Fluctuations and pseudo long range
dependence in network flows:
A non-stationary Poisson process model*

Chen Yu-Dong(陈煜东), Li Li(李  力)†, Zhang Yi(张  毅), and Hu Jian-Ming(胡坚明)

Tsinghua National Laboratory for Information Science and Technology (TNList), Department of Automation, Tsinghua
University, Beijing  100084, China

(Received 30 October 2008; revised manuscript received 14 November 2008)

In the study of complex networks (systems), the scaling phenomenon of flow fluctuations refers to a certain power-
law between the mean flux (activity) $\langle F_i \rangle$ of the $i$-th node and its variance $\sigma_i$ as $\sigma_i \propto \langle F_i \rangle ^\alpha$. Such scaling laws are found to be prevalent both in natural and man-made network systems, but the understanding of their origins still remains limited. This paper proposes a non-stationary Poisson process model to give an analytical explanation of the non-universal scaling phenomenon: the exponent $\alpha$ varies between $1/2$ and 1 depending on the size of sampling time window and the relative strength of the external/internal driven forces of the systems. The crossover behaviour and the relation of fluctuation scaling with pseudo long range dependence are also accounted for by the model. Numerical experiments show that the proposed model can recover the multi-scaling phenomenon.

Keywords: scaling, long range dependence, non-stationary Poisson process
PACC: 0500, 0540

1. Introduction

The study of complex networks has attracted increasing interest since the last decade. Many works have been devoted to the understanding of the dynamic processes of network flows and significant advances have been achieved. However, there are still many questions which remain to be answered thoroughly. Among them the origin of the fluctuations of network flow is of particular interest, which will be addressed in this paper.

The study of the scaling of fluctuations can be dated back to the Taylor’s law initialized in Ref.[5]. Recently, related studies have attracted increasing interest again, since Menezes and Barabasi’s proposal of the so-called scaling analysis. The recent research aims to describe the dynamics of a large number of nodes simultaneously and examine the collective behaviours of the systems. They investigated the coupling between the average flux and fluctuations and found that in complex networks there exists a characteristic coupling between the mean and variance of flux on individual nodes as $\sigma_i \propto \langle F_i \rangle ^\alpha$. Such scaling law was found to hold in diverse natural and man-made systems. Moreover, they claimed that real systems belong to one of two distinct universal classes. The first class yields $\alpha \approx 1/2$, which indicates that system dynamics is dominated by internal factors. Internet and microchip systems belong to this class. The other class yields $\alpha \approx 1$, which indicates that the fluctuations mainly come from external driven forces. Highway traffic, river networks, and the World Wide Web belong to this class.

These interesting findings were quickly followed by further studies. The scaling phenomenon has been confirmed in a wide range of complex networks, e.g. stock markets, gene systems, download networks, software developments, and urban transportation networks. In these systems, a rich variety of exponent $\alpha$ from $1/2$ to 1 can be observed, which approach 1 as the size of sampling time window or the relative strength of the external driven forces to the internal dynamics increases. Some systems even exhibit the so-called multi-scaling behaviour (i.e., the power-law extends to the $q$-th moments of $F_i$) and the crossover (i.e., different groups/types of nodes...
in the same system possess different values of $\alpha$ behaviour.\cite{10,11,14,16} These phenomena are believed to reflect the competition between the external and internal dynamics, as well as the inhomogeneity of the systems’ structure.\cite{11,15,17}

Several models have been proposed to explain the underlying mechanisms that generate the scaling laws,\cite{7,8,16–19} most of which are based on random walk or diffusive dynamic models. These models can reproduce the scaling phenomenon and link the dynamic processes of networks with the associated network topologies. Usually, the following questions are addressed:

1. The scaling law is better to be represented via both explicit analytical expression and numerical simulations.\cite{9,17–21}

2. The non-universal values of $\alpha$ are from $1/2$ to $1$, the effect of the sampling time windows and the internal/external driven forces should be explained in an integrated framework.\cite{9,15,21}

