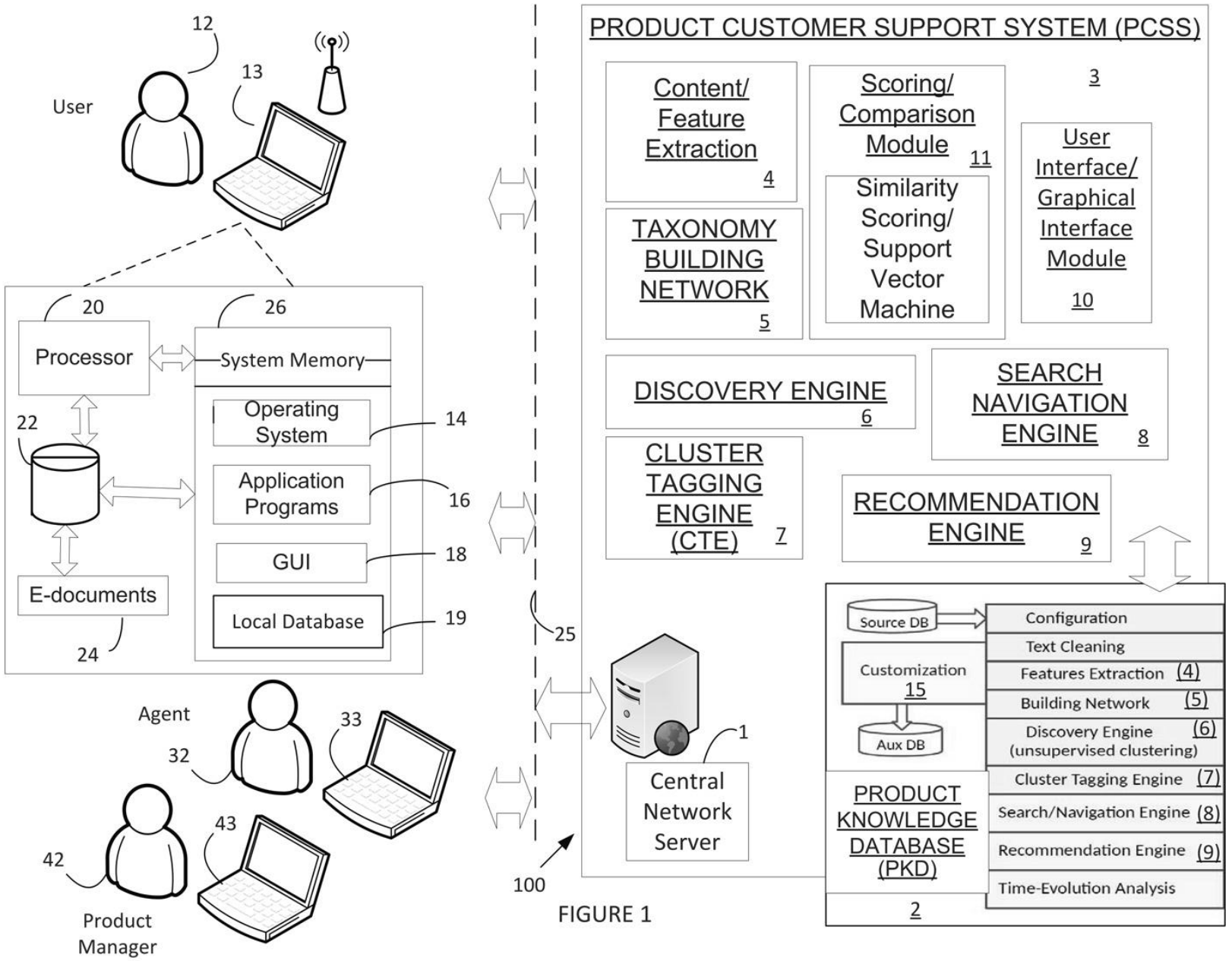
User
12
13
20
26
Processor
System Memory
Operating System
14
22
Application Programs
16
GUI
18
E-documents
Local Database
19
24
Agent
32
33
42
43
Product Manager
25
1
Central Network Server
100
PRODUCT CUSTOMER SUPPORT SYSTEM (PCSS)
3
Content/ Feature Extraction 4
Scoring/ Comparison Module 11
Similarity Scoring/ Support Vector Machine
User Interface/ Graphical Interface Module 10
TAXONOMY BUILDING NETWORK 5
DISCOVERY ENGINE 6
SEARCH NAVIGATION ENGINE 8
CLUSTER TAGGING ENGINE (CTE) 7
RECOMMENDATION ENGINE 9
Source DB
Customization 15
Aux DB
Configuration
Text Cleaning
Features Extraction (4)
Building Network (5)
Discovery Engine (6) (unsupervised clustering)
Cluster Tagging Engine (7)
Search/Navigation Engine (8)
Recommendation Engine (9)
Time-Evolution Analysis
PRODUCT KNOWLEDGE DATABASE (PKD)
2

FIGURE 1

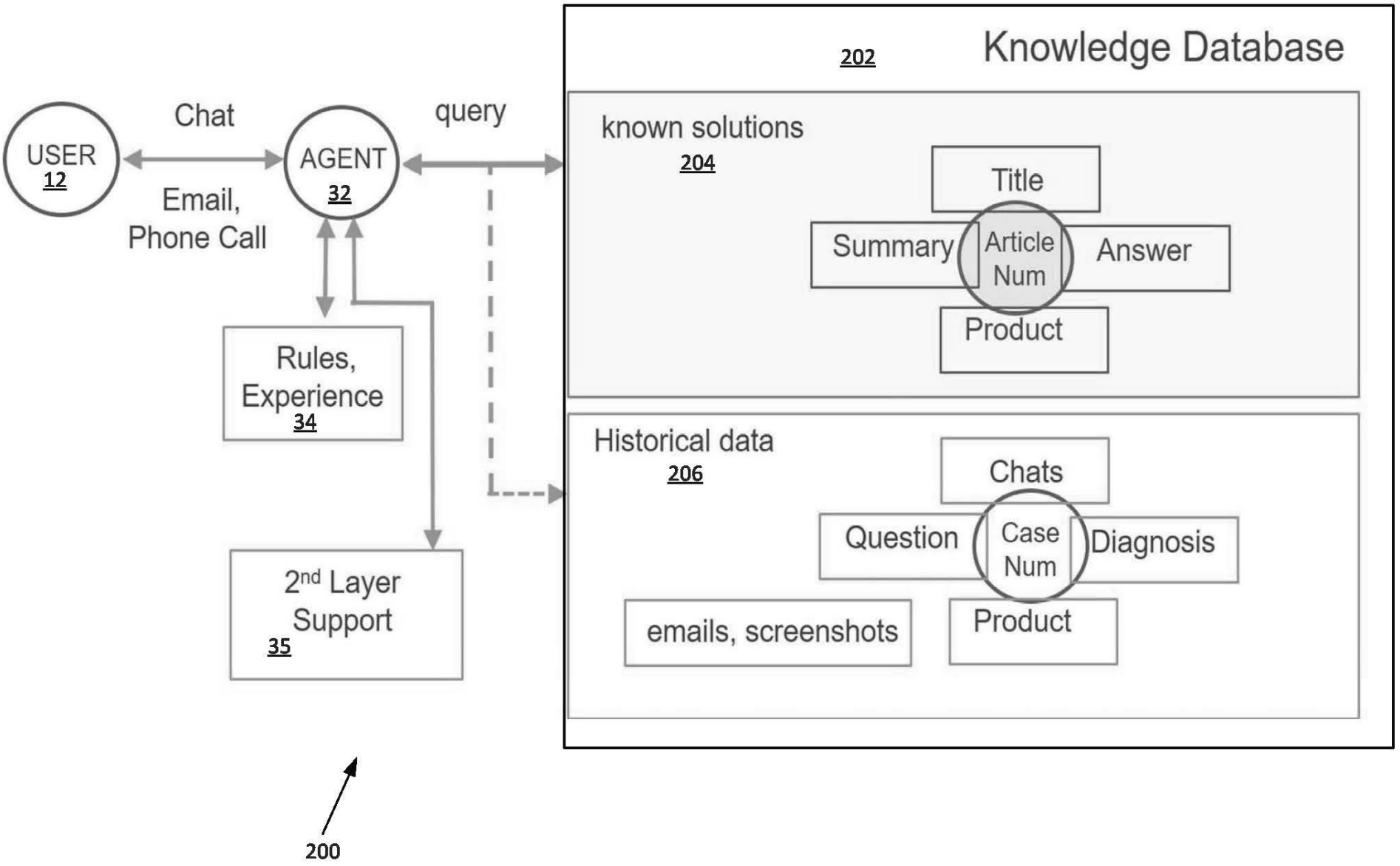
USER
12
Chat
Email,
Phone Call
AGENT
32
query
Rules,
Experience
34
2nd Layer
Support
35
200
202
Knowledge Database
known solutions
204
Title
Summary
Article
Num
Answer
Product
Historical data
206
Chats
Question
Case
Num
Diagnosis
emails, screenshots
Product

FIGURE 2

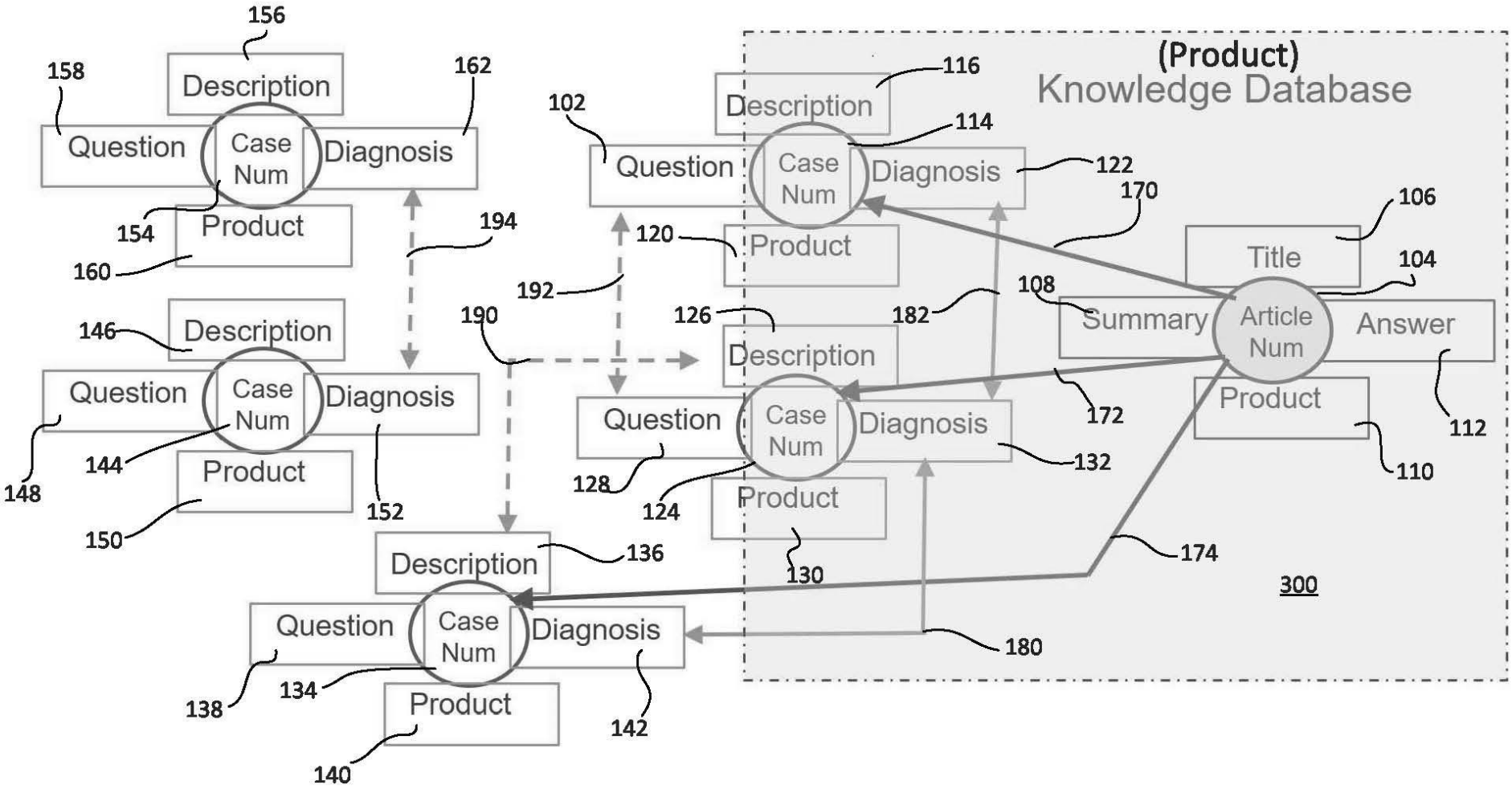
(Product)
Knowledge Database
Description
Question
Case Num
Diagnosis
Product
Title
Summary
Article Num
Answer
300
156
158
162
154
160
194
146
144
148
152
150
102
116
114
122
170
106
104
120
192
190
126
182
108
172
112
110
128
124
132
130
174
136
134
138
142
140
180

