Metoda spektralna - testy

Wawrzeńczyk Adam

13 maja 2019

Zbiory testowe

Przygotowane zostały trzy zbiory testowe:

- **graph** - złożony z kształtów geometrycznych zbiór na wzór grafu K_4 (chmurki gaussowskie i proste).
- **labirynth** - kafelkowy układ na kształt prostego labiryntu
- **windows** - kafelkowy układ złożony z ramki oraz czterech skupień w środku

Wszystkie zbiory są syntetyczne i umieszczone w R^2. W skład zbioru testowego wchodzi plik .R generujący go, plik .data.csv zawierający punkty zbioru i .labels0.csv zawierający etykiety eksperckie.

Wyniki

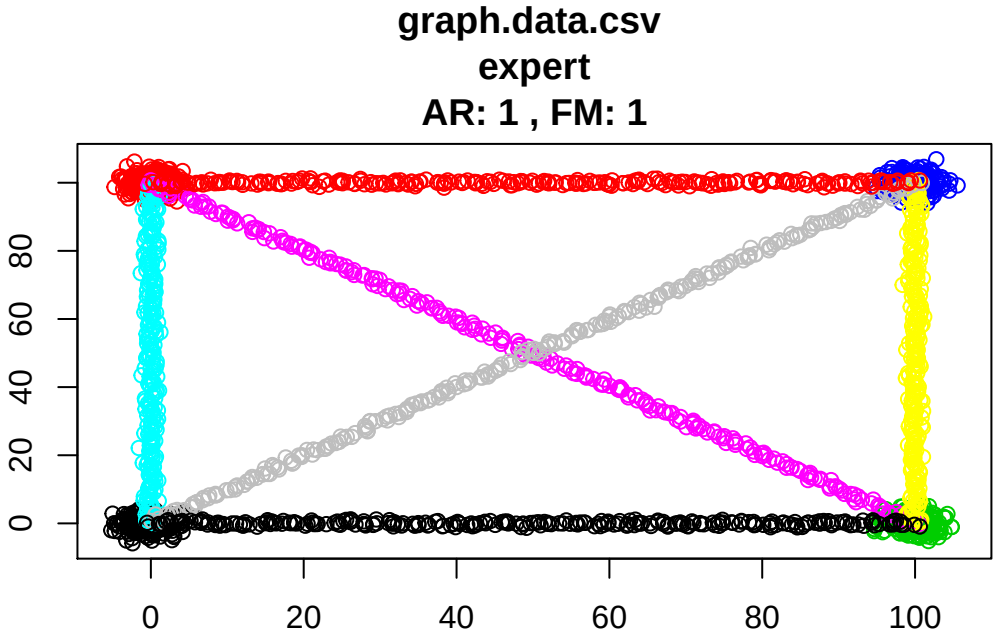
W tej sekcji przedstawione zostaną wyniki dla danych nieustandaryzowanych. Dla każdego zbioru danych zostaną policzone indeksy AR oraz FM. Porównane zostaną one z wynikami otrzymanymi dla innych algorytmów, a następnie wyniki uzyskane przez algorytmy spektralne zostaną przedstawione na wykresach.

