Temporal Association Rules from Motifs

Mauro Milella
Department of Mathematics and Computer Science, University of Ferrara, Italy

Giovanni Pagliarini
Department of Mathematics and Computer Science, University of Ferrara, Italy

Guido Sciavicco
Department of Mathematics and Computer Science, University of Ferrara, Italy

Ionel Eduard Stan
Department of Informatics, Systems, and Communications, University of Milano-Bicocca, Italy

Abstract

A motif is defined as a frequently occurring pattern within a (multivariate) time series. In recent years, various techniques have been developed to mine time series data. However, only a few studies have explored the idea of using motif discovery in temporal association rule mining, and only limited to rules that do not temporally relate motifs with each other. Interval-based temporal association rules have been recently defined and studied, along with the temporal version of the known frequent patterns, and therefore, rule extraction algorithms (such as APRIORI and FPGrowth). In this work, we first define a simple algorithm to discover motifs from a dataset of multivariate time series, and build over them a vocabulary of propositional letters; second, we show how apply the temporal version of association rule discovery on such a vocabulary. With a careful choice of motifs, their lengths, and the thresholds for propositional letter construction, the extracted rules turn out to be very natural, with a very high level of interpretability. We apply our methodology to time series datasets in the fields of hand signs execution and gait recognition, and we discuss the resulting rules from an interpretability point of view.

2012 ACM Subject Classification Theory of computation → Modal and temporal logics; Theory of computation → Theory and algorithms for application domains

Keywords and phrases Motifs, Interval Temporal Logic, Association Rules

Digital Object Identifier 10.4230/LIPIcs.CVIT.2016.23

Acknowledgements We acknowledge the support of the FIRD project Methodological Developments in Modal Symbolic Geometric Learning, funded by the University of Ferrara. Moreover, this research has also been funded by the Italian Ministry of University and Research through PNRR - M4C2 - Investimento 1.3 (Decreto Direttoriale MUR n. 341 del 15/03/2022), Partenariato Esteso PE00000013 - "FAIR - Future Artificial Intelligence Research" - Spoke 8 "Pervasive AI", funded by the European Union under the NextGeneration EU programme

1 Introduction

In machine learning we typically separate between *functional* and *symbolic* learning. The former encompasses all algorithms and strategies for learning a *function* that represents the theory underlying a certain phenomenon. The latter, on the other hand, consists of learning a *logical description* that represents that phenomenon. In the context of modern *explainable artificial intelligence*, natively symbolic approaches maybe preferred over a posteriori extraction of rules from black-box systems: as O'Neil puts it [16], opaque models are the *dark side* of big data. *Classification* and *rule extraction* are, among others, typical problems of machine learning; while classification can be dealt with using both functional and symbolic learning, rule extraction is essentially symbolic. From a logical standpoint, canonical symbolic learning methods are all characterized by being based on propositional logic, and by being designed for static data, in which every instance is described by the value of n attributes. In general, temporal data cannot be successfully dealt with within the same schema in a native way. The standard approach to information extraction from temporal data starts with a pre-processing phase designed to abstract non-static data so that their aspect becomes static again (e.g., by averaging the value of attributes along all dimensions), and that off-the-shelf static symbolic methods can be used. Modal, and in particular temporal symbolic learning [14, 18] takes a different point of view: modal symbolic methods are characterized by being based on modal logic (e.g., temporal logic) so that non-static data can be dealt with natively, and that the extracted knowledge takes the form of interpretable modal logic formulas. So far, modal symbolic learning has been mainly focused on classification, but recent approaches suggested that association rule extraction can benefit of a similar approach [15, 20], with the introduction of the modal adaptation of the known frequent set extraction algorithms, namely APRIORI [1] and FP-Growth [8], to the modal case.

The temporal case is considered a representative example to illustrate the qualities and the characteristics of modal association rules. Nevertheless, the original approach presents some important limitations, originated in the basic definition of temporal alphabet, which may cause difficult-to-interpret association rules.

One way to overcome the limits in alphabet definition towards temporal association rule extraction is to consider motifs. *Motifs* in time series are patterns that are considered interesting because they are frequently occurring, and their name is due to the resemblance of such a concept with its discrete counterpart in computational biology [10]. Motif discovery for time series was introduced in 2003 [5], and several improvements on the original technique were introduced afterwards. A time series motif is naturally interpreted on an interval, and via the use of thresholds one can immediately transfer the idea of motif to the idea of propositional letter that encapsulates a motif. In this way, a natural definition of alphabet emerges, that allows us to use interval temporal logic for temporal rule extraction.

In this paper, we consider the problem of temporal rule extraction from time series using a motif-based alphabet. We adapt the original definition of *local support*, introduced in [15, 20], to accommodate motifs in a suitable way, and we design a simple adaptation of the algorithm ModalAPRIORI [20] with such a modification. Then, we apply our methodology to two temporal datasets, and show how the obtained rules have an immediate natural language translation. Our code is fully open-source and available[1].

2 Background

Motifs and motif discovery. Motif discovery for time series was introduced in 2003 [5], generating quite a body of research. From a practical point of view, motifs have been applied to solve problems in a wide variety of domains such as bioinformatics, speech processing, robotics, human activity understanding, severe weather prediction, neurology and entomology; see, among many others, [2, 21, 22]. From a theoretical one, the research has focused on extensions and generalizations of the original work, especially in the attempt to improve scalability; see [12, 17], among others. Motif discovery falls into two broad classes: *approximate* and *exact* motif discovery; here, we focus on the latter. Virtually every time

[1] See https://github.com/aclai-lab/ModalAssociationRules.jl, part of the project *Sole.jl*.

series data mining technique has been applied to the motif discovery problem, including indexing, data discretization, triangular-inequality pruning, hashing, early abandoning.

▶ **Definition 1.** *A* time series *is a sequence* $T : t_1, \ldots, t_N$ *of* N *real-valued observations. A set of* $\mathcal{T}$ *time series* $T_1, \ldots, T_n$ *is called* multivariate *time series. A region* $T_{ij} : t_i, \ldots, t_j$ *of* $j-i+1$ *consecutive observations in a time series is called a* subsequence*. Given two subsequences* T_i *and* T_k*, their* distance *is the Euclidean distance between the* Z*-normalized version of* T_i *and* T_k*.*

We can take a subsequence and compute its distance to all subsequences in the same time series.

