Topological Analysis of Online Auction Markets

JAMES K. HO

WHAT SHAPE IS AN ONLINE AUCTION MARKET IN?

The Internet is a new medium of communication connecting potential partners in trade worldwide. The initial frenzy over its promises led to grossly exaggerated valuations of business models that were mere transplantations of existing processes to the alternative channel. Now that the bubble has burst (Perkins and Perkins 2001), more sensible and critical thoughts can be turned to true transformations that are creating and nurturing markets of the future. Online auction is one of the very few cases that has held a steady course (Klein 1997, Lucking-Reiley 2000, Turban 1997), as evidenced in the success to date of eBay.com. Founded in September 1995, eBay has become a global trading platform where on any given day, there are more than 16 million items listed across 27,000 categories. In 2003, at least 30 million people bought and sold well over $20 billion in merchandise, so that the entire culture it engenders is now being described as the eBay economy (Hof 2003). It is also the richest source of data for online auction markets, as records of all its transactions are available to the public (on a 2-week rolling basis). We pose the intriguing question of what ‘shape’ a given market is in at a particular moment of development.

This work aims to construct a topological model based only on operational data, without any expert knowledge of the specific auction market, or economic details from the transactions. Using extensive analysis of eBay data, the dimensions for a topology are identified. Microsoft Excel macro programs are designed and used to perform the substantial data mining necessary to extract the information desired. A graphical model is then proposed to visualize this topology, for the purpose of comparing markets at different points of a life cycle. Examples of application of the model are presented.

VISUALIZATION IN HIGH-DIMENSIONAL DATA MINING

Visualization has been a fast developing approach in data-mining (Hoffman and Grinstein 2001) in which graphical models are constructed to provide visual cues for pattern recognition and knowledge discovery from complex data. In the study of financial markets (stock and commodity), the dimension of interest is primarily prices, or the fluctuation thereof. Complexity arises from the large number of instruments involved. The best known examples of visualization models for stock markets are based on the tree-map...
method (Shneiderman 1992, Wattenberg 1999), and the minimum-spanning-tree method (Vandewalle et al. 2001). For auction markets, the game-theoretic dynamics itself gives rise to higher dimensional complexity. And with online auctions removing conventional constraints on time and space, their activities and impact on e-commerce can be expected to grow exponentially. Therefore the availability of operational data from eBay presents unprecedented challenges and opportunities for insight into online auction markets.

We define a topological model for an online auction market to be a simultaneous graphical display of all the dimensions of its relevant database, which provides a geometrical shape as a descriptive, visual statistics of the market. The purpose of this work is to identify the dimensions from available data, and to construct a specific topological model that can help discern market efficiency. We seek a visual cue for whether a market is favorable to buyers or sellers without expert knowledge of the items involved or the prices attained.

THE DIMENSIONS OF ONLINE AUCTION MARKETS

Gross activity

While the number of auctions taking place per day is the most direct measure of the level of activity for any given market, the meaning and significance of this number depends largely on how clearly such a market is delineated. Since eBay provides over 27,000 categories of auction items, it may seem trivial to define a market a priori, e.g., digital cameras, and expect to collect data for such a specific item. Actually, the underlying database is keyword driven so that substantial cross-listing can result for items which may only be remotely related. Even with progressively refined drilling to filter extraneous selections, it is impossible, short of item-by-item scrutiny, to ascertain exact pertinence of the collected data to the prescribed market. For example, even if the search for digital cameras were restricted to a specific brand and model, it would still fail to preclude accessory items such as carrying cases. If a filter is used to block out carrying cases, then it may exclude cameras with carrying cases.

For the purpose of our exploratory studies, this inherent fuzziness in market definition should not pose any significant problem. For more focused applications, additional caution must be exercised.

As we intend to study auction markets at particular points in their development, the time span for the collected data should not be too long. However, to ensure accurate representation of the actual dynamics of the markets, enough data points should be included. From both logistical and statistical considerations, we chose a target size of 500 auctions. Note that these are not samples, but complete records taken over a contiguous interval of time. Four markets are used for illustration throughout: jewellery rings, digital cameras, Swiss watches, US coins. The data are from autumn 2002 over a period of 10 to 20 days. Figure 1 shows the average number of auctions per day. These markets were chosen for their contrast in antiquity versus technology, vanity versus practicality, to reflect the actual diversity of online auction markets on eBay. They were also relatively well-focused for illustrative purposes. Random samples of size 100 indicated a fuzziness factor (percentage of related but irrelevant items) of 0%, 6%, 4% and 0%, respectively.

