International Journal of Future Generation Communication and Networking
Vol. 9, No. 6 (2016), pp. 263-274
http://dx.doi.org/10.14257/ijfgcn.2016.9.6.25

WSN Missing Data Imputing Based on Multiple Time Granularity

Jianfeng Xu [1, 2], Yu Li [3, 2], Yuanjian Zhang [1, 2] and Azhar Mahmood[2]

[1]*Software and Information Engineering School, Tongji University, Shanghai, 201804, China*
[2]*Software school, Nanchang University, Nanchang, Jiangxi, 330047, China*
[3]*North Automatic Control Technology 1nstitute, Taiyuan, Shanxi, 030006, China*
jianfeng.nc@gmail.com, ly4232@hotmail.com
96zhangyj@tongji.edu.cn, azharmahmood8@gmail.com

Abstract

Missing data is a common phenomenon in the data collection process of wireless sensor network (WSN), and the missing data imputing is an important issue of WSN stream data mining. Currently WSN missing data imputing method has little considered about the dynamic characteristics of internal data time structure during the data collection process, which makes data imputing difficult to reflect the real monitoring change objectively. In order to analyze the internal structure and dynamics of WSN time sequence data systematically, with the equivalence relation of the monitored object the time domain can be regarded as a series of integral time granule (ie atomic time point set), a wireless sensor network timing information system (WTIS) is established. The system can reason logically at different time granularity, and a multiple optimal time granularity strategy of WTIS based on hierarchical successive approximation approach is proposed. Finally, based on the research, a multiple optimal time granularity WSN missing data clustering imputing algorithm is proposed. Compared with traditional fixed time granularity missing data imputing algorithm, experiments show that the algorithm can lower error rate when imputing WSN missing data.

***Keywords**: WSN, time series, granularity, clustering*

1. Introduction

In wireless sensor networks, each sensor can be viewed as a data source, producing large amounts of real-time data continuously. The WSN data is a typical time series data [1]. However, due to the flawed environment such as hardware breakdown and transmission congestion, the data quality cannot be guaranteed. Missing data is a common scenario and it hampers the efficiency of data processing.

Currently the missing value imputing algorithm mainly relies on classical machine learning method. WSN missing data imputing methods are mainly based on statistics [2-4], association rules [5-6] and clustering [7-9] algorithm. Statistics method mainly obtains statistical data set through data analysis, and then uses this information to deal with missing values. The most common method is the average imputing method [10]. The main idea of association rules method is to generate frequent item sets which meet certain degree of confidence and support, with this knowledge the missing data is deduced. The main idea of clustering imputing method is to remove the default excessive missing value record initially. For the rest parts, using complete sample as training set, the missing sample as testing set, after training and getting the cluster model, the missing value can be imputed. For example, Zaifei Liao [11] presents a fuzzy k-means clustering algorithm over

ISSN: 2233-7857 IJFGCN

sliding window for the missing value imputation of incomplete data to improve the data quality. However, the aforementioned study mainly focuses on the optimization of methods only, and there is no further discussion for WSN time series data model. General missing data preprocessing methods considered little about temporal properties. The accuracy of data will be improved if the imputed result considered the property of WSN.

This paper is organized as follows, chapter one describes WSN time series data and missing data imputing research. Chapter two discusses WSN time granularity acquisition strategy based on WSN time series information systems. The WSN missing data imputing applications and experimental analysis is described in chapter three. Finally, the full text is concluded.

2. WSN Time Series Data Modeling

WSN time series data stream can be expressed by a series of binary pair $\{(t, f_a(t))\}$, where t represents time, the function $f_a(t)$ represents the value of common attribute a at the time t. In WSN time series data studies, time domain U_T is generally considered as an end point ray, which is isomorphic to axis R^+. Thus, it can also be denoted as T $\{t_0, t_1, t_2, ...\}$

2.1. Time Granulation and WSN Time Series Information System Modeling

Definition 1 :(Atomic time) The measure of indivisible minimum time discretization unit on the time axis T (ie U_T). It is indicated by symbol t_i. where $i \in Z$.

Definition 2 :(Length of time) The size of a period of time $U_{t_i \leftrightarrow t_j}$ that measured by the number of contained atomic time $|U_{t_i \leftrightarrow t_j}|$.

