
- [11] N.B. Thomas, (Encyclopedia of chart patterns, *Hoboken, NJ: John Wiley*, 2005
- [12] A. Bagheri, H.M. Peyhani, and M. Akbari. "Financial forecasting using ANFIS networks with Quantum-behaved Particle Swarm Optimization," *Expert Systems with Applications* vol. 41, no. 14, pp. 6235-6250, 2014.
- [13] J.N. Liu and R.W. Kwong, "Automatic extraction and identification of chart patterns patterns towards financial forecast," *Applied Soft Computing*, vol. 7, pp. 1197–1208, 2007.
- [14] J. Lin, E. Keogh, S. Lonardi, and B. Chiu, "A symbolic representation of time series, with implications for streaming algorithms," In *Proceedings of the 8th ACM SIGMOD workshop on Research issues in data mining and knowledge discovery*, 2003.
- [15] A.M. Canelas, R.F. Neves, and N.C. Horta, "A New SAX-GA Methodology Applied to Investment Strategies Optimization," In *Proceedings of the Fourteenth International Conference on Genetic and Evolutionary Computation Conference - GECCO '12*, pp. 1055 – 1062, 2012.
- [16] A.M. Canelas, R.F. Neves, and N.C. Horta, *Investment Strategies Optimization based on a SAX-GA Methodology*, Springer Science & Business Media, 2013.
- [17] A.M. Canelas, R.F. Neves, and N.C. Horta, "A SAX-GA Approach to Evolve Investment Strategies on Financial Markets based on Pattern Discovery Techniques," *Expert Systems with Applications*, vol. 40, no. 5, pp. 1579–1590, 2013.
- [18] A.M. Canelas, R.F. Neves, and N.C. Horta, "Multi-dimensional Pattern Discovery in Financial Time Series using SAX-GA with Extended Robustness," In *Proceedings of the Fifteenth International Conference on Genetic and Evolutionary Computation Conference - GECCO '13*, pp. 179 – 180, 2013.
- [19] HistData, "Free Forex Historical data," [Online]. Available: <http://www.histdata.com/download-free-forex-data/>.