Local Exceptionality Detection in Time Series Using Subgroup Discovery: An Approach Exemplified on Team Interaction Data

Dan Hudson[1,2](✉), Travis J. Wiltshire[2], and Martin Atzmueller[1]

[1] Semantic Information Systems Group, Osnabrück University, Osnabrück, Germany
{daniel.dominic.hudson,martin.atzmueller}@uni-osnabrueck.de
[2] Department of Cognitive Science and AI, Tilburg University, Tilburg, The Netherlands
t.j.wiltshire@tilburguniversity.edu

Abstract. In this paper, we present a novel approach for local exceptionality detection on time series data. This method provides the ability to discover interpretable patterns in the data, which can be used to understand and predict the progression of a time series. As an exploratory approach, the results can be used to generate hypotheses about the relationships between the variables describing a specific process and its dynamics. We detail our approach in a concrete instantiation and exemplary implementation, specifically in the field of teamwork research. Using a real-world dataset of team interactions we discuss the results and showcase the presented novel analysis options. In addition, we outline possible implications of the results in terms of understanding teamwork.

Keywords: Subgroup discovery · Exceptional model mining · Time series · Teamwork research · Multimodal analysis.

1 Introduction

Methods for local exceptionality detection such as subgroup discovery [2] and its variant exceptional model mining (EMM) [8] are established knowledge discovery techniques for finding interpretable patterns. Basically, they identify patterns relating different attributes of a dataset that are interesting according to some target model, thus providing explicit and *interpretable* rules to associate descriptive properties found in the data instances. Considering time series and/or event data, the investigation of subgroup discovery has been limited, mainly focusing on aggregating/averaging time overall [14] or by considering aggregates on sets of discrete-valued events [17], compared to continuous-valued time series which we consider in this work.

C. Soares and L. Torgo (Eds.): DS 2021, LNAI 12986, pp. 435–445, 2021.
https://doi.org/10.1007/978-3-030-88942-5_34

In this paper, we present a novel approach for performing subgroup discovery and EMM on time series. We propose an extensible approach, in particular relating to feature and target construction on dynamic time series data.

Time series exceptionality analysis is a vast field, e. g., [1,12]. Here, for example, methods for change detection [1], anomaly detection also including symbolic representations [5,9,15] and time series classification are relevant, where typically global approaches are addressed, in contrast to local exceptionality detection which we focus on in this work. An approach based for compressing event logs based on the minimum description length (MDL) principle was presented by [11], making it possible to detect local patterns in temporal data, however without focusing on exceptionality. Compared to this work, which focuses on event sequences as inputs, our approach aims to find meaningful representations of (potentially complex) continuous-valued time series, and assesses the discovered patterns by reference to a target variable rather than compression.

In order to demonstrate our approach, we exemplify its application on a case study conducted in the area of social sensing, wherein team interactions are examined through a multimodal, sensor-based approach. In general, the study of teamwork looks at how groups of multiple individuals work toward a common goal [19] through collaborative team processes [10]. Recent work [18,20] has emphasised the need to understand dynamics within team processes, by embracing methodologies that record teams over time. For example, body movement, along with dynamics of speaking and turn-taking, are well understood to be important social signals used in cooperation and teamwork [16,25]. Although they can be quantified from video and audio recordings, it is difficult to establish the important relationships between these social signals in an empirical way when using multiple time-varying modalities.

Since it is not obvious how to start analysing these time-varying signals, the application of exploratory analysis methods such as subgroup discovery is well suited for such an analysis, to provide first insights and to support hypothesis generation. Our approach leads to interpretable rules which are plausible due to the use of expert knowledge in feature selection (described further in Sect. 3). We therefore choose this case study to showcase our method and to discuss the respective results, using a real-world dataset of 27 video and audio recordings of teams performing a collaborative task, taken from the ELEA corpus [21].
Our contributions are summarised as follows:

1. We present a novel methodological approach as an iterative human-guided process that makes it possible to use subgroup discovery on time series data. We discuss according feature extraction and target construction, e. g., using different time lags, for making the results predictive at different timescales.
2. We showcase our approach through a study in the context of team research. For this analysis, we also introduce a new quality function, adapting the concept of dynamic complexity from team research. In addition, we present a novel subgroup visualisation for multi-dimensional parameter analysis, i. e., the *subgroup radar plot*, also enabling user-guided subgroup assessment.