Definition 3 :(Time granularity attribute) The time attribute which can be equivalent to the time domain which is divided according to the length of specified time, denoted by A_T.

Time domain can divide the time into the same time length according to the length of time (ie the length of a specified time) (the quotient set).

For example, time domain with the range of one year can specify the length of time like the month or day. One year can be divided into 12 months or 365 days.

Definition 4 :(Time granularity) Specific time granularity attribute values (ie specified length of time) is called time granularity, denoted by a_{ti}.

The range of time granularity attribute A_T is denoted by $A_T = \{a_{t1}, a_{t2}, ...,a_{tn}\}$. The granularity attribute a_{ti} is i_{th} time granularity of the time domain. The atomic time granularity is recorded as a_{t1}.

Assuming atomic time granularity is one second, for the next one minute (60 seconds), the time granularity attribute A_T value range in descending order can be described as A_T = {60 seconds, and 59 seconds, ..., ..., 10 seconds, 9 seconds, ... 1 second}.Clearly, the time domain by the N atoms has at most N kinds of different time granularity attribute values

Definition 5 :(Time granule) obtained from each period (time domain) of the time granularity attribute, also known as a time granule, denoted as $*T_{(t_s, a_{ti})}$. t_s is the starting time of the time granule, a_{ti} is the length of the time granule.

Definition 5 shows that time is from time domain, time domain U can be viewed as having a coarse granule of the time.

A_N is common attributes of classical information systems[12] and A_T is time granularity attribute in the time series related attribute set, denoted as A, that is $A = A_T \cup A_N$.

Definition 6: (WSN time information function) In the time series data analysis, the mapping between time granule * T and the granule attribute A_N under a common characteristic values, denoted by $g:(A_N,{*T}_{(t_s,a_{ti})})\to V$,or simply denoted by $g:({*T}_{(t_s,a_{ti})})\to V$ or g. Common characteristic values getting method of g include multi-section site of time series [12], periodically [13], the trend [14].

As is shown in the Figure 1, its abscissa is a time axis, the values of vertical axis locates in the range of $A_{N,}$ any point on the curve TS represent the time point corresponding to A_N, according to the value ati of A_T divides the time domain into three time granules ${*T}_{(t_0,a_{ti})}, {*T}_{(t_i,a_{ti})}, {*T}_{(t_{2i},a_{ti})}$.If each extreme point of time granule (triangle symbol in Figure 1) is selected as the characteristic information, the corresponding feature sequences can be denoted as, $g:(A_N,{*T}_{(t_0,a_{ti})})=\{v_1,v_2,v_3\}$, $g:(A_N,{*T}_{(t_i,a_{ti})})=\{v_4,v_5,v_6\}$, $g:(A_N,{*T}_{(t_{2i},a_{ti})})=\{v_7,v_8,v_9\}$,respectively, T

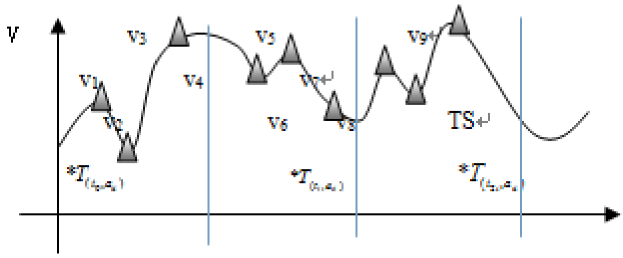

Figure 1. Multi-Section Site of Time Series Time information Function

Generally, researchers construct the corresponding function g according to different requirements. The results obtained by g are substitutes into corresponding data analysis methods, such as time series similarity calculations, association rules, clustering and classification. These are WSN time series data analysis methods, as defined in Definition 7.

Definition 7 :(WSN Time series data analysis method) A method which is used for WSN time series data analysis based on time attribute a_{ti} in given time domain U_T, denoted as $F_{U_T}\left(g\left(A_N,{*T}_{(t_r,a_{ti})}\right),g\left(A_N,{*T}_{(t_{r+a_{ti}},a_{ti})}\right),\cdots,g\left(A_N,{*T}_{(t_{r+K\times a_{ti}},a_{ti})}\right)\right)$, abbreviated as $F(g(t_0,a_{ti})^K)$, where K = 1,2, ...represents the number of time granule a_{ti} which U_T contains.