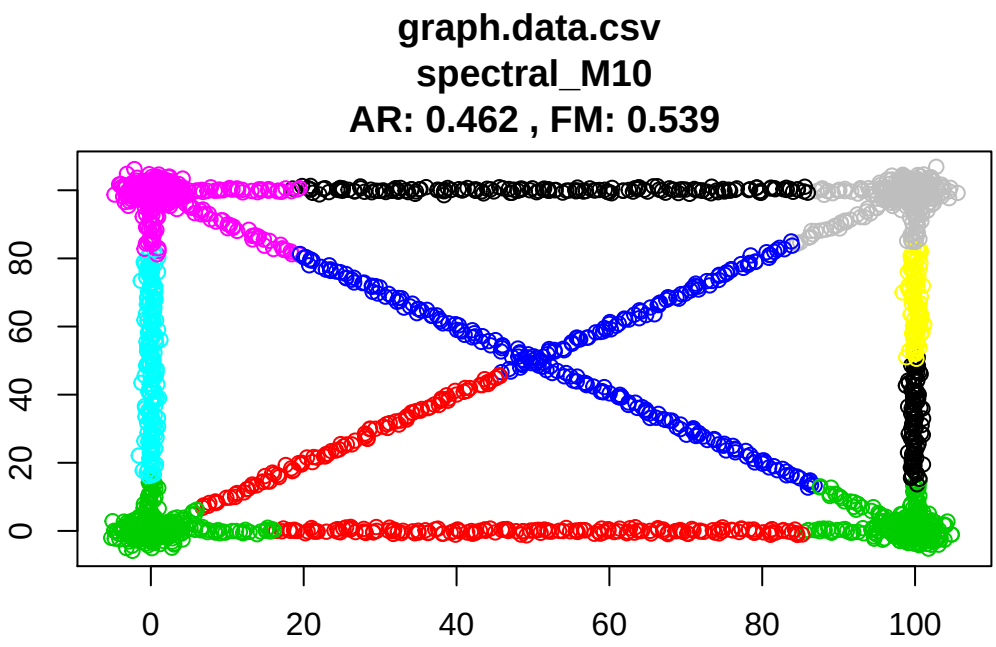
Zbiór graph

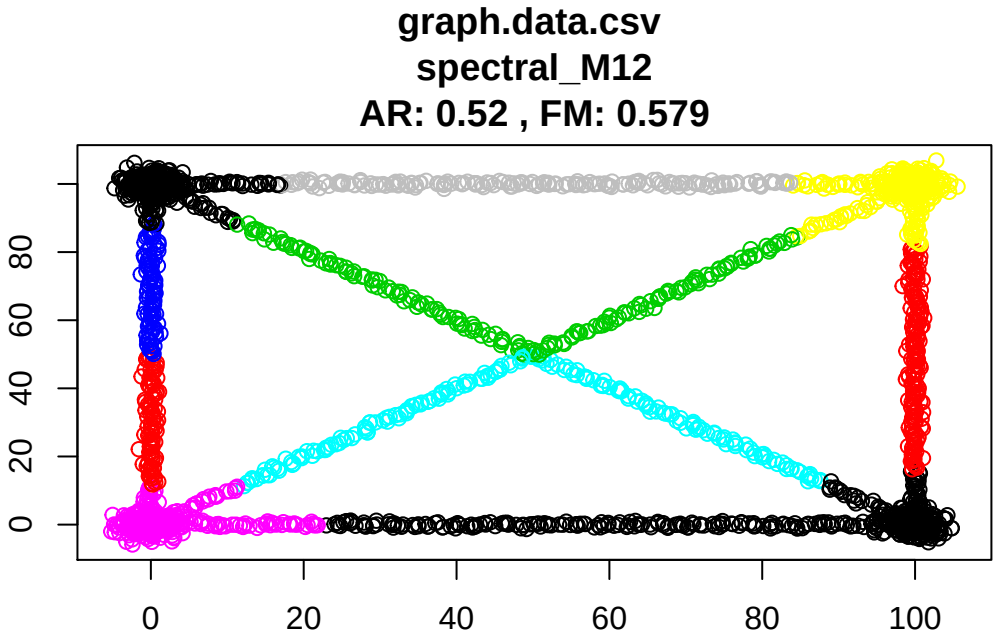
Algorithm	RA	FM
spectral_M30	0.583	0.629
spectral_M20	0.563	0.614
spectral_M2	0.533	0.581
spectral_M15	0.526	0.581
spectral_M12	0.520	0.579
spectral_M8	0.513	0.573
spectral_Mk	0.504	0.564
hclust_wardD	0.477	0.537
genie_0.2	0.472	0.530
genie_0.3	0.472	0.535
fuzzy_5	0.465	0.522
spectral_M10	0.462	0.539
fuzzy_2	0.448	0.515
fuzzy_default	0.445	0.513
spectral_M5	0.435	0.509
hclust_wardD2	0.427	0.501
kmeans	0.425	0.500
genie_0.5	0.416	0.506
hclust_mcquitty	0.413	0.494
hclust_centroid2	0.409	0.492
fuzzy_10	0.407	0.477
hclust_complete	0.387	0.472
hclust_average	0.385	0.472
hclust_centroid	0.378	0.472

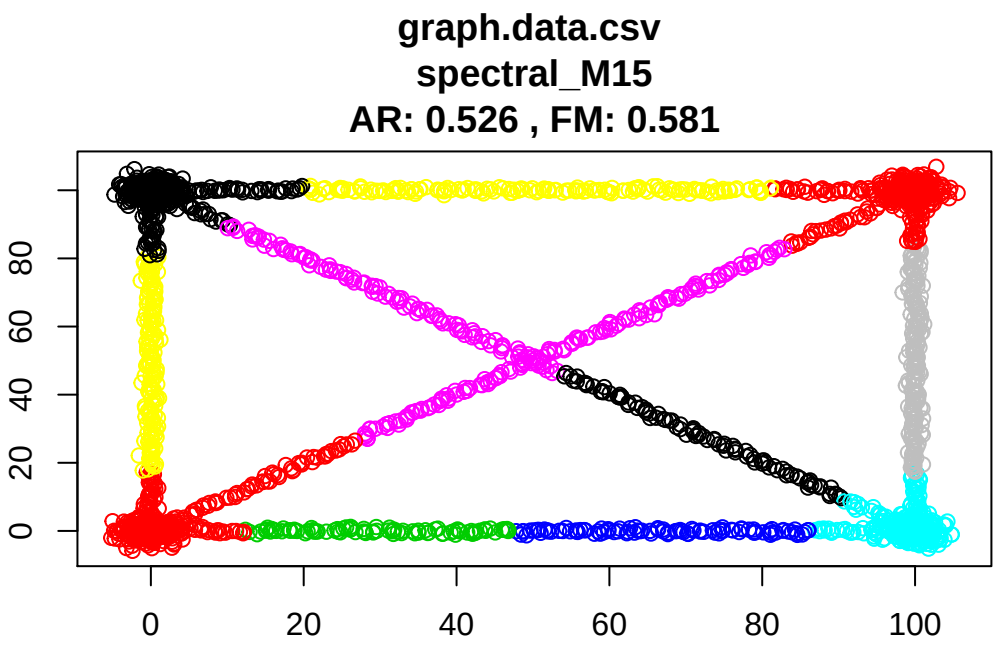
Algorithm	RA	FM
hclust_median	0.374	0.467
genie_0.8	0.238	0.411
hclust_single	0.074	0.328

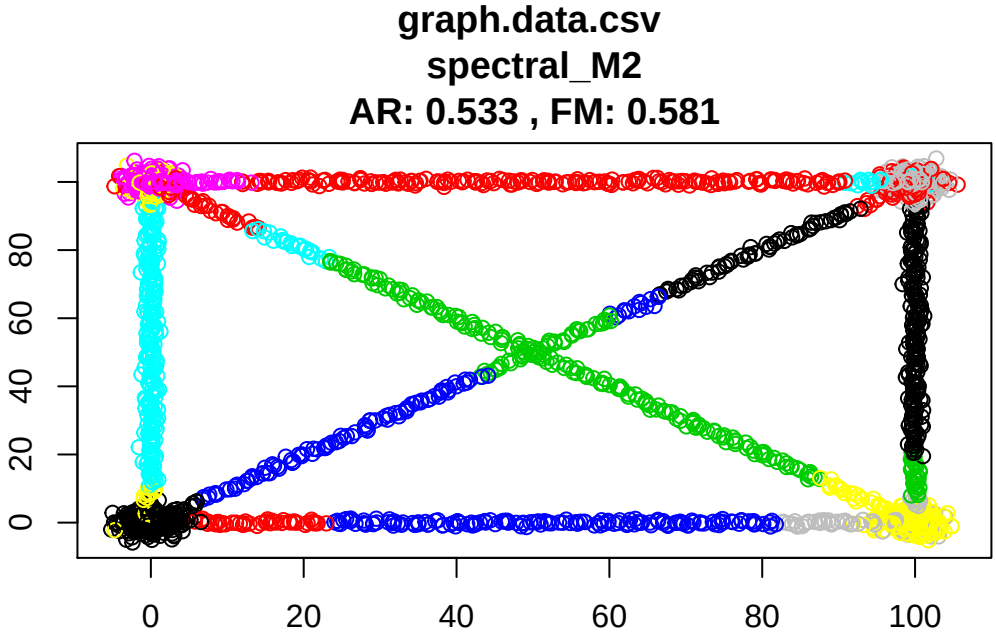
Jest to zbiór, z którym algorytmy spektralne radzą sobie wyjątkowo dobrze - z niektórymi parametrami pokonują nawet algorytm *genie*. Dosyć słabo radzą sobie z nim za to algorytmy hierarchiczne.