3. In our case study, we search for relationships between multimodal data, i. e., body movement and speech in time series. As evaluated by a domain specialist, this gives rise to several meaningful hypotheses that can be investigated in future work in the field of teamwork study.

2 Method

Below, we present our process model for local exceptionality detection in time series, and then we discuss the individual steps of our approach in detail.

Overview. Our proposed approach is visualised in Fig. 1 as a linear workflow, which can be executed in multiple iterations: First, the time series is split into slices, i.e., non-overlapping subsequences of equal length, so that it is possible to investigate moments when a time-varying target variable reaches an extreme value. For each slice, we extract a set of descriptive features. These are (optionally) discretised, e. g., the value of each feature can be converted into 'low', 'medium' and 'high' based on tercile boundaries across the slices. The choice of the appropriate length for a slice should be driven by the application, i. e., to include enough time points to allow a variety of features to be extracted, such as frequency components and estimates of entropy; also, it should be small enough that dynamics of the time series are not likely to change multiple times within a slice. Then, the target variable is prepared. We also propose to investigate the relationship between attributes of the time slices to the target variable at different lags, which necessitates performing the analysis with multiple copies of the dataset (with the lag applied). With a lag of zero, our process discovers subgroups that are informative about how various attributes covary with the target, e. g., for investigating how a system-level process is reflected in multiple variables. With higher lags, the process has a more predictive focus, relating to especially high/low target value at a later point.

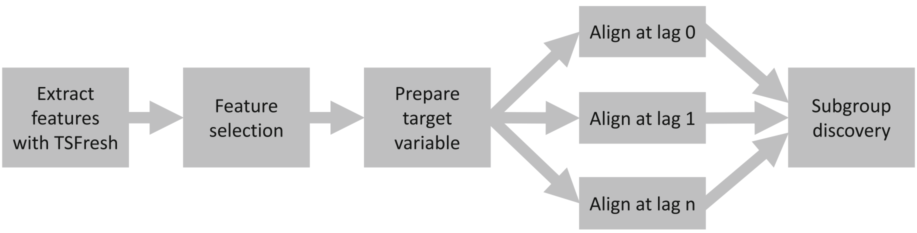


Fig. 1. The workflow of our methodological process to perform subgroup discovery on time series data. With a human-in-the-loop, the workflow can be iteratively applied.

Time Series Feature Extraction. We convert time series into 'slices' which can then individually be summarised with static rather than time-varying attributes. To obtain the necessary features, we use the TSFresh package in

Python [6], which computes a large number of features specifically to summarise time series. Examples of features computed by TSFresh are, e. g.,: (a) mean value, (b) absolute energy, (c) autocorrelation at different lags, (d) Fourier coefficients, (e) binned entropy, (f) sample entropy, (g) root mean square, etc. This approach makes it possible to perform subgroup discovery on the features extracted for each slice, while still retaining some of the variation that is observed as the original time series progresses (since different slices will correspond to different points in time in the original series).

Subgroup Discovery for Local Exceptionality Detection. Subgroup discovery aims at finding a combination of *selectors* or *selection expressions*, in a form similar to rules (e. g., `PropertyA = True` or `PropertyB > 1.5`), which function as membership criteria for a subgroup: any data points that satisfy the criteria are part of the subgroup. A *subgroup description* (or *pattern*) combines selectors into a Boolean formula. For a typical conjunctive description language, a pattern $P = \{sel_1, \ldots, sel_k\}$ is defined out of a set S of selectors $sel_j \in S$, which are interpreted as a conjunction, i.e. $p = sel_1 \wedge \ldots \wedge sel_k$. A *subgroup* corresponding to a pattern then contains all instances d of a database D, i. e., $d \in D$ for which the respective formula for the pattern evaluates to true. Specifying subgroups in this way is useful because the rules are easy to interpret and relate directly to known properties of the data points – also called *instances*.

