VisBiz: A Simplified Visualization of Business Operation

Ming C. Hao, Daniel A. Keim*, Umeshwar Dayal
(ming_hao, umeshwar_dayal)@hp.com; *keim@informatik.uni-konstanz.de
Hewlett Packard Research Laboratories

1. Motivation

Business operations [1] involve many parameters and relationships and are modeled as complex business process workflows. The execution of these business processes generates vast volumes of complex data. Our goal is to use the complex information to analyze and improve business operations.

2. Our Approach

In this poster, we present a new technique, VisBiz, for interactively visualizing business operations. The basic idea of this technique is to visually mining relationships between important operation parameters (attributes) and to map the parameters into visualizations. VisBiz simplifies the complexity by partitioning the operation into multiple attribute circular graphs. VisBiz allows the analysis of business data as follows:

- Use a *business operation* visualization to analyze data distribution and the overall relationships between any set of three selected attributes, as illustrated in Figure 1A.
- Drill down to a two-attribute visualization to observe a subset of business operations (transactions), as in illustrated Figure 1B.

2.1. Business Operation Visualization

VisBiz transforms business attributes to nodes, with the lines between nodes representing a business case on a three-attribute circular graph. The three attributes are used for partitioning the left side, center, and right side of a circle.

VisBiz uses colored lines for specific business metrics. The width of lines represents the number of transactions processed in the node.

The order of the nodes is arranged from bottom to top of the circle based on the value of metrics.

2.2 Attribute Circular Graphs

The three-attribute graph can be drilled down to a two-attribute graph when a user selects one of the nodes from an attribute axis (i.e., Region 6 from attribute 1 in Figure 1A). It shows a subset of business process instances. This graph (as in Figure 1B), contains two half circles. The left and right nodes represent the other two attributes of the business operations (i.e., fraud amount and fraud count in Figure 1B).

2.3 Multiple Attribute Circular Graphs

To analyze a certain business operation problem, VisBiz employs a number of circular graphs. These graphs are linked together to visualize related process flows and relationships.

3. Application – Fraud Analysis

We have experimented with VisBiz for fraud analysis at HP Research Laboratories. For fraud detection, fraud specialists would like to discover new patterns and relationships in the transaction data. Examples of analysis that they typically need include the following:

1. What is the fraud distribution by regions?
2. What is the fraud growing rate in the last three years?
3. Which are the three countries with the highest fraud amount over time?
4. Where does the fraud come from? (Purchase or Cash Transactions?) What are the patterns?
5. Which types of credit cards have the most fraud? Issued by which country?

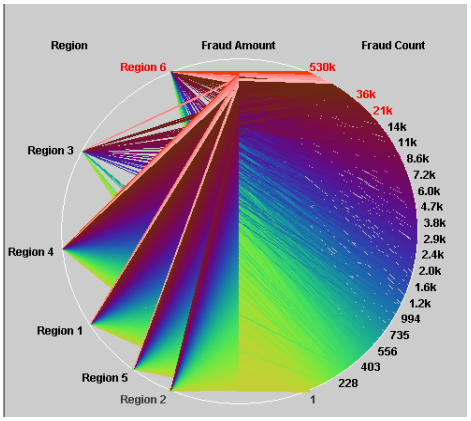


Figure 1A: VisBiz graph using the Region, Fraud Amount and Fraud Count attributes.

- Shows fraud distribution. ordered from bottom to top
- Region 6 has the highest fraud (more red and burgundy)
- Region 2 has the lowest fraud amount (more yellow and green).
- Discovers exceptions: red lines to represent high fraud amount or counts above the top 1%.
- Fraud amount and count has a high correlation.