Net activity

As a categorical measure of the effective activity in an auction market, the proportion of listings that ended with at least one bid is used (see Figure 2). In cases where a reserve price is in effect, the listing is counted whether the reserve is eventually met or not. Listings that generated no bids are obviously not contributing to the sustenance of the market. They may represent items that are of no interest to any buyer, or with unattractively high starting bids. Since there are listing fees, the proportion of no-bid listings should stabilize in the long run. Items of no appeal will cease to be offered. Unrealistic price expectation will be adjusted. In the short term, especially in emerging markets where a critical mass of buyers has yet to develop, or when relatively new items are being introduced, fluctuations of net activity is expected. In the parlance of statistical quality control, no-bid listings can be regarded as ‘defective products’. Appropriate control

![Figure 1. Gross activity: average number of auctions per day](image-url)
charts can be set up to monitor whether the proportion of no-bid listings over time is within normal statistical fluctuation or not. In our examples, the rings market has significantly higher gross activity but lower net activity than digital cameras.

**Participation**

For those auctions that resulted in at least one bid, an aggregate measure of activity is the average number of bids (See Figure 3). Multiple regular bids from the same buyer are counted separately. However, a proxy bid (see below) is counted only once, no matter how many automatic bids it generates in response to subsequent bids by other buyers. In our examples, digital cameras show a significantly higher level of participation than the other three markets.

**Seller diversity**

For a sense of whether a market is dominated by relatively few sellers, or there is a diverse source of supply, the number of auctions in the data collection offered by each individual seller is tracked. Their market shares, as a percentage of total auctions in the data set, are tallied in four categories: i) <1%; ii) 1–3%; iii) 3–10%; and iv) >10% (see Figure 4). The proportions of sellers in these four categories are then used as indicators of seller diversity: the more sellers in the lower categories, the more diverse the sources of supply. The choice of the categories is arbitrary, and can be parameterised to suit more specific purpose of analysis.

Sellers are identified only by their eBay user names, which can be established by a simple process of registration that is free of charge. Therefore, it is conceivable that certain sellers are operating under multiple identities. Also, these indicators are dependent on the particular time frame for the data set, and may not reflect the sellers’ market share on any longer term basis. In all four of our example markets, there are relatively few ‘dealer’-type sellers offering more than 10% of the auctions, while most others are one-time or occasional sellers.

**Seller experience**

Buying and selling over the Internet, dealing with people one will never meet, often under the most frivolous and improbable user names, is certainly not conventional business by any means. Why would anyone send money
to total strangers in the hope that they will deliver the goods? Indeed, early critics were quick to predict how it could never work. The explosive growth of transaction volumes on eBay proves them wrong (Resnick and Zeckhauser 2002). Through a peer evaluation system known as ‘Feedback Ratings’, a culture of trust in cyberspace has taken hold, enabling participants to gauge one another’s trustworthiness based on their own track records. At the completion of any auction, the buyer and seller have the opportunity to post feedback to eBay. The experience of the transaction can be rated ‘Positive’, ‘Neutral’, or ‘Negative’, along with a brief (up to 80 characters) comment. The total number of positive feedbacks from unique users, minus the total number of negative feedbacks from unique users gives the net feedback rating of an eBay user (eBayer in short). This number, given in parentheses after the user name, and updated as new feedbacks are posted, is the single most important indicator of a user’s experience and reputation (Houser and Wooders 2001). The number is linked to the complete track records, including individual comments, for anyone who wishes to examine the details. In addition, on all auction listings, the seller’s standing on positive feedbacks as a percentage of the overall rating, as well as the date and country of registration are displayed.

