Analysis and Visualization of Metrics for Online Merchandising

Juhnyoung Lee, Robert Hoch, Mark Podlaseck, Edith Schonberg, Stephen Gomory

IBM T.J. Watson Research Center
P.O. Box 704
Yorktown Heights, NY 10598
{jyl, rhoch, podlasec, ediths, gomory}@us.ibm.com

Abstract. While techniques and tools for Web marketing are being actively developed, there is much less available for Web merchandising. This paper contributes to the area of Web usage analysis for E-commerce merchandising. First, we categorize areas of analysis for Web merchandising such as product assortment, merchandising cues, shopping metaphors, and Web design features. Second, we define a new set of metrics for Web merchandising, which we call *micro-conversion rates*. The new metrics provide capabilities for examining data about merchandising in online stores, and detailed insight into the effectiveness of different Web merchandising efforts by answering related business questions. Third, we present a set of novel visualizations that explore patterns in micro-conversions in online stores reflecting in customer responses to various Web merchandising efforts. Through an empirical study using look-to-buy data from an online store, we demonstrate how the proposed starfield visualizations can be used to understand the shopping behavior in an online store and the effectiveness of various merchandising schemes it employs. Finally, we discuss the types of data required for this kind of visual analysis of online merchandising, and briefly describe how the data can be collected and integrated from an E-commerce site.

1 Introduction

In just a few years, the Internet has evolved into a significant commerce vehicle, that is, a channel for sales and customer service of virtually every type of business. In order to maximize their return on investment, Web merchants are finding it necessary to thoroughly understand the effectiveness of their sites and to take appropriate action when and where the sites fall short. Web merchants generally analyze their sites' effectiveness from two perspectives: marketing and merchandising.

Marketing on the Web is broadly defined as the activities used to acquire customers to an online store and retain them. Techniques for online marketing include the use of banner ads and email campaigns. Examples of marketing business questions include the followings: Which banner ads generate the most traffic and sales? Which portal sites are pulling in the most qualified traffic? Who are the buyers referred by a particular ad? Web usage metrics for answering these questions are the banner ad *click-through rate*, which is the percentage of viewers who click on a banner ad, and the *conversion rate*, which is the percentage of visitors who purchase from the store. Recently, ad banner *return on investment* (ROI) has become the significant metric for Web marketing. Marketers want to know not only the number of visitors who come to a site from a particular banner ad and purchase from the site, but also how much revenue and profit is generated by these visitors.

Merchandising consists of the activities involved in acquiring particular products and making them available at the places, times, and prices and in the quantity to enable a retailer to reach its goals [3]. The activities also include how and where to display products, and which products to advertise and promote. Online merchandisers are responsible for product assortment and product display, including promotions, cross-selling and up-selling. In large online stores, merchandisers make adjustments to the Web site content continuously, i.e., weekly or even daily. To assist online merchandisers, there are tools available for content management [8], which facilitate modifying the Web site, and also data mining techniques for association rule generation to determine which products are suitable for cross-selling [1]. However, in the area of Web usage analysis, while the needs of reporting and analysis for marketing is being addressed, useful merchandising metrics and tools lag behind. Web page hit counts provide a broad indication of visitor interest. However, keeping track of which products are shown on each Web page can be tedious or impossible, in particular, when page content is highly dynamic and personalized. Furthermore, there is a need to know to what extent interest translates into sales.

This paper focuses on the analysis of merchandising effectiveness in online stores. We define a new set of metrics, which we call *micro-conversion rates*. We show how these metrics provide detailed insight into the success of different Web merchandising, product assortment, and site design strategies. Also, we present novel visualizations of micro-conversion, inspired by the starfield model [2], which can be used to rapidly answer related merchandising questions. The concept of micro-conversion rates is based on banner ad marketing metrics. More specifically, we view an E-commerce site as a collection of advertisements for individual products in the store. From this perspective, we measure click-through rates, conversion rates, and ROI for a broad range of internal Web site features, including cross-sell, up-sell, and promotion. These metrics provide information at the individual product level, as well as on the product attribute and aggregate levels. Our belief is that this kind of information is both actionable and necessary for merchandising success.

The rest of this paper is structured as follows: Section 2 presents a high level set of merchandising usage analysis areas and related business questions. Section 3 defines a new set of metrics for merchandising, referred to as micro-conversion rates. Techniques for visualizing micro-conversion are illustrated and described in Section 4. In Section 5, an empirical study on the starfield visualization technique is described with look-to-buy data from an online store. In Section 6, we discuss the types of data required for this kind of visual analysis of online merchandising, and briefly describe how the data can be collected and integrated from an E-commerce site. In Section 7, related work is evaluated and summarized. Finally, in Section 8, conclusions are drawn and further work is outlined.

2 Web Usage Analysis for Merchandising

Before defining micro-conversion rates, this section first categorizes and describes various aspects of Web merchandising. We identify four areas of usage analysis: product assortment, merchandising cues, shopping metaphor, and Web design features.

