Austrian Journal of Statistics
June 2014, Volume 43/3-4, 229–236.
http://www.ajs.or.at/

Power-Law Random Graphs' Robustness: Link Saving and Forest Fire Model

Marina Leri
Russian Academy of Sciences

Yury Pavlov
Russian Academy of Sciences

Abstract

We consider random graphs with node degrees drawn independently from a power-law distribution. By computer simulation we study two aspects of graph robustness: preserving graph connectivity and node saving in the forest fire model, considering two types of graph destruction: the removal of nodes with the highest degrees and equiprobable node extraction.

Keywords: random graphs, power-law distribution, robustness, simulation modeling, forest fire model.

1. Introduction

The study of random graphs has been gaining interest in the past decades due to the wide use of these models for the description of massive data networks (see e.g. Aiello, Chung, and Lu 2000; Newman, Strogatz, and Watts 2001; Durrett 2007; Hofstad 2011). Such models can be used for representing transport, telephone and electricity networks, social relationships, telecommunications and, of course, the main global network – Internet. While considering these networks it has been noted that their topology could be described by random graphs, with the node degrees being independent and identically distributed (i.i.d.) random variables following the power-law distribution (Faloutsos, Faloutsos, and Faloutsos 1999; Reittu and Norros 2004, etc.).

The structure of present-day complex networks contains many elements, wherefore theoretical research in the field of power-law random graphs includes the study of the limit behaviour of different characteristics of such graphs' structure (Aiello *et al.* 2000; Pavlov 2007; Norros and Reittu 2008, , etc.). Furthermore, one of the important questions raised in the studies of these networks is how their structure and, therefore, functioning change if some of the nodes fall out. That is why one of the important trends in the random graph field has been the study of random graph robustness to different types of breakdowns (see e.g. Cohen, Erez, Ben-Avraham, and Havlin 2000; Bollobas and Riordan 2004; Durrett 2007; Norros and Reittu 2008).

Alongside with the theoretical approach simulation modeling has always been one of the tools for studying random graph objects (Reittu and Norros 2004; Leri 2009, , etc.). In our work

we consider two aspects of graph robustness: link saving or preserving graph connectivity, and node survival, which is closely connected to the study of forest fire models.

2. Power-law graph model

We consider power-law random graphs with the number of nodes that equals N . Node degrees $\xi_1, \xi_2, \ldots, \xi_N$ are i.i.d. random variables drawn from the following distribution:

$$\mathbf{P}\{\xi \geq k\} = k^{-\tau}, \qquad k = 1, 2, \ldots, \quad \tau > 1, \tag{1}$$

For graph construction each node is given a certain degree in accordance with the degree distribution (1). Node degrees form stubs (or semiedges) that are numbered in an arbitrary order. The graph is constructed by joining all the stubs pairwise equiprobably to form links. If the sum of node degrees is odd one stub is added to a random vertex. Obviously these graphs have loops and multiple edges. Such construction gave these graphs one of their names – configuration graphs with i.i.d. degrees (Durrett 2007; Hofstad 2011).

3. Link saving

Research in the last decades (see Faloutsos *et al.* 1999; Reittu and Norros 2004) showed that configuration power-law random graphs with parameter τ of the node degree distribution (1) lying in the interval $(1, 2)$ are deemed to be a good implementation of Internet topology at both rooter and domain levels. That was one of the main reasons for us to study such an aspect of graph robustness as preserving graph connectivity or link saving. This issue is important because it is essential to know how the network structure will be influenced by the destruction of some nodes.

When $\tau \in (1, 2)$ the distribution (1) has finite expectation and infinite variance. Both theory (Reittu and Norros 2004; Durrett 2007; Pavlov 2007) and simulation (Reittu and Norros 2004; Leri 2009) agree on the fact that such graphs contain one so called giant component, which is a connected set of nodes the expectation of which is proportional to the number of graph nodes N.

