Domain Ontology for Time Series Provenance

Lucélia de Souza[1,2], Maria Salete Marcon Gomes Vaz[2,3] and Marcos Sfair Sunye[2]
[1]*Computer Science Department, State University of Center-West, Guarapuava, Paraná, Brazil*
[2]*Informatics Department, Federal University of Paraná, Curitiba, Paraná, Brazil*
[3]*Informatics Department, State University of Ponta Grossa, Ponta Grossa, Paraná, Brazil*

Keywords: OWL, Modules, Time Series Processing, Trend Extraction, Nonstationary Time Series.

Abstract: Time series data are generated all the time with a volume without precedent, constituting themselves of a points sequence spread out over time, usually at time regular intervals. Time series analysis is different from data analysis, given its intrinsic nature, where observations are dependent and the observations order is important for analysis. The knowledge about the data which will be analyzed is relevant in an analysis process, but this knowledge is not always explicit and easy to interpret in many information resources. Time series can be semantically enriched where provenance information using ontologies allows to representing and inferring knowledge. The main contribution of this paper is to present a domain ontology developed by modular design for time series provenance, which adds semantic knowledge and contributes to the choice of appropriate statistical methods for an important step of time series analysis that is the trend extraction (detrending). Trend is a time series component that needs be extracted because it can hide other phenomena, as well as the most statistical methods are developed for stationary time series. With this work, is intended to contribute for semantically improving the decision making about trend extraction step, facilitating the preprocessing phase of time series analysis.

1 INTRODUCTION

The scientific knowledge generation, in several domains, is related with the time series analysis, from which is extracted useful information. Time series data are characterized by way as they were generated and collected, usually at time regular intervals (Chandler and Scott, 2011; Cryer and Chan, 2008).

Time series analysis is usually done in two phases, preprocessing and data analysis, both containing processing steps in order to obtain scientific knowledge. Time series analysis is different from data analysis, given its intrinsic nature, where observations are dependent or correlated and the observations order is important for analysis. Statistical procedures and traditional techniques based on assumptions of independent and identically distributed data are not applied in time series. This way, are necessary different methods of analysis (Cryer and Chan, 2008).

In time series analysis, provenance information, such as *What the observation type of time series?*, *How the time series were generated?*, *What is the decomposition model used?*, *What assumptions were considered?*, *How the time series data can be classified according to assumptions?*, *What the trend type considered?*, among other information, allowing the researcher to interpret the data better and to use appropriate statistical methods, specifically developed regarding its characteristics.

Hair et al (2010) asserts that the knowledge about data that will be analyzed is important in an analysis process. However, according to Hebeler et al (2009), this is not present in several information resources. Such knowledge is not always explicit and easy to interpret. As well as in data analysis, time series also can be semantically enriched where provenance information using ontologies allows representing and inferring knowledge.

This paper describes Time Series Ontology (namespace tso:), a domain ontology (a module in Ontology Web Language - OWL) with the definition of main concepts and relationships involving time series provenance. The proposed ontology adds semantic knowledge in time series, contributing to choose of appropriate statistical methods for an important step of analysis that is the trend extraction, also called detrending (Wu et al, 2007; Meinl, 2011).

Besides of this section, Section 2 describes time

de Souza L., Marcon Gomes Vaz M. and Sfair Sunye M..
Domain Ontology for Time Series Provenance.
DOI: 10.5220/0004886502170224
In *Proceedings of the 16th International Conference on Enterprise Information Systems* (ICEIS-2014), pages 217-224
ISBN: 978-989-758-028-4

series. The Section 3 relates ontologies and the Section 4 reports time series provenance. Section 5 presents the development of the domain ontology. Section 6 brings a comparison with related work. Finally, Section 7 describes the conclusions and future works perspectives.

2 TIME SERIES

Time series are an ordered sequence of observations, usually obtained at time regular intervals (Chandler and Scott, 2011; Cryer and Chan, 2008). Time series are correlated data, being preferable the use of analysis methods specifically developed to this data type. Chandler and Scott (2011) explain that the choice of appropriated methods of analysis depends on questions of interest, where the knowledge about time series data is essential.

The first step in any time series analysis is the careful observation of plotted data over the time. This procedure frequently suggests statistical methods of analysis, as well as the statistics that summarize the information about the data.

