People's Livelihood Appeal Using Big Data Analysis And Mining

Dr.Madhavi Pingili ,Associate Professor & HoD, Department of Information Technology, CMR Engineering College, Hyderabad, Telangana,
E-Mail-id madhavipingili2@gmail.com

Mrs.T.Sarika, Assistant Professor, Department of Information Technology, CMR Engineering College, Hyderabad, Telangana,
E-Mail-id sarika.tuniki@cmrec.ac.in

Abstract - This study presents a time series modelling and technique to partition the time series x(t) of data into long-term change trend L(t), short-term change trend S(t), and occasional change e in order to assess the vast data of people's livelihood appeal (t). Then, using this method, dissect the data of six different livelihood appeals, including unlicensed vendors, industrial noise, sewer cover, academic qualification, out-of-store operation, and public transportation. Combine the data with other data for correlation analysis, identify the underlying cause of the appeal event, and make predictions. The experimental findings demonstrate the value of time series analysis for the analysis of massive data in e-government and for mining the appeal of people's livelihoods.

1. Introduction -

We urgently need to convert data into information and knowledge given the recent rapid expansion in the volume of data in many industries. A statistical approach called time series analysis is used to statistically model a succession of data points in chronological order and identify patterns in data changes [1]. Time series analysis has several uses, including stock analysis and forecasting [2] [3], company sales analysis and forecasting [4] [5], signal analysis [6], traffic flow analysis [7], environmental monitoring [8], aberrant network traffic detection [9], and disease prediction [10], among others. Time series analysis is now used sparingly in big data analysis and mining of government affairs, which is a gap that has to be filled [11]. Traditional government management has struggled to keep up with the demands of the information era [12]. In order to encourage people to reuse or creatively process public data to derive new models and create value, more and more nations around the world are committing to developing a "government data disclosure" strategy [13]. This strategy is intended to help public sector employees make better decisions and increase efficiency. In the management of e-government, big data analysis and mining based on time series analysis are extremely important. This study suggests employing time series modelling and algorithm for large data analysis and mining with a focus on the area of people's livelihood appeal in government issues. The appeal to people's way of life relates to their wants and opinions in relation to all facets of society. This is how the government communicates with the populace. Several local governments focus on the improvement and reform of the appeals procedures, integrate numerous channels for appeals, and provide a standardized platform to manage the enormous data generated by various livelihood

appeals. This study combines time series modelling and algorithms to assess huge data on people's demand for a livelihood across different dimensions in order to respond to current needs and assist service-oriented government decision-making. The time series analysis modelling and algorithm of people's livelihood appeal are introduced in Section 2 of this study. The results of Section 3's detailed analysis of the experimental data on people's appeals for their livelihood are presented. Section 4 concludes by summarizing the research for this study.

2. Modeling and Algorithm Framework

A. Time series additive combination model

The appeal of people's livelihoods is a collection of time-varying sequences, or x(t). Time series changes can be divided into three categories: long-term change trend L(t), short-term change trend S(t), and sporadic change e. (t). An additive combination model is created and expressed by the formula x(t) = L(t) + S(t) + e based on the properties of the sequence (t). There are several factors that influence the short-term change trend S(t), including population changes and dietary patterns. Occasional change e(t) is more pertinent to individual events, such as particular social activities, extreme meteorological phenomena, and significant political events. The long-term change trend L(t) should reflect changes in people's livelihood appeal by month, which is related to economic development, weather changes, etc. It will be easier to assess the source of the occurrence, forecast its tendency, and propose solutions if you look at the pattern of people's livelihood appeal through time. The secret to achieving the aforementioned objectives is to accurately and properly decompose the temporal series of people's livelihood appeal.

B. Time series analysis algorithm design

The steps of the time series analysis algorithm developed in this paper are as follows:

1) Extraction of sporadic change e(t): A sporadic event is made up of a number of outliers that depart from the norm, and it is very valuable for research. In order to distinguish outliers, this study employs the box plot method [14]. The upper bound is designated as QU, and the lower bound as QL. Physically, it means that the majority of the data in all observations fall inside the range [QL, QU]. An outlier is a number in the time series that is either less than or greater than QU + 1.5*(QU - QL). The associated point is the outlier.

