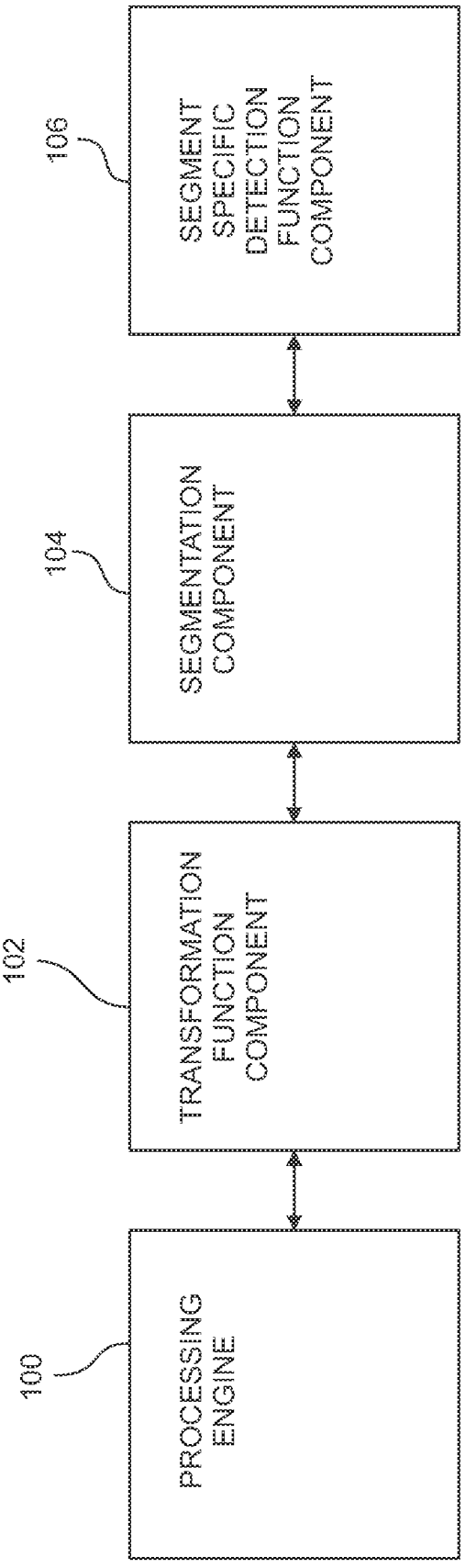

PROCESSING ENGINE
100
TRANSFORMATION FUNCTION COMPONENT
102
SEGMENTATION COMPONENT
104
SEGMENT SPECIFIC DETECTION FUNCTION COMPONENT
106

