Welcome to the SAX (**S**ymbolic **A**ggregate approX**i**mation) Homepage!

SAX is the first symbolic representation for time series that allows for dimensionality reduction and indexing with a lower-bounding distance measure. In classic data mining tasks such as clustering, classification, index, etc., SAX is as good as well-known representations such as Discrete Wavelet Transform (DWT) and Discrete Fourier Transform (DFT), while requiring less storage space. In addition, the representation allows researchers to avail of the wealth of data structures and algorithms in bioinformatics or text mining, and also provides solutions to many challenges associated with current data mining tasks. One example is motif discovery, a problem which we defined for time series data. There is great potential for extending and applying the discrete representation on a wide class of data mining tasks.

SAX was invented by Eamonn Keogh and Jessica Lin in 2002, using funding from NSF Career Award 0237918.

Edward Tufte was kind enough to mention that SAX allows a sparkline like visualization of data. The relevant paper is this one **[pdf]**.

Li Wei has generalized the SAX code to handle the *N/n not equal an integer* case, and to allow alphabet sizes up to 20. Download this zip file for the code and details.

If you want a copy of my SAX time series/Shape tutorial, download this.

Here is a video of Dr. Keogh giving a talk at Google about using SAX for various problems, including shape mining.

Much of the utility of SAX has now been subsumed by *i*SAX , which is a generalization of SAX that allows indexing and mining of massive datasets. Visit the *i*SAX page.

Try this Matlab code snippet:

```
startRange = 2;
stdc= 1;
endRange = 512;

table = cell(endRange-startRange,1);
 for r=startRange:endRange
   table{r-startRange+1} = norminv((1:r-1)/r,0,stdc);
 end
```

This will generate a table of breakpoints from 2 to 512, using std 1.

