On the Periodicity of Time-series Network and Service Metrics

Joseph T. Lizier and Terry J. Dawson
Telstra Research Laboratories
Sydney, NSW, Australia
{joseph.lizier, terry.j.dawson}@team.telstra.com

Abstract—**The presence of an underlying periodicity in time-series network and service metrics has been used as a basis for some recent anomaly detection techniques. These techniques however assume the presence of a periodicity, and would benefit from the concept of a quantitative figure of merit for the strength of a given periodicity in the metric. We survey a number of potential techniques for this purpose, and find none suitable. As such, we construct such a figure of merit to suit our application. Use of the figure of merit allows selection of the most appropriate period for the metric, and we present an efficient automated method for this selection. Furthermore, this figure of merit is a useful indicator of whether periodic analysis for anomaly detection is in fact suitable for the given metric. Finally, we suggest a number of other areas where use of the figure of merit could enhance anomaly detection using periodic analysis.**

I. INTRODUCTION

The presence of an underlying periodicity in various time-series has been cited in several fields, for example information and communications technology (ICT) [1] and climatology [2]. Such periodicity has been seen as an important property primarily from a modeling and predictive perspective; in ICT for example it has been cited as useful for resource scheduling [3]. Additionally, knowledge of an underlying periodicity has found another application in ICT: anomaly detection in network and service metrics, e.g. [1].

The term *network and service metrics* refers to variables used to characterize the performance of a network or service as a function of time. Typical metrics are bytes per minute at a network level, or transactions per minute at a service level. We consider only numeric metrics which are recorded at fixed discrete time intervals.

The term *anomaly detection* describes efforts to detect deviations of a system from normal behaviour. Typical deviations from normal behaviour are network failures and performance degradations. Here we consider anomaly detection focused on a single given network or service metric (i.e. univariate analysis). Any multivariate analysis (either within a system or across systems) is considered to be the task of a higher layer analysis system.

We are interested in the "two dimensional time-series" approach to anomaly detection described in [1], [4] and [5]. In this type of technique statistics are computed for each time point along a *characteristic period*, using a set of historical metric values from previous corresponding points in the period. Anomaly detection is then performed on each incoming metric value using the statistics for that corresponding time point in the period. This anomaly detection technique can benefit from a figure of merit to quantify the strength of a specific period in a given time-series metric. In this paper we seek to define such a figure of merit, and to use it to automatically identify the appropriate characteristic period and evaluate the applicability of the technique to the given metric. This motivation strongly guides our investigations, and as such we explain it in further detail in Section II.

We survey a range of potential methods for computing the figure of merit and identifying the characteristic period. However, as outlined in Section III, this survey does not identify a suitable method.

Therefore, we construct a figure of merit known as the *correlation of periods*, which is described in Section IV. A technique to efficiently search for and identify the characteristic period for a given metric, using the figure of merit, is outlined in Section V. We test the figure of merit and search technique against a number of real sample metrics, described in Section VI. This testing confirms that the methods are suitable for the purposes intended.

Finally, we present several other applications of the figure of merit in Section VII. These applications include the important step of determining whether the given metric is suitable for anomaly detection using the aforementioned type of periodic analysis. For this purpose, we present a guide on the suitability of metrics with various ranges of the figure of merit for the strength of the characteristic period.

II. MOTIVATION: ANOMALY DETECTION USING PERIODIC ANALYSIS

Anomaly detection in ICT systems has attracted significant interest, with recent motivations including security-related intrusion detection [4]. Anomaly detection encapsulates attempts to detect both hard faults (system failures) and soft faults (performance degradations) [6]. The detection of soft faults is considered very important, as it allows for the possibility of fault containment or correction, whereas the detection of a hard fault means that a failure is already present.