The key question is determining which subgroups are interesting, e. g., because they have a particularly high average target value compared to the population mean, as observed for the whole dataset. The interestingness of a pattern is determined by a quality function $q\colon 2^S \rightarrow \mathbb{R}$. It maps every pattern in the search space to a real number that reflects the interestingness of a pattern (or the extension of the pattern, respectively). Many quality functions for a single target feature, e. g., in the binary or numerical case, trade off the size $n = |ext(P)|$ of a subgroup and the deviation $t_P - t_0$, where t_P is the average value in the subgroup identified by the pattern P and t_0 the average value of the target feature in the general population. Thus, standard quality functions are of the form

$$q_a(P) = n^a \cdot (t_P - t_0),\ a \in [0; 1]\,.$$

For binary target concepts, this includes, e. g., a *simplified binomial* function $q_{0.5}$ for $a = 0.5$, or the *Piatetsky-Shapiro* quality function q_1 with $a = 1$, cf. [2]. Recently, [4] described the use of quality functions which also include a term to quantify the dispersion of the target feature within the subgroup, which increases the consistency within subgroups with respect to the target feature.

3 Results: Case Study on Team Interaction Data

Below, we discuss a case study applying our approach in the context of interactive team cognition [7]. We investigate the 'dynamic complexity' of speech amongst team members (described in detail in Sect. 3), which is a method to quantify interaction dynamics that is also sensitive to moments of transition where the dynamics are changing. These moments are potentially to the benefit or the detriment of the team (we provide further discussion in [24]). Subgroup discovery provides interpretable patterns of interesting situations/events, where the dynamic complexity shows *exceptional* local deviations, indicating interesting points in the respective team interaction.

Dataset. The data used in this case study comes from the Emergent Leadership (ELEA) corpus [21], which contains recordings of groups of individuals who have been tasked to work together to rank a list of items given a disaster scenario. In particular, the task was to rank the importance of items such as 'chocolate' and 'newspapers' for the situation in which the group has been stranded in a freezing-cold winter environment. The corpus includes audio recordings from a single microphone in the centre of the room, and video recordings from webcams facing the participants. Both types of recording are available for 27 groups, each consisting of 3–4 participants. Via the video recordings, we quantify body movement during the team task. See [13] for a detailed discussion on how to quantify body movement in this context. Intuitively, we relate the body movement modality to the modality of speech, using the audio recordings to quantify speech dynamics. As a target we use dynamic complexity, cf. Sect. 3 (below), to estimate speech dynamics.

Feature Selection. TSFresh is not domain-specific, and therefore extracts generic features from time series. An important step in our process is to identify which features are potentially relevant and interpretable for the application being considered. We selected a subset of 91 from the 300+ features extracted by TSFresh. At a high level, these features can be categorised as follows:

(a) Descriptive statistics (mean, variance, quantiles, standard deviations); (b) Average and variance of the changes between successive time points; (c) Measures of complexity, such as Lempel-Ziv complexity, as well as multiple forms of entropy; (d) 'Matrix profile' statistics, which can be informative about repetitive or anomalous sub-sequences of the time slices; (e) Measures based on the number of peaks or extreme points; (f) Strength of different frequency components in the signal; (g) Measures based on autocorrelation at different lags; (h) Measures based on how well the data fits a certain (e. g., linear) model.

Features were selected based on potential relevance to body movement in social interactions, through discussion with an expert in interactive team cognition. For example, when considering coefficients of the Fourier transform, we included the magnitude since a large degree of movement at a specific frequency is something that can be visibly interpreted from the video recordings of body movement, but excluded the angle (or phase at the start of the time slice) since this is hard to visually comprehend (without performing further analysis to look for, e.g., synchrony) and therefore seems unlikely to be a usable social signal.

Target Modeling – Dynamic Complexity. As the target variable for subgroup discovery, we focused on the dynamic complexity of the speech recordings. The dynamic complexity measure is used to quantify how complex the behaviour of a system is, and provides us with a way to characterise the dynamics of speech, in a manner which could potentially be useful for (e.g.) detecting moments when a phase transition between patterns of behaviour is likely [22]. It is calculated over a moving window by combining two components, called the *fluctuation* and *distribution*, which respectively correspond to: the degree to which frequent and large oscillations are observed in the window, and, the degree to which the observed values refuse to favour a particular region of the measurement scale (instead being distributed equally across the possible range of values). The product of *fluctuation* and *distribution* gives the *dynamic complexity* of the window. For a detailed discussion we refer to [22,24]. We convert the audio data, which is sampled 40,000 Hz, to a more coarse-grained dataset by computing the energy of each second of audio, which allows us to calculate dynamic complexity on a scale appropriate to interaction behaviour (a timescale of seconds to minutes).

