Matching Detections to Events in Time Series*

Michel Reis[1], Rebecca Salles[2], Geraldo Xexeo[3], Rafaelli Coutinho[1], Eduardo Ogasawara[1]

[1] Centro Federal de Educação Tecnológica Celso Suckow da Fonseca - CEFET/RJ

[2]INRIA

[3]COPPE/UFRJ

michel.reis@aluno.cefet-rj.br, rebecca.pontes-salles@inria.fr

xexeo@cos.ufrj.br,rafaelli.coutinho@cefet-rj.br,eogasawara@ieee.org

Abstract. *SoftED metrics introduce a soft evaluation of event detection methods in time series, incorporating fuzzy logic concepts to provide temporal tolerance in detections. However, these metrics face challenges associating detections with events, especially in cases with multiple associations between detections and events. In this work, we propose structuring this association problem within the graph theory paradigm, approaching it as a bipartite graph matching problem. For this, the Hungarian algorithm is employed to solve the association problem. The results demonstrate the effectiveness of the proposed approach, highlighting the impact of improvements in the associations between detections and events.*

1. Introduction

A time series is a set of observations of a variable of interest collected over time, where the behavior of this series is studied as a function of past data [Hanssens et al., 2003]. In most time series, significant behavioral changes can be observed at some point, characterized as events [Guralnik and Srivastava, 1999]. An event can represent a phenomenon with a defined meaning in a particular domain, making the identification of an event occurrence relevant. Event detection can be done by analyzing time series [Pimentel et al., 2014].

An important aspect of event detection is how they are reported. When events are labeled, event detection is typically characterized similarly to a classification problem. Hence, classification metrics, such as accuracy, recall, precision, and F1, are used [Lavin and Ahmad, 2016; Tatbul et al., 2018]. In a classification context, an incorrect classification is characterized as an error and is duly penalized by standard classification metrics. In this article, these metrics are considered hard metrics.

In an event detection context, an incorrect classification from the perspective of hard metrics may not be characterized as an error when considering the concepts of time and temporal tolerance. In some situations, a detection made sufficiently close to an event can be useful instead of incorrect. Standard classification metrics do not account for this situation because the event detection problem is not a common classification problem. There was a need for soft metrics to incorporate temporal tolerance. Therefore, soft metrics for event detection were created, called SoftED metrics [Salles et al., 2023], which

*The authors thank ao CNPq, CAPES, and FAPERJ for partially sponsoring this research.

associate detections with events. These soft metrics are available in the Harbinger framework [Salles et al., 2020].

The need to associate detections with events brings new challenges to achieving this pairing. This work innovates by noting that this pairing can be mapped to a well-known problem, the bipartite graph matching. To solve the problem, this article incorporates the Hungarian algorithm [Kuhn, 1955] for bipartite graph matching, with a computational complexity of $O(n^3)$, in problem-solving. The results indicate scenarios where such an approach benefits SoftED and the computational improvement in addressing this issue.

In addition to this introduction, this work is organized into four more sections. Section 2 presents the theoretical foundation of SoftED metrics. Section 3 introduces the new method. Section 4 presents the experimental evaluation. Finally, Section 5 provides the conclusions.

2. Soft Metrics for Event Detection

SoftED metrics incorporate concepts of soft computing [Tettamanzi and Tomassini, 2013] to apply temporal tolerance in their measurement. Figure 1 presents the intuition of how soft computing is used to measure event detection. The green circles represent perfectly detected events (event and detection simultaneously). The blue circles represent events without perfect detection. The red circles represent imperfect detections. Hard detections associate 0 or 1 with the detections. Soft evaluation assigns a continuous value between 0 and 1 to the detections, thus weighing the degree to which the detection relates to a specific event [Salles et al., 2023].

