E-commerce network with price comparator sites

Ladislav Beranek
Dept. of Applied Mathematics and Informatics
Faculty of Economics, University of South Bohemia
Ceske Budejovice, Czech republic
beranek@jcu.cz

Radim Remes
Dept. of Applied Mathematics and Informatics
Faculty of Economics, University of South Bohemia
Ceske Budejovice, Czech republic
inrem@ef.jcu.cz

***Abstract* – E-commerce relationships can generally be modeled using a bipartite graph. One part of this chart is made up of customers and the other part consists of e-shops (dealers). Edges show customer activity when visiting websites of various online stores. These links, however, enter online price comparison sites (PCS, comparators, shopbots), such as Heureka. PCS makes it easy to compare the prices of the desired product in different online stores. These comparative pricing not only show the price in different e-shops, but also the assessment of the relevant e-shop made by customers. The issues that this paper addresses are the strategy chosen by the pricing comparator for entering the market, when it is worth closing down online pricing comparison services, etc. In this paper we use network analyzes and simulation methods to model network dynamics to address these issues.**

***Keywords* – network-based inference; simulation; price comparison site, e/commerce.**

I. Introduction

Simultaneously with the development of e-commerce, related business applications have been developed and many new approaches have been widely used, for example, as a recommendation system for various online services or others. Online price comparators, which allow comparison of prices in various online stores, belong to successful applications. These sites are also known as price comparisons, sales points or retailers on the internet. Buyers online use them to get price information or to get user references for the relevant deals. They reduce buyer search costs and help them make decisions by providing price information that is rarely found in the context of physical retail purchases [1].

In recent years, one of the most extensively researched areas are the different types of online networks. Research has, among other things, focused on the characteristics of these networks, their topology, and how links and their types affect the positions of agents in these networks. E.g., what position allows to maximize the influence within a network [2] and vice versa. One of such online systems is the e-commerce shopping and sales network, which can be described using interconnected graphs (networks).

In this paper, we will focus on explaining the formation of the business to customer e-commerce market structure with the emergence of an online price comparator. We will consider two groups of agents, buyers (users) and traders (e-shops). We can model these business networks as bipartite charts that include two sets of buyer and trader nodes (e-shops). The edges link buyers to e-shops. The principles of the network model may also include edges that link merchants to other merchants, but we do not think here. We will now consider a situation where the binder of a commodity that ac-quires a central position enters into the bonds of this bipartite graph. It will then act as an intermediary that will collect and provide users with information on the pricing of different products in a variety of e-shops (merchants). At the same time, the network topology changes.

II. Methods

This article addresses the question: How does a newly arrived information broker (price comparator) integrate into existing bipartite network bindings? How the network will evolve? Will it influence the degree distribution? Our work is based on two assumptions: (a) we are focusing on bi-partite networks that model well e-commerce relationships; and (b) we examine the process of building relation-ships after price comparison sites appear (Figure 1).

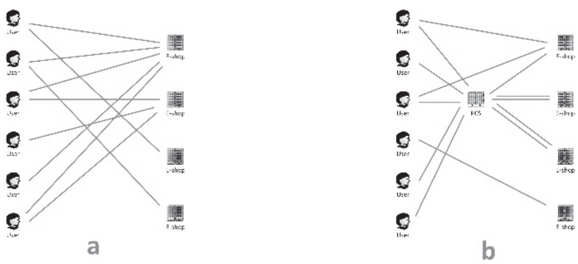

Fig. 1. Model of e-business relationships without (a) and with the presence of PCS (b)

More problems are connected with the described situation: (1) the first one concerns the central position and the way to achieve this position. Here, we can use the frameworks described in [3]. In line with this work, we will consider traders (e-shops) as agents that give newcomers a pricing information about the products offered by these e-shops. The customer will be able to access e-shops via the intermediary's website. (2) The second approach is to model network dynamics. The network will evolve in the form of a sequence of images in discrete time that is determined by the initial network, the custom network evolution path, and the newcomer strategies for adding links. The goal of the newcomer is to adopt a tactic that moves it to the center in a limited number of steps, regardless of dynamic network changes.

A. Related work.

The research of price comparison sites concerned mainly the impact of these comparators on prices of products and services and the sensitivity of Internet customers to the price [4, 5]. Degeratu et al. [6] states that the existence of price comparison sites increases price competition and sensitivity to buyers. We can also look at e-commerce as a network. Social and business networks are increasingly important areas of research in many fields [2, 7, 8, 9]. However, stable equilibrium and models focused primarily on this, while their dynamics and productivity were limited to research. One of the main tasks is to better understand, predict and control their dynamics, including how creates, develops and shapes their behavior and performance [10, 11]. Enough progress

978-1-7281-0450-8/19/$31.00 ©2019 IEEE

has been made in e-commerce applications so far, and e-commerce plays a very important role in the economy. A large number of buyers and sellers collaborate through transactions on websites [12, 13]. These interactions support the development and shape of complex e-commerce structures. Getting a deep look at ecommerce research is of deep and lasting importance.

