RUTGERS
THE STATE UNIVERSITY
OF NEW JERSEY

Exception Prioritization in Continuous Auditing: A Framework and Experimental Evaluation

Pei Li
David Y. Chan
Alexander Kogan

Outline

- **Motivation& Contribution**
- **Exception Prioritization Framework**
- **Experiment**
- **Conclusion & Limitation**

Motivation

- **Enormous Amount of Exceptions**:
  - Number of exceptions can be problematic and overwhelming to an internal audit department. Alles et al. (2006, 2008) finds that the management of exceptions as a pragmatic issue that may derail efficiency gain through automation.
  - Number of exceptions that can be investigated is positively correlated with available audit resources (ex. Labor and Time). (Chan and Vasarhelyi 2011).
- **High Expense of Internal Audit**:
  - According to *2010 Internal Audit Performance Report* from the Maricopa County Internal Audit, an internal audit function can cost more than one million dollars a year to operate (Maricopa County Internal Audit, 2010).

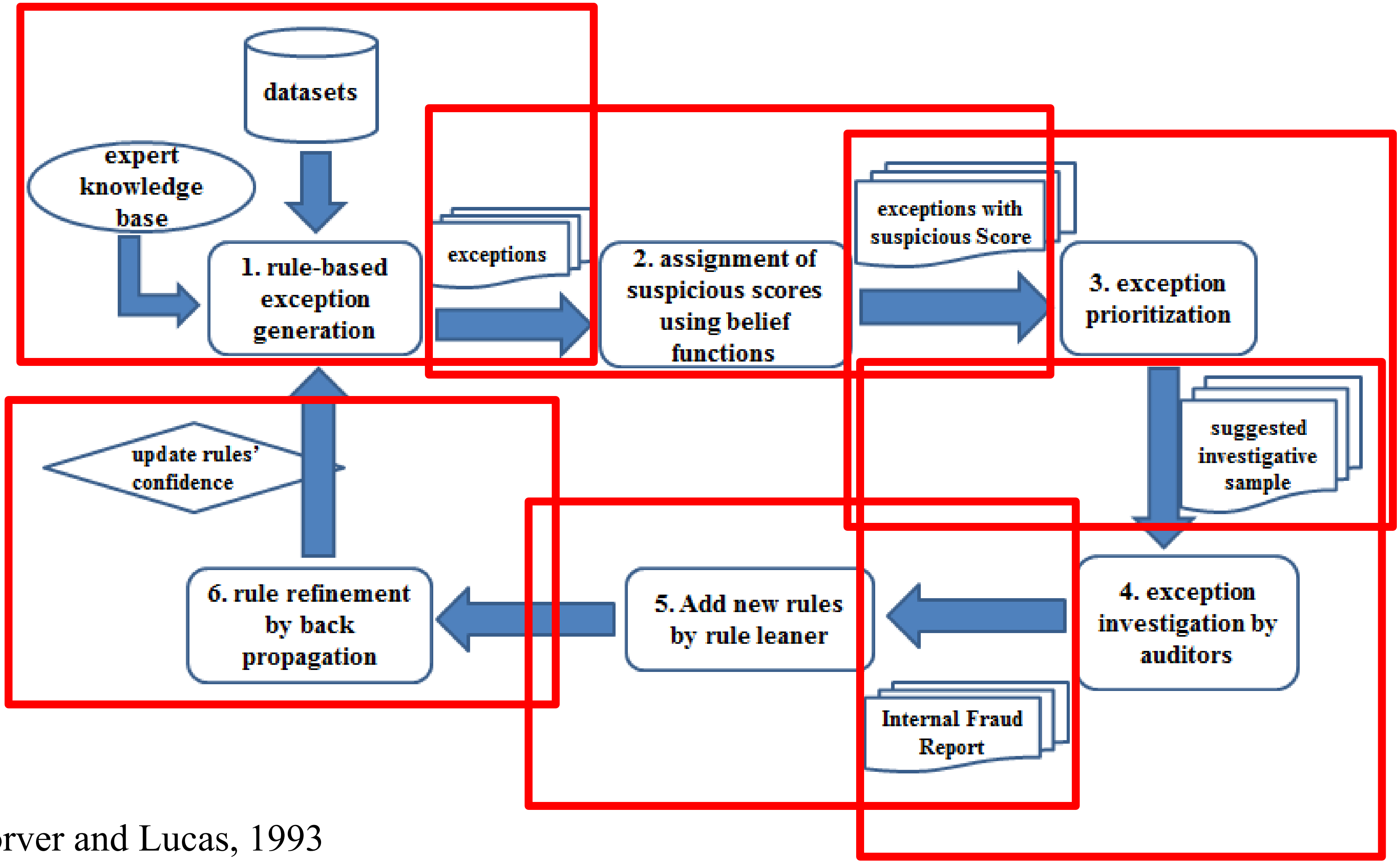
Contribution

- **Exception Prioritization Framework**
  - This paper offer a potential solution by proposing a framework that maximizes an audit department's limited resources.

Reasons for Using Belief Function

- **Belief Functions in Auditing**
  - Srivastava and Shafer (1992) argued that belief functions are useful to represent the auditors’ intuitive understanding of audit risk. (Srivastava and Shafer, 1990; Srivastava et al. 1996; Srivastava et al. 2007)
  - Different areas of risk assessment (Srivastava et al. 2009; Mock et al. 2009; Desai et al. 2010).
  - Perols and Murthy (2012); Walgampaya et al. (2010)