▶ **Definition 2.** *Given an* N*-observations time series* T *and a subsequence* T_{ij}*, the* distance profile D_i *is the vector of all distances between* T_{ij} *and* $T_{k(k+j-i+1)}$*, for every* $1 \leq k \leq N$*.*

Observe that given a time series T and a subsequence T_{ij}, the value of D_i at (near) the point i is (close to) 0, by definition. While a value of 0 in D_i at some point k far away from i is called a *match*, the (near) 0 values in the vicinity of i are referred to as *trivial matches*. Given a parameter s (called *shadow*), the distance profile D_i at points in the interval $[i-s, i+s]$ can be artificially set to ∞ in order to avoid trivial matches. Now, given a time series T of N points and a length $l \leq N$, there are exactly $N-l+1$ points at which a subsequence of length l may start.

▶ **Definition 3.** *Given a time series* T*, a length* l*, and a shadow* s*, the* full distance matrix $M_{T,l,s}$ *(or, simply,* M*, when all parameters are clear from context) is the squared matrix of dimension* $N-l+1$ *whose* (i,j)*-th entry is the distance between the subsequences* $T_{i(i+l-1)}$ *and* $T_{j(j+l-1)}$*. Given the full distance matrix* M*, the* matrix profile P *is the vector that at position* i *contains the value of the minimal distance that any subsequence* $T_{j(j+l-1)}$ *presents with* $T_{i(i+l-1)}$*.*

Two subsequences $T_{i(i+l)}$ and $T_{j(j+l)}$ with a distance less than some threshold α are instances of a *motif*. Computing the matrix profile of a time series enables us to compute its motifs.

The literature on motif discovery is relatively rich, and different libraries implementing the most efficient algorithms for solving this problem are currently being maintained by several groups. The ecosystem of software solutions into which this technique falls is continually expanding; as of 2024 a primitive directive in Matlab to efficiently calculate matrix profiles was introduced. In this work, we integrated into such ecosystem the framework *Sole.jl*[2], an end-to-end open source Julia solution for symbolic learning and reasoning with non-tabular data, such as multivariate time series, enriching it with *MatrixProfile.jl*[3], that includes an efficient implementation of STAMP [25], for the computation of the matrix profile of a time series. In this work, we focus on a particular kind of motif called *snippet*; snippets can be seen as the *most representative* motifs in a time series [11], that is, they can be sorted by how much of a time series is explained by each snippet.

Modal association rules extraction. Rule-based methods are ubiquitous in modern machine learning and data mining applications [7], that aim at learning regularities in data which can be expressed in the form of *if-then* rules. In this paper, we are interested in descriptive (or association) rules, that collectively cover the instances' space. Fixed

[2] https://github.com/aclai-lab/Sole.jl
[3] https://github.com/baggepinnen/MatrixProfile.jl

HS modality	Definition w.r.t. the interval structure			Example
$\langle A \rangle$ (after)	$[i,j]R_A[i',j']$	$\Leftrightarrow$	$j = i'$	
$\langle L \rangle$ (later)	$[i,j]R_L[i',j']$	$\Leftrightarrow$	$j < i'$	
$\langle B \rangle$ (begins)	$[i,j]R_B[i',j']$	$\Leftrightarrow$	$i = i' \wedge j' < j$	
$\langle E \rangle$ (ends)	$[i,j]R_E[i',j']$	$\Leftrightarrow$	$j = j' \wedge i < i'$	
$\langle D \rangle$ (during)	$[i,j]R_D[i',j']$	$\Leftrightarrow$	$i < i' \wedge j' < j$	
$\langle O \rangle$ (overlaps)	$[i,j]R_O[i',j']$	$\Leftrightarrow$	$i < i' < j < j'$	

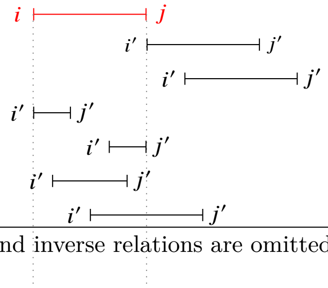

Table 1 Allen's relation and their notation within HS; equality and inverse relations are omitted.

an alphabet $\mathcal{P} = \{p_1, \ldots, p_k\}$, a *propositional rule* is an object of the type $\rho : X \Rightarrow Y$, where $X \subset \mathcal{P}$ is called *antecedent*, $Y \subset \mathcal{P}$ is called *consequent*, and $X \cap Y = \emptyset$. The use of the non-logical symbol $\Rightarrow$ emphasizes the fact that association rules are not logical implications. Following the standard literature, interesting association rules are extracted via statistical considerations: a set of propositional literals, typically named *items* or *itemset* in this scenario [1], $X \cup Y$ is considered interesting when it is *frequent*, that is, if it occurs more often than a predetermined threshold referred to as *minimum support*, and a rule $X \Rightarrow Y$ is extracted if the ratio between the support of X and that of $X \cup Y$ is higher than another predetermined threshold known as *minimum confidence*. Most association rule mining techniques leverage the computation of support and confidence measures (i.e., the support-confidence framework) to mitigate the combinatorial explosion due to itemsets generation, which is a #P problem [23, 24], but many of the generated rules could still be filtered out as non-interesting; on top of objective measures, the practice suggests that the interestingness level of a rule should always be double-checked by an expert. However, a wide variety of objective interestingness measures and statistical tests can be used to reveal important insights about the rules; such measures are based on the definition of support and can augment the support-confidence framework to further reduce the number of rules in the output.

The idea of extending association rules to the case of non-tabular data, generally represented as *modal* data, that is, in which every instance is described as graph, was first introduced in [15, 20], in which the generalization of the known rule extraction algorithms APRIORI and FP-Growth were proposed (respectively *ModalAPRIORI* and *ModalFP-Growth*). From this general setting, it is possible to specialize the rule definition and extraction to the (qualitative) temporal case, on the line of [3].

Temporal association rule extraction. While classical association rules are designed for propositional patterns to emerge (i.e., items are not temporalized, and their co-presence is assumed to be contemporary), temporal association rules are designed to generalize this idea to patterns with a temporal component. The natural choice to describe temporalized co-occurrence of events or patterns with a duration is interval temporal logic. The standard choice for temporal logic of intervals is the modal logic defined over Allen's interval-interval relations. Given a linear order D, an *interval* on D is defined as a pair $i < j$, where $i, j \in D$. Allen's relations relate any two intervals in one of thirteen possible binary relations, as shown in Tab. 1, where we omit equality and inverse relations for simplicity.