For computer experiments we built a simulation model of power-law random graphs (Leri 2009) based on an algorithm introduced by Tangmunarunkit, Govindan, Jamin, Shenker, and Willinger (2002) using a pseudo random generator "Mersenne twister" (see Matsumoto and Nishimura 1998). Previously we showed (Leri 2009) that the structure of these graphs dramatically changes with the variations of the value of the node degree distribution parameter τ even within this small interval $(1, 2)$, but is much less dependent on the graph size N. With the value of parameter τ close to 1, the graph will be more connected and more than 95% of all graph nodes will form the giant component. On the other hand, the closer the value of parameter τ is to 2, the fewer nodes there will be in the giant component, i.e. only a half of all graph nodes. This does not however imply a significant growth of the other components. For example, the fraction of nodes in the second component will be rather small in comparison with the giant one. Even at its most, it will not exceed a little more than 1% of all graph nodes. In fact, other components will grow not in size, but in number. In particular, the structure of these power-law random graphs is one of the reasons why is it interesting to see how this structure changes when some graph nodes are removed.

In our work we consider two types of breakdowns: "random breakdown" when graph nodes are removed equiprobably, and "target attack", which means a removal of nodes with the highest degrees. For simulations we took graphs of ten sizes N from 500 to 5000 and 9 values of parameter τ from the interval $(1, 2)$ with a step of 0.1 (for each pair (N, τ) 100 graphs were generated to form statistical data). The graph destruction process looks as follows. When a chosen node is destroyed, all the links going out of this node are also removed. And then all isolated nodes are taken away.

Let $\eta_1, \eta_2, \ldots, \eta_s$ be random variables that are equal to the sizes of graph components in decreasing order (η_1 – the size of the giant component, η_2 – the size of the second component, etc.), where s is the total number of components. A graph is deemed destroyed when following event A occur: $\{\eta_1 \leq 2\eta_2\}$. Hence, when the size of the second biggest component becomes greater or equal to half the size of the giant component, the graph is considered destroyed.

Simulation results allowed us to derive regression dependencies of node percentages in the giant and the second biggest components (η_1 and η_2, respectively) and the total number of components s on the graph size N, the parameter of the node degree distribution τ and the percentage of removed nodes r.

In the case of "random breakdown" the following relations were found:

$$\eta_1 = 129 - 36\ \tau - 1.1\ r,$$

$$\eta_2 = 2 - 0.25 \ln N + 0.42\ \tau - 0.017 \ln r,$$

$$\frac{s}{N} = -0.18 + 0.2\ \tau - 0.004\ r \ln \tau.$$

The determination coefficients (R^2) of these regression models are equal to 0.98, 0.7 and 0.98, respectively. The percentage of removed graph nodes has to be confined within the following bounds: $100/N \leq r \leq 117 - 32.7\tau$. Here in after the lower bound implies the removal of one node, and the upper bound means that the extraction of a higher percent of nodes will lead to complete graph breakdown. The results show that the percentage of nodes in the giant component does not depend on the graph size N and the percent of nodes in the second component will not exceed 2% of the graph size.

The following regression dependencies were derived for "target attack" on the nodes with the highest degrees:

$$\eta_1 = 130 - 46\ \tau - 9\ r,$$

$$\eta_2 = 4.36 - 0.44 \ln N + \tau + 0.4 \ln r,$$

$$\ln s = -3.3 + \ln N + 2.3 \ln \tau + 0.1\ r,$$

with determination coefficients 0.95, 0.6 and 0.98, respectively, and the percentage of removed nodes confined within the following bounds: $100/N \leq r \leq 14 - 5.15\tau$. Here again, the percent of nodes in the giant component does not depend on the graph size N, and the percentage of nodes in the second component will not exceed 4% of the graph size.

Below are the results of the estimation of the regression dependence of the probability $\mathbf{P}\{\mathrm{A}\}$ of graph destruction on the percentage of removed nodes r and parameter τ with $R^2 = 0.84$ and $R^2 = 0.76$, respectively.

For "random breakdown":

$$\mathbf{P}\{\mathrm{A}\} = \begin{cases} 0, & r < 37/\sqrt{\tau}, \\ -0.2 + 1.5 \cdot 10^{-4} \tau r^2, & 37/\sqrt{\tau} \leq r < 89/\sqrt{\tau}, \\ 1, & r \geq 89/\sqrt{\tau}, \end{cases}$$

For the "target attack":

$$\mathbf{P}\{\mathrm{A}\} = \begin{cases} 0, & \ln r < 1.85 - \tau, \\ -0.38 + 0.06 r e^{\tau}, & 1.85 - \tau \leq \ln r \leq 3.13 - \tau, \\ 1, & \ln r > 3.13 - \tau, \end{cases}$$