- *the performance **SAX** enables is amazing, and I think a real breakthrough.* Cromwell, Timepedia.org
- *I am a researcher at AT&T, and I have recently been working on a time-seri had tried several different approaches, but so far Ive had the most success*
- ***SAX** represents the state-of-the-art in time series data streams analysis due*
- *In our current research the **(SAX)** symbolic representation of Lin and Keog* Kruse, Steinbrecher and Moewes
- *In order to characterize the expression waveforms we follow the basic **SAX***
- *(for) predictive morphology variations detection in Electrocardiogram Sign **(SAX)**..* Syed, Indyk and Guttag JLMR 2009
- *we explicitly focus on the **SAX** representation, which also provides some sig*
- *We compare wafer patterns using distance measures in **SAX** and provide se*
- *To circumvent the limitations of our previous work, we now rely on a simila* et. al 2007.
- ***(SAX** based VizTree is).. a way to do such analysis more systematically* Edv
- ***SAX**, and more recently **iSAX** ,has gained traction as the de facto represent*
- *(our work on shape mining is) relying on a general-purpose symbolic representation such as **SAX**..* Wang and Candan.
- *The method is based on the notion of so called time signature of the clusters, introduced in Lin & Keogh and obtained us*
- *"Our goal is to identify those transcripts that share significant components of their expression patterns. In order to do so*
- *we take **SAX** Motif developed by Keogh in order to support a medical expert in discovering interesting knowledge.* Kitag
- *In order to symbolize a street data, we utilize the **SAX** approach.* Jalili and Alipour.
- *...we examine another interesting query, the Time Relaxed Spatiotemporal Trajectory Join (TRSTJ)... we address the TRS*
- *We have decided to use **SAX** (to detect sophisticated attack tools).. **SAX** is a recent and popular method with interesting*
- *It's easier to analyze TS in this **(SAX)** form!* Silvia Miksch
- *..we are currently using (Lin and Keoghs **SAX**) approach to creating discrete data from continuous data.* Amy McGovern
- *(to find repeated patterns in protein unfolding data) ...we adopted a two step approach called **SAX*** Ferreira et al.
- *..we use **SAX** and Keogh's **Tarzan** algorithm to do anomaly detection in network traffic.* Kyoji Umemura et. al.
- ***SAX** has already prove efficient in a large variety of domains* Fabian Pouget, Telecom Paris.
- ***SAX** representation of abstracted data makes analysis (of anterior-posterior center of pressure) more easy and accurate.*
- *Our Symbolic Transformation (based on **SAX** method) can be use to discover novel gene relations by mining similar sub*
- *..we use **SAX** bitmap matrices to compute an anomaly score for acoustic signals, enabling the extraction of bird vocaliza*
- *Using the time-series data as an input, it takes too much computation amount to extract motifs from the human motion in*
- ***SAX** has the advantage of dimensionality and noise reductions. It also allows real valued data to remain the original cha*
- ***SAX** demonstrates some promising properties for the field of anomaly detection in a marine engine.* Morgan, Liu, Turnbu
- *..motivated by recent advances in the symbolic representation of streaming data **(SAX)**, effectively reduces the dimension*
- *(we use **SAX** to create a) ..secure multiparty protocol for the privacy preserving pattern discovery problem.* Costa da Silv
- *By using **SAX** with the sensor network data, we are able to detect such complex patterns with good accuracy .. **SAX** is a*
- *we apply the **(SAX** based) motif discovery approach the analysis of responses obtained by tactile stimulation of different*
- *We extend the **SAX** approach.. to support Query-by-Singing/Humming.* Duda, Nurnberger and Stober (2007).
- *We use an algorithm based on **SAX** (Symbolic Aggregate approXimation) to discover human skills..* Makio, Tanaka, and
- *symbolic aggregate approximation **(SAX)** outperform other dimensionality reduction techniques like singular value deco*
- *Portions of our work have been inspired by Symbolic Aggregate Approximation **(SAX)**..* Cohen, Bjornsson, Temple, Banl
- *we apply a technique that has demonstrated success with the interpretation of univariate data, named **SAX** to visualize p*
- *The continuous attributes are transformed into ordered categories using the transformation technique presented in **SAX**.*
- *We argue that symbolic representations (in particular **SAX**) are comparatively superior in the data analysis of time serie*
- *With a **(SAX)** string of symbols describing the trend, analysis is greatly simplified. Field, Stirling , Naghdy and Pan. ACI*
- *We will often apply this symbolization approach, using the methodology of **SAX**, with a primary goal of reducing the nun*
- *We have shown that a **SAX** method is particularly useful..* McGregor et al. 2009
- *In current research the **(SAX)** symbolic representation of Lin and Keogh wins out even over well-known approximations.*
- *The **SAX** representation of images renders itself as a better alternativeto the common histogram representation specially*
- *(For the problem of) identification of Global Transcriptional Dynamics we have elected to explore the basic principles o*
- *We have developed Tiny**SAX** - an efficient implementation of SAX using integer arithmetic that makes the application of*
- ***SAX** is a clever way to discretize signals..* Castro and Ferreira 2007.
- *A method of prediction .. on **SAX** utilising an alternative modelling approach known as Markov chaining.* Morgan and L
- *We validated the SAX technique using reflectance indices coming from the band 3 (near infrared) imaging AVHRR-NOO*
- *The Symbolic Aggregate approximation **(SAX)** technique, enables the use of existing text-retrieval techniques and algori*
- *We treat **SAX** as a black box to which we pass a numeric sensor sequence and it returns a reduced string representation.*
- ***SAX** is a time series representation that has proved to be the state-of-the-art technique in time series representation an*
- *We applied the **SAX** representation using string lengths of 32.. (for Collaborative Context Prediction)* Voigtmann et al. 2
- *Specifically, our approach uses aperiodicity information to identify articulations, histograms to compute the density of t*
- We use five different data sets and classifiers to show that a combination of **SAX** similarity features and basic features ir
- *We use **iSAX** (indexable Symbolic Aggregate approXimation) to enable interactive searching (<2 sec/example) of a large*
- *We converted the data into propositional time series by applying the **SAX** algorithm...* Kerr et al. IJCAI11
- *It should be noted that even the classification results of the **SAX** feature set outperform the basic feature results by sever*
- *We shall use the **SAX** symbolization in computing (gait)..* Parshad et al 2011
- *We use **SAX** to measure movement asymmetry symptoms associated with Parkinson's disease.* Sant' Anna. 2011 IEEE T
- *This approach,called **SAX**, has also been shown to improve clustering due to the smoothing effect of dimensionality redu*
- *each morphological profile is then normalized and discretized. For this purpose, the Symbolic Aggregate approXimation*

Papers by Keogh and collaborators that use SAX. (in random order)

In [1] we show how to use SAX to find time series *discords* which are unusual time series. In [2] we consider a special case of SAX, which has an alphabet size of 2, and a word size equal to the raw data, and show that we can use this bit-level representation for a variety of data mining tasks. In [3] we show how to use SAX to create time

series bitmaps, which allow visualization of time series data directly within a standard GUI such as MS Windows. In [4] we further show how to use time series bitmaps to do anomaly detection. In [5] we show that SAX can support parameter-lite data mining of time series, including classification and clustering. In [7] we show that SAX can replace standard representations of time series (i.e DWT, DFT) for all classic data mining problems including classification, clustering and indexing. We first used SAX to find time series motifs (exactly, and somewhat fast) in [9], and later showed a blinding fast probabilistic algorithm in [8]. In [10] we tentatively showed how to use SAX to meaningfully cluster time series streams. In [12] we show an application of SAX to a shape mining problem, and in [11] we generalize the time series bitmap concept to more general datasets. In [13] we show how to use SAX to find approximately duplicated shapes (shape motifs) in large databases. Paper [14] is a journal paper reviewing SAX first two years. Paper [15] shows how to find motifs under uniform scaling. Paper [16] introduces *i*SAX. Paper [17] shows how to do SAX on resource limited sensors.