We evaluated two variations of dynamic complexity, captured as a target attribute, which we constructed in the spirit of EMM. First, we model the dynamic complexity as a Gaussian distribution of values, and use the z-score normalised mean as the target variable to determine which subgroups are most interesting. Second, we perform a linear regression of the dynamic complexity against time as a target model, and use the resulting slope as the target attribute. In both cases, the quality function we use to rank subgroups is the simple binomial quality function ($q_{0.5}$, see above), which tends to favour smaller subgroups with a more extreme target value. To balance this, we set the minimum subgroup size to 20 so that they do not become too small to be meaningful.

Results. As stated earlier, we are able to perform subgroup discovery at different time lags, making the task predictive when using lags greater than zero, or an exploration of the relationships between variables at lag zero. First, we discuss the 0-lag results, with slices of 1 min. A selection of five subgroups is presented in Table 1. The subgroups are also visualised as subgroup radar plots in Figs. 2(a), indicating the most important quality parameters, and 3(a), additionally showing how the subgroups differ according to 5 key selector variables. In general, these presented novel subgroup visualisations (Figs. 2, 3) allow a seamless overview–zoom–detail cycle, according to the *Information Seeking Mantra* by Shneiderman [23]: Overview first (macroscopic view), browsing and zooming

Table 1. A selection of subgroups discovered using the SD-Map algorithm [3] at a lag of 0 min: subgroup pattern, a textual description, size (S) and mean z-score ($\varnothing$).

	Pattern	Description	$\|S\|$	$\varnothing$
1	`mean__change_quantiles__f_agg_"mean"__isabs_False__qh_0.6__ql_0.2=low`, AND `mean__longest_strike_below_mean=high`, AND `mean__quantile__q_0.8=low`	The value of changes around the mean (after values have been restricted to remain between the 0.2 and 0.6 quantiles) is low across the team. Team members tend to have at least one long sequence of values below the mean. The 0.8 quantile also tends to be low.	21	1.137
2	`mean__lempel_ziv_complexity__bins_100=low`, AND `mean__longest_strike_below_mean=high`, AND `mean__quantile__q_0.8=low`	The Lempel-Ziv measure of complexity is neither high nor low across the team. Team members tend to have at least one long sequence of values below the mean. The 0.8 quantile also tends to be low.	25	1.026
3	`mean__longest_strike_below_mean=high`, AND `mean__mean=low`	Team members tend to have at least one long sequence of values below the mean. The average value of movement is generally low across the team.	38	0.81

(mesoscopic analysis), and details on demand (microscopic focus) – from basic quality parameters to subgroup description and its combination. This allows a human-in-the-loop approach, which proved beneficial for our domain specialist when inspecting the results of our knowledge discovery methodology.

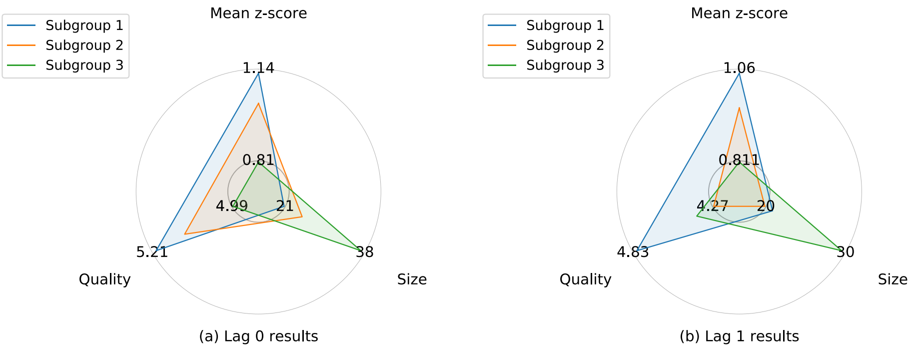
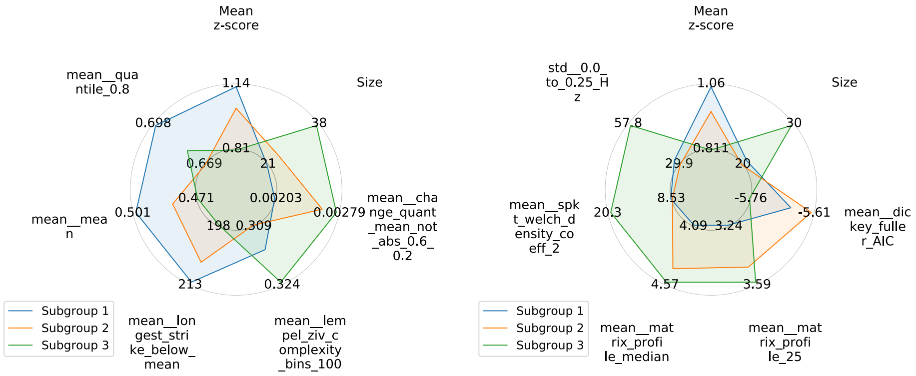