The first analysis area, *product assortment*, deals with whether the products in an online store appeal to the visitors. If the product assortment is not optimal, the merchants may adjust, for example, brands, quality, selection, inventory or price of the products they carry. Examples of business questions related to product assortment include the following: What are the top sellers for a specific period of time, e.g., this week? What is the conversion rate for a particular department? In what frequencies and quantities are products purchased? What characterizes the products that end up being abandoned? Analyzing the effectiveness of individual products by finding answers to these questions provides some insight into product assortment. However, there may be some other merchandising problems that are resulting in poor sales. Complete understanding requires looking at other metrics and the remaining three areas of analysis, which deal with site design effectiveness.

Merchandising cues are different ways Web merchants present and/or group their products to motivate purchase in online stores. Examples of merchandising cues are cross-sells, up-sells, recommendations, and promotions. A *cross-sell* in an online store is a hyperlink which refers the visitor to a Web page marketing an item complementary in function to the item marketed on the current Web page. An *up-sell* hyperlink refers the visitor to a Web page presenting a similar but more upscale item. A *recommendation* hyperlink highlights product pages that are likely to be of interest to the shopper based on knowledge of the shopper and the behavior of a larger population. A *promotion* hyperlink refers a visitor to a product page from a "What's Hot" page or a high traffic area such as the "Home" page for informing, persuading and/or reminding the shoppers about a product and/or other aspects of the site. Online merchants need to understand the effectiveness of the merchandising cues in their stores in terms of traffic and sales driven by them. Examples of business questions related to merchandising cues include the followings: How much cross-sells and up-sells contribute to gross revenue? What are the best performing cross-sell pairs? And worst? What is the overall conversion rate for cross-sells? How much do promotions contribute to gross revenue? Which promotions are generating the most sales? At which levels in the site hierarchy are the best promotions located?

Shopping metaphors in an online store are different ways that shoppers use to find products of interest. Examples include browsing through the product catalog hierarchy, various forms of searching, and configuration for "build-to-order" type products. The effectiveness of different shopping metaphors in the store is a concern for online merchants. Like merchandising cues in online stores, shopping metaphors are associated with hyperlinks on Web pages. This allows one to categorize and group together hyperlinks in an online store by their types of merchandising cue and shopping metaphor. Examples of business questions related to shopping metaphors in online stores include the followings: What generates the most sales value, e.g., search or browsing? How much does search contribute to gross revenue? What is the conversion rate for search?

Other merchandising aspect that can be used to categorize hyperlinks are their design features such as media type (e.g., image or text), font (if text), size, color, location. The effectiveness of Web deign features presents another area of analysis for merchandising. Examples of business questions related to Web design features include: What are the features of links customers most frequently click? What are the features of links customers most frequently buy from? What are the parts of Web pages customers most frequently buy from? Do products sell better in the upper left corner?

Just as Web marketing uses banner ads and/or referral sites to attract customers from external sites to an online store, Web merchandising uses hyperlinks and image links within the store to lead customers to click to Web pages selling products. Web merchants employ a variety of merchandising schemes associated with hyperlinks. From this perspective, the problem of tracking and measuring the effectiveness of different merchandising strategies in an online store can be partitioned into three sub-problems:

- classifying each hyperlink by its merchandising purposes,
- tracking and measuring traffic on the hyperlinks and analyzing their effectiveness (e.g., profitability), and
- attributing the profit of the hyperlinks to their merchandising cue type, shopping metaphor type, and design features.

The analysis of the effectiveness of marketing strategies is conducted in a similar way by using the metrics such as click-through rates and ad banner ROI described earlier. The only difference is that the originating hyperlinks in marketing efforts are presented and controlled in external sites.

3 Micro-Conversion Rate

Traditionally, the conversion rate of an online store is the percentage of visitors who purchase from the store. While this measure is useful for evaluating the overall effectiveness of the store, it does not help understand the possible factors affecting the sales performance. The notion of a micro-conversion rate extends the traditional measure by considering the four general shopping steps in an online store, which are:

- *product impression*: the view of hyperlink to a Web page presenting a product.
- *click-through*: a click on the hyperlink and view the Web page of the product.
- *basket placement*: the placement of the item in the shopping basket.
- *purchase*: the purchase of the item - completion of a transaction.

Basic micro-conversion rates are computed for each adjacent pair of measures resulting in the following three rates:

- *look-to-click rate*: how many product impressions are converted to click-throughs.
- *click-to-basket rate*: how many click-throughs are converted to basket placement.

- *basket-to-buy rate*: how many basket placements are converted to purchases.

Note that the first of these, look-to-click rate, is similar to the click-through rate used for measuring the amount of traffic on banner ads. Also note that the micro-conversion rates relate the traffic-related measure to sales which happen later in the shopping process. By precisely tracking the shopping steps in this manner, it is possible to spot exactly where the store loses the most customers.

