Quantum Intelligence, Inc.

3375 Scott Blvd, Suite 424
Santa Clara, CA, 95054

Phone 408-252-7746
Fax 760-875-8766

White Paper

Quantum Intelligence System for Retail

By

Ying Zhao, Ph.D. and Charles C. Zhou, Sc.D.

ying_zhao@quantumintelligence.com

charles@quantumintelligence.com

August, 2002

Introduction

Quantum Intelligence System (QIS) is based on the concept of integrating quantum intelligence, local and small knowledge of a complex process, into a global intelligence. QIS integrates a collection of state-of-the-art information mining and optimization technologies into a single framework. The process of QIS is shown in Figure 1. QIS first employs the best practices in database and data warehousing area to integrate the important information from diversified databases, then employs innovative information mining techniques to search for patterns, trends, dependency and anomalies, and report the discovery in front of decision makers. Decision makers can then act on the integrated power of such knowledge.

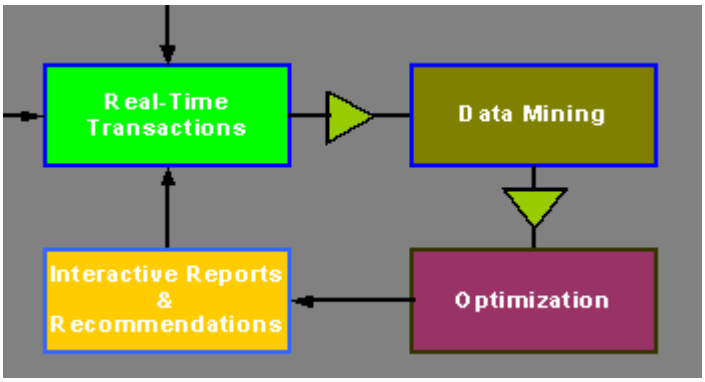

Figure 1: QIS sifts through transaction data directly, integrates data mining and optimizations and delivers interactive reports, visualizations and recommendations for a business process.

QIS for Retail is a retail application using Quantum Intelligence System to optimize the overall process of revenue management through targeted marketing, cross sell and effective promotion. Retailers always try to understand their customers. Tailoring and targeting their customers' need is a critical area for their success. Companies invest lots of money to understand who are their loyal customers, who are their heavy spenders and who likely buy what kind of products. Retailers understand the values of product affinities and they often use cross-sell and up-sell techniques to increase market basket size and revenue. Retailers also frequently use promotions to increase revenue. These promotions are often in the form of discount pricing of carefully selected products over a fixed period of time. Retailers have long realized that different products respond with significantly different sensitivities for their promotions, which is often referred as price elasticity.

All these factors affect the whole process of revenue management. How to leverage the power of all the aspects of revenue management is the area that QIS could provide great potential for retailers. QIS provides a decision support platform that is able automatically and quantitatively to consider all these factors from transaction data. QIS then uses the knowledge to optimize a sequence of future events and actions in real-time to maximize revenue and profit. By using QIS, retailers can act on integrated power of targeted marketing, cross-sell, up-sell and effective promotion, increase basket sizes and eventually increase the overall revenue/profit of their business.

Why Quantum Intelligence System?

Discovery-driven data mining, a process to automatically extract patterns, rules and knowledge from large databases, has drawn much attention recently for business intelligence for various industries. Statistical inferencing methods, artificial intelligence, machine learning and pattern recognition techniques have been applied in mining business data for solving business problems. Although many of these approaches can provide a multidimensional view of large databases, they usually focus on the isolated components of a process and often stop short at the discovery phase. Quantum Intelligence

System combines data mining and optimization techniques into a single framework and focuses on the actionable aspect of business intelligence. It has advantages over other data mining systems as follows

- Focus on time series of behavior, event or activities of revenue management.
- Optimize the process as a whole through targeted marketing, cross-sell to promotion effectiveness.
- Actionable: knowledge can be directly recommended for future promotions.
- Real-time monitoring and recommendation: update models and search/match new information in real-time