Fig. 1

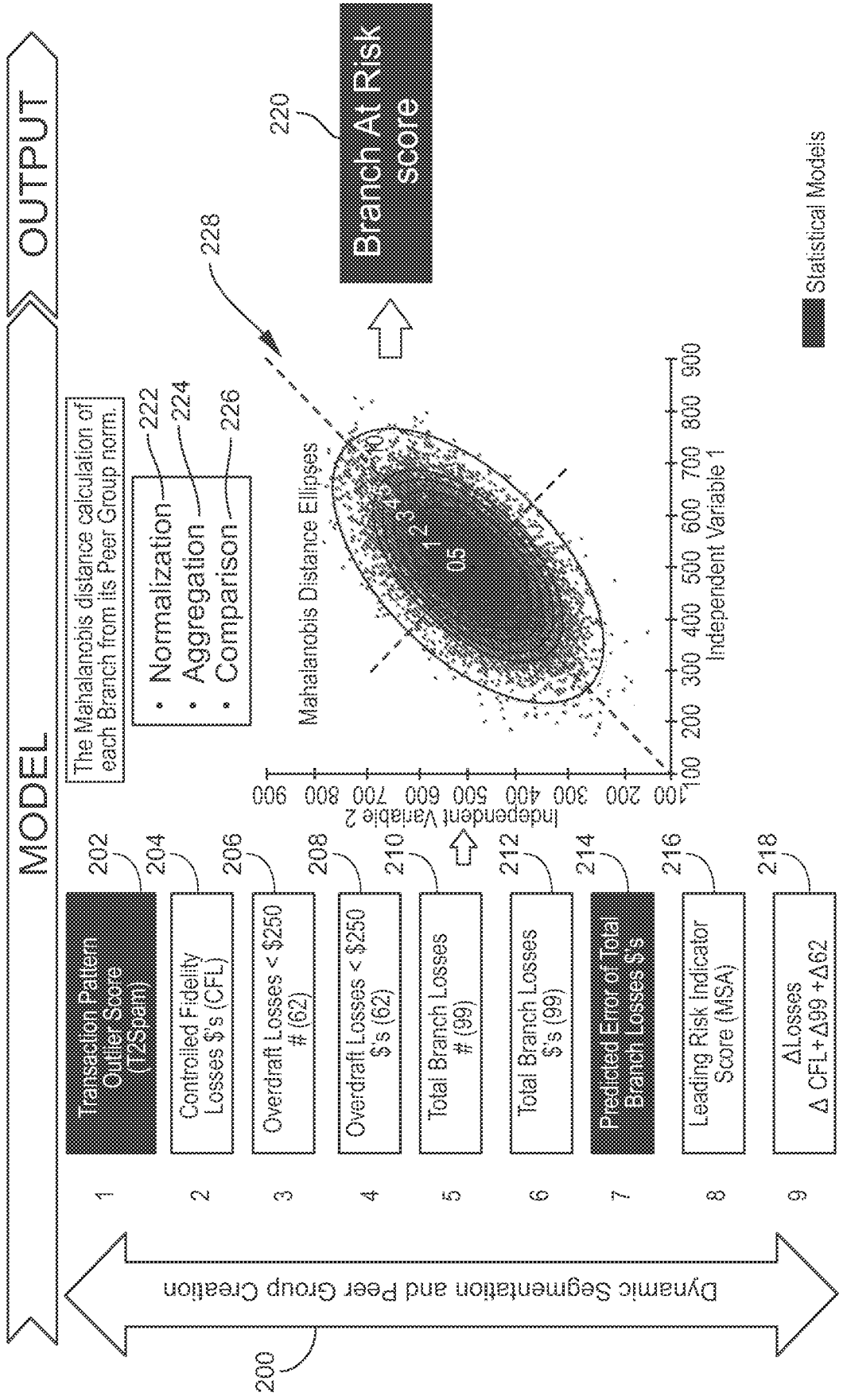

MODEL
OUTPUT
200
Dynamic Segmentation and Peer Group Creation
1
2
3
4
5
6
7
8
9
202
Transaction Pattern Outlier Score (T2Span)
204
Controlled Fidelity Losses $'s (CFL)
206
Overdraft Losses < $250 # (62)
208
Overdraft Losses < $250 $'s (62)
210
Total Branch Losses # (99)
212
Total Branch Losses $'s (99)
214
Predicted Error of Total Branch Losses $'s
216
Leading Risk Indicator Score (MSA)
218
ΔLosses Δ CFL + Δ99 + Δ62
The Mahalanobis distance calculation of each Branch from its Peer Group norm.
222
224
226
• Normalization
• Aggregation
• Comparison
228
Mahalanobis Distance Ellipses
Independent Variable 2
100 200 300 400 500 600 700 800 900
Independent Variable 1
100 200 300 400 500 600 700 800 900
220
Branch At Risk Score
Statistical Models

