Interactive Visual Analysis of Anomalous Accounts Payable Transactions

Dr. Kishore Singh & Prof. Peter Best

Department of Accounting, Finance & Economics
Griffith University

Introduction

- Modern ERP systems record several thousands of transactions daily
- Difficult to find a few instances of anomalous activities among legitimate transactions
- CA/CM systems perform substantial analytics, but may produce lengthy reports → information overload
- Approaches that reduce the burden of excessive information are more likely to contribute to the overall effectiveness of the audit process
- We address this issue by demonstrating the use of visualization to present information graphically

Visualization

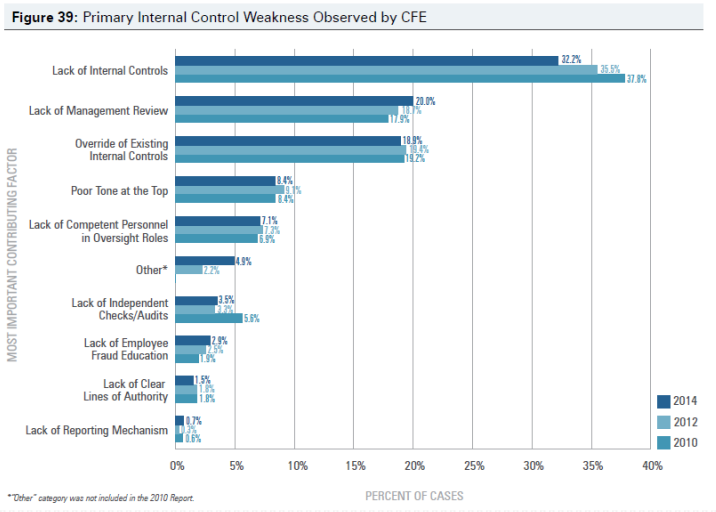
- Any technology that enable users to 'see' information - helps them better understand and put it into an appropriate context
- Patterns, trends and correlations that may go undetected in text-based data → exposed and recognised with less effort.
- High volume data - visualized as a collection of points in two-dimensional space

Framework for Research

- The visualizations developed in this paper are based on node-link diagrams
- Each node is shown as a point, circle, polygon, or some other graphical object, and each edge is shown as a line or curve connecting the two nodes
- Nodes are places in two-dimensional space, and edges represent relationships between the nodes
- Why node-link diagrams? They simplify identification of relationships
- Goal - create a representation that makes underlying data understandable and visually appealing

Anomaly Detection in Accounts Payable

- Key methods to detect
  - violations in segregation of duties
  - known fraud schemes
- *We focus on the former*



- ACFE (2014) - key indicators for frauds are lack of internal controls or an ability to override existing internal controls
- E.g. - to perpetrate a vendor fraud an employee - creates shell company and submit fictitious invoices for payment
- To successfully perpetrate this scheme requires violation of segregation of duties by creating (or modifying) vendor master records, and entering invoices for payment

Question 1

- *What types of visualizations may assist auditors in discovering potential anomalies in accounts payable transaction data?*

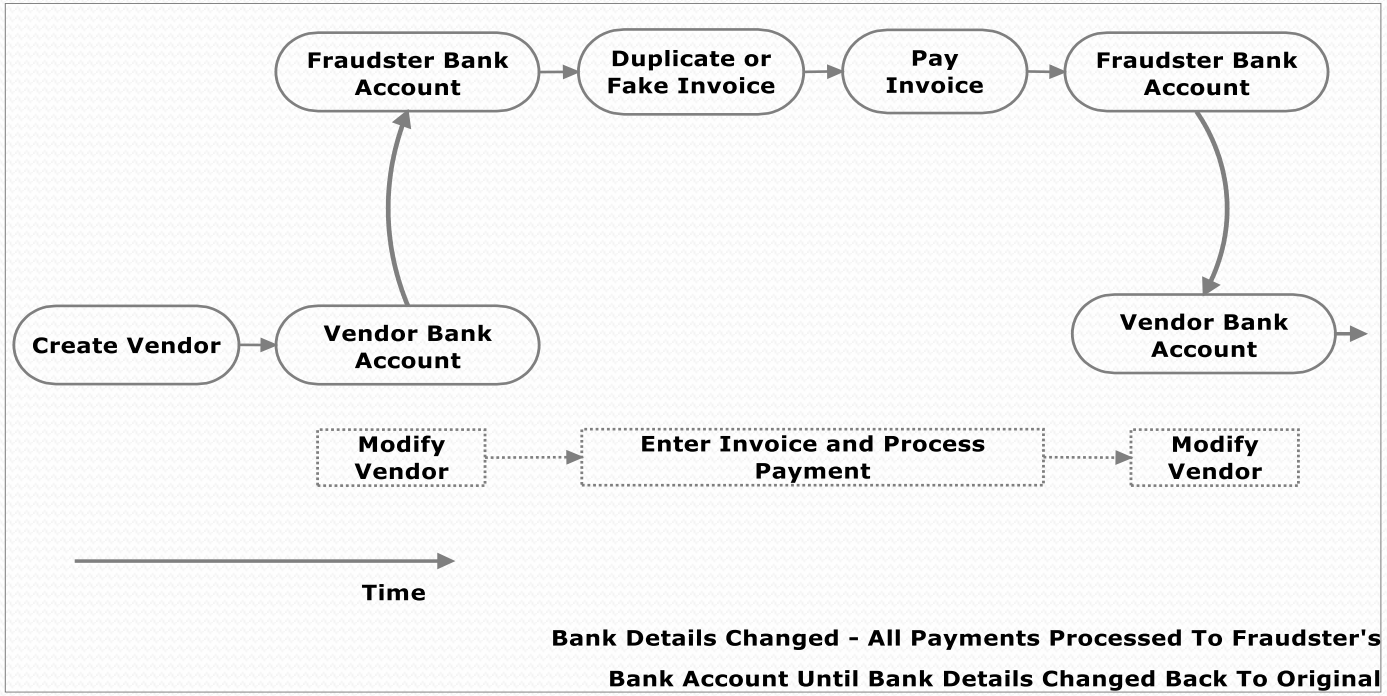
- Little and Best (2003) proposed the following two separation of duties principles for accounts payable
  - separation of master record maintenance from transaction entry
  - separation of payments and cheque entry from invoice data entry
- Motivation - users that have these authorizations are capable of creating shell companies and paying fictitious invoices without being detected.