QIS Data Model

QIS starts with a careful integration of raw data from various perspectives of a business process. It follows the best practice in the area of data warehousing to integrate different data sources that can be represented in a data model. A data model is used to describe how critical and relevant data from different resources associated to each other. QIS uses raw transactions as the core data. Other related dimensions are joined with the transaction data through a *star schema* as shown in Figure 2. A star schema is a common design used by data warehousing professionals which consists of

- A single *fact* table containing the lowest granularity of detail relevant to the analysis
- One or more *dimension* tables containing related information that can be joined to the fact table through keys identifying important business concepts.

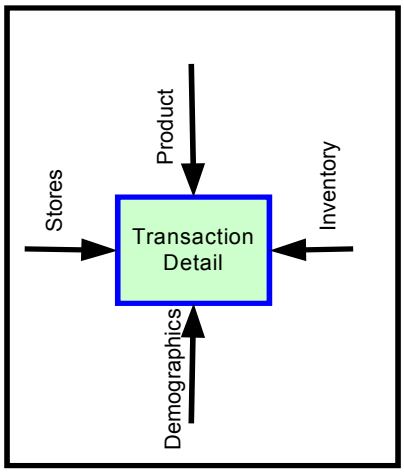

Figure 2 QIS data model

The QIS data model is a virtually centralized data repository that is potentially able to link all aspects of the revenue management efforts of a retailer. The database may contain information of transactions, products, promotions, inventories, customer demographics and store demographics that are linked as time series. QIS focuses on the time series of transactions, activities and events. QIS pre-processes the combined data as follows:

- Integrate all possible data, make logic connections and sequence them by time stamps
- Extract, transform and load (ETL) related data into a relational database for analysis
- Report initial statistics of all data elements for data quality checking
- Replace and predict missing data
- Transform data and generate derived variables

QIS Analytic Capabilities

QIS contains a collection of state-of-the-art information mining and optimization technologies into a single framework. It includes the following functionalities: “Predict, Score, Link and Act” as indicated in Figure 3

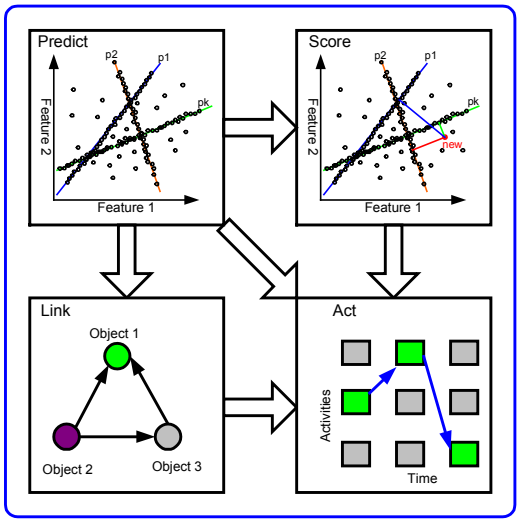

Figure 3: QIS includes "Predict, Score, Link and Act" modes that are integrated with multiple algorithms.

- **Predict** is to establish accurate and often non-linear correlation among data attributes. For example, to answer the questions, such as, who are the loyal customers? (correlation between loyalty and demographics); who are the responders to mail campaigns?(correlation between response and demographics); what kind of products are effectively promoted? (correlation between promotion effectiveness and product attributes).
- **Score** is to classify, categorize and correlate new information to the previously analyzed groups to see if how the new information is related to the historic knowledge. It can be used, for example, to pull out the people who are likely to response to a new marketing campaign; to forecast, the promotion effectiveness for a new promotion or new product.
- **Link** is to discover associations and relationships among objects and entities that can be, for example, product affinities and promotion effectiveness. It can be used to address the problem such as cannibalization when a promoted product cross-sells or cannibalizes the sales of substitutes.
- **Act** is to automatically perform a large set of "what if" analyses based on quantum knowledge discovered in the data and then to provide simulations of proactive planning and course of actions for future. The underlying technique based on structured dynamic programming.