Fig. 2. Visualisations of how subgroups differ according to quality, mean z-score, and size. The results are shown at: (a) a lag of 0 min, and (b) a lag of 1 min.

Overall, our results show that it is possible to discover relatively small subgroups (size 20–40 compared to a population size of 327) whose dynamic complexity has on average a z-score of around 1, as the mean, averaged across members of the subgroup. Since the outputs are interpretable, it is possible to speculate about what they mean in the context of body movement and speech dynamics. Many of the subgroups, such as those shown in Table 1, suggest that low amounts of movement might be indicative of complex speech dynamics, particularly if there is a long sequence of low values. This perhaps suggests that while speech dynamics are becoming chaotic, the team members become more still – for example moving less as they focus more on the discussion. This is a hypothesis generated from the data which further work could seek to verify. In order to consider the impact of the window size, we also performed subgroup discovery using 30-second slices of the time series. This uncovered subgroups which often used the same rules, e.g., stating that the change around the mean between the 0.2 and 0.6 quantiles be low across the team, and, that the mean and various quantiles should also be low. There were some differences, especially that these subgroups incorporated more rules concerning variability between team members with respect to their number of values below the mean/above the mean and their Lempel-Ziv complexity, suggesting that certain types of imbalance in the teams may also help to identify moments of high dynamic complexity in speech. Overall, changing the window size in this way did not have a large impact on the discovered subgroups.

Next, we discuss results when applying a lag of 1 min, discovering ways to predict high complexity in speech dynamics from the body movement signals a minute earlier. A selection of subgroups are presented in Table 2, Fig. 2(b),



(a) Lag 0 results (b) Lag 1 results

Fig. 3. Visualisations of how subgroups differ according to the mean z-score, the size, and 5 key selector variables – with (a) a lag of 0 min, and (b) a lag of 1 min.

and Fig. 3(b), to give an idea of how exceptional the subgroups are compared to the population overall. Like with the 0-lag results, it appears to be possible to discover subgroups of around 20–40 members which have an average z-score of close to 1. The features used to define subgroups, however, are different. Some of the subgroups, such as the first two listed in Table 2, suggest that the team members might have similar, low values for the low-frequency movement components during the minute preceding a period of high dynamic complexity. Features based on 'matrix profile' statistics are used to define many subgroups. Looking at these features, it seems that high dynamic complexity can be expected following a period when the body movement signal does not have clear repetitions in its structure. This could indicate that complexity and a lack of pattern in body movement is predictive of chaotic speech dynamics shortly thereafter. This is another example of a data-driven hypothesis that future confirmatory work could verify.

Furthermore, besides the mean dynamic complexity of speech, we performed the analysis with two related target concepts. Specifically, we considered (1) the slope when conducting a linear regression of the dynamic complexity against time, and (2) the change between successive windows. The results in this case would be informative about periods when the complexity of speech dynamics is increasing. Many of these subgroups included rules that complexity of body movement (measured through Lempel-Ziv, Fourier entropy and binned entropy) is generally high across the team, and some subgroups also indicated that a large number of peaks in the slices of body movement was related to increasing dynamic complexity in speech. This could indicate that complexity in speech increases fastest while movement is already complex. From this analysis, the subgroups also suggested a relationship between increasing speech complexity and a medium-strength frequency component in body movement at 0.5–0.75 Hz, and also medium values for the mean and various quantiles. Results when using

Table 2. A selection of subgroups discovered using the SD-Map algorithm [3] at a lag of 1 min: subgroup pattern, a textual description, size (S) and mean z-score ($\varnothing$).