▶ **Definition 4.** *Let $\mathcal{X} = \{A, \overline{A}, L, \overline{L}, B, \overline{B}, E, \overline{E}, D, \overline{D}, O, \overline{O}, =\}$. The well-formed formulas of the* Halpern and Shoham Modal Logic of Time Intervals *(*HS*, for short) are built from an alphabet $\mathcal{P}$, the classical connectives $\vee$ and $\neg$, and a modality for each Allen's interval relation,* *as follows:*

$$\varphi ::= p \mid \neg\varphi \mid \varphi \vee \varphi \mid \langle X \rangle \varphi,$$

where $p \in \mathcal{P}$ *and* $X \in \mathcal{X}$*. The other propositional connectives and constants (i.e.,* $\psi_1 \wedge \psi_2 \equiv \neg\psi_1 \vee \neg\psi_2, \psi_1 \rightarrow \psi_2 \equiv \neg\psi_1 \vee \psi_2$ *and* $\top = p \vee \neg p$*), as well as, for each* $X \in \mathcal{X}$*, the* universal modality $[X]$ *(e.g.,* $[A]\varphi \equiv \neg\langle A \rangle \neg\varphi$*), can be derived in the standard way.*

A formula of HS is interpreted over an *interval model*, defined as pair $\mathcal{T} = \langle D, V \rangle$[4], where $V : I(D) \rightarrow 2^{\mathcal{P}}$ is a *valuation function* from the set $I(D) = \{[i,j], i < j, i,j \in D\}$ of all intervals that can be built on D.

▶ **Definition 5.** *Given an interval model* $\mathcal{T}$ *based on the alphabet* $\mathcal{P}$*, an interval* $[i,j]$ *in it, and a formula* φ *of HS, we say that* φ *is satisfied by* $\mathcal{T}$ *at* $[i,j]$*, and we denote it by* $\mathcal{T},[i,j] \Vdash \varphi$*, if and only if:*

$$\begin{array}{lll}
p \in V([i,j]) & \textit{iff} & \varphi = p; \\
\mathcal{T},[i,j] \not\Vdash \psi & \textit{iff} & \varphi = \neg\psi; \\
\mathcal{T},[i,j] \Vdash \psi \textit{ or } \mathcal{T},[i,j] \Vdash \xi & \textit{iff} & \varphi = \psi \vee \xi; \\
\mathcal{T},[i',j'] \Vdash \psi \textit{ for some } [i',j'] \textit{ such that } [i,j]R_X[i',j'] & \textit{iff} & \varphi = \langle X \rangle \psi
\end{array}$$

Interval temporal logic gives us a way to naturally describe temporal association rules, as time series can be naturally seen as interval models. Let $\mathfrak{T} = \{\mathcal{T}_1, \mathcal{T}_2, \ldots, \mathcal{T}_m\}$ be a set of (multivariate) time series, or *temporal dataset*, and fix a propositional alphabet $\mathcal{P}$; let us also assume that each time series in $\mathfrak{T}$ is based on the same temporal domain D. Each single multivariate time series $\mathcal{T}$ (e.g., the temporal history of an hospitalized patient) is a collection of n time series $T_1, \ldots, T_n$ (e.g., the collection of all time-changing values that describe an hospitalized patient); elements of $\mathcal{P}$ are naturally associated to a specific time series (e.g., the truth value of *high fever* is associated to the, time-changing, body temperature of an hospitalized patient). In this way, if $\mathcal{T}$ is a multivariate time series (i.e., an interval model), $[i,j]$ is an interval, and φ an interval formula, the notion of $\mathcal{T},[i,j] \Vdash \varphi$ can be interpreted as the notion of *the multivariate time series* $\mathcal{T}$ *satisfies* φ *at* $[i,j]$.

▶ **Definition 6.** *Given an alphabet* $\mathcal{P}$*, a* temporal literal *(or* temporal item*)* λ *is defined by the following grammar:*

$$\lambda ::= p \mid \langle X \rangle \lambda \mid [X]\lambda,$$

where $p \in \mathcal{P}$ *and* $X \in \mathcal{X}$*. The set of all possible temporal items is denoted by* $\Lambda_{\mathcal{P}}$*.*

A *temporal itemset* is a set of temporal items, and temporal rules antecedents and consequents are temporal itemsets. The notion of support is generalized to the temporal case, as follows.

▶ **Definition 7.** *Let* $\mathfrak{T}$ *be a temporal dataset,* $\Lambda_{\mathcal{P}}$ *be the set of temporal items built on the alphabet* $\mathcal{P}$*, and let* $X \subseteq \Lambda_{\mathcal{P}}$ *be an itemset. The* local support *of* X *on some instance* $\mathcal{T} \in \mathfrak{T}$ *is defined as the relative frequency of* X *holding on all the possible intervals:*

$$lsupp(\mathcal{T}, X) = \frac{|\{[i,j] \in I(D) \mid \mathcal{T},[i,j] \Vdash X\}|}{|I(D)|},$$

[4] We intentionally use the same symbol for interval model and multivariate time series, to convey the idea that the latter can be interpreted as the former.

and, given a certain minimum local support threshold $s_l \in (0,1] \subset \mathbb{R}$, the global support *of X on $\mathfrak{T}$* relatively to *s_l is defined as:*

$$gsupp_{s_l}(\mathfrak{T}, X) = \frac{|\{\mathcal{T} \in \mathfrak{T} \mid lsupp(\mathcal{T}, X) \geq s_l\}|}{|\mathfrak{T}|}.$$

ModalAPRIORI and *ModalFP-Growth* can be used to extract temporal patterns as a particular case of modal patterns; an itemset X is said to be *frequent* on $\mathfrak{T}$ or *globally frequent* if its global support is greater than a minimum global support threshold s_g, and after extracting all the globally frequent itemsets and using them to generate rules, we need to judge which of the latter are association rules and which are not, employing, as already mentioned, suitable interestingness measures. This has been suggested, for example, in [15], where it is proposed to use standard temporal information extraction functions to devise a suitable propositional alphabet. The primary characteristics of such information extraction functions is that they can be applied to an interval of time and reduce the subseries in that interval to a scalar value. In this work, we show how, in general, this strategy can be improved by replacing scalars with motifs (and, specifically, snippets) for alphabet building; however, this requires adaptation to the mining algorithms.