Fig. 2

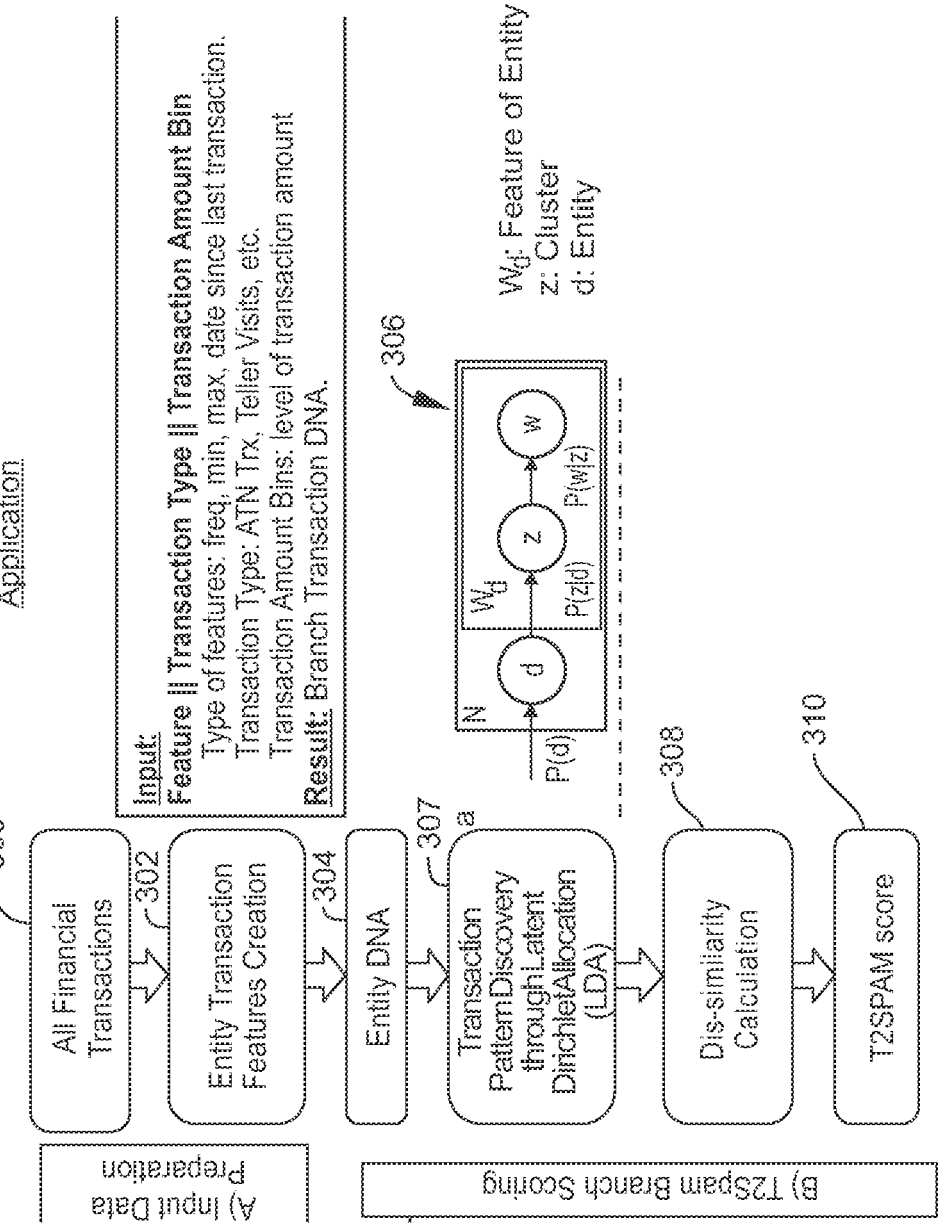

Generic Methodology
Application
300
301
302
304
307
306
308
310
305
A) Input Data Preparation
All Financial Transactions
Entity Transaction Features Creation
Entity DNA
Transaction PatternDiscovery throughLatent DirichletAllocation (LDA)
Dis-similarity Calculation
T2SPAM score
B) T2Spam Branch Scoring
Input:
Feature || Transaction Type || Transaction Amount Bin
Type of features: freq, min, max, date since last transaction.
Transaction Type: ATN Trx, Teller Visits, etc.
Transaction Amount Bins: level of transaction amount
Result: Branch Transaction DNA.
N
d
z
w
P(d)
P(z|d)
P(w|z)
W_d
W_d: Feature of Entity
z: Cluster
d: Entity
Multi-dimensional variable attributes, transaction type, and transaction amount group.
Unsupervised hidden topic discovery process.
Distance calculation based on normalized latent class.
The mean (equivalent to the sum) of the dissimilarities of the observation to their closest medoid is used as a measure of the quality of the clustering.

Fig. 3

Application

Feature Description	# Feature
Frequency of transaction amount (10 bins) & transaction types (153)	1,530
Frequency of latency periods (31) by transaction amount (10)	310
Frequency of latency periods (31) by transaction type (153)	4,743
Frequency of latency periods	31
Frequency of transaction on each day	30
Frequency average transaction amount (10) by type (153)	1,530
Frequency transaction variance (10) by type (153)	1,530
Frequency transaction variance (10) by type (153)	1,530
Frequency min amount (10) by type (153)	1,530
Total Features: Branch Level	**12,764**