	Pattern	Description	$\|S\|$	$\varnothing$
1	`mean__augmented_dickey_fuller__attr_"teststat"__autolag_"AIC"=medium`, AND `mean__spkt_welch_density__coeff_2 =low`, AND `std__0.0_to_0.25_Hz=low`	The signal is neither relatively well-modelled nor relatively poorly-modelled by a process with a unit root, according to the Augmented Dickey-Fuller test. The strength of the frequency component at 0.234 Hz is low, and there is low variability of the strength of frequency components between 0.0 Hz and 0.25 Hz among team members.	21	1.062
2	`mean__matrix_profile__feature_"25"__threshold_0.98=high`, AND `mean__spkt_welch_density__coeff_2=low`, AND `std__root_mean_square=low`	The 0.25 quantile of the similarity of subwindows within the signal to other subwindows within the signal is low, suggesting that a reasonable proportion of subsequences (of the time series) are unusual (not repeating). The strength of the frequency component at 0.234 Hz is low. How well the time slices can be modelled by a linear progression is not varied across the team members.	20	0.964
3	`mean__matrix_profile__feature_"median"__threshold_0.98=high`, AND `std__change_quantiles__f_agg_"var"__isabs_True__qh_0.6__ql_0.4=low`, AND `std__quantile__q_0.9=low`	The median similarity of subwindows within the signal to other subwindows within the signal is low, suggesting that subsequences (of the time series) tend to be relatively unusual (not repeating). The absolute value of changes around the mean (after values have been restricted to remain between the 0.4 and 0.6 quantiles) is consistent across the team. The 0.9 quantile also has low variability.	30	0.811

the change between successive as a target variable also suggest that the frequency components at 0.0–0.25 Hz should be low. The reliability and significance of these frequency components as indicators of increasing speech complexity could be investigated through future work.

4 Conclusions

In conclusion, we have presented a novel approach for local exceptionality detection in time series using subgroup discovery – as a workflow that can be applied in multiple iterations for modeling features and the target, respectively. With these, it is possible to identify features which are strongly associated with or highly predictive of a target variable. We demonstrated the approach via a case study of analysing team interaction data, matching body movement information to a measure of the dynamics of speech. Among other things, this showcased the hypothesis-generating capabilities of our approach. Future research could consider possible refinements, e. g., by considering to extract features from overlapping windows at different offsets as an alternative to non-overlapping windows.

Acknowledgements. The research leading to this work has received funding by the Dutch Research Council (NWO), project KIEM ICT ODYN (ENPPS.KIEM.019.016), and by the German Research Foundation (DFG), project "MODUS" (grant AT 88/4-1).

