Journal of Theoretical and Applied Information Technology
31st October 2012. Vol. 44 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

# INFORMATION FLOW ANALYSIS BETWEEN TRAFFIC FLOW TIME SERIES BASED ON THE TRANSFER ENTROPY

[1]YUECHEN CHEN, [2]PENGJIAN SHANG
[1]School of Science, Beijing Jiaotong University
Beijing 100044, P. R. China
E-mail: [1]pjshang@bjtu.edu.cn

**ABSTRACT**

In this paper, we employ transfer entropy as our analysis tool of traffic flow. We select the passenger flow volume of Beijing Subway Line 1 as our data source. The research aims at finding out the strength and the direction of information transfer of the passenger flow volume.

**Keywords:** *Traffic flow, Passenger flow volume, Transfer entropy, Information flow, Beijing subway.*

## 1. INTRODUCTION

Economists are employing more and more tools for physics, to solve the economic problem [1, 2, 3]. In the system Theory, entropy [4, 5] is a significant parameter to quantify the degree of irregularity or uncertainty of the system. The system tends to be more chaotic when the entropy increases. Information is a decrease in uncertainty. This concepts origins from Physics and is now applied to time series analysis. If a time series is regular, the previous value(s) may contain the information of the next adjacent value(s), and decrease the uncertainty of the next value. This is widely used in statistics.

When investigating interacted time series, transfer entropy [6] would be an effective method. Using transfer entropy to measures the interaction between time series has become popular in economy researches [7, 8], commonly between financial indexes, e.g. stock price. The transfer entropy not only measures the information flow of the system`s internal structure, but also captures the direction of the information flow [6, 9, 10]. In this paper, the research would discover the internal relationship of the Traffic flow time series, and further to evaluate the traffic network.

In recent years, Beijing is building more and more subway lines. Currently, 15 subway lines is operating (up to Apr. 2012), 11 of them are newly accomplished in the last 5 years. But few researches of the passenger flow volume have been done. Studying the passenger flow volume (PFV) to obtain the PFV between individual stations; find the "key stations" of PFV; quest the possible irrelevant structures of the subway networks; help us to make further decision of subway line construction.

Through transfer entropy, this paper focuses on information flow between 7 major stations of the Beijing Subway Line 1 (BSL1), directionally and quantitatively, to find out the relationship of the PFV; in the section 2, a brief introduction of the transfer entropy would be given; in the section 3, an overview of BSL1 is provided; in the section 4, the result by calculate the transfer entropy is obtained; in section 5, conclusion is summarized, and future works are mentioned.

## 2. METHOD OF ANALYZING

First of all, we will explain the definition of Shannon entropy.

Shannon entropy is a classic definition of entropy. For a stochastic discrete variable x, the Shannon entropy H(x) is defined as:

$$H(x) = -\sum_{i=1}^{n} p_i \log p_i \qquad P(x = x_i) = p_i \quad (1)$$

The H(x) describes the uncertainty of the stochastic variable x. When x equals to a particular value as surely, the H(x) equals zero. Reversely, H(x) reaches the maximum when pi = 1/n.

Consider two processes I, J, the definition of transfer entropy from J to I is:

$$T_{J\to I} = \sum p(i_{t+1}, i_t^{(k)}, j_t^{(k)}) \log \frac{p(i_{t+1} | i_t^{(k)}, j_t^{(k)})}{p(i_{t+1} | i_t^{(k)})} \quad (2)$$

Journal of Theoretical and Applied Information Technology
31st October 2012. Vol. 44 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Where $i_t$ and $y_t$ represent the states at time t of I and J, respectively. In terms of relative entropy, it can be rephrased as the distance from the assumption that J has no influence on I, when:

$$p\left(i_{t+1}\middle|i_t^{(k)}, j_t^{(l)}\right) = p\left(i_{t+1}\middle|i_t^{(k)}\right) \quad (3)$$

One may rewrite Equation (2) as:

$$\begin{aligned} T_{J\to I} &= H\left(i_{t+1}^{(k+1)}, j_t^{(l)}\right) - H\left(i_t^{(k)}, j_t^{(l)}\right) \\ &\quad - H\left(i_{t+1}^{(k+1)}\right) + H\left(i_t^{(k)}\right) \\ &= hI(k; t) - hIJ(k, l; t). \end{aligned} \quad (4)$$

## 3. OVERVIEW OF THE SUBWAY LINE 1

Beijing Subway line 1(BSL1) is a west-east direction subway, 31 kilometers long, west to the Pingguoyuan Station, which is outside Beijing Fifth Ring road, east to the Sihui East Station. Now a sketch map of BSL1 is given below (up to Apr. 2012).