Since the feedback rating is the net result of positives minus negatives, and not affected by neutrals, it is insufficient on its own to represent the reliability of the user. The same number may conceal any percentage of negatives or neutrals. Therefore, we prefer to use it as a surrogate measure of the user’s experience — the positive experience as perceived by other users to be exact. This experience covers all transactions involving the user on eBay, not just in the market under study. The feedback ratings of sellers identified in the data set is classified into four arbitrarily chosen ranges: [0–5], [6–25], [26–100], (>100), to represent four levels of experience — from ‘Beginners’ to ‘Veterans’. The proportions of sellers at each level are compiled for each market (see Figure 5). In our examples, all four markets are dominated by experienced sellers, although the distributions are less skewed for digital cameras and Swiss watches.

### Buyer diversity

For a sense of the distribution of buyer participation over the auctions in the data set, the number of auctions in which each user placed a bid is tracked. Typically, each identified user participated only in a small percentage of the total number of auctions. Therefore the absolute number, rather than the percentage, is used. Again, four ranges are arbitrarily chosen: [1], [2–5], [6–10], (>10]. The proportions of buyers falling within these ranges are recorded (Figure 6). For the markets in our examples, most buyers participated in only a single auction, and very few were involved in five or more.
This is completely analogous to the measure of seller experience. The feedback ratings of buyers identified in the data set are classified into the same four categories, and the proportions of buyers at each level are compiled for each market (Figure 7). Apart from specialized professional sellers with feedback ratings in the thousands, and exclusively for selling, a typical eBay user with moderate experience would have participated as both buyer and seller. When referring to a user’s experience, as either a buyer or seller in any particular auction, we actually invoke the overall experience. Also, as any user’s rating might have changed over the duration of the data set in question, the latest rating is used in the compilation of the experience profiles for the data set. With further data mining, it is possible to isolate a user’s experience in specific markets, which can in turn be used as a proxy measure of the user’s expertise in such markets. This extension is beyond the scope of the present work and is not used in the topological analysis. In our examples, buyer experiences are more evenly distributed than seller experiences in all cases. Note that for coins, experienced buyers dominated the others, unlike the other three markets.

Matching

Since the intent of auction markets is for competing bidders to arrive at a fair value for the item offered, an auction that ended with a single, winning bid does not make for an interesting or typical case. However, it does signal a unique match in supply and demand, although this match is typically imperfect. The item may not have enough appeal to generate significant interest among potential buyers. The eventual buyer, not considering the item as an ideal find, may enter the bid as a ‘low ball’ strategy, hoping for a better than expected bargain. Had the item been indeed an ideal find for the buyer, the seller must have set the starting bid lower than what the buyer would have been willing to pay. In either case, the auction, though resulting in a transaction, fails to provide a reliable indicator of the supply and demand, as well as a fair market value for the item. Just like auctions with no bids, these single bid auctions should be a self-regulating subset of the given market, having equilibrium levels over the long term. Over the short run, the proportion of auctions with a single bid can be regarded as a measure of how well sellers have studied the market, for example by reviewing records of completed auctions, before listing items and setting prices for the starting bid (Kamins et al.).
On the average, about a quarter of the auctions in our dataset were single-bid matches, which seems significantly high, especially for coins at 35.3% (see Figure 8).

Duelling

The opposite situation to single-bid matching is when there is keen competition. Aside from the total number of bids, it is of interest to gauge the degree of ‘head-to-head’ competition. While the actual dynamics of such bidding patterns, in terms of who is challenging and who is responding, can get quite complex and combinatorial, a tractable criterion can be defined for the occurrence of ‘duelling’ between two pre-eminent bidders. We consider duelling to take place when the two bidders with the highest number of bids together account for the majority (or some other pre-set level) of the bids, provided the total number of bids exceeds some pre-set threshold (6 is used throughout this study). The last provision is to exclude trivial cases, for example when there are only two bidders with one bid each. It should be remarked that neither of the bidders involved in such duelling needs to be the eventual winner. However, any amount of duelling at any point in an auction should contribute to raising the perceived value of the item, as it is well-known that the heat of competition may lead bidders beyond rational valuations (Bajari and Hortacsu 2003). In any case, legitimate duelling in any form and extent is a welcome phenomenon for a seller, as well as a sign of vitality of the actual auction market. Using the parameters set for this study: the two most active bidders placing 50% or more of the total number of bids of six or more, the proportion of auctions with duelling in each data set is compiled (Figure 9). There was a significant amount of duelling in all four of our example markets, especially for digital cameras at 53.8%.