1. **E. Keogh, J. Lin and A. Fu (2005). HOT SAX: Efficiently Finding the Most Unusual Time Series Subsequence. In Proc. of the 5th IEEE International Conference on Data Mining (ICDM 2005), pp. 226 - 233., Houston, Texas, Nov 27-30, 2005. [pdf].** More info on *discords* and HOT SAX is here Also KAIS journal paper.
2. **Ratanamahatana, C., Keogh, E., Bagnall, T. and Lonardi, S. (2005). A Novel Bit Level Time Series Representation with Implications for Similarity Search and Clustering. PAKDD 05. [pdf]** Also DMKD journal paper.
3. **Kumar, N., Lolla N., Keogh, E., Lonardi, S. , Ratanamahatana, C. A. and Wei, L. (2005). Time-series Bitmaps: A Practical Visualization Tool for working with Large Time Series Databases . In proceedings of SIAM International Conference on Data Mining (SDM '05), Newport Beach, CA, April 21-23. pp. 531-535 [pdf]**
4. **Li Wei, Nitin Kumar, Venkata Nishanth Lolla, Eamonn Keogh, Stefano Lonardi, Chotirat Ann Ratanamahatana (2005). Assumption-Free Anomaly Detection in Time Series. In Proc. of the 17th International Scientific and Statistical Database Management Conference (SSDBM 2005), Santa Barbara, CA, U.S.A., June 27-29, 2005.**
5. **Keogh, E., Lonardi, S. and Ratanamahatana, C. (2004). Towards Parameter-Free Data Mining. In proceedings of the tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Seattle, WA, Aug 22-25, 2004. [pdf, slides]** Also DMKD journal paper.
6. **Lin, J., Keogh, E., Lonardi, S., Lankford, J. P. & Nystrom, D. M. (2004). Visually Mining and Monitoring Massive Time Series. In proceedings of the tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Seattle, WA, Aug 22-25, 2004. [pdf ,slides]** Also Information Visualization journal paper.
7. **Lin, J., Keogh, E., Lonardi, S. & Chiu, B. (2003) A Symbolic Representation of Time Series, with Implications for Streaming Algorithms. In proceedings of the 8th ACM SIGMOD Workshop on Research Issues in Data Mining and Knowledge Discovery. San Diego, CA. June 13. [pdf, slides]**
8. **Chiu, B. Keogh, E., & Lonardi, S. (2003). Probabilistic Discovery of Time Series Motifs. In the *9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.* August 24 - 27, 2003. Washington, DC, USA. pp 493-498. [Expanded Version pdf]**
9. **Patel, P., Keogh, E., Lin, J., & Lonardi, S. (2002). Mining Motifs in Massive Time Series Databases. In proceedings of the 2002 IEEE International Conference on Data Mining. Maebashi City, Japan. Dec 9-12.**
10. **Keogh, J. Lin, and W. Truppel. (2003). Clustering of Time Series Subsequences is Meaningless: Implications for Past and Future Research. In *proceedings of the 3rd IEEE International Conference on Data Mining* . Melbourne, FL. Nov 19-22. pp 115-122. [pdf]** Also KAIS journal paper.
11. **Eamonn Keogh, Li Wei, Xiaopeng Xi, Stefano Lonardi, Jin Shieh, Scott Sirowy (2006). Intelligent Icons: Integrating Lite-Weight Data Mining and Visualization into GUI Operating Systems. ICDM 2006. [pdf]**
12. **Li Wei, Eamonn Keogh and Xiaopeng Xi (2006) SAXually Explict Images: Finding Unusual Shapes. ICDM 2006. [pdf].** Now a Data Mining and Knowledge Discovery Journal paper.
13. **Xiaopeng Xi, Eamonn Keogh, Li Wei, Agenor Mafra-Neto (2007). Finding Motifs in a Database of Shapes. SIAM International Conference on Data Mining.**
14. **Jessica Lin, Eamonn Keogh Li Wei and Stefano Lonardi (2007) Experiencing SAX: a Novel Symbolic Representation of Time Series. DMKD Journal.**
15. **Dragomir Yankov, Eamonn Keogh, Jose Medina, Bill Chiu, and Victor Zordan (2007). Detecting Motifs Under Uniform Scaling. SIGKDD 2007. [pdf]** Supporting webpage with video and datasets.
16. **Jin Shieh and Eamonn Keogh** (2008). ***i*SAX: Indexing and Mining Terabyte Sized Time Series**. SIGKDD 2008. Also DMKD journal paper.
17. **Shashwati Kasetty, Candice Stafford, Gregory P. Walker, Xiaoyue Wang, Eamonn Keogh** (2008). **Real-Time Classification of Streaming Sensor Data.** 20th IEEE Int'l Conference on Tools with Artificial Intelligence. [pdf]
18. **Alessandro Camerra, Themis Palpanas, Jin Shieh and Eamonn Keogh** (2010) ***i*SAX 2.0: Indexing and Mining One Billion Time Series.** ICDM 2010 [pdf]

Selected papers by others that use SAX.