The most common method of anomaly detection for ICT metrics is threshold analysis, where an alarm is raised when the metric exceeds (or falls below) a static threshold value. Most network elements have agents to perform such tests [6]. Statistical process control (SPC) is a useful method, less commonly used for anomaly detection in ICT systems. SPC improves on static threshold analysis by incorporating statistical analysis of the metric over time to calculate a region for normal operation, which can be updated. An example application of SPC techniques to ICT metrics is given in [7]. A wide variety of other mechanisms for anomaly detection in ICT metrics have

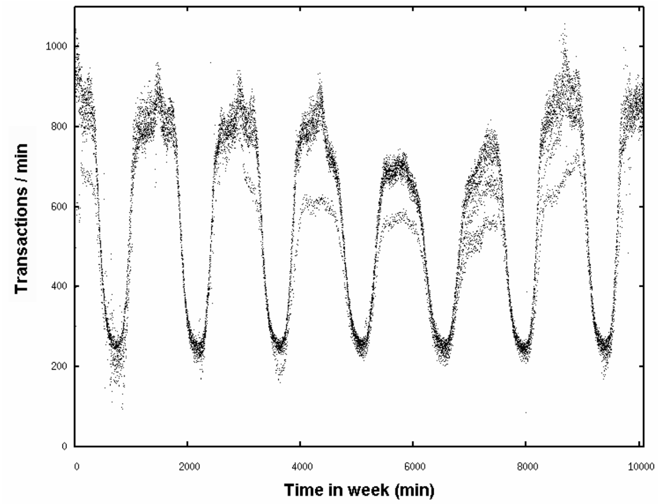

Figure 1. Transactions per minute versus time in week for "Transaction system 1". 8 consecutive weeks are superimposed. 1 week = 10080 min.

been reported, ranging from monitoring regularities in individual user behaviour [8] to microscopic self-similarity [9].

By definition, anomaly detection requires an (at least implicit) understanding of the normal behaviour of the metric under analysis. Many network and service metrics exhibit an underlying periodicity. For example, Fig. 1 shows transactions per minute on a highly used system over a number of weeks. There are clear similarities between corresponding days of the week, as well as between each day. Where such periodicity exists, it is sensible for an anomaly detection technique to incorporate it in a definition of normal behaviour of the metric.

To this end, several recently published anomaly detection techniques use the underlying periodicity of a given metric to determine normal behaviour. Data mining approaches to this are out of scope for our purposes (e.g. [10] defines temporal association rules between properties of transactions). We are concerned with techniques which compute a range for the normal behaviour of a numeric metric based on the underlying periodicity observed in historical values of the metric. We refer primarily to the work of Burgess *et al.* [1], [4], [5]. Related work includes that of Ho *et al.* [6], [11] and to a lesser extent Brutlag [12]. In [4], for a given metric a weighted mean and standard deviation are computed for each time point along a weekly interval, using historical values from the corresponding time in previous weeks. These weighted means and standard deviations are then used to construct a time-dependent range of expected normal behaviour; incoming values outside this range are labeled as anomalies. We shall refer to this as the anomaly detection technique of interest.

The above techniques all utilize a single characteristic period. Ho *et al.* [6] and Brutlag [12] assume their characteristic periods to be one week and one day respectively. Burgess *et al.* [1] describe that the daily and weekly periods are the only significant patterns in Fourier analysis and autocorrelation. They conclude that the weekly period was stronger than the daily period because of the larger variation between days of the week, and because weekends introduce larger uncertainty.

However, these methods did not seek to quantify the strength of the characteristic period in the metrics under analysis. As such, they were unable to make a quantitative decision on whether a daily, weekly or other period was most appropriate for a given metric. Furthermore, [5] gives an example of a metric with no obvious periodicity, and questions whether this type of periodic analysis is appropriate for that data set.

We conclude that the anomaly detection technique of interest would benefit from the concept of a figure of merit for the underlying periodicity in a given network or service metric. In the following sections, we seek to define the figure of merit and explore how it can be used to address the above issues.

III. Survey of Candidates for Figure of Merit Computation

In this section, we survey candidates for computing a figure of merit for the strength of a given periodicity in a metric. Some candidates are perhaps more strongly aligned to a search for the characteristic period than a definition of its strength, but we include these for completeness. We restrict our scope to discrete time signals, where the characteristic period is an integer multiple of the basic unit of time.

