# Complex Networks in Recommendation Systems

MASSIMILIANO ZANIN[1,*], PEDRO CANO[1], JAVIER M. BULDÚ[2], OSCAR CELMA[1]
[(1)] Music Technology Group
Universitat Pompeu Fabra
C/ Ocata 1, 08003 Barcelona
SPAIN

[(2)] Nonlinear Dynamics and Chaos Group,
Universidad Rey Juan Carlos
Calle Tulipán s/n, 28933 Móstoles, Madrid
SPAIN

[(*)] massimiliano.zanin@hotmail.com

*Abstract:* - Complex network theory was boosted in 1967 thanks to the experiment of Milgram: since then, the application of this special kind of graphs has given fruitful results in social science problems, from sexual disease control to music communities identification. When focusing on the problem of recommending items to a user (i.e. a customer of an e-store), the underlying transaction data can be seen as a complex network (specifically, a bipartite network): inside this structure, information about customer tastes is codified and can be of good use for future suggestions.

*Key-Words:* - Complex networks, recommendation systems, collaborative filtering

## 1 Introduction

The first real experiment related to the idea of complex networks was realized by S. Milgram in 1967 [1]. He sent several packages to randomly chosen people, asking them to forward it, by hand, to a given person; if those people didn't know directly that person, they should "mail this folder to a personal acquaintance who is more likely than you to know the target person" [1]. The result is the small-world hypothesis: the mean distance between two arbitrary people (that is, the diameter of this social graph) was calculated by Milgram to be as small as 6.

Since then, complex networks have been applied to a great variety of problems, mainly related with social interactions: *e.g.,* sexual disease control [2, 3] or music topology [4, 5]. Other interesting aspects of the real world have been studied with this tool: railway and subway networks [6, 7], streets networks [8], or vulnerability of infrastructure networks [9].

Complex networks theories can be applied to an important problem in computational systems: recommendation. To understand the basis of this problem, we should consider the musical market and go back in time to the sixties; in that decade, there were music stores where shop attendants usually suggested musical groups (that is, items) to customers, using the previous purchases to guess personal tastes. Nowadays, with the increasing number of "virtual" or electronic stores, a user has to face with a huge quantity of items; and, moreover, he has no help while surfing within this ocean of information. In this context recommendation systems have born, as a personalized choice assistant.

## 2 Properties of complex networks

First we must introduce some basic ideas about complex networks. As its name indicates, a network is a group of nodes connected between them by links. When the network is big enough, some complex behavior appears; for instance, the small-world effect: in a big net, the mean distance between any pair of nodes can be very small.

Following are some properties that are useful when studying a recommendation system.

### 2.1 Preferential attachment

The most well known model of network growth is the Barabási-Albert (BA) model, first defined in [10]. The main element is the preferential attachment: when creating a new link, nodes with a higher number of connections (i.e. a higher degree) acquire links at higher rates than other nodes. In other words, the probability that a link will connect a new node $j$ with another existing node $i$ is linearly proportional to the actual degree of $i$:

$$p(j \rightarrow i) = \frac{k_i}{\sum_{j=1}^{N} k_j}$$

where $N$ is the number of nodes and $k_j$ is the degree of the node $j$. This growth model has been successfully applied to information structures like the World Wide Web and the scientific citation network [11]: in both examples, a very popular node is more liked to be linked by new nodes of the network, leading to a *rich gets richer* effect.

### 2.2 Centrality

The concept of centrality was first introduced to quantify the importance (or influence) of a person in a social network [12]. Many measures have been defined in this context; the simplest is the degree: the more a node is connected with other nodes, the more that node should be central for the network.

A more complex approach to evaluate the centrality of a node is the PageRank algorithm [13], developed to quantify the importance of web pages. The underlying idea is to let an agent move through the graph, and make it jump to a random node with some probability $p$: the more the agent visits a node, the more centrality it has.

### 2.3 Bipartite graphs

When we are dealing with recommendation systems, we can identify two distinct kinds of nodes: users (or customers) and items (goods to be sold). They can be represented as two groups of nodes lying in two parallel planes, connected each other with certain links.