In [A] the authors *"New approaches for representing, analyzing and visualizing complex kinetic mechanisms"*, they note "*The procedure is based on the methodology recently proposed by (Lin and Keogh) for the analysis of multi-dimensional time series*". Papers [B.C,D,E] use SAX and random projection (see [8] above) to discover motifs in telemedicine time series. In paper [F] the authors convert plamprint to time series, then to SAX, then they do biometric recognition. Paper [G] says "*we take Motif developed by Keogh in order to support a medical expert in discovering interesting knowledge*". Paper [H] uses SAX and random projection (see [8] above) to mine motion capture data. Paper [I] uses SAX to find repeated patterns in motion capture data. Paper [J] uses SAX to find rules in time series. Paper [K] uses SAX to find motifs of unspecified length. Paper [L] uses SAX to find repeated patterns in robot sensors. Androulakis et. al. [M] uses SAX for electing Maximally Informative Genes to. Enable Temporal Expression Profiling. In paper [N] the authors us SAX to do Spatiotemporal Trajectory Joins. In [O] the authors use SAX motifs to *"analyze respiration wave during cello performance*" Paper [P] uses SAX to *"detect multi-headed stealthy attack tools".* Paper [Q] is using SAX to " *Understand the formation of tornadoes*"! Paper [R] uses SAX and time series motifs to for the *Selection of Informative Genes in Time-Course Gene Expression Data*. Paper [S] makes the minor extensions to [8] above, to allow it to handle the multidimensional case. Ph.d Thesis [T] uses SAX for a variety of tasks in network traffic analysis. Paper [U] uses SAX to do *Anomaly Detection in Network Traffic*. Paper [V] uses SAX to do prediction of severe weather phenomena such as tornados, thunderstorms, hail, and floods. Paper [W] uses a modification of SAX to discover novel gene relations by mining similar subsequences in time-series microarray data. Paper [X] uses SAX for classification of environmental sounds. [Y] uses SAX for financial data mining. Paper [Z] uses SAX for motif discovery. Paper [AA] uses SAX to find motifs in motion capture data. paper [AB] uses SAX based motifs to mine system call sequences. Paper [AB] uses SAX to classify control chart patterns. Papers [AD] and [AE] extend SAX for segmentation of time series into natural episodes. Paper [AF] uses SAX to find anomalies in SAX in a marine engine. Paper [AG] uses SAX for the selection of informative genes. paper [AH] uses SAX to detect complex events in wireless sensor networks. Paper [AI] uses SAX to mine MRIs. Paper [AJ] uses SAX to mine motion capture data. Paper [AK] uses SAX for privacy-preserving discovery of frequent patterns in time series. Paper [AL] uses SAX to find association rules in time series. Paper [AM] uses SAX and Vistree to find patterns in CPU traces. Paper [AN] uses SAX for similarity search. Paper [AO] uses a SAX-like (but not SAX) approach for assessing the wellbeing of unsupervised, vulnerable individuals. Paper [AP] uses SAX for characterizing the mechanism of action of anti-inflammatory drugs. Paper [AQ] uses SAX to visualize patterns that may differentiate between medical conditions such as renal and respiratory failure. Paper [AR] uses SAX to Understand malicious internet traffic by mining honeypot traces. Paper [AS] uses SAX to mine ECG data. Paper [AT] uses SAX to tokenize for gestures. Paper [AU] uses SAX for palm print biometrics. Paper [AV] uses SAX for ECG pattern recognition on mobile devices.Paper [AW] uses SAX for large-scale network traffic analysis. Paper [AX] uses SAX for robotic motion segmentaion. Paper [AY] uses SAX for as part of a complexity measure for nonstationary signals. Paper [AZ] uses SAX for mining color distributions in images. Paper [BA] uses SAX for measuring brain states. Paper [BB] uses SAX for mining hurricane data. Paper [BC] uses SAX for quality control in semiconductor manufacturing. Paper [BD] uses SAX as an input to a Markov prediction system. Paper [BE] uses SAX for panic disorder treatment. Paper [BF] uses SAX for analog-circuit fault diagnosis using three-stage preprocessing and time series data. Paper [BG] uses SAX for Mining closed flexible patterns in time-series databases. Paper [BH] uses SAX for clustering industrial heating telemetry. Paper [BI] uses SAX for a computational resource advisory system. Paper [BJ] uses SAX to mine human gait data. Paper [BK] (in Portuguese) uses SAX to mine river levels. Paper [BL] augments SAX for the identification of informative genes in replicated microarray experiments. Paper [BM] uses SAX (actually *i*SAX) for mining motifs. Paper [BN] uses SAX for shape mining. ...and so on...