2) Extraction of the long-term trend of change L(t): The original sequence has two parts: the long-term change trend L(t) and the short-term change trend S. This sequence is derived by subtracting the occasional change e(t) from the original sequence (t). According to universal consensus, the long-term change trend L(t) ought to be able to capture the non-stationary monthly variation in the appeal of people's means of subsistence. In order to extract the long-term changing trend L, this paper uses the moving average filtering method [15]. (t). The sliding window function's length is 30 points, reflecting the month-by-month change trend.

3) Extraction of the short-term change trend S(t): The original sequence can be stripped of the long-term change trend L(t) and the occasional change trend e(t), resulting in the short-term

change trend S(t). It is commonly accepted that the stationary component of the time series, the short-term changing trend S(t), has a specific periodicity.

3. Experimental and Analysis

A. Experimental data

This paper's data came from an open platform run by a local government. The data collection includes 39,788 public appeal submissions made between February 1, 2018 and May 5, 2019 via WeChat, email, hotline, Weibo, and other methods. The primary data fields' descriptions following deduplication processing are shown in Table 1.

TABLE I. Description of the main data fields

Seq	*Field Name*	*Description*
1	CREATE_TIME	Creation Time
2	DISTRICT_ID	District ID
3	STREET_ID	Street ID
4	COMMUNITY_ID	Community ID
5	EVENT_TYPE_ID	Event Type ID
6	MAIN_TYPE_ID	Main Type ID
7	SUB_TYPE_ID	Sub Type ID
8	INTIME_ARCHIVE_NUM	On Schedule Flag
9	DISPOSE_UNIT_ID	Processing Department ID

We pre-processed the data set, extracted data from it, and compiled the key fields of each appeal, including: event time, event location (including street and community), event kind, processing department, and whether or not to be finished on schedule.

Six primary categories of people's livelihood appeal were identified through statistical analysis: unauthorized vendor, industrial noise, sewer cover, academic qualification, out-of-store operation, and public transit. Figure 1 displays the time series for these six kinds.

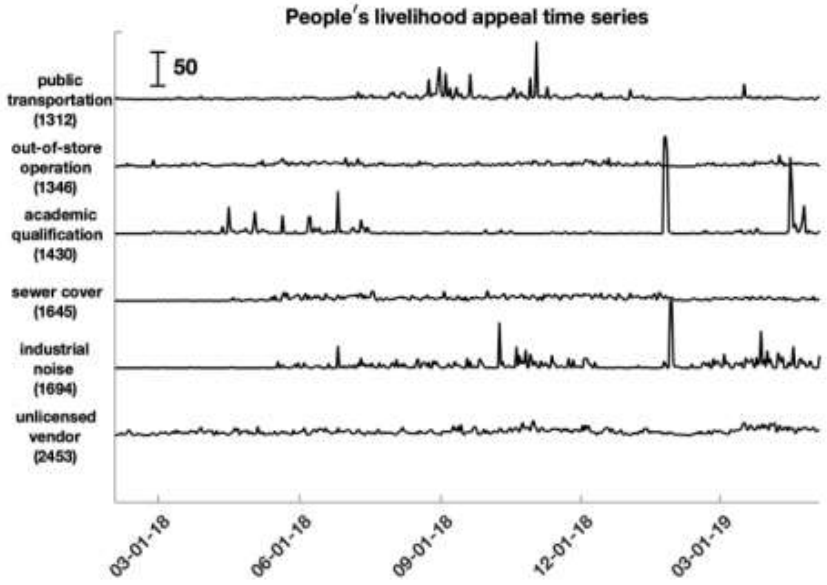

Fig.1. Time series of six types of people's livelihood appeal

This work employs the time series algorithm suggested in section II to dissect the six types of people's livelihood appeal, merges additional data to examine the causes, and offers practical solutions in order to more thoroughly research the time series change of the people's livelihood appeal.