3 Extracting Temporal Association Rule from Motifs

Alphabet discovery in time series: naïve approach. In order to generalize association rule mining to the temporal scenario, we need to enhance the support-confidence framework by integrating a suitable strategy to capture relevant temporal aspects from data. The notion of *feature extraction function* is naturally applied to this case. Given a multivariate time series $\mathcal{T} = \{T_1, \ldots, T_n\}$, one can consider a set of functions $\mathcal{F} = \{F_1, \ldots, F_k\}$, where each function F is defined as $F : \mathbb{R}^d \to \mathbb{R}$ for some natural value $d \leq N$, and then apply such functions to all intervals. Via arbitrarily selecting thresholds α, β, one can then create an alphabet

$$\mathcal{P} = \{\alpha \leq F(T) \leq \beta \mid F \in \mathcal{F}, T \in \mathcal{T}, \alpha \in \mathbb{R} \cup \{-\infty\}, \beta \in \mathbb{R} \cup \{+\infty\}\}.$$

Letters in $\mathcal{P}$ are immediately interpreted over intervals, as follows: given an interval $[i, j]$, F is applied to the segment $T_{i,j}$ to obtain a scalar value that is then compared with α and β. Theoretically speaking, this definition allows mining temporal association rules; indeed, given an interval model $\mathcal{T}$, an alphabet $\mathcal{P}$ and the set of temporal items $\Lambda_{\mathcal{P}}$ built over $\mathcal{P}$, the local support semantics is well-defined.

Unfortunately, this approach may introduce strong bias during model checking, leading to a sterile mining.

▶ **Example 8.** Consider Fig. 1, in which the feature extraction function *maximum* is applied to every subsequence of the time series depicted on the left. As a result, we obtain the same scalar for every interval including $[2, 3]$.

Behaviours such as the above one constitute a major obstacle towards interpretability, possibly leading to association rules with promising interestingness levels which, however, encode trivialities. More refined feature extraction functions (e.g., *hctsa* [6], *catch22* [13]) do not mitigate this problem, since it is inherently caused by our alphabet definition.

Alphabet discovery in time series: an approach with motifs. In order to avoid degenerate situations such as the one described in the above example, we can modify the definition of alphabet. To this end, we first extend the notion of representative motifs for

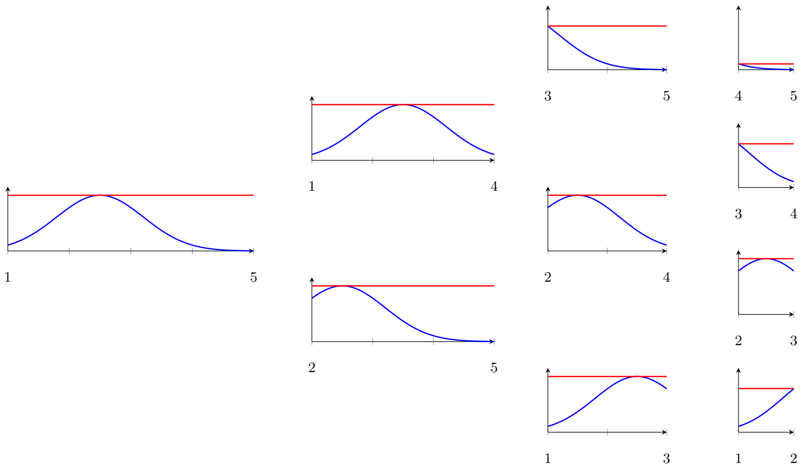

Figure 1 Maximum feature extraction function applied to the time series in blue on the left, and to all its subintervals. The value of such function, which is highlighted in red, is identical in every interval including $[2, 3]$.

motif extraction from a temporal dataset, via a generalized full distance matrix $M_{\mathfrak{T},N,s}$. This allows us to obtain an collection of representative motifs from a temporal dataset $\mathfrak{T}$:

$$\mathcal{A}_{\mathfrak{T}} = \{\mu \mid \mu \text{ is representative in } \mathfrak{T}\}.$$

Then, fixed a distance function δ and a constant α, we define a motif-based temporal alphabet as:

$$\mathcal{P} = \{\delta(T, \mu) \leq \alpha \mid T \in \mathfrak{T}, \mu \in \mathcal{A}_{\mathfrak{T}}, \alpha \in \mathbb{R}\}.$$

As in the previous case, letters in $\mathcal{P}$ have an intuitive and immediate interpretation on intervals: a given $p \in \mathcal{P}$ is true on a given segment T_{ij} if and only if the behaviour of T within the segment is close enough to the motif μ encapsulated by p.

▶ **Example 9.** Consider the example in Fig. 2. The two individual time series represent, respectively, the position of the right hand on the y-axis (T_1) and on the x-axis (T_2), in the context of a body-tracking data analysis exercise whose aim is to describe some specific movement in the most rigorous manner. Assume that motifs μ_1 and μ_2 had been extracted from a temporal dataset $\mathfrak{T}$ of similar movements. We define two temporal literals $p ::= \delta(T_1, \mu_1) \leq 1.0$ and $\lambda ::= \langle B \rangle \delta(T_2, \mu_2) \leq 1.0$, where δ is the Z-normalized Euclidean distance and $\langle B \rangle$ is the Allen's relation *begins*, and consider an hypothetical association rule $A ::= p \Rightarrow \lambda$. As it can be seen, such a rule can easily be translated to natural language, without being limited by the hidden redundancy of a naïve alphabet definition: *when the right-hand moves away from the hip to the right at an approximate speed of* 0.9 *distance units per time unit, for two consecutive time units, then the movement has started with the same hand being at the highest point.*

Temporal association rules discovery with motifs. In order to lift the motif-based definition of alphabet to the idea of association rule extraction, we need to be able to compute the support of temporal items and itemsets built on it. Unfortunately, simply applying our original approach based on local support does not provide a suitable solution.