QIS Architecture

QIS is a Java-based and web-enabled as illustrated in Figure 4. It integrates data sources through Java Database Connectivity (JDBC). The user interfaces are designed to be ease-of-use for decision makers.

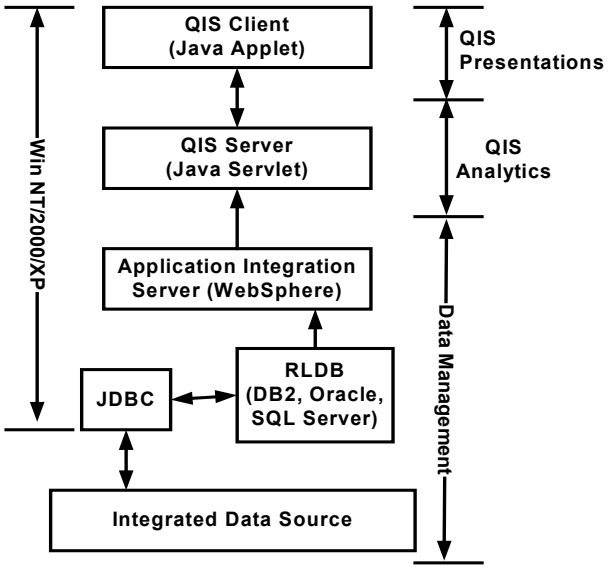

Figure 4: QIS is a web-enabled system connected to integrated data sources through JDBC (Java Database Connection) and built on top of an application server such as WebSphere.

QIS Examples

Targeted Marketing

The retailer in this example wanted to run a direct mail campaign to stimulate holiday sales and increase revenues. Traditionally, the retailer had employed Recency, Frequency and Monetary (RFM) exclusively to generate name lists for direct mail campaigns. RFM is a standard tool in the direct mail industry that leverages purchase history to identify profitable customers for mail campaigns. Although RFM is a proven and well-accepted method, it is not without drawbacks. First, RFM does not leverage demographic information about customers. It is therefore limited to activating existing customers, and it ignores potentially useful information. Second, RFM is pre-disposed to select those customers who have purchased recently and to ignore dormant customers who are good candidates for re-activation. We found that QIS predictive models identified high-spending customers who were neglected by traditional RFM analysis. Therefore, combining RFM and QIS predictive models generate more powerful campaign lists as shown in Figure 5. This map, created with Quantum Intelligence System illustrates that Quantum Intelligence System's predictive model identified high-spending customers who were neglected by traditional RFM analysis. The scales represent the demi-deciles of rank (i.e. bins of 5%) with each model. Quantum Intelligence System's model is on the x-axis and RFM is on the y-axis. Color indicates the actual campaign result in terms of average revenue per mail. Note that the Quantum Intelligence System model identified many high-spending customers (orange and red) in the first column, despite the fact that RFM ranked them poorly.