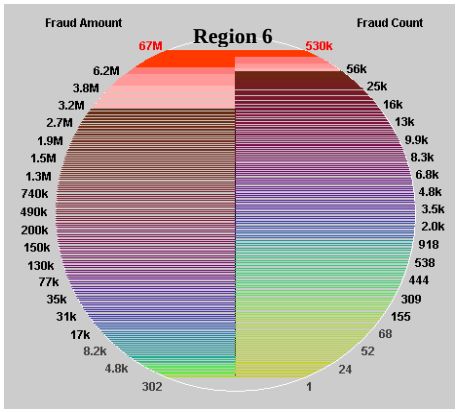
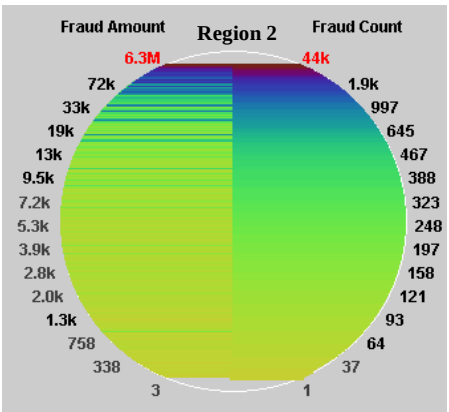


Figure 1B: Generated when the analyst clicks on Region 6 in Figure 1A.

- Shows high fraud amount and fraud count (more lines are burgundy).
- Shows more exceptions (more red)



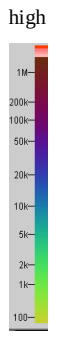


Figure 1C: Generated when the analyst clicks on Region 2 in Figure 1A.

- Shows low fraud amount and fraud count (more green and yellow).
- By comparing with the flows displayed in Figures 1B, Region 6 has the least fraud transactions and the highest fraud amount and count (less lines and more blue, burgundy, and red).

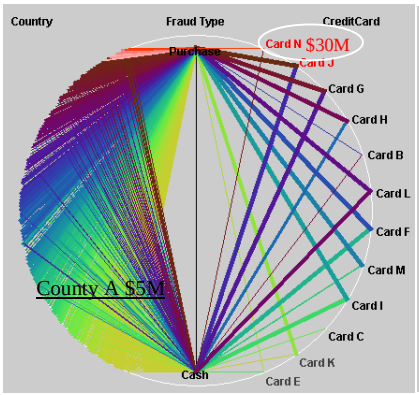


Figure 2A: Year of 2000

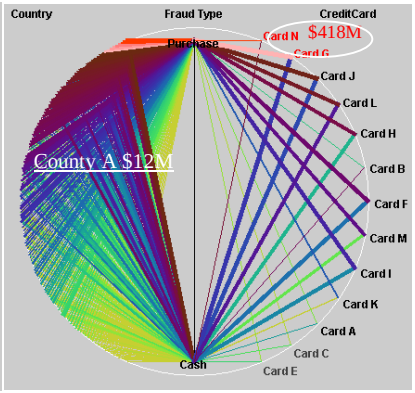


Figure 2B: Year of 2001

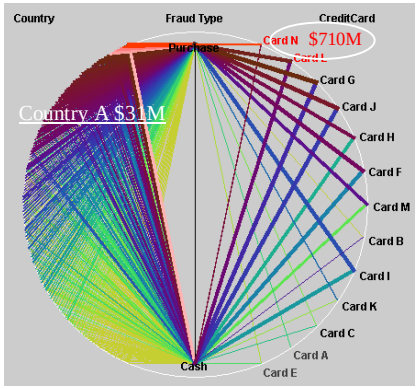
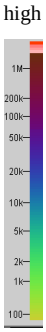


Figure 2C: Year of 2002



3.1 Fraud Distribution Analysis

To address the first question in Figure 1A, we select the three most related attributes – region, fraud amount, and fraud count. The attribute 1 nodes comprise the Region (1-6) and reside on the left side of the circle. The attribute 2 nodes are the fraud amounts and reside in the middle axis of the circle. The attribute3 nodes are the fraud counts and reside on the right side of the circle. The linked lines represent the connections between the nodes. The color represents the fraud amount. For fast identification, nodes are ordered by fraud amount from bottom to top (highest) on the circle. The analyst clicks on a node to show the relationships with the other two attributes. (i.e., fraud amount and count) as illustrated in Figure 1B and Figure 1C.

3.2 Fraud Growth Rate Analysis

To answer the second and third questions, we select three different attributes - country, fraud type and credit card - and map them on a circular graph over time as shown in Figures 2A, 2B, and 2C. VisBiz shows fraud growing patterns over three years. By comparing the flows displayed in figures 2A, 2B, and 2C, the analyst can quickly determine that the year 2002 has the most fraud transactions and the highest fraud amount (more lines, more blue and burgundy). The fraud rate has grown threefold from 2000 to 2002.