3. The crossover phenomenon should be accounted for.\cite{9,20,21}

However, we are still interested in finding an even more simpler model other than random walk/diffusive dynamic models to answer the above questions. In the authors’ earlier attempt, a non-stationary Poisson process model was proposed in Ref.\cite{20} to overcome the above shortcomings. This model could explain, both analytically and numerically, the transition of $\alpha$ from $1/2$ to $1$, as well as its dependence on the sampling time windows and external/internal driven forces. A similar model also appeared in Ref.\cite{21}. However, both Refs.\cite{20} and \cite{21} assumed that the length of the intervals between two consecutive jumps of the arriving rate equals that of the sampling time window. In order to better model the phenomena in real systems, it is necessary to relax this restriction and allow different size of the jumping intervals and the sampling time windows.

This paper extends the model in Ref.\cite{20} to more general cases and provides a more comprehensive study. In particular, analytical solutions are provided for $\alpha$’s dependence on sampling time windows and external/internal driven forces, as well as for the crossover phenomenon. Moreover, the connection of fluctuation scaling with pseudo long range dependence is explained. The multi-scaling behaviours are also investigated numerically. The proposed model is able to capture the essential mechanism that drives the fluctuations of flow and offers a unified and concise picture for the various phenomena observed in real traffic flows.

2. The non-stationary Poisson process model

Suppose that the network system consists of $N$ nodes, which will be indexed as $i = 1, ..., N$ in the rest of this paper. The arriving flow $f_i(t)$ of the $i$-th node at the $t$-th time interval is assumed to behave as a non-stationary Poisson process

$$\Pr(f_i(t) = n) = \frac{e^{-\lambda_i(t)}\lambda_i(t)^n}{n!},$$

where $n = 1, 2, ..., \Pr$ denotes the abbreviation of the probability.

Moreover, the varying process of the arriving rate $\lambda_i(t)$ is assumed to change according to a network-wide finite-state deterministic or stochastic process. The system will enter one state during a certain time interval $\tau$ and then enter another state, which as a result changes the arriving rate as well as the arriving flux of the Poisson process.

For instance, the widely used Markov-modulated Poisson process (MMPP) model is a typical example of such models. It is a doubly stochastic process where the intensity of a Poisson process is defined by the state of a Markov chain. The Markov chain can therefore be said to modulate the Poisson process, and thus comes the name MMPP in many literatures.\cite{22–27} Figure 1 shows a Poisson process modulated by a two-state Markov process, in which we can observe the shift of average arriving flow every $\tau$.

![Fig.1. Examples of a Poisson process modulated by a two-state Markov process.](image-url)
Before discussing the characteristics of such models, let us introduce some phrases to distinguish the different time intervals employed in the rest of this paper.

1. Intrinsic time interval denotes the minimum discernable (from the viewpoint of observer) time interval in the model, which is determined only by the non-stationary Poisson process. And we use the integer number set \( \{t = 1, 2, 3, \ldots, T\} \) to index the intrinsic time points.

2. Endogenous jump time interval denotes the lasting time after the process enters one state and before it leaves this state, which is determined only by the non-stationary Poisson process, too. Here, we assume it to be an integer multiple of the intrinsic time interval and use the integer number set \( \{j = 1, 2, 3, \ldots, T/\tau\} \) to index the jump time points.

3. Exogenous sampling time interval denotes the length of the sampling window (from the viewpoint of observer), which can be changed by the observer. We assume it to be an integer multiple of the intrinsic time interval, too. We use the symbol \( \Delta t \) to represent it and the integer number set \( \{s = 1, 2, 3, \ldots, T/\Delta t\} \) to index the sampling time points.

Suppose the dynamic varying process of the arriving rate \( \lambda_i(t) \) can be depicted as

\[
\lambda_i(t) = k_i \lambda(j), \quad \text{for} \quad (j-1)\tau + 1 \leq t \leq j\tau, \quad (2)
\]

where \( k_i \) is a scale coefficient which accounts for the factors controlling the relative magnitude of flow on node \( i \). For example, in a random walkers model, \( k_i \) stands for the degree of node \( i \).\[{21}\]

\( \lambda(j) \) represents the network-wide stochastic scalar parameter at the \( j \)-th jumping interval. We assume that all the nodes in the network systems follow a synchronized transition tempo so as to reproduce the network-wide fluctuations in a simple way. Studies show that we can relax this assumption to reproduce even more complex phenomena.