References

1. Aminikhanghahi, S., Cook, D.J.: A survey of methods for time series change point detection. Knowl. Inf. Syst. **51**(2), 339–367 (2016). https://doi.org/10.1007/s10115-016-0987-z
2. Atzmueller, M.: Subgroup discovery. WIREs DMKD **5**(1), 35–49 (2015)
3. Atzmueller, M., Puppe, F.: SD-map – a fast algorithm for exhaustive subgroup discovery. In: Fürnkranz, J., Scheffer, T., Spiliopoulou, M. (eds.) PKDD 2006. LNCS (LNAI), vol. 4213, pp. 6–17. Springer, Heidelberg (2006). https://doi.org/10.1007/11871637_6
4. Boley, M., Goldsmith, B.R., Ghiringhelli, L.M., Vreeken, J.: Identifying consistent statements about numerical data with dispersion-corrected subgroup discovery. Data Min. Knowl. Disc. **31**(5), 1391–1418 (2017). https://doi.org/10.1007/s10618-017-0520-3
5. Chandola, V., Banerjee, A., Kumar, V.: Anomaly detection for discrete sequences: a survey. IEEE TKDE **24**(5), 823–839 (2010)
6. Christ, M., Braun, N., Neuffer, J., Kempa-Liehr, A.W.: Time series feature extraction on basis of scalable hypothesis tests (tsfresh-a python package). Neurocomputing **307**, 72–77 (2018)
7. Cooke, N.J., Gorman, J.C., Myers, C.W., Duran, J.L.: Interactive team cognition. Cogn. Sci. **37**(2), 255–285 (2013)
8. Duivesteijn, W., Feelders, A.J., Knobbe, A.: Exceptional model mining. Data Min. Knowl. Disc. **30**(1), 47–98 (2015). https://doi.org/10.1007/s10618-015-0403-4
9. Ramirez, E., Wimmer, M., Atzmueller, M.: A computational framework for interpretable anomaly detection and classification of multivariate time series with application to human gait data analysis. In: Marcos, M., et al. (eds.) KR4HC/TEAAM-2019. LNCS (LNAI), vol. 11979, pp. 132–147. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-37446-4_11
10. Fiore, S.M., Smith-Jentsch, K.A., Salas, E., Warner, N., Letsky, M.: Towards an understanding of macrocognition in teams: developing and defining complex collaborative processes and products. Theor. Iss. Ergon. Sci. **11**(4), 250–271 (2010)
11. Galbrun, E., Cellier, P., Tatti, N., Termier, A., Crémilleux, B.: Mining periodic patterns with a MDL criterion. In: Berlingerio, M., Bonchi, F., Gärtner, T., Hurley, N., Ifrim, G. (eds.) ECML PKDD 2018. LNCS (LNAI), vol. 11052, pp. 535–551. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-10928-8_32
12. Hamilton, J.D.: Time Series Analysis. Princeton University Press, Princeton (2020)
13. Hudson, D., Wiltshire, T.J., Atzmueller, M.: Multisyncpy: a python package for assessing multivariate coordination dynamics. PsyArXiv (2021)
14. Jin, N., Flach, P., Wilcox, T., Sellman, R., Thumim, J., Knobbe, A.: Subgroup discovery in smart electricity meter data. IEEE TII **10**(2), 1327–1336 (2014)
15. Keogh, E., Lin, J., Fu, A.: Hot Sax: efficiently finding the most unusual time series subsequence. In: Proceedings ICDM. IEEE (2005)
16. Kim, T., Chang, A., Holland, L., Pentland, A.S.: Meeting mediator: enhancing group collaboration using sociometric feedback. In: Proceedings Conference on Computer supported cooperative work, pp. 457–466. ACM (2008)

17. Knobbe, A., Orie, J., Hofman, N., van der Burgh, B., Cachucho, R.: Sports analytics for professional speed skating. Data Min. Knowl. Disc. **31**(6), 1872–1902 (2017). https://doi.org/10.1007/s10618-017-0512-3
18. Kozlowski, S.W., Chao, G.T.: Unpacking team process dynamics and emergent phenomena: challenges, conceptual advances, and innovative methods. Am. Psychol. **73**(4), 576 (2018)
19. Salas, E., Dickinson, T.L., Converse, S.A., Tannenbaum, S.I.: Toward an understanding of team performance and training. In: Swezey, R.W., Salas, E. (eds.) Teams: Their Training and Performance, pp. 3–29. Ablex Publishing, Norwood (1992)
20. Salas, E., Reyes, D.L., McDaniel, S.H.: The science of teamwork: progress, reflections, and the road ahead. Am. Psychol. **73**(4), 593 (2018)
21. Sanchez-Cortes, D., Aran, O., Mast, M.S., Gatica-Perez, D.: A nonverbal behavior approach to identify emergent leaders in small groups. IEEE Trans. Multimedia **14**(3), 816–832 (2011)
22. Schiepek, G., Strunk, G.: The identification of critical fluctuations and phase transitions in short term and coarse-grained time series–a method for the real-time monitoring of human change processes. Biol. Cybern. **102**(3), 197–207 (2010)
23. Shneiderman, B.: The eyes have it: a task by data type taxonomy for information visualizations. In: Proceedings IEEE Symposium on Visual Languages, pp. 336–343 (1996)
24. Wiltshire, T.J., Hudson, D., Lijdsman, P., Wever, S., Atzmueller, M.: Social analytics of team interaction using dynamic complexity heat maps and network visualizations. arXiv preprint arXiv:2009.04445 (2020)
25. Wiltshire, T.J., Steffensen, S.V., Fiore, S.M.: Multiscale movement coordination dynamics in collaborative team problem solving. Appl. Ergon. **79**, 143–151 (2019)