In addition to these three rates, there is another useful metric that aggregates the four shopping steps. By looking at this rate, online merchants can tell if a product is overexposed or underexposed and take action to change the presentation of the product. We call this rate look-to-buy rate:

- *look-to-buy rate*: what percentage of product impressions are eventually converted to purchases.

The micro-conversion rate extends the traditional measure by considering the marketing or merchandising purposes associated with hyperlinks viewed in the first shopping step described above, i.e., product impression. In this way, the micro-conversion rate is related to strategies for marketing and merchandising, and can be used for evaluating the effectiveness of different merchandising aspects of the store.

Fig. 1 illustrates in a bar chart an example of micro-conversion rates for three different types of merchandising cues, i.e., promotions, recommendations, and cross-sells. Visitors are seeing twice as many impressions of promotions (40,000) than cross-sell impressions (20,000), and twice as many cross-sell impressions than those of recommendations (10,000). However, the look-to-click rate for recommendations (18%) is twice as high as for either promotions or cross-sells. Additionally, recommendations are resulting in a relatively high look-to-buy rate (2%). This means that the recommendation engine is relatively effective at personalization. On the other hand, this example shows that promotions in this store are not effective. Of the visitors who place a promoted item in a shopping basket, 30% of them purchase the product. However, the click-through rate for promotions is 10% and the click-to-basket rate is only 2.5%, so the look-to-buy rate is 0.075%, which shows poor overall performance, and an over-exposure of the promoted items. Finally, the look-to-buy rate for cross-sells in this example is about 0.5%.

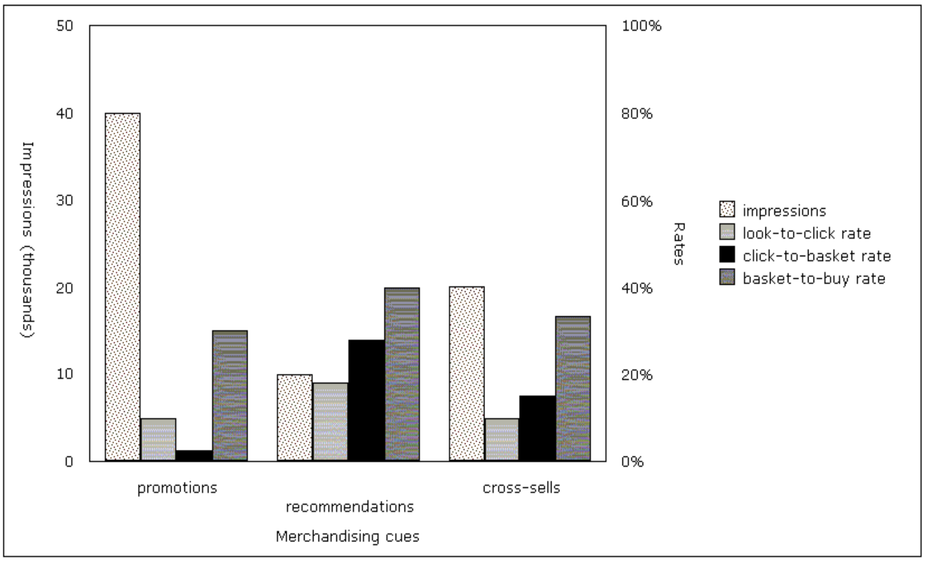

Fig. 1. Example micro-conversion rates

As illustrated in this example, micro-conversion rates can be calculated for individual merchandising cue types, and also, for individual products, individual shopping metaphor types, individual design features, and individual banner ads. As a result, all the individual hyperlinks pointing to product pages in various forms and purposes, internal or external to the site can be analyzed. Unlike the traditional conversion rate, which gives just one number for the entire site, these micro-conversion rates provide insight into a rich set of information pertaining to the analysis areas described in section 2. For instance, they can be computed for individual products to measure the product performance in the site. The rates for individual products can then be rolled up to give the rates for categories of the products, and then again all the way up to the entire site. Table 1 presents sample business questions on merchandising that can be addressed by micro-conversion rates for each analysis area.

Product Assortment	Look-to-click	Is a product's exposure optimized for the current level of customer interest?
	Click-to-basket	Is the detailed information about a product appropriate?
	Basket-to-buy	What kinds of products are abandoned in shopping cart?
Merchandising Cues	Look-to-click	What cross-sells are working best? Worst?
	Click-to-basket	Is a cross-sell more likely to be placed in a shopping basket if the first item has already been placed there?
	Basket-to-buy	Are the customers who responded to a cross-sell any more or less likely to abandon a product in a shopping cart?
Shopping Metaphors	Look-to-click	Are customers finding what they want from the search engine?
	Click-to-basket	Do customers who found a product through the search engine want the same amount of product detail as those that found it by browsing?
	Basket-to-buy	Are the customers who responded to a search result any more or less likely to abandon a product in a shopping cart?
Design Features	Look-to-click	Are visitors clicking more image links than text links?
	Click-to-basket	Are there product links that are misleading?
	Basket-to-buy	Where are the problems in the check-out process?