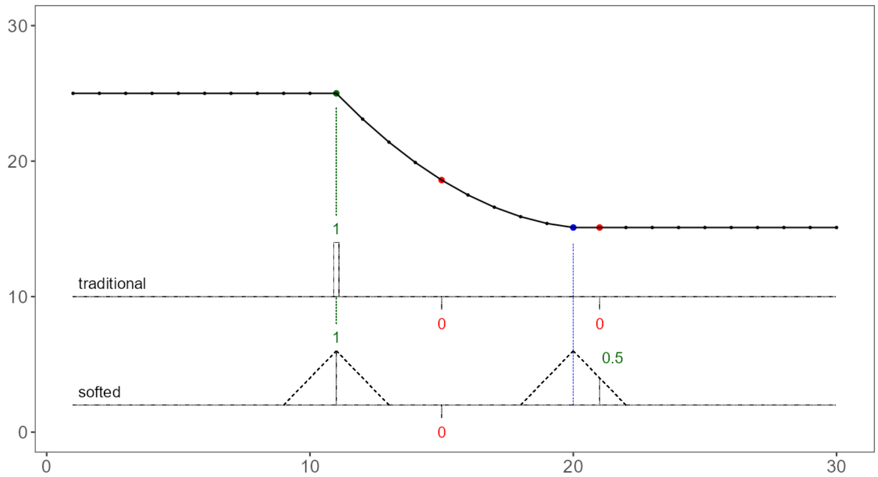

Figure 1. Hard and soft evaluation of event detection [Salles et al., 2023]

The relevance degree of a detection concerning a specific event is defined by a triangular membership function inspired by fuzzy sets [Zadeh, 1965]. This technique meets the demand for more flexible evaluations of event detections. Expanding the detection evaluation area of an event creates an integrity problem. As can be seen in Figure 2, situations occur where (a) a detection could be related to many events, or (b) an event could be related to many detections.

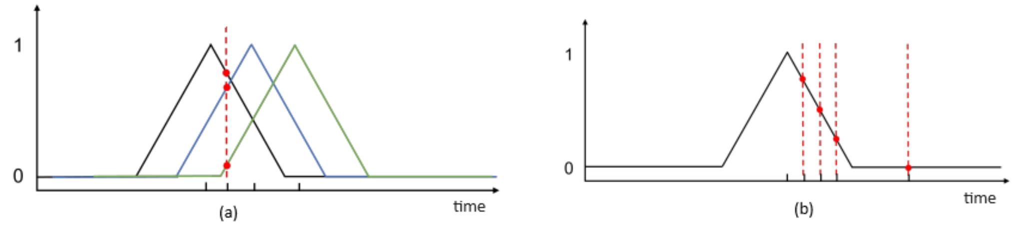

Figure 2. Integrity problem [Salles et al., 2023]

To maintain integrity with standard metrics, each detection was associated with the event that provided the highest score according to the evaluations of the membership functions of each event. After that, the remaining events and detections go through the same process, and so on, until events or detections have no association [Salles et al., 2023]. This solution partially resolves the problem, as shown in Figure 2. In (a), a detection associated with the event that best evaluates it; the other events receive an evaluation of 0. In (b), the highest-evaluated detection is associated with the only event that occurred; the other detections receive an evaluation of 0.

Despite this approach to resolving the integrity problem, another problem occurs when membership functions overlap and more than one detection. As seen in Figure 3, Detection 1 associates with Event 2, creating the detection-event pair $[(1, 2)]$ with a value of 0.8. This approach was not the most appropriate, as seen when a detection always associates with the nearest event. It causes other potentially better associations to be disregarded. The detection-event pairs $[(1, 1), (2, 2)]$ produce the value $0.6 + 0.4 = 1.0$, suggesting these would be better associations.