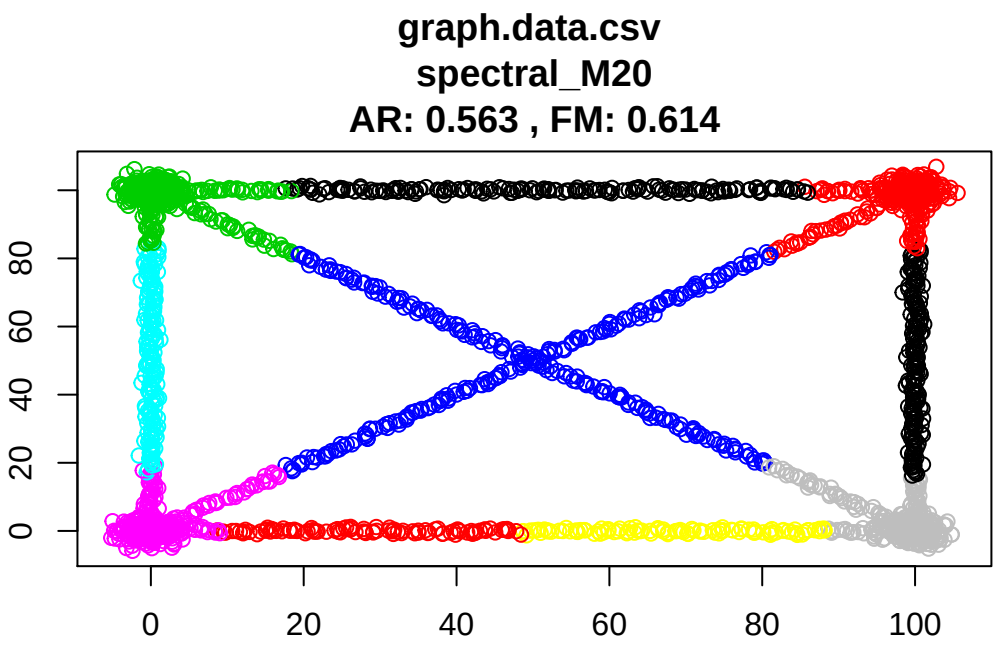


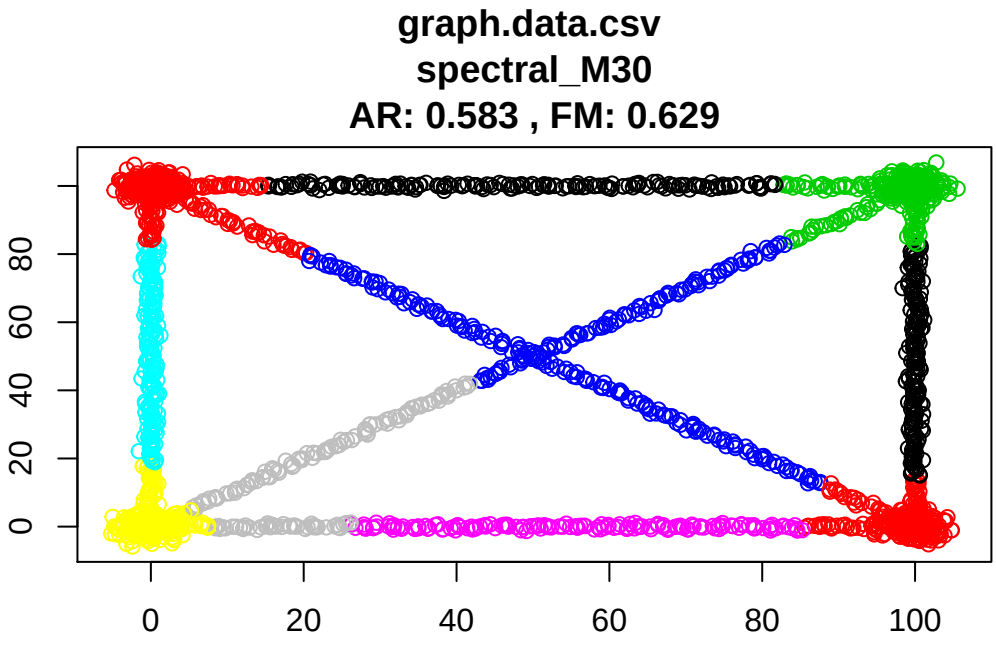

graph.data.csv
spectral_M10
AR: 0.462 , FM: 0.539
0
20
40
60
80
0
20
40
60
80
100