A. Androulakis, I. P. (2005). *New Approaches for Representing, Analyzing and Visualizing Complex Kinetic Mechanisms*. . In proceedings of the 15th European Symposium on Computer Aided Process Engineering. Barcelona, Spain. May 29-June 1.

B. Silvent, A., Dojat, M. & Garbay, C. (2004). *Multi-level Temporal Abstraction for Medical Scenario Construction*. International Journal of Adaptive Control and Signal Processing.

C. Silvent, A. S., Carbay, C., Carry, P. Y. & Dojat, M. (2003). *Data, Information and Knowledge for Medical Scenario Construction.* In proceedings of the Intelligent Data Analysis In Medicine and Pharmacology Workshop (IDAMAP 2003). October. Protaras, Cyprus.

D. F. Duchene, C. Garbay, V. Rialle, "*Mining heterogeneous multivariate time-series for learning meaningful patterns: Application to home health telecare,*" Research Report 1070-I, Institut d Informatique et Mathematiques Appliquees de Grenoble (IMAG), Grenoble, France, 2004.

E. F. Duchene and C. Garbay, *Apprentissage de motifs temporels, multidimensionnels et heterogenes -- Application a la telesurveillance medicale,* Conference francophone sur lapprentissage automatique (CAP), Nice, France, 31 mai - 3 juin 2005. Presses Universitaires de Grenoble.

F. Chen, J. S., Moon, Y. S. & Yeung, H. W. (2005). *Palmprint Authentication Using Time Series*. In proceedings of the 5th International Conference on Audio- and Video-Based Biometric Person Authentication. Hilton Rye Town, NY. July 20-22.

G. Kitaguchi, S. (2004). *Extracting Feature based on Motif from a Chronic Hepatitis Dataset.* In proceedings of the 18th Annual Conference of the Japanese Society for Artificial Intelligence (JSAI). Kanazawa, Japan. June 2-4.

H. Tanaka, Y. & Uehara, K. (2004). *Motif Discovery Algorithm from Motion Data*. In proceedings of the 18th Annual Conference of the Japanese Society for Artificial Intelligence (JSAI). Kanazawa, Japan. June 2-4.

I. Celly, B. & Zordan, V. B. (2004). *Animated People Textures*. In proceedings of the 17th International Conference on Computer Animation and Social Agents (CASA 2004). July 7-9. Geneva, Switzerland.

J. Ohsaki, M., Sato, Y., Yokoi, H. & Yamaguchi, T. (2003). *A Rule Discovery Support System for Sequential Medical Data In the Case Study of a Chronic Hepatitis Dataset*. ECML 2003

K. Tanaka, Y. & Uehara, K. (2003). *Discover Motifs in Multi Dimensional Time-Series Using the Principal Component Analysis and the MDL Principle*. In proceedings of the 3rd International Conference on Machine Learning and Data Mining in Pattern Recognition. pp.252-265.

L. Koji Murakami Yoshikazu Yano Shinji Doki Shigeru Okuma (2004). *Behavior extraction from a series of observed robot motion* . RoboMec2004

M. Androulakis, I.P., J. Wu, J. Vitolo and C. Roth, *Selecting maximally informative genes to enable temporal expression profiling analysis*, Proceedings of Foundations of Systems Biology in Engineering, Santa Barbara, CA, (2005)

N. P. Bakalov, M. Hadjieleftheriou, V. J. Tsotras, (2005). *Time Relaxed Spatiotemporal Trajectory Proc. of the ACM International Symposium on Advances in Geographic Information Systems*(ACM-GIS),Bremen, Germany, November 2005.

O. Keita Kinjo Tomonobu Ozaki Keigo Sawai Koichi Furukawa (2005) *Knowledge acquisition from time series data by association rule and network analysis.* The 19th Annual Conference of the Japanese Society for Artificial Intelligence, 2005

P. F. Pouget, G. Urvoy-Keller, and M. Dacier *Time Signatures to detect multi-headed stealthy attack tools* In 18th Annual FIRST Conference Baltimore, Maryland, USA June 2006.

Q. Amy McGovern, Univ. of Oklahoma, Norman, OK; and A. Kruger, D. Rosendahl, and K. Droegemeier. (2007) *Understanding the formation of tornadoes through data mining*. Fifth Conference on Artificial Intelligence Applications to Environmental Science.

R. Eric Yang and Ioannis Androulakis (2006) *Selection of Informative Genes in Time-Course Gene Expression Data.* AIChE 2006.

S. David Minnen, Thad Starner, Irfan Essa, and Charles Isbell. *Improving Activity Discovery with Automatic Neighborhood Estimation*. In Proceedings of the Twentieth International Joint Conference on Artificial Intelligence (IJCAI), 2007.