This means that in the "random breakdown" case (see Figure 1), for example, an estimated probability of graph destruction equals 0 when $\tau = 1.1$ for all $r < 35.3\%$, and when $\tau = 1.9$ for $r < 26.9\%$. And $\mathbf{P}\{\mathrm{A}\} = 1$ for $r > 84.8\%$ when $\tau = 1.1$, and when $\tau = 1.9$ for $r > 64.5\%$. On the other hand in the case of "target attack" (see Figure 2) $\mathbf{P}\{\mathrm{A}\} = 0$ when $\tau = 1.1$ for all $r < 2.12\%$, and when $\tau = 1.9$ for $r < 0.95\%$. And $\mathbf{P}\{\mathrm{A}\} = 1$ for $r > 7.6\%$ when $\tau = 1.1$, and when $\tau = 1.9$ for $r > 3.4\%$.

The results show that power-law random graphs with parameter $\tau \in (1, 2)$ are much more vulnerable to "target attacks" than to "random breakdowns". In order to destroy such a graph by deleting high-degree nodes it is enough to remove $3 - 7\%$ of them. If however graph nodes

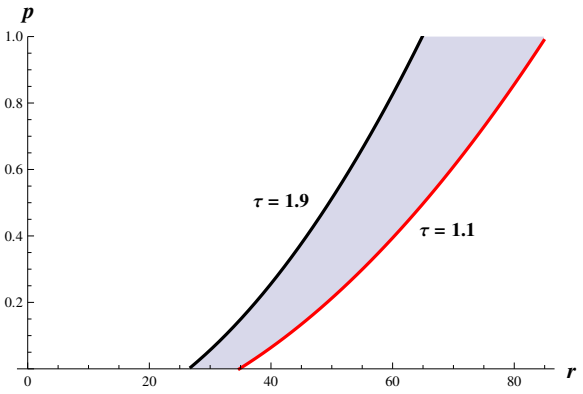

Figure 1: Probability of graph destruction for "random breakdown".

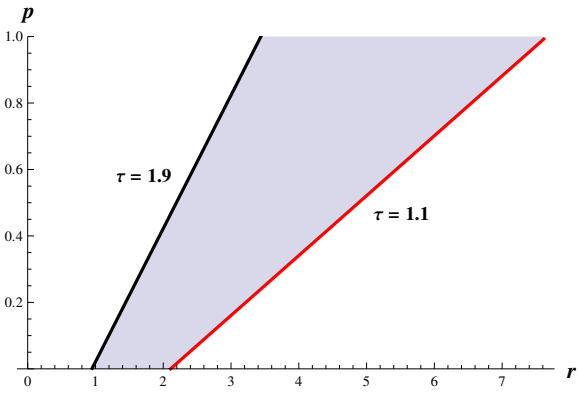

Figure 2: Probability of graph destruction for "target attack".

are broken randomly, it may be not ruined even if more than 60% of its vertices had been removed. Furthermore, robustness of these random graphs strongly depends on the value of parameter τ. In both breakdown cases the graph proved to be more resistant if the value of τ was closer to 1 and more vulnerable as it moved closer to 2.

4. Node survival – forest fire model

The second aspect of power-law graphs' robustness we are considering here is node survival. This issue branched off the studies of forest fire models (see e.g. Drossel and Schwabl 1992; Bertoin 2012). Let us consider graph nodes as trees on a certain area of a real forest. Two nodes are connected if a fire can move on related trees from one tree to another (for proper implication it looks more like a crown fire). So, we pose the question of finding how many trees should initially be set on a certain area to ensure their maximum survival in case of a fire. This approach could be used not only for modeling forest fire dynamics. It also has other applications (Bertoin 2011), including modeling banking system defaults in order to minimize their negative effects (see e.g. Annakov 2008; Arinaminparty, Kapadia, and May 2012).