Proposed Framework

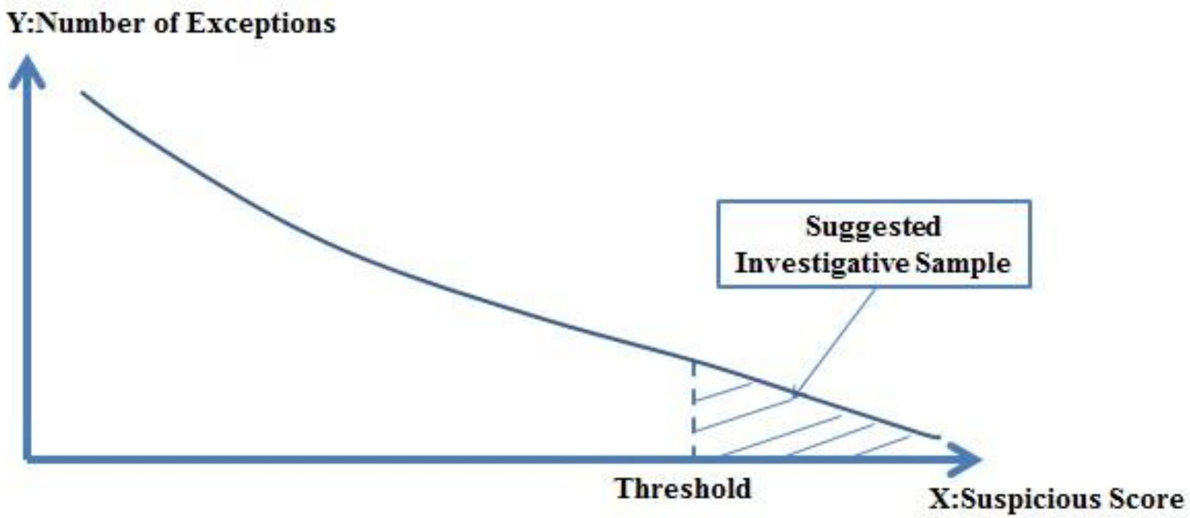


Korver and Lucas, 1993

Cohen,1995; Furnkranz and Widmer,1994

Rumelhart et al. 1986; Mahoney and Mooney,1993; Towell and Shavlik,1994

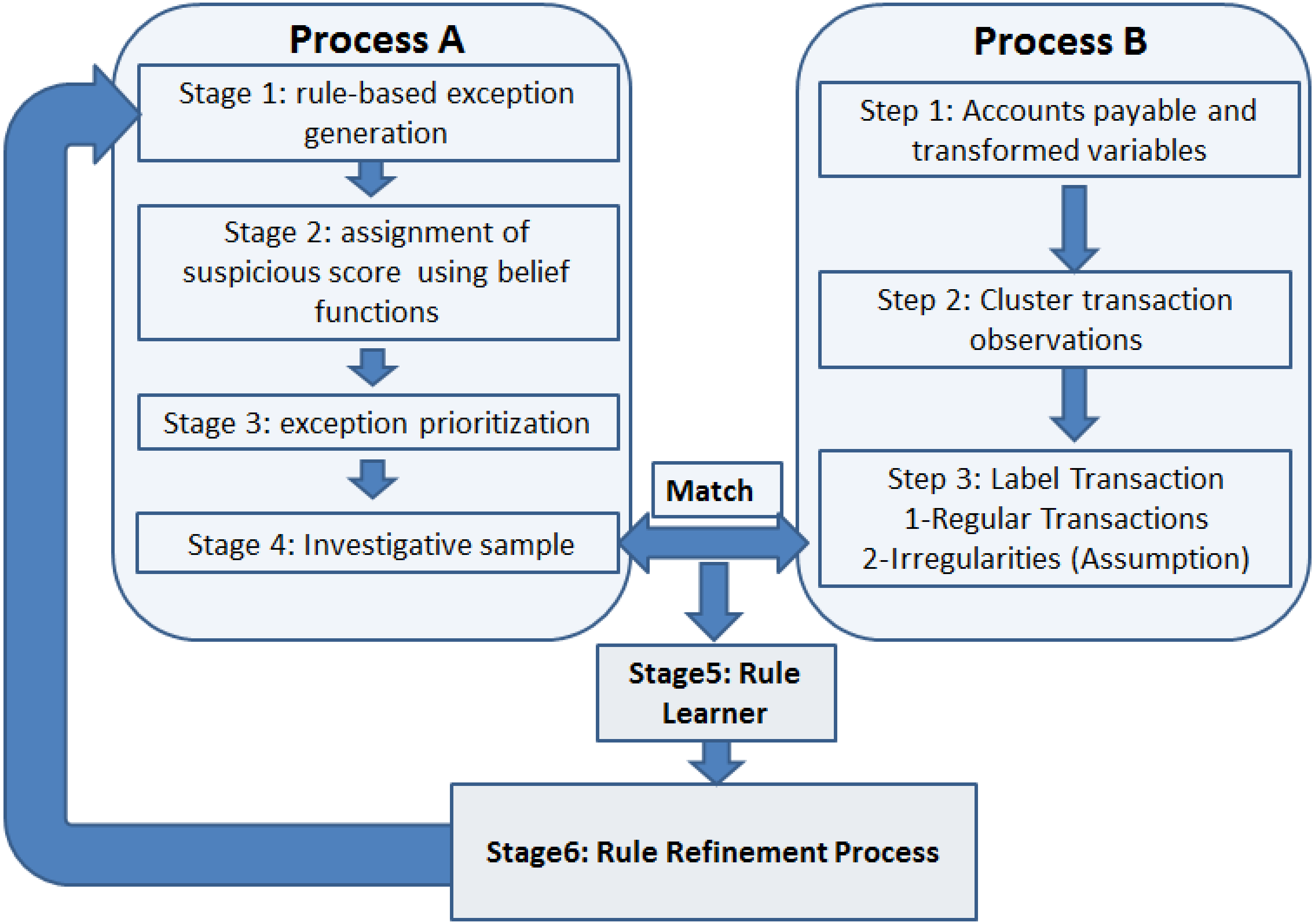
Exception Prioritization



Belief Function

- Dempster's rule is utilized to combine the confidence levels of rules for indicating fraud.
- Since the rules are independent, normalization is not required.
- If transaction *t* violates several rules, its suspicious score will be:
- We interpret the *Belief* function as the suspicious score for transaction *t*,$\mathbf{Bel}_{\mathrm{t}}(\sim f)$.
- $\mathbf{Bel}_{\mathrm{t}}(\sim f) = 1 - \prod_{R_i \in A_t}(1 - r_i)$