FIGURE 3

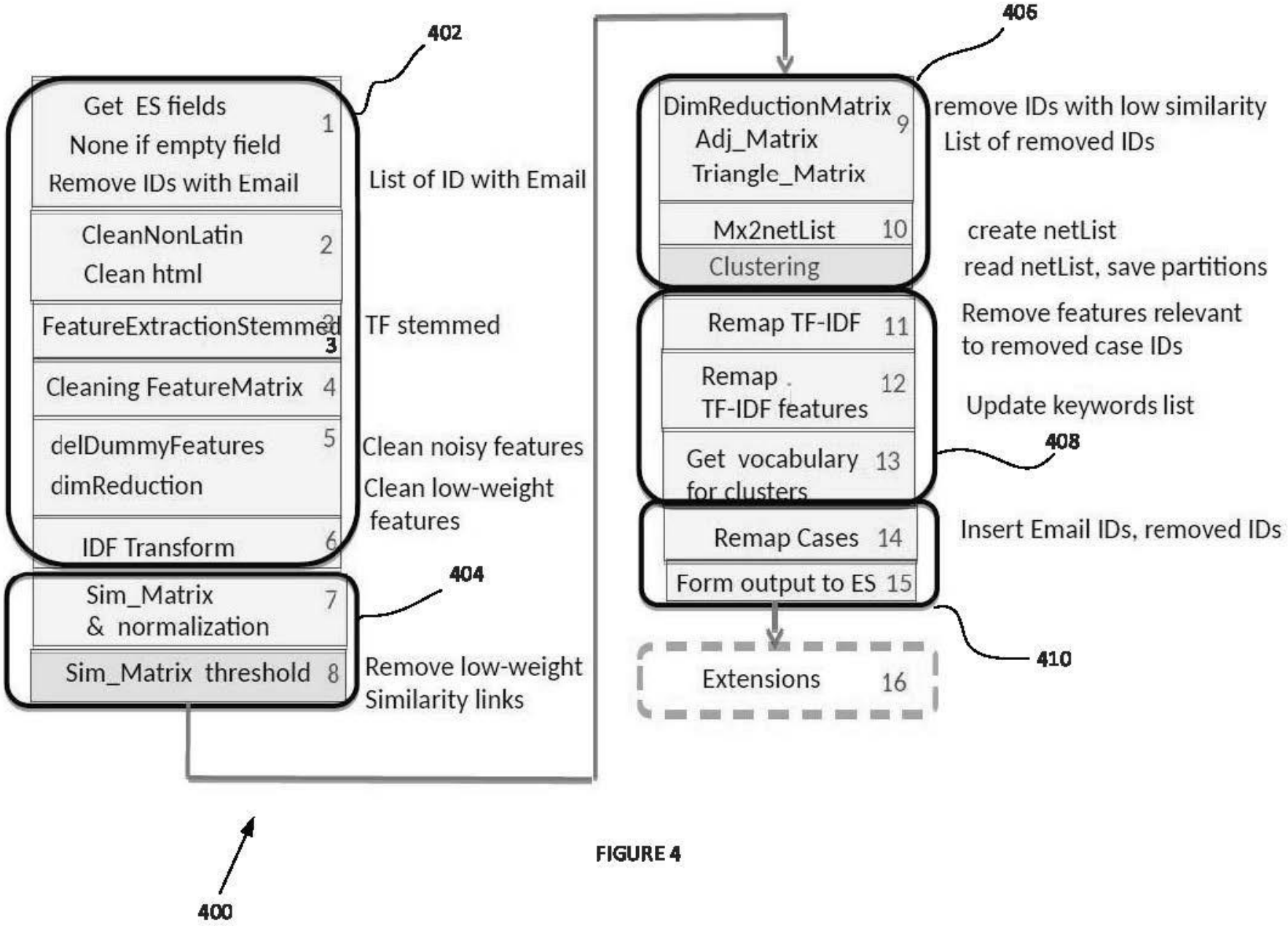
402
Get ES fields
None if empty field
Remove IDs with Email
1
List of ID with Email
CleanNonLatin
Clean html
2
FeatureExtractionStemmed
3
TF stemmed
Cleaning FeatureMatrix 4
delDummyFeatures
dimReduction
5
Clean noisy features
Clean low-weight features
IDF Transform 6
404
Sim_Matrix
& normalization
7
Sim_Matrix threshold 8
Remove low-weight Similarity links
406
DimReductionMatrix
Adj_Matrix
Triangle_Matrix
9
remove IDs with low similarity
List of removed IDs
Mx2netList 10
Clustering
create netList
read netList, save partitions
Remap TF-IDF 11
Remap TF-IDF features 12
Remove features relevant to removed case IDs
Update keywords list
Get vocabulary for clusters 13
408
Remap Cases 14
Form output to ES 15
Insert Email IDs, removed IDs
410
Extensions 16
400

FIGURE 4

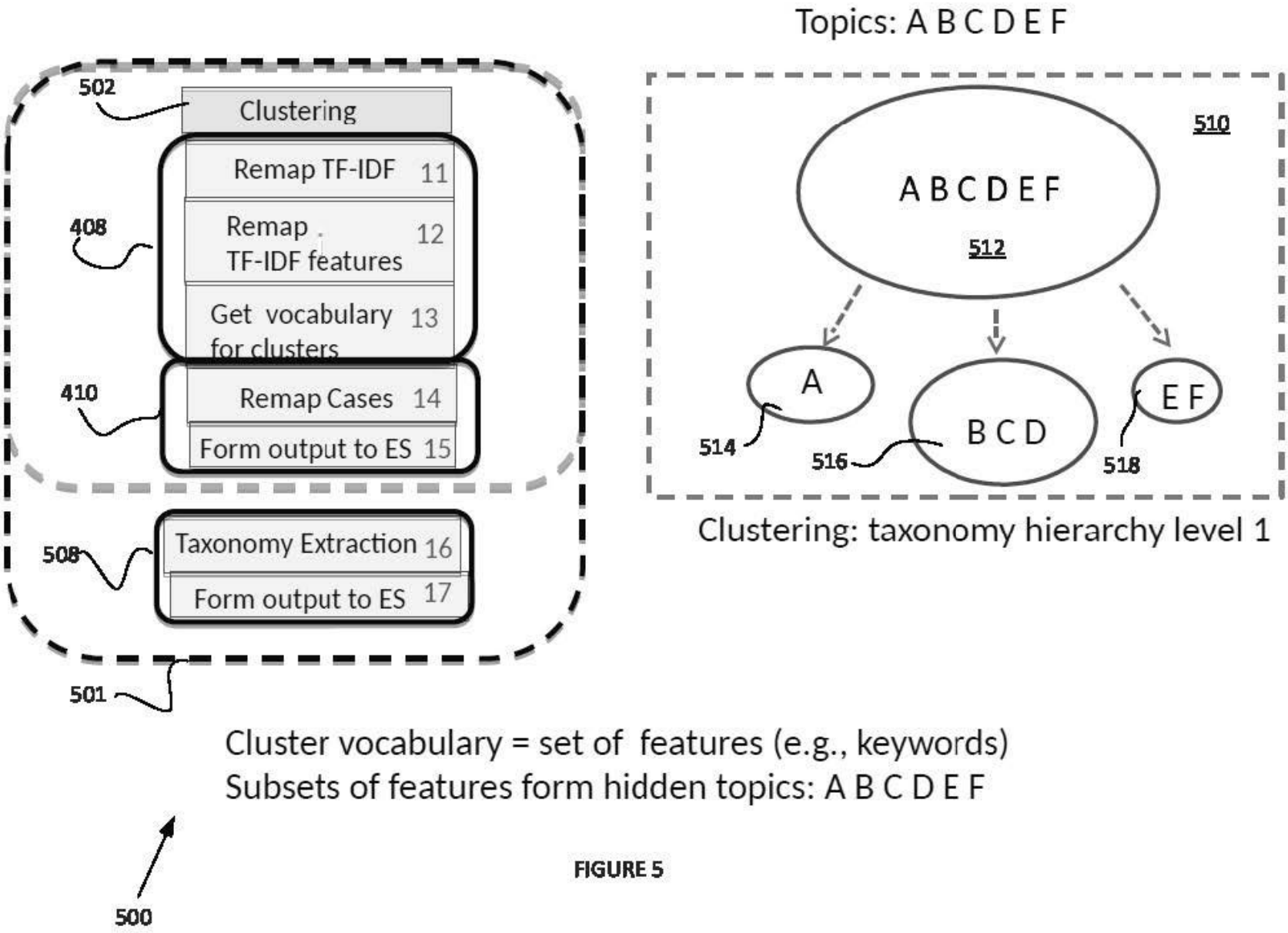
Topics: A B C D E F
502
Clustering
Remap TF-IDF 11
408
Remap TF-IDF features 12
Get vocabulary for clusters 13
410
Remap Cases 14
Form output to ES 15
508
Taxonomy Extraction 16
Form output to ES 17
501
510
A B C D E F
512
A
B C D
E F
514
516
518
Clustering: taxonomy hierarchy level 1
Cluster vocabulary = set of features (e.g., keywords)
Subsets of features form hidden topics: A B C D E F
500

FIGURE 5

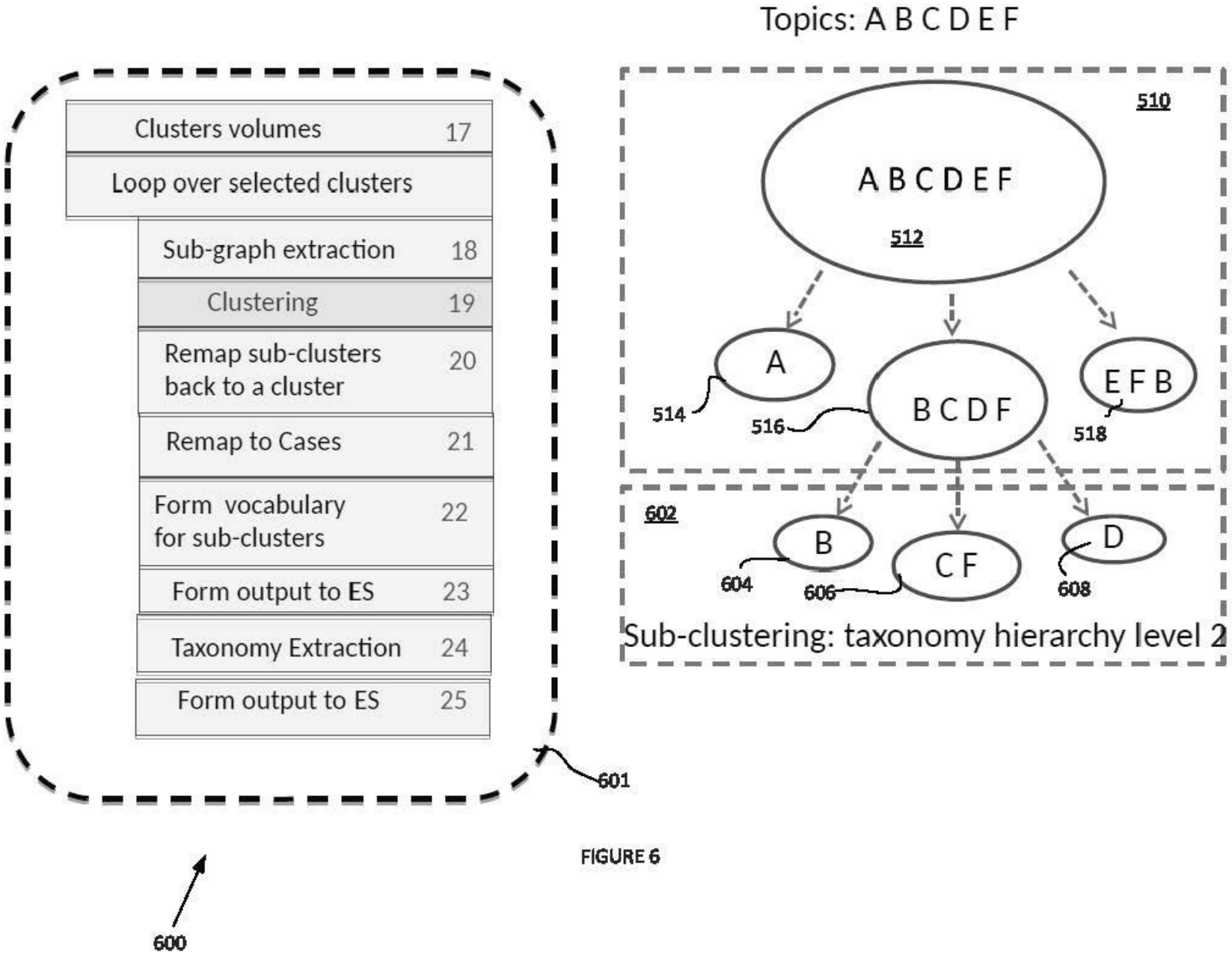
Clusters volumes 17
Loop over selected clusters
Sub-graph extraction 18
Clustering 19
Remap sub-clusters back to a cluster 20
Remap to Cases 21
Form vocabulary for sub-clusters 22
Form output to ES 23
Taxonomy Extraction 24
Form output to ES 25
601
600
Topics: A B C D E F
510
A B C D E F
512
A
514
B C D F
516
E F B
518
602
B
604
C F
606
D
608
Sub-clustering: taxonomy hierarchy level 2

FIGURE 6

Customer Support System
702
Product
704
Product 1
Product 2
Product 3
Product 4
Symptoms
706
Application Issue 710
Connectivity 712
HW Failure 714
Content Exploration 716
How to use 718
Diagnosis
708
Application/ Desktop/ Malfunction 720
Application/ Desktop/ Func.N.A. 722
Application/ Desktop/ Crash 724
Application/ Desktop/ How to use 726
Application/ Infrastructure/ Crash 728
Application/ Infrastructure/ Malfunction 730
Application Infrastructure/ Platform 732
Connectivity/ Networking/ Platform 734
Connectivity/ Networking/ Source 736
Content Expansion 738
Performance/ Networking/ Platform 740
Performance/ Networking/ Cust.Manag 742
700

FIGURE 7

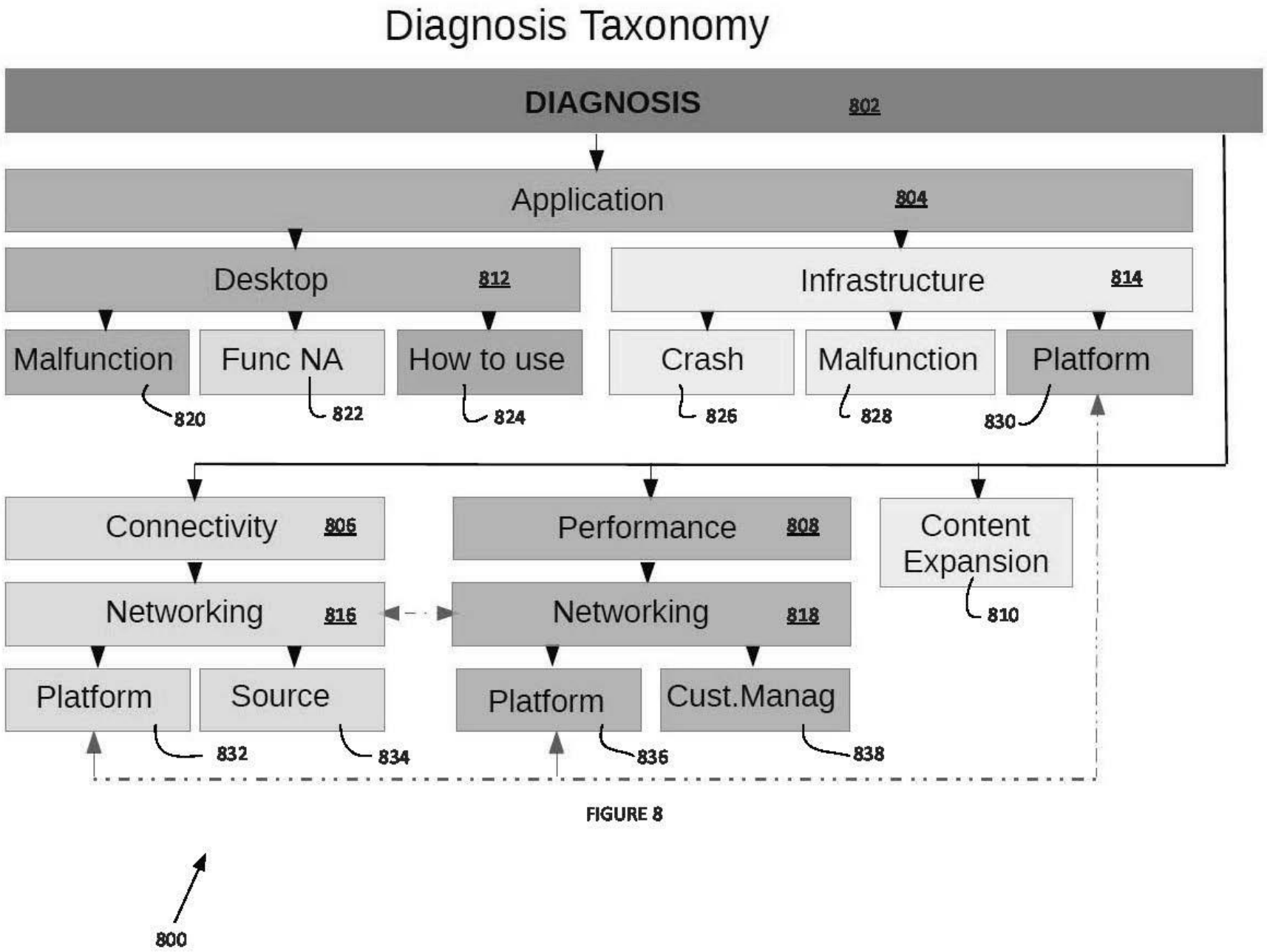
Diagnosis Taxonomy
DIAGNOSIS
802
Application
804
Desktop
812
Infrastructure
814
Malfunction
Func NA
How to use
Crash
Malfunction
Platform
820
822
824
826
828
830
Connectivity
806
Performance
808
Content Expansion
810
Networking
816
Networking
818
Platform
Source
Platform
Cust.Manag
832
834
836
838
800