Most of the work examining the development of network formation is based on gaming-theoretical research and takes into account the balance between rational agents [14]. Graph chart definition is the standard role of graph theory as well as eccentricity. Both scales belong to the family of center-based indexes based on distance. The issue of the core / peripheral structure is dealt with, for example, in [15]. In principle, a particular center can be identified for any real network. The observation of such stupendous structures is based on economics where the world is divided between industrial, "core" nations and agricultural, "peripheral" nations [16]. Similar structures are subsequently recorded, for example, in social networks and business networks [17]. Core agents that are hubs have many benefits, such as information control and resource control. A key feature of the nucleus, in addition to its central location and density, is stability over time [18].

Our article complements these work by examining modeling of relationship changes and related strategy within e-commerce processes after a price comparison site appears.

B. Building a structure in a dynamic network

In general, the dynamic network evolves in discrete time, that is, $G_i = (V_i, E_i)$ is a network instance at a certain time point $i \geq 0$. We define a set of vertices of the dynamic network G as a set $V_G = \cup_{i \in N} V_\mathrm{i}$. Since G can contain infinitely many time stamps, the set of V_G can be infinite. For each vertex $v \in V_G$, the set of neighbors $E_G (v)$ is $\{u \in V_G \mid vu \in E_i,\ i \in N\}$. Since individuals have only limited number of transactions, the set $\mathrm{E_G}$ (v) will be finite for all v. Thus, the graph $(V_G, \cup_{i\in\mathrm{N}}E_i)$ remains locally the final graph.

We will define two limitations: First, basically, any addition or removal of vertices or edge edges can occur. However, we focus on a simpler form of dynamics in this paper, when we only consider adding vertices or edge edges. Secondly, we will only show the network when dynamics modeling occurs when updates occur. The exact meaning should correspond to the actual application scenario. When a customer finds one reseller and clicks (buy or buy the product), he does not search for the goods back on the search engine.

The E-commerce network is modeled as a bipartite graph G (U, V, E) whose vertices can be divided into two disjoint and independent sets U and V such that each edge uv joins a vertex u in U with vertex v in V. We designate the set of edges as E. In our concept, U denotes a set of buyers (users), V denotes a set of e-shops (see Figure 1).

Every user u ∈ U gains the benefit from the G network in terms of communicating information about prices of goods and services connected (information on goods transport, etc.). Although e-shops can only connect to their immediate users, they also use indirect communication streams with those agents (users, e-shops) with which their direct neighbors are connected, etc. The flow of communication or knowledge flows from other agent shrinks away from players represented by the spatial depreciation rate of $0 < \delta < 1$, which captures the idea that the value that results from the connection k_j is proportional to the distance between these two agents. Furthermore, there exists an "internal value" $w_{ij} \geq 0$, that agent i gives agent j. In the following cases, it is assumed that all users and all e-shops are identical except for one so-called "key-player" k, which provides higher value than any other agent (this could be interpreted as leading technology or an e-commerce platform provider).

Imagine an external agent now trying to be an information medium between the two parts of the bipartite network. This means passing information about the prices of certain products to individual e-shops by users. At the beginning, this agent has to choose a way to take the appropriate position. This means ensuring that a certain group of e-shops (referred to as initial seed) initially opts for this agent's information service and provides information about the prices of their products.

More formal description. Let $G_1 = (U_1, E_1)$, $G_2 = (U_2, E_2)$ are graphs (U, V may or may not overlap), we denote the operation of unification with the symbol $\oplus$, i.e., $G_1 \oplus G_2$ indicates the network $(V_1 \cup V_2, E_1 \cup E_2)$. Next, we will mark the new intermediate agent with the sign w.

For any subset $S \subseteq V$ we define $S \otimes w$ as $(S \cup w, \{vw \mid v \in S\})$. So $G \oplus (S \otimes w)$ is the resultant network obtained after the integration of w to G by building the connection between w and every vertex in S (i.e., the middle and a set of e-shops that initially provide information on the prices of their products).

We designate a set of initial seed as $S \subseteq V$. Through the contacts with the members of the initial seed set, w obtains access to the network and a certain central position in the network. S must be of a certain magnitude so that w can act as an intermediary. Generally, this is the problem of a minimum set of brokers that is NP-complete [2].