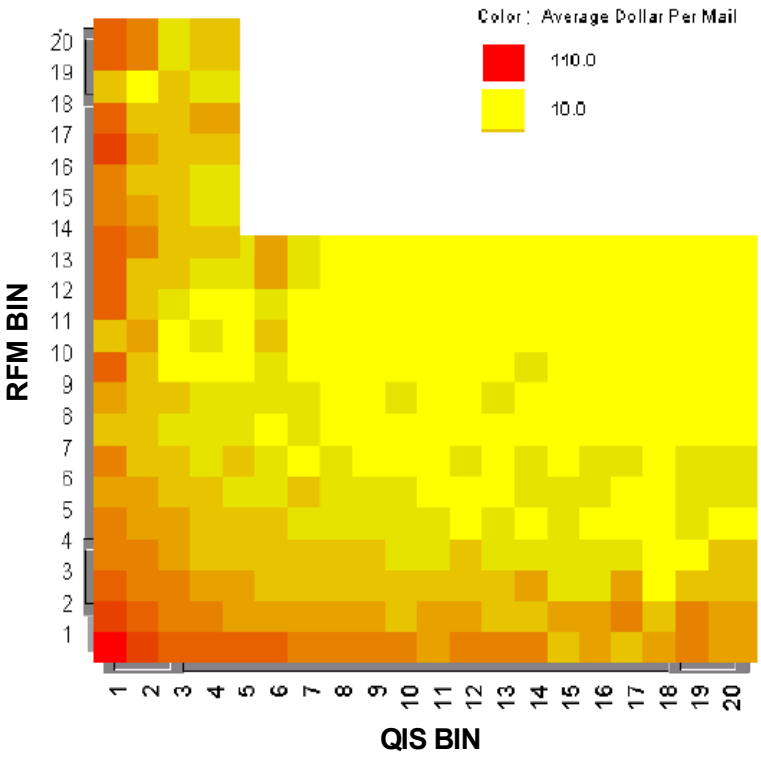

Figure 5: Targeted marketing using QIS

Cross-Sell

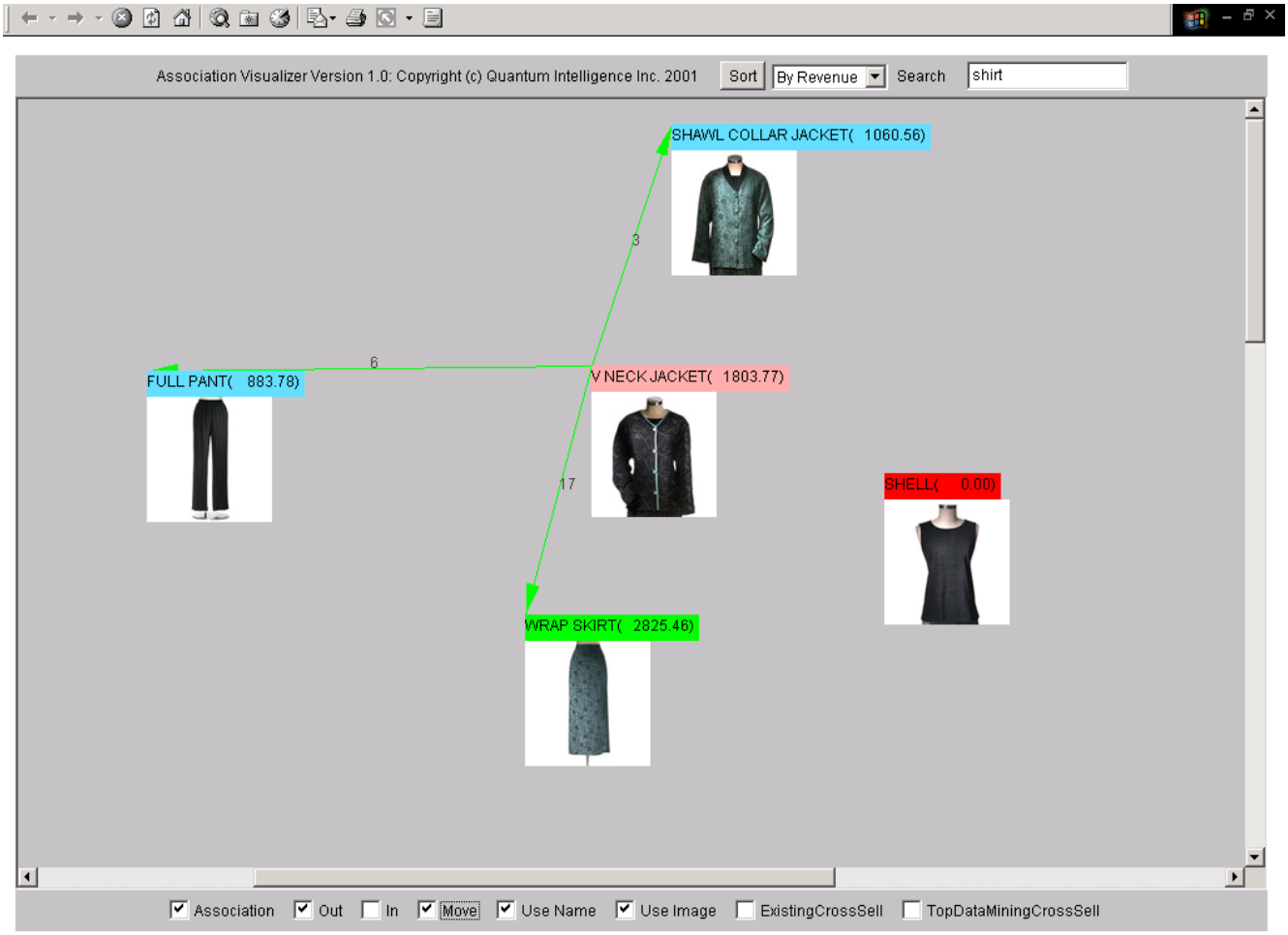

Figure 6: Cross-Sell Using QIS

E-commerce presents better opportunities for cross-sell and up-sell. In a brick-and-mortar store, re-arranging shelves to put associated products for cross-sell is an expensive operation while web sites provide an ideal platform for refreshing such recommendations periodically without physically re-arranging the items in the store. A web retailer used QIS to look for new trends of outfits that go together that were embedded in the data, many of which retailer experts might not have thought of. Further more, the retailer used QIS to evaluate the effectiveness of their existing cross-sells by seeing if there was enough evidence that people actually took the recommendation and bought the recommendations as shown Figure 6. Products highlighted in red (SHELL) are displayed as cross-sells for the V NECK JACKET (in pink) on the website, however the two were never bought together. Other products FULL PANT and SHAWL COLLAR JACKET highlighted in blue were purchased with the V NECK JACK (with in 6 and 3 orders respectively), but not recommended as cross-sells in the website previously. They were only discovered from the data. The WRAP SKIRT highlighted in green was displayed as a cross-sell and indeed 17 WRAP SKIRTS were purchased together with the V

NECK JACKET. The cross-sell recommendation of the WRAP SKIRT worked better than the SHELL.

Promotion Effectiveness

We have also worked with a large US grocery store and used QIS for promotion effectiveness analysis for their historic promotions. The promotion effectiveness is characterized by many aspects, for example, how effective does each individual item response to various price discount (price elasticity)? If an item is promoted, does it eat away the sales of what substitutes (cannibalization)?