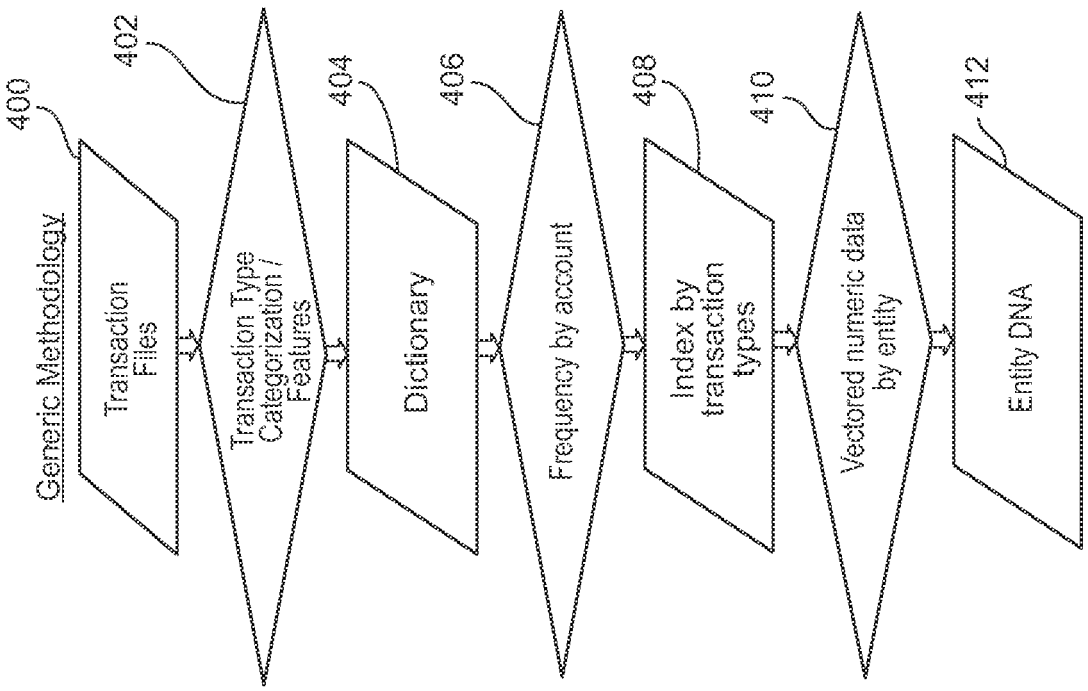


Fig. 4

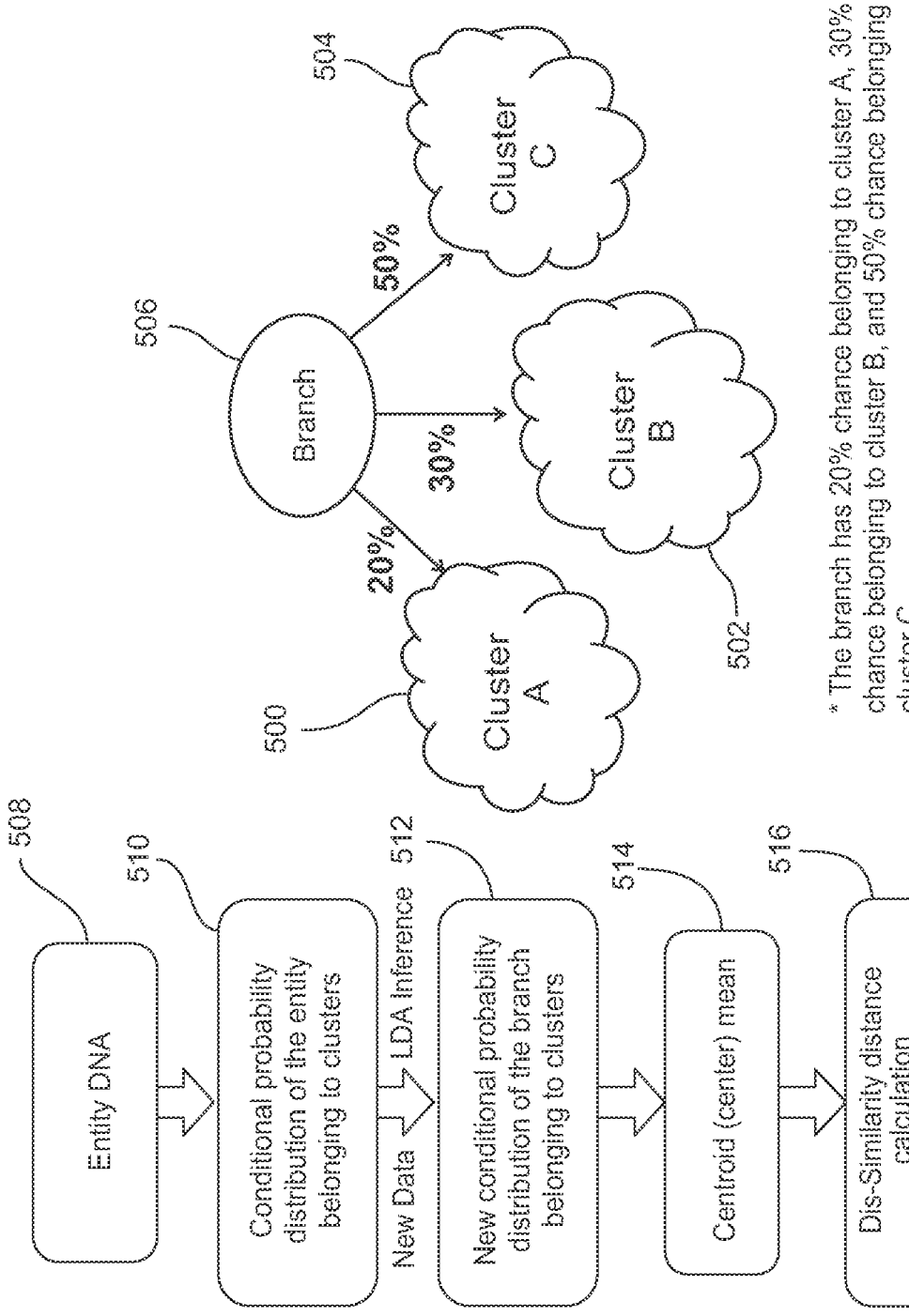

Entity DNA
Conditional probability distribution of the entity belonging to clusters
LDA Inference
New Data
New conditional probability distribution of the branch belonging to clusters
Centroid (center) mean
Dis-Similarity distance calculation
T2spam score
Multiple probability distribution
Using text mining approach for data mining
Identify the nearest cluster.
Calculate the distance between the entity and its cluster's centroid.
508
510
512
514
516
Branch
506
20%
30%
50%
Cluster A
Cluster B
Cluster C
500
502
504
* The branch has 20% chance belonging to cluster A, 30% chance belonging to cluster B, and 50% chance belonging to cluster C.

Fig. 5

Generic Methodology
Dependent Variables
Independent Variables
Peer Groups
Predicted Errors by Peer Groups
604
606
600
602
Application
Independent Variables
Total Checking
Total Liability Balances
Total Assets
Teller Transactions
Teller FTEs
Total Headcount
ATM Count
T2spam
Branch Manager Tenure
Asst. Branch Mgr Tenure
Business Banker Tenure
Personal Banker Tenure
Universal Banker Tenure
Teller Tenure
...
Dependent Variables
Overdraft Losses (62)
Controllable Fidelity Losses (CFL)
Total Branch Losses (99)
Del_62
Del_CFL
Del_99
Example of Peer Group Generation:
a.g1 <- subset(a, Total_Checking_Accts>=2872.5 & Teller_FTE>=4.725)
a.g2 <- subset(a, Total_Checking_Accts>=2872.5 & Teller_FTE<=4.725
& abs(Total_Liabilities_MM)>=99.39046)
.....