Visualizations to Detect AP Fraud

- The following node-link visualizations are produced in this study to detect violations in SoDs:
  - Users performing vendor maintenance, entering invoice and processing payments
  - Users performing vendor maintenance and processing payments
  - Users performing vendor maintenance and entering invoices
  - Users entering invoices and processing payments

Special Case

- Modify existing legitimate vendor – change vendors banking details temporarily to fraudulent account, process payment, revert vendors banking details to the original values (flipping)



Visualizations to detect special case

- Vendors sharing bank accounts – if an employee sets up shell company to perpetrate vendor fraud and use a common account to have payments sent to, then amongst the visualization of vendor bank accounts, it will appear that both a legitimate vendor and one or more other vendors shared the same bank account at some point

- Vendors with multiple bank accounts – should an employee temporarily or permanently modify an existing legitimate vendors banking details (for genuine or fraudulent reasons), then these changes visually appear as though the vendor had more than one bank account at some point

- Time line analysis for vendor bank account changes (relates to vendors with multiple bank accounts) –list of transactions that are processed to any or all listed bank accounts that a vendor had at some point

Finding Collusion

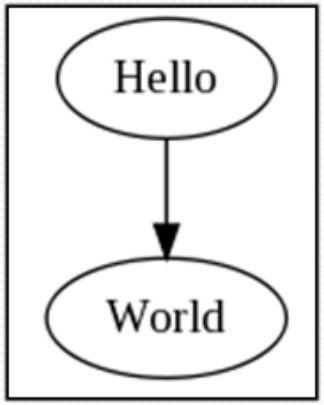
- Challenging - no “silver bullet”
- Employees collude to overcome well-designed internal controls
- Visualizations produced in this study have the potential to highlight such activities which may assist an auditor in directing their investigations

Question 2

- *How can a dataset be used to dynamically produce visualizations without user intervention?*
- Pre-processed data → source data for visualizations
- Visualizations produced in Graphviz
- Open source graph visualization software
- Uses the DOT language to describe graphs
- In DOT - three types of objects
  - Graphs
  - Nodes
  - Edges
- Graphs may be undirected or directed

How does it work?

- Several layout programs available in Graphviz
- Take descriptions of graphs written in DOT (syntax), and produce diagrams
- For example
  - This syntax: digraph G {Hello->World}
  - Produces:

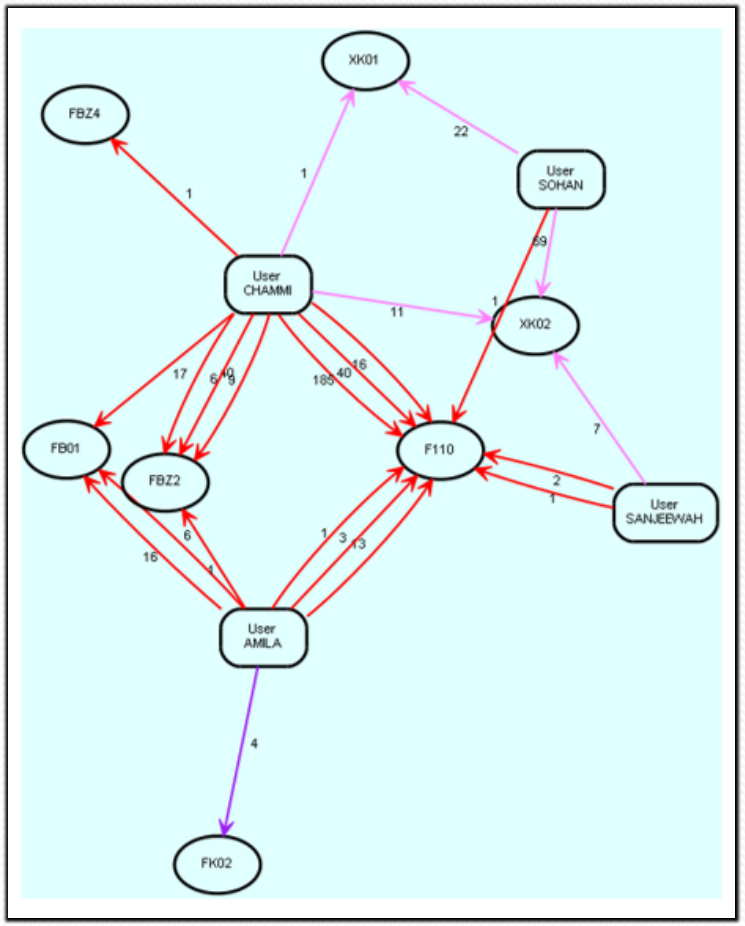


Question 2 version 2

- *How can a pre-processed data set be used to dynamically create DOT code which may be used to produce node-link visualizations in Graphviz?*
- DOT is simple yet complex
  - Several attributes need to be defined for graphs, nodes and edges
  - Nodes → ellipses, boxes, records or plaintext (no outline)
  - Node → polygon or record-based
  - Default node label is its name
  - Node and edge labels need to be set explicitly
  - Multi-line labels are possible
  - Colour attributes can be specified for nodes and edges
  - Other characteristics - orientation, size, spacing and placement are all configurable

DOT Example

- This visualization demonstrates relationship among users and types of transactions they perform
- It requires 74 lines of DOT code
- More complex visualizations may potentially contain hundreds or thousands of lines of DOT code that may vary from one visualization to the next



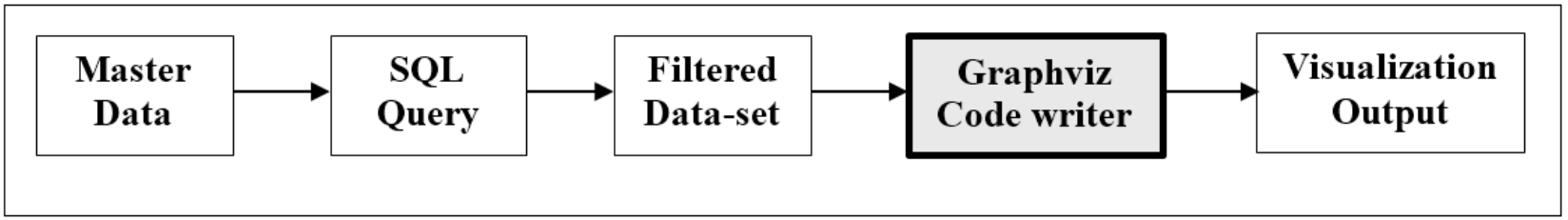
DOT Code for Example

- A section of the DOT code to produce the previous visualization

```
digraph G { bgcolor=lightcyan ranksep=3; ratio=auto; rotate=0;
overlap="false"; "Hub" [label = " ", fontname = "arial", fontsize = 8,
shape = "oval", style = "rounded", color = "transparent" ]  ;
"AMILA" [label ="User\nAMILA", fontname = "arial", fontsize = 8,shape
= "Mrecord", style = "bold", color = "black" ] ;
…
"F110" [label ="F110", fontname = "arial", fontsize = 8,shape =
"oval", style = "bold", color = "black" ] ;
…
"AMILA"  -> "F110" [ label="  1", penwidth=1.5, fontname =
"arial",fontsize = 8, color="red",arrowhead="vee" ] ;
```