An obvious starting point is Fourier analysis. The Fast Fourier Transform (FFT) generally shows strong daily and weekly components for our sample metrics, as expected. Typically, the daily component greatly outweighed the weekly component, even where the weekly period was stronger from a self-similarity perspective (e.g. the metric in Fig. 1). This is a sensible result since the FFT computes the strength of sinusoidal components, yet given our focus on periodic self-similarity it is highly undesirable. Fourier analysis did however show promise in suggesting candidate periods of interest.

Another suggestion from a traditional perspective is the autocorrelation function, as by definition it evaluates a given data set's self-similarity after various shifts. Fig. 2 shows the autocorrelation of the metric plotted in Fig. 1. The autocorrelation plot of this strongly periodic metric (which is similar to Fig. 2 in [1]), oscillates as the metric is shifted along itself, reaching maxima after each shift of another day. On top of the daily oscillations, a weekly oscillation can be seen, with the peak after 7 days being the largest of the daily peaks, and the 3 and 4 days shifts being the smallest (of the shifts of less than a week). This elevation of the weekly period over the daily period fits with our intuitive understanding of the periodicity of this metric. Also, the property of the autocorrelation function in returning a value in [-1, 1] is quite appealing for the figure of merit. However, for a given shift, the autocorrelation function evaluates the similarity between data points separated by one shift. For application to anomaly detection, comparison to data points several shifts away is desirable. Also, skewing of the function is experienced where the length of input data is not an integer multiple of the characteristic period. Similarly, windowing effects are seen when using the FFT as a short-cut for the autocorrelation computation in an efficient time (via the Wiener-Khinchin Theorem [13]). Furthermore, methods would be required for differentiating meaningful periods from the high autocorrelation values at low shifts, and differentiating the

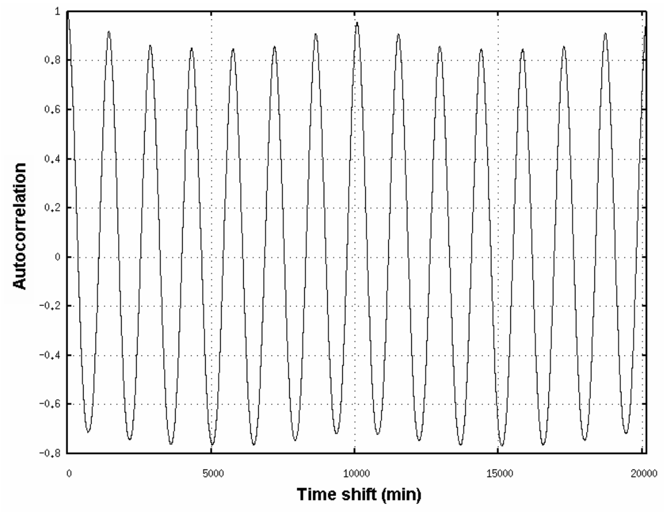

Figure 2. Autocorrelation of transactions per minute versus time shift for "Transaction system 1". The autocorrelation function was computed over 16 weeks of data. 1 week = 10080 min.

characteristic periods from integer multiples of themselves (which have similar autocorrelation values). Again, this function does not appear directly useful, though it shows promise for indirect use.

Several authors have reported data mining techniques for detecting periodicity, e.g. [3], [14], [15]. Typically this involves the discretization of the given data set, replacing numeric values of the data with the symbol for the corresponding interval, and performing a mining analysis for periodicity in the symbol series. From our perspective such approaches are too coarse, as the smoothing function of the discretization dilutes any numeric measure on the strength of the underlying periodicity.

We also note the reporting of a *square coherence statistic* in [2]. However this statistic is computed from a frequency domain perspective so does not give a true measure of periodicity from a time domain similarity perspective.