Experiment-Design



RUTGERS

Data Sample

Variables in Transaction Data

Variables	Category	Mean	Minimum*	Maximum*
Vendor_ID	Character	N/A	N/A	N/A
Vendor_Name	Character	Missing	Missing	Missing
Invoice_No	Character	N/A	N/A	N/A
Voucher_Description	Character	N/A	N/A	N/A
Invoice_Date	Numeric	3/31/2009	1/13/2000	6/24/2010
Amount	Numeric	8073.59	-46656.5	3435664
Tax_Amount	Numeric	308.3	-4241.5	44000
Goods_Amount	Numeric	7765.26	-42415	3435664
Voucher_No	Character	N/A	N/A	N/A
Invoice_Type	Character	N/A	N/A	N/A
Due_Date	Numeric	3/15/2010	2/12/2000	8/17/2010
Full_Pay_Status	Character	N/A	N/A	N/A
Date_Full_Payment_Due	Numeric	4/9/2010	4/1/2009	6/28/2010
Payment_Date	Numeric	4/30/2010	10/31/2008	6/30/2010
IDRUNPAY	Numeric	N/A	N/A	N/A
GL Account	Character	N/A	N/A	N/A
EXPAMTT	Numeric	704166	107000	70400000
Bank ID	Character	N/A	N/A	N/A
Payment ID	Numeric	N/A	N/A	N/A

*Minimum: if the variable is data format, it means the first date

*Maximum: if the variable is data format, it means the last data

89,712 Transactions

Expert-based Rules

Rules Overview	
Confidence level	Rules
Low	Missing Disbursement Date
Low	Missing Invoice Type
Low	Invoices with invalid GL account information
High	Disbursements Posted without Invoices
Low	Round Amount Disbursements (by line item)
Low	Round Amount Disbursements (by invoice)
Medium	Keywords Search
Medium	Outlier Analysis - Disbursements to Vendor
Medium	Outlier Analysis - Disbursements by G/L Account
High	Invoices with Vendors that Do Not Appear on the Vendor Master List
Low	Duplicate Invoices/Disbursements
N/A*	Payment Date vs. Due Date Analysis
N/A*	Payment Date vs. Invoice Date Analysis
N/A*	Gaps in Voucher Number Sequence
High	Payment is a Negative Amount
High	Payment is a Zero Dollar Amount
Medium	Payments Made on the Weekend
Medium	Payments Made on a Holiday
Low	Vendors with multiple invoices per day
	*N/A: Internal Auditors do not provide weight for a certain rule

Process B

Clustered Instances

0	17985	(20%)
1	2421	(3%)
2	1437	(2%)
3	1269	(1%)
4	2179	(2%)
5	629	(1%)
6	1094	(1%)
7	252	(0%)
8	7913	(9%)
9	18588	(21%)
10	1878	(2%)
11	328	(0%)
12	1876	(2%)
13	19974	(22%)
14	6366	(7%)
15	86	(0%)
16	5437	(6%)

Log likelihood: 37.81504

- Assumption: Treat irregularities as Fraud
- Irregularities Definition
  - Type 1: Cluster with small populations (less than 0.5%)
  - Type 2: Observation with low probability of being a member of a cluster (<0.55)
- Total amount of fraud is 2,933
  - 3.269% in the total population.
  - Type 1: 1,295.
  - Type 2: 1,638.

Experiment

- Two Processes
  - training process
  - testing process
- Split Data
  - Training: 67,284 (75%) with Fraud Observation 2,196
  - Testing: 22,428 (25%) with Fraud Observation 737
- Fix the investigative sample size to ten percent of the population, 6,728 observations.
- Trained: 13 cycles.