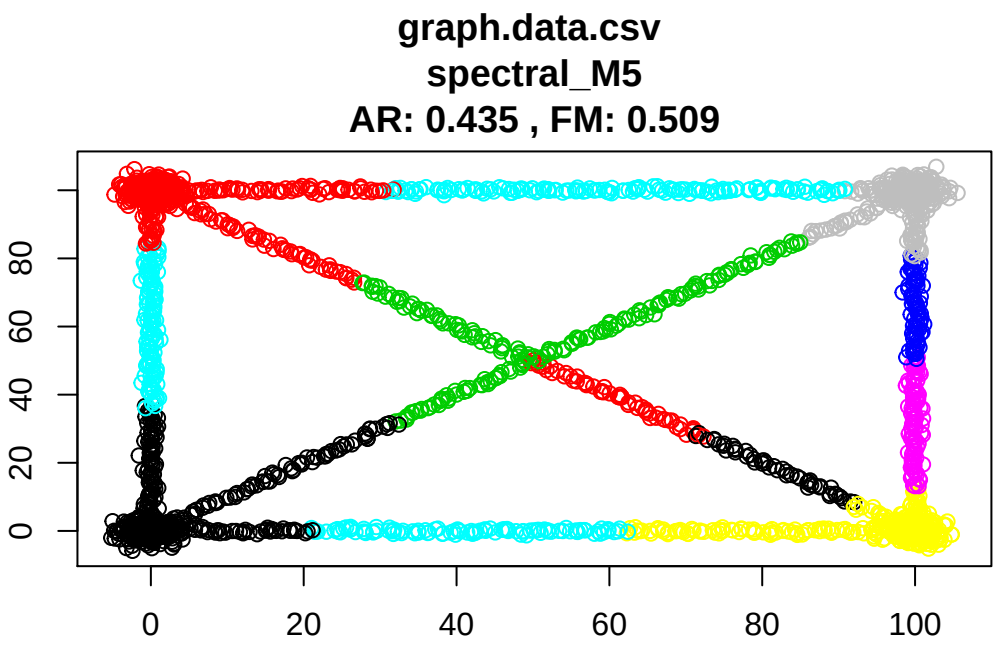

graph.data.csv
spectral_M12
AR: 0.52 , FM: 0.579
0
20
40
60
80
0
20
40
60
80
100

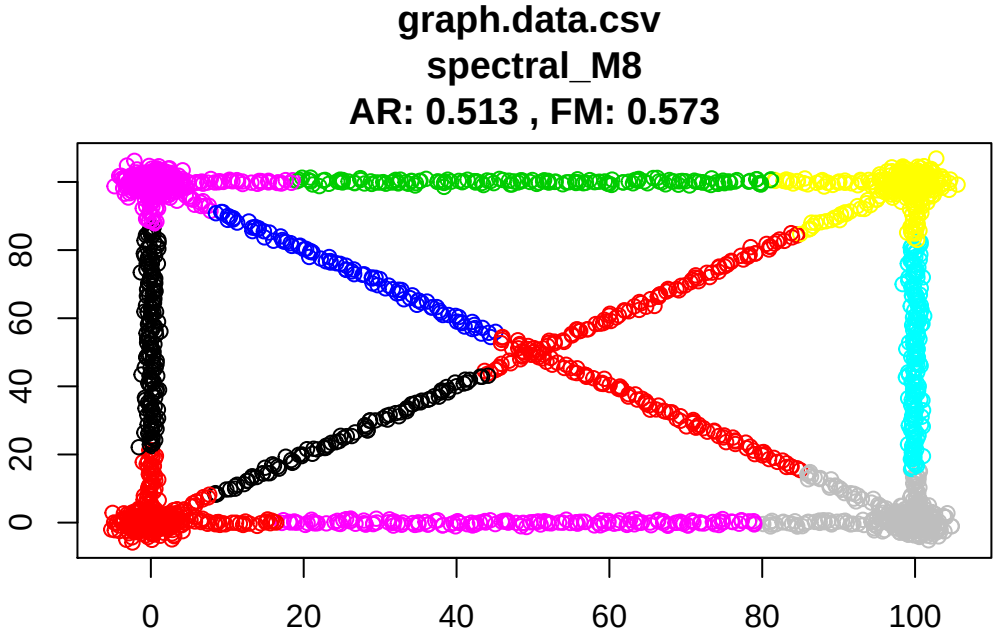

graph.data.csv
spectral_M15
AR: 0.526 , FM: 0.581
0
20
40
60
80
0
20
40
60
80
100

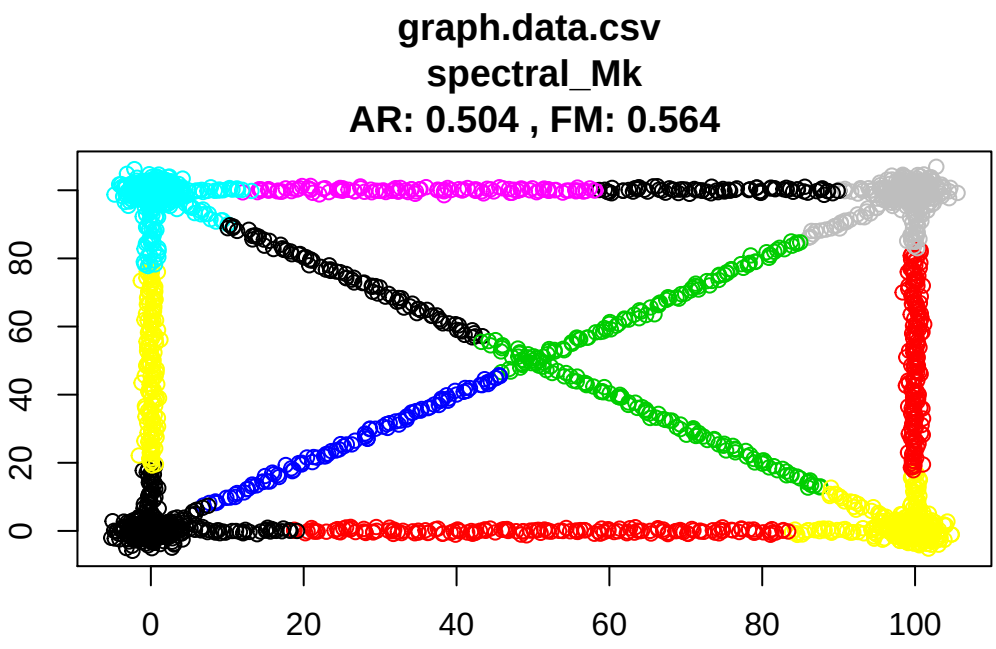

graph.data.csv
spectral_M2
AR: 0.533 , FM: 0.581
0
20
40
60
80
0
20
40
60
80
100


graph.data.csv
spectral_M20
AR: 0.563 , FM: 0.614
0
20
40
60
80
0
20
40
60
80
100


graph.data.csv
spectral_M30
AR: 0.583 , FM: 0.629
0
20
40
60
80
0
20
40
60
80
100