FIGURE 8

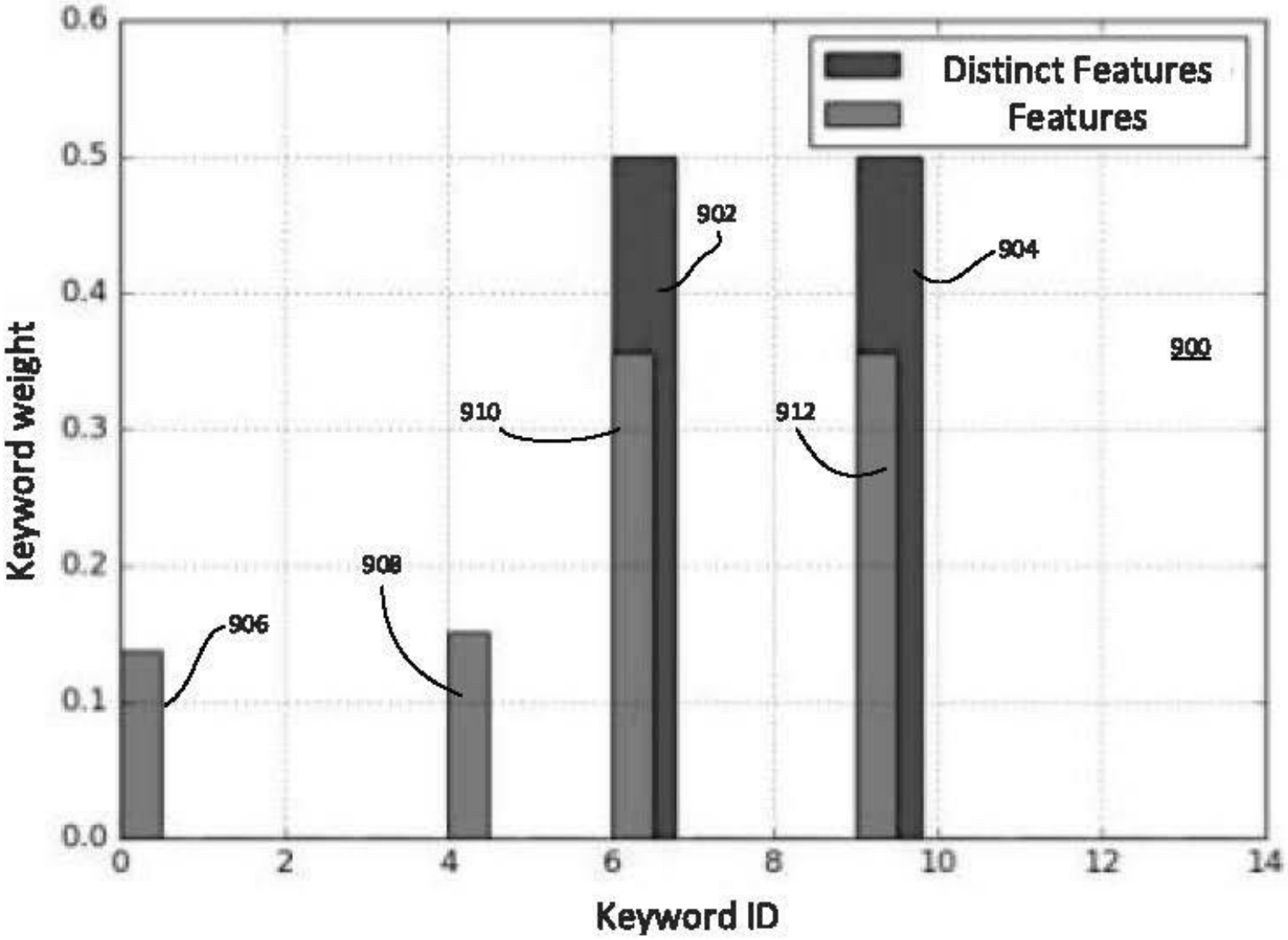
Distinct Features
Features
902
904
900
910
912
908
906
Keyword weight
Keyword ID
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0
2
4
6
8
10
12
14

FIGURE 9

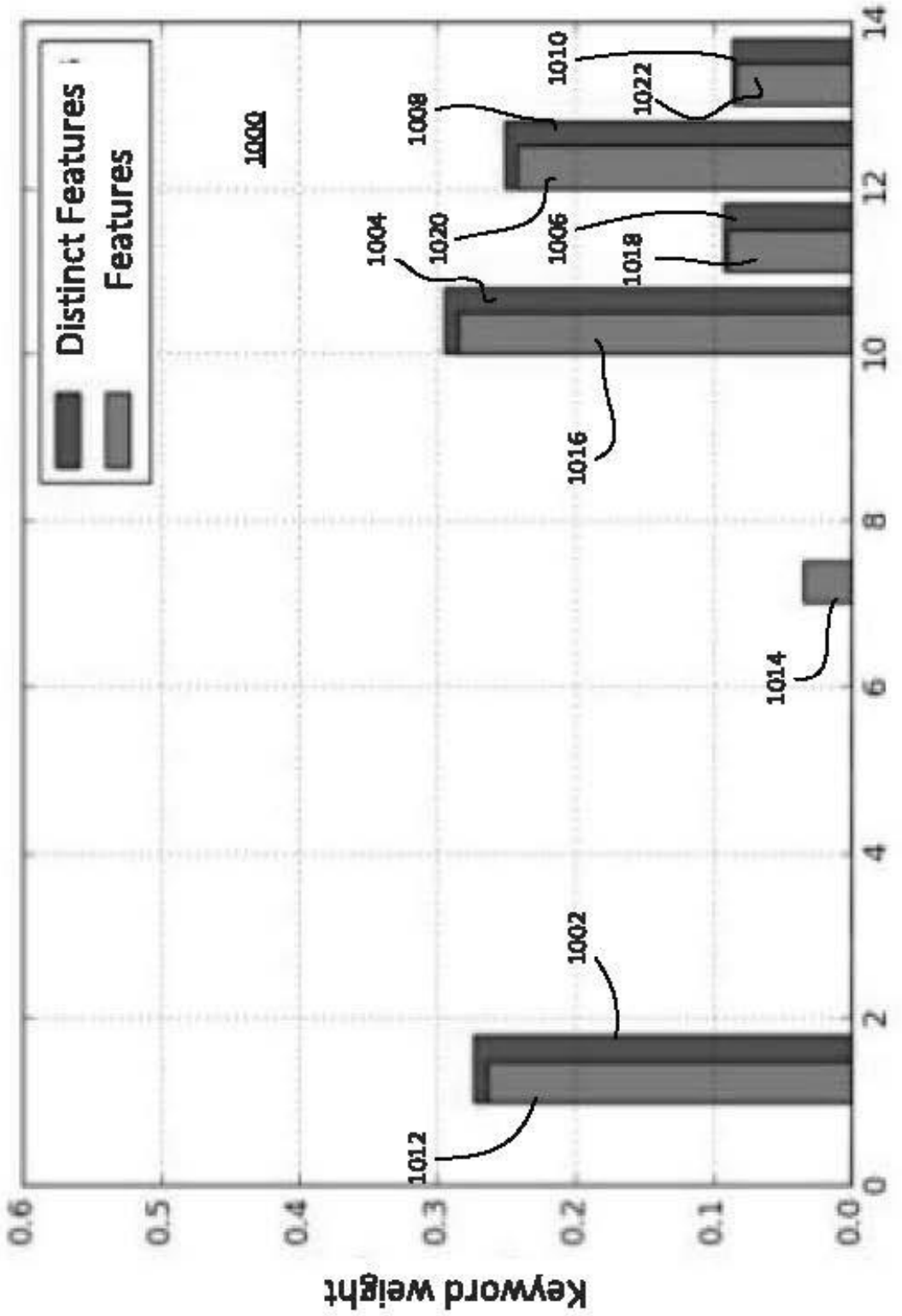
Distinct Features
Features
1000
1002
1004
1006
1008
1010
1012
1014
1016
1018
1020
1022
Keyword weight
Keyword ID
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0
2
4
6
8
10
12
14

FIGURE 10

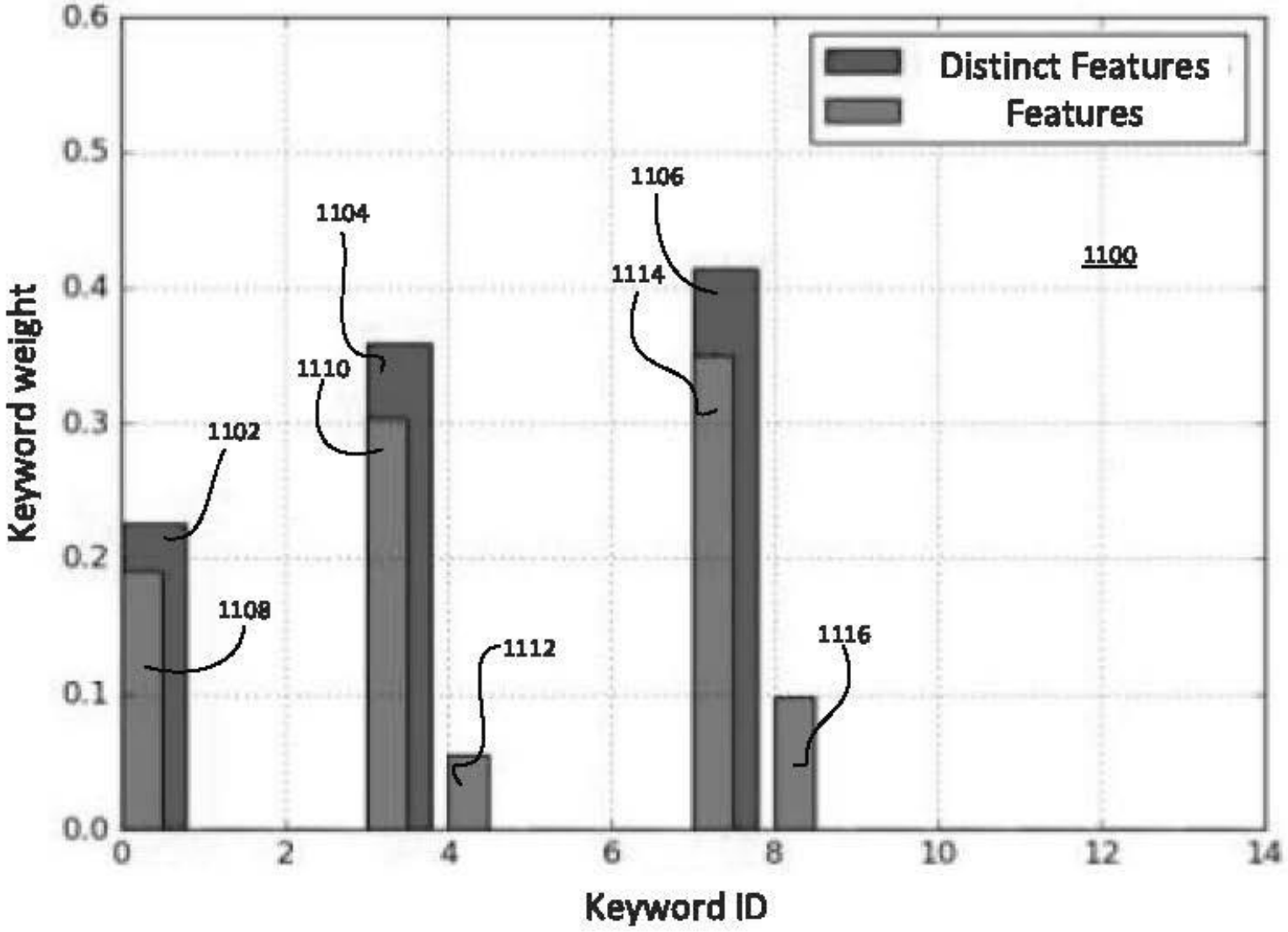
Distinct Features
Features
1100
1102
1104
1106
1108
1110
1112
1114
1116
Keyword weight
Keyword ID
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0
2
4
6
8
10
12
14