T. Fabian Pouget (2005). *Distributed System of Honeypot Sensors*. Telecom Paris

U. Masayuki Okabe, Taeko Miwa, Kyoji Umemura (2006) *Anomaly Detection in Network Traffic based on String Analysis.* IC2006.

V. McGovern, Amy, and Rosendahl, Derek H., and Kruger, Adrianna, and Beaton, Meredith G., and Brown, Rodger A., and Droegemeier, Kelvin K. (2007) *Understanding the formation of tornadoes through data mining*. Fifth Conference on Artificial Intelligence and its Applications to Environmental Sciences at the American Meteorological Society annual conference.

W. Vincent Shin-Mu Tseng, L. C. Chen and J. J. Liu (2006) *Discovering Novel Gene Relations by Mining Similar Subsequences in Time in Time-Series Microarray Data.* in Proc. Intl Workshop on Science of Artificial, Taiwan, 2005

X. *Automated Ensemble Extraction and Analysis of Acoustic Data Streams*. Technical Report MSU-CSE-06-40. December 2006. Eric P. Kasten and Philip K. McKinley Stuart H. Gage

Y. *Application Research of a New Symbolic Approximation Method-SAX in Time Series Mining* (2006) LIU Yi,BAO De-pei,YANG Ze-hong. COMPUTER ENGINEERING AND APPLICATIONS 2006 Vol.42 No.27

Z. *Motif Detection Inspired by Immune Memory* (2007) William Wilson, Phil Birkin, and Uwe Aickelin

AA. *Motion motif extraction from high-dimensional motion information.* Araki , Arita and Taniguchi 2006

AB. Wilson Will, Feyereisl Jan and Aickelin Uwe (2007): *Detecting Motifs in System Call Sequences*, Proceedings of the 8th International Workshop on Information Security Applications (WISA 2007), Lecture Notes in Computer Science, pp, Jeju, Korea

AC. K. Lavangnananda, and C. Wongwattanakarn (2007) *Utilizing Symbolic Representation and Evolutionary Computation in Classification of Control Chart Patterns*. Soft Computing in Industrial Applications 2007

AD. T. Armstrong and T .Oates. RIPTIDE: *Segmenting Data Using Multiple Resolutions.* In the Proceedings of the 6th IEEE International Conference on Development and Learning (ICDL), 2007.

AE. T. Armstrong and T. Oates. UNDERTOW: *Multi-Level Segmentation of Real-Valued Time Series*. In the Proceedings of the 22nd AAAI Conference on Artificial Intelligence (AAAI) (student abstract), 2007.

AF. Time Discretisation Applied to Anomaly Detection in a Marine Engine. Morgan, Liu, Turnbull, and Brown 2007.

AG. Yang, E., F. Berthiamume, M. L. Yarmush, and I. P. Androulakis. *SeLection of INformative Genes via Symbolic Hashing Of Time Series* .Proceedings of the Joint 9th International Symposium, Processing Systems Engineering and 16th European Symposium, 2006.

AH. M. Zoumboulakis and G. Roussos, *Escalation: Complex Event Detection in Wireless Sensor Networks* ,in Proceedings of 2nd European Conference on Smart Sensing and Context (EuroSSC), 23-25 Oct 2007, Lake District, UK.

AI. M. Fabri, G. Mascioli, G. Palonara, A. M. Perdon, S. R. Viola (2007) *Activation and delay in FMRI brain signals of selective attention.* in Proceedings of Int. IJCNN07 Workshop on Neurodynamics, Orlando, Florida, USA, August 17, 2007.

AJ. Kosuke Makio, Yoshiki Tanaka, and Kuniaki Uehara (2007) *Discovery of Skills from Motion Data.* New Frontiers in Artificial Intelligence

AK. Da Silva, J.C.; Klusch, M. (2007): *Privacy-Preserving Discovery of Frequent Patterns in Time Series.* Proceedings of the 7th Industrial Conference on Data Mining ICDM, Leipzig, Germany, Springer.

AL. *Discovery Association Rules in Time Series Data*. Kittipong Warasup and Chakarida Nukoolkit

AM. Ooi Boon Yaik Chan Huah Yong Fazilah Haron (2006) *CPU Usage Pattern Discovery Using Suffix Tree. Distributed Frameworks for Multimedia Applications,* 2006.

AN. C*ombining SAX and Piecewise Linear Approximation to Improve Similarity Search on Financial Time Series.* Hung, Nguyen Quoc Viet Anh, Duong Tuan Information Technology Convergence, 2007. ISITC 2007.

AO. Julia Hunter and Martin Colley (2007) *Feature Extraction from Sensor Data Streams for Real-Time Human Behaviour Recognition*. PKDD2007

AP. *Analysis of Regulatory Interaction Networks from Clusters of Co-expressed Genes* (2008) E. Yang et al

AQ. *Visualizing Multivariate Time Series Data to Detect Specific Medical Conditions.* Ordonez et al. AMIA 2008.