Pingguoyuan - (6 more station between) – Gongzhufen (At the west third ring road) - (3 more station) -

Fuxingmen (Transfer station to the Line 2, which travels around the central part of Beijing) -

Xidan (Transfer station to the Line 4) - (3 more station) -

Dongdan (Transfer station to the Line 5) – Jianguomen (Transfer station to the Line 2 – the loop line) - (1 more station) -

Guomao (At east third ring road; Central business district; Transfer station to the line 10) - (1 more station) -

Sihui - Sihui East (Both of these two stations are Transfer station to Batong Line, Which extended to the east, leads to a satellite city – Tongzhou)

*Figure 1 Overview of BSL1*

## 4. DATA ANALYZING

The PFV 2011 of 7 major stations of BSL1: Pingguoyuan, Gongzhufen, Fuxingmen, Dongdan, Jianguomen, Guomao and Sihui East were picked as our data source. Every 15 minutes, we record the total number of passengers entering or exiting the station.

In order to eliminate the interference of the first train or the last train, the data before 6 a.m. and after 11 p.m. are eliminated. After 6 a.m. the first train from all direction (including the concerned line at transfer station) has arrived at the station and the last train to all direction has not departed before 11 p.m.

Data are obtained during Jul. 2011, 68 inflowing and 68 outflowing data every single day every station, 29512 totally.

For each station, the data of inflowing are divided into 3 category labeled high, medium and low. Here, the thresholds are the 1st quartile and the 3rd quartile.

To pursue the relationship between the increase and decrease of the PFV of the 7 major station, we also calculate the change of the PFV ($d_t = i_{t+1} - i_t$) and using the same method divide them into 3 category. The same work was done to process the data of outflowing PFV, and the Total PFV.

In the next part we would examine the result.

The transfer entropy between 7 stations, Pingguoyuan (PGY), Gongzhufen (GZF), Fuxingmen (FXM), Dongdan (DD), Jianguomen (JGM), Guomao (GM), Sihuidong (SHD), is given below. First we examine the transfer entropy of inflowing PFV, which is given by the Table 1:

*Table 1 : Transfer entropy of Inflowing PFV*

| From \ to | PGY | GZF | FXM | DD | JGM | GM | SHD | Mean |
|---|---|---|---|---|---|---|---|---|
| PGY | 0 | 0.0573 | 0.0452 | 0.0253 | 0.0214 | 0.0323 | 0.1076 | 0.0482 |
| GZF | 0.0364 | 0 | 0.0501 | 0.0675 | 0.0496 | 0.0377 | 0.0639 | 0.0509 |
| FXM | 0.0298 | 0.0194 | 0 | 0.0866 | 0.0903 | 0.0649 | 0.0461 | 0.0562 |
| DD | 0.0287 | 0.0088 | 0.0763 | 0 | 0.0595 | 0.0545 | 0.0512 | 0.0465 |
| JGM | 0.0301 | 0.0224 | 0.0606 | 0.0802 | 0 | 0.0522 | 0.034 | 0.0466 |
| GM | 0.0258 | 0.0317 | 0.088 | 0.0636 | 0.0759 | 0 | 0.0441 | 0.0548 |
| SHD | 0.0125 | 0.0489 | 0.0411 | 0.0157 | 0.0201 | 0.0409 | 0 | 0.0299 |
| Mean | 0.0272 | 0.0314 | 0.0602 | 0.0565 | 0.0528 | 0.0471 | 0.0578 | |

Noticing the value that is significantly big in this table, they are PGY to SHD, FXM to DD & JGM, JGM to DD, GM to FXM. Especially PGY, the transfer entropy from PGY to SHD is much higher than PGY to others. This can be explained by the geographical position difference: FXM, JGM and GM are considered to be the central district of Beijing, while the SHD and PGY are near to residential area.