A time series can have the following classical components (Meinl, 2011): Seasonal, Cyclic, Irregular and Trend. In environmental sciences, the trend is defined by Chandler and Scott (2011) as a long-term temporal variation in statistics properties of a process. However, the period of trend is dependent of each application.

In environmental applications, among the possible reasons for trend analysis stands out the analysis of systems, where long-term changes can to obscure aspects of real interest (Chandler and Scott, 2011). In this case, the first step in the time series analysis is the identification and extraction of trend to clearly understand the inter-relationships of data. This paper is related with this reason, where trend is clearly defined and needs to be extracted, otherwise to make difficult the knowledge discovery process.

To many analyses, trend extraction is an important step because it can to hide other phenomena, as well as the most statistical methods are developed for stationary time series (Montesino-Pouzols and Lendasse, 2010). It making a stationary time series means to extract all the deterministic features, such as statistics measures of mean and variance, in such way that correlations turn themselves independent over time.

The need of modeling these and others characteristics must be considered in order to determine an adequate strategy in the analysis (Chandler and Scott, 2011). The definitions and the main features of the time series presented contribute for the choice of appropriate statistical methods. Such descriptions are considered in development of the domain ontology as way of adding semantic knowledge. The next section describes Ontologies used to generation of time series provenance.

3 ONTOLOGIES

Ontology is a formal and explicit specification of a shared conceptualization (Borst 1997). Kiryakov (2006) defines ontology as a set of classes representing concepts in domain, relationships between concepts, axioms used to modeling restrictions and rules and instances of classes, constituting a knowledge base.

Guarino (1998) classifies ontologies according to the generality level, where domain ontologies constitute vocabularies about a generic domain. In this work the domain of ontology is related with time series provenance.

Although Resource Description Framework - RDF and OWL were not developed to given support for numeric concepts, because they depend on schema definition and are based on the eXtensible Markup Language - XML, the set of upper level ontologies Semantic Web for Earth and Environmental Terminology - SWEET <http://sweet.jpl.nasa.gov/ontology/> is a good example to integrate mathematical knowledge with scientific application domains (Lange, 2013). In this work some statements from SWEET are reused and extended.

In the Semantic Web defined by Berners-Lee (2001), the W3C Standard defines the effort of the Linking Open Data (LOD) community where is increased the number of data providers that publishing and interlinking data on the web.

The actual web of data consists of billion of RDF triples, in several domains. The DBpedia knowledge base <http://wiki.dbpedia.org/About> is a central hub of interlinking of the emerging data web, which extracts structured information from Wikipedia and makes them available on the web. In this work, some instances of the classes of knowledge base are associated with definitions from DBpedia, allowing semantic interoperability.

The modularization involves identifying one or more modules in ontology. A module is considered as one significant and self-contained part of ontology. Although there is no universal way for modularization, the choosing of a particular technique must be guided by application requirements (Suárez-Figueroa et al, 2012). The next section describes four use cases related with time series provenance.

4 TIME SERIES PROVENANCE

There are many definitions for provenance in literature (Moreau, 2010). Usually, it refers to origin or source of something. The knowledge about the data source in domains where the volume of data that need to be analyzed is excessive, it contributes to prove the correctness of resultant data, enabling the understanding about how the data were generated. Tan (2007) comments that provenance information is considered as important as the result itself.

Besides of ways *coarse-grain* and *fine-grain* to generate provenance (Tan, 2007), another approach to provenance considers the use of semantic information based on ontologies, modeling concepts and relationships used in the generation of provenance, contributing with inferences to discovering of implicit knowledge by means of languages as RDF and OWL. This approach provides as advantages the semantic description of the context, improvements in queries and proofs of origin and looking of interoperability of generated data (Moreau et al, 2011).

In this work, time series provenance contributes to the researcher to get information about the origin and other characteristics of time series, helping and facilitating the decision making about the use of appropriate statistics procedures. In the following, are presented four use cases involving real time series data, which are related in Section 5.

- *Use Case 1:* in this case, it is important for the researcher to known, by instance, if discrete nonstationary time series presenting extreme events such as outliers (Hair et al, 2010). This information contributes to choose of an appropriate statistic method for use, as a robust method (Chandler and Scott, 2011).