It can be observed from definition 7, the shape of * T is the most important factor in time series data analysis method when common attribute A_N and WSN time information function g is given. However, the current WSN time series data analysis research pays more attention to optimize the F, whereas there is little consideration about * T in time series data analysis effects. This paper argues that the study of WSN time series data analysis includes time domain U_T, time granule * T, time granularity attribute A_T, common attributes A_N, time information function g. It is necessary to discuss from the perspective of information systems.

Definition 8: WTIS (WSN Time Information System) WTIS = (U_T, A, n, V, g, f) where U_T is the time domain, A = $A_N \cup A_T$ representative the set of attributes, of which A_N is the common attribute, A_T is the time granularity attribute, n is the number of sensors, which is the number of time series. In time granularity attribute A_T = {a_{t1}, a_{t2}, ...,a_{tn}} ,different values forms different U_T equivalent division granularity. f is a mapping function, $f:A\times U_T\to V$,V is the set of values according to atomic time, the time granular characteristic value can be obtained from WSN time information function $g\left({*T}_{(t_r,a_{ti})}\right)$.

The following WTIS table shows the situation in different time granularity for time domain U_T.

Table 1. WTIS Table

<table>
<tr><td></td><td colspan="6">U_T
t_0 t_1 t_2 t_3 t_4 t_5 t_6 T</td></tr>
<tr><td>a_{t1}</td><td>$*T$</td><td>$*T$</td><td>$*T$</td><td>$*T$</td><td>$*T$</td><td>$*T$</td></tr>
<tr><td>a_{t2}</td><td colspan="2">$*T$</td><td colspan="2">$*T$</td><td colspan="2">$*T$</td></tr>
<tr><td>a_{t3}</td><td colspan="3">$*T$</td><td colspan="3">------</td></tr>
<tr><td>...</td><td colspan="6">------</td></tr>
</table>

In Table 1, the first row represents the time domain U_T, where t_i represents atomic granule. t_0 is the starting point, the left of starting point is not discussed.

In table 1 the left column represents different granularity attribute set $\{a_1, a_2, ...\}$, a_{t1} has one atomic granule, a_{t2} has two atomic granules, a_{t3} has three atomic granules. The size of the atomic granules is depending on the application demands.

In table 1,A_{ti},i = 1,2,3 ...... Different row $\{*T_{(t_i,a_{ti})}\}$ represents different granularity in the U_T. Each small box $*T_{(t_i,a_{ti})}$ represents a time granule.

2.2. The Optimal Time Granularity based on WTIS

2.2.1. WTIS Optimal Time Granularity

As can be observed from the structure $F_{U_{t_a \leftrightarrow t_b}}\left(g(t_a, a_{ti})^K\right)$, the selection of the time granularity have an important influence on the F when other factors are constant. Therefore we can use F (g) to present a novel perspective of the optimal time granularity, described in definition 9, definition 10 and definition 11 respectively.

Definition 9: (Optimal time granularity) In a certain time domain, a granule *T is applied to WSN granular data analysis function F, through the testing it can get the optimal effect. Then this *T is defined the optimal time granularity.

Definition 10: (Single optimal time granularity WTIS) The WTIS based on the same optimal time granularity is defined single optimal time granularity WTIS.

Many scholars used different time granularity [15-16] to game and get the optimal time granularity. The idea is as follows: In WTIS, selecting different time granularity and comparing the effects of different time granularity according to the same method F, the best data analysis effect is the optimal time granularity.

Definition 11: (Multiple optimal time granularity WTIS) In some WTIS, there are many different subdomains for the whole domain; each subdomain has its optimal time granularity. It is defined multiple optimal time granularity WTIS.

In order to get better data analysis effect, the time granularity need to be adjusted according to the applications flexibly; it can be explained by multiple optimal granularities. For example, in oil drilling, different geological structures, it needs to choose different time window to analyze sample values at different drilling depths.

2.3.2. Multiple Optimal Time Granularity Acquisition Strategy

For WSN time series with complex features, multiple optimal time granularity problem is an important issue to solve.