FIGURE 11

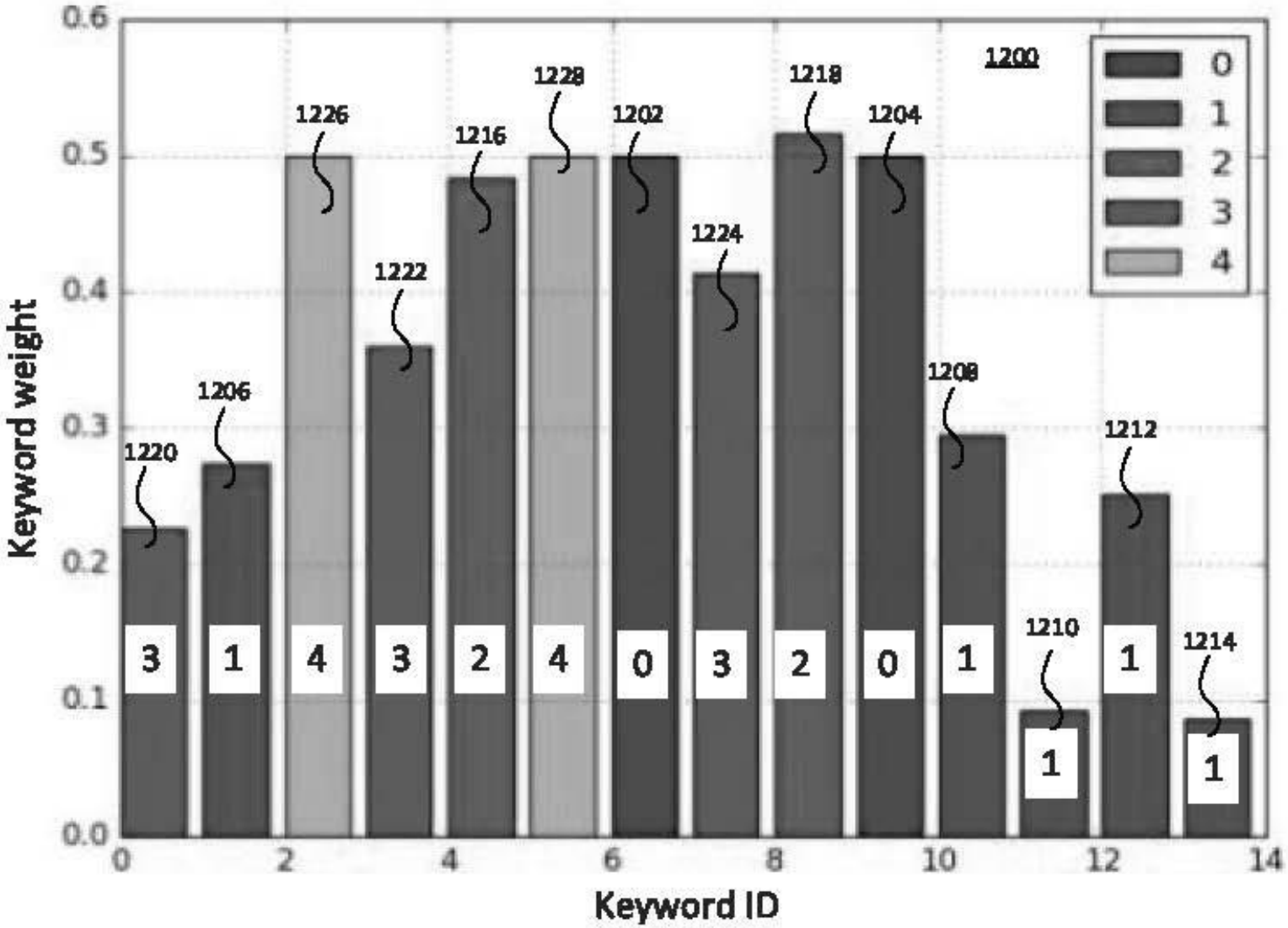
1200
0
1
2
3
4
Keyword weight
Keyword ID
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0
2
4
6
8
10
12
14
1220
1206
1226
1222
1216
1228
1202
1224
1218
1204
1208
1212
1210
1214
3
1
4
3
2
4
0
3
2
0
1
1
1
1

FIGURE 12

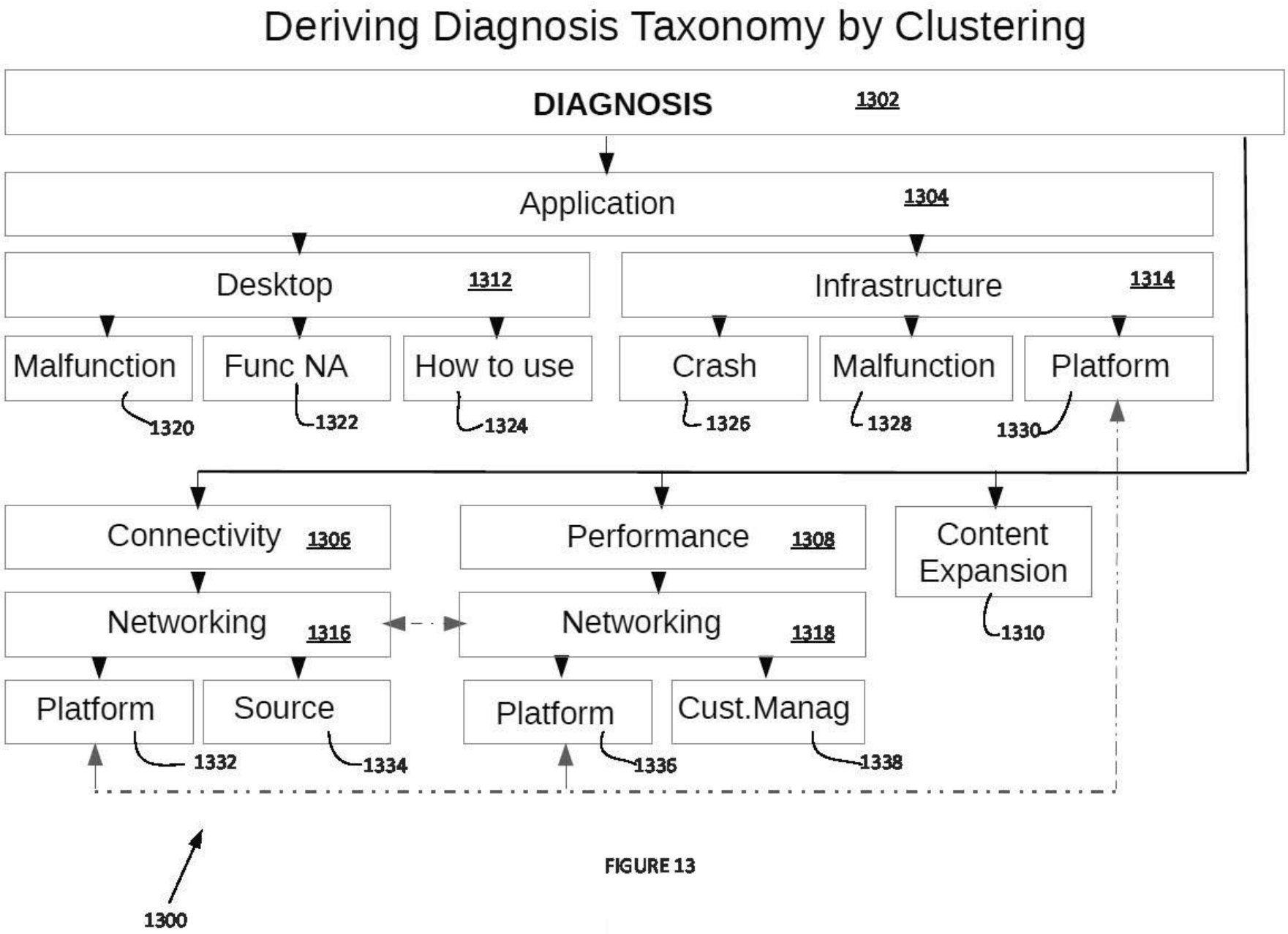
Deriving Diagnosis Taxonomy by Clustering
DIAGNOSIS
1302
Application
1304
Desktop
1312
Infrastructure
1314
Malfunction
Func NA
How to use
Crash
Malfunction
Platform
1320
1322
1324
1326
1328
1330
Connectivity
1306
Performance
1308
Content Expansion
1310
Networking
1316
Networking
1318
Platform
Source
Platform
Cust.Manag
1332
1334
1336
1338
1300

FIGURE 13

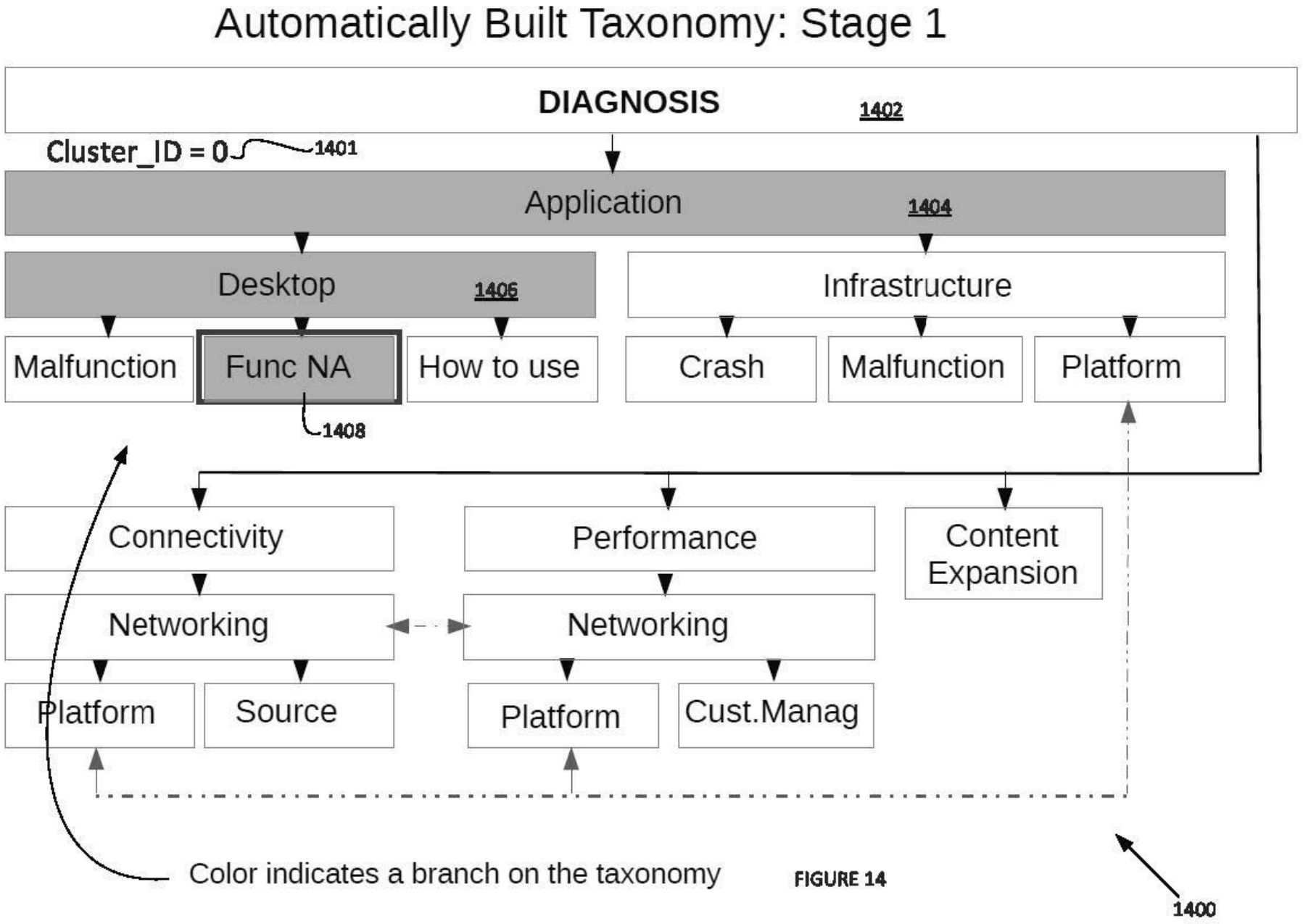
Automatically Built Taxonomy: Stage 1
DIAGNOSIS
1402
Cluster_ID = 0
1401
Application
1404
Desktop
1406
Infrastructure
Malfunction
Func NA
How to use
Crash
Malfunction
Platform
1408
Connectivity
Performance
Content Expansion
Networking
Networking
Platform
Source
Platform
Cust.Manag
Color indicates a branch on the taxonomy
1400

FIGURE 14

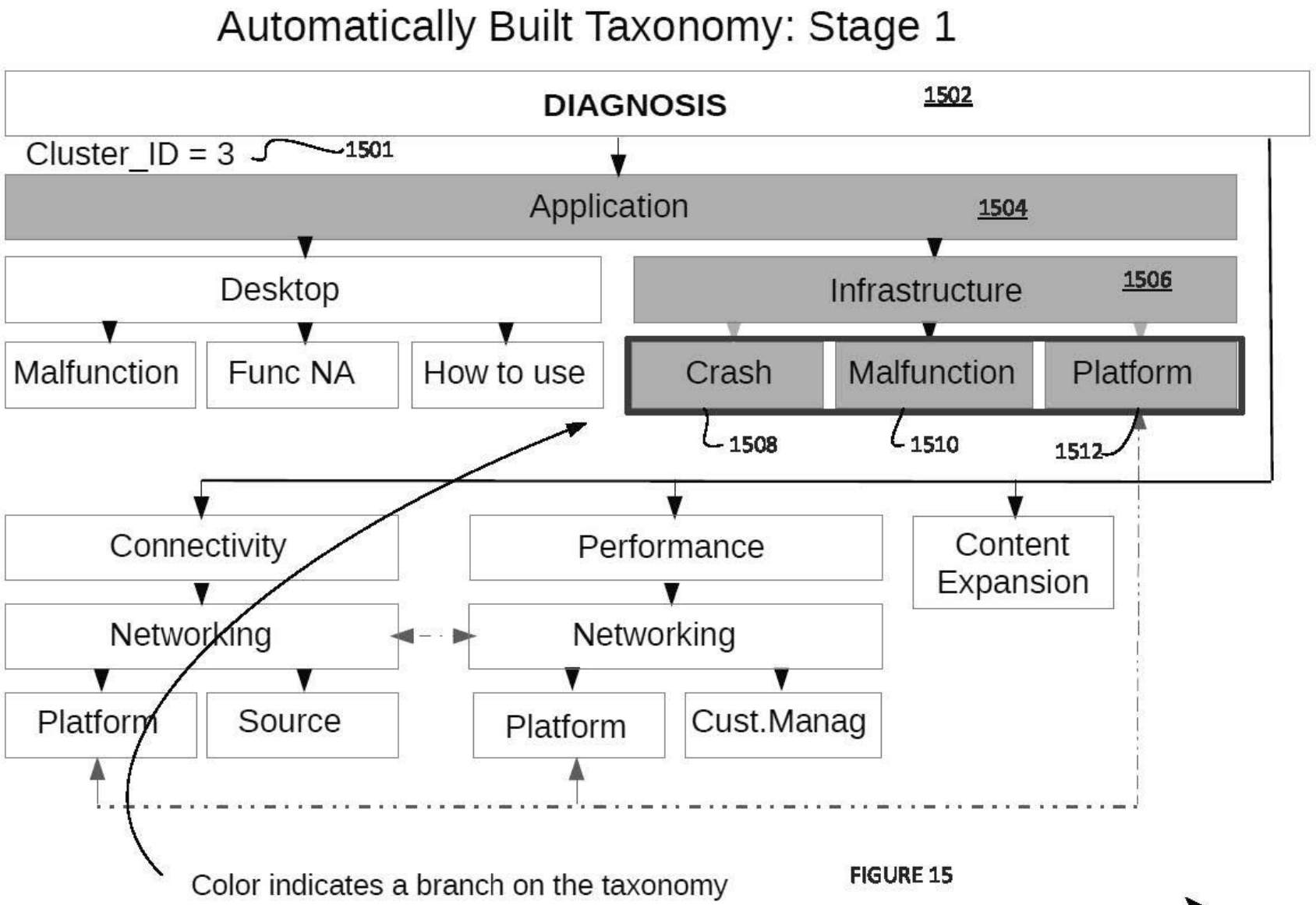
Automatically Built Taxonomy: Stage 1
DIAGNOSIS
1502
Cluster_ID = 3
1501
Application
1504
Desktop
Infrastructure
1506
Malfunction
Func NA
How to use
Crash
Malfunction
Platform
1508
1510
1512
Connectivity
Performance
Content Expansion
Networking
Networking
Platform
Source
Platform
Cust.Manag
Color indicates a branch on the taxonomy
1500