- *Use Case 2:* another case is when the researcher can identify the appropriate statistical technique considering the observation type (regularly or irregularly sampled). For analysis, the autocorrelation function needs multiple pairs of observations to quantify the serial dependence. In this situation, the time series need to be regularly spaced and with little bit of missing data. This approach is not appropriate for observations in intervals highly irregular, needing alternatives techniques (Chandler and Scott, 2011).

- *Use Case 3:* also it is important to the researcher to know the generator process and in which statistic measure occurs the nonstationarity of the time series, that is, in the mean or in the variance. The nonstationarity in the mean can be removed by differencing, for instance. However, in inhomogeneous time series in the variance (heteroscedastic time series), to reduce nonstationarity, other transformations in the data are needed (Wei, 2006).

- *Use Case 4:* another case is about the time series decomposition model. This information contributes to researcher to choose appropriate statistical methods to trend removal of the time series. For instance, if the time series were decomposed additively (Yaffee and McGee, 2000), the estimated trend is subtracted from data. To a multiplicative decomposition model, this is done by the division of time series by values of trend. In the next section the ontology for time series provenance is presented, developed as an OWL module.

5 ONTOLOGY FOR TIME SERIES PROVENANCE

The methodology for development of domain ontology is based on Ontology Development 101 (Noy et al., 2001). In parallel to this classic methodology, the modularization of the ontologies was considered (Suárez-Figueroa et al, 2012), where ontologies are developed in separated parts, from the self-contained form, being important to a sub-domain or task, allowing scalability. Applied to Ontology Engineering, modularity is central for reducing the complexity of understanding and maintenance, querying and reasoning over modules (Kutz and Hois, 2012).

The modular design describes Time Series Ontology (namespace tso:) related to time series provenance and Detrend Ontology (namespace do:) which describes statistical methods for trend estimation. These modules (ontologies) are used in the Detrend Provenance Model (namespace dpm:) (De Souza et al, 2014) that reuse and extend the Open Provenance Model - OPM (namespace opmo:) (Moreau et al, 2011) as means of generating semantic knowledge about detrending time series. The Detrend Ontology and the Detrend Provenance Model are not described in this paper.

It was identified, along with experts, a set of competence questions that the ontology should be able to answer, involving intrinsic features about time series data and its components. They were identified based on conceptual W7 Model (Ram and Liu, 2009), which contributes to define, capture and to use data provenance, presenting seven inter-connected elements: *What?*, *When?*, *Where?*, *How?*, *Who?*, *Which?* and *Why?*. These elements can be used to track events that affect the data during its lifetime. This provenance model is general and extensible for capture provenance semantics for data in different domains (Ram and Liu, 2009).

From these questions, the classes and its relationships were identified, as well as the instances. Restrictions on the classes and relations are declared using axioms and/or rules, providing semantics and allowing inferences by a reasoner in the knowledge base. The elements from ontologies are represented in this paper between parentheses.

The reuse from ontologies SWEET utilizes namespaces of its sub-ontologies. For instance, the subclass (phen:StochasticProcess) is declared as a disjoint subclass of (tso:NonStochasticProcess). The object property (rela:hasPhenomena) is reused and (tso:hasDynamicalPhenomena) was created and extended in (tso:hasStochasticProcess) and (tso:hasNonStochasticProcess) which are declared as disjoint properties.

The Time Series Ontology describes time series related to nonstationary processes, which presenting trends. These time series are the rule and are not the exception in several application domains.

In relation to scope, are not included statistical methods to transform the time series, which are modeled in detrend module. About scalability, on the one hand, the ontology is extensible due to reuse of the triples from set of Ontologies SWEET, which can be extended based on these ontologies.

The Classes Diagram developed in Ontograf Plugin from Protégé 4.1, presents the main classes related with the class (tso:TimeSeriesData) and the respective provenance elements of W7 Model (Figure 1). The classes represent the main definitions and features of time series, including: processes that generated them and time related; analysis type associated and assumptions considered; observation type done and how they can be classified according to assumptions, knowledge domain, collection, scientific instrument or generator software related; models and types of decomposition of time series and its components, as well as the event component and the mathematical property associated.