In this paper, we mainly focus on the steady state of this process, since it determines the steady distribution of the scalar parameter \( \lambda(j) \) and thus shapes the mean and variance of flux \( f_i(t) \) at the \( i \)-th node. Suppose that we have altogether \( M \) possible values of the scalar parameter \( \lambda(j) \), which are denoted as \( \{\lambda_m | m = 1, \ldots, M\} \); and each \( \lambda_m \) corresponds to a certain state of this process. Moreover, we assume that \( \lambda(j) \) have the following stationary distribution

\[
\Pr(\lambda(j) = \lambda_m) = p_m. \quad (3)
\]

In practice, we usually deal with sampled flows series instead of the original series \( f_i(t) \). Given a sampling window of length \( \Delta t \), the observed flow data at the \( s \)-th sampling index is the summation of the original flow values within the sampling window:

\[
F_i^{\Delta t}(s) = \sum_{t=(s-1)\Delta t+1}^{s\Delta t} f_i(t), \quad \text{for} \quad s = 1, \ldots, T/\Delta t. \quad (4)
\]

Under the assumption of ergodicity, the time average \( \langle \cdot \rangle \) of the flow series equals its equilibrium value. Thus the observed average flow at node \( i \) can be written as

\[
\langle F_i^{\Delta t} \rangle = \frac{1}{T/\Delta t} \sum_{s=1}^{T/\Delta t} F_i^{\Delta t}(s) = \frac{\Delta t}{T} \sum_{s=1}^{T/\Delta t} \sum_{t=(s-1)\Delta t+1}^{s\Delta t} f_i(t)
\]

\[
= \frac{\Delta t}{T} \sum_{t=1}^{T/\tau} f_i(t) = \frac{\Delta t}{T} \sum_{j=1}^{T/\tau} \sum_{i=(j-1)\tau+1}^{j\tau} f_i(t)
\]

\[
= \frac{\Delta t}{T} \sum_{j=1}^{T/\tau} k_i \lambda(j) \tau = k_i \Delta t \frac{1}{T/\tau} \sum_{j=1}^{T/\tau} \lambda(j)
\]

\[
= k_i \Delta t \langle \lambda \rangle, \quad (5)
\]

where \( \langle \lambda \rangle = \frac{1}{T/\tau} \sum_{j=1}^{T/\tau} \lambda(j) = \sum_{m=1}^{M} \lambda_m p_m. \)

To derive the variance \( \sigma_i^{\Delta t^2} \) of the observed flow, we need to distinguish two cases in terms of sampling window size \( \Delta t \).

I) \( \Delta t \leq \tau \) (see Fig.2). Further assume that \( \tau = \kappa \Delta t, \kappa \in \mathbb{N}, \) thus the arriving rate keeps constant in each sampling window. Noting that the \( s \)-th observed data satisfies \( F_i^{\Delta t}(s) \sim P(k_i \lambda_j \Delta t) \) for

---

**Fig.2. Diagram of Case I), where \( \Delta t \leq \tau \).**
The scaling of fluctuations and pseudo long range dependence reproduced by the proposed model

Based on Eqs. (10) and (11), the scaling law of the proposed model can be analytically explained. Particularly, we will discuss the three scaling phenomena which have been frequently observed both in simulations and real systems in previous literature. The connection to pseudo long range dependence will also be investigated.