The Solution

- Graphviz code writer – black box solution



- Requires filtered/pre-processed dataset
- Six step process

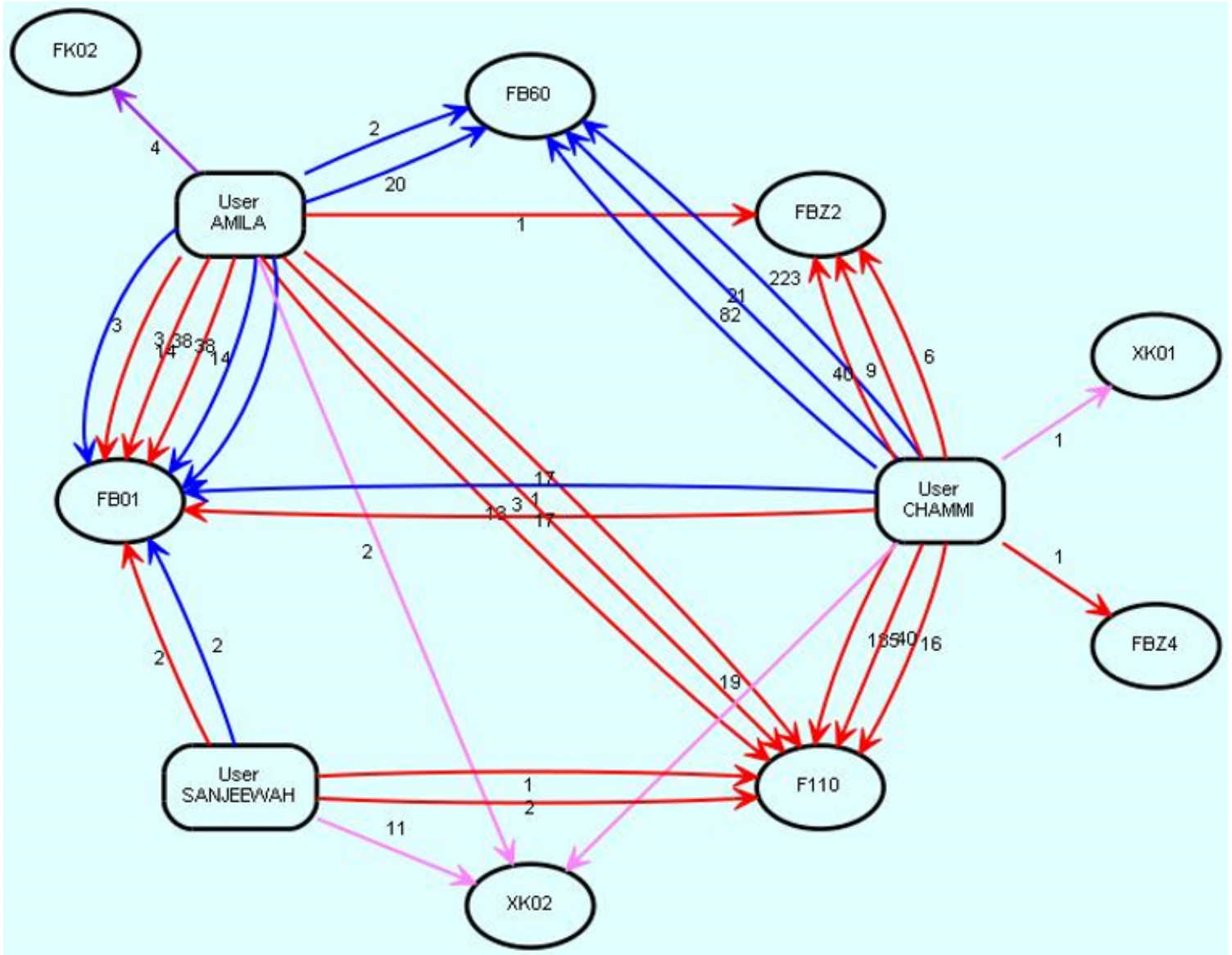
GraphViz Code writer process

```
digraph G { bg
overlap="false
shape = "oval"
"AMILA" [label
= "Mrecord", s
…
"F110" [label
"oval", style
…
"AMILA"  -> "F
"arial",fontsi
```

- Step 1
  - Read pre-processed data into code writer
- Step 2
  - Define type of graph (e.g. directed), preconfigure attributes
- Step 3
  - Extract all user nodes from dataset and preconfigure their attributes (SQL Select)
- Step 4
  - Extract all transaction nodes from dataset and preconfigure
- Step 5
  - Find associations between user and transaction nodes, establish edges, preconfigure attributes
- Step 6
  - Export DOT file for use buy layout program

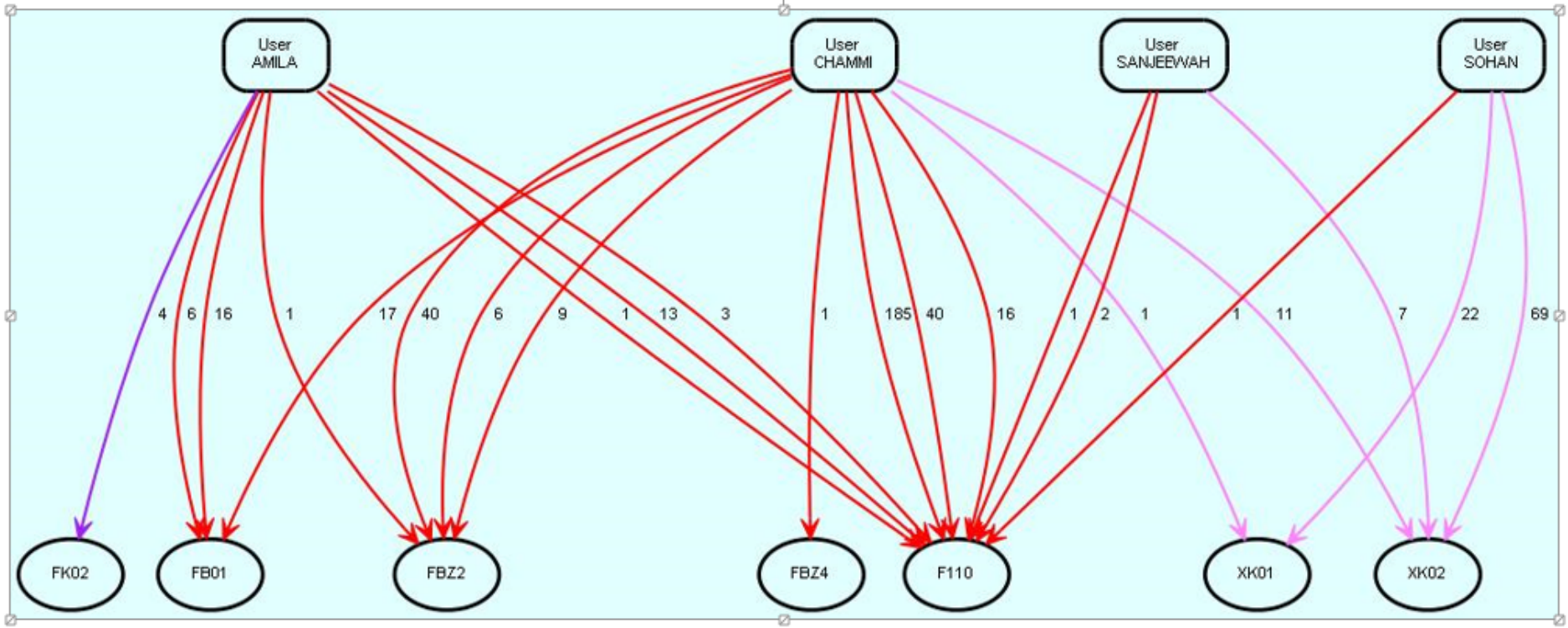
Implementation and Testing

- Tested on SAP ERP system of a large organization
- They provided a sample of accounting transaction data which included between 500,000 and 800,000 individual transactions across the various data tables, for a six month period
- Investigation
  - Violations in SoDs
  - Anomalies relating to vendor bank accounts