All classes are noted by means of the tag (rdfs:comment), identifying which is the source of the definition. This contributes to the understanding of the concepts, as well as allows us to known which is the provenance of the definitions. The data properties and object properties are also noted by means of this tag. Also the tag (rdfs:label) is used for labeling the elements from this module.

The association of instances with DBpedia resources allow, besides knowing its provenance, obtaining more information about the data. This way, it is possible to obtain semantic interoperability about such concepts with LOD.

The ontology presents a classification as the assumptions declared by means of defined rules. Below some rules are presented that infers knowledge about the time series.

```
tso:TimeSeriesData(?x),
tso:hasObservationType(?x,
tso:Regularly_Spaced) ->
tso:hasTimeSeriesAssumption(?x
, tso:Homogeneity)                                   (1)
```

```
tso:TimeSeriesData(?x),
tso:hasTimeSeriesAssumption(?x
, tso:Homogeneity) ->
tso:hasObservationType(?x,
tso:Regularly_Spaced)                                (2)
```

```
tso:TimeSeriesData(?x),
tso:hasTimeSeriesAssumption(?x
, tso:Homogeneity) ->
tso:HomogeneousTimeSeries(?x)                        (3)
```

```
tso:TimeSeriesData(?x),
tso:hasTrendType(?x,
tso:Deterministic_Trend) ->
tso:NonStationaryTimeSeries(?x
)                                                    (4)
```

The two first rules are related with the observation type of time series (regularly or irregularly spaced) and its classification, as the type of time series related. If the observation is regularly spaced, the time series are declared as presenting the Homogeneity Assumption.

Also if the researcher declares that the time series presenting the Homogeneity Assumption, they are inferred how regularly spaced time series. The opposite also occurs, when the observation is irregularly spaced, the time series are inferred as presenting the Heterogeneity Assumption.

According to third rule, if time series present the Homogeneity Assumption, they are classified as being of the type Homogeneous, belonging to class (tso:HomogeneousTimeSeries). The same occurs with the Heterogeneity Assumption.

When the time series presenting some type of trend, for instance deterministic, according to fourth rule, the same are inferred in the class (tso:NonStationaryTimeSeries), where the trend

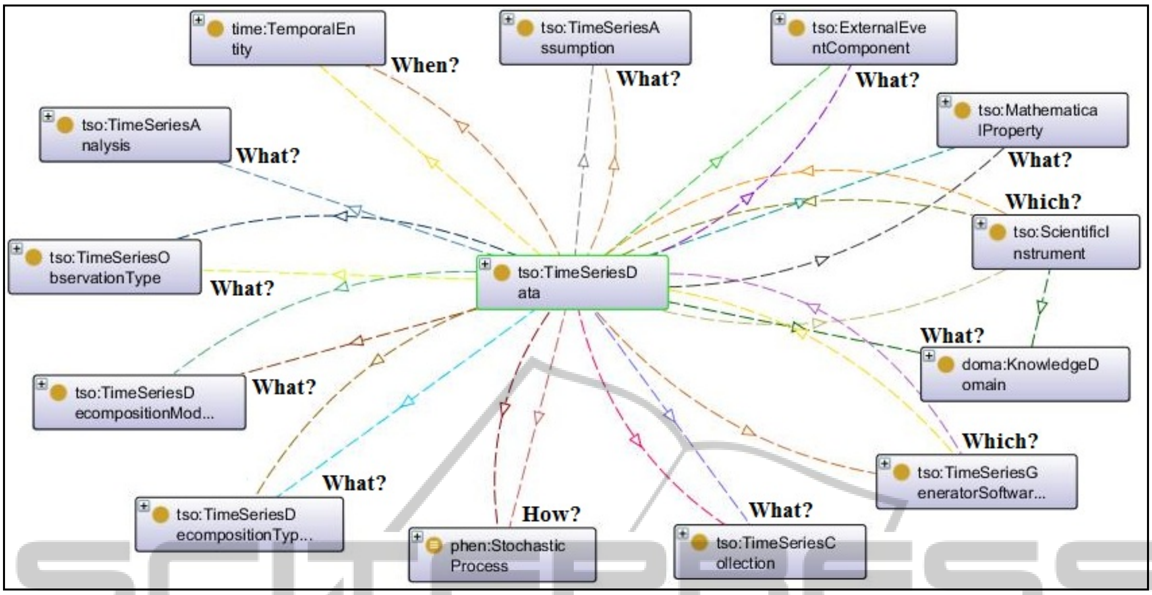

Figure 1: TimeSeriesData Class and relationships with conceptual elements of the W7 Model.

component needs be extracted because it can hide other phenomena, as well as the most statistical methods are developed for stationary time series.