RUTGERS

TABLE 8
Testing Process*

Cycle	Number of Fraud in Investigative Sample*	Percentage of Detected Fraud	Mean/Median of Suspicious Score of Fraud in the Investigation Sample**	Mean/Median of Suspicious Score of Normal Transactions in the Investigation Sample**	Mean/Median of Suspicious Score of Fraud in the Whole Sample **	Mean/Median of Suspicious Score of Normal Transactions in the Whole Sample **
1	271	36.77%	0.8203 (0.8125)	0.7442 (0.7188)	0.6023 (0.5781)	0.3981 (0.4375)
2	301	40.84%	0.7808 (0.7605)	0.7136 (0.6835)	0.5766 (0.5420)	0.3710 (0.4044)
3	299	40.57%	0.7602 (0.7260)	0.6790 (0.6456)	0.5491 (0.5019)	0.3430 (0.3692)
4	299	40.57%	0.7398 (0.6898)	0.6431 (0.6031)	0.5240 (0.4770)	0.3142 (0.3395)
5	302	40.98%	0.7136 (0.6482)	0.6006 (0.5596)	0.4944 (0.4462)	0.2828 (0.3013)
6	305	41.38%	0.6836 (0.6276)	0.5538 (0.5093)	0.4610 (0.3925)	0.2499 (0.2619)
7	309	41.93%	0.6499 (0.6029)	0.5036 (0.4596)	0.4252 (0.3485)	0.2169 (0.2193)
8	303	41.11%	0.6264 (0.5821)	0.4525 (0.4043)	0.3910 (0.3166)	0.1834 (0.1771)
9	317	43.01%	0.5855 (0.4922)	0.4040 (0.3537)	0.3565 (0.2648)	0.1521 (0.1377)
10	315	42.74%	0.5695 (0.4608)	0.3654 (0.3188)	0.3334 (0.2538)	0.1261 (0.1023)
11	313	42.47%	0.5540 (0.4446)	0.3296 (0.2923)	0.3075 (0.2176)	0.1008 (0.0702)
12	312	42.33%	0.5399 (0.4145)	0.2975 (0.2621)	0.2863 (0.1941)	0.07886 (0.0524)
13	308	41.79%	0.5304 (0.3859)	0.2701 (0.2273)	0.2658 (0.1615)	0.0608 (0.0235)

*The updated confidence level of rules in each iterative run generated in training process are applied into the testing set. The size of the investigative Sample is fixed to ten percent of the population, 2,242 obervations. The number of fraud in the testing sample is 737.

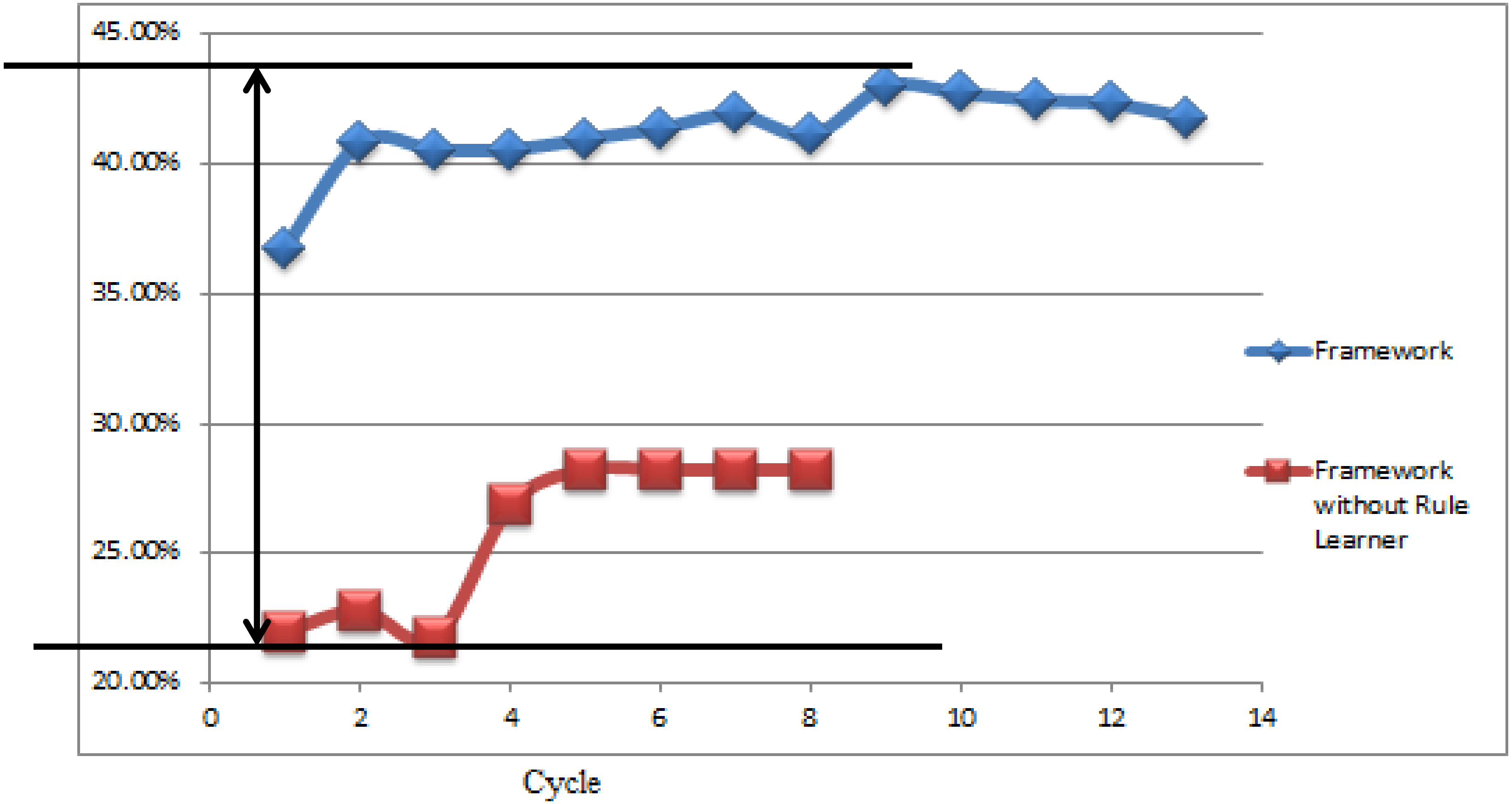
**The number of fraud placed in the investigative sample are shown. Each row summarizes the suspicious scores of fraud and normal transactions both in the whole sample and the investigative sample.