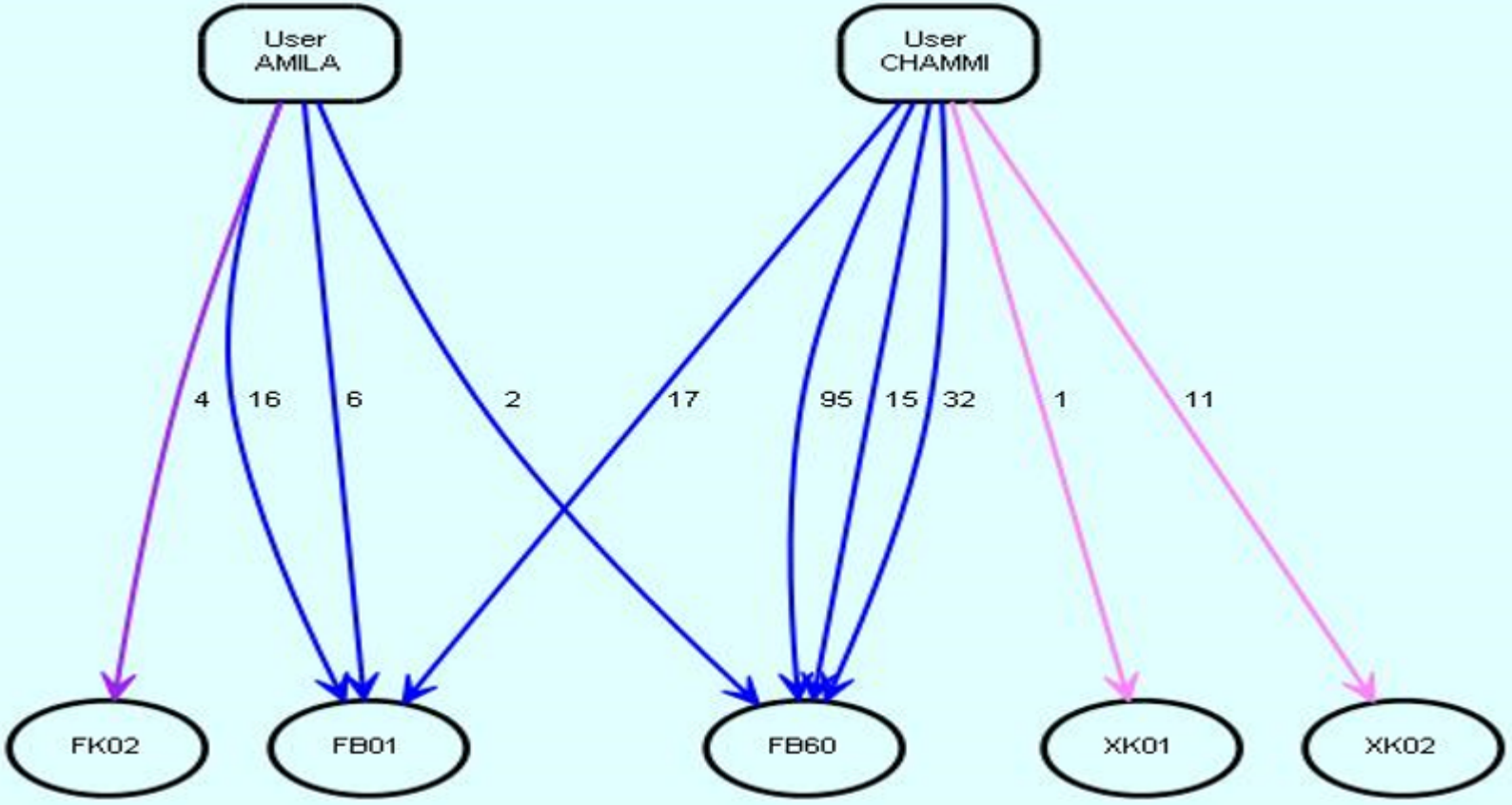
Users performing vendor maintenance, entering invoice and processing payments

(Note: multiple edges from a user node to a specific transaction code node indicate that the user has entered the same transaction type across multiple company codes within the SAP ERP system)