In this part of our work we consider the same configuration power-law random graphs with node degree distribution (1), but with parameter $\tau > 1$ with no the upper bound. Since we assume that the area of a forest is limited, we have to also restrain the number of trees growing there. So, to specify the graph topology, let graph vertices be placed in the nodes of a square lattice sized 100×100. Links connect nodes in the "closest neighbour" manner, so in a fully packed graph every inner tree (node) has 8 adjacent neighbours. This does not mean that in the following study we consider power-law random graphs with node degrees no higher than 8. The fact is that on the one hand high node degrees and, all the more, an average node degree have low probabilities, and on the other hand the graph may contain multiple edges that raise the probability of fire transfer from one neighbour to the other. That is why we introduce the lattice only to determine the relation between the initial number of nodes in the area and the parameter τ of the node degree distribution (1). If an average node degree i is less than 8, some graph links are missing. Figure 3 shows a couple of examples of lattice-graph topology for two average inner node degrees.

Taking into consideration that graph node degrees are defined by distribution (1), and having determined the dependency between an average node degree i and parameter τ on the interval $i \in (1, 8]$ as $i = \zeta(\tau)$ (where ζ is a Riemann zeta function) (see Table 1), we found that graph size $N \leq 10000$ is related to parameter τ by the following regression function (see Figure 4) with $R^2 = 0.97$:

$$N = 9256\ \tau^{-1.05}. \tag{2}$$

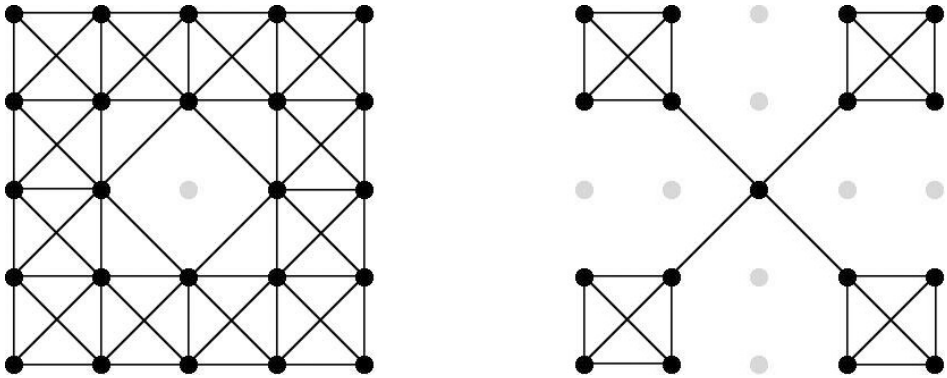

Figure 3: Lattice graph topology for $i = 7$ and $i = 4$, respectively.

Table 1: Calculated values of τ and N for different i.

i	1.01	1.21	1.33	1.42	1.5	1.6	2	2.66	3	4	5	6	7	8
τ	6.75	2.96	2.53	2.32	2.19	2.05	1.73	1.49	1.42	1.29	1.23	1.18	1.16	1.13
N	3350	3600	3750	3900	4000	4200	5000	4780	4489	5578	6700	8350	8911	10000

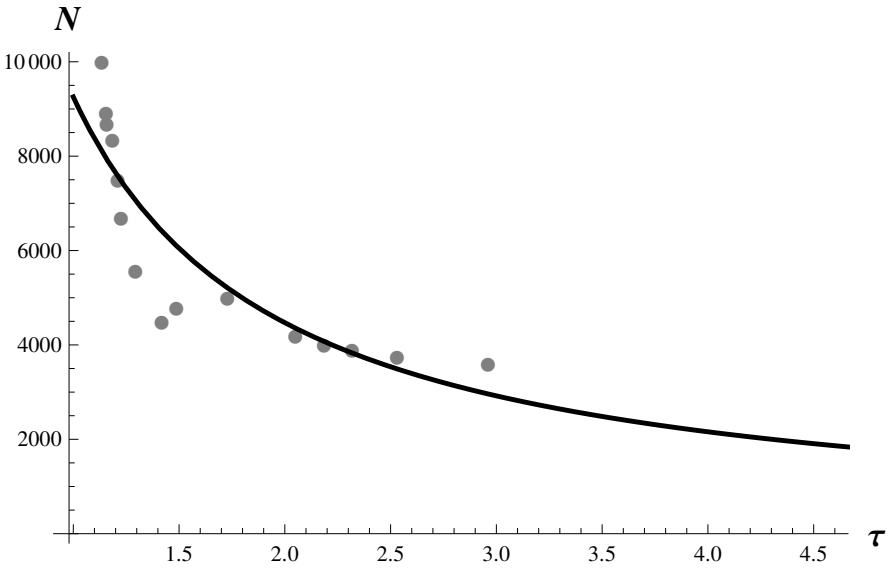

Figure 4: Regression relationship between N and τ.