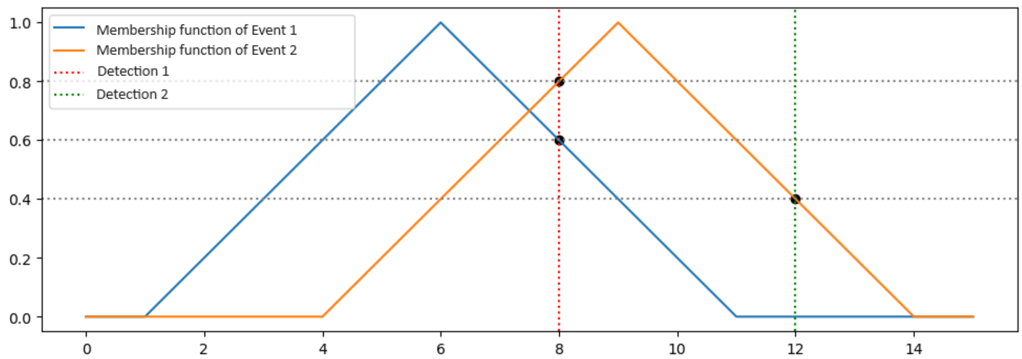

Figure 3. Associations between events and detections

3. Method

The problem of evaluating event detections using soft metrics can be structured within the graph theory paradigm and understood as a bipartite graph matching problem [Bollobás, 1979]. In the proposed context, a time series contains a set of detections $D = \{d_1, d_2, ..., d_n\}$ and a set of events $E = \{e_1, e_2, ..., e_m\}$. Consider a matrix M where each element M_{ij} represents the result of the membership function of e_j associated with detection d_i. This relationship can be represented by a bipartite graph $G = (D, E, A)$, where each edge $(d_i, e_j) \in A$ has a value corresponding to M_{ij}.

A matching in G is a subset of edges $P \subseteq A$ such that no edges share a common vertex, i.e., each vertex in D is matched with exactly one vertex in E. Given the matrix M, our goal is to find a matching that maximizes the total value given by the sum of the weights of the edges included in the matching. To solve this problem, this work uses the Hungarian algorithm to find the minimum cost matching in a cost matrix C, which is the negation of the matrix M representative of the graph G. As a result, this operation returns the subset of edges P, which represents the optimal detection-event pairs.

The mapping algorithm consists of the following steps:

1. **Initialization**: Create a cost matrix C whose initial values are equal to the negation of the matrix M. Subtract the smallest value in each row from all elements in that row, then subtract the smallest value in each column from all elements.
2. **Zero Coverage**: Cover all zeros in the cost matrix using the smallest possible number of horizontal and vertical lines.
3. **Matrix Adjustment**: If the number of covered rows and columns equals the number of rows or columns in the matrix, a maximum weighted matching has been found. Otherwise, subtract the smallest uncovered value from all uncovered elements and add that value to the elements covered by two lines.
4. **Repetition**: Repeat until maximum weighted matching is found.

After this process, the M_{ij} values are extracted for each pair $(d_i, e_j) \in P$ so that the edge set P is mapped to the detection results vector ds. This vector is used to create the soft versions of TP, FP, TN, and FN to calculate soft metrics. In Table 1, $|X|$ denotes the total length of the time series, and m represents the total number of events in the time series. This definition fully complies with the principles of SoftED [Salles et al., 2023].

Table 1. Formalization of SoftED metrics

$\text{TP}_s = \sum_{i=1}^{n} ds(d_i)$	$\text{FN}_s = m - \text{TP}_s$
$\text{FP}_s = \sum_{i=1}^{n} (1 - ds(d_i))$	$\text{TN}_s = (\lvert X \rvert - m) - \text{FP}_s$

Any matching formed by selecting pairs from this operation's resulting zeros is optimal [Kuhn, 1955]. The execution time of this adaptation of the Hungarian algorithm to solve the problem in this approach is $O(q^3)$, with $q = max(|D|, |E|)$.

4. Experimental Evaluation

The experiments were conducted using the Harbinger library [Salles et al., 2020] in RStudio. The computational environment was composed of an Intel Xeon W3-2423 processor with 512GB of RAM and 12 cores, running Ubuntu 22.04 LTS. A qualitative evaluation was first performed on a synthetic series to demonstrate the method's operation, followed by a quantitative evaluation on real datasets. Each test was assessed using three metrics: Hard (hard metric), SoftED Trad (traditional version), and SoftED Par (paired version).

4.1. Qualitative Evaluation

Table 2 presents an interval of a synthetic series in which three events and three detections occur. Each combination of Event and Detection will receive a value determined by the membership function of each event, as shown in Table 3.