Figure 2 shows the extension from SWEET Ontology related with the class (phen:StochasticProcess), extended with the classes (tso:NonStationaryProcess), (tso:StationaryProcess) and its respective subclasses. According to a defined rule, if the nonstationarity occurs in the mean and the generator process of time series is considered (tso: Stationary_Difference_Process), the inference done is (tso:Stochastic_Trend). In the same way, if the generator process of time series is (tso:Stationary_Trend_Process), the inference upon the trend type related is (tso:Deterministic_Trend), done by Pellet reasoner (Sirin et al, 2007).

The information about the trend type is relevant for the researcher to choose appropriate statistical methods for detrending time series. In Figure 2(a), the trend can be removed by Differencing, where the same can be subtracted from time series. And in Figure 2(b), the trend can be estimated by adjusting of deterministic method as Regression Analysis that fits a model to the trend, which is subsequently removed from them.

Considering the Uses Cases (Section 4), Figure 3 presents a query about Use Case 1, showing which nonstationary time series and its mathematical property presenting some extreme event component. Such knowledge contributes for the researcher to choose statistical methods specifically developed for these features as a robust method. The association of the instances with DBpedia allows getting more knowledge, besides contributing to semantic interoperability.

About Use Case 2, Figure 4 shows features of time series such as observation type, observation interval and missing data percent. This knowledge allows quantifying the autocorrelation function that measures the dependence among sucessive observations.

In Use Case 3 (Figure 5), the knowledge about in which process and statistical measure the nonstationarity occurs allows choosing an appropriate method to extract it. About the Use Case 4 (Figure 6) describes about decomposition model, bringing knowledge about the way as the trend can be removed, in this case, by subtraction due to additive decomposition model.

The ontology was evaluated by Ontologists and Experts of time series area. For evaluation's applicability, were developed the following documents: commitment term, describing the purpose of the evaluation and about ethical questions; list of competence questions; feedback from evaluators about nomenclature, sources of definitions and concordance in relation to the reuse. The feedback from evaluators was analyzed and considered in this module.

The semantic knowledge about time series provenance, contributes meaningfully with the analysis process, as the choice of appropriate statistical methods that considers its characteristics.