FIGURE 15

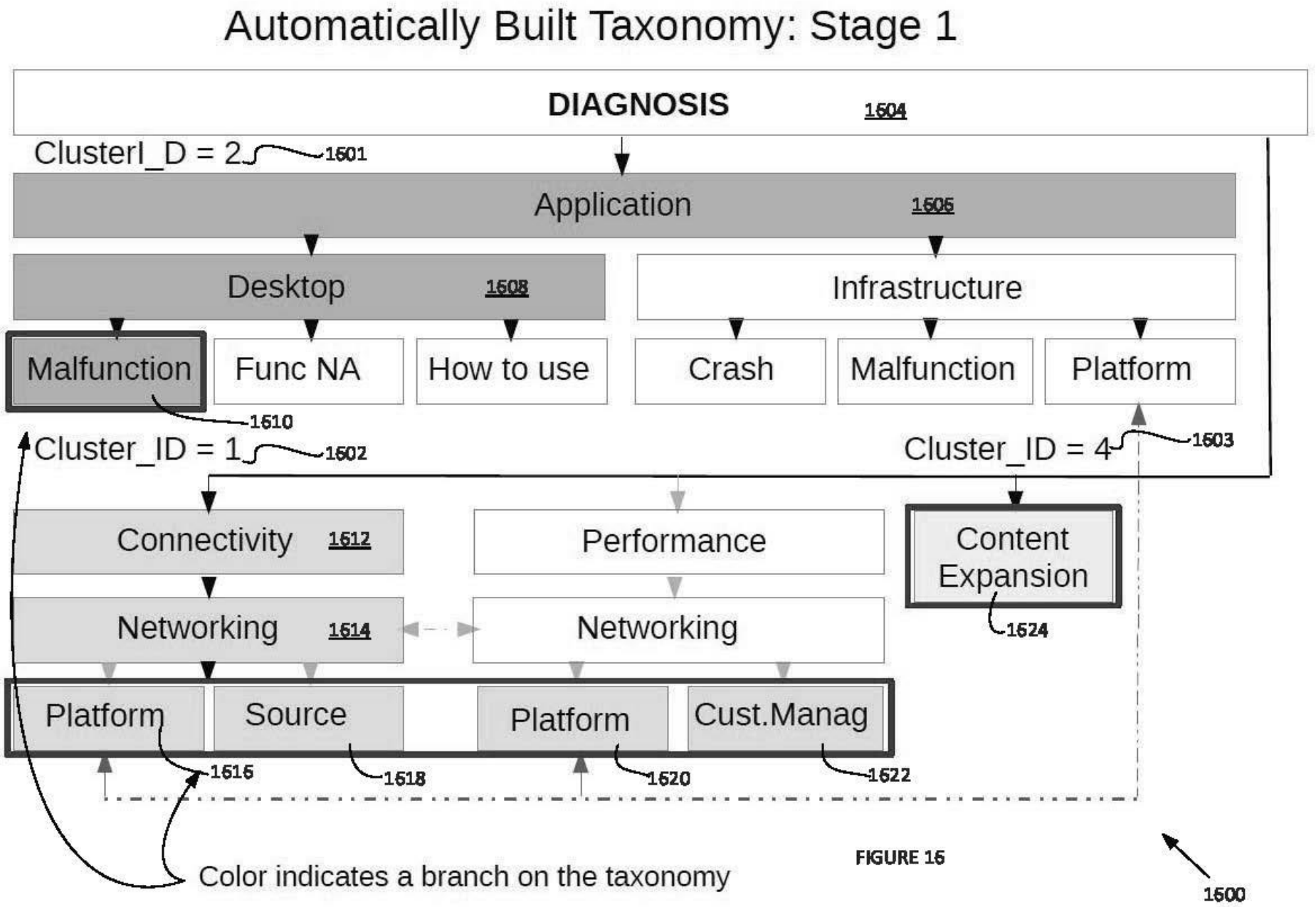
Automatically Built Taxonomy: Stage 1
DIAGNOSIS
1604
ClusterI_D = 2
1601
Application
1606
Desktop
1608
Infrastructure
Malfunction
Func NA
How to use
Crash
Malfunction
Platform
1610
Cluster_ID = 1
1602
Cluster_ID = 4
1603
Connectivity
1612
Performance
Content Expansion
1624
Networking
1614
Networking
Platform
Source
Platform
Cust.Manag
1616
1618
1620
1622
Color indicates a branch on the taxonomy
1600

FIGURE 16

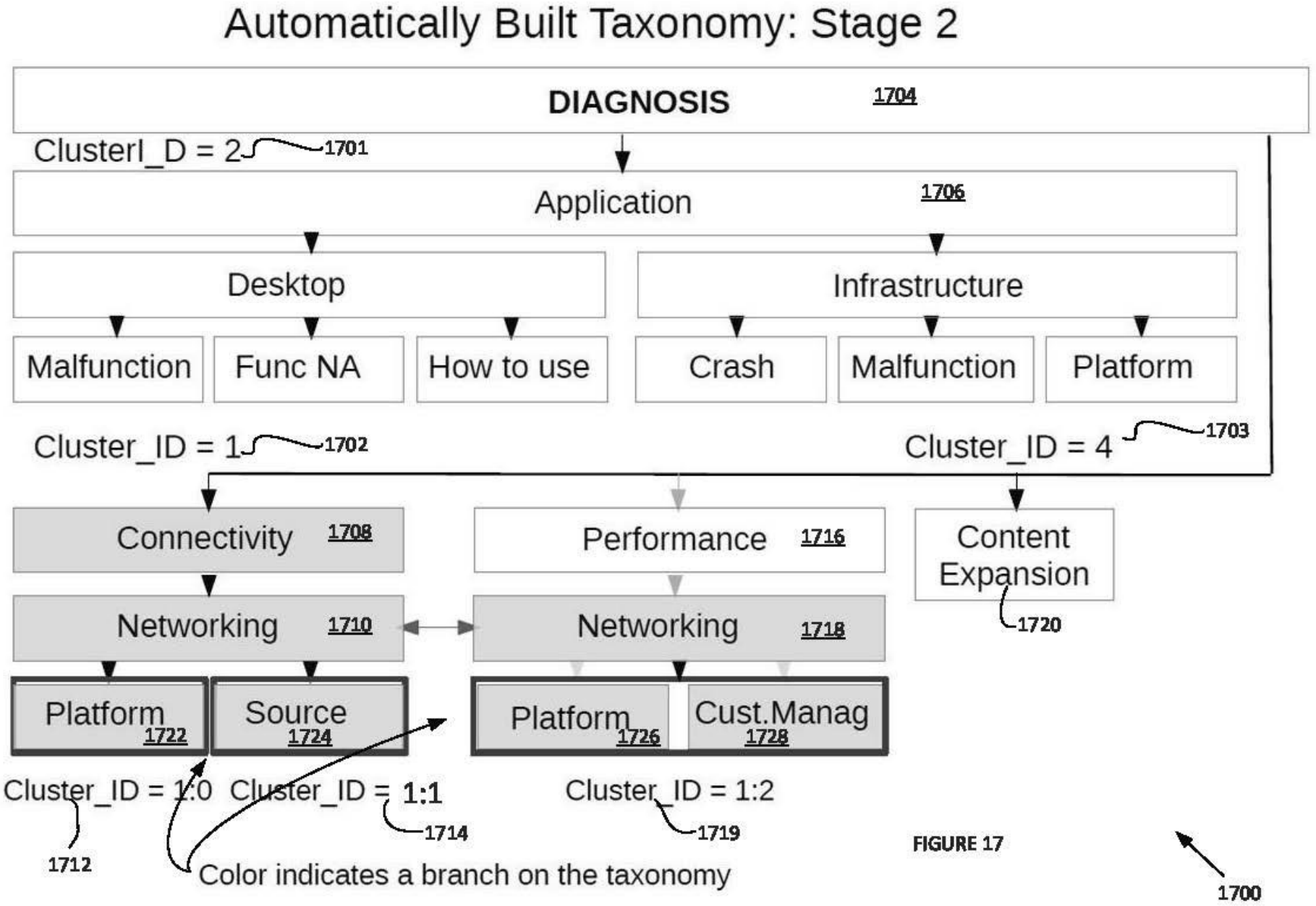

FIGURE 17

1700

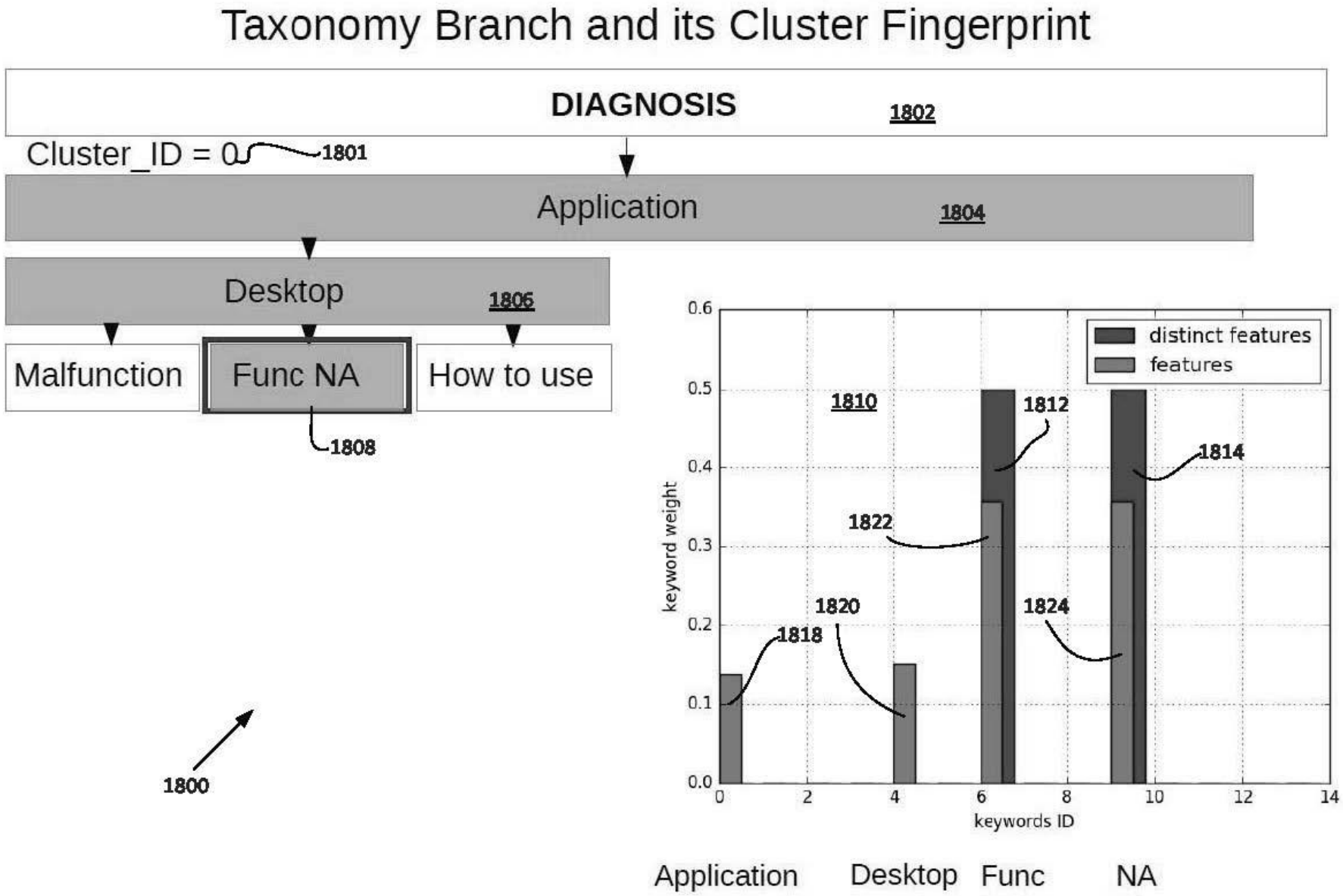
Taxonomy Branch and its Cluster Fingerprint
DIAGNOSIS
1802
Cluster_ID = 0
1801
Application
1804
Desktop
1806
Malfunction
Func NA
How to use
1808
1800
distinct features
features
1810
1812
1814
1822
1820
1824
1818
keyword weight
keywords ID
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0
2
4
6
8
10
12
14
Application
Desktop
Func
NA

FIGURE 18

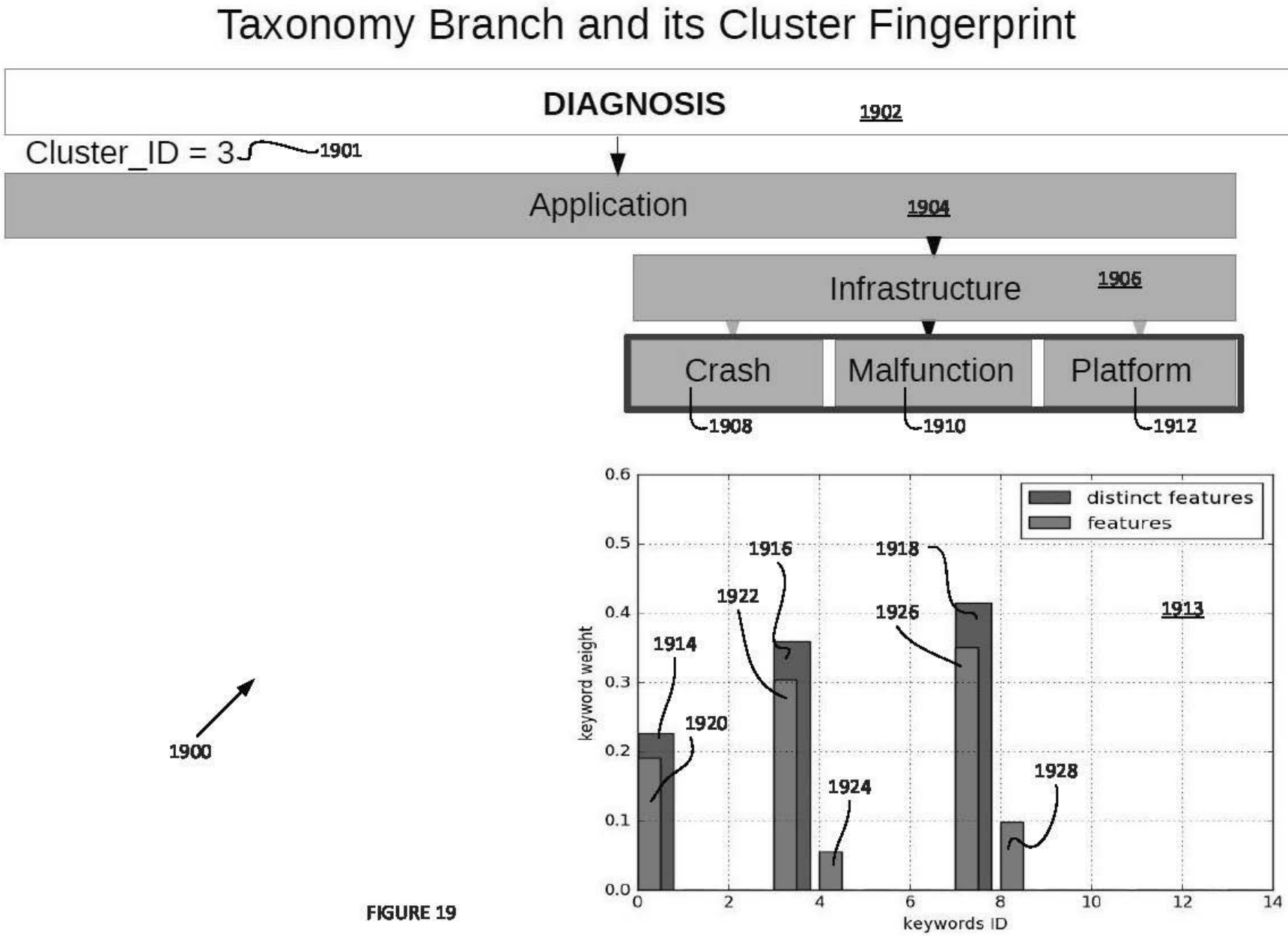
Taxonomy Branch and its Cluster Fingerprint
DIAGNOSIS
1902
Cluster_ID = 3
1901
Application
1904
Infrastructure
1906
Crash
Malfunction
Platform
1908
1910
1912
1900
1913
1914
1916
1918
1920
1922
1924
1926
1928
distinct features
features
keyword weight
keywords ID
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0
2
4
6
8
10
12
14

