Alternate approach to Time Series reduction

Chee Keong Wee
Science and Engineering Faculty,
Queensland University of Technology
Brisbane, Australia
Chee.wee@qut.edu.au

Richi Nayak
Science and Engineering Faculty,
Queensland University of Technology
Brisbane, Australia
r.nayak@qut.edu.au

***Abstract* - Piecewise aggregate approximation is one of the methods used to reduce the dimensionality of time series by means of equal time frame, but it tends to have a lower accuracy. This investigation proposes an alternative to the PAA which can capture the essence of dimensionality reduction and minimise the loss of results.**

Keywords – time series, compression

I. INTRODUCTION

Time series is ubiquitous amongst all domains of industries, research, education, and medicine. The amount of data produced by modern equipment and machines are a constant challenge to maintain, interpret, store, and extract. There are several dimensionality reduction methods available such as Discrete Fourier Transform(DFT), Discrete Wavelet Transform(DWT), Singular Value Decomposition (SVD), Adaptive Piecewise Constant Approximation (APCA) and Piecewise aggregate approximation (PAA)(1-4). What PAA does is to segment a time series of *n* size into smaller vector of the desired period *M*, $\bar{X} = (\bar{x}_1, \bar{x}_2, \ldots\ldots, \bar{x}_M)$, where $M \leq n$. The $\bar{X}$ is calculated as;

$$\bar{x}_i = \frac{M}{n} \sum_{j=n/M(i-1)+1}^{(n/M)i} xj$$

Adaptive Piecewise Constant Approximation(APCA) is enhancement of PAA where it is more adaptive to the data values and so it will have shorter period for area with high activity (5). So, the vector is:

$$\bar{X} = (\{\bar{x}_1, r_1\}, \{\bar{x}_2, r_2\} \ldots\ldots, \{\bar{x}_M, r_M\})$$

Where,

$$x_i = mean(x_{r_{i-1}+1} \ldots x_{ri}) \text{ (5).}$$

The paper is organized as followed; in section 2, we described about our proposal for our alternate method of time series reduction. In section 3, we test the techniques against a series of time series data to show the various results that are derived from PAA versus those that are generated from our proposed method and with variation on the reducing parameters. In section 4, we discuss about the results and a conclusion.

II. PROPOSAL

Our proposal is an alternate variance to both PAA and APCA(4, 5). The objective is to achieve higher dimensionality reduction but at the same time, the amount of variation can be controlled in the time series reduction. The key is to specify the value of reduction - either by value or percent.

The algorithm will compute the mean for a predefined vector of time '*w*'. By using the mean of vector *X* at *w* length, it is compared with the next data point of w+1. If the difference is less than the reduction value, it will absorb into the vector and a new mean will be calculated. This new mean will then be compared to the next data point at time: $TS_{(w+\Delta w)}+1$, and if the value is greater or less than the mean $(X_{(w+\Delta w)})$, then a new vector will be formed. The cycle repeats until it reached the entire length of the time series as followed;

$$TS = ([w_1, a_1] \ldots . [w_1, a_m])$$

Where w is the varying time window, and *a* is the average of all the data points in the time window, *w*.

$$a_i = \frac{1}{w + \Delta w} \sum_{j=i}^{i+w+\Delta w} xj$$

We traverse through *w1 +1*, sampling the value of *y1+/- D1*. We reached the limit of *w1* when we find *y(1+Δw) >< y1(+/- D1). w1* and *y1* are recorded and then proceed on again for *w2*, resetting *w2=w0.* The algorithm is as followed;

1. The first initial width, *w0*, is defined.
2. The first average, *average_0*, is initialized with the same value as *ts[1],* where *ts* is the time series.
3. The reduction value, *d,* is set.
4. The counter for the window: *counter_w,* is set to zero
5. The routine loops through the entire length of *ts.*
6. The difference between *average_0* and the next data point, *ts[i]* is compared.
7. If the different is less than the reduction value stated, the value of *ts[i]* is added to *sum_w*.

8. *counter_w* is then incremented.
9. *a0*, the new average is calculated for the new window. Otherwise, if the difference if more than the reduction, Both the sum and the average will be initialized to the current *ts[i]* value.
10. The counter will be reset.

III. Results

The data used for this test is time series from a substation's power distribution in a local Brisbane suburb and it is in 15mins interval. The first test is to use PAA method with fixed intervals. This chart shows that the result for each individual window periods is taken at average without consideration on the difference between the upper and lower bound values of the data points as shown in figure 1.