Table 2. Values of the synthetic series in the interval [48-59] for event and detection

	48	49	50	51	52	53	54	55	56	57	58	59
Event	F	F	**T**	**T**	F	F	**T**	F	F	F	F	F
Detection	F	F	**T**	F	F	**T**	F	F	F	F	F	**T**

Table 3. Matrix with values of the membership function (k = 10)

	Event 1	Event 2	Event 3
Detection 1	1.0	0.9	0.6
Detection 2	0.7	0.8	0.9
Detection 3	0.1	0.2	0.5

The values calculated by the membership function are influenced by the temporal tolerance constant k. This constant determines which area around the event is considered for calculating this function. This work addresses situations where more than one detection occurs in an area where more than one membership function operates, as shown in Figure 3. The method's effectiveness is more evident in scenarios where this overlap occurs, as it improves detection-event associations in this context.

As seen in Table 3, the pairing algorithm adopted in this article chooses the detection-event pairs $[(1,1),(2,2),(3,3)]$ whose values are $[(1.0),(0.8),(0.5)]$ and the total is 2.3. The previous approach would choose the pairs $[(1,1),(2,3)]$ whose values are $[(1.0),(0.9)]$ and the total is 1.9.

Table 4 compares the evaluation metrics results in this synthetic series. It is clear that, in a case where the reported overlap exists, the new approach produces better associations, enabling a better evaluation.

Table 4. Comparison of the metrics results in synthetic series

Metric	Hard	SoftED Trad	SoftED Par
Accuracy	0.96	0.97	0.98
Precision	0.33	0.63	0.76
Recall	0.33	0.63	0.76
F1	0.33	0.63	0.76

4.2. Quantitative Evaluation

For the evaluation on real datasets, samples of ten time series of global stock indices were used: IBOV, AUS200, DE30, DJ, ESP35, HK50, JP225, NAS, S&P500, and UK100. Twenty thousand observations from the year 2019 were extracted for each series. The series was pre-processed with differentiation and z-score, the distribution of the samples can be viewed in Figure 4. The datasets were trained on seven different event detection algorithms in the Harbinger library: ARIMA, FBIAD, DTW, K-Means, FFT, GARCH, and Wavelet. In total, 70 tests and 210 evaluations were performed, with $k = 3$ for the evaluations with soft metrics. The execution time of the evaluations was also calculated.

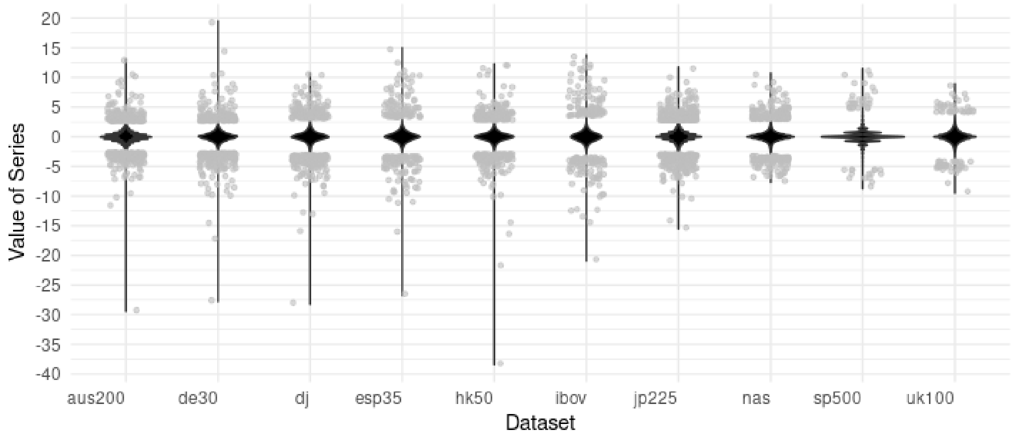

Figure 4. Violin plots of each dataset with jittered points highlighting outliers

The average execution time of the Hard evaluations was recorded as 0 seconds, likely because the algorithm operates in $O(N)$ time, simply iterating through the array once to compare elements. The SoftED Trad approach had an average execution time of 11.77 seconds, compared to 5.84 seconds for the SoftED Par approach. Figure 5 also shows that the method is considerably faster than the previous version and offers slightly superior results, possibly due to the rarity of the conflicts it addresses. The choice of these series was due to the fact that the spacing between events had a good probability of creating the situations contemplated by this work.