Figure 7 shows an example of the QIS interactive visualization of product and promotion affinities. It shows what other non-promoted items are in the same baskets with a promoted item. There are many some interesting observations

- For example, "PEPSI SODA" with 20% off ("|0.2" stands for 20% discount in Figure 7) increased it own revenue by $1842.31. Meanwhile, it also increased the revenue of the non-promoted item "7-UP SODA" by $208.05, "A&W SODA ROOT BEER" by $62.92, "SPRITE SODA" by $133.33, "COCA COLA CLASSIC SODA" by $242.16, so on and so forth. The lift that is shown in the link has the following meaning: compared with all the baskets as a whole, the baskets that contain "PEPSI SODA" with a promotion (20% off) are 5.37 times more likely to have the non-promoted "7-UP SODA" and 3.95 times more likely to have the non-promoted "A&W SODA ROOT BEER", etc. Therefore, for this "PEPSI SODA" promotion, it did not actually cannibalize the sales of some of the substitutes, but rather cross-sold them. The products in the group with a blue color represent the products whose sales were positively affected by the promotion.
- There are products such as "TORTILLA CHIPS", "LAYS CHIPS", and "FRITOS CORN CHIPS" which are not the substitutes but are associated products in the positively affected group (blue one). This is consistent with a common observation that soda and chips that are often sold together.
- The products in the red group represent the products whose sales were negatively affected by the promotion. In other words, products in the red group were cannibalized by the promotion. The cannibalized products include substitutes such as "MOUNTAIN DEW SODA","PEPSI DIET SODA","DR PEPPER SODA", etc.