B. Occasional change analysis

Fig. 2 depicts the sporadic alternation of occurrences. Table 2 lists the overall number of incidents, along with the frequency of sporadic events. The findings indicate that the six forms of appeal may be classified into two groups based on the proportion of occasional events: unlicensed vendor, sewer cover, and out-of-store operation have very low proportions of occasional events and are therefore reasonably stable. The other three categories of appeal—industrial noise, academic achievement, and public transportation—all have pronounced peaks. Rare occurrences make up a significant share of the overall number of events, and the percentage of people with academic degrees is even higher at 65.80%. It is obvious that understanding the root causes of sporadic events is crucial for identifying trends and lowering the frequency of such occurrences.

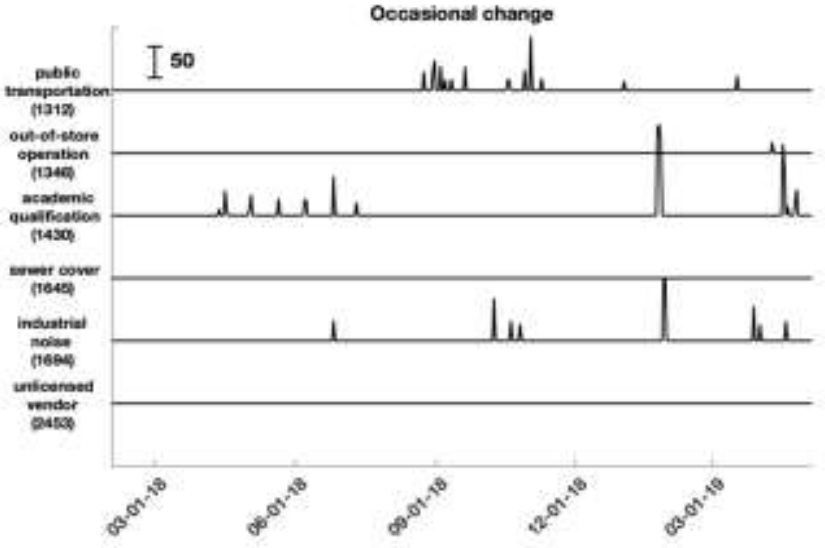

Fig.2. Occasional change of six types of people's livelihood appeal

TABLE II. Proportion of occasional change of major appeals

	Unlicensed vendor	*Industrial noise*	*Sewer cover*	*Academic qualification*	*Out-of-store operation*	*Public transportation*
Total Event	2453	1694	1645	1430	1346	1312
Occasional Event	0	474	0	941	17	427
Proportion	0%	27.98%	0%	65.80%	1.26%	32.55%

The outcomes of big data analysis using the time series decomposition algorithm suggested in this research are explained in the sections that follow using the example of a degree. The time series diagram of academic qualification appeals is shown in Fig. 3.