AR. Almotairi, Saleh I. et al (2007) *Extracting Inter-arrival Time Based Behaviour from Honeypot Traffic using Cliques.*

AS. Kulahcioglu B., Ozdemir S., Kumova B.I., *Application of Symbolic Piecewise Aggregate Approximation (PAA) Analysis to ECG Signals*, The 17th IASTED International Conference on Applied Simulation and Modelling (ASM 2008) .

AT. *Tokenization for Gesture Space Modelling.* Aaron Licata, Alexandra Psarrou 13th International Conference on Applications of Natural Language to Information Systems, Doctoral Symposium (NLDB'08-DS)

AU. *Using SIFT Features in Palmprint Authentication*. Jiansheng Chen Yiu-Sang Moon. ICPR'08

AV. Eirik Aanonsen and Rune Fensli, *Pattern recognition on mobile devices*. 2006

AW. *Applying multiple time series data mining to large-scale network traffic analysis* Weisong He,; Guangmin Hu,; Xingmiao Yao,; Guangyuan Kan,; Hong Wang,; Hongmei Xiang 2008.

AX. *Motion segmentation for humanoid control planning*. Matthew Field, David Stirling , Fazel Naghdy, Zengxi Pan. ACRA 2008

AY. Erik M. Bollt, Joseph D. Skufca, Stephen J McGregor, *Control Entropy: A Complexity Measure for Nonstationary Signals*, Mathematical Biosciences and Engineering, 6 1 1-25 (2009)

AZ. *Effective Image Mining by Representing Color Histograms as Time Series* (2007) Zaher Al Aghbari. Journal of Advanced Computational Intelligence and Intelligent Informatics

BA. *Combining Electroencephalograph and Functional Near Infrared Spectroscopy to Explore Users' Mental Workload*. Leanne M. Hirshfield. 2009.

BB. *Range Queries over Trajectory Data with Recursive Lists of Clusters: a case study with Hurricanes data*. GISRUK 2009

BC. *Full-line FDC Diagnosis System via Signals Similarity Measure. Tzu-Cheng* Lin and Shir-Kuan Lin. AEC/APC Symposium Asia 2009

BD. *Predicting Future States with n-Dimensional Markov Chains for Fault Diagnosis*. Morgan and Lui, IEEE Transactions on Industrial Electronics.

BE. *Intelligent panic disorder treatment by using biofeedback analysis and web technologies*. Shie et al International Journal of Business Intelligence and Data Mining 5(1) 2010

BF. *Analog-Circuit Fault Diagnosis Using Three-Stage Preprocessing and Time Series Data. Mining.* Weisong He, Hongmei Xiang and Jingyuan 2009