graph.data.csv
spectral_M5
AR: 0.435 , FM: 0.509
0
20
40
60
80
0
20
40
60
80
100

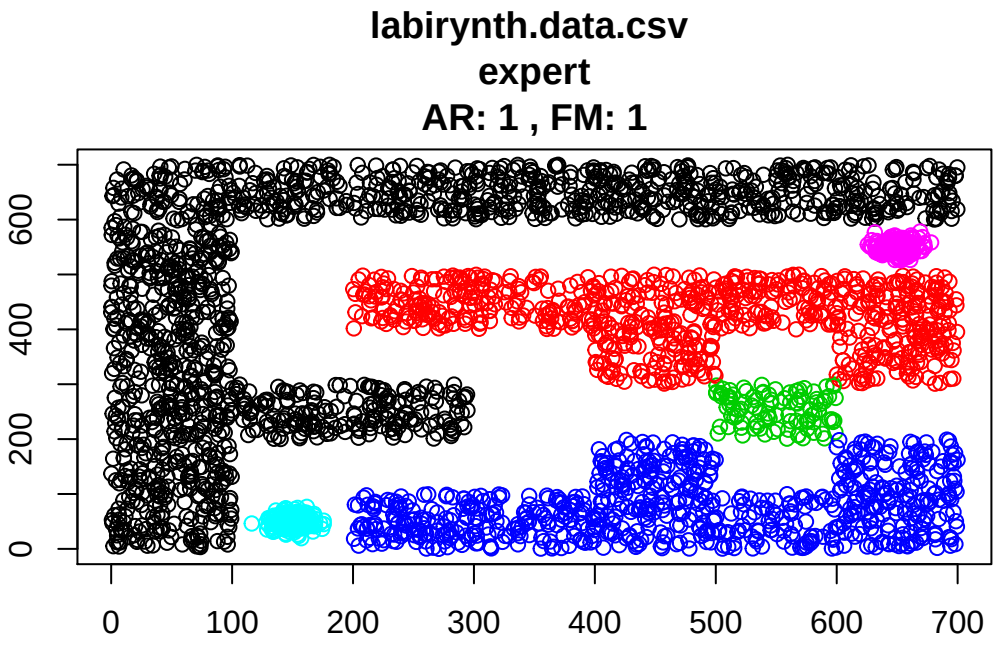

graph.data.csv
spectral_M8
AR: 0.513 , FM: 0.573
0
20
40
60
80
0
20
40
60
80
100

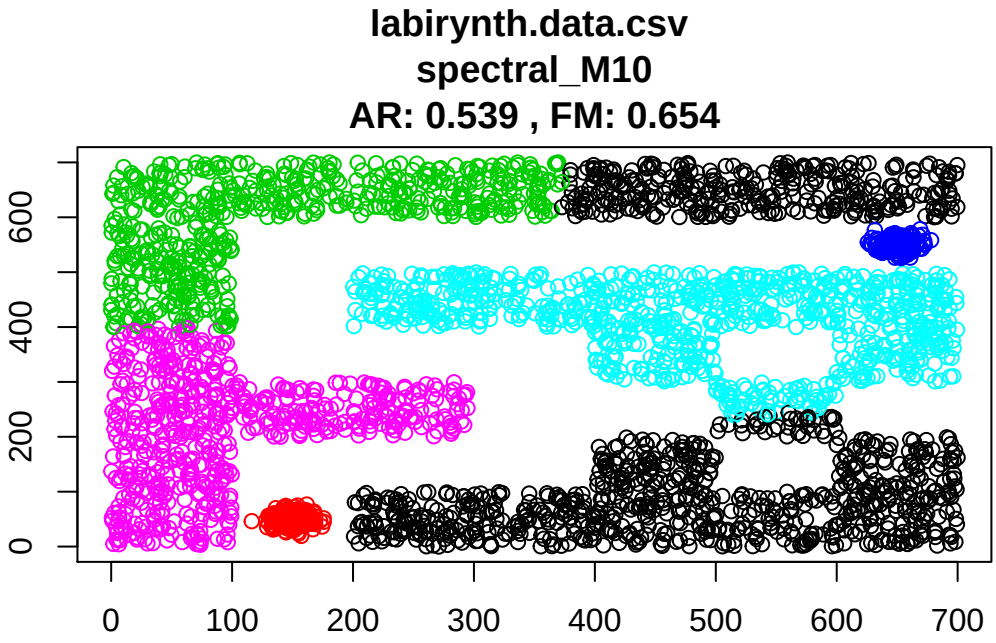

graph.data.csv
spectral_Mk
AR: 0.504 , FM: 0.564
0
20
40
60
80
0
20
40
60
80
100

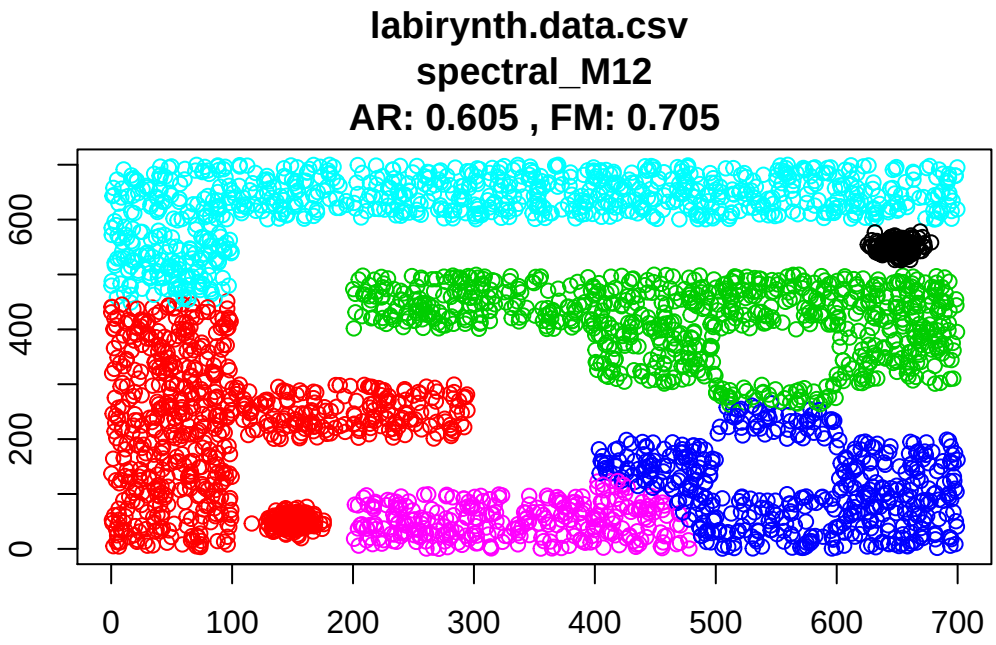
Zbiór labirynth

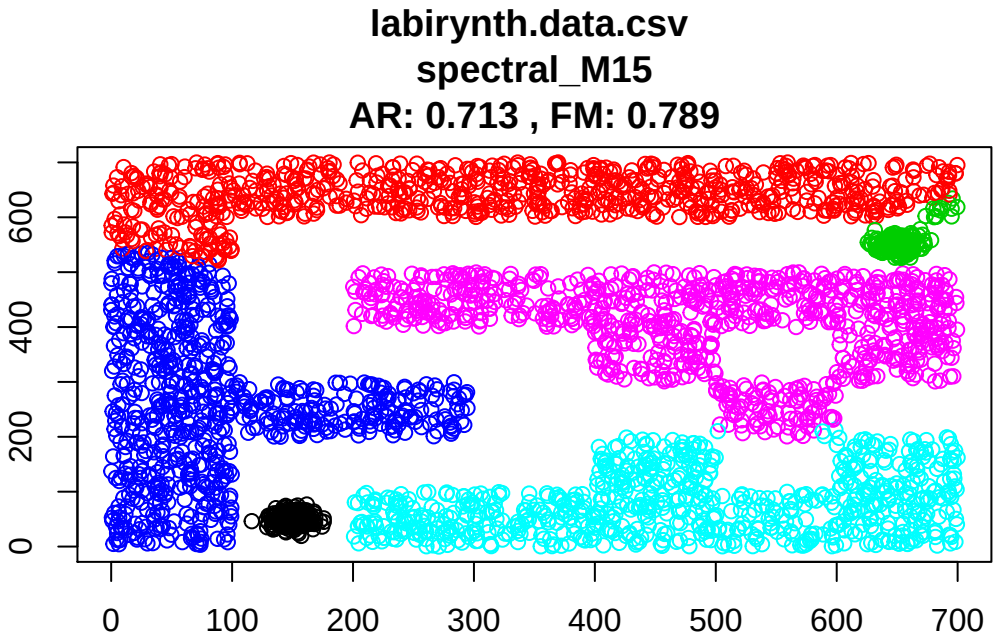
Algorithm	RA	FM
hclust_single	0.761	0.848
genie_0.8	0.745	0.835
spectral_Mk	0.719	0.794
genie_0.5	0.718	0.793
spectral_M15	0.713	0.789
spectral_M12	0.605	0.705
genie_0.2	0.585	0.689
genie_0.3	0.585	0.689
spectral_M20	0.566	0.675
spectral_M5	0.552	0.663
spectral_M10	0.539	0.654
spectral_M8	0.472	0.603
spectral_M30	0.437	0.574