Table 1. Micro-conversion rates for sample business questions

4 Visualizations of Micro-Conversion Rates

Visualizations of micro-conversion can help provide insight into the kinds of merchandising questions raised in Section 3. The product visualization, a scatterplot graph, shown in Fig. 2 augments the interactive starfield model [2], a general-purpose analysis tool useful for finding patterns in multidimensional data. With the associated tree controls given in Fig. 3, users can filter on hierarchical dimensions found in the E-commerce domain such as product taxonomy (pictured here) and site architecture. Selection of one or more branches of the tree causes the products under that branch to be pictured in the graph area. The color key associated with a particular branch in the tree can be inherited from a parent (the default) or overridden with a color unique to that child.

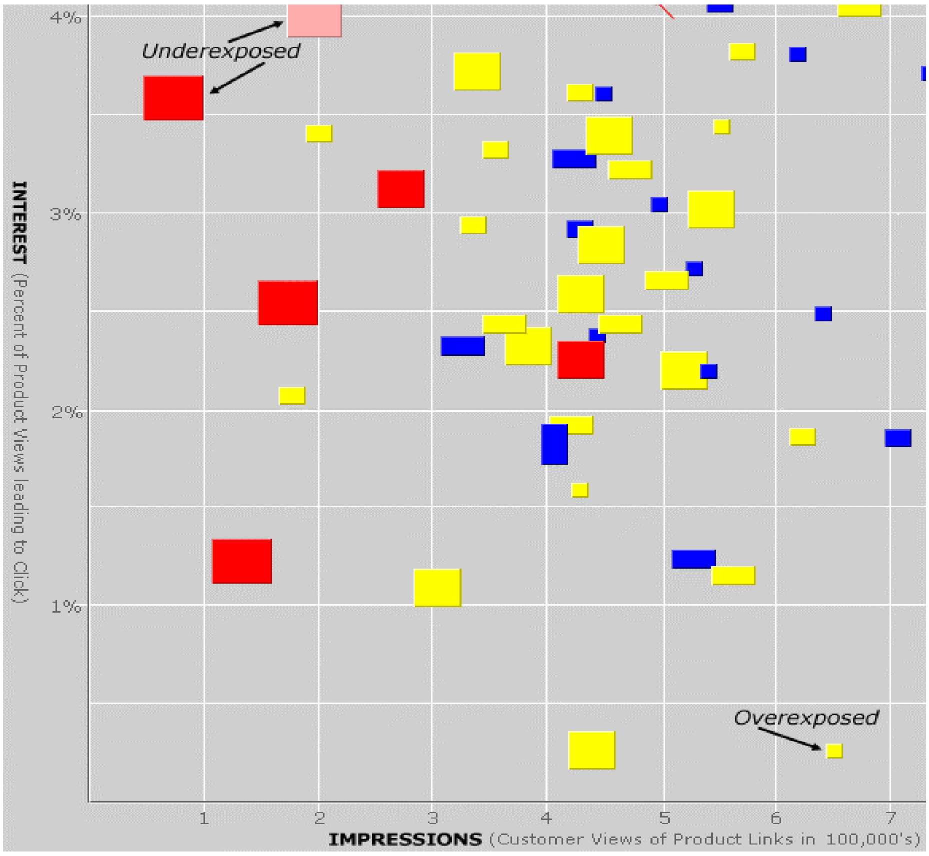

Fig. 2. Product visualization based on the starfield model

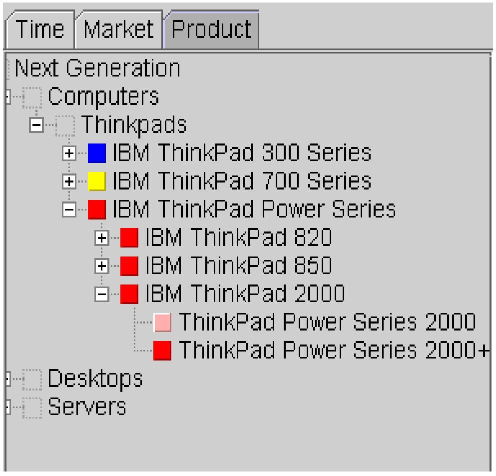

Fig. 3. Tree control

Each rectangle or *glyph* in the graph space represents a distinct product (a certain brand and type of T-shirt, for example, but not its colors or sizes). The color of each glyph in this example corresponds to the product's department, as indicated by the color key on the tree control described above. The area occupied by a glyph describes a product's relative significance: width is an indicator of the product's price, and height, its relative (profit) margin.