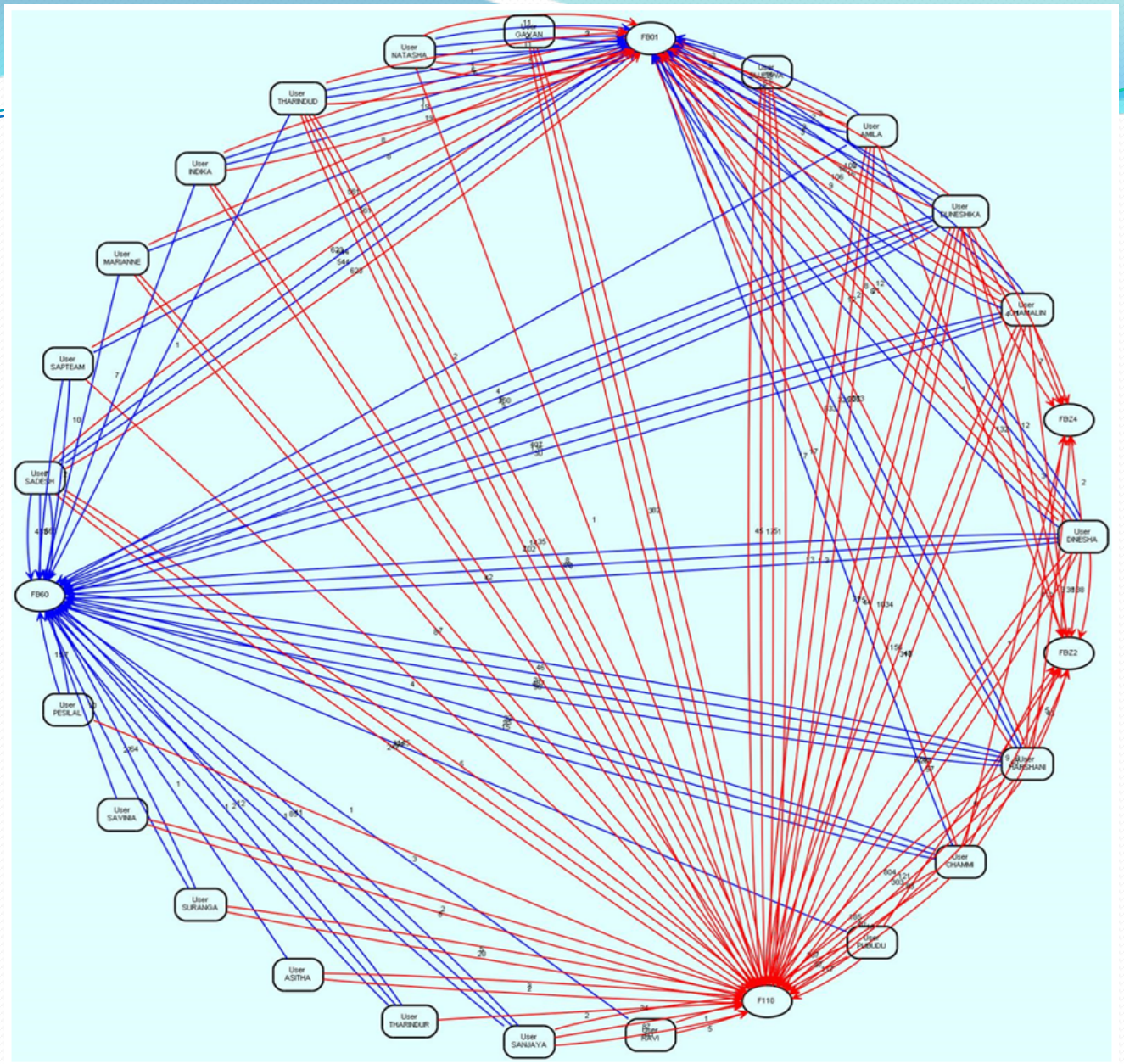
Users performing vendor maintenance and processing payments

(Note: multiple edges from a user node to a specific transaction code node indicate that the user has entered the same transaction type across multiple company codes within the SAP ERP system)



Users performing vendor maintenance and entering invoices

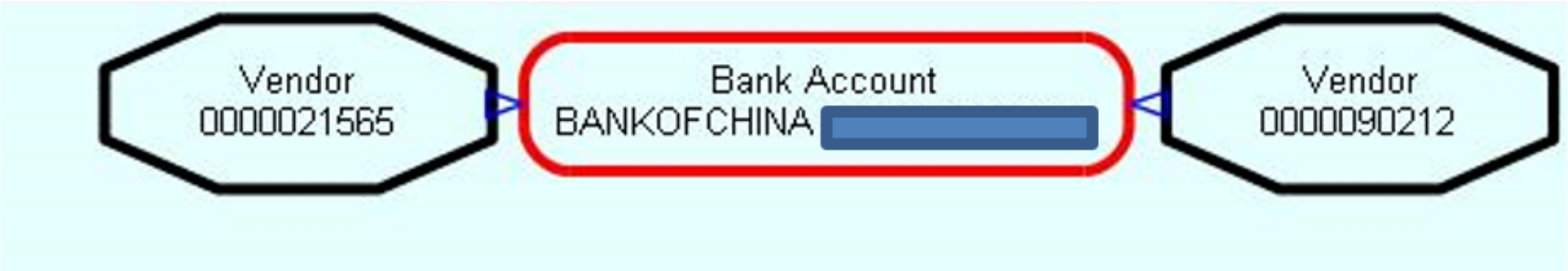
(Note: multiple edges from a user node to a specific transaction code node indicate that the user has entered the same transaction type across multiple company codes within the SAP ERP system)



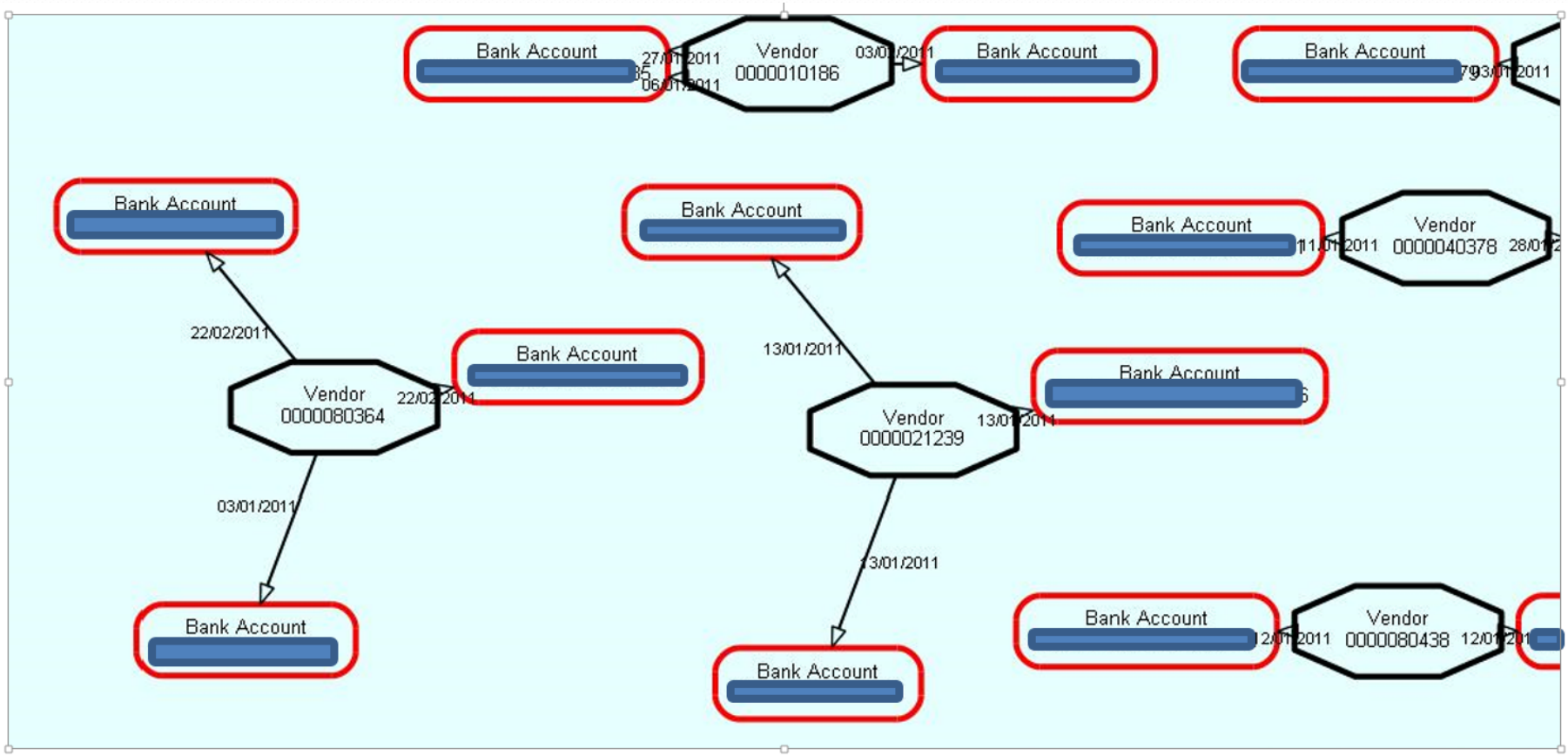
Users entering invoices and processing payments