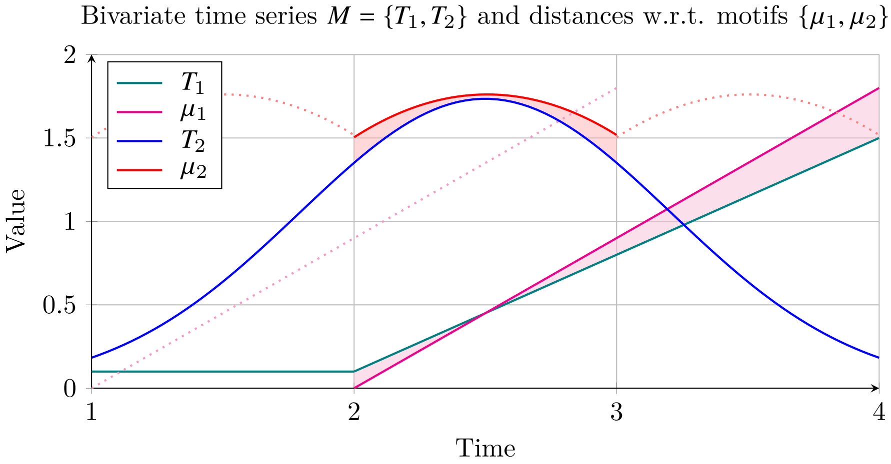

Figure 2 Bivariate time series consisting of T_1 and T_2, respectively in blue and green, and two motifs μ_1 and μ_2, respectively in red and magenta. Since $|\mu_1| = 1$, the euclidean distance $\delta(T_1, \mu_1)$ is calculated for all the 1-length intervals of T_1. When the distance is small enough, it is highlighted with a colored area. Similar applies for T_2 and μ_2.

► **Example 10.** Consider, again, the situation depicted in Fig. 2, and the problem of computing the local support of p, λ, and $\{p, \lambda\}$ as defined in Ex. 9. Since motifs have a predetermined length in terms of time units, p, for instance, would have a counter-intuitive upper bound of $\frac{2}{6}$, since it cannot hold on intervals of length different from 2.

Computing the local support of an item as the ratio between the number of intervals that satisfy that item and the number of intervals that may potentially satisfy it would be the simplest solution to the above problem; for items without temporal modal connectives, in particular, this requires to examine only intervals of the same length as the motif encapsulated by it. This solution, however, may not be optimal for temporal items that are not literals, as it may cause an unwanted effect similar to the one that emerged with the naïve alphabet. A more elaborate solution requires the idea of anchored itemset.

► **Definition 11.** *An itemset X is said to be* anchored *if and only if contains at least one non-temporal item, and all non-temporal items in it are based on motifs of the same length; the subset $\Omega \subset X$ of non-temporal items is called* anchor *of X. Let $l(\Omega)$ denote the length of the interval on which an anchor may hold.*

In this way, the length of intervals that may potentially satisfy the entire itemset is fixed, allowing us to define the frequency of that itemset in a truly representative way.

► **Definition 12.** *Let $\mathfrak{T}$ be a temporal dataset, $\Lambda_{\mathcal{P}}$ be the set of temporal items built on the motif-based alphabet $\mathcal{P}$, let $X \subseteq \Lambda_{\mathcal{P}}$ be an anchored itemset, and let $\Omega \subset X$ be its anchor. The* motif-based local support *of X on some instance $\mathcal{T} \in \mathfrak{T}$ is defined as:*

$$mblsupp(\mathcal{T}, X) = \frac{|\{[i,j] \in I(D) \mid T, [i,j] \Vdash X\}|}{|\{[i,j] \in I(D) \mid j - i + 1 = l(\Omega)\}|},$$

and, given a certain minimum motif-based local support threshold $s_l \in (0,1] \subset \mathbb{R}$, the motif-based global support *of X on $\mathfrak{T}$* relatively to s_l *is defined as:*

$$mbgsupp_{s_l}(\mathfrak{T}, X) = \frac{|\{\mathcal{T} \in \mathfrak{T} \mid mblsupp(\mathcal{T}, X) \geq s_l\}|}{|\mathfrak{T}|}.$$

Using the modified version support, and computing the other interesting measures based on it, we now perform a series of rule extraction experiments.

4 Experiments

In our experiments, we consider two well-known public datasets concerning human action and gait recognition, namely NATOPS [19] and HuGaDB [4]; given the intuitive nature of the data, extracted rules will be interpretable even to the non-expert, unlike other scenarios. Although both datasets are labeled and designed for classification, our objective is to describe common patterns within the same class, and to express them in natural language.

The experimental setup is as follows. First of all, we establish the values for the minimum local and global support, and for the minimum confidence, respectively s_l, s_g and c. Given a dataset, we focus on all the instances belonging to a specific class and concatenate (each time series of) its instances. We then proceed to extract the most representative motifs from each time series by leveraging the algorithm *Snippet-Finder* proposed in [11]; to this end, we set reasonable lengths for potential motifs. After the motif extraction phase, we define an alphabet $\mathcal{P}$ that takes into account each motif μ, the Z-normalized Euclidean distance δ, and a threshold α, the latter being the s_g-th percentile of the values obtained by computing δ between μ and all subintervals of the corresponding time series. We define the set of temporal items $\Lambda_{\mathcal{P}}$ by considering every Allen's relation, with the exception of $\langle L \rangle$, which tends to introduce trivial redundancies in local support computation (i.e., fixed an anchor, the intervals covered by *later* relation are significantly more numerous than those covered by the other relations).

We choose to mine frequent itemsets by leveraging ModalAPRIORI, generating all emerging association rules starting from *closed* itemsets, that is, itemsets whose none of their immediate supersets have exactly the same support, and filtering out the less interesting ones using both confidence and lift. *Lift*, in particular, is defined as the ratio between the confidence of a rule and the global support of its consequent; it assesses the degree to which the occurrence of the antecedent "lifts" the occurrence of the consequent, that is, how much they are positively correlated (i.e., lift is greater than one) or independent [9]; it has an immediate modal and temporal counterpart, given that its only based on support.

In order to keep the amount of generated rules below a reasonable number, we limit the length of each frequent itemset, as well as the length of both the consequent and antecedent of each rule, and impose the latter to be anchored.

