A Organ Systems' Components

Table 1: Physiological and biological biomarkers recorded as individual features, and the criteria resulting in the corresponding organ dysfunction

Organ Dysfunction	Response	Abbreviation	Failure criteria
Cardiovascular	Systolic blood pressure (SBP)	Sb	<90 mmHg
	SBP_{max}* - Systolic BP	Sd	>40 mmHg within an 8-hour period
	Mean arterial pressure (MAP)	Mp	<65 mmHg
Renal	Creatinine	Cr	>1.2 mg/dL
	(Creatinine - C_{base}**)/(C_{base})	Cd	>50% from initial creatinine
	Blood Urea Nitrogen (BUN)	Bu	>20 mg/dL
Hematopoietic	WBC	W	<4,000 cells/mL
	Platelet	Pl	<100,000 cells/mL
Metabolic	Lactate	La	>2.0 mmol/L
Gastrointestinal	Bilirubin	Bi	>2 mg/dL
Respiratory	Fraction of inspired oxygen (FiO_2)	Fi	>21%
	Pulse oximetry (SpO_2)	Px	<90%
	SpO_2/FiO_2	Or	<421
	Oxygen (O_2) Source	Os	Mechanical ventilation required (bilevel positive airway pressure (BiPAP) or continuous positive airway pressure (CPAP) or ventilator)
Central Nervous	Glasgow Comma Score	Gc	<14
	Glasgow Best Verbal Response	Gv	<5
Biomarkers	Procalcitonin	pc	>0.15 ng/mL
	C-Reactive Protein	cr	>8 mg/L
	Erythrocyte Sedimentation Rate	sr	>20 mm/hr

*: Maximum systolic blood pressure for each observation within 8-hour windows.
**: Initial creatinine value observed in each visit.

Table 2: The performance [mean (standard deviation)] of algorithms adopted for identification of sepsis patients in the four data sets of the study. The features used are statistics of individual responses (Ind) and their combinations with age, gender, and medical history of patients (Ind^+). ADA: AdaBoost, DT: Decision tree, NB: Naïve Bayes, LG: Logistic regression, kNN: k-nearest neighbors, RF: Random forest, SVM: Support vector machine, SVM^p: Support vector machine with polynomial kernel function.