So multiple optimal time granularity WTIS can be described as follows: Use time domain U_T, time series data analysis methods F and time granularity information function g, to find a set of time subdomains $U_{t_a \leftrightarrow t_b}$, to make F achieve the desired effect λ.

Multiple optimal time granularity algorithm is described as follows:

Algorithm 1 GenerateTimeDomain
//Input: Time Domain(Granular) U, Candidate Equivalent Division D
//Output: K Time subdomain
{TD_1,TD_2,……,TD_j}=**GenerateTimeDomain**(U,D)
1. cur=U_begini=1,
2.**while**(cur<U_end)
3. $TD_i = {}^*T_{(cur,a_{tD})}$
4. i=i+1 cur=cur+D;
5. **end while**
6. **return** {TD_1,TD_2,……,TD_j}

Algorithm 2 FindOptimalGranularity
//Input: Time Domain (Granular) $U_{T,}$ Candidate Equivalent Division D, WSN time information function g, WSN Time series data analysis method
(U_t, $d_k \in D$, *flag*)=**FindOptimalGranularity**(U_t,D,g,F)
1. index=0 flag=false temp=0
2.**for** each d_i in D do
3.{TD_1,TD_2,……,TD_j}=**GenerateTimeDomain(U_t,d_i)**
4. **if**($AVG\left(F(g(TD_j)\right) \le \lambda,\ j=1,2,\cdots$)
5. flag=true; **break**;
6. **else**
7. **Max(** $F(g(TD_j)),\ j=1,2,\ldots$) ; index=i;
8 **end if else**
9. **end for**
10.**if**(flag)
11.append (U_t,D_{index}) to TR
12.**end if**
13.**return** (U_t,D_{index},flag)

Algorithm 3: MOTG(short for Multiple Optimal time granularity)
Input: time domain U_T; time series data mining method F; the time granularity information function g; the candidate equivalent set D = {d_1, d_2, ... ,, d_n} (1 = d_1<d_2<...... <d_n); threshold λ of time series data mining evaluation F
Output: Multiple optimal time granularity WTIS of selected time domain, $TR=\{(U_j,d_k)|U_j \subseteq U_T \wedge d_k \in D \wedge j,k=1,2,\cdots\}$
1. **if** (!flag)// initialize domain
2. $\{TD_j|TD_j \subseteq U_T \wedge j=1,2,\cdots\}$=**GenerateTimeDomain**($U_T, d_k$)
3.append ($\{TD_j|TD_j \subseteq U_T \wedge j=1,2,\cdots\}$) toG// G is a candidate solution set
4. **end if**
5. **if**(!flag)
6. **while(**(Num=|G|)>0) // presence of the candidate domain
7. **for** i=1 to Num do
8. (G_i, $d_k \in D$, *flag*) =**FindOptimalGranularity**(G_i,D,g,F)
9. **if** (!flag)// to find the optimal granularity by loop iterations
10. $\{TD_j|TD_j \subseteq G_i \wedge j=1,2,\cdots\}$=**GenerateTimeDomain**($G_i$, d_k)
11. append $\{TD_j|TD_j \subseteq G_i \wedge j=1,2,\cdots\}$ to G
12. **else**
13. append (G_i) to TR
14. **endif else**
15. **end for**
16. **end while**
17.**end if**

3. WSN Optimal Time Granularity Acquisition and Missing Data Imputing Applications

3.1. Experiments

3.1.1. Data Sources

The experiment use the WSN data collected by 44 SLRF-GW-WiFi temperature and humidity sensors in Nanchang University Library in July 2014, sampling interval is 5 seconds. Each field of collected data is shown in Table 2.

Table 2. Field Description of WSN Stream Data

Field Name	Type	Length	Field meaning	Description
dataid	Long Integer	50	dataID	
rssi	Text	50	RSSI	signal strength
gm	Text	50	light	
tem	Text	50	Temperature	
hum	Text	50	Humidity	
nodename	Text	50	sensor node name	
intime	Text	50	collected Time	YYYY-MM-DD HH:MM:SS

3.1.2. WSN Missing Data Imputing Algorithm Model

WSN missing data imputing algorithm model is shown in Figure 2.