BG. *Mining closed flexible patterns in time-series databases* . Huei-Wen Wu, Anthony J.T. Lee 2010
BH. *Clustering Techniques Applied to Multiple-Models Structures*. Silva , Becerra & Calado 2009
BI. *CPU Usage Pattern Discovery Using Suffix Tree For Computational Resource Advisory System* . Ooi, Boon Yaik (2006)
BJ. *Human Gait Data Mining by Symbol Based Descriptive Features.* 2009. Ergovic, Tonkovic, and Medved
BK. *Analise de padroes sequenciais em serie historica do rio Paraguai.* Anais 2 Simposio de Geotecnologias no Pantanal, Corumba, 7-11 novembro 2009, p.323-332. Laurimar Goncalves Vendrusculo, Stanley Robson de Medeiros Oliveira, Julio Cesar Dalla Mora Esquerdo, Joao Francisco Goncalves Antunes
BL. *A New Symbolic Representation for the Identification of Informative Genes in Replicated Microarray Experiments.* (2010) Jeremy D. Scheff, Richard R. Almon, Debra C. DuBois, William J. Jusko, and Ioannis P. Androulakis
BM. *Multiresolution Motif Discovery in Time Series.* Nuno C. Castro and Paulo Azevedo, SDM 2010
BN. *Relevant shape contour snippet extraction with metadata supported hidden Markov models.* Wang and Candan. CIVR 10.
BO. *Multiple Kernel Learning for Heterogeneous Anomaly Detection: Algorithm and Aviation Safety Case Study.* Santanu Das, Bryan Matthews, Ashok Srivastava, ; Nikunj Oza, NASA Ames Research Center (For the continuous data, each time series was SAX transformed)
BP. *Beating the baseline prediction in food sales: How intelligent an intelligent predictor is?* Indre Zliobaite (we used quadruple SAX representation of the series..)
BQ. *A 3D Visualization Technique for Large Scale Time-Varying Dat*a. Maiko Imoto, Takayuki Ito (We apply SAX for symbolic character representation of time-varying values)
BR. *An Animated Multivariate Visualization for Physiological and Clinical Data in the ICU*. IHI 10 . Ordonez et al.
BS. Michael Zoumboulakis, George Roussos: *Complex Event Detection in Extremely Resource-Constrained Wireless Sensor Networks*. MONET 16(2): 194-213 (2011)
BT. RA-SAX: Resource-Aware Symbolic Aggregate approXimation for Mobile ECG Analysis, Hossein Tayebi , Shonali Krishnaswamy ,Agustinus Waluyo, Abhijat Sinha , , Mohamed Gaber .
BU. Can Temporal and Spatial Patterns of Dynamc Terrain State Properties be Determined Using a Symbolic Aggregate ApproXimation (SAX) Approach? FRANKENSTEIN ICMG 2011
BV. *Event Detection using Archived Smart House Sensor Data obtained using Symbolic Aggregate Approximation.* A. Onishi and C. Watanabe. PDPTA 2011.
BW. *3D Time-Varying Data Visualization Method Technique Featuring Symbolic Aggregate approximation* ,M. Imoto, T. Itoh, IEEE Pacific Visualization 2011.
BX. *Legato and Glissando identification in Classical Guitar*. Ozaslan and Arcos 2010. 7th Sound and Music Computing Conference
BY. *Attack Based Articulation Analysis of Nylon String Guitar*. Ozaslan and Arcos CMMR2010
BZ. Voigtmann, C.; Lau, S. L. & David, K. (2011), An Approach to Collaborative Context Prediction, in '2011 IEEE International Conference on Pervasive Computing and Communications Workshops. IEEE
CA. *Improving the Classification Accuracy of Streaming Data Using SAX Similarity Features.* Pekka Siirtola et al Pattern Recognition Letters.
CB. *Discovering Patterns for Prognostics: A Case Study in Prognostics of Train Wheels*. Chunsheng Yang and Sylvain Letournea. IEA/AIE (1) 2011: 165-175
CC. *MAGIC 2.0: A Web Tool for False Positive Prediction and Prevention for Gesture Recognition Systems.* Daniel Kohlsdorf, Thad Starner, Daniel Ashbrook: FG' 11, 2011
CD. *Activity Recognition with Finite State Machines.* Wesley Kerr, Anh Tran and Paul Cohen IJCAI 2011
CE. Rana D. Parshad, Stephen J. McGregor, Michael A. Busa, Joseph D. Skufca, Erik Bollt *A statistical approach to the use of Control Entropy identifies differences in constraints of gait in highly trained versus untrained runner*s, Mathematical Sciences and Engineering (MBE) (2011)
CF. Heeyoul Choi, Chen Yu, Olaf Sporns and Linda Smith, "*From Data Streams to Information Flow: Information Exchange in Child-Parent Interaction,*" The Annual Meeting of the Cognitive Science Society (CogSci 2011), Boston, MA. July 20-23, 2011.
CG. *Unsupervised Discovery of Motifs under Amplitude Scaling and Shifting in Time Series Databases*. Tom Armstrong and Eric Drewniak.Lecture Notes in Computer Science, 2011, Volume 6871, Machine Learning and Data Mining in Pattern Recognition, Pages 539-552
CH. Pekka Siirtola, Heli Koskim&#65533;ki, Ville Huikari, Perttu Laurinen, Juha R&#65533;ning: *Improving the classification accuracy of streaming data using SAX similarity features.* Pattern Recognition Letters 32(13): 1659-1668 (2011)
CI. F. Ciompi, O. Pujol, S. Balocco et al., &#65533;Automatic Key Frames Detection in Intravascular Ultrasound Sequences,&#65533; in MICCAI Workshop in Computing and Visualization for (Intra)Vascular Imaging (CVII), 2011
CJ. Takuma Nishii, Tomoyuki Hiroyasu, Masato Yoshimi, Mitsunori Miki, Hisatake Yokouchi: *Similar subsequence retrieval from two time series data using homology search*. SMC 2010: 1062-1067
CK. Anita Sant&#65533;Anna et al (2011). *A New Measure of Movement Symmetry in Early Parkinson&#65533;s Disease Patients Using Symbolic Processing of Inertial Sensor Data*. IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING
CL. Hyokyeong Lee, Asher Moody-Davis, Utsab Saha, Brian M. Suzuki, Daniel Asarnaw, Steven Chen, Michelle Arkin, Conor R. Caffrey, and Rahul Singh, Quantification and Clustering of Phenotypic Screening Data using Time-Series Analysis for Chemotherapy of Schistosomiasis, BMC Genomics,
CM. *Unsupervised Simultaneous Learning of Gestures, Actions and their Associations for Human-Robot Interaction.* Yasser Mohammad Toyoaki Nishida Shogo Okada IROS 2009: 2537-2544