Fig. 6

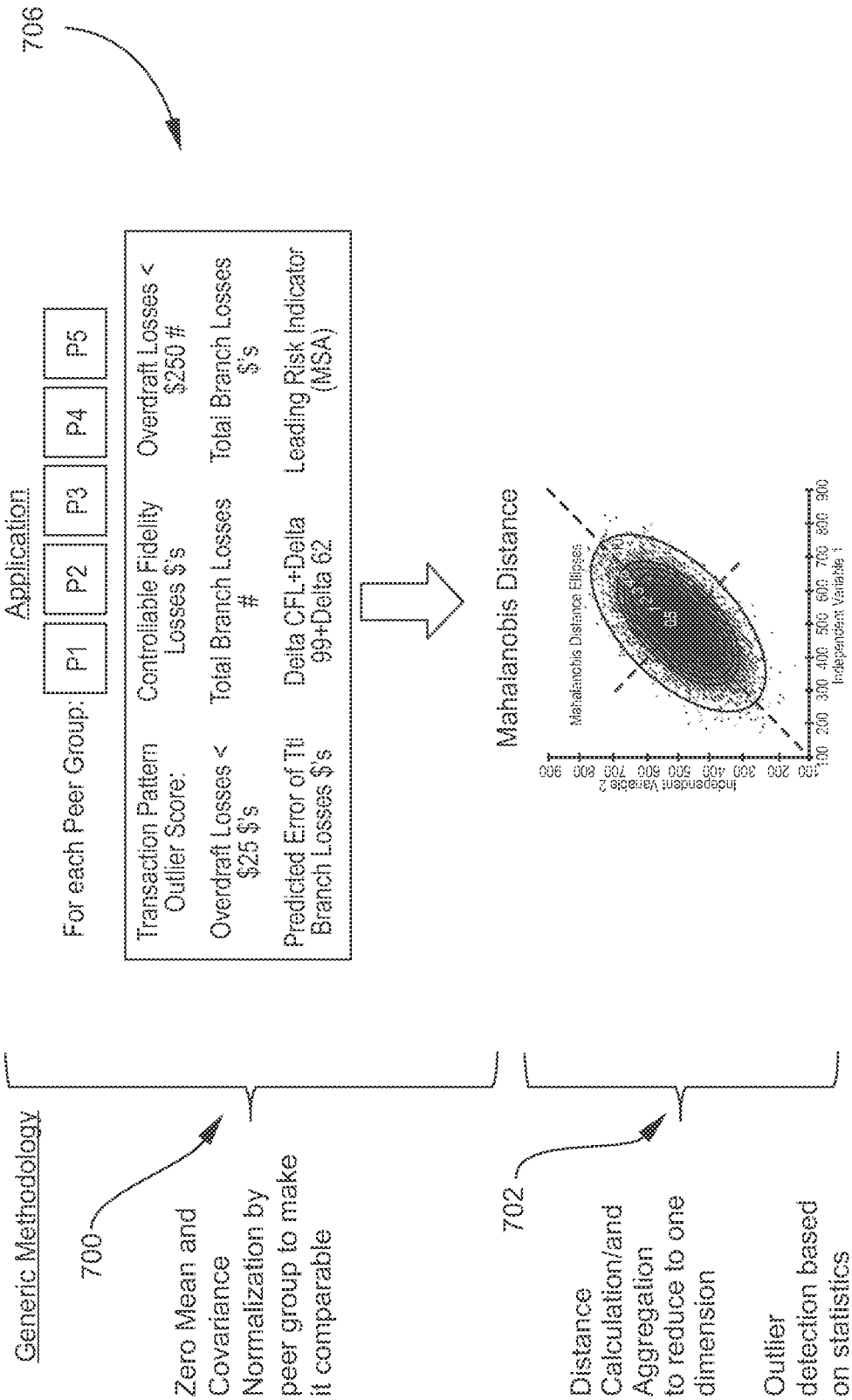

706
Application
For each Peer Group:
P1
P2
P3
P4
P5
Transaction Pattern Outlier Score:
Controllable Fidelity Losses $'s
Overdraft Losses < $250 #
Overdraft Losses < $25 $'s
Total Branch Losses #
Total Branch Losses $'s
Predicted Error of Tti Branch Losses $'s
Delta CFL+Delta 99+Delta 62
Leading Risk Indicator (MSA)
Mahalanobis Distance
Mahalanobis Distance Ellipses
Independent Variable 2
Independent Variable 1
Generic Methodology
700
Zero Mean and Covariance Normalization by peer group to make it comparable
702
Distance Calculation/and Aggregation to reduce to one dimension
Outlier detection based on statistics

Fig. 7

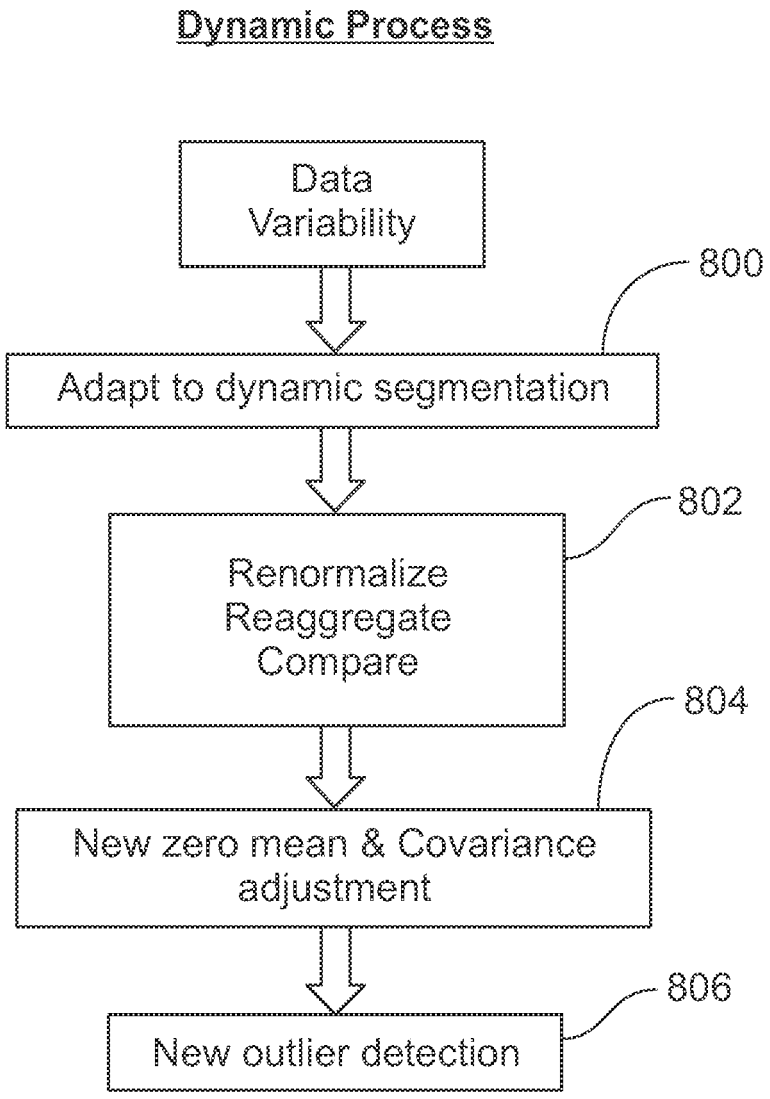

Dynamic Process
Data Variability
800
Adapt to dynamic segmentation
802
Renormalize
Reaggregate
Compare
804
New zero mean & Covariance adjustment
806
New outlier detection