FIGURE 19

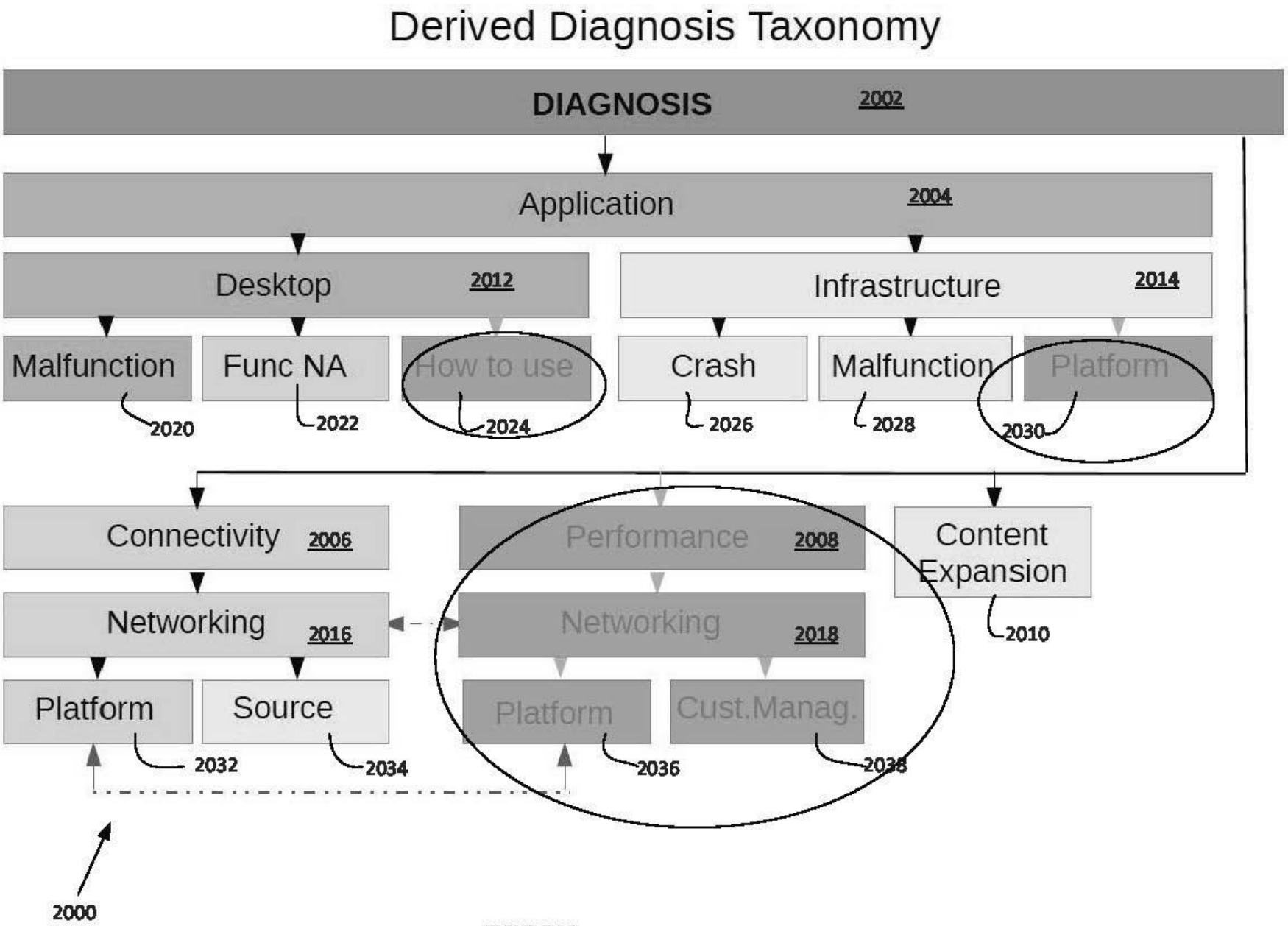
Derived Diagnosis Taxonomy
DIAGNOSIS
2002
Application
2004
Desktop
2012
Infrastructure
2014
Malfunction
Func NA
How to use
Crash
Malfunction
Platform
2020
2022
2024
2026
2028
2030
Connectivity
2006
Performance
2008
Content Expansion
2010
Networking
2016
Networking
2018
Platform
Source
Platform
Cust.Manag.
2032
2034
2036
2038
2000

FIGURE 20

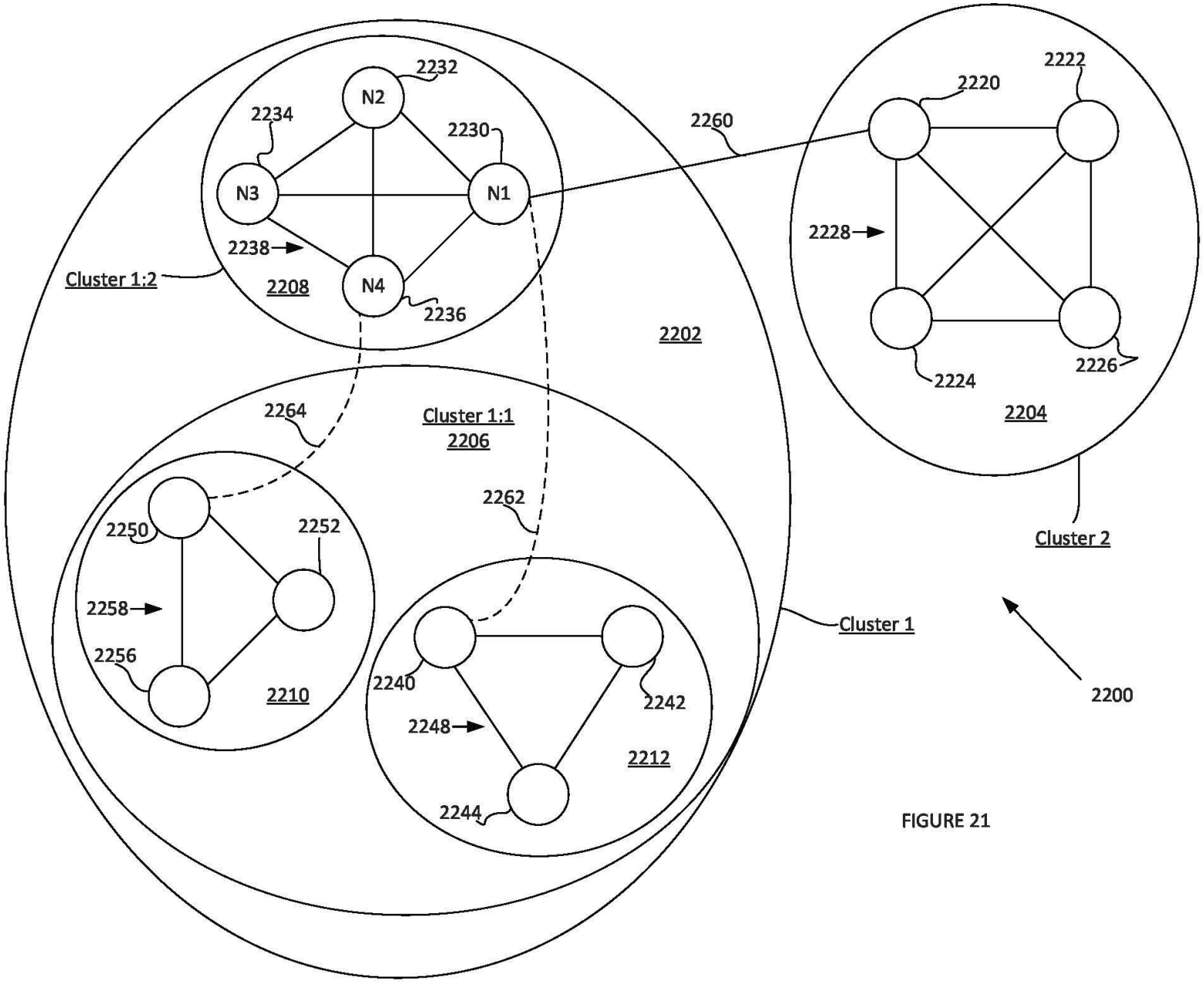
2232
N2
2234
2230
N3
N1
2238
Cluster 1:2
2208
N4
2236
2202
2260
2222
2220
2228
2224
2226
2204
Cluster 2
2264
Cluster 1:1
2206
2262
2250
2252
2258
2256
2210
2240
2242
2248
2212
2244
Cluster 1
2200