An interesting technique in [16] describes computing the strength of a periodicity n via the *singular value decomposition* (SVD) operation applied to the data series as an $m \times n$ matrix. (m being the number of periods of length n). The strength is given as the ratio of the first two singular values s_1 / s_2. The computation for a given periodicity is said in [16] to require $O(mn^2)$ time; since $n > m$ in general for our periods of interest, this is quite significant. Also, multiplicative differences in the shape of each period do not have an adverse effect on the strength, which is undesirable. The range of the strength value $[0,\infty)$ is not as desirable as $[-1, 1]$, but is still useful.

IV. Correlation of Periods Figure of Merit

We have constructed our own figure of merit for the strength of a period n in a given metric. This technique, the *correlation of periods* method, has been tailored for our anomaly detection purpose, and involves computing a correlation score between periods in the history of the metric. The steps of the method are as follows:

1. For a given length of historical metric values, take the largest integer number m of the given period n, and form into an $n \times m$ matrix (i.e. periods are arranged in columns).
2. Compute the linear correlation coefficient r_{ij} between each distinct period (i.e. between each column) i and j. Then scale the coefficient to account for the different amplitudes (encapsulated by α) and means (encapsulated by ρ) of the periods; i.e. the formula for r in [13] is scaled to become:

$$r_{ij} = \frac{\sum_{t=1}^{n}(x_i(t)-\bar{x}_i)(x_j(t)-\bar{x}_j)}{\alpha}\cdot\rho$$

$$\alpha = Max\left(\sum_{t=1}^{n}(x_i(t)-\bar{x}_i)^2, \sum_{t=1}^{n}(x_j(t)-\bar{x}_j)^2\right). \quad (1)$$

$$\rho = \frac{Min(\bar{x}_i,\bar{x}_j)}{Max(\bar{x}_i,\bar{x}_j)}$$

3. Average these correlation scores to obtain the figure of merit *FOM(n,m)* for the strength of the given period n over m periods:

$$FOM(n,m) = \frac{2}{m(m-1)}\sum_{i=1}^{m}\sum_{j=i+1}^{m} r_{ij}\ . \quad (2)$$

This technique addresses many of the shortcomings of the previously described methods, with respect to our application to anomaly detection. It takes into account how each instance of the period relates to each other instance, not only their neighbors in time, which is very well aligned with our purpose. The use of an integer number of the given period only (as also done in [16] and [2]) removes the problem of skewing due to an incomplete period. Also, the figures of merit for short periods are not unreasonably large (unlike those of the Autocorrelation function). Furthermore, this figure of merit is defined on the range [-1, 1], which was previously noted as desirable. Finally, the computation of this figure of merit requires $O(m^2n)$ time (there are $O(m^2)$ correlation computations, each performed in $O(n)$ time) and generally $n > m$ for periods of interest. This is less efficient than an autocorrelation computation for a given shift value at $O(mn)$, however provides a more thorough overall measure of periodic self-similarity.

On the virtues of these properties, we adopt the correlation of periods method as our figure of merit for the strength of an underlying periodicity in a network or service metric.

The technique is also somewhat extensible to cater for potential variations in the relevant anomaly detection techniques. References [4] and [5] describe computing the anomaly detection statistics from the historical values of the metric using exponential decay of the contribution of each week. This could be reflected in a similar decay of the contribution of each correlation score to the figure of merit, based on the relative times of the two periods (to each other and to the present time). Also, the anomaly detection technique of interest does not incorporate a linear trend into the range of expected values. If this were to be done, then the computation of the figure of merit should involve the removal of the linear trend from the historical window. The linear trend would then be incorporated into the range of expected values in the future.

V. Efficient Search for the Characteristic Period

Intuitively, the characteristic period of the metric should be the period with the highest figure of merit. For our application, the realistic major candidates for the characteristic period are one day or one week, so perhaps only these candidates need have their figures of merit evaluated. Nonetheless, it is useful to understand how the characteristic period would be identified if this assumption were invalid or for other applications. Appendix I shows that an exhaustive search computing the figure of merit for all possible periods will have asymptotic complexity $O(N^2 \log N)$, where N is the length of the historical window of metric values (note: $N \geq nm$ for any given period n and number of periods m). It is desirable to find a search method more efficient than this.