The x-axis and the y-axis of the starfield graph can represent any two of the micro-conversion rate metrics. In the example in Fig. 2, the user configured the axes to analyze product exposure relative to customer interest ("Are the right products being promoted? How do I optimize the exposure of all my products to maximize my revenue stream?"). The x-axis thus represents raw impressions, that is, the number of times a hyperlink to a product was served. (Product hyperlinks can occur anywhere on the commerce site: the home page, category pages, search result pages, as well as other product pages.) The y-axis represents the percentage of impressions that resulted in a click-through (that is, of the number of customers that saw a hyperlink to this product, the percentage that clicked on the link).

The scatterplot graph makes evident the heavy over-promotion of a product represented by the small glyph in the lower right quadrant. While it has had more impressions than almost any other product, its click-through is almost the lowest. To make matters worse, it is a low-priced, low-margin product. Its exposure could be reduced by moving its promotion to a less-trafficked page, or eliminating it entirely.

On the other hand, the large glyphs in the upper left quadrant represent products that are under-exposed. Although links to these products have few impressions, a relatively high percentage of customers are clicking on them. This level of interest might be maintained if the number of impressions of these product links were incremented. If that is not the case, it is possible that the products are niche products, appealing only to the small group of customers that are specifically looking for them (left-handed joy-sticks, for example). Therefore, depending on the nature of the products in the upper-left corner, one might chose to display or to promote them more, and then monitor the results carefully.

Reconfiguring the graph space allows one to explore other questions. For example, one might reassign the x-axis to represent click-throughs and the y-axis to represent the percentage of click-throughs that resulted in a product being placed in the shopping basket. Products with a high click-through rate, but low basket-placement rate would occupy the upper left quadrant of the graph. These are products in which customers were interested enough to click on, but not interested enough to consider buying. Causes to explore here include the quality of the information on the product detail page, surprise pricing, or misleading product links.

5 A Case Study: Visualizing Look-to-Buy Rates of an Online Store

Note that the starfield visualization in Fig. 2 used a set of mock-up data to demonstrate diverse look-to-click characteristics of individual products. In order to understand its applicability and usefulness, we have performed an empirical study on the proposed visualization technique with a number of micro-conversion data sets available from several operating online stores. We have tested the data for different combinations of micro-conversion steps. To obtain the visualizations, we used simplified but still useful versions of scatterplot that are available from several office software suites.

Fig. 4 provides a visualization example from the study, a product visualization based on a data set of 531,873 hits (minus hits on graphic files) and 7,584 basket placements acquired from an online computer store over a period of one week in 1999. Fig. 4 shows the products from three departments of the online store: Consumer, Business, and Accessory. The number of products in the Consumer department is 55, that of the Business department 30, and that of the Accessory department 157. Data for several other departments in the online store is not shown in this figure for the sake of clarity.

In this figure, bubbles instead of rectangles are used to represent individual products. As in Fig. 2, the size of each bubble describes a product’s relative significance. In this case, the diameter of each bubble indicates the price of the product represented. (Data about profit margin of products were not available for this study.) The color of each bubble represents department to which this product belongs The x-axis of the graph represents raw impressions, i.e., the number of times a Web page that containing a hyperlink to a product is served. The y-axis represents the percentage of impressions that resulted in a basket placement, i.e., of the number of visits that saw a hyperlink to this product, the percentage that eventually added this product to the shopping cart. Note that the percentage of impressions leading to basket placements (i.e., look-to-basket rate), not the raw orders, is used to normalize the data over a fixed range, that is, between 0 and 100.

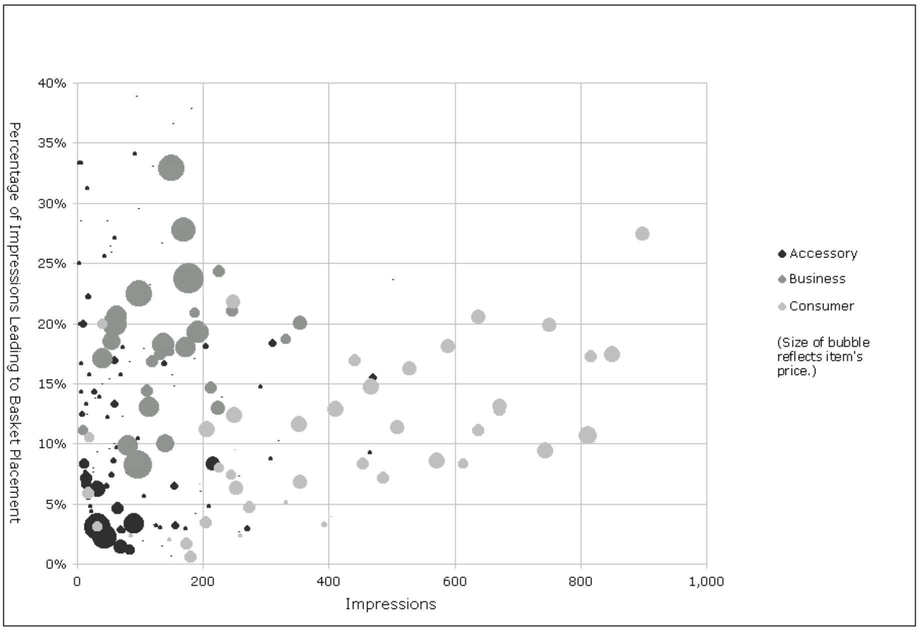

Fig. 4. Scatterplot of look-to-basket rates of an online store

Fig. 4 illustrates that the look-to-basket rates of products from three different departments show vastly different characteristics. There are several possible causes. Either the store displays the products differently, resulting in products from some departments shown to more visitors than others, or the shopping behavior of visitors is different for the various departments of products. Identifying the shopping behavior of customers and what is important for customers of different departments, Web merchandisers can upgrade the Web design and merchandising schemes to adjust for different characteristics of individual departments.