Algorithm	gap (hours)	Features	Accuracy	Recall	Precision
ADA	3	Ind	0.71 (0.02)	0.68 (0.03)	0.69 (0.03)
		Ind^+	0.75 (0.02)	0.73 (0.02)	0.72 (0.03)
	6	Ind	0.70 (0.02)	0.67 (0.03)	0.68 (0.03)
		Ind^+	0.73 (0.02)	0.72 (0.02)	0.70 (0.03)
	12	Ind	0.69 (0.02)	0.66 (0.03)	0.69 (0.03)
		Ind^+	0.73 (0.02)	0.71 (0.03)	0.72 (0.03)
	24	Ind	0.67 (0.02)	0.64 (0.03)	0.67 (0.04)
		Ind^+	0.71 (0.02)	0.69 (0.04)	0.71 (0.03)
DT	3	Ind	0.65 (0.02)	0.65 (0.03)	0.60 (0.02)
		Ind^+	0.67 (0.02)	0.66 (0.03)	0.63 (0.03)
	6	Ind	0.63 (0.02)	0.63 (0.03)	0.60 (0.03)
		Ind^+	0.65 (0.02)	0.65 (0.03)	0.62 (0.03)
	12	Ind	0.63 (0.03)	0.63 (0.04)	0.61 (0.03)
		Ind^+	0.64 (0.02)	0.64 (0.04)	0.62 (0.03)
	24	Ind	0.60 (0.02)	0.60 (0.04)	0.59 (0.04)
		Ind^+	0.63 (0.02)	0.63 (0.03)	0.62 (0.03)
NB	3	Ind	0.66 (0.02)	0.42 (0.03)	0.72 (0.03)
		Ind^+	0.69 (0.02)	0.53 (0.03)	0.72 (0.03)
	6	Ind	0.64 (0.02)	0.38 (0.03)	0.72 (0.04)
		Ind^+	0.67 (0.02)	0.51 (0.07)	0.72 (0.04)
	12	Ind	0.61 (0.02)	0.33 (0.05)	0.72 (0.04)
		Ind^+	0.66 (0.03)	0.45 (0.05)	0.73 (0.05)
	24	Ind	0.61 (0.02)	0.37 (0.04)	0.69 (0.05)
		Ind^+	0.65 (0.03)	0.49 (0.03)	0.70 (0.04)
kNN	3	Ind	0.67 (0.02)	0.60 (0.03)	0.65 (0.03)
		Ind^+	0.70 (0.02)	0.58 (0.03)	0.71 (0.02)
	6	Ind	0.66 (0.02)	0.61 (0.03)	0.64 (0.03)
		Ind^+	0.68 (0.02)	0.59 (0.03)	0.69 (0.03)
	12	Ind	0.65 (0.02)	0.60 (0.04)	0.65 (0.03)
		Ind^+	0.68 (0.02)	0.56 (0.03)	0.70 (0.03)
	24	Ind	0.62 (0.03)	0.57 (0.04)	0.62 (0.04)
		Ind^+	0.65 (0.02)	0.54 (0.03)	0.68 (0.03)
LG	3	Ind	0.72 (0.02)	0.66 (0.03)	0.70 (0.03)
		Ind^+	0.75 (0.02)	0.73 (0.02)	0.72 (0.03)
	6	Ind	0.70 (0.02)	0.65 (0.03)	0.68 (0.03)
		Ind^+	0.74 (0.02)	0.72 (0.02)	0.72 (0.03)
	12	Ind	0.70 (0.02)	0.65 (0.03)	0.70 (0.03)
		Ind^+	0.74 (0.02)	0.72 (0.02)	0.73 (0.03)
	24	Ind	0.67 (0.02)	0.61 (0.03)	0.68 (0.03)
		Ind^+	0.72 (0.02)	0.70 (0.03)	0.71 (0.03)
RF	3	Ind	0.73 (0.02)	0.72 (0.03)	0.69 (0.03)
		Ind^+	0.76 (0.02)	0.78 (0.03)	0.71 (0.03)
	6	Ind	0.71 (0.01)	0.71 (0.03)	0.69 (0.02)
		Ind^+	0.74 (0.02)	0.76 (0.02)	0.71 (0.03)
	12	Ind	0.72 (0.02)	0.71 (0.03)	0.71 (0.03)
		Ind^+	0.75 (0.02)	0.76 (0.03)	0.72 (0.03)
	24	Ind	0.69 (0.02)	0.68 (0.03)	0.69 (0.04)
		Ind^+	0.73 (0.03)	0.74 (0.04)	0.71 (0.03)
SVM	3	Ind	0.72 (0.02)	0.69 (0.03)	0.70 (0.03)
		Ind^+	0.76 (0.02)	0.77 (0.02)	0.72 (0.02)
	6	Ind	0.71 (0.02)	0.68 (0.03)	0.69 (0.03)
		Ind^+	0.75 (0.02)	0.76 (0.02)	0.71 (0.03)
	12	Ind	0.71 (0.02)	0.69 (0.03)	0.70 (0.03)
		Ind^+	0.75 (0.02)	0.76 (0.02)	0.72 (0.03)
	24	Ind	0.68 (0.02)	0.65 (0.03)	0.68 (0.04)
		Ind^+	0.73 (0.02)	0.74 (0.04)	0.71 (0.03)
SVM^p	3	Ind	0.66 (0.02)	0.41 (0.03)	0.74 (0.04)
		Ind^+	0.71 (0.02)	0.52 (0.03)	0.77 (0.03)
	6	Ind	0.65 (0.02)	0.41 (0.03)	0.74 (0.04)
		Ind^+	0.70 (0.02)	0.53 (0.03)	0.77 (0.04)
	12	Ind	0.64 (0.02)	0.39 (0.03)	0.74 (0.04)
		Ind^+	0.68 (0.02)	0.48 (0.03)	0.77 (0.04)
	24	Ind	0.62 (0.02)	0.37 (0.03)	0.72 (0.05)
		Ind^+	0.66 (0.02)	0.47 (0.04)	0.75 (0.04)