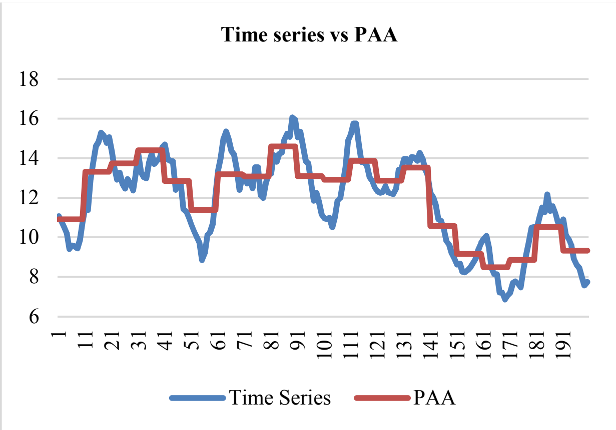
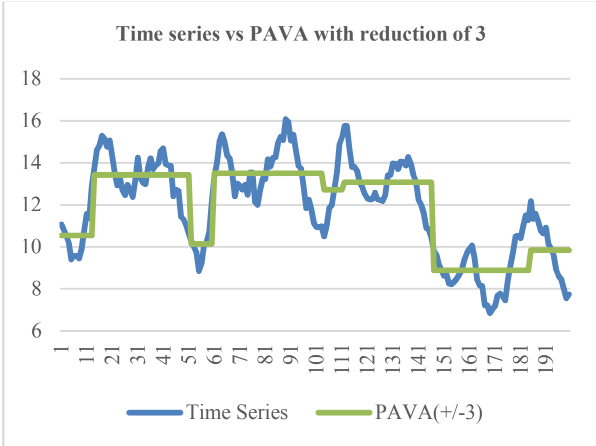


Fig. 1. – times series and PAA results

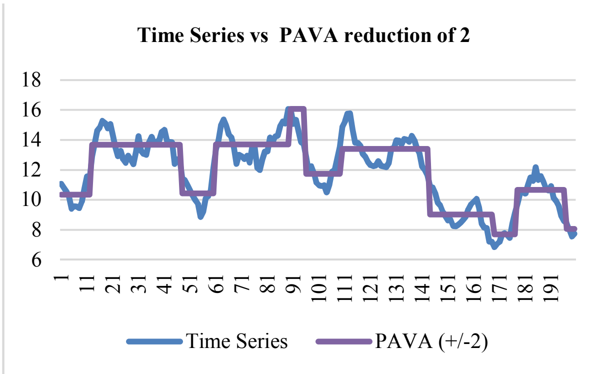
For the PAVA test, a reduction value of 3 is used. As figure 2 shows, it has taken a wider time frame to sample the data points, there are much fewer windows frames with more aggressive would result in losing too many details in the pursuit of dimension reduction. The representation of the reduced TS of 200 data points contains 7 vectors.



Results with reduction of 3, reduced time series is represented as; TS=[{13,10.540},{38,13.417},{9,10.137},{43,13.501},{8,12.719},{35,13.074},{38,8.879}]

Fig. 2 – Time series reduction with PAVA and a reduction of 3

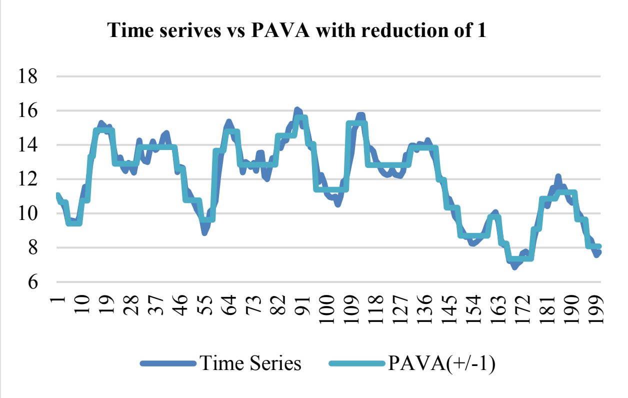
In figure 3, there are more windows periods and the means are more sensitive to the variation in the data points' values. There is still some loss in the details, but it contains more as compared to the higher reduction of 3. So the TS of 200 data points is now represented by 10 vectors.



Results with reduction of 2, reduced time series is represented as; TS=[{12,10.345},{35,13.668},{13,10.429},{28,13.688},{7,14.940},{14,11.740},{34,13.397},{25,9.021},{9,7.6866}, {18,10.669}]

Fig. 3. – Time series reduction with PAVA and a reduction of 2

In Figure 4, the reduction value has reduced to 1. And with greater granularity, the number of vectors has increased to 36. As the chart shows, PAVA is now more responsive to the change in the data points' values. So, it uses more window vectors to represent the time series.