For simulations we used a subset of configuration graphs the number of nodes in which is specified by relation (2). We assume that the graph destruction process (or fire) starts from some chosen node. As the first node is set on fire, it passes on along incident links to the connected nodes with a given probability $0 < p \leq 1$. Let's call it the probability of link destruction. This means that each link becomes inflammable with a probability p, and therefore a connected node is also set on fire. Otherwise a link becomes fire resistant and the node connected through this link remains intact. This does not mean however that the fire cannot reach this node via a parallel link (if any) with the same probability p.

The fire spreads over the graph until there appear inflammable links, and all burnt nodes and links are removed from the graph when it stops. The aim is to find the optimal values of parameter τ that secure maximum survival of nodes, and to find how they depend on the probability of link destruction.

We consider two cases of fire startup: "random fire start" when the first node to be removed is chosen equiprobably, and "targeted fire start" with fire starting from a node with the highest

degree. Let g be the number of nodes remaining in a graph after a fire. Figure 5 and Figure 6 show the results obtained for $p = 1$ in both breakdown cases, respectively.

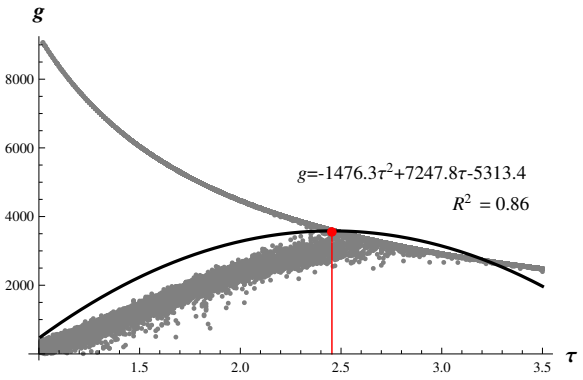

Figure 5: Relation between the number of remaining nodes g and parameter τ ("random fire start", $p = 1$).

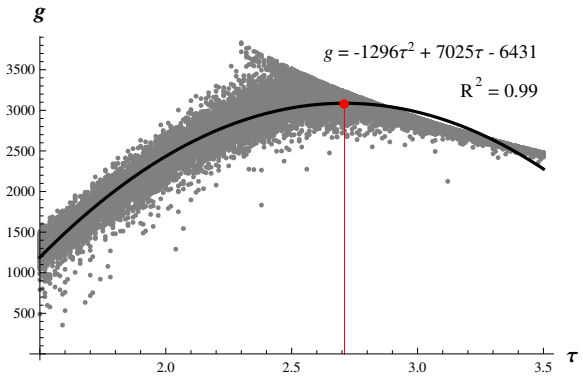

Figure 6: Relation between the number of remaining nodes g and parameter τ ("targeted fire start", $p = 1$).

This means, for example, that in the "random" case the number of remaining nodes becomes maximal for the power-law graphs with parameter $\tau = 2.45$. The initial graph size N then equals 3605 and an average number of nodes remaining after the fire is $g \approx 3580$.

For both cases of breakdown start were found regression dependencies of the number of nodes remaining in a graph g on τ and the probability of fire spread p. Below we give these models for the cases of "random" start and "target" start, respectively:

$$g = 6008.8 - 1915.3\ p - 217.4\ \tau^2, \qquad (R^2 = 0.91);$$

$$g = 2938 - 894.2 \ln p - 74.5\ \tau^2, \qquad (R^2 = 0.95).$$

Obviously, the number of remaining nodes decreases as p increases. Relations were found that describe the dependencies of g on τ for different values of p and dependencies of g on p for different τ. This allowed to find the relation between $\tau_{max} = \tau_{max}(p)$ of parameter τ for which g reaches its maximum g_{max} on p and, thus, find the values of $g_{max} = g_{max}(p)$. For example, for $p = 1$ and $p = 0.6$ under "random" start $\tau_{max}(1) = 2.46$, $\tau_{max}(0.6) = 1.1$, $g_{max}(1) = 3585$, $g_{max}(0.6) = 4468$, and under "target" start $\tau_{max}(1) = 2.74$, $\tau_{max}(0.6) = 2.23$, $g_{max}(1) = 3216$, $g_{max}(0.6) = 3995$.