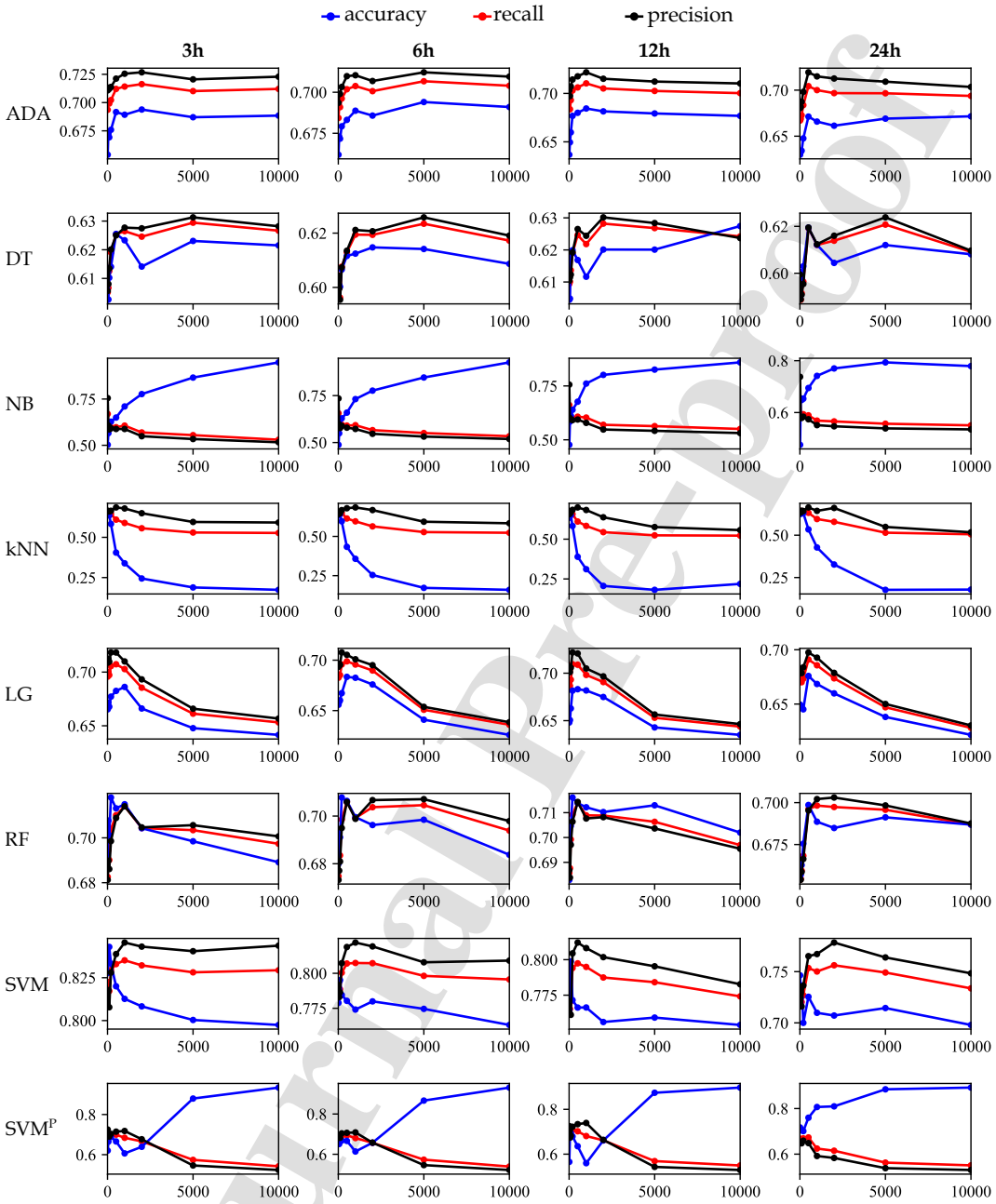

Figure 1: The performance of algorithms adopted for identification of sepsis patients in the four data sets of the study. The features used are statistics of individual responses, patients' age, gender, and medical history, frequent temporal patterns, and frequent evolving patterns. ADA: AdaBoost, DT: Decision tree, NB: Naïve Bayes, LG: Logistic regression, kNN: k-nearest neighbors, RF: Random forest, SVM: Support vector machine, SVM^{P}: Support vector machine with polynomial kernel function.

B Methods - Further Details

B.1 Frequent Temporal Patterns

It is very well-known that mining frequent patterns in a data set of networks requires verification of graph and subgraph isomorphism. Two networks $N_1 = (V_1, E_1)$ and $N_2 = (V_2, E_2)$ are isomorphic if there is a bijective function $\mathcal{I}$ such that:

$$\{v_i, v_j\} \in E_1 \iff \{\mathcal{I}(v_i), \mathcal{I}(v_j)\} \in E_2 \tag{2}$$

Also, for a given $N_1 = (V_1, E_1)$, $N_2 = (V_1, E_1)$ would be a subgraph of N_1 if and only if:

$$N_2 \subseteq N_1 \iff V_2 \subseteq V_1 \wedge E_2 \subseteq E_1 \tag{3}$$

The subgraph isomorphism test or verification is defined as given two networks or *CIG*s, e.g., $C_1 = (V_1, E_1)$ and $C_2 = (V_2, E_2)$, whether we can find a subgraph of C_1 isomorphic to C_2. Both graph and subgraph isomorphism problems are computationally expensive. We use the lexicographic ordering approach proposed in [34] to minimize the negative impact of the graph isomorphism problem. Also, the occurrence lists are used to avoid the direct evaluation of the subgraph isomorphism problem by only investigating the *CIG*s supporting one candidate's parent. The pseudocode of mining frequent temporal patterns is provided in Algorithm 1. Note that in this algorithm, first, a data set of *CIG*s, DS^*, associated with temporal networks in *DS* is created. Then, DS^* is sent to Algorithm 2 in which the frequent temporal patterns are detected and returned.

B.2 Frequent Evolving Patterns