RUTGERS

Testing Process

FIGURE 7

Testing Process: Percentage of Fraulent Transactions Located in Investigative Sample



This figure plots the percentage of fraudulent transaction placed in the investigative sample during the tesing process. It also plots the percentage of fraud in the investigative sample using the framework without Rule Learner

RUTGERS

Limitation

- Assume all the rules are independent (Korver and Lucas 1993).
- Use artificially generated fraudulent data.

Conclusion

- It provides a framework to maximizes an audit department's limited resources.
- The proposed framework uses belief functions to prioritize exceptions.
- Identified fraudulent transactions increasing from 21.98% to 43.01%

Thank you for the time and valuable comments

Process B

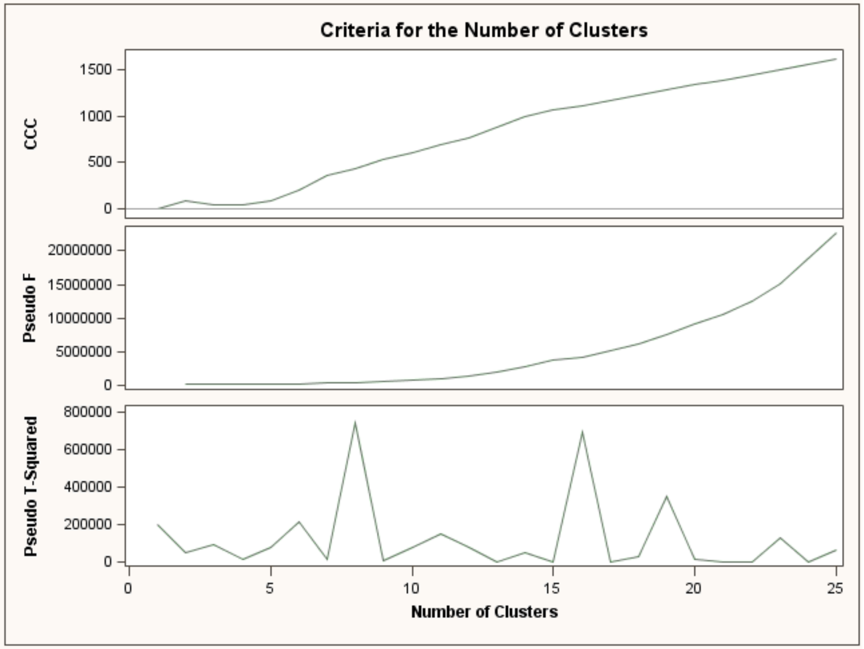
Step1: Accounts payable and transformed variables

	Transaction Variables
1	Invoice_Date
2	Tax_Amount
3	Goods_Amount
4	GL Account
5	Payment ID