hclust_mcquitty	0.382	0.529
kmeans	0.379	0.524
hclust_average	0.356	0.507
fuzzy_5	0.355	0.504
hclust_wardD2	0.338	0.493
hclust_wardD	0.320	0.478
hclust_centroid	0.307	0.477
hclust_centroid2	0.303	0.467
hclust_complete	0.303	0.462
fuzzy_10	0.284	0.448
fuzzy_default	0.279	0.442
fuzzy_2	0.273	0.438
hclust_median	0.253	0.442
spectral_M2	0.113	0.332

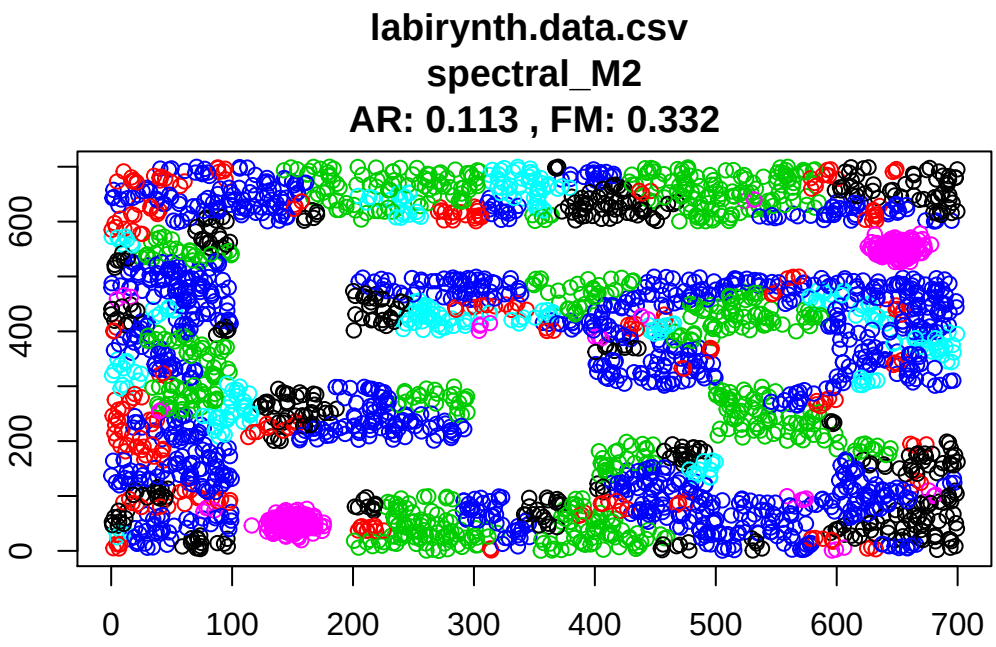
Ten zbiór danych sprawia algorytmom spektralnym nieco większe problemy. Jedna z odmian nadal jest w czołówce, jest jednak gorsza od algorytmów *hclust_single* oraz *genie_0.8* - co ciekawe, są to algorytmy, które zwykle nie mogą pochwalić się dobrymi wynikami.

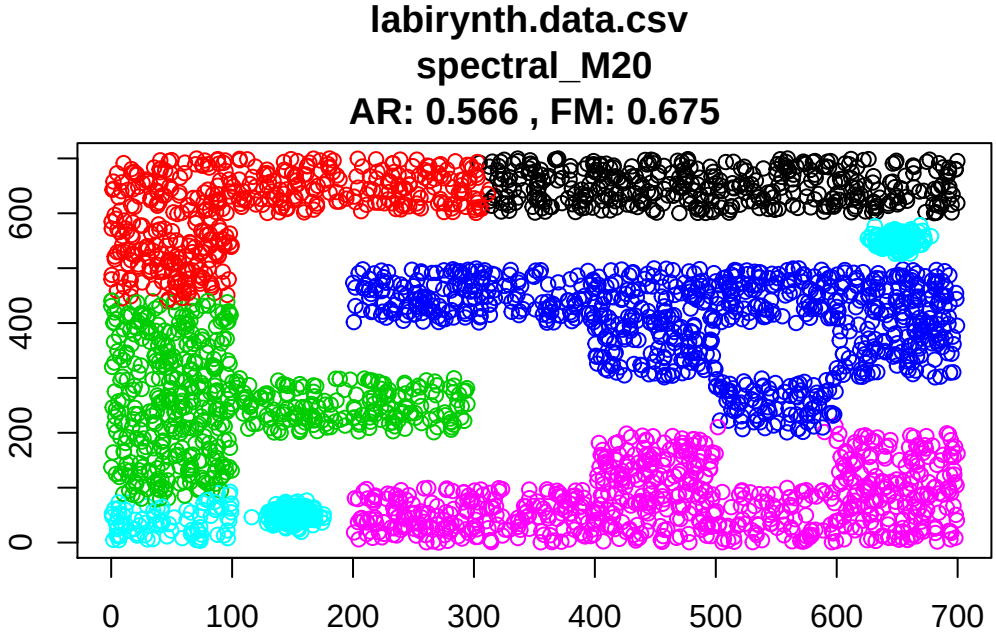

labirynth.data.csv
expert
AR: 1 , FM: 1
0
200
400
600
0
100
200
300
400
500
600
700