Averagely, FXM station has the biggest transfer entropy to others. SHD has the lowest transfer entropy to others. PGY has the lowest transfer entropies.

Journal of Theoretical and Applied Information Technology
31st October 2012. Vol. 44 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

The transfer entropies of outflowing PFV are given by Table 2. This exhibits some different properties. Averagely, the SHD has the highest transfer entropy from others, while the PGY is the second.

This also happens in the Transfer entropy of total PFV.

*Table 2: Transfer Entropy Of Outflowing PFV*

| From \ to | PGY | GZF | FXM | DD | JGM | GM | SHD | Mean |
|---|---|---|---|---|---|---|---|---|
| PGY | 0 | 0.0175 | 0.0325 | 0.0328 | 0.0156 | 0.0178 | 0.089 | 0.0342 |
| GZF | 0.0726 | 0 | 0.0193 | 0.0266 | 0.0407 | 0.0596 | 0.1054 | 0.0541 |
| FXM | 0.0738 | 0.0771 | 0 | 0.0374 | 0.0901 | 0.0839 | 0.0908 | 0.0755 |
| DD | 0.0709 | 0.0836 | 0.0859 | 0 | 0.1039 | 0.0754 | 0.0724 | 0.082 |
| JGM | 0.0798 | 0.0647 | 0.0366 | 0.0308 | 0 | 0.063 | 0.0841 | 0.0598 |
| GM | 0.0877 | 0.0518 | 0.035 | 0.0344 | 0.0535 | 0 | 0.1007 | 0.0605 |
| SHD | 0.0616 | 0.0395 | 0.0393 | 0.0392 | 0.0207 | 0.0278 | 0 | 0.0380 |
| Mean | 0.0744 | 0.0557 | 0.0414 | 0.0335 | 0.0541 | 0.0546 | 0.0904 | |

*Table 3: Transfer Entropy Of Total PFV*

| From \ to | PGY | GZF | FXM | DD | JGM | GM | SHD | Mean |
|---|---|---|---|---|---|---|---|---|
| PGY | 0 | 0.036 | 0.0369 | 0.0526 | 0.0349 | 0.0428 | 0.1031 | 0.051 |
| GZF | 0.0791 | 0 | 0.0645 | 0.0514 | 0.0697 | 0.075 | 0.0695 | 0.0682 |
| FXM | 0.0489 | 0.0158 | 0 | 0.0409 | 0.1006 | 0.1022 | 0.0398 | 0.058 |
| DD | 0.0486 | 0.0338 | 0.0716 | 0 | 0.0849 | 0.0922 | 0.0342 | 0.0609 |
| JGM | 0.0414 | 0.0098 | 0.0377 | 0.0344 | 0 | 0.0528 | 0.0346 | 0.0351 |
| GM | 0.0441 | 0.0094 | 0.0463 | 0.0308 | 0.0426 | 0 | 0.028 | 0.0335 |
| SHD | 0.0388 | 0.0436 | 0.0375 | 0.0518 | 0.0484 | 0.0564 | 0 | 0.0461 |
| Mean | 0.0501 | 0.0247 | 0.0491 | 0.0437 | 0.0635 | 0.0702 | 0.0515 | |

This time GZF have the biggest transfer entropy of Total PFV to others.

For each cell of the "from A station to B station" figure, we calculate the average of the 3 data of transfer entropy (PFV inflowing, PFV outflowing, total PFV inflowing, from station A to station B), and express them by brightness of the color in the corresponding cell. Then, we derive the Fig.2.

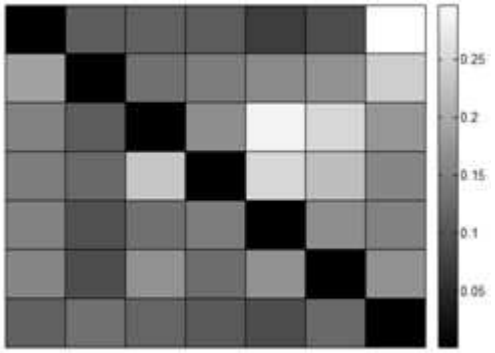


*Figure 2 Transfer Entropy Of PFV*

In Fig. 2, from top to bottom, and left to right, the corresponding stations are PGY GZF FXM DD JGM GM SHD. The brighter part represents higher transfer entropy,

Then we turn to check the transfer entropy between the change (increase /decrease) of PFV, Which is given by the Table 4, Table 5, Table 6.