(Note: multiple edges from a user node to a specific transaction code node indicate that the user has entered the same transaction type across multiple company codes within the SAP ERP system)

Vendors sharing bank accounts

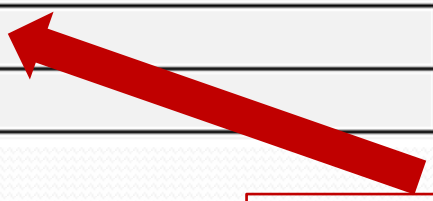


Vendors with multiple bank accounts

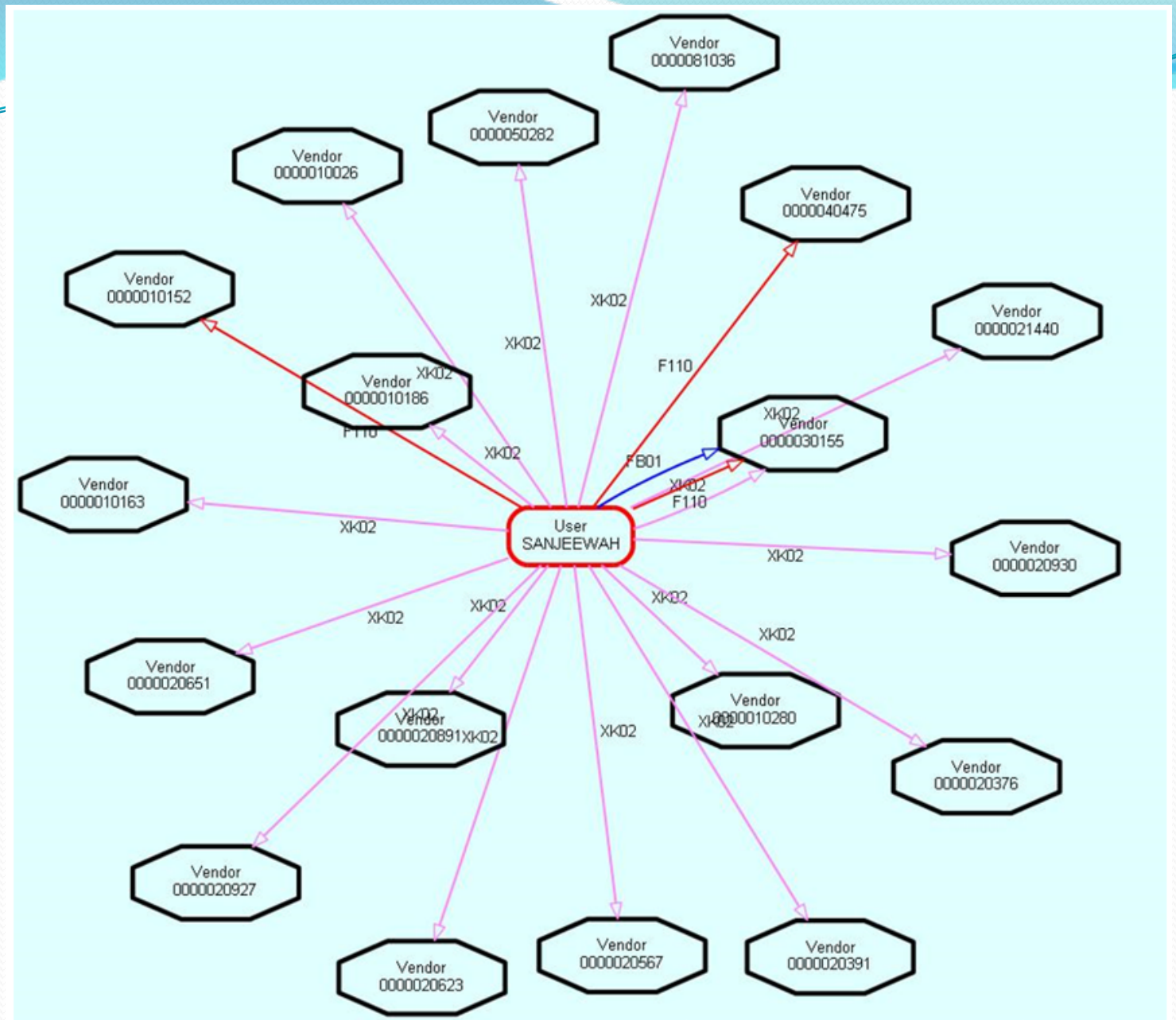


Time line analysis for vendor bank account changes

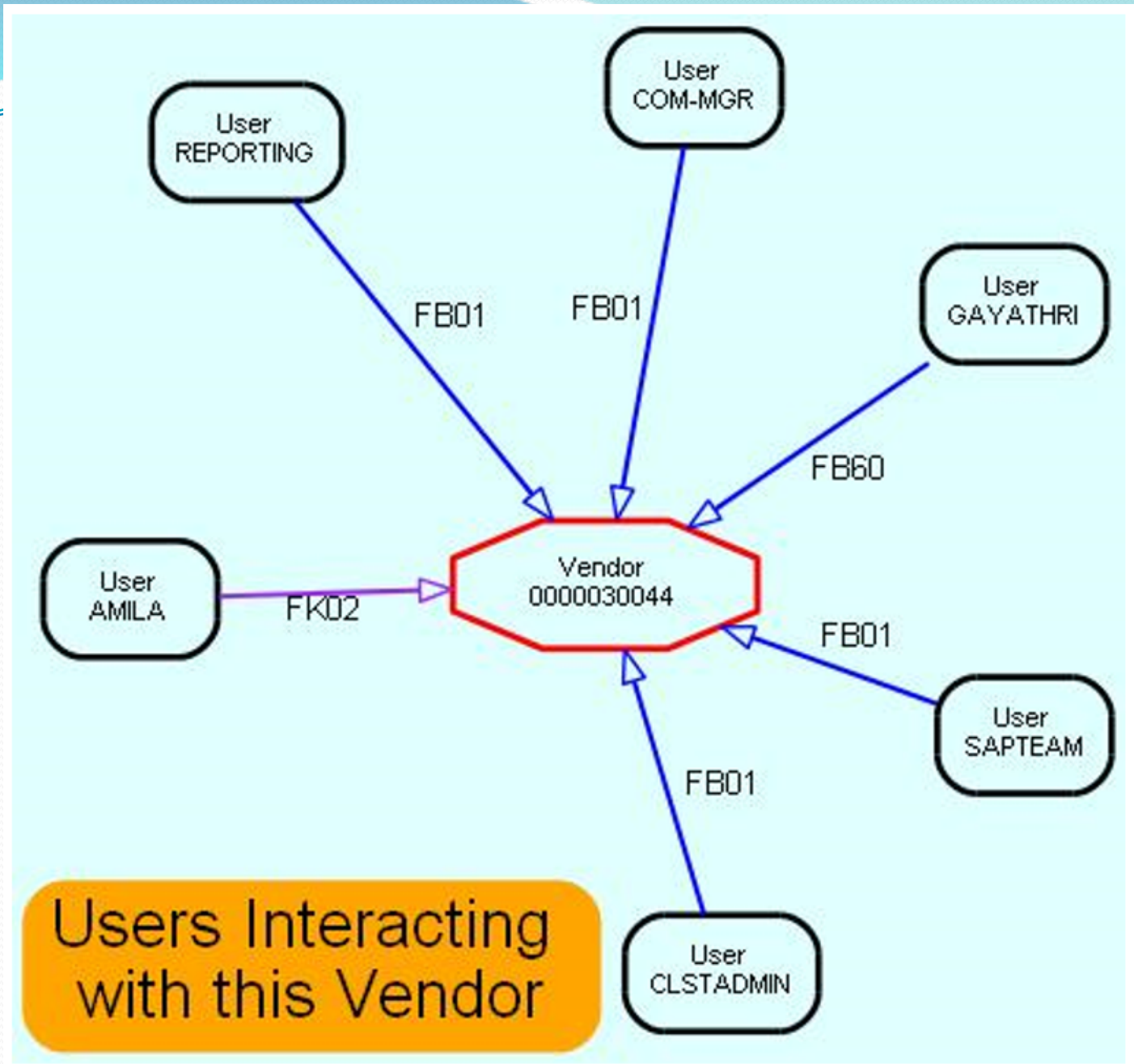
Date	User	Tcode	Amount	Bank Details	Doc. No
6/01/2011	SOHAN	XK02		[redacted]5000	00474973300
11/01/2011	INDIKA	FB60	$718.15		19000161190