Results with reduction of 1, reduced time series is represented as;
TS=[{0,12.1724},{4,10.653},{5,9.569},{3,11.226},{2,13.318},{7,14.855},{9,12.888},{14,13.857},{3,12.5896},{6,10.773},{5,9.624},{1,10.690},{1,12.132},{2,13.669},{5,14.773},{1,13.4981},{13,12.833},{7,14.534},{4,15.607},{3,14.050},{1,12.840},{11,11.382},{2,13.144},{5,15.259},{16,12.814},{10,13.831},{3,11.9526},{5,10.327},{11,8.693},{4,9.787},{3, 8.236},{9,7.342},{3,9.080},{6,10.86},{7,11.230},{4, 9.637}]

Fig. 4. – time series with PAVA and a reduction of 1

Table 1 shows the mean squared error (MSE) between PAA and PAVA with three different reduction values. While PAVA of higher reduction scored higher than PAA, the other two results with reduction of less than 2 have better scores.

TABLE 1 – MSE OF PAA AND PAVA WITH 3 REDUCTION VALUES

	PAA	PAVA(+/-3)	PAVA(+/-2)	PAVA(+/-1)
MSE	0.923691	1.016908	0.778338	0.409634

The table in figure 5 showed the time series with the results from PAA and the PAVA with three set of deviating values. PAA abides strictly by the time vectors and averages out the data points within, thus the value of standard reduction in these vectors will be high especially when that is high fluctuation of values(Anstey, 2007 #66).

The PAVA with different reduction values produced different numbers of vectors, each with a mean that is responsive to the individual vectors as what the reduction dictates.

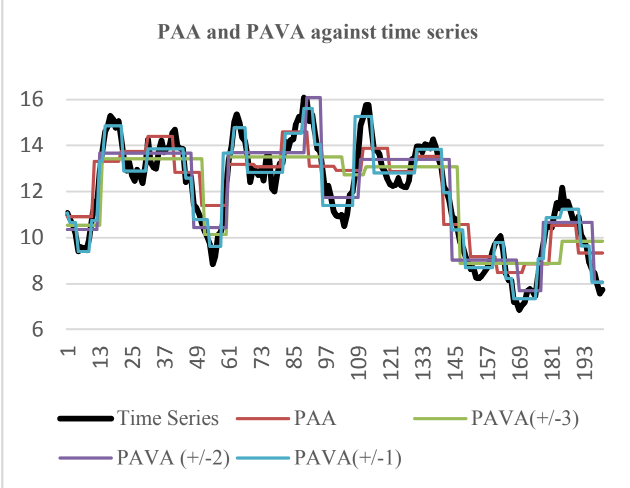


Figure 5 – time series with PAA and PAVA with all three reduction values

IV. DISCUSSION AND CONCLUSION

In this work, we have shown that the PAVA method is another alternate and viable time series reducing method to APCA, plus it has the feature to throttle the reduction with a basic parameter. Like APCA, PAVA can offer excellent reduction on time series where the data points have few fluctuations, but it will require more vectors to represent those data point with higher fluctuation (6). The reduced time series can be transposed easily back to a raw time series with minimum loss in details.(5)

The key advantage of PAVA over the rest lies in the reduction parameter which allow users to choose and achieve a balance between information reduction vs the amount of details lost. But one of the shortcoming is the variances of time windows between a group of time series that have been reduced with PAVA method and different reducing factors. Comparison among them will require computation to either to derive precise values based on a common time intervals or revert them back to the original raw time series before comparison can be done.

One of the future work to create a matrix of both varying time windows and reduction value to achieve much better dimension reduction while minimising the detail loss. Another option is to compare the moving average of window period to the previous one against the next data point in the time series to minimize the possibility of runaway gradual descend to global maxima. The vectors derived from a PAVA of a typical time series will not be the same with another time series that have been reduced by the same PAVA method. To compare two reduced time series using PAVA, we may need to explore future possibility of using Dynamic Time Warp (DFT) – PAVA ensemble method to achieve this.

REFERENCES:

1. Bianchi T, Piva A, Barni M. On the implementation of the discrete Fourier transform in the encrypted domain. IEEE Transactions on Information Forensics and Security. 2009;4(1):86-97.
2. Acharya T, Chakrabarti C. A survey on lifting-based discrete wavelet transform architectures. Journal of VLSI signal processing systems for signal, image and video technology. 2006;42(3):321-39.
3. Wall ME, Rechtsteiner A, Rocha LM. Singular value decomposition and principal component analysis. A practical approach to microarray data analysis: Springer; 2003. p. 91-109.
4. Guo C, Li H, Pan D. An Improved Piecewise Aggregate Approximation Based on Statistical Features for Time Series Mining. KSEM. 2010;10:234-44.
5. Keogh E, Chakrabarti K, Pazzani M, Mehrotra S. Locally adaptive dimensionality reduction for indexing large time series databases. ACM Sigmod Record. 2001;30(2):151-62.
6. Anstey JS, Peters DK, Dawson C, editors. An improved feature extraction technique for high volume time series data. Proceedings of the 4th conference on IASTED International Conference: Signal Processing, Pattern Recognition, and Applications; 2007.