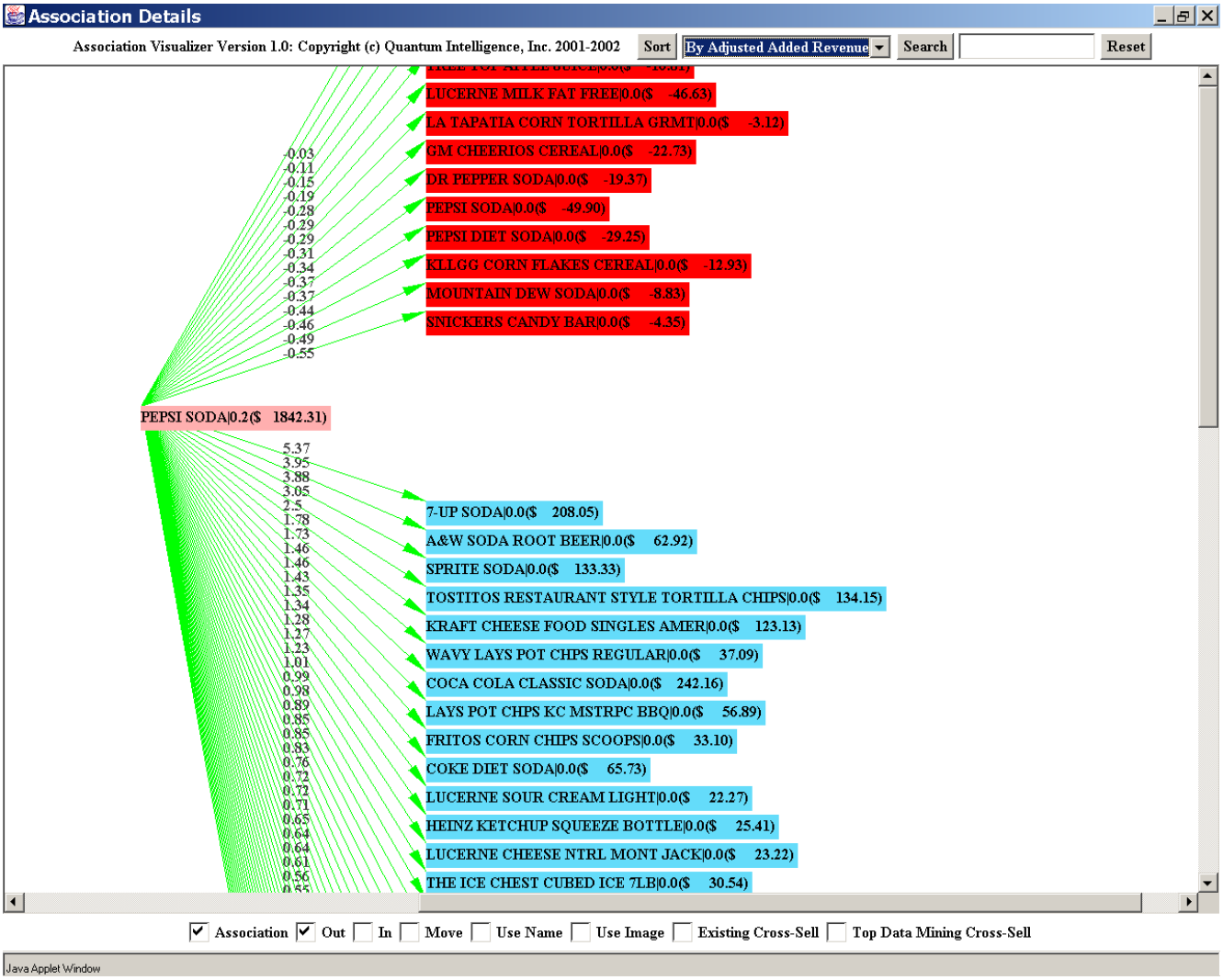

Figure 7: Visualization of QIS for promotion effectiveness

Acknowledgements

Authors want to thank Martin Raffauf from IBM Solution Partnership Center at San Mateo, California. The center provided extensive support and testing facilities for porting QIS to DB2, iSeries, and WebSphere, etc.