Furthermore, periods at integer multiples of the characteristic period are likely to display peaks in the figure of merit at values similar to that of the characteristic period. This is also observed in [16] with respect to the SVD based method. It is desirable for the search method to distinguish between these and identify only the characteristic period.

To address these issues, we have constructed a more efficient method to search for the characteristic period using the figure of merit. This method harnesses advantageous properties of the Fourier transform and autocorrelation function, and is composed of the following steps:

1. Take the autocorrelation function of the given window of historical metric values. This is done harnessing the Wiener-Khinchin Theorem [13] to compute the autocorrelation function via the FFT in $O(N \log N)$ time.
2. Take the FFT of the autocorrelation function. As Fig. 2 shows, the important candidate periods produce an oscillation in the autocorrelation function, so will produce peaks in its FFT (at corresponding frequencies). (Note: Step 1 can be computationally eliminated, because step 2 is computed first via the Wiener-Khinchin Theorem shortcut).
3. Select a number of local maxima from this FFT as frequencies of interest. This can be done by selecting a fixed number of the most significant local maxima, or all local maxima with a minimum spectral content.
4. For each frequency of interest:
   a. Convert the frequency of interest to the corresponding range of periods of interest.
   b. For each period of interest, compute the figure of merit for the strength of the period.
   c. Store the period of interest with the maximum figure of merit for this frequency of interest. This period becomes a candidate period. If this period is at the edge of the range of periods of interest, continue evaluating figures of merit beyond the range in order to find a local maximum (This accounts for skewing in the FFT).
5. From the candidate periods, select that with the largest figure of merit as the characteristic period.

As discussed in Section VII, if the selected characteristic period has a figure of merit below a certain threshold, it should be concluded that no characteristic period exists.

Appendix I shows that this algorithm scales as $O(N \log N)$ for an increase in the time length of historical data analyzed. This is a better scaling than the exhaustive search. However, the algorithm is also shown to scale as $O(N^2 \log N)$ for an increase in the density of points per unit time. While this scaling is no better than the exhaustive search, the performance of this algorithm is better by a significant constant factor, since by definition it evaluates the figure of merit for a finite proportion of possible periods. We conclude that the algorithm is a definite improvement in efficiency over an exhaustive search.

This method also effectively eliminates the candidacy of integer multiples of the characteristic period, since they will not correspond to a major frequency component in the FFT. Similarly, harmonics of the characteristic period's frequency are likely to produce only candidate periods with insignificant figures of merit. This is a significant improvement as this issue was highlighted in related work, e.g. [16]. Also, if the daily and weekly periods are the only candidates evaluated, this method provides some insight into whether the weekly period is strong in its own right, or only as an integer multiple of the daily period. This amounts to whether the weekly period produces a major frequency component of the FFT of the autocorrelation.

Finally, it is important to note that this efficient search method is not specific to our correlation of periods figure of merit. That is to say, it could be used with any related figure of merit method, e.g. the SVD based method [16].

VI. Results of Application to Sample Data Sets

In order to assess the suitability of the *correlation of periods* figure of merit, we applied it to several real sample metrics. The results are shown in Table I. For each metric, the computation was made over various lengths of historical data (from 1 month up to 1 year in some cases) and from various starting points in time; this results in a range of figures of merit for each metric.

For each data set here (aside from the "Single internet customer") the weekly period was computationally stronger than the daily period, though the difference was more pronounced in some metrics than others. This was expected from a visual inspection of the data sets. This is by no means a universal result – there are data sets where the daily period will be the characteristic period (e.g. that shown in [12]).

A reasonable spread of figures of merit is displayed by our metrics. There is no "correct" answer to compare the results with; however a visual inspection of the metrics confirms that the relative ranking of the metrics' periodicity by the figure of merit seems qualitatively correct. For example, the figures of merit for the weekly period of "Transaction System 1" indicate a strongly periodic metric – this is confirmed by the visualization of this metric in Fig. 1. Also, the negligible figures of merit for the weekly and daily periods in the "Single internet customer" metric were confirmed in that no periodicity was evident in a visual inspection of the metric.