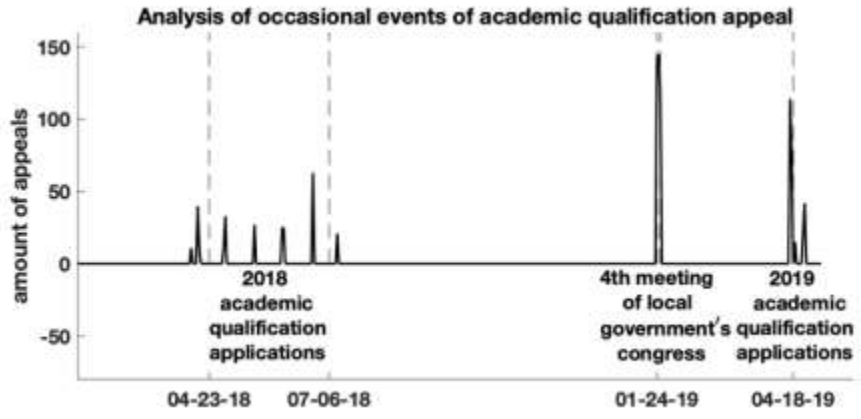

Fig.3. Analysis of occasional events of academic qualification appeal

The timing of significant political gatherings and social events pertaining to education is indicated by the dotted line in Figure 3. The application time for elementary school degrees in 2018 is from April 23 to July 6, and the application period for elementary school degrees in 2019 is from April 18 to July 12, respectively. From January 24 to January 26, 2019, the local administration hosted the fourth meeting of the People's Congress. The graph demonstrates unequivocally that there is a significant association between significant political gatherings and sporadic events that contribute significantly to academic qualification. As everyone is aware, there are extremely few primary and secondary schools in a certain region, and the admission process is also rather challenging. Each family is required to submit a substantial amount of supporting documentation in order to earn the necessary entrance points. It was inevitable that there would be several communications with the education department throughout this process, which increased the amount of complaints and suggestions. Moreover, on January 24, 25, and 26, 2019, respectively, 143, 145, and 119 appeals for academic qualifying peaked. We have noted that the fourth meeting of the People's Congress was hosted by the local government from January 24 to January 26, 2019. This crucial political gathering has an impact on peoples' daily lives. At the same time, we observed that just 80.8% of academic qualification appeals were completed on time. Therefore, we discovered the following as the causes of the sudden rise in academic qualification appeals from January 24 to January 29, 2019: because academic qualification appeals had not been properly resolved in a timely manner, people took the chance to voice their grievances and suggestions when the local government held a meeting of the People's Congress, which resulted in the sudden rise from January 24 to 29, 2019. We go on to develop predictions about academic qualification appeals based on the analysis mentioned previously. The conflict between primary and secondary school degree applications still remains since the discrepancy between the degree capacity and the number of school-age children in a given area has not been substantially resolved. While the application period for school degrees runs from April 18 to July 12, 2019, academic qualification issues are still a concern. In January 2020, when the People's Congress meets, the demand for academic credentials will once more reach its height.

C. Long-term change trend analysis

Six different forms of people's livelihood appeal are depicted in Fig.4 as a long-term shift trend. Unlicensed vendor and industrial noise share several similarities that lead us to believe that these two sorts of appeal might have a common source. We retrieved the economic monitoring report from February 2018 to April 2019 from the official website of the development and reform bureau of a local government and used the information to determine the industrial production of the local government during that time.