24/01/2011	INDIKA	FB60	$1,800.00		19000161190
27/01/2011	SANJEEWAH	XK02		[redacted]5000	00048291325
3/02/2011	SOHAN	XK02		[redacted]0714	00048493355
9/03/2011	INDIKA	FB60	$422.15		19000161191
15/03/2011	INDIKA	FB60	$119.29		19000161192
24/03/2011	INDIKA	FB60	$357.00		19000161193
21/04/2011	INDIKA	F110	$1,406.70		15000025082
21/04/2011	INDIKA	FB01	$315.15		19000161195



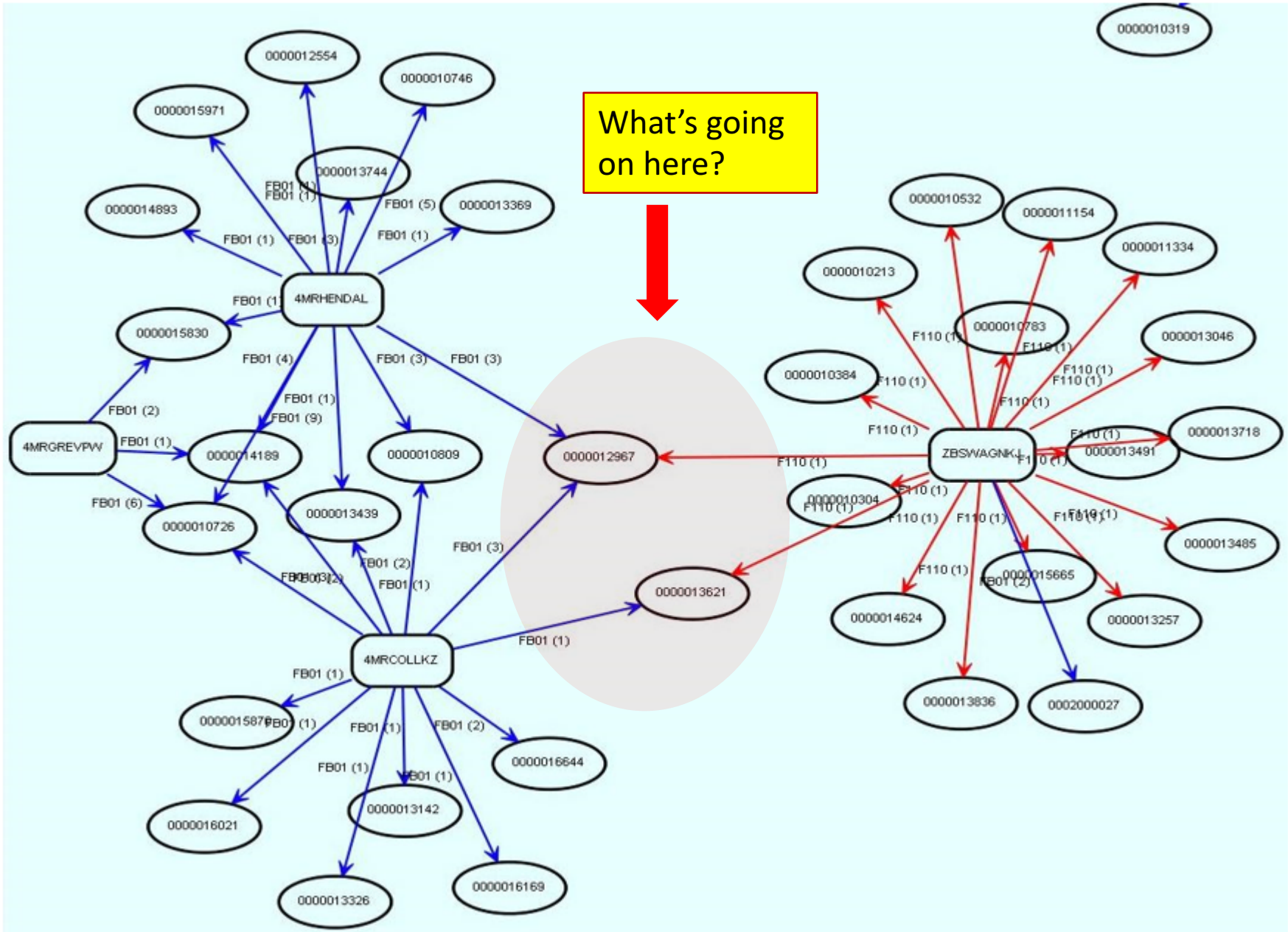
This is a payment



Detailed activities of a single risky user



Targeting a specific vendor to identify which users that have interacted with the vendor