The fundamental events considered in this study are defined as follows:

- **Birth ($\circ\!\!-$):** A birth event is defined as the appearance of a pattern of physiological and biological biomarkers. The responses contributing to the birth events should all appear in the user-defined connectivity radius τ from one another.
- **Expansion ($\prec$):** An expansion event is defined as the continuous growth of the pattern of physiological and biological biomarkers. The number of responses contributing to an expansion event be more than the user-defined size threshold σ. Also, expansion should happen gradually, defined by both the σ and a user-defined continuity radius parameter δ. In other words, the growth should happen in a window larger than continuity radius δ. Over the expansion window, the size of the pattern increases monotonically.
- **Merge ($\vdash$):** A merge event is characterized by joining two connected patterns (each with a size more than σ) to form one connected component. The patterns contributing to the merge event are initially disconnected.
- **Contraction ($\succ$):** A contraction is a continuous reduction in the size of the pattern. Similar to the expansion event, the number of responses removed from the pattern should be more than a user-defined size threshold σ. Their removal should be continuous, spanning over a window longer than a user-defined continuity radius parameter δ. In the window over which the contraction event happens, no growth is observed.
- **Split ($\dashv$):** A split event is defined as transforming one connected component into two or more connected components. Each connected component should be larger than the user-defined size threshold σ and become disconnected after the split event.
- **Death ($-\!\!\circ$):** A death event is defined as the disappearance of physiological and biological biomarkers' measurements from a patient's record. The responses contributing to the death event are all in a user-defined connectivity radius τ from one another.