*Table 4: Transfer Entropy Of Inflowing PFV Change*

| From \ to | PGY | GZF | FXM | DD | JGM | GM | SHD | Mean |
|---|---|---|---|---|---|---|---|---|
| PGY | 0 | 0.0296 | 0.0183 | 0.0202 | 0.0132 | 0.0199 | 0.054 | 0.0259 |
| GZF | 0.0239 | 0 | 0.039 | 0.0237 | 0.0646 | 0.0398 | 0.0381 | 0.0382 |
| FXM | 0.0222 | 0.0515 | 0 | 0.0474 | 0.0963 | 0.0747 | 0.0336 | 0.0543 |
| DD | 0.0109 | 0.0242 | 0.0395 | 0 | 0.0388 | 0.0304 | 0.0129 | 0.0261 |
| JGM | 0.011 | 0.0508 | 0.0561 | 0.0399 | 0 | 0.0514 | 0.0329 | 0.0403 |
| GM | 0.0292 | 0.0368 | 0.0452 | 0.0323 | 0.0676 | 0 | 0.0411 | 0.042 |
| SHD | 0.0347 | 0.0428 | 0.0214 | 0.0217 | 0.0315 | 0.0199 | 0 | 0.0287 |
| Mean | 0.022 | 0.0393 | 0.0366 | 0.0309 | 0.052 | 0.0393 | 0.0354 | |

*Table 5: Transfer Entropy Of Outflowing PFV Change*

| From \ to | PGY | GZF | FXM | DD | JGM | GM | SHD | Mean |
|---|---|---|---|---|---|---|---|---|
| PGY | 0 | 0.026 | 0.0205 | 0.0205 | 0.0081 | 0.0104 | 0.0362 | 0.0203 |
| GZF | 0.0349 | 0 | 0.0614 | 0.0386 | 0.0555 | 0.063 | 0.0411 | 0.0491 |
| FXM | 0.0242 | 0.0658 | 0 | 0.0696 | 0.0802 | 0.0809 | 0.0415 | 0.0604 |
| DD | 0.0186 | 0.0559 | 0.0609 | 0 | 0.0431 | 0.0563 | 0.0374 | 0.0454 |
| JGM | 0.0277 | 0.0683 | 0.0721 | 0.0469 | 0 | 0.0799 | 0.0294 | 0.054 |
| GM | 0.0277 | 0.0621 | 0.0571 | 0.0504 | 0.0889 | 0 | 0.0445 | 0.0551 |
| SHD | 0.0415 | 0.0339 | 0.022 | 0.0279 | 0.0243 | 0.0331 | 0 | 0.0304 |
| Mean | 0.0291 | 0.052 | 0.049 | 0.0423 | 0.05 | 0.0539 | 0.0383 | |

*Table 6: Transfer Entropy Of Total PFV Change*

| From \ to | PGY | GZF | FXM | DD | JGM | GM | SHD | Mean |
|---|---|---|---|---|---|---|---|---|
| PGY | 0 | 0.0533 | 0.0154 | 0.0293 | 0.0309 | 0.0181 | 0.0648 | 0.0353 |
| GZF | 0.0696 | 0 | 0.093 | 0.0713 | 0.1005 | 0.0703 | 0.0831 | 0.0813 |
| FXM | 0.0445 | 0.0984 | 0 | 0.0972 | 0.125 | 0.1069 | 0.091 | 0.0938 |
| DD | 0.0502 | 0.061 | 0.0909 | 0 | 0.0741 | 0.069 | 0.0713 | 0.0694 |
| JGM | 0.0526 | 0.0691 | 0.094 | 0.0741 | 0 | 0.0877 | 0.0582 | 0.0726 |
| GM | 0.0777 | 0.0823 | 0.0991 | 0.0861 | 0.1167 | 0 | 0.0941 | 0.0927 |
| SHD | 0.0812 | 0.065 | 0.0581 | 0.0574 | 0.0524 | 0.0492 | 0 | 0.0606 |
| Mean | 0.0626 | 0.0715 | 0.0751 | 0.0692 | 0.0833 | 0.0669 | 0.0771 | |

From Table 4-6, we found that FXM is dominating. The transfer entropy of PFV change from FXM is the biggest one. So we can esteem FXM as a information source as BSL1.