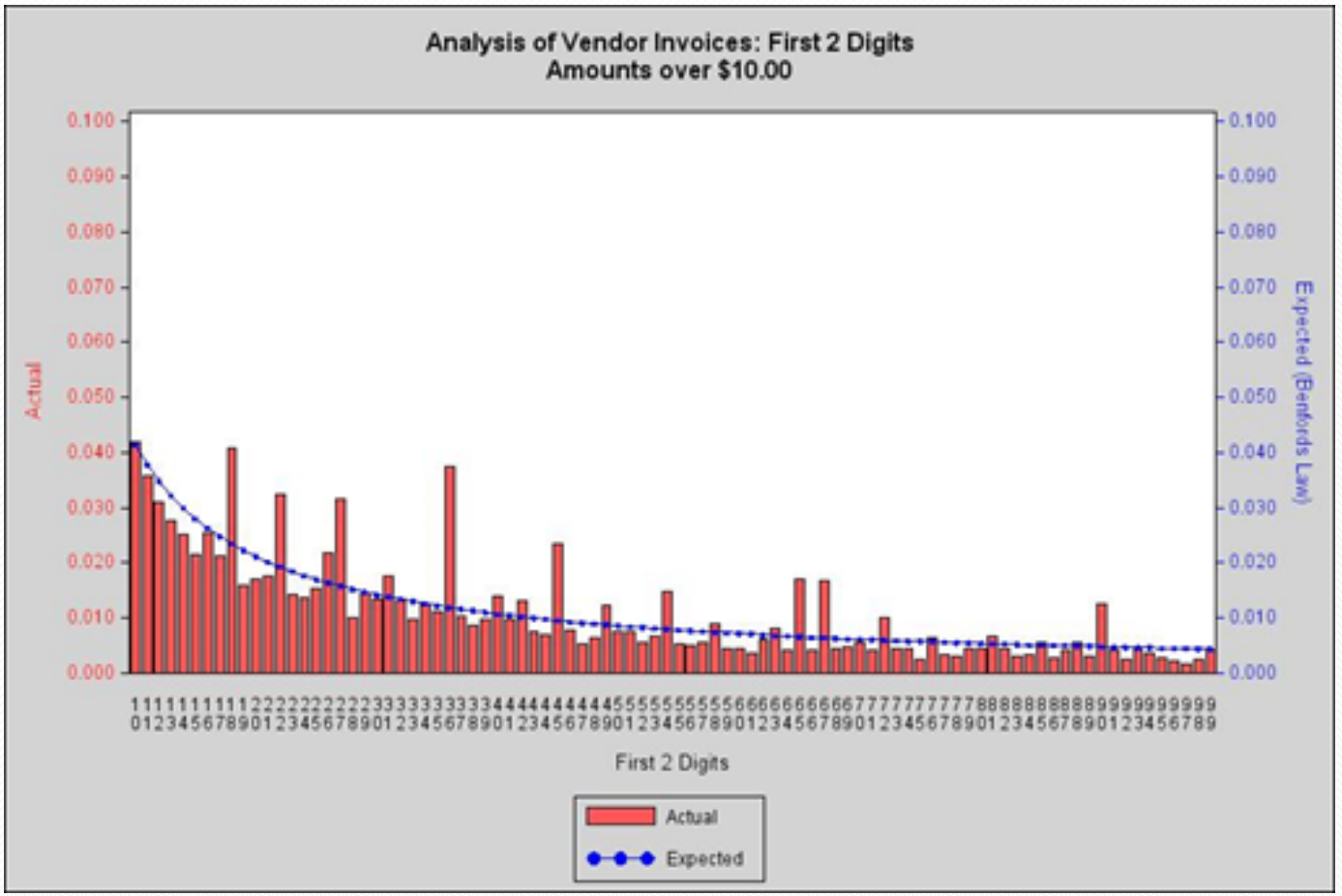


Potential to "see" relationships among multiple users and common vendors

Benford’s Law: Law of Large Numbers

- **Benford's law** of large numbers, gives expected frequencies of digits in numerical data.
- Analysis of the first two digits for vendor invoices revealed large deviations at **11, 22, 27, 36, 45, 54** and **67**.
- Other smaller deviations were also observed but appeared insignificant.
- **36** was selected as this was the largest**.** The investigation revealed 1217 invoice transactions, all containing **36** as the first two digits.
  - Several **identical** amounts appeared to have been recorded for the **same vendors**. These transactions were entered by different users. A follow up investigation was conducted and several duplicate invoices were discovered. *(Further details of this investigation were not provided by the organization).*

Benford's Analysis



Validation

- Reviewed by the Executive Director – Information Systems Audit of a top international accounting firm, stated: '…*Automated fraud detection software can provide internal auditors with a tool to efficiently assess the presence of fraud within an organization…. In general, I found the functionality of the tool to be useful. The user interface would require a minimal level of training and some level of understanding of the SAP application, which is a reasonable constraint. The graphs and visualizations clearly communicated a message for the reader.*'

Feedback from auditing practitioners

- Feedback from a panel of auditing practitioners was very positive.
- They found the visualizations easy to understand, and useful in aggregating large volumes of data.
- Visualizations were seen as enabling identification of relationships or patterns in data that would otherwise be difficult in textual data.
- Overall, the panel rated the visualizations as innovative and important tools in a fraud investigator's toolkit

Visualizations (charts & diagrams) *(Questionnaire scale 1 to 7)*	**Mean**	**Variance**	**Std. Dev.**
Easy to understand	5.87	0.87	0.92
Useful in aggregating large amount of information	6.09	0.54	0.73
Enables effective exploration of graphical data	6.13	0.57	0.76
Enables identification of relationships or patterns in data that are otherwise difficult to do in textual data	6.17	0.60	0.78
Enhances investigation and analysis for potential fraud	6.22	0.54	0.74
Are an innovative way of presenting information	6.35	0.42	0.65
Are an important tool in a fraud investigators toolkit	6.04	0.77	0.88
N=23			

Conclusion

- New and evolving opportunities for fraudsters
- Thousands of transactions daily generate thousands of lines of data in ERP system - novel approaches required to leverage the amount of data
- Hidden among gigabytes of data may possibly be fraudulent transactions - near impossible to detect.
- Forensic analysts and auditors seeking new and innovative methods to discover fraud
- Complete fraud detection is challenging - no “silver bullet”
- Visualization, when combined with other methodologies, may improve an auditor’s ability to identify suspicious activities not otherwise identifiable, and to encourage further investigations.

References

- K Singh & P Best (2016) Interactive visual analysis of anomalous accounts payable transactions in SAP enterprise systems. *Managerial Auditing Journal* 31 (1), 35-63
- Little, A. & Best, P. J. (2003) A framework for separation of duties in an SAP R/3 environment *Managerial Auditing Journal* 18 (5), 419-430
- ACFE (2014) Report to the Nation on Occupational Fraud and Abuse, http://www.acfe.com/rttn. *Accessed:* 2 June 2014