Stashing

In markets for items whose value will increase with scarcity, such as collectibles, certain buyers may set out — to whatever extent — to ‘corner’ the market. They hope to build up a stock and profit from higher valuations in the future. Or there may be professional dealers with brick-and-mortar sales outlets who have managed to arbitrage actual price differentials between eBay and the conventional marketplace. In either case, ‘stashing’ activities would be reflected in frequent purchases of the same or similar items within the time frame of the dataset by such buyers. It is interesting to observe on eBay...
revealing user names such as ‘xxxmanyyy’, where ‘xxx’ is the item in question, and ‘yyy’ is some extension to distinguish from others of like mind. Obviously, it would be impossible to ascertain the intent and purpose of any transaction in general, and in particular, where the purchase stands relative to actual fair market value. A stasher with foresight may be able to pick up bargains, for example from single ‘matching’ bids, before the item garners popular interest. Or he or she may outbid others in ‘duelling’ bids, perhaps overpaying for the moment, but recognizing the profit potential for the future. Our topological approach specifically circumvents such economic details, and simply relies on the operational data to derive an indicator of stashing, namely, the highest percentage of the auctions in the dataset won by any single user. It should be remarked that if for any reason stashers wish to disguise their purchasing patterns, they can easily assume multiple user identifies. In our examples, the evidence of stashing (Figure 10) is most apparent for coins, and least for rings, again quite in keeping with the nature of the items.

Sniping

The specific design of online auction will affect the participants’ behaviour, and hence the topological shape of the market. Given the gaming situation, there are optimal strategies, rational as well as perceived. Users, regardless of demographic background, tend to adapt quickly to the game, and device their own ‘winning’ strategies. Three parameters for the eBay online auction system lead to an interesting and significant phenomenon. First, it is a second price auction, which means the winner does not pay his or her highest bid, but rather that of the second highest bidder plus a prescribed increment. For example, with an increment of $1, if the highest bid is $110 and the second highest is $100, then the winner pays $101. Second, it is a fixed deadline auction, which terminates at a precise, posted moment. Third, with few exceptions (known as private auctions), the list of bidder identities, though not the bid amounts, is posted throughout the auction. For a combination of reasons, there is advantage for bidders to wait until the very end to enter a first, and hopefully winning bid. This is known as ‘sniping’ and has been well observed and studied (Roth and Ockenfels 2002).

The appearance of a bid by a user who is perceived as an expert in the market can instill confidence in other bidders for the valuation of the item. By hiding one’s interest and intent to bid until the last second, the knowledgeable user can avoid sharing such expertise and raising the price of the item. It takes time for the out-bidden to respond with a higher bid, as in the sequential process of duelling. A last minute bid can preclude further challenges and terminate the auction at a lower winning price than attainable otherwise. Even without any game theoretic knowledge, users of all ages quickly learn of its potential advantage. Many practice it for the added psychological satisfaction of a surreptitious victory. While it may take nerve and patience to execute this strategy manually, there are actually software programs and online services to enter such bids automatically. All one needs to do is to specify the eBay auction (by its unique number), and the amount of the bid (Steiner 2002).

For our analysis (see Figure 11), we consider an auction to be won by sniping if the winning bid was placed within one minute from termination, and it was the first and only bid placed by the bidder. Note that a bidder who placed an earlier bid or bids may still return later and use sniping to win, but such non-pure sniping strategies are not tracked in this study. Also, there is no accounting of failed sniping attempts, as these would have been rejected by the system on arrival as being out-bidden. In our examples, there were higher incidence of sniping with digital camera and coins than with rings and watches.