Thus, in order to secure a maximum of unburnt trees in some specified territory in the case of a fire (either "random" or "targeted") the topology of their layout has to correspond to the topology of power-law random graph with parameter τ of node degree distribution (1) between values 2.4 and 2.7. Such graph will represent a multicomponent structure with no giant connected component with an average node degree 1.2 through 1.4. As for the difference between graphs robustness in the two breakdown cases, the graph will be more robust and more nodes will survive in a fire in the case of a "random" start than in the case of a "target" start.

5. Acknowledgments

The study was supported by the Russian Foundation for Basic Research, grant 13-01-00009 and by the Strategic Development Programme of the Petrozavodsk State University for years 2012–2016. Also we would like to thank professor A.M. Zubkov (Steklov Mathematical Institute of RAS) for a constructive discussion of the problem.

References

Aiello W, Chung F, Lu L (2000). "A random graph model for massive graphs." *Proc. of the 32nd Annual ACM Symposium on Theory of Computing*, pp. 171–180.

Annakov B (2008). "Bank crisis and forest fire: What's in common?" *URL http://www.empatika.com/blog/agent_modeling_forest_fire*. In Russian.

Arinaminparty N, Kapadia S, May R (2012). "Size and complexity model financial systems." *Proceedings of the National Academy of Sciences of the USA*, **109**, 18338–18343.

Bertoin J (2011). "Burning cars in a parking lot." *Commun. Math. Phys.*, **306**, 261–290.

Bertoin J (2012). "Fires on trees." *Annales de l'Institut Henri Poincare Probabilites et Statistiques*, **48**(4), 909–921.

Bollobas B, Riordan O (2004). "Robustness and vulnerability of scale-free random graphs." *Internet Mathematics*, **1**(1), 1–35.

Cohen R, Erez K, Ben-Avraham D, Havlin S (2000). "Resilience of the Internet to Random Breakdowns." *Phys. Rev. Lett.*, **85**, 4626–4628.

Drossel B, Schwabl F (1992). "Self-organized critical forest-fire model." *Phys. Rev. Lett.*, **69**, 1629–1632.

Durrett R (2007). *Random Graph Dynamics*. Cambridge Univ. Press, Cambridge.

Faloutsos C, Faloutsos P, Faloutsos M (1999). "On power-law relationships of the Internet topology." *Computer Communications Rev.*, **29**, 251–262.

Hofstad R (2011). *Random Graphs and Complex Networks*. Eindhoven University of Technology.

Leri M (2009). "Modelling of random graphs of Internet-type." *Surveys in Applied and Industrial Mathematics*, **16**(5), 737–744. In Russian.

Matsumoto M, Nishimura T (1998). "Mersenne Twister: A 623-dimensionally equidistributed uniform pseudorandom number generator." *ACM Trans. on Modeling and Computer Simulation*, **8**(1), 3–30.

Newman M, Strogatz S, Watts D (2001). "Random graphs with arbitrary degree distribution and their applications." *Phys. Rev. E.*, **64**, 026118.

Norros I, Reittu H (2008). "Attack resistance of power-law random graphs in the finite mean, infinite variance region." *Internet Mathematics*, **5**(3), 251–266.

Pavlov Y (2007). "The limit distribution of the size of a giant component in an Internet-type random graph." *Discrete Mathematics and Applications*, **17**(5), 425–437.

Reittu H, Norros I (2004). "On the power-law random graph model of massive data networks." *Performance Evaluation*, **55**, 3–23.

Tangmunarunkit H, Govindan R, Jamin S, Shenker S, Willinger W (2002). "Network topology generators: degree-based vs. structural." *Proceedings of the SIGCOMM'02*, pp. 147–159.

Affiliation:

Marina Leri and Yury Pavlov
Institute of Applied Mathematical
Research of the Karelian Research Centre
Russian Academy of Sciences
11, Pushkinskaya str.
Petrozavodsk Karelia
185910, Russia
E-mail: leri@krc.karelia.ru and pavlov@krc.karelia.ru

Austrian Journal of Statistics
published by the Austrian Society of Statistics
Volume 43/3-4
June 2014

http://www.ajs.or.at/
http://www.osg.or.at/

Submitted: 2013-11-25
Accepted: 2014-03-10