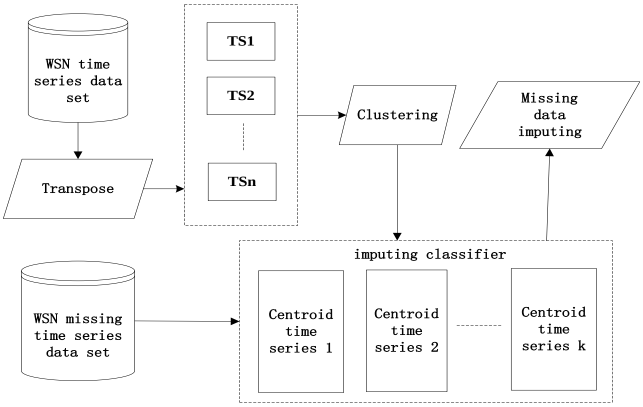

Figure 2. WSN Missing Data Imputing Algorithm Model

WSN missing data imputing algorithm model is shown in Figure 2, the algorithm steps are as follows:

Step 1: First make the collected data into discrete data, then use SPSS tools to transpose into time series table and get data collected by the K sensor nodes in an area;

Step 2: Experiment use the random method to generate missing data, the missing value is set to 0, and the missing data is regarded as testing set, intact data is regarded as training set.

Step 3: For WSN time series data in the training set, use k_means and improved x_means [17] clustering respectively to obtain centroid time series and the type of each time series;

Step 4: Matching the testing set and centroid sequence to find the most similar category;

Step 5: Use different granularity 1, 2, 4, 6(that is 1, 2, 4, 6 aliquots of the time domain, $U_T/1$, $U_T/2$, $U_T/4$, $U_T/6$) time series to perform the experiments; compare the results of the experiments. Imputing missing values using centroid clustering time series of corresponding category, and calculate the imputing accuracy;

Step 6: Evaluate the effect of the testing set data mining. In order to verify the accuracy of the predicted missing data, compare the imputing value with the actual value of the missing position and calculate prediction error. The error rate is calculated as follows:

Error $= |A - B|/B * 100\%$, where A represents the predicted value, B represents the actual value (3.1)

3.2. Analysis of Experimental Results

Experiments are in K-means (K = 3 and K = 7) and X-means (d = 0.1, d = 0.2, d = 0.05, d = 0.5 and d = 1, d is the distance time series under the measurement of given formula, to get different X value) clustering methods with single optimal time granularity experiments in different granularity time series 1,2,4,6, while compared with multiple optimal time granularity experiment and get histogram.

Single optimal time granularity experimental error rate are shown in Table 3 and Table 4:

Table 3. Error Rate of WSN Humidity Data Imputing Algorithm

granularity\average error rate	1	2	4	6
K-means (K=3)	15.83%	21.81%	15.09%	17.12%
K-means (K=7)	15.54%	19.00%	11.29%	12.83%
X-means (d=0.1)	7.51%	13.19%	10.71%	10.20%
X-means (d=0.2)	7.73%	20.80%	11.63%	10.20%
X-means (d=0.05)	4.32%	10.97%	11.09%	9.46%
X-means (d=0.5)	7.54%	20.86%	15.86%	15.92%
X-means (d=1)	18.15%	21.66%	15.86%	15.92%

Table 4. Error Rate of WSN Temperature Data Imputing Algorithm

granularity\average error rate	1	2	4	6
K-means (K=3)	32.76%	32.84%	7.49%	7.91%
K-means (K=7)	32.08%	28.77%	5.67%	4.87%
X-means (d=0.1)	1.99%	29.43%	5.29%	6.40%
X-means (d=0.2)	2.11%	32.51%	6.83%	7.05%
X-means (d=0.05)	2.03%	5.25%	4.74%	5.92%
X-means (d=0.5)	32.34%	29.01%	6.83%	7.05%
X-means (d=1)	32.15%	37.70%	6.83%	7.05%

The experiments select humidity and temperature as two normal attributes to do WSN data error analysis and comparison. For WSN missing data imputing, the choice of different time granularity lead to different experimental results. For different experimental environment (K-means (K = 3 and K = 7) and X-means (d = 0.1, d = 0.2, d = 0.05, d = 0.5 and d = 1), optimal time granularity is different (as is shown in Table 3 and Table 4, underlined data is optimal granularity for current experiment parameters). Additionally, it can be inferred from Table 3 and Table 4 that X-means group can achieve better average accuracy compared to the K-means group relatively. The increase of d value in X-means group can deteriorate the overall accuracy. It is interesting to find out that the change of d is inversely proportional to the thickness of optimal time granularity.