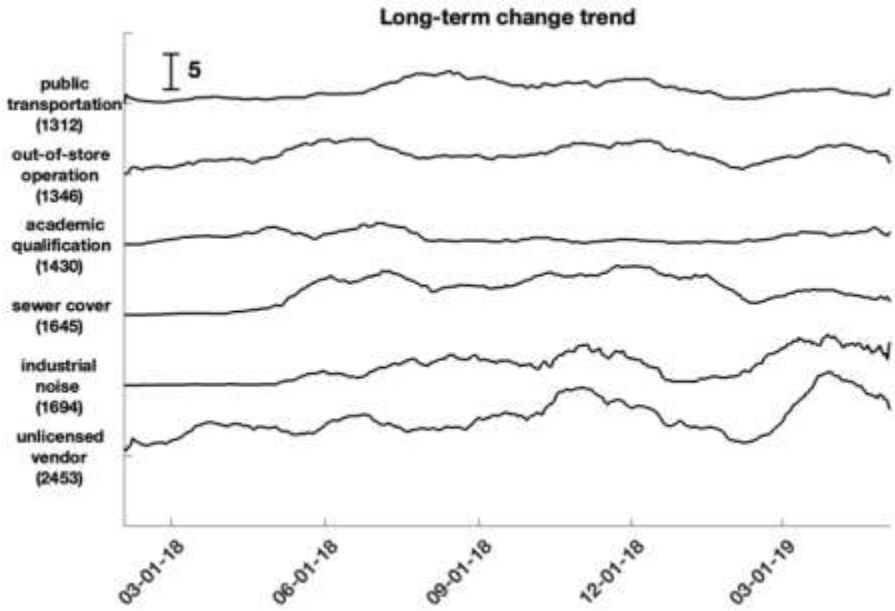

Fig.4. Long-term change trend of six types of people's livelihood appeal

According to Fig. 5, the local government's industrial production curve and the appeals curve for unlicensed vendors and industrial noise are extremely comparable to one another, with the local government's curve roughly 60 days behind the other. Table 3 displays the results of Pearson correlation analysis and t-paired significance testing after lag time effects were taken into account for the industrial product, appeals for unlicensed vendor, and industrial noise. According to the findings, there is some degree of positive link between them. The p-value indicates that the calculation findings are significant; the correlation coefficients are 0.5045 and 0.5788, respectively.

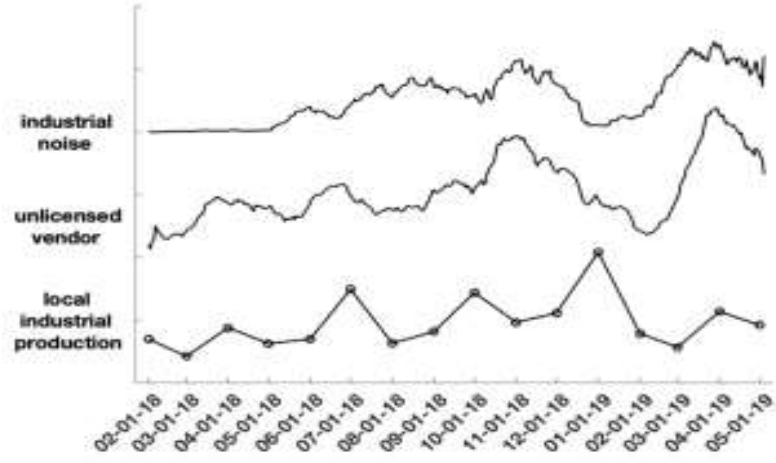

Fig. 5. The curve of unlicensed vendor and industrial noise appeals and industrial production of a local government

Table 3 Correlation analysis between unlicensed vendor and industrial noise appeals and industrial production of a local government

	industrial production	
	correlation coefficients	***the p-value of significance test***
Amount of unlicensed vendor appeals	0.5045	<0.05
Amount of industrial noise appeals	0.5788	<0.05

We go into more detail about the reasons behind the two appeals—unlicensed vendor and industrial noise—based on the results mentioned above. A particular location has a lot of industrial parks, and there is a strong correlation between industrial activity and noise levels. At the same time, a sizable group of unlicensed vendors congregates close to the factory to serve industrial workers with breakfast, lunch, and dinner catering services. It is evident that there is some connection between industrial activity and the rise of unlicensed vendors. In general, the value of industrial production will appear in the statistics reports after the production activities. Also, this explains why the industrial production curve trails behind that of unlicensed vendors and industrial noise. The technique above allows us to forecast the appeals for unlicensed vendors and industrial noise based on overall industrial production. The number of appeals for unauthorized vendors and industrial noise will rise in line with the anticipated development in industrial production. To improve the supervision of production activities and effectively reduce industrial noise, we advise completing administrative tasks two months in advance. At the same time, actively promoting the availability of livelihood options like restaurants close to the factory can also lessen the likelihood of complaints about unlicensed vendors.