Retailing

On eBay, sellers are provided with an option to offer the item at a fixed price. This is known as Buy-It-Now (BIN). Whether this removes the auctioning element
altogether depends on the minimum bid specified. If the starting bid must match the BIN price, then it is strictly a retailing offer. If the starting bid is lower and there is no reserve price, then the BIN option is available until the first bid is placed, at which point, the listing reverts to a regular auction. If there is a reserve price, then the BIN option remains available until the reserve is met. While this option clearly detracts from the true spirit of online auctions, its viability is certainly self-regulating, and its effectiveness affords us another glimpse of the shape of the market. The proportions of auctions won by Buy-It-Now in our datasets are compiled (Figure 12). Not surprisingly, it is highest for standardized items such as digital cameras, and lowest for collectibles such as coins. Note that the count of single bid matching auctions discussed previously excludes the ones ended with Buy-It-Now. The main difference is that BIN is an outright purchase, while matching is a regular auction with the single bidder having to wait till the end to learn that there was no competition for the item. In our example for digital cameras, which are standardized products with well-known valuations, the proportion of BIN was relatively high at 13.7%. Whereas for coins, which are collectibles with more subjective valuations, a relatively low rate of BIN at 3.5% was observed.

Proxy

As e-commerce is supposed to help remove the restriction of space and time in business processes, it makes good sense not to expect participants of online auctions to monitor their progress in real time and on a continual basis. One way to relieve users of any excessive attention span is the use of proxy bidding. On eBay, a user can submit a maximum bid of what he or she is willing to pay for the item. Then, as necessary, the system will place the appropriate incremental bid for the user against other incoming bids, until the maximum is reached. Theoretically, if every bidder can submit a true valuation of the item: exactly what it is worth to the bidder, no more, no less, then all is well. If the bid wins, that is wonderful for the bidder. If it does not, even just by 10 cents, there should be no regrets, since in principle the bidder is not willing to pay a penny more. The seller can also rest assured that the best price is obtained for the item, since every bidder revealed his or her true valuation. That is why eBay, as promoter of such markets for its own interest, is the strongest advocate of proxy bidding. However, human psychology does not function quite that way in reality. When outbid by the marginal increment of say one dollar, a bidder would inevitably lament that he or
she would have gone that much further, not recognizing that the opponent is likely to go even higher. Also, there is a tendency to avoid or at least defer the revelation of one’s true valuation in the hope of winning with less. This effect is compounded by the apparent advantage of sniping discussed above. Therefore, the use of proxy bidding is probably out of necessity and convenience as the bidder cannot monitor the actual auction, more than out of theoretical conviction of the rationality of the strategy. In any case, we track the proportion of auctions in our dataset won by proxy bids (Figure 13). These are indicated by winning bids that are not the last ones placed chronologically. The relatively high proportions observed in our example markets, up to 41% for digital cameras, reflect the fact that many bids placed near the end of an auction, as long as they beat at least one other that came later, are considered winning proxy bids by definition.

A TOPOLOGICAL MODEL OF ONLINE AUCTION MARKETS

Our topological model is based on the star plot for displaying multivariate data with an arbitrary number of dimensions (Chambers et al. 1983). Each data point is represented as a star-shaped figure (or glyph) with one ray for each dimension. As the resulting shapes depend on the configuration of the dimensions, we further analyse the observations along the dimensions identified above in an effort to present a visual model (see a generic example in Figure 14) of the shape of online auction

![Figure 13. Proxy: percentage of auctions won by proxy bids](image)

![Figure 14. Topology of online auction market](image)
markets. This is done by dividing the dimensions into a buyer-seller dichotomy.

Gross and net activities are significant for inter-market comparisons, and neutral to buyers and sellers. We map gross activities, on some relative scale for all online auction markets, on the upper vertical axis, and use it as the radius of a circle to frame the visual model. Net activities, as percentage of auctions with bids, are mapped on the same vertical axis. Participation, expressed as the average numbers of bids per auctions, is mapped on the lower vertical axis. This way, the size of the circle represents the size of the market relative to all others, whereas the span between the net activity and participation indicates its vibrancy.

Buyer dimensions

The seller profiles, in terms of diversity and experience, are buyer dimensions. For diversity, we make the assumption that dominance in market share by relatively few sellers does not provide as much sourcing opportunities for buyers as when there are more sellers. Analysis of data from over 30 markets led us to propose the use of the proportion of sellers offering 1% or less of the auctions as the metric for diversity. For experience among sellers, the rationale is that an even distribution is favourable to buyers. The skewness of any observed distribution can then be normalized against the even distribution, which is top of the scale. In particular, the maximum of 0.25 which equals 0.0039 is normalized to 1.0 so that a distribution of, say, (0.4, 0.3, 0.2, 0.1) will be normalized to 0.61 on this scale. The two seller profile dimensions are mapped radially in the top-right quadrant in Figure 14.