First, we observed that most of the Consumer products receive high number of impressions, i.e., between 180 and 900, compared with products of other departments. It may be the case that the store promotes Consumer products more in high-trafficked pages to lure the customers to the Consumer department. It is also speculated that the customers of the Consumer department browse more and review more information about products than those of other departments. Web merchandisers of the Consumer department can use this information to improve their Web design and sales. For example, they may want to improve product information in the department because it is what the customers of the department look for. Also, unlike the Business products (as will be explained later), well-presented impressions are particularly important for the Consumer products.

It is important to note that the look-to-basket rates of the Consumer products generally increase as their impressions increase, i.e., overall average about 3% increase in look-to-basket rate for each impression. Also important is that, despite this general tendency, the Consumer products show a wide range of individual difference in their impressions and look-to-basket rates, that is, the bubbles are widely scattered across both x- and y-axis. For example, two Consumer products which receive roughly the same number of impressions (615), show significant difference in their look-to-basket rates: one has about 21%, and the other about 7%, i.e., about 300% difference. Based on this observation, Web merchandisers can make decisions on which products are better candidates for promotions and/or recommendations.

Unlike the Consumer products, only few from the Business department receive more than 200 impressions. In general, the Business products have relatively low impressions, but high look-to-basket rates. It is probable that the store promotes the Business products less than the Consumer products. Also, it is speculated that the customers of the Business department already know what they want to buy when they come to the store. Hence, they do not tend to browse the store for product information. Instead, they use tools such as search engine to quickly find products that they want to buy. This unique shopping behavior of the Business customers may cause the relatively high look-to-basket rates of the Business products, despite their relatively high price (indicated by the size of the bubbles). Again, Web merchandisers of this department can use this information to improve their sales. For instance, Web merchandisers may want to improve their shopping mechanisms (other than browsing) to help the Business purchasers rapidly find products they want to buy under various conditions. They may also be able to increase sales by marketing to the more potential customers in the segment.

There are more products in the Accessory department than in the other two departments, but most of them are low-priced as their size indicates. Most receive relatively small number of impressions, probably because the store does not heavily promote the Accessory products. It is interesting to note that several high-priced Accessory products show both low impressions and low look-to-basket rates, while most low-priced Accessory products show relatively high look-to-basket rate. This indicates that when shopping for accessory items, customers become more sensitive to price.

In this case study, we have suggested various possible interpretations of the data in this visualization. Ultimately, the merchandisers who know their own products are best qualified to assign meaning and take action. The role of visualization tools is to help the merchandisers understand data by displaying complex data in a clear way.

6 Data Requirements

In this section, we briefly describe several data requirements for the proposed starfield visualizations for E-commerce. While some source data are readily available from most E-commerce sites, others are not and need to be collected with some special tools. Also, the collected data has to be integrated to compute the micro-conversion rates before visualizing them to provide insight into the merchandising effectiveness of online stores.

First, the visualization of merchandising effectiveness based on micro-conversion rates requires the combination of the site traffic data (i.e., access requests) and sales data. In most E-commerce sites, the two types of data are typically stored in separate storage systems in different structures: the traffic data in Web server logs in a file format, and the sales data in the database of commerce server. The commerce server database also contains information about customers and products (including product taxonomy) that may be useful and interesting to visualize with micro-conversion rates. It is important to tie together data from the two different sources with a common key and to construct an integrated database system or a data mart system for business visualization.

Second, computing micro-conversion rates also requires product impression data. Capturing product impression data involves tracking content of served Web pages, which is more challenging when Web pages are dynamic. Currently, the standard Web server logging mechanism does not capture the content of Web pages. One possible method is to enhance the Web server logging as a way to dynamically parse the content of served Web pages and extract useful data such as product impressions and information on hyperlink types. The ability of dynamically scanning Web pages as they are served is critical for tracking Web usage, because more and more Web pages are dynamically created from databases and contain personalized and adaptive content. A simple example of a dynamically created Web page is a search result page commonly found in online stores.