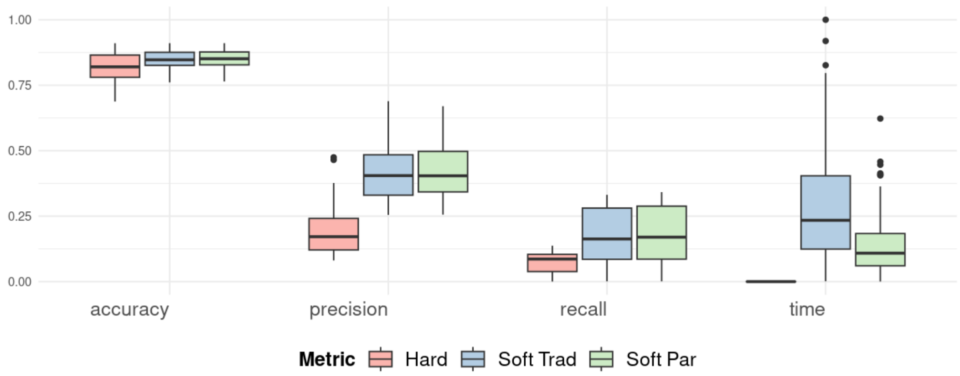

Figure 5. Comparison of metrics in accuracy, precision, recall, and time

5. Conclusion

This article evaluated event detection methods in time series using SoftED metrics, which incorporate fuzzy logic to provide temporal tolerance. We proposed a new approach for associating detections with events based on the Hungarian algorithm, solving the problem as a bipartite graph matching. The main results show that this new methodology slightly improved the quality of detection-event associations in scenarios with overlapping membership functions and multiple detections and significantly reduced execution time compared to the previous method. For future work, we plan to review and adapt the Hungarian algorithm to optimize the time complexity of SoftED Par. This optimization will facilitate the effective application of the method in large-scale or real-time environments.

References

Bollobás, B. (1979). *Graph Theory*, volume 63 of *Graduate Texts in Mathematics*. Springer, New York, NY.

Guralnik, V. and Srivastava, J. (1999). Event Detection from Time Series Data. In *Proceedings of the Fifth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '99, pages 33–42, New York, NY, USA. ACM.

Hanssens, D. M., Parsons, L. J., and Schultz, R. L. (2003). *Market Response Models: Econometric and Time Series Analysis*. Springer, Boston, Mass., 2nd edition.

Kuhn, H. W. (1955). The Hungarian method for the assignment problem. *Naval Research Logistics Quarterly*, 2(1-2):83–97.

Lavin, A. and Ahmad, S. (2016). Evaluating real-time anomaly detection algorithms - The numenta anomaly benchmark. In *Proceedings - 2015 IEEE 14th International Conference on Machine Learning and Applications, ICMLA 2015*, pages 38 – 44.

Pimentel, M. A., Clifton, D. A., Clifton, L., and Tarassenko, L. (2014). A review of novelty detection. *Signal Processing*, 99:215 – 249.

Salles, R., Escobar, L., Baroni, L., Zorrilla, R., Ziviani, A., Kreischer, V., Delicato, F., Pires, P. F., Maia, L., Coutinho, R., Assis, L., and Ogasawara, E. (2020). Harbinger: Um framework para integração e análise de métodos de detecção de eventos em séries temporais. In *Anais do Simpósio Brasileiro de Banco de Dados (SBBD)*, pages 73–84. SBC.

Salles, R., Lima, J., Coutinho, R., Pacitti, E., Masseglia, F., Akbarinia, R., Chen, C., Garibaldi, J., Porto, F., and Ogasawara, E. (2023). SoftED: Metrics for Soft Evaluation of Time Series Event Detection. SoftED, arXiv, 2023-04-04.

Tatbul, N., Lee, T. J., Zdonik, S., Alam, M., and Gottschlich, J. (2018). Precision and recall for time series. In *Advances in Neural Information Processing Systems*, volume 2018-December, pages 1920 – 1930.

Tettamanzi, A. and Tomassini, M. (2013). *Soft Computing: Integrating Evolutionary, Neural, and Fuzzy Systems*. Springer Science & Business Media.

Zadeh, L. (1965). Fuzzy sets. *Information and Control*, 8(3):338 – 353.