B.3 Missing-Data Imputation

Due to missingness in available data, some responses may not be observed for patients in sepsis and non-sepsis subpopulations. Therefore, in prediction models including statistics of individual responses, some of the values might be missing. As most machine learning algorithms assume that the input data is complete, we must impute the missing values. For this purpose, we use a similarity-based missing data imputation method similar to the one proposed in [48]. For any patient with a missing value, we find the most similar patient with a known value for the corresponding response to the patient of interest and

impute the missing value with the known value of the other patient. The patients' similarities are computed based on the cosine similarity of each pair of patients' known values.

C Results - Further Discussion

The number of frequent patterns, $|s|$, detected for the other two isomorphism definitions and their sizes, $|E|$, for different values of frequency thresholds of f are shown in Figure 2 for the four data sets. Also, Figure 3 shows the number of frequent patterns detected for the four different noise tolerance duration considered for the two different definitions of isomorphism for DS^3.

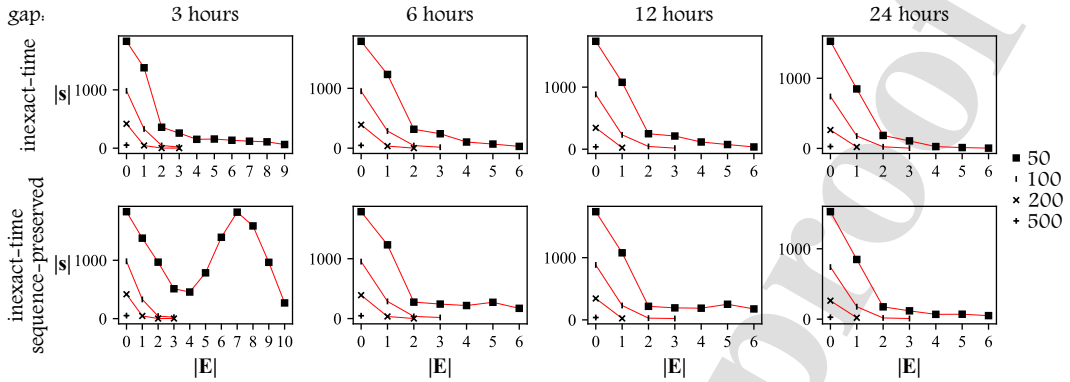

Figure 2: The number $|s|$ and size $|E|$ of frequent temporal patterns detected from the four data sets DS^3, DS^6, DS^{12}, and DS^{24} for two isomorphism definitions at different support threshold.

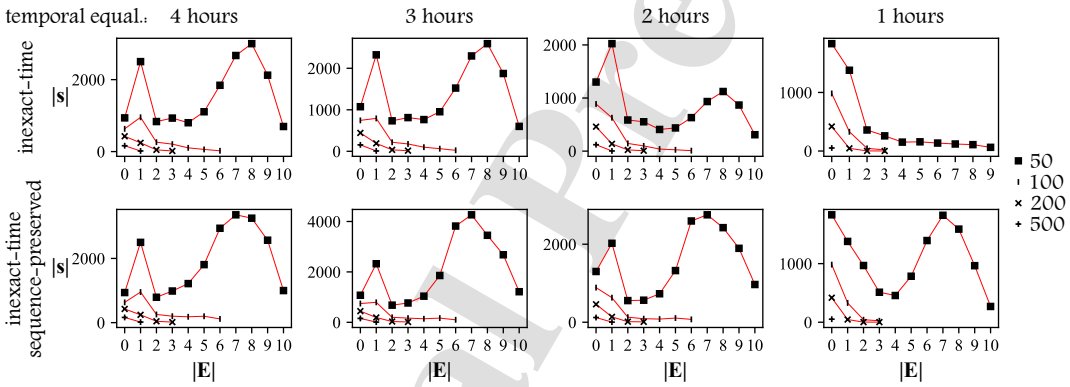

Figure 3: The number $|s|$ and size $|E|$ of frequent temporal patterns detected from the data sets with 3-hour window, DS^3, for different temporal noise tolerance thresholds and two isomorphism definitions at different support threshold.