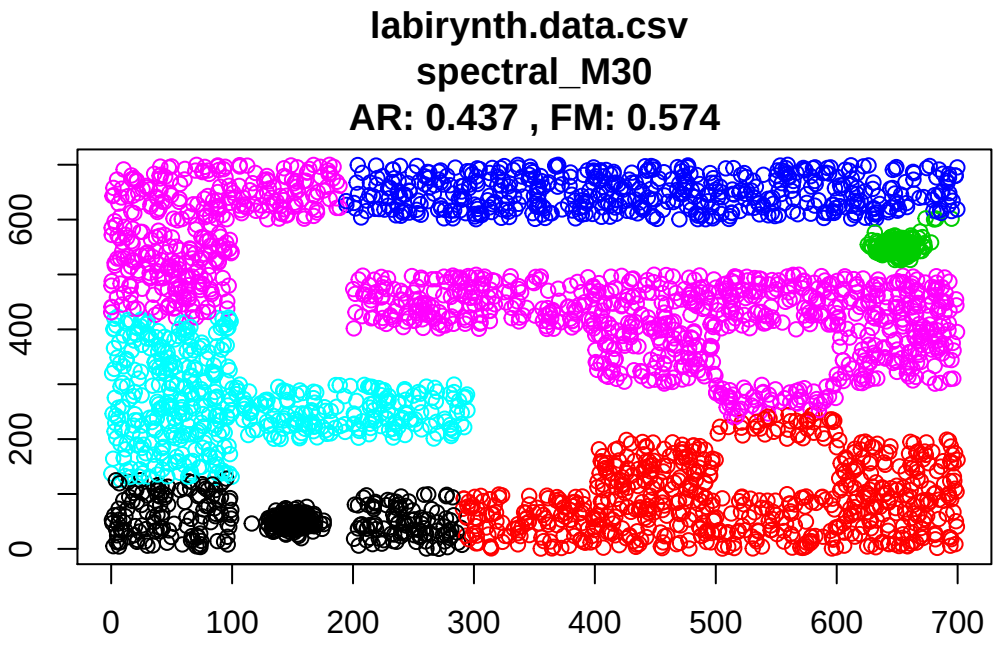

labirynth.data.csv
spectral_M10
AR: 0.539 , FM: 0.654
0
200
400
600
0
100
200
300
400
500
600
700

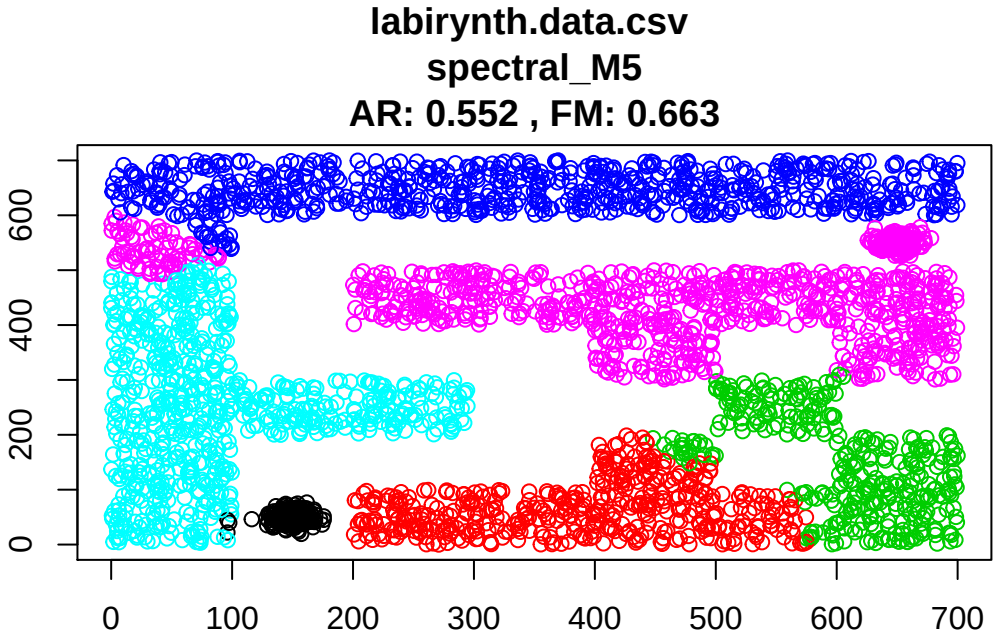

labirynth.data.csv
spectral_M12
AR: 0.605 , FM: 0.705
0
200
400
600
0
100
200
300
400
500
600
700