1) The effect of the external and internal forces

The time-variation of \( \lambda(j) \) can be regarded as the result of external driven force; thus the ratio \( \text{Var} \lambda / \langle \lambda \rangle \) reflects its relative strength to internal dynamic factors. If the external driven force is weak, we have

\[
\langle s_x \rangle = \frac{T}{\Delta t} \sum_{s=1}^{T/\Delta t} A_x = \langle \lambda \rangle, \text{ we can simplify Eq. (7)}
\]

\[
\langle s_x \rangle^2 = k_1 \Delta t \langle A_x \rangle + k_2^2 \langle A_x \rangle^2 \left[ k_3 \text{Var} A_x + \langle A_x \rangle^2 \right]
\]

\[
= \langle F_i^{\Delta t} \rangle + \frac{k_3 \text{Var} A_x}{\langle A_x \rangle^2} \langle F_i^{\Delta t} \rangle^2.
\]
\[ \text{Var} \lambda / \langle \lambda \rangle \ll 1/(k_i \Delta t); \] thus the second term on the right-hand side of Eq.(11) can be omitted, yielding \( \alpha = 1/2 \). On the other hand, if the external driven force is strong and \( \text{Var} \lambda / \langle \lambda \rangle \gg 1/(k_i \Delta t) \), the first term can be omitted, yielding \( \alpha = 1 \). For intermediate values of \( \text{Var} \lambda / \langle \lambda \rangle \), the scaling law actually breaks down,\(^{28}\) but we still have effective values of \( \alpha \) ranging between 1/2 and 1.

Particularly, we are interested in network systems containing two or more driven forces (some are global/system-wide and external; some are local/district-wide). To separate these might help distinguishing which effect is dominant at a certain time, which will further assist us to take appropriate actions/controls. For example, in a previous report,\(^{15}\) we found that the scaling phenomena in urban ground traffic network might be explained as the interaction between the nodes’ internal dynamics (i.e. queuing at intersections, car-following in driving) and the changes of the external (network-wide) traffic demand (i.e. the every day increase of traffic amount during peak hours and shocking caused by traffic accidents). This may allow us to further understand the mechanisms governing the transportation system’s collective behaviour.

2) The influence of the sampling window

If the size of the sampling window \( \Delta t \ll \langle \lambda \rangle / (k_i \text{Var} \lambda) \), we can omit the second term on the right-hand side of Eq.(11), which contains the quadratic form of \( \Delta t \), and thus \( \alpha = 1/2 \).\(^{1}\) If \( \Delta t \gg \langle \lambda \rangle / (k_i \text{Var} \lambda) \), we can omit the first term on the right-hand side of Eq.(11) and thus \( \alpha = 1 \). For intermediate values of \( \Delta t \), the effective \( \alpha \) lies between 1/2 and 1.

3) The crossover behaviour

For nodes with small \( k_i \), the second term on the right-hand side of Eq.(11) can be omitted and it yields \( \alpha = 1/2 \). For nodes with large \( k_i \), the first term on the right-hand side of Eq.(11) can be omitted and thus \( \alpha = 1 \). If the system has heterogeneous structure and consists of a broad range of \( k_i \) across the nodes, it can be expected that different subsystems (i.e., different groups of nodes with different \( k_i \) and thus different \( \langle F_i^\Delta \rangle \)) have different values of exponent \( \alpha \).

4) The coexistence with pseudo long range dependence

Besides the scaling law of fluctuations, the long range dependence represents another kind of power-law, which couples the variance and the size of the sampling window

\[ \sigma_i(\Delta t) \propto \Delta t^{H(i)}, \]

where \( H(i) \) denotes the well-known Hurst exponent.

The coexistence of these two power-laws has received considerable interest, since they describe the behaviour of the same standard deviation \( \sigma_i(\Delta t) \). Previous studies showed that these two power-laws coexisted in some systems (e.g., stock markets), and several models were developed to explain such phenomenon.\(^{17,28,29}\)

In the proposed model, the power-law in Eq.(12) can be derived directly from Eq.(12), with \( H(i) \) ranging between 1/2 and 1, which recovers the possible coexistence of long range dependence (actually “pseudo long range dependence” here) and the scaling of fluctuations. Moreover, equation (11) predicts that \( H(i) \to 1/2 \) when \( \text{Var} \lambda / \langle \lambda \rangle \to 0 \), which reflects the uncorrelated nature of a simple Poisson process. On the other hand, when \( \text{Var} \lambda / \langle \lambda \rangle > 0 \), the time-variation of the arriving rate induces long-range dependency to the modulated Poisson process and results in \( H(i) > 1/2 \).