	Rule Related Variables
1	Miss_Date
2	Miss_Invoice_Type
3	Invalid_GL
4	Miss_Invoice_No
5	RoundAmount_Line
6	RoundAmount_Invoice
7	FraudKeyword
8	FCPAKeyword
9	Outlier_Disbur
10	Outlier_GL
11	Invalid_Vendor
12	Duplicate
13	Payment_Due_Date
14	Payment_Invoice_Date
15	GAP_Voucher_No
16	Payment_Negative_Amount
17	Payment_Zero_Amount
18	PaymentOnWeekend
19	Payment_On_Holiday
20	MultiInvoice_Per_Day

Step 2: Clustering Transaction Data

- Determine the Number of Clusters
  - Using PROC CLUSTER in SAS software
  - SAS provide the three statistics for estimating the number of clusters.
    - cubic clustering criterion (CCC)
    - pseudo F statistic (PSF)
    - pseudo t2 statistic
  - Potential Number of Cluster:
    - 7, 9, 13, 15, 17,20



Stage5 Rule Learner

- Unbalance Issue
  - under-sampling method (Chawla et al. 2002)
  - oversampling (Chawla et al. 2002)
  - combination of oversampling and under-sampling (Chawla et al. 2002)
  - cost-sensitive classifier (Zadrozny et al. 2003),
  - meta-cost classifier (Domingos, 1999).

Table 4

Comparison of Methods Dealing with Unbalanced Dataset

	Original	Undersample	Oversampling	Combination	Cost-Sensitive	MetaCost Classifier
TP rate*	0.1832	0.9701	0.3772	0.8847	1.0000	0.5890
FP rate **	0.0003	0.2136	0.0092	0.1698	1.0000	0.0382

***TP rate:** it measures the proportion of fraudulent observations which are correctly detected as fraudulent transactions.

****FP rate:** it measures the proportion of non-fraudulent observations are correctly identified as such.

Original: do not use any method to deal with the unbalanced issue. It works as a comparison.

Stage5 Rule Learner

Table 5

Undersample Methond

Rate *	**1.0**	**1.1**	**1.2**	**1.3**	**1.4**	**1.5**	**1.6**	**1.7**	**1.8**	**1.9**
TP rate **	0.965	0.970	0.969	0.959	0.934	0.929	0.879	0.878	0.729	0.807
FP rate ***	0.202	0.214	0.187	0.208	0.143	0.157	0.145	0.121	0.116	0.097

***Rate:** it is the ratio of undersampling the training set between the fraudulent class and the non-fraudulent class.

****TP rate:** it measures the proportion of fraudulent observations which are correctly detected as fraudulent transactions.
*****FP rate:** it measures the proportion of non-fraudulent observations are correctly identified as such.

Training Process

TABLE 7
Training Process*

Cycle	Number of Fraud in Investigative Sample**	Percentage of Detected Fraud	Mean(Median) of Suspicious Score of Fraud in the Investigation Sample**	Mean(Median) of Suspicious Score of Normal Transactions in the Investigation Sample**	Mean(Median) of Suspicious Score of Fraud in the Whole Sample **	Mean (Median) of Suspicious Score of Normal Transactions in the Whole Sample **
1	787	35.84%	0.832 (0.841)	0.747 (0.718)	0.607 (0.578)	0.400 (0.437)
2	831	37.84%	0.802 (0.799)	0.715 (0.683)	0.581 (0.541)	0.373 (0.405)
3	854	38.89%	0.778 (0.785)	0.681 (0.645)	0.555 (0.501)	0.345 (0.371)
4	887	40.39%	0.752 (0.738)	0.645 (0.603)	0.530 (0.477)	0.316 (0.339)
5	887	40.39%	0.730 (0.696)	0.603 (0.559)	0.502 (0.446)	0.284 (0.301)
6	888	40.44%	0.705 (0.656)	0.556 (0.513)	0.469 (0.392)	0.251 (0.262)
7	915	41.67%	0.670 (0.643)	0.506 (0.461)	0.435 (0.348)	0.218 (0.224)
8	898	40.89%	0.650 (0.653)	0.455 (0.406)	0.402 (0.316)	0.185 (0.183)
9	918	41.80%	0.617 (0.613)	0.406 (0.356)	0.368 (0.264)	0.153 (0.149)
10	959	43.67%	0.588 (0.508)	0.368 (0.320)	0.346 (0.253)	0.127 (0.117)
11	964	43.90%	0.571 (0.469)	0.332 (0.298)	0.321 (0.217)	0.102 (0.089)
12	963	43.85%	0.557 (0.429)	0.300 (0.262)	0.300 (0.194)	0.079 (0.061)
13	960	43.72%	0.546 (0.393)	0.272 (0.230)	0.281 (0.161)	0.062 (0.042)