Finally, it is important to classify and identify hyperlinks by their merchandising purposes, so that later to attribute the profit of the hyperlinks to their merchandising cue type, shopping metaphor type, and design features. For this purpose, Web pages and hyperlinks in an online store need to be tagged with semantic labels describing their merchandising features. Semantic labels of a hyperlink may include, for example, a product label, a cross-sell or promotion label, and a tag indicating where the product is being displayed. Such semantic labels for hyperlinks in a site may be explicitly provided in a form of meta-data during the site creation. If this is not the case, semantic labels need to be inferred from various sources such as the file name and/or path portion of URLs, types, values and orders of parameters in URLs of dynamic Web pages, and the location of a hyperlink in the page.

E-Commerce Intelligence (ECI) is an ongoing project at IBM T. J. Watson Research Center. The architectural goal of the ECI project is to provide an analysis environment that is rich in data expressed in terms that are comfortable for the business analysts to answer the business questions. The visualization of micro-conversions with the starfield model is a reporting and analysis metaphor provided by the ECI system. The flow of data for the starfield visualization in the system is illustrated in Fig. 5.

7 Related Work

There are several commercial services and software tools for evaluating the effectiveness of Web advertising such as banner ads, in terms of traffic and sales driven by them [7]. They use metrics such as click-through rates and ad banner ROI. The objective of these tools is different from our work; they focus on understanding the effectiveness of advertising, while our work focuses on the effectiveness of merchandising. The techniques used for Web advertising tracking tools are not directly applicable to tracking and measuring the merchandising effectiveness in online stores.

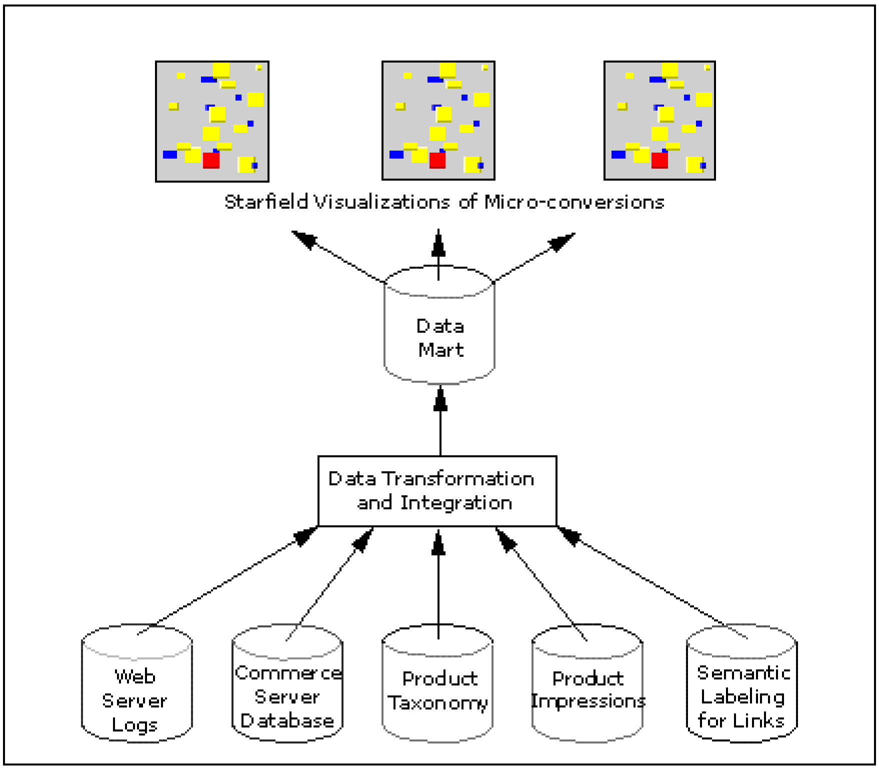

Fig. 5. Data requirements for visualizing micro-conversions

There are several commercial Web server log analysis tools available [12]. Most of these Web log analysis tools generate reports mainly on Web traffic and system measurements. It is difficult, though not impossible, to report on the effectiveness of specific marketing and merchandising efforts because those tools primarily rely on information in Web server logs which are hard to interpret and extract data useful for measuring the business efforts.

Recently, Web usage analysis by applying data mining techniques to Web server logs has been actively studied [5, 6]. Some of the data mining algorithms that are commonly used in the studies are association rule generation, sequential pattern generation, and clustering. Also, there has been some work done on loading Web usage data into a data cube structure in order to perform both data mining and traditional OLAP (On-Line Analytical Processing) operations such as roll-up and drill-down of the data [4, 13]. These studies formally defined generic data hypercubes for Web usage data and provided designs for exploratory analysis and reporting. Some of the work showed how data mining techniques are used on the data model in electronic commerce scenarios. While these studies and our work share similar objectives of finding interesting aspects of Web usage of an online store which are potentially useful for improving marketing and merchandising strategies on the site, they address different types of business questions and may be used in a complementary way.