Single-bid matching, though not an indicator of the vitality of a true auction market, does work in the favour of buyers. It meets their demand at opening bid prices, which are likely to be bargains. The success of winning bids by sniping also puts buyers at an advantage. As demonstrated by Roth and Ockenfels (2002), the winning bid prices can be expected to be undervalued. Buy-It-Now options, if acceptable to buyers, necessarily reflect favourable prices, and so can be ruled a buyer dimension. These three buyer dimensions, all expressed as percentage of total auctions, are mapped in the lower right quadrant in Figure 14.

Seller dimensions

The buyer profiles, in terms of diversity and experience, are seller dimensions. Analogous to the buyer dimensions we make the assumption that diversity in participation and an even distribution in experience among buyers are favourable to sellers. For diversity, we use the proportion of buyers participating in a single auction as the metric. For experience, the skewness of the distribution is normalized against the even distribution as described above. These two dimensions are mapped in the top left quadrant in Figure 14.

Dueling, which signals competitive bidding to raise the winning price is obviously to the advantage of sellers. Stashing by buyers with ulterior motives, which are not necessarily rational in the actual market, is also a seller dimension. Finally, proxy bidding can both attract bidders who would otherwise be too busy to monitor an auction, as well as elicit true valuations from potential buyers to help maximize the final price, and hence a seller dimension. These three seller dimensions, all expressed as percentage of total auctions, are mapped in the lower left quadrant in Figure 14.

Examples of applications

The graphical display of the topological model of online auction markets provides a visual aid in describing the ‘shape’ of a market at any particularly stage of its development. The dimensions (or attributes) deemed advantageous to buyers and sellers are mapped on the right and left sides of a circle, respectively. The relative areas of the two sides of the star glyph provide a visual cue to market efficiency: a larger right side favours buyers, and vice versa. By themselves, such snapshots encapsulate market characteristics, especially when large numbers of markets are to be studied. However, their potential utility is in comparative analysis, as demonstrated by the following two examples.

The first example is from inter-brand comparisons. Datasets for two different brands of digital cameras: Sony and Nikon were studied in autumn 2002. The topological models are displayed in Figure 15. The Sony brand had both higher gross and net activities, hence a larger market. However, the Nikon brand had higher participation rate, with an average of 19.7 bids per auction compared with Sony’s 7.2. Aside from this, the markets had very similar shape, except for buyers experience. Sony had a significantly more even distribution than Nikon. Closer examination of the data revealed that nearly half of the bidders for Nikon were beginners.

The second example is from studying a given market over time. Datasets for travel and vacation packages from November 2002 and October 2003 were compared, with results shown in Figure 16. While there was significant growth in the size of the market, we observed that both net activity and participation remained relatively low. The general shape on the buyers side remained stable. However, there was movement on the sellers side, with substantial increase in duelling, but fewer winning bids by proxy. On balance, the topological comparison suggested that the market had improved for sellers.
DISCUSSION

We presented a topological model of online auction markets based only on operational data, without any expert knowledge of the specific items, or economic details from the transactions. A graphical rendition of the model gives shape to such markets, facilitating the visualization of characteristics such as market efficiency, and changes that may favour either buyers or sellers. Conceptually, the star glyph maps market topology into iconography. Research is ongoing to quantify this approach by optimizing the dimensional configuration to maximize the resolution of the buyer–seller dichotomy. While the topology is completely independent of market economics, information on the latter can be used in either a knowledge-base or neural-net approach to identify shapes of interest. With a shape, we can now take snapshots as illustrated, and eventually produce movies of online auction markets as they develop. This can be a useful tool for experimental economists and architects of electronic markets (O’Toole 2001), for the simulation and empirical study of innovative designs, as well as the monitoring of emerging markets in the future.
ACKNOWLEDGMENT

Santosh Bukitgar designed the MS Excel macros for the data mining used in this work, with a contribution from Litian Zhang. Yan Lu served as research assistant and coordinated several MBA and EMBA classes at UIC in Fall 2002, Spring 2003, and Fall 2003, which contributed to the data collection and analysis in their term projects.

References