FIGURE 21

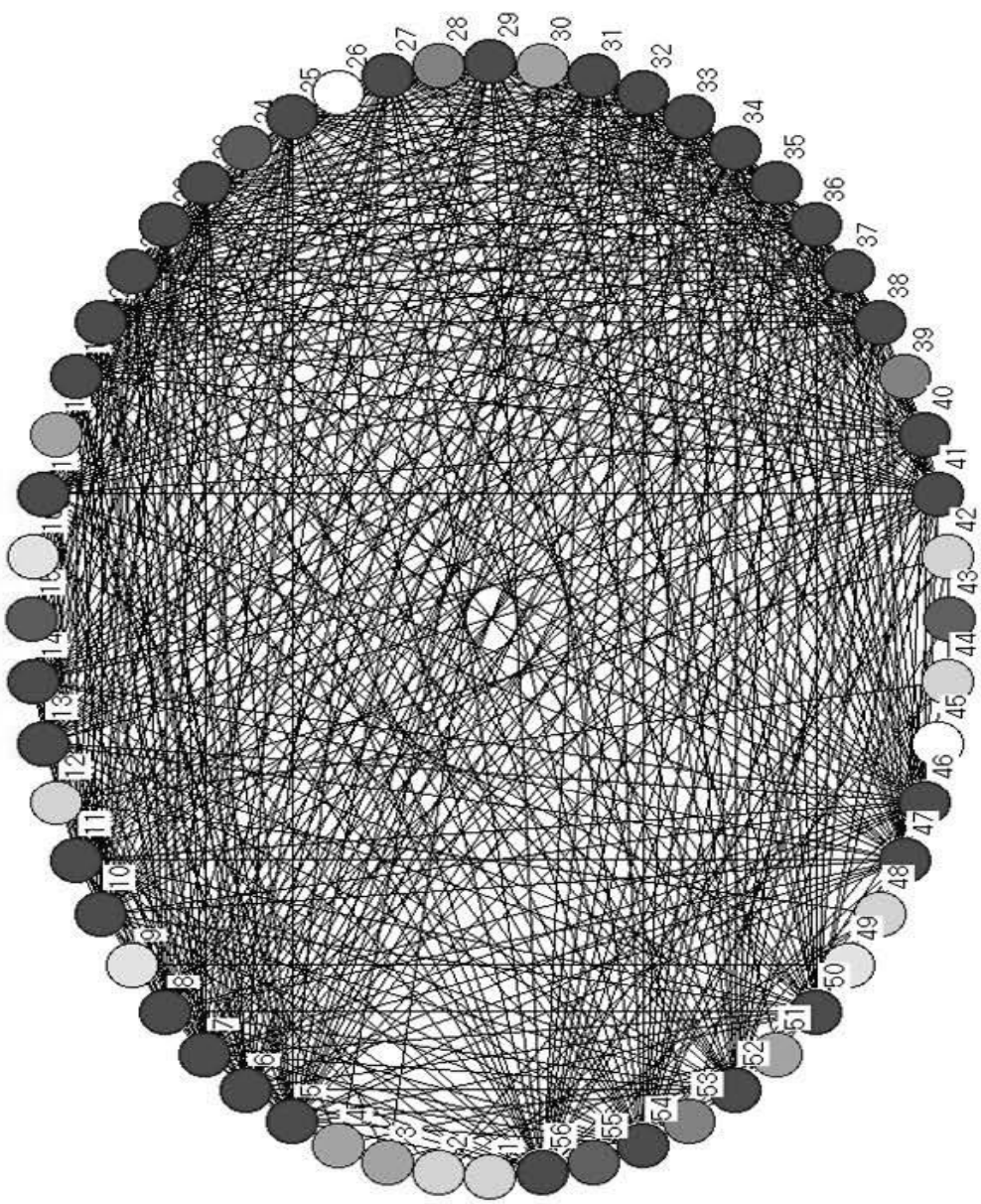

FIGURE 22

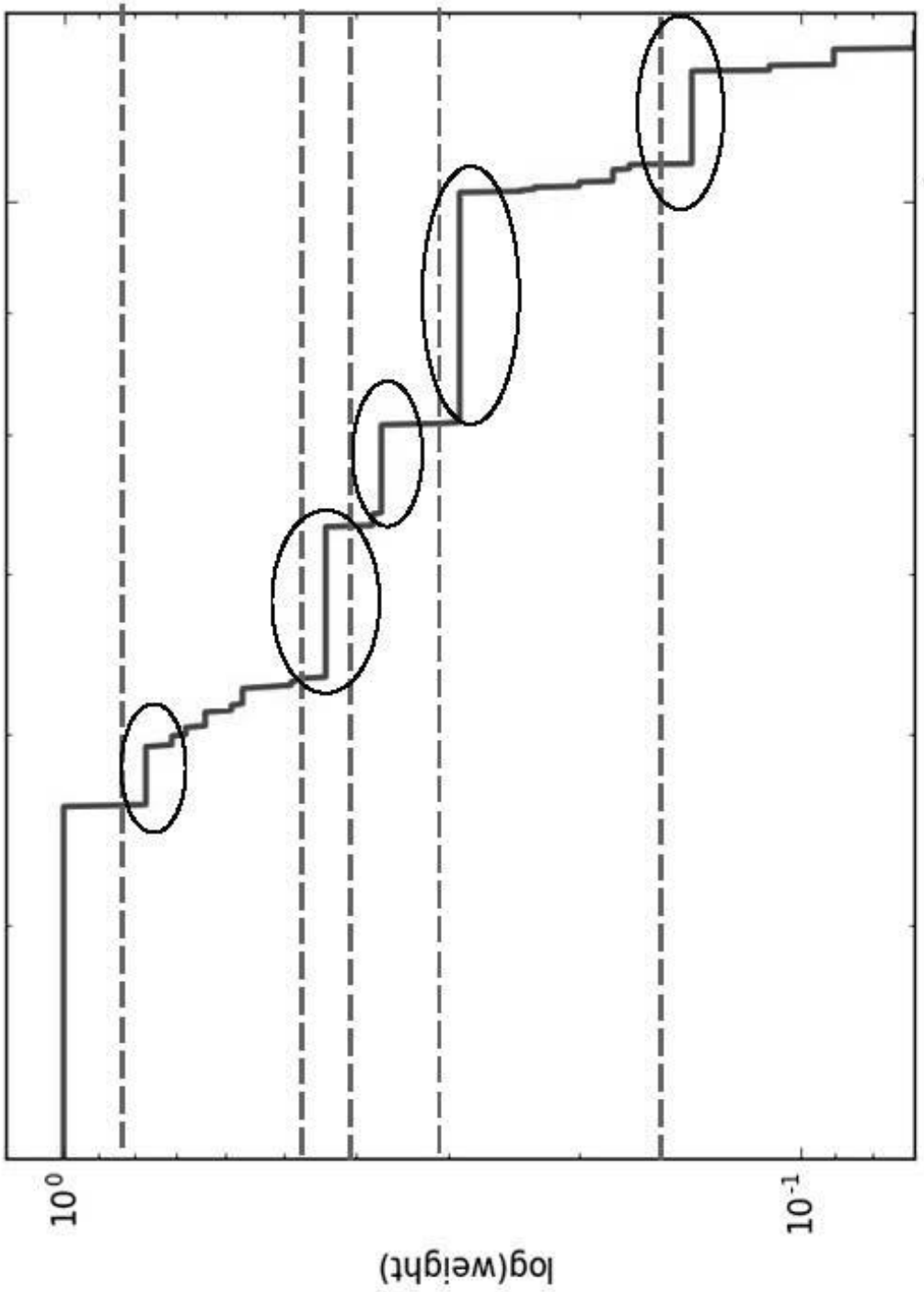

FIGURE 23

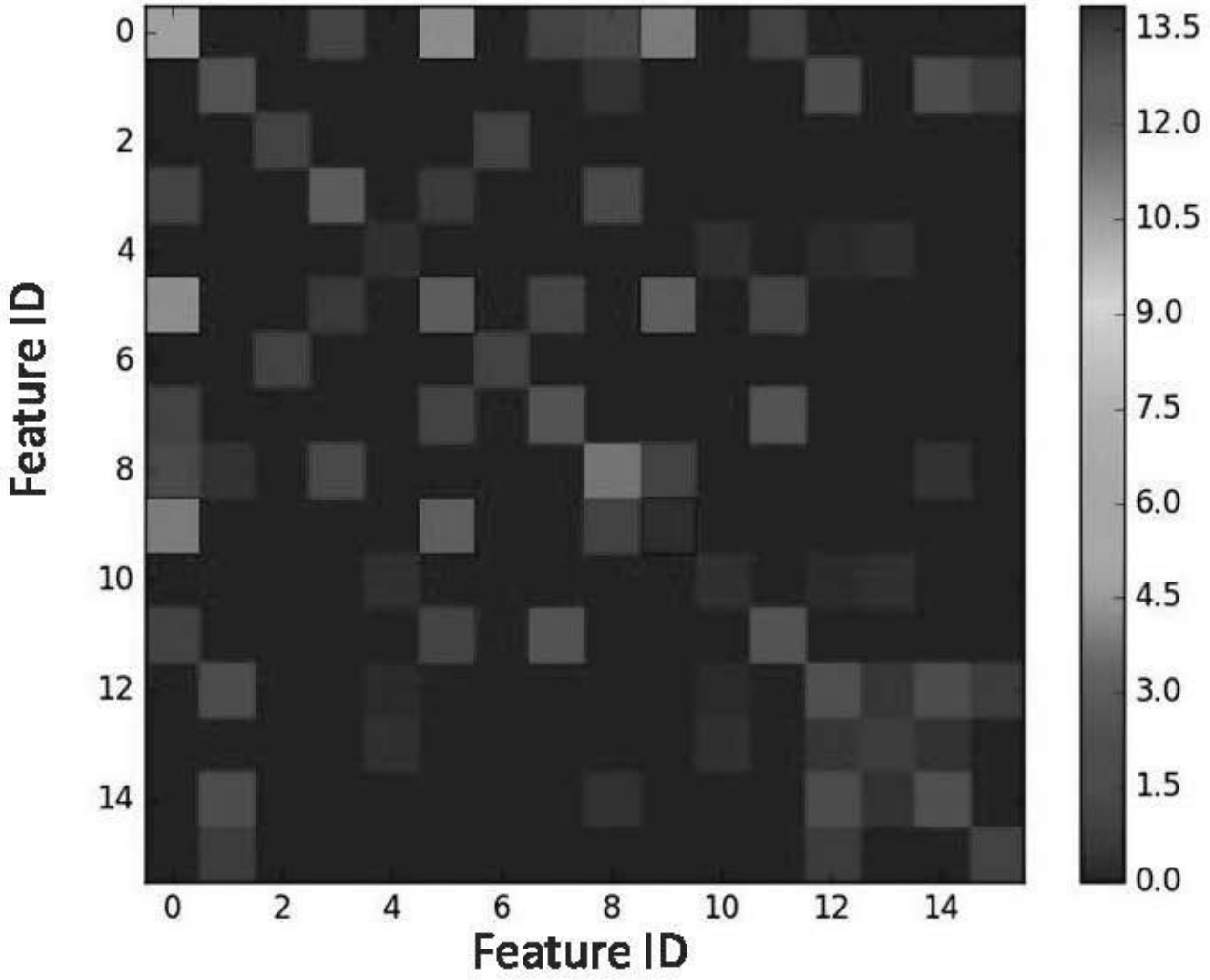
Feature ID
0
2
4
6
8
10
12
14
Feature ID
0
2
4
6
8
10
12
14
13.5
12.0
10.5
9.0
7.5
6.0
4.5
3.0
1.5
0.0

FIGURE 24

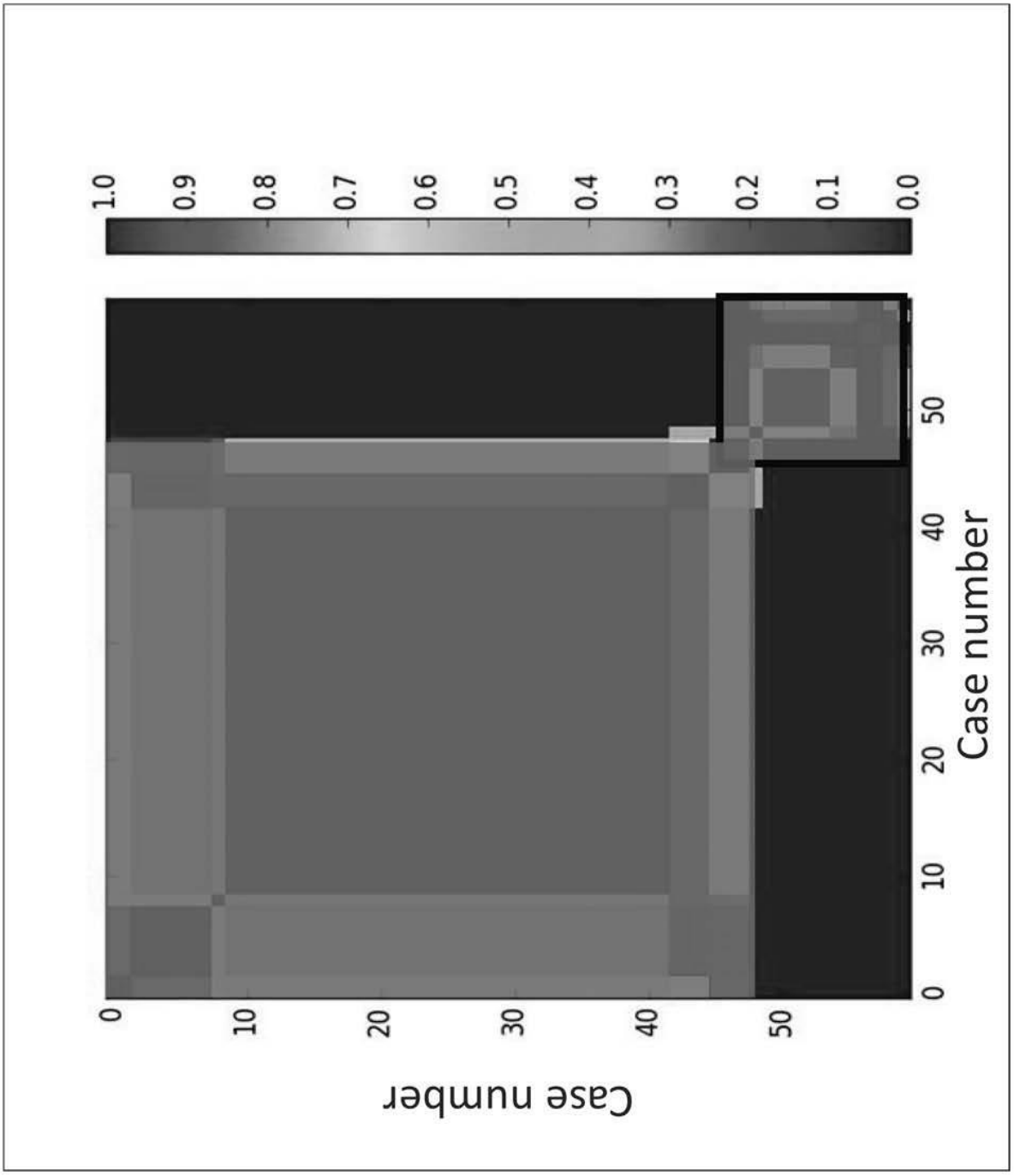
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
10
20
30
40
50
Case number
Case number

FIGURE 25

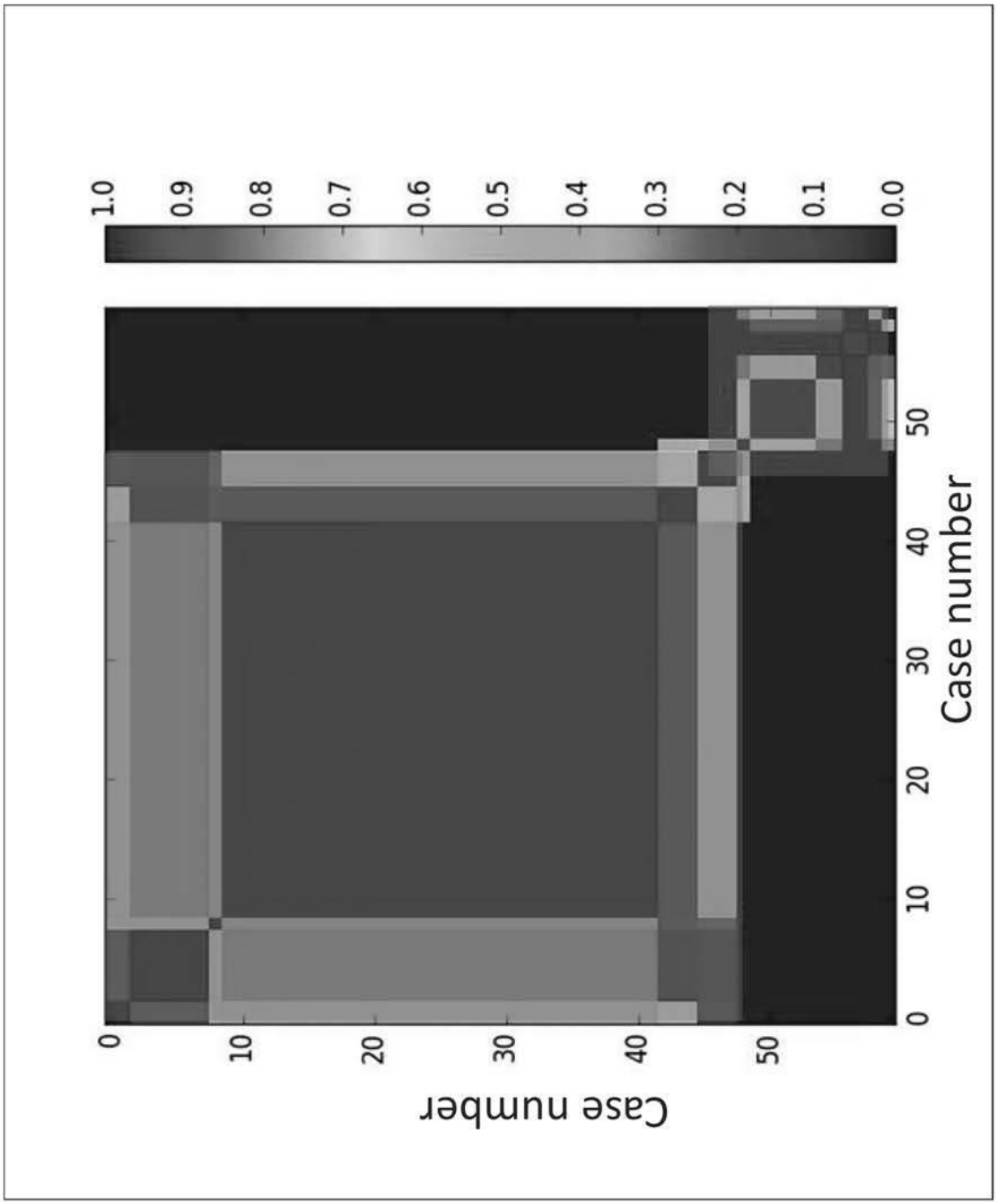
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
10
20
30
40
50
Case number
0
10
20
30
40
50
Case number