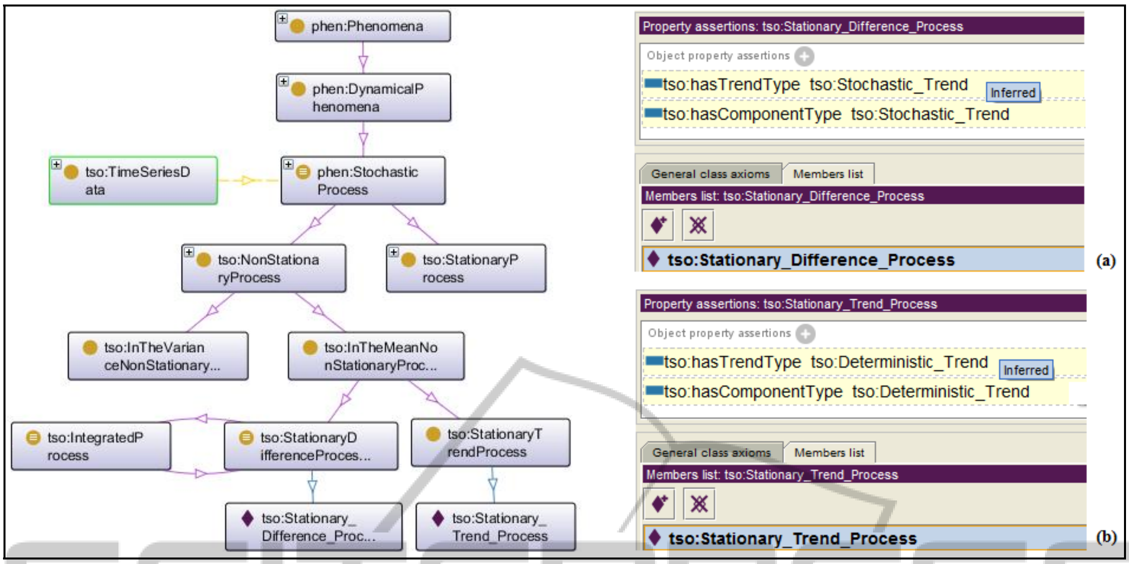

Figure 2: TimeSeriesData class and its relationship with the phen:StochasticProcess class.

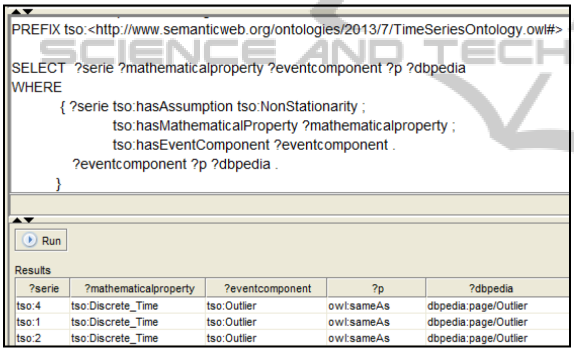

Figure 3: Features of nonstationary time series showing extreme events, associated with DBpedia.

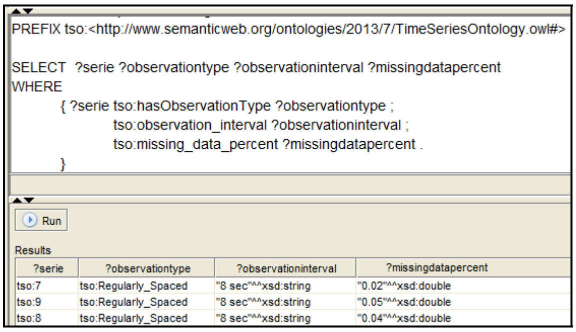

Figure 4: Observation type, missing data percent and observation interval of time series.

6 RELATED WORKS

Henson et al (2009) presents an ontological representation of time series observations for

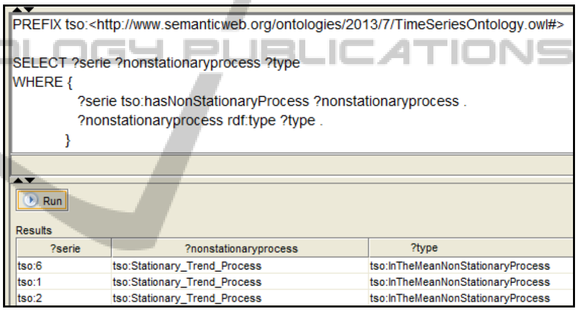

Figure 5: Generator process and which statistical measure the nonstationarity occurs in time series.

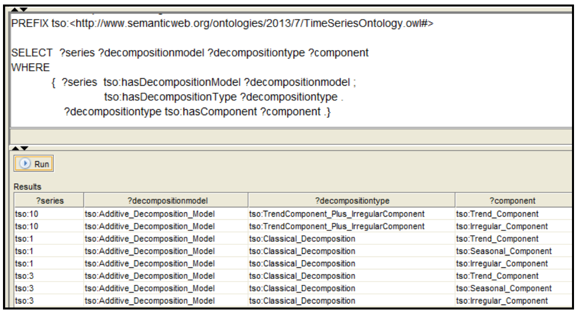

Figure 6: Model and type of Decomposition and components related.

Semantic Sensor Web. It is described as time series observations can be modeled in ontology, in order to solve problems related to integration and queries. This work is related to the modeling of observations of time series from O&M XML Model. It is presented as OWL allows data restrictions better than XML, and also promotes semantic interoperability. Nevertheless, this ontology describes observations and measurements and does not model intrinsic characteristics of time series.

The concept of semantic time series is defined by Bozic (2011), where technologies from Semantic Web are combined with time series processing models, making possible its use in new applications. However, this work proposes a generic language of processing for simulation and modeling of the semantic time series.