The vertices corresponding to 17 numeric responses are labeled with $\ell = (\mu, M, \sigma, R)$. However, considering that some of these responses are precisely measured by sensors, we also add some noise tolerance in the labeling approach. For this purpose, we discretize each of μ, M, R, and σ measurements at different intervals and then create the labels. The number of intervals considered for statistics discretization is 2, 5, 10, and 20. In other words, for example, for a response r and a discretization parameter $d = 5$, using the values of r for all the patients in the data set, five equal-width intervals are created. The intervals are ordered and labeled from 1 to 5. Then, we replace the r for each patient and interaction window with the corresponding interval's label instead of using the exact value of r. The same approach is applied to the M, R, and σ. Please note, as the number of intervals goes to ∞, the labels close to the actual measurements. However, for large values of labeling discretization parameters, the frequent patterns become rare due to the sensor-based measurement precision. The results of the implementation of frequent temporal mining over patients' networks for different values of labeling categories for DS^3 are shown in Figure 4.

For the two inexact-time (i) and inexact-time sequence-preserved (is) versions of the isomorphism, the algorithm could find many frequent patterns with a significantly large number of patterns for the latter definition. Figure 2 shows the number of frequent temporal patterns identified for the four data sets for different frequency thresholds and the two inexact-time (i) and inexact-time sequence-preserved (is)

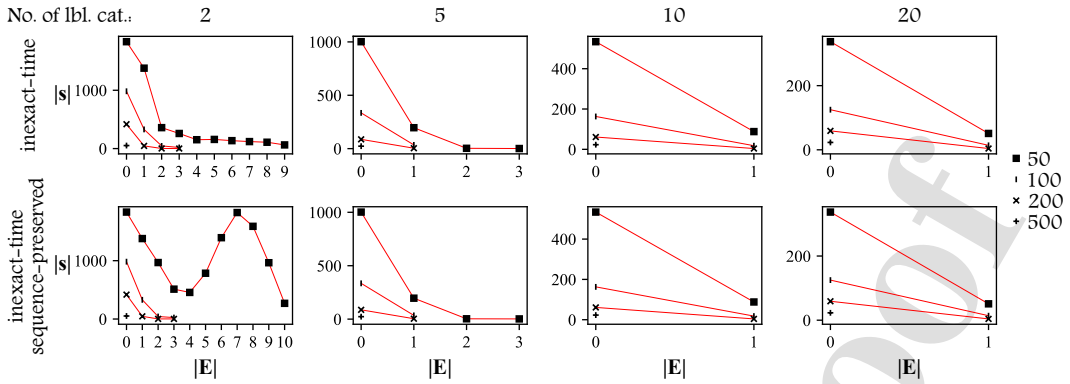

Figure 4: The number $|s|$ and size $|E|$ of frequent temporal patterns detected from the data sets with 3-hour window, DS^3, for different numbers of labeling categories and two isomorphism definitions at different support threshold.

versions of the isomorphism. This figure shows that the number of patterns is higher for lower gap values (e.g., 3 hours before sepsis onset). It can be attributed to two facts. First, the number of patients in DS^3 is higher, and it decreases as the length of the window before sepsis onset increases. Secondly, as we reduce the gap before the sepsis onset, we have a wider window of hospitalization records of patients. Therefore, it is more probable that the patterns infrequent in, for example, DS^{12} become frequent in the same set of patients in DS^3. Also, as we increase the frequency threshold, the number of frequent patterns consistently decreases. Note that frequent patterns composed only of vertices are shown with edges of size zero in this figure.