Fig. 8

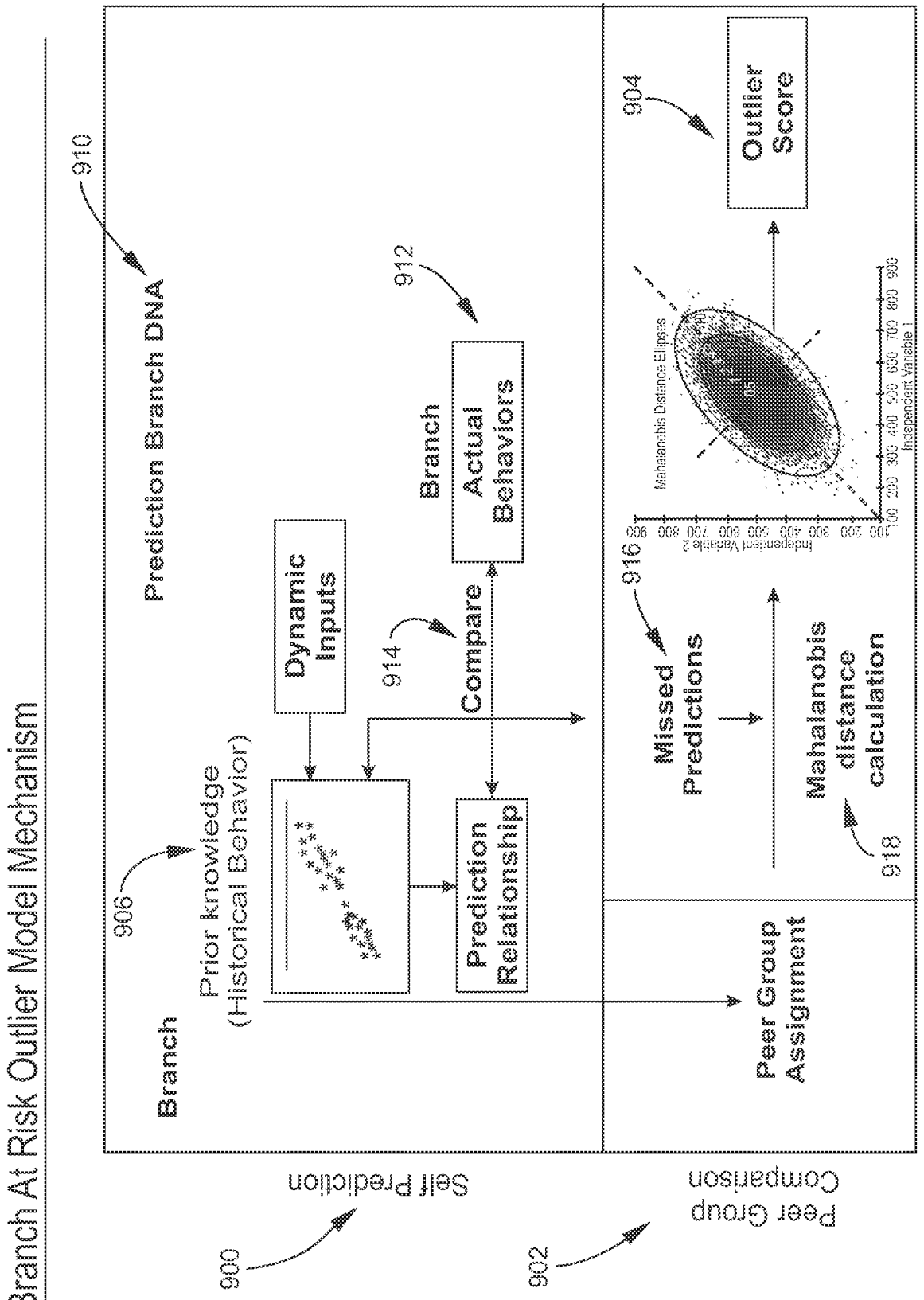

Branch At Risk Outlier Model Mechanism
900
Self Prediction
902
Peer Group Comparison
910
Prediction Branch DNA
Branch
906
Prior knowledge (Historical Behavior)
Dynamic Inputs
914
Compare
912
Branch Actual Behaviors
Prediction Relationship
Entity Signature
Peer Group Assignment
916
Missed Predictions
Mahalanobis distance calculation
918
Outlier Determination
904
Outlier Score
Mahalanobis Distance Ellipses
Independent Variable 2
Independent Variable 1