Table 5. Error Rate of Multiple Optimal Time Granularity Imputing Algorithm

granularity\average error rate	humidity	temperature
K-means (K=3)	14.21%	6.735%
K-means (K=7)	9.85%	4.16%
X-means (d=0.1)	3.36%	2.105%
X-means (d=0.2)	4.52%	2.865%
X-means (d=0.05)	3.46%	2.23%
X-means (d=0.5)	7.67%	6.72%
X-means (d=1)	14.49%	6.35%

Multiple optimal time granularity experimental error rate is shown in Table 5. With multiple optimal time granularity algorithm, the better results can be obtained. The comparative experiments are shown in Figure 3 and Figure 4. In different parameters, for the humidity data the error rate of multiple granularity experiment is lower than single granularity experiments, and for the temperature data the error rate of multiple granularity experiment is lower than single granularity experiments except X-means (d = 0.2) and X-means (d = 0.05) two experiment results.

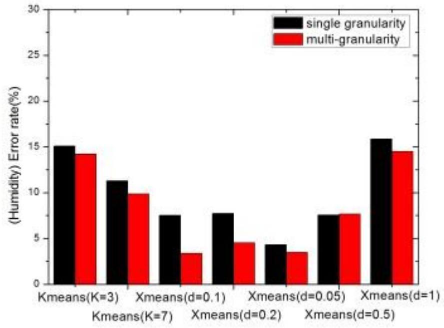

Figure 3. Single and Multiple Optimal Time Granularity Experimental Comparison of Humidity Data

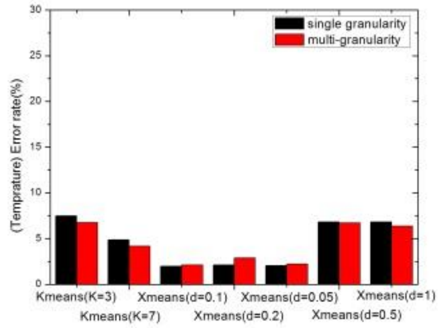

Figure 4. Single and Multiple Optimal Time Granularity Experimental Comparison of Temperature Data

The experimental results show, the selection of time granularity will influence the results to some extent, the thought of granularity not only exists in the number of clusters, the number of time series slices will also affect missing data WSN imputing effect. Multiple optimal time granularity imputing algorithm for optimal WSN time series data structure and dynamics phenomenon can be made to impute missing data at a lower error.

4. Conclusion

This paper deals with time series modeling problem of wireless sensor networks through multiple granularity theory, establishes WTIS model, presents WTIS multiple optimal time granularity acquisition strategies according to WSN stream data application, and applies to WSN missing data imputing algorithm and achieved good results. In the future, the research will further carry out the discussions of dynamic bending, concept drift and the missing data imputing method with statistics, association rules. Currently, experiment has achieved initial success, but there are still some problems like large amount of calculation, not intelligent enough to choose the parameters. Both the experiment and WTIS reasoning models need to be improved.