For both *i* and *is* isomorphism definitions, we needed to determine a noise tolerance threshold. This threshold is used to consider two responses equal if they have identical response values but close enough unequal duration. We consider four different thresholds, from 1 hour to 4 hours. Figure 3 shows that for both *i* and *is* isomorphism definitions increasing this threshold consistently increases the number of patterns detected.

Similar to the temporal duration of measurements that it would be rare to identify patients with identical duration for the same responses, the results showed that it is also rare to find a group of responses, all with the same values, to be in common in many patients. However, when we categorized the responses' measurements into multiple categories, we could identify larger frequent patterns as shown in Figure 4. Note that a larger number of categories implies values closer to the actual values, while a smaller number of labeling categories allows for more noise tolerance in the responses' measurements.

The evolving patterns characterize evolution events for different temporal patterns over time. Figure 5 shows that for wider gaps before sepsis onset, in general, we have a larger number of frequent evolving patterns (with some exceptions between DS^6 and DS^{12}). Increasing the min_{freq} threshold consistently reduces the number of evolving patterns. Also, we observed that the increase of size parameter σ reduces the number of evolving patterns. However, the connectivity and continuity radii might impact the numbers of frequent evolving patterns differently. The impacts of these two parameters on the number of patterns depend on the topological structure of physiological and biological biomarkers and their temporal relationships with one another.

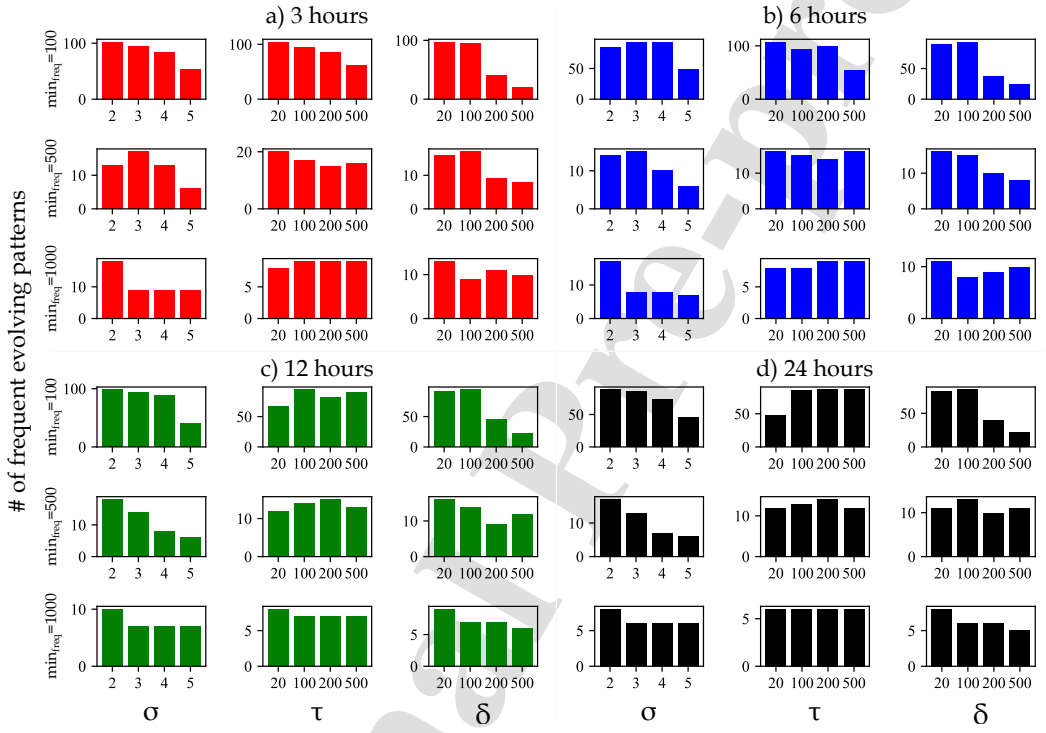

Figure 5: The number of frequent evolving patterns identified in the four data sets of the study for different values of parameters σ, τ, δ, and min_{freq}.