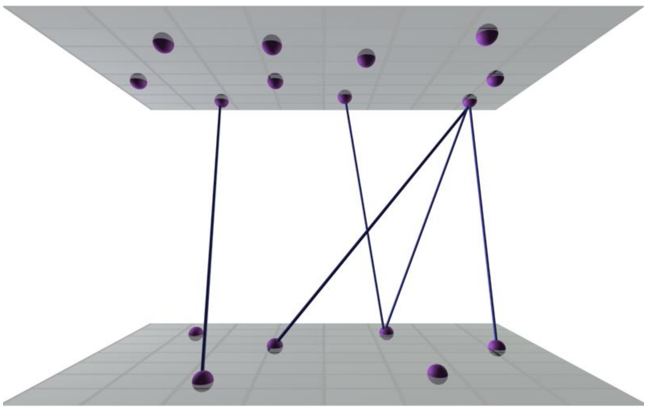

*Fig. 1: A simple representation of a bipartite graph, where two classes of nodes are placed in different planes.*

Such a structure is not directly usable by an algorithm: instead, two complementary projections can be calculated. In the first one, nodes represent items, and two nodes are connected if at least one user brought both items; on the other side, the second represents a user-projection, where users are connected if both have brought the same item in the past.

## 3 Building a recommendation system

The first step to implement complex network theories into a recommendation system is to define a basic algorithm, to be used as the ground to compare any other obtained result. One of the most implemented systems is the *item-based* strategy, that is explained below.

The initial data of the problem is that the system must recommend an item to a target user using the information encoded in the previous brought items. Those items are compared with the others in the network, and the most *similar* are chosen and recommended [14].

Now the problem is to define the *similarity* between two items: there are several ways to compute this measure, while the most popular is the *cosine-based similarity*. For each item, a vector of length $N$ is created, where $N$ is the total number of users; the $n^{th}$ element of the vector has a value of 1 if the $n^{th}$ user brought that item in the past, and a value of 0 otherwise. The distance between two items is then defined as the cosine of the angle drawn by two vectors:

$$dis_i(j) = \cos(\vec{i}, \vec{j}) = \frac{\vec{i} \cdot \vec{j}}{|\vec{i}| \cdot |\vec{j}|}$$

Items with lower distance should be recommended first, as there should be closer to the user's tastes.

### 3.1 Introducing *trendiness*

From a sociological point of view, preferential attachment is like some trendiness force: for instance, a movie that is well known in the market would have a greater probability to be chosen by a user. When the underlying market has such characteristic, we may expect an improvement in the recommendations if preferential attachment is somehow implemented.

With this aim, one should turn over the basic algorithm: now, it would recommend items that have been brought by compatible users, that is, users that share tastes with the target customer. Distance between users is easy to be calculated: it's like the item case, with the only difference that the vectors now represent connections between a user and its items.

The final *compatibility* value of an item is defined as the sum of the compatibility of every user that brought the item in the past:

$$comp(l, user) = \sum_j (1 - dis(j))$$

where $l$ is the item to be recommended to the user, and $j$ accounts for users that have connections with $l$.

From the above, it is easy to see how the more users are connected with $l$, the higher the compatibility value expected. At the same time, the tastes of the target user are taken in consideration: an item connected with users with different tastes (and therefore with high distance) would not be recommended.

### 3.2 The *Aging* problem

The traditional approach when implementing a recommendation system is to use the whole dataset: this is because it is commonly accepted that the more information is included, the better the output will be. Nevertheless, when considering *trendiness* in the recommendation, this could lead to mistakes. One item can have a high popularity at a time $t_0$, but it can lose it that after a certain time $t_1$; if we look at the entire dataset, that item will have a great number of connections, although its popularity in the moment of the recommendation is low.

To control the aging of the data inside the network, the easiest method is to delete all data older than a certain number of days, a number that should be estimated computationally. Moreover, this strategy has a great advantage: by reducing the total amount of information handled, the computation time is reduced drastically.