These qualitative results provide evidence that the correlation of periods figure of merit is an appropriate measure of the strength of a given periodicity in a metric, from the perspective of the anomaly detection technique of interest.

TABLE I. FIGURES OF MERIT FOR SAMPLE NETWORK AND SERVICE METRICS

Metric Name	Sampling rate interval	Figure of Merit range	
		Weekly period	***Daily period***
Transaction system 1 – trans. / min	5 min	0.86 to 0.96	0.70 to 0.77
Transaction system 2 – trans. / min	1 min	0.64 to 0.72	0.30 to 0.36
Telephony platform 1 – calls / min	1 min	0.86 to 0.97	0.59 to 0.67
IP router 1 – bytes / min in	15 min	0.51 to 0.57	0.24 to 0.27
IP router 2 – bytes / min in	15 min	0.61 to 0.67	0.25 to 0.30
Single internet customer 1 – bytes / min in	5 min	0.0010	0.0018

In addition to the suitability of the figure of merit computations, the efficient search for the characteristic period was observed to function well. For all sample metrics except for "Single internet customer", the technique successfully identified one week as the characteristic period of the metric.

An example search over 26 weeks of data of the "Telephony platform 1 – calls/min" metric (with a 1 minute sampling rate, giving 10080 time points per week) took approximately 12 minutes. It must be noted that the search technique is currently implemented in scripts for the Octave mathematical environment [17], which are run in a single threaded mode. While the time for the code to run is not prohibitive, it can certainly be improved with parallel figure of merit calculations in compiled rather than scripted code.

It was noted that a shift in to or out of daylight saving during the historical data affected the figure of merit computations where the timestamp used was a Universal Time Co-ordinate (UTC) or offset equivalent. The use of local time co-ordinates (that shift with daylight saving) was found to uphold the identification of the expected period.

Finally, notice that the sample data sets here include network elements in both telephony and IP as well transaction systems. The results demonstrate the broad applicability of the figure of merit across various types of metrics, and importantly the applicability of the anomaly detection technique of interest.

VII. APPLICATIONS FOR THE FIGURE OF MERIT

The primary application for the figure of merit is identification of the characteristic period of the metric under analysis. However, there are a number of areas to which the utility of the figure of merit can be extended. Some have been explored here, others are left to future work.

An important use of the figure of merit is in determining whether the anomaly detection technique of interest is valid for the given metric. If the metric does not display a strong periodicity, then other anomaly detection techniques such as traditional threshold analysis or SPC may be more appropriate. In attempting to apply an anomaly detection technique based on [4] and [5] to our sample data sets, we have made observations for the following approximate ranges of the figure of merit for the characteristic period:

1. Greater than 0.75 – a figure of merit in this range indicates strong periodicity. This type of technique appears highly appropriate for such metrics.
2. Between 0.5 and 0.75 – a figure of merit in this range indicates moderate periodicity. This type of technique appears to have some applicability here. Its application could be improved by using a longer window of historical data to compute the statistics, and with a large degree of smoothing or "local averaging" (as labeled in [4]) of the statistics across neighboring points in the period.
3. Less than 0.5 – a figure of merit in this range indicates mild periodicity down to no periodicity (around zero) and anti-periodicity (represented by negative scores). The technique does not appear applicable for such metrics.

Prior to the anomaly detection technique of interest being applied to any given metric, the figure of merit for the characteristic period should be evaluated (as per the above criteria) to establish whether the technique is appropriate.

The figure of merit could also be used to determine the length of historical metric values used for computing the statistics for each time point in the period. Since the figure of merit calculation indicates the strength of the period during a given window of historical data, calculations over various window lengths could be used to detect process changes and guide selection of the window length.

Furthermore, the figure of merit could add to conjecture over the appropriate sampling rate or counting interval to use in monitoring the metric. Reference [1] suggests that the sampling rate could be inferred from the autocorrelation half-life, and that no significant changes are expected over intervals of 5 minutes or so. Further insight could be gained by computing the figure of merit over various counting intervals for a given metric, and using its variation to guide the selection of counting interval.