D. Short-term change trend analysis

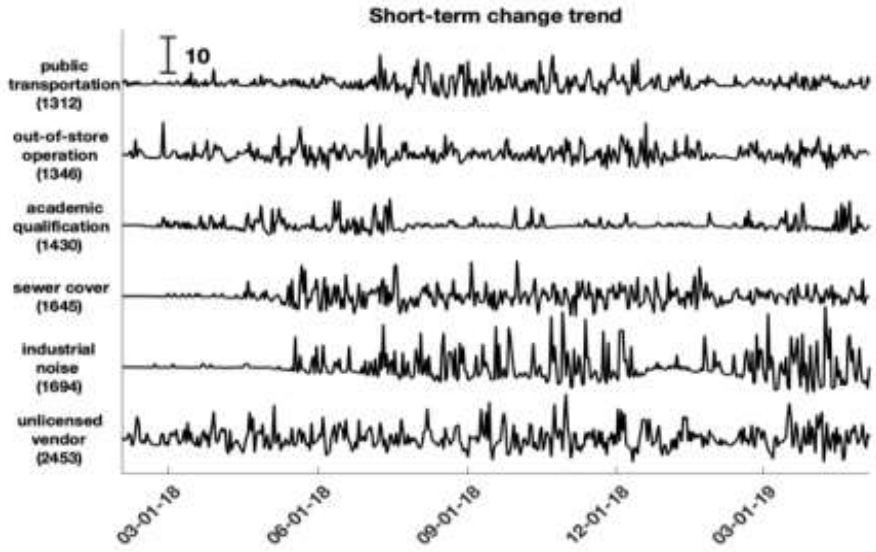

Fig.6. Short-term change trend of six types of people's livelihood appeal

The attraction of people's livelihoods is shown to alter quickly in Fig. 6. It demonstrates that the short-term alterations have a cyclical pattern and last for roughly 7 days. The six main sorts of appeal are listed in this paper in the following order: Sunday, Monday, Tuesday, Wednesday, Thursday, Friday, and Saturday. This is done in an effort to further explore this subject. The findings indicate that people's working and living times have an impact on short-term changes. The appeal for an unlicensed vendor does not change significantly over the course of a week; the appeal for an out-of-store operation slightly declines during the working day; the appeals for industrial noise, sewer cover, academic qualification, and public transportation significantly change over the course of the week; and there are more appeals on weekdays and fewer on weekends. Thus, we urge government agencies to suitably improve their management of appeals on weekends in addition to aggressively managing them throughout the week in response to appeals relating to commercial activities like out-of-store business and unlicensed vendor. The employment of artificially intelligent technologies (such as service robots and intelligent customer service) can help with this.