It should be pointed out that the long-range dependency discussed above is indeed the so-called ‘pseudo’ long range dependency (PLRD) instead of mathematically rigorous LRD. Usually, LRD means that the decay of the autocorrelation function is hyperbolic and decays slower than exponentially for all the time lag (time scale). However, long range dependency properties, heavy tail distributions, and all other characteristics of real time series (i.e. internet traffic, ground vehicular traffic flow) are meaningful only over a limited range of time scales. Many recent approaches\(^{24–26}\) had illustrated that appropriately constructed Markov models appear to be a viable modelling tool in the context of modelling LRD time series over several time scales. Particularly, the MMPP models are shown to provide good matches of LRD properties under large time scales. The key idea behind is obtaining a model that has fast transients on a local scale and evolves with a slower time constant between different groups of states. In other words a model that has some implicit form of memory, which is the origins of such PLRD. The model proposed here yields some kind of PLRD also via this way.

Besides, the proposed model also indicates clear non-universality, which means that the Hurst expo-

\(^{1}\)As pointed out in Ref.\(^{18}\), in the limit of infinitesimal sampling window such that \( \langle F_i^\Delta \rangle \ll 1 \), \( \alpha = 1/2 \) can be always recovered regardless of the underlying arriving process.
ments of different nodes may vary even though the underlying Poisson mechanisms for these nodes are the same. Thus we should be careful when we apply concepts like scaling and universality to complex systems.\cite{30}

4. Numerical experiments

In this section, some numerical experiments are designed to verify the proposed model.

In order to observe the influence of the strength of the external and internal forces, the corresponding parameters of the model are set as $T = 100000$, $\tau = 1$, $\Delta t = 1$. $N = 20$, $k = [1, 2, 3, ..., 20]$. $M = 10$, the ratio $\text{Var} \lambda/\langle \lambda \rangle$ is controlled to vary from 0.01 to 200. In the first test, the transition of $\alpha$ from $1/2$ to 1 under different $\text{Var} \lambda/\langle \lambda \rangle$ can be clearly observed from Fig.4. Similar results can be found that $\Delta t > \tau$ cases and are thus omitted here.

![Fig.4. The transition behaviour of $\alpha$ with respect to the competition of the external/internal driven forces produced by simulation.](image)

In the second test, the corresponding parameters of the model are set as $T = 1000000$, $\tau = 1$, $\Delta t = 1$, $M = 10$, $p_m = 1/10$ for $m = 1, ..., 10$, $\lambda_m = [1, 2, 3, ..., 10]$. $N = 20$, $k = [0.5, 0.5^2, ..., 0.5^5, 1, 2, 5^2, ..., 5^{10}]$ in the third test. The result is plotted in Fig.6, which shows clear crossover behaviour.

![Fig.5. The transition behaviour of $\alpha$ with respect to the sampling time window produced by simulation.](image)

![Fig.6. The crossover behaviour produced by simulation.](image)

Another important phenomenon numerically reproduced by our model is the multi-scaling behaviour, which is the extension of the original scaling law to the $q$-th-order central moments

$$\sigma_{i,q} = \langle |F_i^{\Delta t} - \langle F_i^{\Delta t} \rangle|^q \rangle \propto \langle F_i^{\Delta t} \rangle^{\alpha(q)}. \quad (13)$$
Generally, this multi-scaling law (i.e. a dependence of \( \alpha(q) \) on \( q \)) is assumed to be related with the multi-fractality of time series.\[17\] In the fourth test, the corresponding parameters of the model are set as \( T = 10000, \tau = 10, M = 10, p_j = 1/10 \) for \( j = 1, \ldots, 10 \), \( \lambda_j = [2, 3, \ldots, 11] \). \( N = 20, k = [1, 2, \ldots, 10] \). Figure 7 plots the three \( q-\alpha(q) \) plots for \( \Delta t = 1, \Delta t = 10 \) and \( \Delta t = 100 \), respectively, in which \( \alpha(q) \) heavily depends on \( q \).