Other two works of the same author are Bozic and Winiwarter (2013) and (Bozic and Winiwarter, 2012). The first is an extension of the second, presenting a showcase about semantic time series process. The showcase presents the functionality of the Time Series Semantic Language (TSSL), demonstrating as this technology can improve the time series processing by the usage of a dedicated language in a community building. This work shows high practical impact in the time series process, giving in new data source for applications of semantic web.

Bozic and Winiwarter (2012) propose Time Series Semantic Language – TSSL, a generic community building language for semantic time series, allowing observing data flux as data sensor with additional information, tagging postings of scientists with a specific search topic. The TSSL architecture supports high level of expressivity, user-friendly syntax, extensibility, allowing significant data models. It presents the Time Series Processor - TSP, which coordinates all workflow process of time series. The main contribution of this paper is the semantic time series processing, related to time series data and also its meaning, creating new information by means of links among different data structures.

Comparing with related works, the same do not describe intrinsic features of time series data. The OWL module proposed in this work can be extended and contributes to the researcher to know and understanding better the data (time series), facilitating the decision making about as turn them stationary, improving semantically an important step of preprocessing phase of time series analysis that is the trend extraction.

7 CONCLUSIONS

The scientific contribution of this paper presents two aspects. First, about ontology engineering, is presented a study of case related to modular design of ontologies, presenting the development of a module related to time series provenance, developed by separated way of related statistical procedures.

In this case, the modularity decrease the modeling complexity, facilitate individual evaluation, promoting the reuse and extensibility of modules. Another contribution of this nature is the generation of provenance information for a special type of data, characterized by showing temporal dependence of observations. Also, the definition of the competence questions using expressions of the W7 model, contribute to obtain knowledge about time series provenance.

Second, the contribution that stands out is in time series area, enriching them semantically, improving the analysis process, facilitating the decision making about appropriate statistical procedures and contributing for the scientific knowledge generation.

The applicability of proposal domain ontology is related with nonstationary time series, which presenting trends. In development of ontology, was considered the orthogonality of concepts, where the same are decomposing in its components parts, facilitating its extensibility.

The proposal ontology presents as main advantages: the generation of provenance information about time series; the reuse of statements from set of Ontologies SWEET, allowing semantic interoperability and extensibility; the definition of classes using nomenclature from bibliographies of time series analysis, contributing to the understanding of concepts, which can be visualized by an online documentation developed; the association, when applicable, of the instances with DBpedia enables to increase the ontology definitions and contributes for semantic interoperability with linked open data; and the use of rules allows to infer more semantic knowledge. The ontology also contributes to the choice of the appropriate statistical methods, facilitating the decision making in detrending step.

With the modular development is necessary to select subject for module composition. In this work, data and methods are considered as separated modules, which are combined in the detrend provenance model. Also, in this context, stands out that the wide range of time series analysis area turns it difficult the understanding about concepts and its relationships.

The main contribution of this paper is based on intersection of the following key topic of search: time series and its components, provenance, ontologies and the semantic web, resulting in the generation of domain ontology for time series provenance, as a means to enrich them semantically, allowing logic inferences and the development of queries.

As future works stands out the development of an online environment using the proposal ontology, contributing to facilitate and enrich semantically the trend extraction step of preprocessing phase of the time series analysis.

REFERENCES

Berners-Lee, T., Hendler, J., Lassila, O., 2001. The semantic web. *Scientific American*, 284(5):34–43.

Borst, W. N., 1997. *Construction of Engineering Ontologies*. Thesis. University of Tweenty Centre for Telematica and Information Technology, Enschede, Nederland.

Bozic, B., 2011. Simulation and Modeling of Semantically Enriched Time Series. *19th International Congress on Modeling and Simulation*, Perth, Australia.

Bozic, B., Winiwarter, W., 2012. Community Building Based on Semantic Time Series. *iiWAS* 2012:213-222.

Bozic B., Winiwarter, W., 2013. A Showcase of Semantic Time Series Processing. *IJWIS*. Volume 9. Number 2.

Chandler, R., Scott, M., 2011. *Statistical Methods for Trend Detection and Analysis In the Environmental Sciences*. First Edition. John Wiley & Sons, Ltd.