Experiment 1: NATOPS. Each NATOPS instance is a time series of 51 timestamps, representing the x, y, z coordinates of sensors placed on various body parts of subjects performing aircraft handling signals. Since the time between two consecutive timestamps is approximately five hundredths of a second, we decided to extract the top 5 motifs with length 10 and the top 3 with length 20, in order to capture qualitatively appreciable patterns; shorter subsequences would bring little informativeness, while longer one would be too coarse.

NATOPS signals are standardized in the Naval Air Training and Operating Procedures Standardization (NATOPS) manual. The dataset includes several classes, but we choose focus on three for simplicity, namely *I have command*, *Not clear* and *Locked wings*, whose typical signals are depicted in Figure 3.

In Tab. 2, rules are encoded as follows. First, each literal is formatted in a compact manner, omitting any reference to motifs and thresholds, as well as the distance function: x, y, z indicates coordinates, with subscripts indicating body parts (r is right, l is left, h is

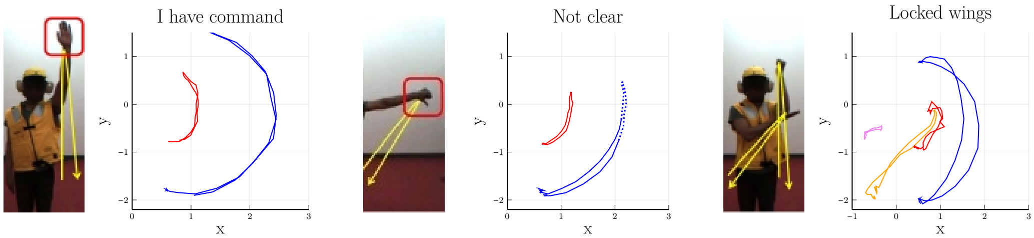

Figure 3 Three typical movements of NATOPS, namely *I have command*, *Not clear* and *Locked wings*. The x-axis encodes left (low values) and right positions for each considered body part of the operator, while the y-axis encodes down (low values) and high; the z-axis is not represented for simplicity. In the leftmost figure, the operator traces the blue arc with his right hand, starting from the rest position, up until stretching the arm above the head, while the right elbow follows as indicated by the red signal. In the middle figure, both right hand and elbow follows a movement similar to that just described, but the right hand stops at the height of the shoulder before returning to rest position; during the part of the movement highlighted with the blue dots, the thumb is oriented toward the floor. In the rightmost figure, the blue and red signals are similar to those in the leftmost figure, but the left arm, in orange, pivots on the left elbow, in violet, to touch the right elbow with the left hand.

Rule	Target class	Measures
$Y_{re}^{up} \wedge Z_{rh}^{front} \Rightarrow \langle O\rangle x_{rh}^{left\&invert} \wedge \langle B\rangle y_{re}^{rest\&up}$	I have command	$c = 1.0,\ l = 6.70$
$Z_{re}^{front} \wedge \langle A\rangle Z_{rh}^{retract} \Rightarrow \langle A\rangle Z_{re}^{retract}$	I have command	$c = 1.0,\ l = 4.0$
$Y_{re}^{up\&down} \Rightarrow \langle D\rangle x_{rh}^{rest\&right} \wedge \langle E\rangle y_{re}^{down}$	I have command	$c = 0.48,\ l = 3.22$
$X_{rh}^{right} \wedge X_{re}^{right} \Rightarrow \langle B\rangle y_{rt}^{rest\&down}$	Not clear	$c = 0.81,\ l = 5.40$
$Y_{rt}^{down} \wedge \langle A\rangle x_{rh}^{left} \Rightarrow \langle O\rangle y_{rh}^{rest} \wedge z_{re}^{front}$	Not clear	$c = 0.60,\ l = 3.27$
$Y_{le}^{up} \wedge X_{re}^{right} \wedge \langle D\rangle y_{rh}^{up} \Rightarrow Y_{lh}^{up} \wedge \langle D\rangle z_{le}^{retract}$	Lock wings	$c = 0.86,\ l = 4.29$
$Z_{lh}^{rest} \wedge \langle D\rangle y_{lh}^{down} \Rightarrow \langle E\rangle x_{lh}^{left}$	Lock wings	$c = 1.0,\ l = 3.75$

Table 2 Association rules extracted from NATOPS dataset, their associated class and the value of confidence and lift meaningfulness measures. Each literal is formatted in a compact manner, omitting any reference to motifs and thresholds, as well as the distance function. First, x, y, z indicates coordinates, with subscripts indicating body parts (r: right, l: left, h: hand, e: elbow, t: thumb) and where superscripts give an intuition about the movement captured by the motif underlying the literal. Second, an uppercase literal denotes a longer movement (20 time unites, approximately 1 second), and a lowercase one denotes a shorter one (10).

hand, e is elbow, and t is thumb), and superscripts give an intuition about the movement captured by the motif underlying the literal. Second, uppercase (resp., lowercase) coordinates denote the length of the underlying motif: uppercase indicates 20 time units (approximately 1 second), and lowercase indicates 10 time units (approximately $\frac{1}{2}$ of a second).

Let us the most interesting association rules in Tab. 2, which show high confidence and lift and would not probably be naturally deducible from an inattentive high-level description of the corresponding movement. The first rule of the class *I have command* can be read *whenever the right hand of the operator is completely stretching in front of him/her and their*

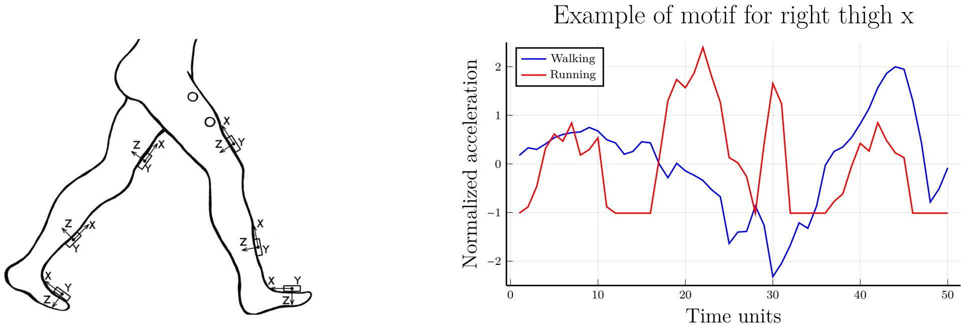

Figure 4 (Left) Location of sensors in HuGaDB. (Right) Two normalized motifs of length 50, referring to the right thigh x-axis and extracted from Walking and Running classes. In the former class, the thigh accelerates up and down more gently w.r.t. Running, where the signal is steeper.