Fig. 9

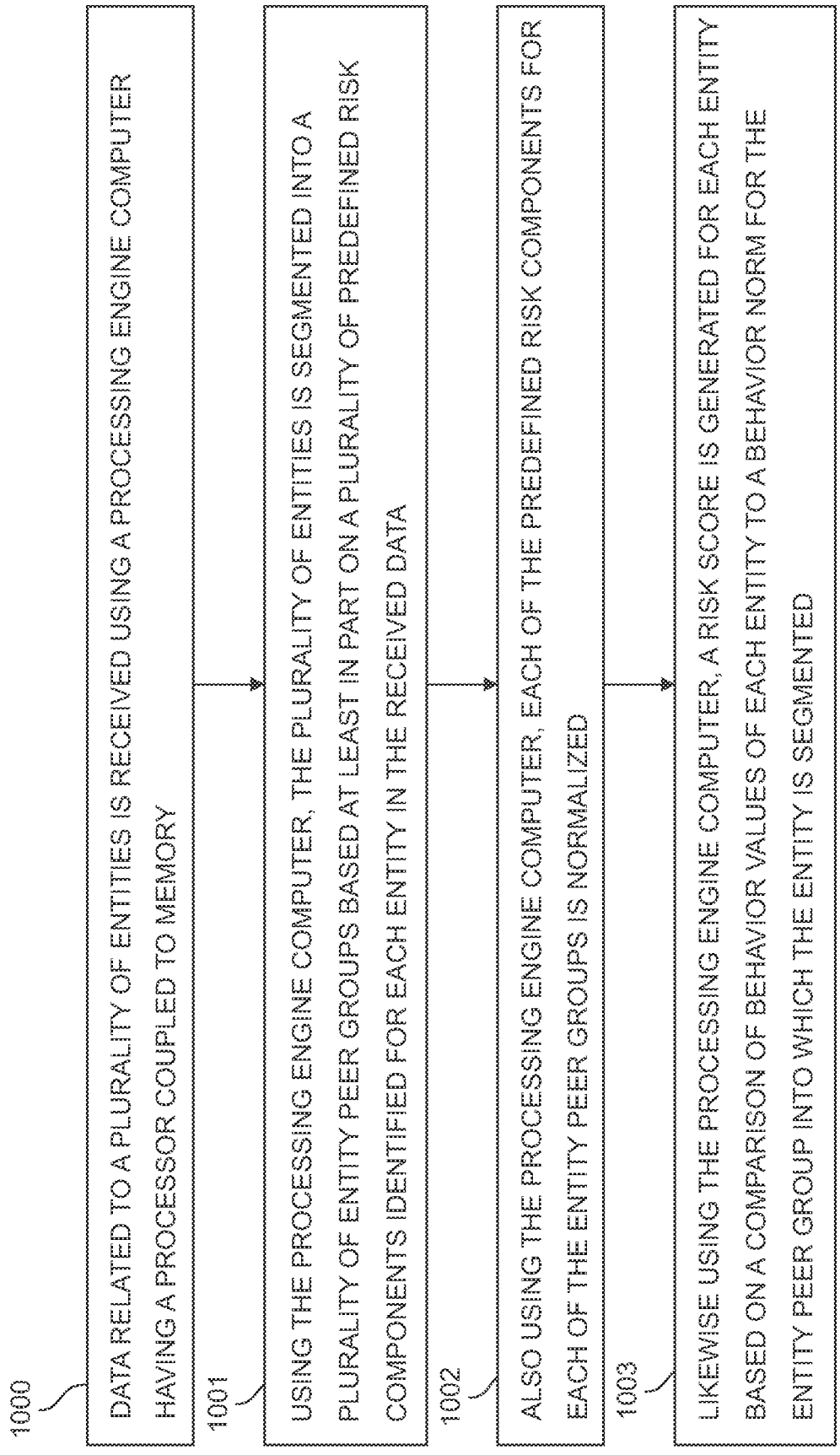

1000
DATA RELATED TO A PLURALITY OF ENTITIES IS RECEIVED USING A PROCESSING ENGINE COMPUTER HAVING A PROCESSOR COUPLED TO MEMORY
1001
USING THE PROCESSING ENGINE COMPUTER, THE PLURALITY OF ENTITIES IS SEGMENTED INTO A PLURALITY OF ENTITY PEER GROUPS BASED AT LEAST IN PART ON A PLURALITY OF PREDEFINED RISK COMPONENTS IDENTIFIED FOR EACH ENTITY IN THE RECEIVED DATA
1002
ALSO USING THE PROCESSING ENGINE COMPUTER, EACH OF THE PREDEFINED RISK COMPONENTS FOR EACH OF THE ENTITY PEER GROUPS IS NORMALIZED
1003
LIKEWISE USING THE PROCESSING ENGINE COMPUTER, A RISK SCORE IS GENERATED FOR EACH ENTITY BASED ON A COMPARISON OF BEHAVIOR VALUES OF EACH ENTITY TO A BEHAVIOR NORM FOR THE ENTITY PEER GROUP INTO WHICH THE ENTITY IS SEGMENTED

Fig. 10