Of course, the figure of merit could be used to quantitatively investigate the prevalence of periodicity in various types of network and service metrics. There is some evidence that periodicity is more prevalent in metrics which represent aggregate data rather than data from individual users or from systems with low volume or sporadic use. It could be said that periodicity is a population property – while individual users are guided by similar periodic influences, the periodicity is only strongly evident when examining the group behaviour. Reference [1] notes that stronger trends were observed in metrics from hosts that experience more usage, and that the strength of periodicity depends on how strongly the metric is coupled to the periodic force of the system users. These observations are echoed in our own quantitative results here – the metrics exhibiting the strongest periodicity were macroscopic metrics (e.g. total transactions per minute) from systems handling the largest number of users. Also, the only metric shown here without any significant periodicity was from a single user system.

Finally, it is worth noting that the anomaly detection technique of interest only harnesses a single characteristic period. Related methods however could incorporate multiple characteristic periods (be they harmonically related as one day and one week, or otherwise). The figure of merit could be useful in identifying the multiple periods, perhaps being computed for each candidate period after the influence of the preceding period is removed from the metric.

VIII. Conclusion

We have identified the need for a better understanding of periodicity in network and service metrics for an anomaly detection technique of interest. Subsequently, we constructed the correlation of periods figure of merit for the strength of the periodicity in such metrics. This figure of merit can be used to identify the characteristic period of the metric, and an efficient search technique was presented for this purpose. Another important application of the figure of merit is in quantitatively evaluating whether the anomaly detection technique of interest is appropriate for a given metric. Other uses for the figure of merit were discussed, and are potential areas for future work. Future work will also include probing the noise immunity of our figure of merit and search technique, using generated sample data sets.

Appendix I. Asymptotic Complexity of Searches for the Characteristic Period

An exhaustive search for the characteristic period involves computations of the figure of merit over the potential periods n = 1 to $N/2$. Each figure of merit evaluation has complexity $O(m^2 n) = O(N^2/n)$. So, the total number of operations in the exhaustive search is:

$$\begin{aligned}
\sum_{n=1}^{N/2} \frac{N^2}{n} &= N^2\left[1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\frac{1}{5}+\frac{1}{6}+\frac{1}{7}+..+\frac{1}{N/2}\right] \\
&\le N^2\left[\begin{array}{l} 1+\left(\frac{1}{2}+\frac{1}{2}\right)+\left(\frac{1}{4}+\frac{1}{4}+\frac{1}{4}+\frac{1}{4}\right)+.. \\ ..+\left(\frac{1}{2^q}+\ldots+\frac{1}{2^q}\right) \end{array}\right] \qquad (3) \\
&\le N^2\left[1+(1)+(1)+(1)+\ldots+(1)\right] \\
&= N^2(1+q),
\end{aligned}$$

where $q = \lfloor log_2 N \rfloor - 1$. That is, the asymptotic complexity of the exhaustive search for the characteristic period is $O(N^2 \log N)$. In a practical situation, the figure of merit may not be evaluated for periods larger than say $N/3$ or smaller than 3, however this does not alter the asymptotic complexity.

The run time of the efficient search is dominated by the loop over the frequencies of interest in steps 3 and 4. We assume the algorithm selects a fixed number of maxima from the frequency spectrum. The number of periods of interest corresponding to this fixed number of frequencies is directly proportional to the density d of samples in the time series metric, i.e. $O(d)$. Now, the figure of merit evaluation for each period of interest costs $O(m^2 n)$, however since we don't know which periods will be periods of interest we estimate the average cost as that of the exhaustive search averaged over each period, i.e. $O(N \log N)$. So, the run time of the efficient search scales as $O(dN \log N)$ – for an increase in the time length of historical metric data this scales as $O(N \log N)$; for an increase in the density of samples it scales as $O(N^2 \log N)$.

References

[1] M. Burgess, H. Haugerud, S. Straumsnes, T. Reitan, "Measuring system normality", *ACM Trans.Computer Systems*, vol..20, no.2, pp. 125-160, May 2002.