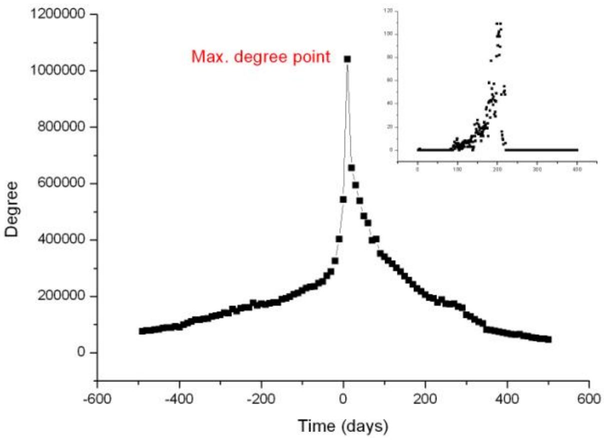


*Fig. 2: Mean evolution of the popularity of an item (NetFlix data [15]); note how there is a peak of maximum popularity, and how after the peak the popularity decreases significantly.*

### 3.3 Weighting links

Following the line of considering time within the implementation, we may go further including the age of the link inside the recommendation. For this purpose, we can assign a weight defined by a function $W$. Despite several functions can be defined, we use the simplest one, a linear decay with time:

$$W(i) = \begin{cases} 1, a_i > \beta \\ 1 + \dfrac{\beta - a_i}{\beta} \alpha, a_i \leq \beta \end{cases}$$

where $a_i$ is the age of the link $i$, and $\alpha$ and $\beta$ are two constants that wheight the importance of the age after a certain time period defined by $\beta$. For each link, we are creating a number that give more relative importance to links with less than $\beta$ days. The new compatibility function should be modified as follows:

$$comp(l, user) = \sum_j (1 - dis(j)) \cdot W(j \rightarrow l)$$

where $(j \rightarrow l)$ corresponds to the link connecting the user $j$ to item $l$.

## 4 Looking for results

After explain which strategies based in complex networks theory can be implemented, we need a way to measure and compare the performance of each algorithm. For a random set of operations, the algorithm should create a list of recommendations, ordered by compatibility; then, the real item (i.e., the one that the user buys after the time of recommendation) of the operation is searched in the list, and a *score* is calculated according to its position:

$$Score = 1 - \frac{Pos_{item}}{\#items}$$

The more the target item is in the upper part of the list (thus representing a good recommendation), the more the *score* approximates to 1.

The three algorithms based on complex networks have been checked with the NetFlix dataset [15]. This is a wide collection of movies rated by users: for our purpose, we have selected only operations where a user has rated 5/5 a movie, and constructed the bipartite network with this links.

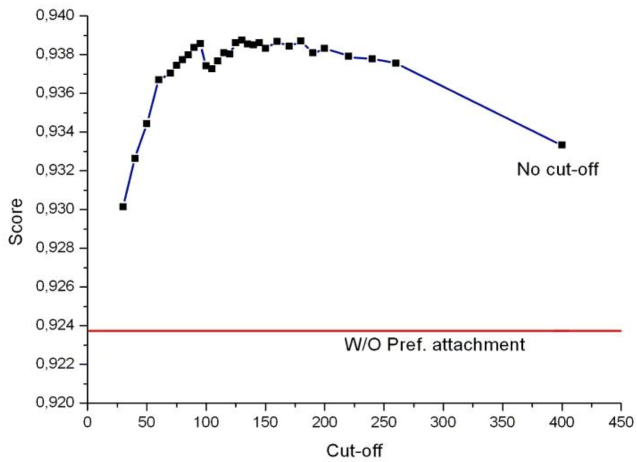

*Fig. 3: Mean score for different aging windows*

In Fig. 3, we plot the evolution of the *score* as a function of the cut-off time (i.e., the maximum number of day that a link can have); clearly, we observe a maximum around 120 days. This fact indicates that using more information introduces mistaken in trendiness datasets reduces the performance of the recommendation. At the same time, short time windows delete too much data, and the score is also reduced. Only for intermediate windows the score increases, leading to an improvement in the performance of the algorithm compared to the basic case, that does not include preferential attachment.