Cryer, J. D., Chan, Kung-Sik, 2008. *Time Series Analysis With Applications in R.* Second Edition. Springer.

De Souza, L., Vaz, M. S. M. G., Sunye, M. S., 2014. Modular Development of Ontologies for Provenance in Detrending Time Series. Accept for presentation in *ITNG* 2014, 7-9 April. Nevada, Las Vegas, EUA.

Guarino, N., 1998. Formal Ontology and Information Systems. Nicola Guarino, editor, *Proceedings of FOIS'98*, Trento, Italy, 6-8 June 1998, pages 3–15, Amsterdam. IOS Press.

Hair, J. F. Jr., Black, W.C., Babin, B.J., Anderson, R. E., 2010. *Multivariate Data Analysis*. 7th edition. Pearson Prentice Hall.

Hebeler, J., Fisher, M., Blace, R., Perez-Lopez, A., Dean, M., 2009. *Semantic Web Programming*. John Wiley & Sons Inc., Chichester, West Sussex, Hoboken, NJ.

Henson, Cory, Neuhaus, H., Sheth, A., P., Thirunarayan, K., Buyya, R., 2009. An Ontological Representation of Time Series Observations on the Semantic Sensor Web. *ESWC* 2009. Herkalion, Greece.

Kiryakov, A., 2006. Ontologies for Knowledge Management. DAVIES, J. et al. (Eds). *Semantic Web Technologies: trends and research in ontology-based systems*, pages 115–138.

Kutz, O., Hois, J. 2012. Modularity in ontologies. Guest Editorial. *Applied Ontology*. 7, 109-112. IOS Press.

Lange, C., 2013. Ontologies and Languages for Representing Mathematical Knowledge on the Semantic Web. *Semantic Web*, vol. 4, nr 2, pages 119-158.

Meinl, T., 2011. *A Novel Wavelet Based Approach for Time Series Data Analysis*. PhD Thesis. Ref. Prof. Dr. Svetlozar Rachev. Karlsruhe.

Montesino-Pouzols, F., Lendasse, A., 2010. Effect of Different Detrending Approaches on Computational Intelligence Models of Time Series. *IJCNN*, pages 1–8. IEEE.

Moreau, Luc, 2010. The Foundations for Provenance on the Web. *Foundations and Trends in Web Science*. Volume 2, Numbers 2-3 (2010). Pages 99-241.

Moreau, L., Clifford, B., Freire, J. Futrelle, J. Gil, Y., Groth, P., Kwasnikowska, N., Miles, S., Missier, P., Myers, J., Plale, B., Simmhan, Y., Stephan, E., den Bussche, J. V., 2011. The Open Provenance Model Core Specification (v1.1). *Future Gener. Comput. Syst.*, 27(6):743–756.

Noy, N. F., McGuinness, D, L., 2001. Ontology Development 101: A Guide to Creating Your First Ontology. *Development*, 32(1):1–25.

Ram, S., Liu, J., 2009. A New Perspective on Semantics of Data Provenance. *SWPM*, 2009.

Sirin, E., Parsia, B., Grau, B.C., Kalyanpur, A., Katz, Y. 2007. Pellet: A practical OWL-DL Reasoner. *Web Semant.*, 5(2):51–53.

Suárez-Figueroa, M. C., Gómez-Pérez, A., Motta, E., Gangemi, A. 2012. *Ontology Engineering in a Networked World*. Berlin, Springer.

Tan, W.C., 2007. Provenance in Databases: Past, Current, and Future. *IEEE Data E. Bull.*, 30(4):3-12.

Wei, W. S. W., 2006. *Time Series Analysis. Univariate and Multivariate Methods*, 2nd edition. Pearson Education.

Wu, Z., Huang, N. E., Long, S. R., Peng, Chung-Kang, 2007. On the Trend, Detrending, and Variability of Nonlinear and Nonstationary Time Series. *Proc. of the National Academy of Sciences*, 104(38):14889–14894.

Yaffee, R. A., McGee, M., 2000. *Introduction to Time Series Analysis and Forecasting with Applications of SAS and SPSS.* Acad. Press, Inc. Orlando, FL, USA.