Creating a relationship with agents of the initial seed set is a dynamic process. Since the relationship requires effort and time, the set of these are iterated, where edges are added at one after the other. Behind the dynamic network G, it would act while the network is developing [19].

Incorporating w into the bipartite network is based not only on your own actions, but depends on the actions of other network agents. At each time stamp, the relationship w and e-shops evolve, and this affects another part of the bipartite network - the user. There are two types of updates, and e-shops send information to the node w, that is to say in our concept of edge creation to w, and users (buyers) create new edges to node w. The last update can be described similarly as creating a relationship with e-shops, i.e., formally for any subset $R \subseteq U$ we define $R \otimes w$ as $(R \cup w, \{uw \mid u \in R\})$. So $G \oplus (R \otimes w)$ is the resultant network obtained after the integration of w to G by building the connection between w and every vertex in R (i.e., the middle and the set of users (buyers).

Correct the initial network $G_0 = (V_0, U_0, E_0)$ and the new incoming w not belonging to V_0 or U_0. For $k, \ell \in N$ is the integration process (IP) is a dynamic network $I = G_0, G_1, G_2,$ ... where $\forall i \geq 0$

$$G_{\mathrm{i}+1} = G_\mathrm{i} \oplus (S_\mathrm{i} \otimes w) \oplus (R_\mathrm{i} \otimes w)$$

where S_i is a subset of vertices in V_i and R_i is a subset of U_i.

Conceptually, IP (integration process) can be seen as an iterative interaction between w and a network. The progress from iteration $i \geq 0$ to $i+1$ will change in the network by (i) "connecting" the parts U_i a V_i of a graph G_i and (ii) adding an edge between w and vertices in S_i and R_i. The sequence of edges (i) and (ii) is called the evolution of the trace and the sequence of S_1, S_2, ..., is called the newly arrived strategy of a newcomer w. IP is uniquely determined by the initial network G_0 and network actions (in the form of tracing evolution). We will define these processes by the following rules:

The network grows as follows. At any time $t > t_0$, a new user or e-shop is added. A new edge is added by the following rules:

a) a percentage $\alpha \in [0; 1]$ of them at random;

b) the rest by linear preferential attachment according to e-shop's degree, so the higher the current links to a e-shop, the higher the probability to be visited by user;

c) when user visit PCS, the two edges are created, from user to PCS and from PCS to e-shop with the use of linear preferential attachment.

The question arises as to how a newcomer (PCS) can choose his strategy during the initial integration process to get the more link as possible. To analyze the factors that impart tactical performance, we run following adapted models of dynamic networks in our simulation:

- Dynamic Barabási-Albert (BA) model [20]. This well established dynamic model takes the parameter d ∈ N and adds a new vertex to each time stamp that randomly associates d vertices with the preferential attachment mechanism. With multiple iterations, the graph develops a scale-free property, but does not achieve a highly clustered core.
- Dynamic Jackson-Rogers (JR) model. This model proposed by [21] simulates stochastic friendship making among an agent population. An agent may link with a friend of friends or a random individual. At each timestamp, the model randomly samples for every vertex v a set $S_1(v)$ of m nonadjacent vertices from the entire network, and another set $S_2(v)$ of m vertices who are at distance 2 from v ($S_1(v)$ and $S_2(v)$ may not be disjoint). It then builds edges between v and every vertex in $S_1(v) \cup S_2(v)$ with probability p. As argued in (Jackson and Rogers, 2007), the model meets most of the desired properties such as scale-free and small-world properties. The value $m \approx d/4p$ relies on p and an expected average degree $d \in N$ which are parameters of the model. We pick $p = \{0.25, 0.5, 1\}$ to resemble the fitted values on the real-world networks in [21].
- Dynamic rich club. The rich club is a "go-to" model of a core / peripheral structure that develops a dense central core with a thin edge [22]. At each time stamp, the process adds a new vertex with probability $\alpha \in [0, 1]$ (and links it to a random vertex), or a relationship between two existing vertex with a probability of 1 - α. We choose in the second case a random source $w \in V$ and connect it to the target in the following way: For each $k \in N$ we set $[k] = \{v \in V \mid \deg(v) = k\}$; the probability that $z \in [k]$ is $\alpha\ k\ [k]$. The probability α, calculated as $\alpha = 2(N + 1) / (Nd + 2)$, depends on the target diameter of the d scale and the N-size graph, which are the model parameters.
- Dynamic onions. The onion is a core / periphery structure, but unlike in a rich club, the peripheral vertex are joined here, forming the e or several layers surrounding the core, reminiscent of very elastic meshes, for example, criminal rings [23]. The original model of a static onion creates a network with a fixed distribution of degrees of power law $q(k) \sim k - \gamma$ (where $\gamma \in R$ depends on the average degree d). We modify this model by following ways. In each time interval, (1) we add a new vertex v, whose degree is $\deg(v) = k$ with the probability $q(k)$; (2) to add v to G while preserving the degree distribution, create a pool of "half-edges" initially containing k studs attached to v; (3) randomly severe k existing edges into $2k$ vertex which are added to L; (4) repeatedly "join" random pairs of vertex v,w in L to form edge vw with probability $p(vw) = (1+3|sv-sw|)-1$, taking care to avoid self-loops and duplicates, until $L = \emptyset$ (Wu and Holme, 2011).