[2] R. Lund, H. Hurd, P. Bloomfield, R. Smith, "Climatological time series with periodic correlation", *J. Climate*, vol. 8, no. 11, pp. 2787-2809, Dec. 1995.

[3] A. Andrzejak, M. Ceyran, "Characterizing and predicting resource demand by periodicity mining", *J. Network and Systems Management*, vol. 13, no. 2, Jun. 2005.

[4] M. Burgess, "Two dimensional time-series for anomaly detection and regulation in adaptive systems", in *Proc. 13th IFIP/IEEE Int. Workshop Distributed Systems: Operations and Management (DSOM) (Lecture Notes in Computer Science Vol.2506)*, Montreal, Oct. 2002, pp.169-180. Springer Verlag.

[5] M. Burgess, "Probabilistic anomaly detection in distributed computer networks", *Science of Computer Programming*, in press.

[6] L. L. Ho, D. J. Cavuto, S. Papavassiliou, A. G. Zawadzki, "Adaptive and automated detection of service anomalies in transaction-oriented WAN's: network analysis, algorithms, implementation, and deployment", *IEEE J. Selected areas in Communications*, vol. 18, no. 5, pp. 744-757, May 2000.

[7] N. Ye, S. Vilbert, Q. Chen, "Computer intrusion detection through EWMA for autocorrelated and uncorrelated data", *IEEE Trans. Reliability*, vol. 52, no. 1, pp. 75-82, Mar. 2003.

[8] A. Seleznyov, "A methodology to detect temporal regularities in user behavior for anomaly detection", in Dupuy, M., Parandinas, P. (eds) *Trusted information: The new decade challenge: Proc. 16th Int. Conf. Information Security (IFIP/Sec '01)*, Paris, Jun. 2001, pp. 339-352. Kluwer Academic Publishers.

[9] W. H. Allen, G. A. Marin, "On the self-similarity of synthetic traffic for the evaluation of intrusion detection systems", in *Proc. 2003 Symp. Applications and the Internet (SAINT 2003)*, Orlando, Jan. 2003, pp. 242-248. IEEE Comput. Soc.

[10] Y. Li, N. Wu, S. Wang, S. Jajodia, "Enhancing profiles for anomaly detection using time granularities", *J. Computer Security*, vol. 10, no. 1-2, pp. 137-157, 2002.

[11] L. L. Ho, D. J. Cavuto, S. Papavassiliou, A. G. Zawadzki, "Adaptive Anomaly Detection in Transaction-Oriented Networks", *J. Network and Systems Management*, vol. 9, no. 2, pp. 139-160, Jun. 2001.

[12] J. D. Brutlag, "Aberrant Behavior Detection in Time Series for Network Monitoring", in *Proc. 14th Systems Administration Conf. (LISA)*, New Orleans, Dec. 2000, pp. 139-146. USENIX Assoc.

[13] W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery, *Numerical Recipes in C, 2nd Edition*. Cambridge: Press Syndicate of The University of Cambridge, 1992, pp. 496-500, p. 636.

[14] S. Ma, J. L. Hellerstein, "Mining partially periodic event patterns with unknown periods", in *Proc. 17th Int. Conf. on Data Engineering (ICDE), Heidelberg*, April 2001, pp. 205-214. IEEE Comput. Soc.

[15] C. Berberidis, I. Vlahavas, W. G. Aref, M. Atallah, A. K. Elmagarmid, "On the discovery of weak periodicities in large time series", *in Proc. 6th European Conf. Principles of Data Mining and Knowledge Discovery (PKDD) (Lecture Notes in Computer Science vol. 2431)*, Helsinki, August 2002, pp. 51-61. Springer Verlag.

[16] P. P. Kanjilal, J. Bhattacharya, G. Saha, "Robust method for periodicity detection and characterization of irregular cyclical series in terms of embedded periodic components", *Phys. Rev. E*, vol. 59, no. 4, pp. 4013-4025, Apr. 1999.

[17] J. W. Eaton, GNU Octave, 1998. Available: http://www.octave.org/