References

[1] T. C. Fu, "A review on time series data mining", Engineering Applications of Artificial Intelligence, vol. 24, no. 1, **(2011)**, pp. 164–181.
[2] R. J. A. Little and D. B. Rubin, "Statistical analysis with missing data", John Wiley & Sons, **(2014)**.
[3] H. A. E. Sharif, R. S. V. Teegavarapu, "Evaluation of spatial interpolation methods for missing precipitation data: Preservation of Spatial Statistics", Proceedings of the World Environmental and Water Resources Congress, **(2012)**, pp. 3822-3832.
[4] Q. Huang, L. Wei and H. Huo, "Research on independent college teachers' teaching ability based on factor analysis In SPSS", Mathematical Modelling and Engineering Problems, vol. 1, no. 1, **(2014)**, pp. 25-28.
[5] L. Zhou and S. Yau, "Association rule and quantitative association rule mining among infrequent items", Proceedings of the 8th international workshop on Multimedia data mining: (associated with the ACM SIGKDD 2007), ACM, **(2007)**, p. 9.
[6] T. Herawan and M. M. Deri, "A soft set approach for association rules mining", Knowledge-Based Systems, vol. 24, no. 1, **(2011)**, pp. 186-195.
[7] Y. F. Zhang, Y. T. Qian, T. Y. Liu and S. Y. Wu, "Data mining clustering algorithm research and application", Advanced Materials Research, **(2014)**, pp. 926-930.
[8] B. Mirkin, "Clustering: a data recovery approach", CRC Press, **(2012)**.
[9] R. Jin, C. Kou and R. Liu, "Biclustering algorithm of differential co-expression for gene data", Review of Computer Engineer Studies, vol. 1, no. 1, **(2014)**, pp. 7-14.
[10] R. J. A. Little and D. B. Rubin, "The Analysis of Social Science Data with Missing Values", Sociological Methods and Research, vol. 18, **(1990)**, pp. 292-326.
[11] Z. Liao, X. Lu, T. Yang and H. Wang, "Missing Data Imputation: A Fuzzy K-means Clustering Algorithm over Sliding Window", Fuzzy Systems and Knowledge Discovery, 2009, FSKD '09, Sixth International Conference, vol. 3, **(2009)**, pp.133-137.
[12] W. Liu, B. Chen and R. A. Swartz, "Investigation of Time Series Representations and Similarity Measures for Structural Damage Pattern Recognition", The Scientific World Journal, **(2013)**, pp. 248-349.
[13] Y. Q. Zhang and X. Wan, "Statistical fuzzy interval neural networks for currency exchange rate time series prediction", Applied Soft Computing, vol. 7, no. 4, **(2007)**, pp. 1149-1156.
[14] C. Damle and A. Yalcin, "Flood prediction using time series data mining", Journal of Hydrology, vol. 333, no. 2, **(2007)**, pp. 305-316.
[15] R. K. Jain, K. M. Smith, P. J. Culligan and J. E. Taylor, "Forecasting energy consumption of multi-family residential buildings using support vector regression: investigating the impact of temporal and spatial monitoring granularity on performance accuracy", Applied Energy, vol. 123, no. 3, **(2013)**, pp. 168–178.
[16] R. Dong and W. Pedrycz, "A granular time series approach to long-term forecasting and trend forecasting", Physica A: Statistical Mechanics and its Applications, vol. 387, no. 13, **(2008)**, pp. 3253-3270.
[17] D. Pelleg and A. W. Moore, "X-means: Extending K-means with Efficient Estimation of the Number of Clusters", ICML, **(2000)**, pp. 727-734.

Authors

Jianfeng Xu, he was born in 1973. He received the M.S. degree in computer science from Nanchang University in 2006. Now he is an associate professor at Software College of Nanchang University and study at P.H.D in Tongji University. His research interests include granular computing, machine learning, rough set etc. He is a member of IEEE and CCF.

Yu Li, he was born in 1988. Now he is working in north automatic control technology institute. He received software engineering M.S. degree at Nanchang University in 2015. His research interest includes data mining, granular computing, artificial intelligence and machine learning. He has published in Journals like Information Technology Journal and Journal of Chinese Computer Systems.

Yuanjian Zhang, he was born in 1990. He was born in 1990. He received software engineering M.S. degree at Nanchang University, Nanchang, China in 2015. Now he is studying in Computer Science and Technology from Software and Information Engineering School, Tongji University, China as a P.H.D candidate. His research interests include granular computing, machine learning and data mining. He has published in Journals including Journal of Software Engineering, Journal of Nanjing University (Natural Sciences) and Journal of Frontiers of Computer Science and Technology. He is a member of CCF.

Azhar Mahmood, he is an Assistant Professor of Computer S cience Department at SZABIST Islamabad Pakistan. He received his doctorate in Computer Applied Technology from HUST Wu han China. His research interests include Wireless Sensor Netwo rks, Data Mining, Distributed Systems, and published more than 15 articles in various reputed international journals and conferen ces.