III. Research results

In this section, the numerical results are presented in Figure 2. All the results are the average of 10 simulations for different realization of e-business networks under the same parameters. The step of all generated networks all reach 4000 nodes. We have also did the simulation to 1000 steps. A network with 4000 nodes could give us a nice description for asymptotic distribution. The degree distribution of authors is far from a power law distribution. The simulations are consistent with the scale free property observed from empirical data.

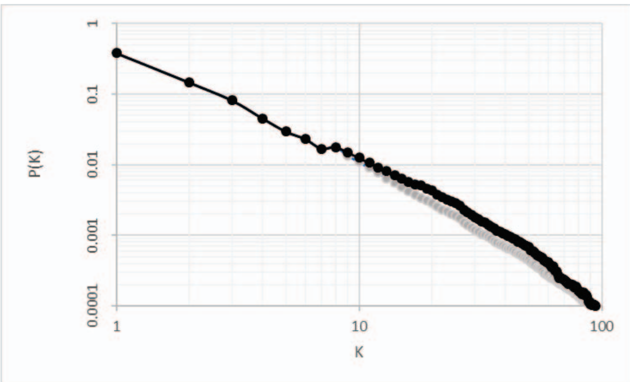

Fig. 2. Degree distribution without the presence of a price comparison site (gray color) and with the presence of a price comparison site (black color)

Numerical simulations indicate that this network evolves into a scale-invariant state with the probability that a node has k edges (here e-shops, see Figure 2) following a power-law with an exponent $c = 1.43$ The scaling exponent is independent of m, the only parameter in the model. Fit-get-richer mechanisms are proposed which is better adapted to model certain networks where topological properties are essentially determined by ''physical'' information intrinsically related to the role played by each node in the network, such as the ability of an individual, the content of a web site, or the innovation of a scientific article.

In this paper, we focus on B2C e-commerce market network which involves three kinds of nodes, i.e., e-business website nodes and buyer nodes with the presence of price comparison sites. In the network model with two layers, each buyer node connects to at least one website node, even much more. We neglect the direct internal relations in each layer. It implies that the competitions among websites exist by way of attracting buyers. The theoretical background is drawn from niche competition theory. In the theory, businesses that have the same niche (resources, customers or market share) compete with each other. Figure 2 shows the course of the degrees of price comparison sites. The position depends on the strategy of during integration process to get into the network center (seed sets, the time of entering into e-commerce relationship.

IV. Conclusions

So far there are very limited studies about e-commerce market with a network science perspective. Our study is a multi-disciplinary research intersected by marketing, economy and network science. We believe that the study of a real e-commerce market network will lead to further research that will reveal the hidden mechanism of economic and social system. Based on our empirical analysis of e-commerce market, we modify and expand evolutionary mechanisms of network evolving model. Our model can not only reveal how the structure of e-commerce market in the presence of price comparison site takes shape.

The fitness in our model is static. However, in real market, websites are growing and adjusting themselves, and dynamic fitness should be researched in future. In addition, the prerequisite for our research is that social, economic and technical environments are almost stable for a time. Environmental impact should be considered for future work.

Network dynamics modeling poses a number of challenges. Many future works remain: (a) it is a natural question to explore dynamic models where separate links are added; (b) there is a difference between the concept of core network and center [16]. The future question would be to investigate tactics that instead of the newcomer to the kernel, rather than the center of the network; (c) the community structure is another predominant property on a scale and the same issue could be focused on dynamic models of community structure; (d) transferring from a single agent's tactic to a population of agents can formulate and explore game theoretical models of networking based on social capital concepts.

A further challenge arises from the using real networks data about e-commerce. While available online networks can include thousands or millions of users, and thus give strong statistical correlations, detailed information on why users form links is usually lacking. Thus, it is difficult to distinguish links arising from prior similarity from influence of linked individuals creating similar preferences. In our future work, we want to explore further insight into the network of interest and make a more dynamic analysis of the network possible.