elbow goes all the way up on the y-axis, the same elbow started the movement range in a rest position and, near the end of the movement range, the operator's right hand is moving to the left, but will soon change direction. The first rule of the *Not clear* class can be translated to *when both the right hand and the right elbow of the operator are completely stretching to the right, the thumb starts being relaxed but it is rotated downward after a couple of tenths of a second*; the second rule of the same class, instead, describes the descending part of the movement. Finally, regarding the class *Lock wings*, its first rule tells us that *if the right hand completes its upward movement during the (slower) upward movement of the left elbow and the opening of the right elbow to the right, then the left hand is going up too and, while doing so, the left elbow slightly retracts on the z-axis w.r.t. the hip*; the second rule tells us that when the left elbow aligns with the hip while the left hand is going down, then the right hand is moving to the left in the last half second of the movement.

Experiment 2: HuGaDB Data in the Database for Human Gait Analysis consists of combined activities performed by various performers and recorded continually; for instance, a participant might have walked, then ran and finally sat down. Data is collected from a body sensor network of six wearable inertial sensors (accelerometers and gyroscopes) located on the right and left thighs, shins and feet, and two EMG sensors placed on both quadriceps to measure muscle activity, as shown in Fig. 4 (left-hand side). Different activities are segmented through different labels and, among all the possible gait segments, we focus on *Walking* and *Running* gaits.

We manipulated the dataset to obtain a set of instances where each time series has 100 timestamps and the variables are limited to the x and z axis for feet and thighs; operationally, the y axis can be ignored, considering our target gaits. Upon observing of the signals, we decided to extract the top 3 representative motifs of lengths 25 and 50 time units. An example of two extracted motifs of length 50, both referring to the right thigh in the two different gaits, are depicted in Fig. 4 (right-hand side): the motif extracted from *Walking* gait consists of a gentle acceleration up and down, while the motif for *Running* gait is faster and steeper.

Similarly to the case of NATOPS, we summarize some of the most promising association rules in Table 3. The first two rules concern the class *Walking*, and states that *when the right foot accelerates forward (backwards), then the left (right) thigh accelerates upward immediately after.* Note that, since confidence is not so high, these rules may depend on the personal gait of a performer. For example, both rules do not hold if the walk is performed by keeping the

Rule	Target class	Measures
$x_{rf}^{front} \Rightarrow \langle A \rangle x_{lt}^{up}$	Walking	$c = 0.43,\ l = 3.21$
$x_{lf}^{back} \Rightarrow \langle A \rangle x_{rt}^{up}$	Walking	$c = 0.33,\ l = 2.50$
$X_{rf}^{front} \Rightarrow \langle D \rangle x_{rt}^{up} \wedge \langle D \rangle x_{rt}^{down}$	Walking	$c = 1.0,\ l = 1.67$
$X_{lf}^{front} \Rightarrow \langle D \rangle x_{lt}^{up} \wedge \langle D \rangle x_{lt}^{down}$	Walking	$c = 1.0,\ l = 1.50$
$Z_{rt}^{fbx2} \wedge Z_{rf}^{up\&down} \wedge \langle D \rangle z_{lf}^{down\&stop} \Rightarrow \langle D \rangle x_{rf}^{fast\_front}$	Running	$c = 1.0,\ l = 7.0$
$X_{lf}^{bfx2} \wedge \langle D \rangle x_{lt}^{down\&up\&down} \Rightarrow Z_{lt}^{bfx2} \wedge X_{rf}^{fbx2}$	Running	$c = 0.67,\ l = 2.33$
$x_{lf}^{back\&front} \Rightarrow \langle A \rangle z_{rt}^{front\&behind}$	Running	$c = 0.25,\ l = 1.75$

Table 3 Association rules extracted from HuGaDB dataset, their associated class and the value of confidence and lift meaningfulness measures. As before, literals codify their meaning in a compact manner. First, x, z indicates the axis, with subscripts indicating body parts (r: right, l: left, f: foot, t: thigh). Second, superscripts suggest the movement, and in particular, $fbx2$ indicates the double repetition of a sudden frontal acceleration followed by a strong backward acceleration (and similarly for $bfx2$). As before, uppercase literals denote longer motifs (50 time units), and lowercase ones shorter motifs (25).

toes low. This symmetry emerges also in the next two rules, which could be rewritten as *while the right (left) foot accelerates forward, the right (left) thigh accelerates up and down at a certain point.*

In the case of the class *Running*, the first rule shows very high confidence and lift, thus resulting independent of the personal idiosyncrasies of each candidate. It could be translated to *if the right thigh springs off forward and backward two times, the right foot goes up and down and, at a certain point, the left foot accelerates downward and then stays still for approximately 10 time units, then it means that the right foot considerably accelerates forward at a certain point.* The second rule highlights the complementarity between the left leg, iteratively accelerating backwards and forward, and the right foot accelerates in opposite directions. The last rule describes a trait for a particular kind of run, that is, a light ride in which the right foot waits for the left foot, instead of moving complementarily at the same time.

5 Conclusions

In this paper, we addressed the limitations of existing temporal symbolic learning methods for rule extraction by introducing a motif-based approach to temporal alphabet definition. By leveraging motifs-based frequently recurring patterns in time series, we proposed a more interpretable and structurally robust framework for mining temporal association rules. Our method resolves the biases introduced by naive alphabet definitions and enhances the semantic clarity of extracted rules. We formalized the concept of anchored itemsets and introduced a novel definition of motif-based local and global support, ensuring that the mined patterns are both meaningful and computationally tractable. Experimental validation on temporal datasets demonstrates the expressiveness and interpretability of the extracted rules, showing promise for applications in explainable temporal data analysis.

References

1 R. Agrawal and R. Srikant. Fast Algorithms for Mining Association Rules in Large Databases. In *Proceedings of 20th International Conference on Very Large Data Bases (VLDB)*, pages 487–499, 1994.