FIGURE 26

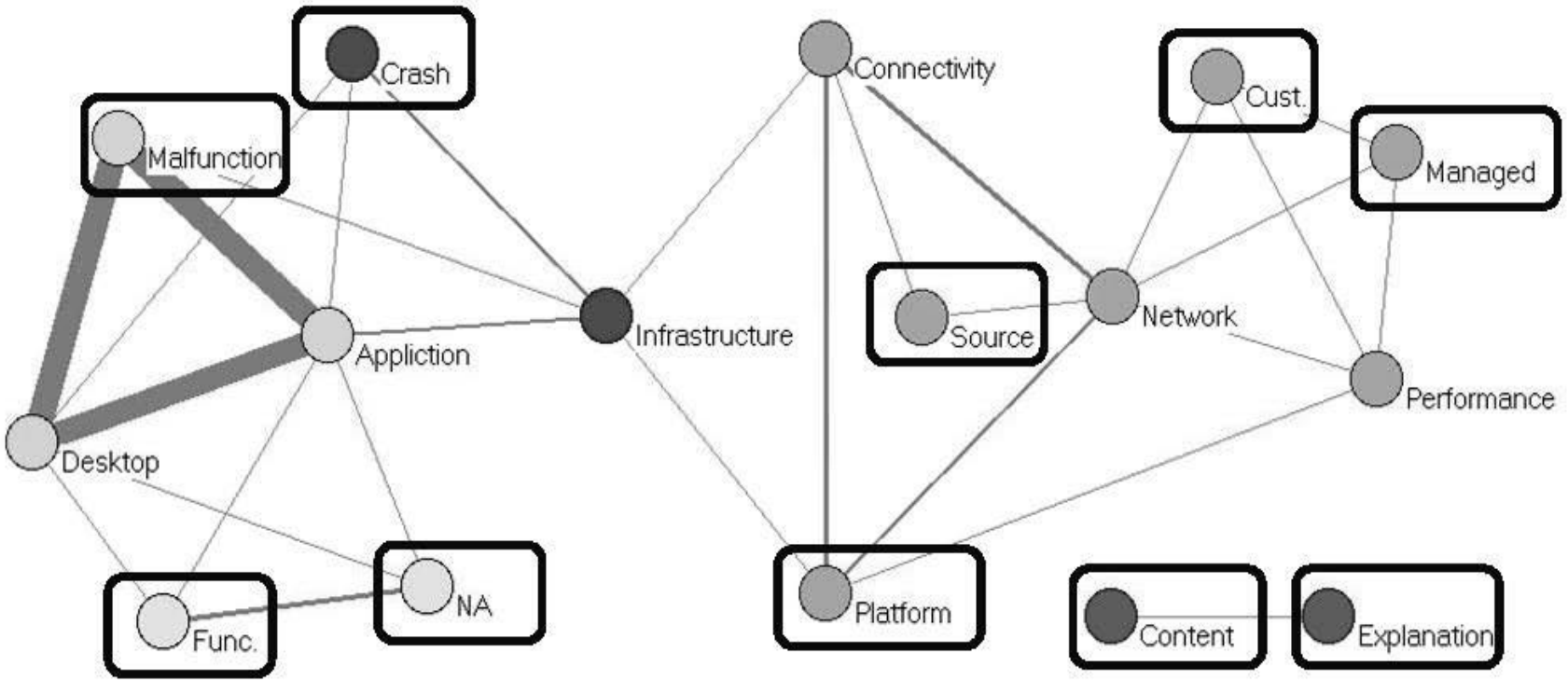
Crash
Malfunction
Appliction
Desktop
Func.
NA
Infrastructure
Connectivity
Platform
Source
Network
Cust.
Managed
Performance
Content
Explanation

FIGURE 27

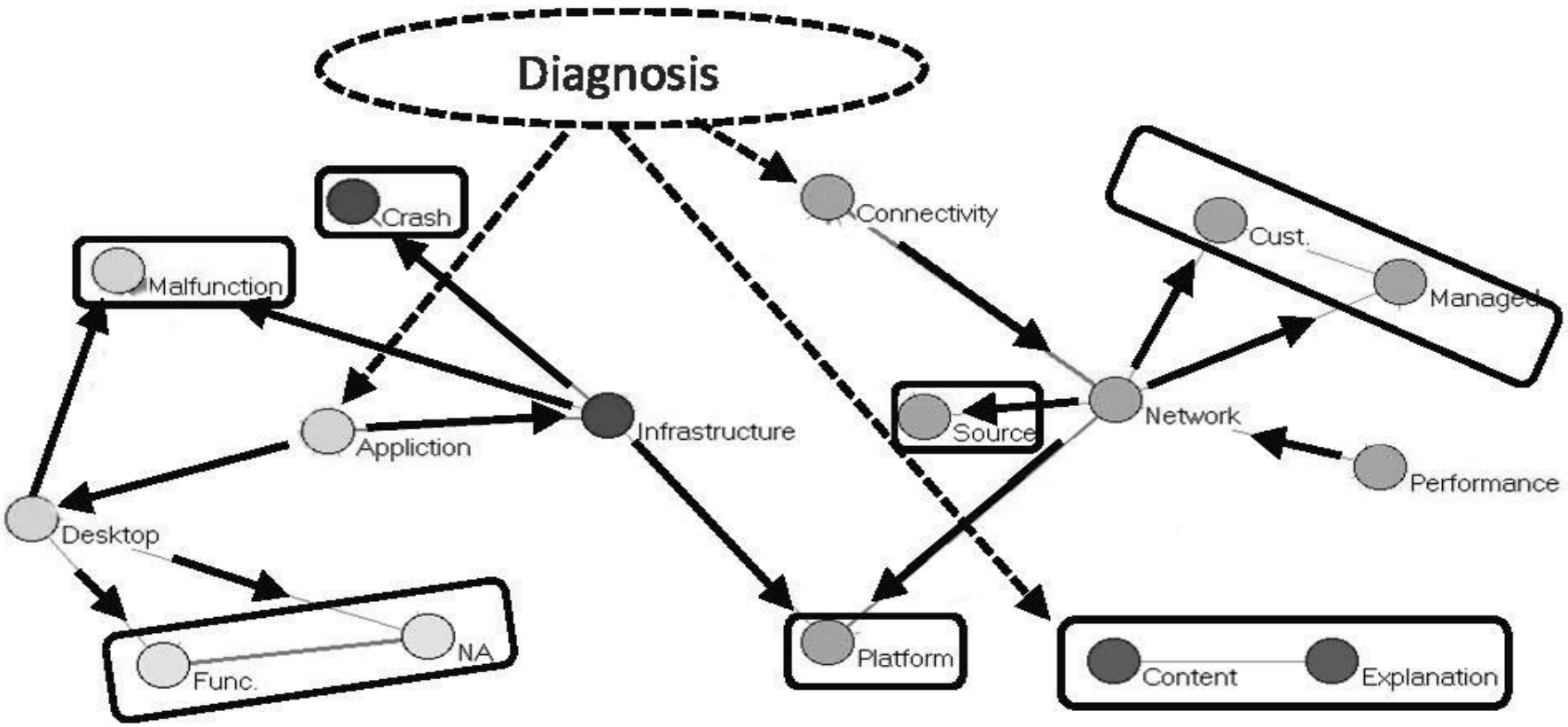
Diagnosis
Crash
Connectivity
Malfunction
Cust.
Managed
Appliction
Infrastructure
Source
Network
Performance
Desktop
NA
Func.
Platform
Content
Explanation

FIGURE 28

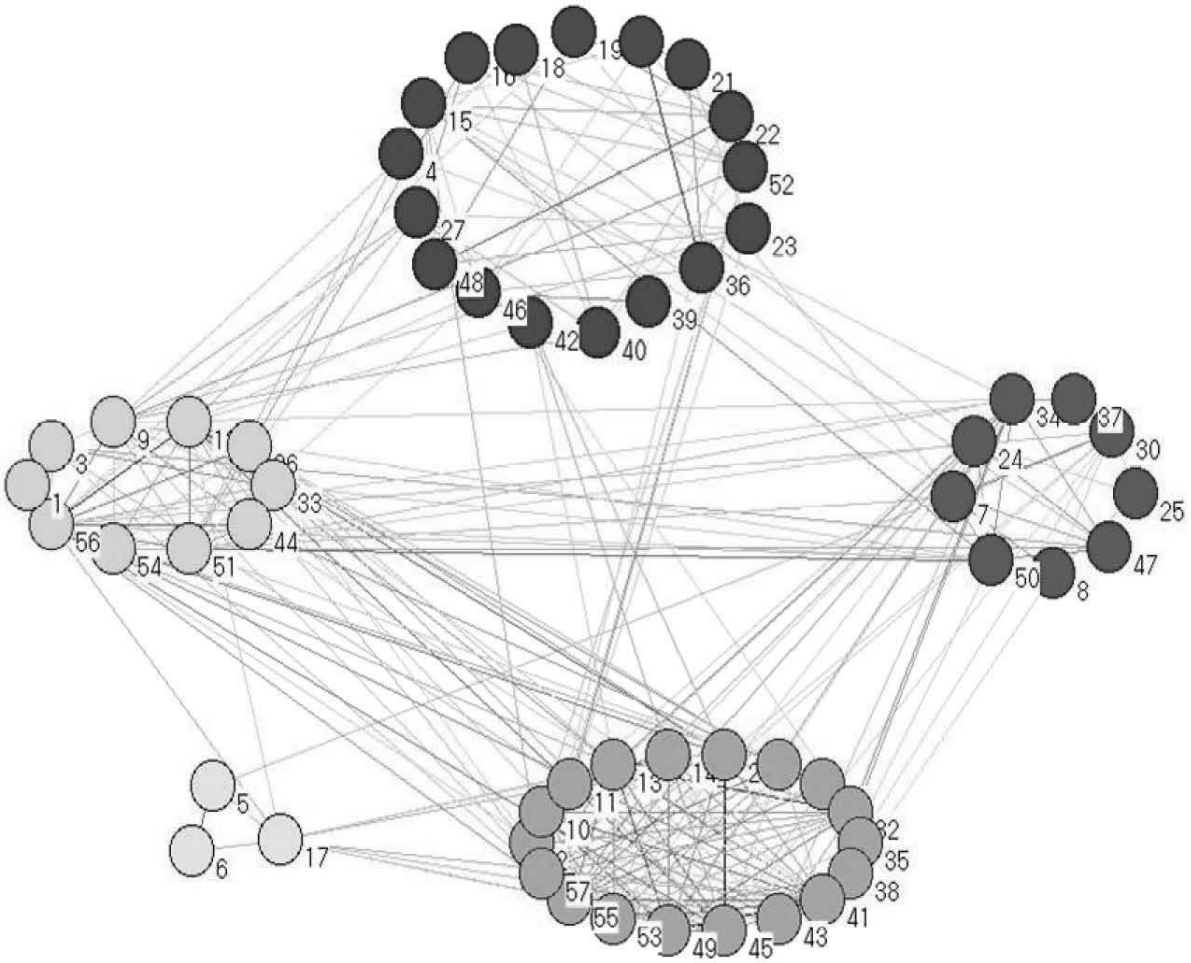
19
18
16
21
15
22
4
52
27
23
48
36
46
39
42
40
3
9
1
33
56
44
54
51
34
37
30
24
25
7
47
50
8
5
17
6
13
14
11
10
32
35
57
38
55
41
53
49
45
43

FIGURE 29

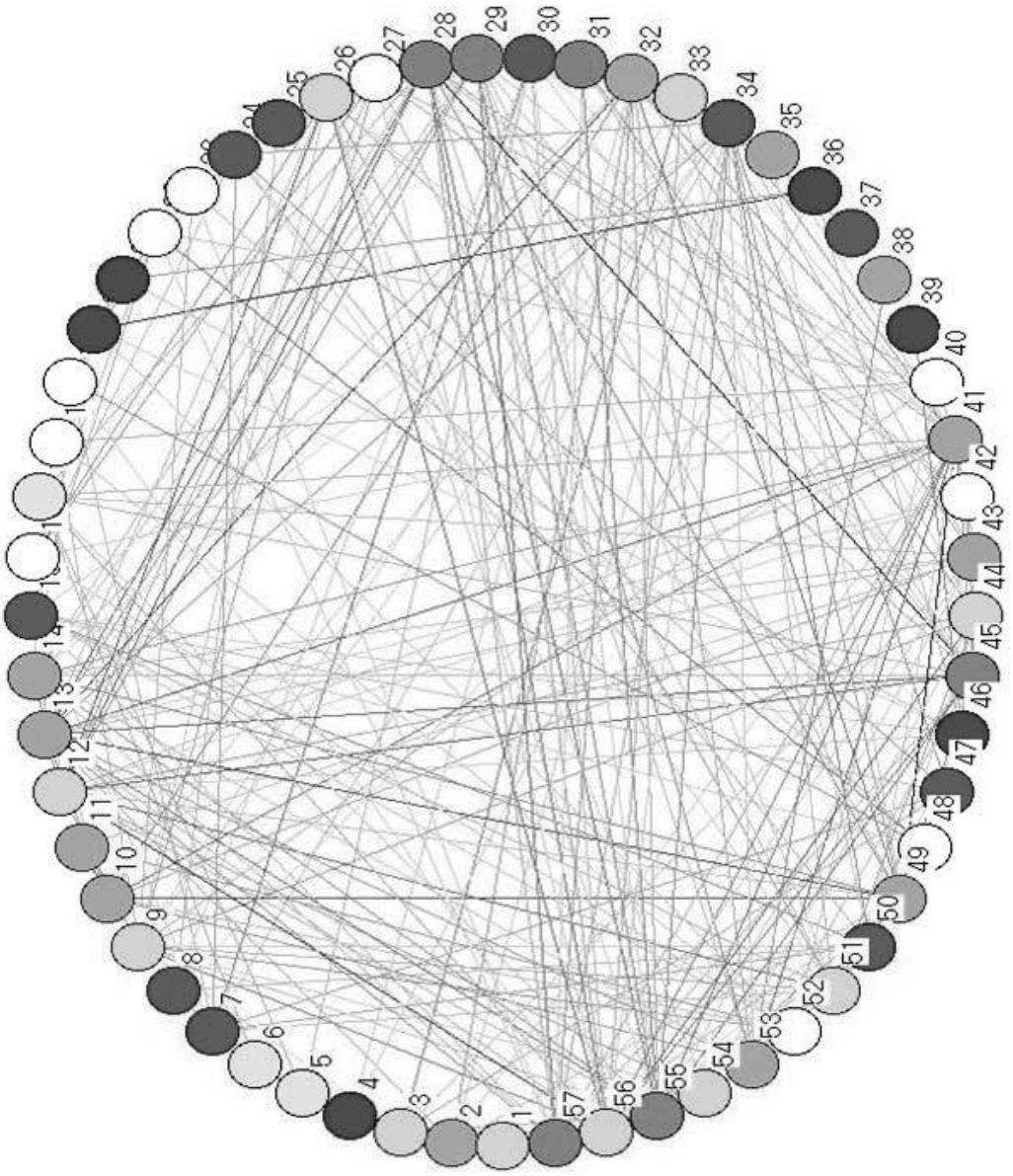

FIGURE 30