5. Conclusion

In summary, a non-stationary Poisson process model is introduced in this paper to explain the scaling of fluctuations in complex systems. The scaling exponent \( \alpha \) with non-universal values between 1/2 and 1 can be analytically derived. The influences of the sampling time window and the external/internal force ratio, the crossover behaviour, the multi-scaling phenomenon, and the connection with long range dependency are also naturally explained. The model is verified by numerical simulations.

Our work also sheds light on the well-known debate on whether Poisson process models are applicable in modelling complex network systems which often exhibit burstiness and long range dependency.\[31\]–\[33\] The results presented in this paper support the argument that the non-stationary Poisson process are still powerful in modelling the multi-scale behaviours of complex time-variant systems with complicated interactions between external and internal dynamics.

References

[1] Albert R, Barabási A-L 2002 Rev. Mod. Phys. 74 47
[2] Newman M E J 2003 SIAM Rev. 45 167
[3] Boccaletti S, Latora V, Morenod Y, Chavez M and Hwang D U 2006 Phys. Rep. 424 175
[4] Gao Z, Sun H and Wu J 2007 Chin. Phys. Lett. 24 1284
[5] Taylor L R 1961 Nature 199 732
[6] de Menezes M A and Barabási A-L 2004 Phys. Rev. Lett. 92 028701
[7] de Menezes M A and Barabási A-L 2004 Phys. Rev. Lett. 93 068701
[8] Eisler Z, Kertesz J, Yook, S H and Barabási A-L 2006 Europhysics Lett. 69 664
[9] Eisler Z, Kertesz J 2008 Adv. Phys. 57 89
[10] Raffaelli G and Marsili M 2006 J. Stat. Mech. Theory Exp. 08 L08001
[11] Nacher J C, Ochiai T and Akutsu T 2005 Mod. Phys. Lett. B 19 1139
[12] Han D D, Liu J G and Ma Y G 2008 Chin. Phys. Lett. 25 765
[13] Valverde S 2007 Europhys. Lett. 77 12002
[14] Chen Y, Li L, Zhang Y and Jin X 2008 Chin. Phys. Lett. 25 1912
[15] Meloni S, Gómez-Gardeñes J, Latora V and Moreno Y 2008 Phys. Rev. Lett. 100 208701
[16] Rossi R 1987 Austral. Telecommun. Res. 21 53
[17] Fischer W and Meier-Hellstern K 1992 Performance Evaluation 18 149
[18] Andersen A T and Nielsen B F 1998 IEEE J. Selected Areas in Communications 16 719
[19] Salvador P, Valadas R and Pacheco A 2003 Telecommun. Syst. J. 23 123
[20] Muscariello L, Mellia M, Meo M, Ajmone Marsan M and Lo Cigno R 2008 Comput. Commun. 28 1385
[21] Bosch G, Greiner S, de Meer H, Trivedi K S and Trivedi K S 2006 Queueing Networks and Markov Chains: Modeling and Performance Evaluation with Computer Science Applications 2nd ed. (New York: Wiley)
[22] Eisler Z and Kertesz J 2006 Phys. Rev. E 73 046109
[23] Liu F, Ma Z, Ren, Shan X, Zhang J and Ma Z 2004 Acta Phys. Sin. 53 373 (in Chinese)
[24] Eisler Z and Kertesz J 2006 Econophysics of Stock and Other Markets Chatterjee A, Chakrabarti B K eds. (Milan: Springer) 49
[25] Paxson V and Floyd S 1995 IEEE/ACM Transactions on Networking 3 226
[26] Cao J, Williams S C, Lin D and Don X S 2001 ACM SIGMETRICS Performance Evaluation Review 29 102
[27] Karagiannis T, Molle M, Faloutsos M and Brodian A 2004 Proceedings of INFOCOM