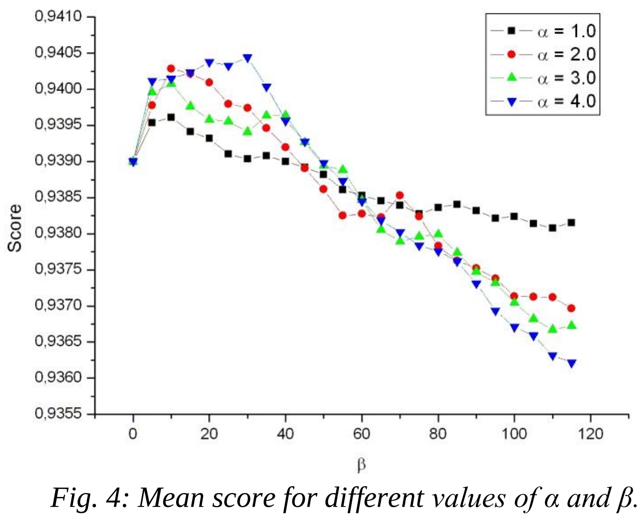

*Fig. 4: Mean score for different values of α and β.*

In Fig. 4, *score* is plotted for different values of $\alpha$ and $\beta$: again, by fine-tuning these two parameters, we can find a region where the mean recommendation is improved.

## 5 Conclusion

Complex networks are a new field of investigation that has been successfully applied to a great variety of problems, mainly in social relations. When facing the challenge of constructing a good recommendation algorithm, that could guide the customer through a great variety of items in an e-store, complex networks can help us in improving the result.

When the underlying market has important aging and trendiness components, a better approximation of user tastes can be obtained by including a preferential attachment strategy. Moreover, it needs only a part of the entire dataset, thus allowing a shorter computational time: this question is especially important in real time recommendation systems.

*References:*


[1] S. Milgram, “The small world problem”, *Psychol. Today*, No. 1, 1967, pp. 61-67.

[2] R.M. May and A.L.Lloyd, “Infection Dynamics on Scale-free Networks”, *Phys. Rev. E*, No. 64, 2001.

[3] R. Pastor-Satorras and A. Vespignani, “Epidemic dynamics in finite size scale-free networks”, *Phys. Rev. E*, No. 65, 2002.

[4] R. Lambiotte and M. Ausloos, “On the genre-fication of music: a percolation approach", *Eur. Phys. J. B*, No. 50, 2006, pp. 183-188.

[5] Juyoung Park, Oscar Celma, Markus Koppenberger, Pedro Cano and Javier M. Buldú, “The Social Network of Contemporary Popular Musicians”, *Int. J. of Bifurcation and Chaos*, 17, 2281-2288 (2007).

[6] P. Sen, S. Dagupta, A. Chatterjee, P.A. Sreeram, G. Mukherjee and S.S. Manna, “Small-world properties of the Indian Subway”, *Phys. Rev. E*, No. 67, 2003.

[7] V. Latora and M. Marchiori, “Is the Boston subway a small-world network?”, *Physica A*, No. 314, 2002, pp. 109-113.

[8] M. Rosvall, A. Trusina, P. Minnhagen and K. Sneppen, “Networks and Cities: An information Perspective”, *Phys. Rev. Lett.*, No. 94, 2005.

[9] Vito Latora and Massimo Marchiori, “Vulnerability and protection of infrastructure networks”, *Phys. Rev. E*, No. 71, 2005.

[10] A.-L. Barabási, R. Albert, “Emergence of Scaling in Random Networks”, *Science*, No. 286, 1999.

[11] D.J. de S. Price, “Networks of Scientific Papers”, *Science*, No. 149, 1965, pp. 510-515.

[12] S. Wasserman, K. Faust, *Social Networks Analysis*, Cambridge University Press, 1994.

[13] S. Brin and L. Page, “The anatomy of a large-scale hypertextual web search engine”, *Computer Networks and ISDN Systems*, No. 30, 1998, 107-117.

[14] Badrul Sarwar, George Karypis, Joseph Konstan and John Riedl, “Item-based Collaborative Filtering Recommendation Algorithms”, *In Proc.*

*of the 10th International World Wide Web Conference*, 2001.
[15] Available at *www.netflixprize.com*