Interactive starfield visualizations were studied to explore Web server log data in [9]. The study argued that these visualizations provide capabilities for examining data that exceed those of traditional Web log analysis tools, by combining two-dimensional displays of access requests, color and size coding for additional attributes, and facilities for zooming and filtering. While the study introduced a series of interactive starfield visualizations to explore Web traffic data across various dimensions, it used only Web server log data as its data source. In comparison, the work presented in this paper envisions how interactive starfield visualizations can be used to explore commerce server data and how they can be interpreted to provide insight into merchandising effectiveness of online stores.

8 Concluding Remarks

This paper has presented a set of Web usage analysis requirements in the area of online merchandising. It also presented the details of tracking and analyzing the effectiveness of merchandising in online stores by using the novel concept of micro-conversion rate. Then we presented techniques for visualizing the micro-conversions, which can be used to rapidly answer merchandising questions. Through an empirical study, we demonstrated how the proposed starfield visualizations can be used to understand the shopping behavior in an online store and the effectiveness of various merchandising schemes it employs. We also discussed the types of data required for this kind of visual analysis of online merchandising, and briefly described how the data can be collected and integrated from an E-commerce site.

The ideas presented in this paper are currently being implemented in a business intelligence system that manages and integrates diverse data from an E-commerce site, and provides capabilities for examining and exploring the data. As part of this effort, we have designed a data model, and developed a number of software tools for collecting and managing E-commerce data. Those components are required to prepare data that will be shown in the visualizations presented in this paper.

We extend our work on visualization of E-commerce data with a number of other data visualization ideas including parallel coordinates which map multi-dimensional data into a two-dimensional display. Unlike the starfield model, parallel coordinates places axes parallel rather than perpendicular to each other, allowing many axes to be placed and seen [10]. This mapping procedure has unique geometric properties and useful relationships to the original space.

Our work for analyzing online store merchandising presented in this paper can be extended in many ways to address different questions on both business and system effectiveness of online stores. First, the idea of classifying hyperlinks, labeling them with semantic vocabularies, and aggregating them, can be generalized and applied to other types of business questions in areas such as online marketing and operation. One example is clustering customers by their shopping behavior measured by types of hyperlinks they click. In this case, hyperlinks in an online store need to be categorized and labeled to distinguish characteristics of shoppers' behavior. Also, the metrics for merchandising effectiveness presented in this work can be adjusted and extended for the use in new shopping paradigms in the Internet such as online auction and dynamic pricing. Finally, the approach of data visualization-based exploratory analysis used in this work can be combined with a different but complementary approach, that is, data mining. Data mining can help sample data clusters interesting for visualization, and/or find patterns in navigation paths and identify product associations.

References

1. S. Anand, J. Hughes, D. Bell, and A. Patrick, "Tackling the Cross-Sales Problem Using Data Mining," *Proc. Of the 1st Pacific-Asia Conference on Knowledge Discovery and Data Mining*, 1997, 331-343.
2. C. Alhberg and B. Schneiderman, "Visual Information Seeking: Tight Coupling of Dynamic Query Filters with Starfield Displays," *ACM CHI Conference on Human Factors in Computing Systems*, 1994, pp. 313-317.
3. B. Berman and J. R. Evans, *Retail Management: A Strategic Approach*, 7th Edition, Prentice-Hall, Inc., 1998.
4. A. G. Büchner and M. Mulvenna, "Discovering Internet Marketing Intelligence through Online Analytical Web Usage Mining," *SIGMOD Record*, 27(4):54-61, December 1998.
5. M. S. Chen, J. S. Park, P. S. Yu, "Data Mining for Traversal Patterns in a Web Environment," *Proc. of the 16th International Conference on Distributed Computing Systems*, 1996.
6. R. Cooley, B. Mobasher, and J. Srivastava, "Data Preparation for Mining World Wide Web Browsing Patterns," *Journal of Knowledge and Information Systems*, 1(1), 1999.
7. S. V. Haar, "ROI Shows Marketers The Money," *InteractiveWeek*, March 22, 1999, http://www.interactive-week.com.
8. M. E. Hardie, J. C. McCarthy, M. MacKenzie, and R. Scheier-Dolberg, "Content Management Crisis," *The Forrester Report*, Volume One, Number Eleven, January 1997.
9. H. Hochheiser, and B. Schneiderman, "Understanding Patterns of User Visits to Web Sites: Interactive Starfield Visualizations of WWW Log Data," *Technical Report*, CS-TR-3989, Department of Computer Science, University of Maryland, 1999.
10. A. Inselberg, "The Plane with Parallel Coordinates," *The Visual Computer* 1, 1985, 69-91.
11. M. Monticino, "Web Analysis: Stripping Away the Hype," *IEEE Computer* 13(12), December 1998, 130-132.
12. T. Wilson, "Web Site Mining Gets Granular," *InternetWeek*, March 29, 1999, http://www.internetwk.com.
13. O. R. Zaiane, M. Xin, J. Han, "Discovering Web Access Patterns and Trends by Applying OLAP and Data Mining Technology on Web Logs," *Proc. of Advances in Digital Libraries Conference*, 1998.