2 A. Balasubramanian, J. Wang, and B. Prabhakaran. Discovering multidimensional motifs in physiological signals for personalized healthcare. *IEEE Journal of Selected Topics in Signal Processing*, 10(5):832–841, 2016.

3 D. Bresolin, E. Cominato, S. Gnani, E. Muñoz-Velasco, and G. Sciavicco. Extracting interval temporal logic rules: A first approach. In *Proc. of the 25th International Symposium on Temporal Representation and Reasoning (TIME)*, volume 120 of *LIPIcs*, pages 1 – 15. Dagstuhl, 2018.

4 Roman Chereshnev and Attila Kertész-Farkas. Hugadb: Human gait database for activity recognition from wearable inertial sensor networks. In *Analysis of Images, Social Networks and Texts - 6th International Conference, AIST 2017, Moscow, Russia, July 27-29, 2017, Revised Selected Papers*, pages 131–141. Springer, 2017.

5 B.Y. Chiu, E.J. Keogh, and S. Lonardi. Probabilistic discovery of time series motifs. In *Proc. of the 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 493–498. ACM, 2003.

6 Ben D Fulcher and Nick S Jones. hctsa: A computational framework for automated time-series phenotyping using massive feature extraction. *Cell systems*, 5(5):527–531, 2017.

7 J. Fürnkranz, D. Gamberger, and N. Lavrac. *Foundations of Rule Learning*. Cognitive Technologies. Springer, 2012.

8 J. Han, J. Pei, and Y. Yin. Mining Frequent Patterns without Candidate Generation. In *Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data*, pages 1–12, 2000.

9 Jiawei Han, Micheline Kamber, and Jian Pei. *Data Mining: Concepts and Techniques, 3rd edition*. Morgan Kaufmann, 2011.

10 G.Z. Hertz and Gary Stormo. Identifying dna and protein patterns with statistically significant alignments of multiple sequences. *Bioinformatics*, 15:563–577, 1999.

11 S. Imani, F. Madrid, W. Ding, S. Crouter, and E. Keogh. Matrix profile xiii: Time series snippets: A new primitive for time series data mining. In *Proc. of the IEEE international conference on big knowledge (ICBK)*, pages 382–389. IEEE, 2018.

12 Y. Li, L. Hou U, M. Lung Yiu, and Z. Gong. Quick-motif: An efficient and scalable framework for exact motif discovery. In *Proc of the 31st IEEE International Conference on Data Engineering (ICDE)*, pages 579–590. IEEE, 2015.

13 Carl H Lubba, Sarab S Sethi, Philip Knaute, Simon R Schultz, Ben D Fulcher, and Nick S Jones. catch22: Canonical time-series characteristics: Selected through highly comparative time-series analysis. *Data mining and knowledge discovery*, 33(6):1821–1852, 2019.

14 Federico Manzella, Giovanni Pagliarini, Guido Sciavicco, and Ionel Eduard Stan. Interval temporal random forests with an application to COVID-19 diagnosis. In Carlo Combi, Johann Eder, and Mark Reynolds, editors, *28th International Symposium on Temporal Representation and Reasoning, TIME 2021, September 27-29, 2021, Klagenfurt, Austria*, volume 206 of *LIPIcs*, pages 7:1–7:18. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2021. URL: https://doi.org/10.4230/LIPIcs.TIME.2021.7, `doi:10.4230/LIPICS.TIME.2021.7`.

15 M. Milella, G. Pagliarini, G. Sciavicco, and I.E Stan. Modalfp-growth: Efficient extraction of modal association rules from non-tabular data. In *Proc. of the 25th Italian Conference on Theoretical Computer Science (ICTCS)*, volume 3811 of *CEUR*, pages 241 – 254. CEUR-WS.org, 2024.

16 C. O'Neil. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown Publishing Group, 2016.

17 P. Schäfer and U. Leser. Motiflets - simple and accurate detection of motifs in time series. In *Proc. of the 48th International Conference on Very Large Data Bases Vancouver*, volume 16, pages 725–737, 2022.

18 Guido Sciavicco and Ionel Eduard Stan. Knowledge extraction with interval temporal logic decision trees. In Emilio Muñoz-Velasco, Ana Ozaki, and Martin Theobald, editors, *27th International Symposium on Temporal Representation and Reasoning, TIME 2020, September 23-25, 2020, Bozen-Bolzano, Italy*, volume 178 of *LIPIcs*, pages 9:1–9:16. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2020. URL: https://doi.org/10.4230/LIPIcs.TIME.2020.9, doi:10.4230/LIPICS.TIME.2020.9.

19 Y. Song, D. Demirdjian, and R. Davis. Tracking body and hands for gesture recognition: NATOPS aircraft handling signals database. In *Proc. of the 9th IEEE International Conference on Automatic Face and Gesture Recognition (FG)*, pages 500 – 506, 2011.

20 I.E Stan, G. Sciavicco, E. Muñoz-Velasco, G. Pagliarini, M. Milella, and A. Paradiso. On modal logic association rule mining. In *Proc. of the 23rd Italian Conference on Theoretical Computer Science (ICTCS)*, volume 3284 of *CEUR*, pages 53 – 65. CEUR-WS.org, 2022.

21 Y. Tanaka, K. Iwamoto, and K. Uehara. Discovery of time-series motif from multi-dimensional data based on MDL principle. *Machine Learning*, 58(2-3):269–300, 2005.

22 A. Vahdatpour, N. Amini, and M. Sarrafzadeh. Toward unsupervised activity discovery using multi-dimensional motif detection in time series. In *Proc. of the 21st International Joint Conference on Artificial Intelligence (IJCAI)*, pages 1261–1266, 2009.

23 L. G. Valiant. The Complexity of Enumeration and Reliability Problems. *SIAM Journal on Computing*, 8(3):410–421, 1979.

24 G. Yang. The complexity of mining maximal frequent itemsets and maximal frequent patterns. In *Proceedings of the 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, pages 344–353, 2004.

25 Y. Zhu, Z. Zimmerman, N.S. Senobari, C-C.M. Yeh, G.J. Funning, A. Mueen, P. Brisk, and E.J. Keogh. Matrix profile II: exploiting a novel algorithm and GPUs to break the one hundred million barrier for time series motifs and joins. In *Proc. of the 16th International Conference on Data Mining (ICDM)*, pages 739–748. IEEE, 2016.