Color Lines are ordered by fraud amount. The countries with the highest fraud amount over time are in red at top of the left side of the circles.

3.3 Fraud Source Analysis

Figure 3A and Figure 3B are generated when the analyst drills down from Figure 2C on attribute 2 nodes (fraud source: Purchase and Cash). VisBiz answers the fourth question on where the fraud comes from (fraud source) and the patterns. Purchase has more red, burgundy, and blue color lines than Cash transactions. In Purchase, most fraud amounts are in the medium range (blue and green lines). Cash transactions have many small fraud amounts (yellow lines). Figure 3A has more lines (transactions) than Figure 3B. Most fraud comes from Purchase rather than Cash.

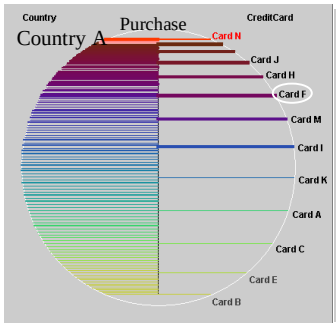


Figure 3A: Country and Purchase Fraud Analysis

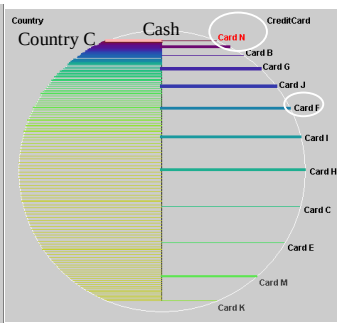


Figure 3B: Country and Cash Fraud Analysis

3.4 Credit Card Usage Analysis

Figure 4 answers the fifth question on credit card usage from a re-arranged three-attribute (Fraud Type, Credit Card, and Country) circular graph: Card N has the highest fraud amount (colored red). Card F has the most purchase transactions and is used by many countries (more lines). Country A has the most fraud with red lines and at the top of the circle. Most credit cards used by country A have the highest fraud amount (red lines). Cards N, L, G, and J have highest fraud amounts and most fraud from Purchase.

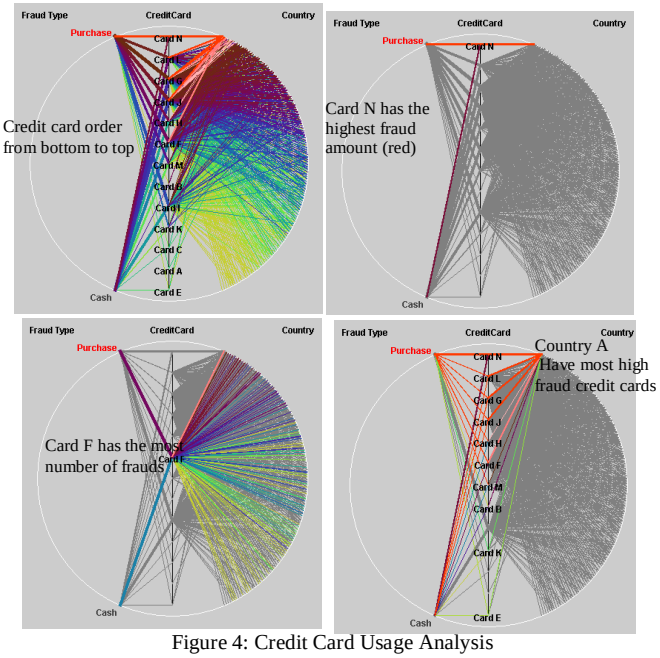


Figure 4: Credit Card Usage Analysis

Using the above information, the company is able to place strict control on certain countries, such as country A and certain credit cards, such as Card N (highest fraud). After knowing the sources of the frauds, the company will be able to take preventative actions.

4. Conclusion

In this poster, we simplify the visualization of complex business processes by partitioning the business operation into many linked attribute circular graphs. This visualization provides identification of key patterns and relationships. We have applied VisBiz to other applications, such as service contract analysis and SARS disease analysis. The result shows significant advantages of the VisBiz technique in finding patterns and exceptions.

Reference: [1] E_Bizinsights http://www.bizinsights.com