References

[1] E. Brynjolfsson, M.D. Smith, “Frictionless Commerce? A Comparison of Internet and Conventional Retailers”, Management Science, vol. 46, pp. 563–585, 2000.

[2] S. Borgatti, D. Halgin, “On network theory”, Organization science, vol. 22, no. 5, pp. 1168 – 1181, 2011.

[3] A. Moskvina, J. Liu, “How to build your network? a structural analysis”, In: IJCAI-2016, AAAI Press, pp. 2597–2603, 2016.

[4] M. Takayuki, W. Tsutomu, “Why Are Product Prices in Online Markets Not Converging?:, PLOS ONE, vol. 8, no. 8, pp. 1 – 7, 2013.

[5] B.K. Pathak, “Comparison Shopping Agents and Online Price Dispersion: A Search Cost based Explanation”, Journal of Theoretical and Applied Electronic Commerce Research, vol. 7, no. 1, pp. 64 – 76, 2012.

[6] A.M. Degeratu, A. Rangaswamy, J. Wu, “Consumer choice behavior in online and traditional supermarkets: The effects of brand name, price, and other search attributes”, International Journal of Research in Marketing, vol. 17, no. 1, pp. 55–78, 2001.

[7] D. Easley, J. Kleinberg, Networks, Crowds, and Markets. Cambridge: Cambridge University Press, 2010.

[8] V. Lamanauskas, V. Slekiene, A. Balog, C. Pribeanu, “Exploring the Usefulness of Social Networking Websites: a Multidimensional Model”, Studies in Informatics and Control, vol. 22, no. 2, pp. 175-184, 2013.

[9] M. Drlik, L. Beranek, “The role of the online social network in informal learning”, In: Proceedings of the 12th Conference on Efficiency and Responsibility in Education (ERIE). Czech Univ. Life Sci. Prague, June, 2015, pp. 81 -86, 2015.

[10] J.E. Rauch, “Does Network Theory Connect to the Rest of Us? A Review of Matthew O. Jackson's Social and Economic Networks”, Journal of Economic Literature, 2010, pp. 980-986, 2010.

[11] F. Schweitzer, G. Fagiolo, D. Sornette, “Economic Networks: The New Challenges”. Science, pp. 325 – 422, 2009.

[12] L. Beranek, V. Nydl, “The Use of Belief Functions for the Detection of Internet Auction Fraud”. In: Proceedings of the 31st International Conference Mathematical Methods in Economics 2013. Jihlava: College of Polytechnics Jihlava, 2013, pp. 31 – 36, 2013.

[13] L. Beranek, P. Tlusty, R. Remes, “The Usage of Belief Functions for an Online Auction Reputation Model”. In: Proceedings of the 28th Conference on Mathematical Method in Economy 2010. Economical faculty. University of South Bohemia, Ceské Budejovice, Czech Republic, September 2010, pp. 49-54, 2010.

[14] M. Jackson, “Social and economic networks”. Princeton university press, 2010.

[15] S. Borgatti, M. Everett, “Models of core/periphery structures”, Social networks, vol. 21, no. 4, pp. 375–395, 2000.

[16] P. Krugman, A. Venables, “Globalization and the inequality of nations”, The Quarterly Journal of Economics, vol. 110, no. 4, pp. 857–880, 1995.

[17] D. Fricke, T. Lux, “Core–periphery structure in the overnight money market: evidence from the e-mid trading platform”, Computational Economics, vol. 45, no. 3, pp. 359–395, 2015.

[18] P. Rombach, M. Porter, J. Fowler, P. Mucha, “Core-periphery structure in networks” (revisited), SIAM Review, vol. 59, no. 3, pp. 619–646, 2017.

[19] B. Yan, Y. Chen, J. Liu, “Dynamic relationship building: Exploitation versus exploration on a social network”. In International Conference on Web Information Systems Engineering, Springer 2017, pp. 75–90, 2017.

[20] R. Albert, A.-L. Barabási, “Statistical mechanics of complex networks”, Reviews of modern physics, vol. 74, no 1, pp. 47, 2002.

[21] M. Jackson, B. Rogers, “Meeting strangers and friends of friends: How random are social networks?”, American Economic Review, vol. 97, no. 3, pp. 890–915, 200).

[22] P. Csermely, A. London, L. Y. Wu, B. Uzzi, “Structure and dynamics of core/periphery networks”, Journal of Complex Networks, vol. 1, no. 2, pp. 93–123, 2013.

[23] Z.X. Wu, P. Holme, “Onion structure and network robustness”, Physical Review E, vol. 84, no. 2, 026106, 2011.