About The Authors: Dr. Ying Zhao and Dr. Charles Zhou

Dr. Ying Zhao received her Ph.D. in Statistics and Artificial Intelligence from MIT, 1992. She is the President of Quantum Intelligence, Inc.
Dr. Charles Zhou received his Sc.D. from MIT, 1991. He is the Vice president of Quantum Intelligence, Inc.

Reference

Agrawal, R., Imielinski, T., Swamim, A. (1993). ``Mining Associations between Sets of Items in Massive Databases", *Proc. of the ACM SIGMOD Int'l Conference on Management of Data*, Washington D.C., May 1993, 207-216.

Charles Zhou, Ying Zhao (2002) “Words of a Feather”, DB2 Magazine, Vol. 7, #2, 2002, p.10. Also available at http://www.db2magazine.com/db_area/archives/2002/q2/miner.shtml

Ying Zhao, Charles Zhou and Jim Bainbridge, "Data into Information", invited paper for AS/400 News UK, Desktop, Workgroup & Enterprise Solutions for AS/400 Managers, September 1998.

Ying Zhao, et al (2001). “Increasing Campaign ROI with Blue Martini Marketing”, Blue Martini Internal Report, 2001

Ying Zhao, et al (2001). “Understanding, Targeting and Interacting with Customers”. Blue Martini Internal Report, 2001.

Ying Zhao, Charles Zhou and Jim Bainbridge, "Data into Information", invited paper for AS/400 News UK, Desktop, Workgroup & Enterprise Solutions for AS/400 Managers, September 1998.

George John, Ying Zhao "Mortgage Data Mining", pp 456-463, Proc. Of The 1997 International Conference on Financial Engineering, March, 1997, New York.

Ying Zhao, R. Schwartz, J. Sroka and J. Makhoul, ``Hierarchical Mixtures of Experts Methodology Applied to Continuous Speech Recognition", pp. 3443-3446, *Proc. of The 1995 International Conference on Acoustics, Speech, and Signal Processing*.

Ying Zhao and C. G. Atkeson, ``Implementing Projection Pursuit", the IEEE Transactions on Neural Networks, March, 1996.

Ying Zhao, R. Schwartz and J. Makhoul, ``Segmental Neural Net Optimization for Continuous Speech Recognition". In Advances in Neural Information Processing Systems 6, J. D. Cowan, G. Tesauro and J.Alspector, eds. Morgan Kaufmann Publishers, San Mateo, 1994.

G. Zavaliagkos, Ying Zhao, R. Schwartz and J. Makhoul, ``A Hybrid Segmental Neural Net/Hidden Markov Model System for Continuous Speech Recognition", vol.2, no. 1, pp.151-160, IEEE Transactions on Neural Networks, 1993.

Ying Zhao, ``On Projection Pursuit Learning", Ph.D. dissertation. Department of Mathematics and Artificial Intelligence Laboratory, M.I.T., 1992.

Ying Zhao and C. G. Atkeson, ``Some Approximation Properties of Projection Pursuit Learning Networks", In Advances in Neural Information Processing Systems 4. J. E. Moody, S. J. Hanson and R. P. Lippmann (eds.), San Mateo, CA, Morgan Kaufmann Publishers, 1991.

Ying Zhao and C. G. Atkeson, ``Projection Pursuit Learning", in Proc.of the 1991 International Joint Conference on Neural Networks, Seattle, WA, 1991.

Charles Zhou, Co-Author of “Siebel 7 Using DB2 UDB”, 276 pages, published by Vervante, April 2002. Also available at http://www.redbooks.ibm.com/pubs/pdfs/redbooks/sg246415.pdf

Charles Zhou, Die Yield Improvement Using Data Mining Technology for Pentium II, publication of Intel, 1998