Similarly, we sum the Transfer entropy of the PFV inflowing change, PFV out flowing change and total PFV change, and depict the transfer entropy by the brightness of the color. The consequence is given at Fig.3. Then we sum the six transfer entropy together, to see the total situation of transfer entropy.

From Fig.4 we can also extract FXM station from others. The averagely have the biggest transfer entropy to others, so we can esteem the FXM as an information source as the PFV of the BSL1, especially to the nearby urban stations.

## 5. CONCLUSION AND FUTURE WORKS

In this passage, we employed transfer entropy as our tool, to analyze the passenger flow volume between 7 major stations. We derived the

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

information source of the Beijing Subway Line 1, However, Beijing is constructing more and more subway lines, in this paper we did not discuss the transfer entropy between stations that located on different subway lines. Only by calculating the transfer entropy between different lines can we correctly evaluate the Whole subway network. This would be a potential future work direction. In addition, whether the transfer entropy is effective in other kind traffic flow analysis is still to be tested.

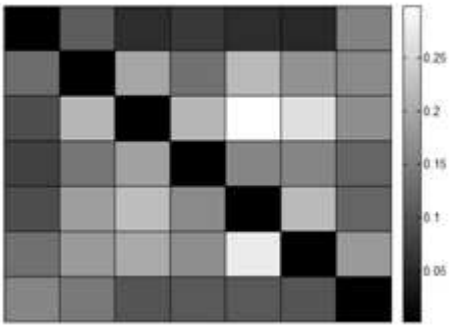

*Figure 3 Transfer Entropy Of PFV Change*

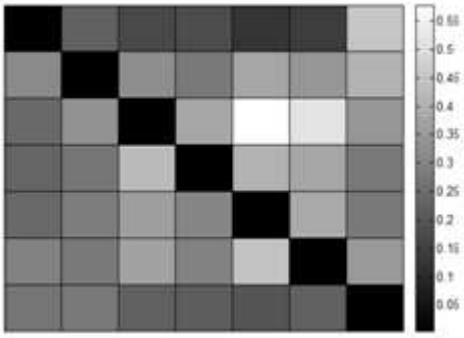

*Figure 4 Transfer Entropy Of PFV & PFV Change*

## 6. ACKNOWLEDGMENTS

The financial supports from the funds of the China National Science (61071142), the Beijing National Science (4122059) and the National High Technology Research Development Program of China (863 Program) (2011AA110303) are gratefully acknowledged.

## REFRENCES:

[1] W.B. Arthur, S. N. Durlauf, D. A. Lane, The Economy as an Evolving Complex System II, Perseus Books, 1997

[2] R. N. Mantegna, H. E. Stanley, Introduction to Econophysics, Cambridge University Press, 2000

[3] B. B. Mandelbrot, Quant. Finance 1 (2001) 124.

[4] C. E. Shannon and W. Weaver, The Mathematical Theory of Information University of Illinois Press, Urbana, IL,1949.

[5] P. Billingsley, Ergodic Theory and Information Wiley, New York, 1965.

[6] T. Schreiber, Phys. Rev. Lett, Measuring Information Transfer, 85 (2000) 461.

[7] Seung Ki Baek, Woo-Sung Jung, Okyu Kwon, Hie-Tae Moon, Transfer Entropy Analysis of the Stock Market, http://arxiv.org/abs/physics/0509014v2

[8] Okyu Kwon, Jae-Suk Yang, Information flow between composite stock index and individual stocks, Physica A,387(2012)2851

[9] L. Kullmann, J. Kert´esz, K. Kaski, Phys. Rev. E, Time-dependent cross-correlations between different stock returns: A directed network of influence, 66 (2002) 026125.

[10] T. Schreiber, J. Phys. A - Math. Gen. 23 (1990) L393.