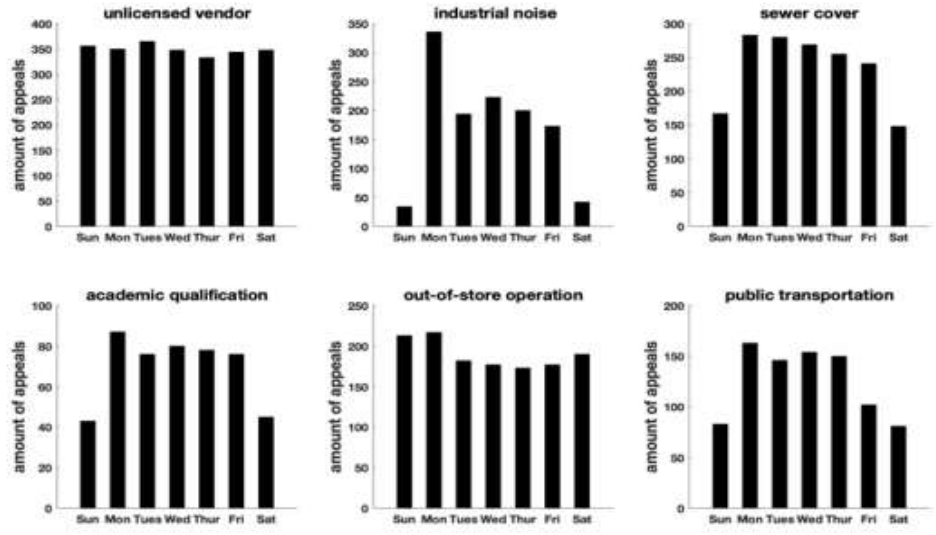

Fig.7 Distribution of six types of appeal during a week

4.Conclusion

Time series analysis is a methodology that applies statistical methods to find the pattern of data changes, and it is significant in our daily lives. There are currently only a few number of large data mining and analysis projects that utilize the time series technique. In this study, an additive combination model was used to evaluate the time series data of 39,788 people's livelihood pleas that the public reported via WeChat, email, hotline, Weibo, and other channels on government open platforms. The three sub-categories of long-term change trend, short-term change trend, and sporadic change are used to break down the change of time series. The data set is subjected to big data analysis and mining utilizing the time series analysis method. We have identified the shifting trends in people's livelihood appeal by experiments and analyses, examined the reasons for the occurrences, and provided predictions. It was proven that the time series analysis method works, and this information could be used by the government to conduct research, make choices, and address issues related to people's livelihoods.

References

[1] Anthony Bagnall, Jason Lines, Aaron Bostrom, James Large, Eamonn Keogh. "The great time series classification bake off: a review and experimental evaluation of recent algorithmic advances." Data Mining & Knowledge Discovery, 2016, vol.31.3, pp.606-660.

[2] Tang.Xiangyu, C.Yang, and J.Zhou. "Stock Price Forecasting by Combining News Mining and Time Series Analysis." 2009 IEEE/WIC/ACM International Conference on Web Intelligence, WI 2009, Milan, Italy, 15-18 September 2009, Main Conference Proceedings ACM, 2009.

[3] Zahid Iqbal and Shaikh Abdul Hamid. "Stock Price and Operating Performance of ESOP Firms: A Time-Series Analysis." Quarterly Journal of Business & Economics, 2000, vol.39.3, pp.25-47.

[4] Bruce J. Summers. "A time series analysis of business loans at large commercial banks." Economic Review May,1975, pp.8-14.

[5] Esad Jakupović, Vladimir Stojanović, Sanel Jakupović. "Time Series Analysis of New Cars Sale in Bosnia and Herzegovina". Applied Mechanics & Materials, 2015,vol. 806, pp.287-293.

[6] Manolakis, Dimitris, Bosowski, Nicholas, Ingle, Vinay K. "Count Time-Series Analysis: A Signal Processing Perspective". IEEE Signal Processing Magazine, vol.36.3, pp.64-81.

[7] Menon P K , Sweriduk G D , Bilimoria K D . "New Approach for Modeling, Analysis, and Control of Air Traffic Flow". Journal of Guidance Control & Dynamics, 2004, vol.27.5, pp.737-744.

[8] Faht G , Silva M R D , Pinheiro A. "Environmental Monitoring Model for a Drainage Basin Obtained through Spectral Analysis of Time Series". Water Environment Research, 2012, vol.84.8, pp.662-672.

[9] Yadav R K , Balakrishnan M. "Comparative evaluation of ARIMA and ANFIS for modeling of wireless network traffic time series". EURASIP Journal on Wireless Communications and Networking, 2014, vol.2014.1, pp.1-8.

[10] Kane M J , Price N , Scotch M. "Comparison of ARIMA and Random Forest time series models for prediction of avian influenza H5N1 outbreaks". BMC Bioinformatics, 2014, vol.15.1, pp.276.

[11] Zhang Min, Wu Yusong, Huo Chaoguang. "Research Hotspots and Tendency Analysis of E-government in China ". Journal of Intelligence. 2015, vol.2, pp.137-141.

[12] Kim G H , Trimi S , Chung J H . "Big-data applications in the government sector ". Communications of the Acm, 2014, vol.57.3, pp.78-85.

[13] Bertot, John C, Gorham, Ursula, Jaeger, Paul T. "Big data, open government and e-government: Issues, policies and recommendations". information polity, 2014, vol.19.1, pp.5-16.

[14] Schwertman N C , Owens M A , Adnan R . "A simple more general boxplot method for identifying outliers". computational statistics & data analysis, 2004, vol.47.1, pp.165-174.

[15] Singh, Omkar, Sunkaria, Ramesh Kumar. "ECG signal denoising based on Empirical Mode Decomposition and moving average filter". international journal of medical engineering & informatics, 2013, vol.6.6, pp.1-6.