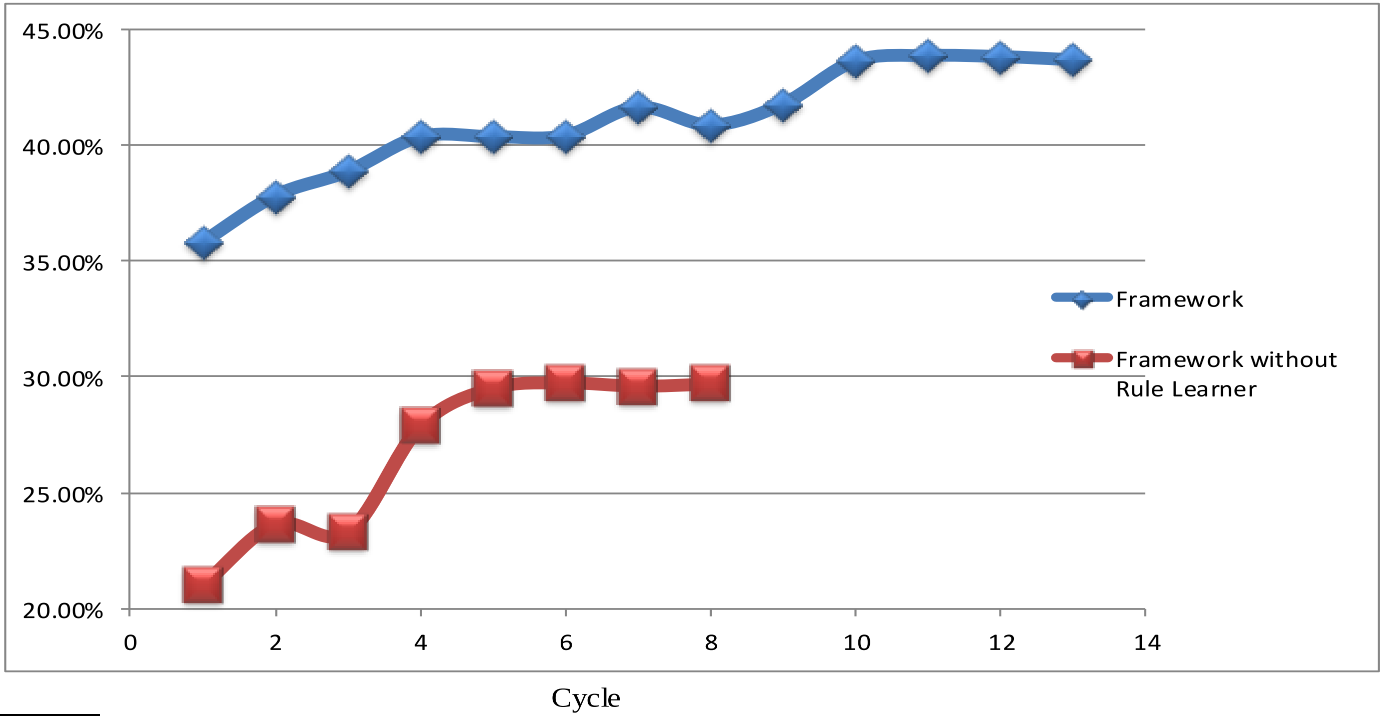
*The size of the investigative Sample is fixed to ten percent of the whole population, 6,728 obervations. The number of fraud in the training set is 2,196.

**The number of fraud placed in the investigative sample are shown. Each row summarizes the suspicious scores of fraud and normal transactions both in the whole sample and the investigative sample.

Training Process

FIGURE 6

Training Process: Percentage of Fraulent Transactions Located in Investigative Sample



This figure plots the percentage of fraudulent transaction placed in the investigative sample during the training process. It also plots the percentage of fraud in the investigative sample using the framework without Rule Learner