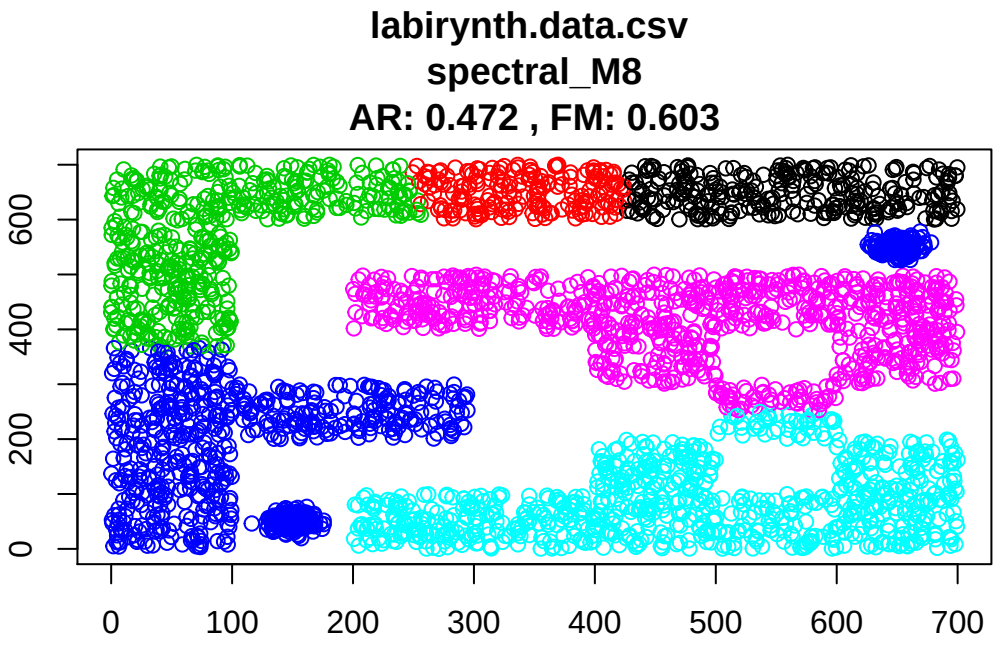

labirynth.data.csv
spectral_M15
AR: 0.713 , FM: 0.789
0
200
400
600
0
100
200
300
400
500
600
700

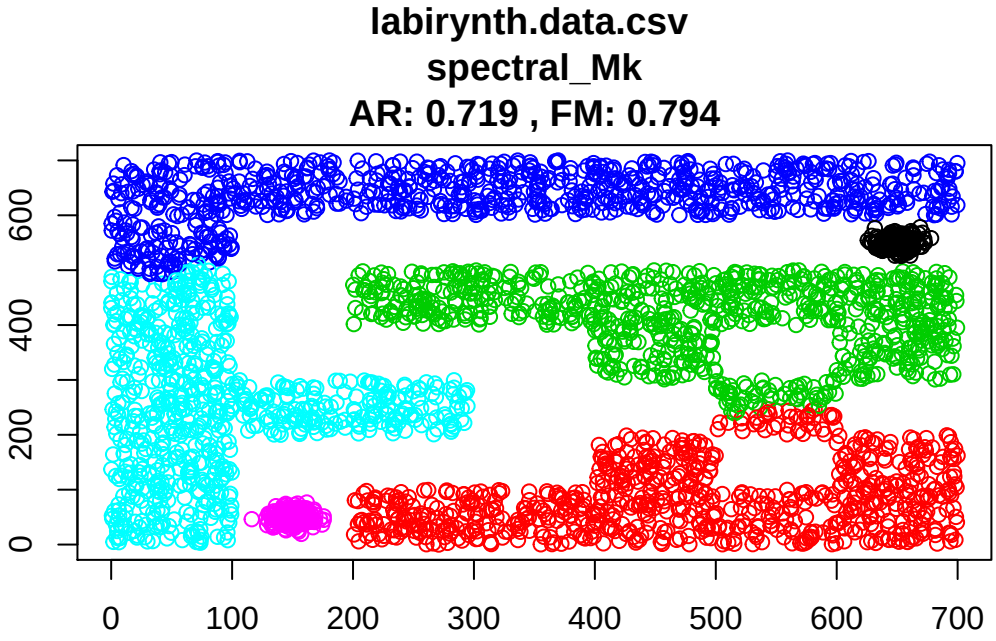

labirynth.data.csv
spectral_M2
AR: 0.113 , FM: 0.332
0
200
400
600
0
100
200
300
400
500
600
700


labirynth.data.csv
spectral_M20
AR: 0.566 , FM: 0.675
0
200
400
600
0
100
200
300
400
500
600
700


labirynth.data.csv
spectral_M30
AR: 0.437 , FM: 0.574
0
200
400
600
0
100
200
300
400
500
600
700


labirynth.data.csv
spectral_M5
AR: 0.552 , FM: 0.663
0
200
400
600
0
100
200
300
400
500
600
700

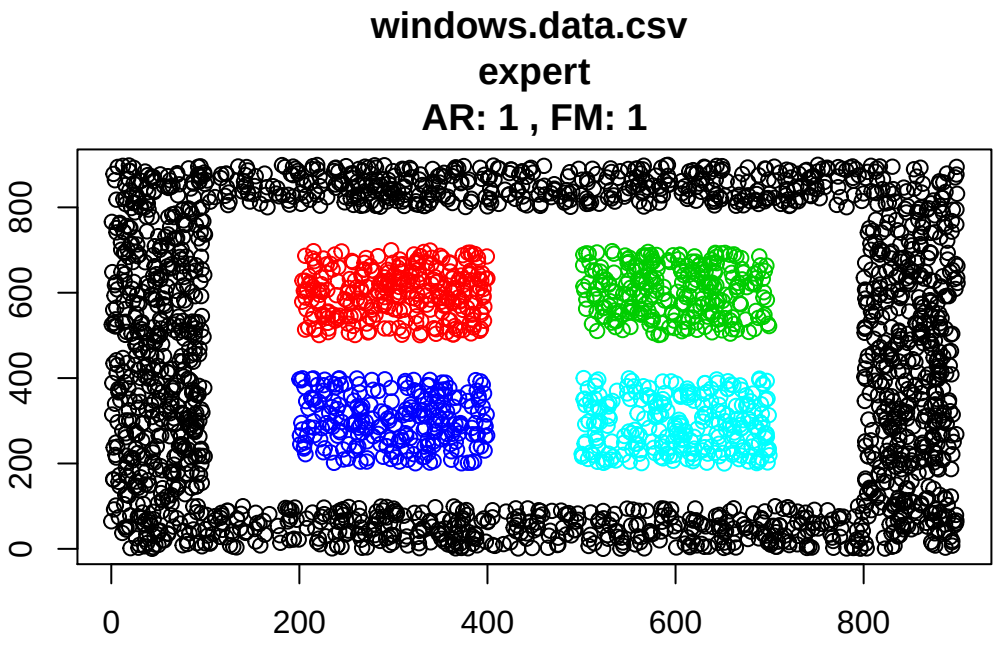

labirynth.data.csv
spectral_M8
AR: 0.472 , FM: 0.603
0
200
400
600
0
100
200
300
400
500
600
700

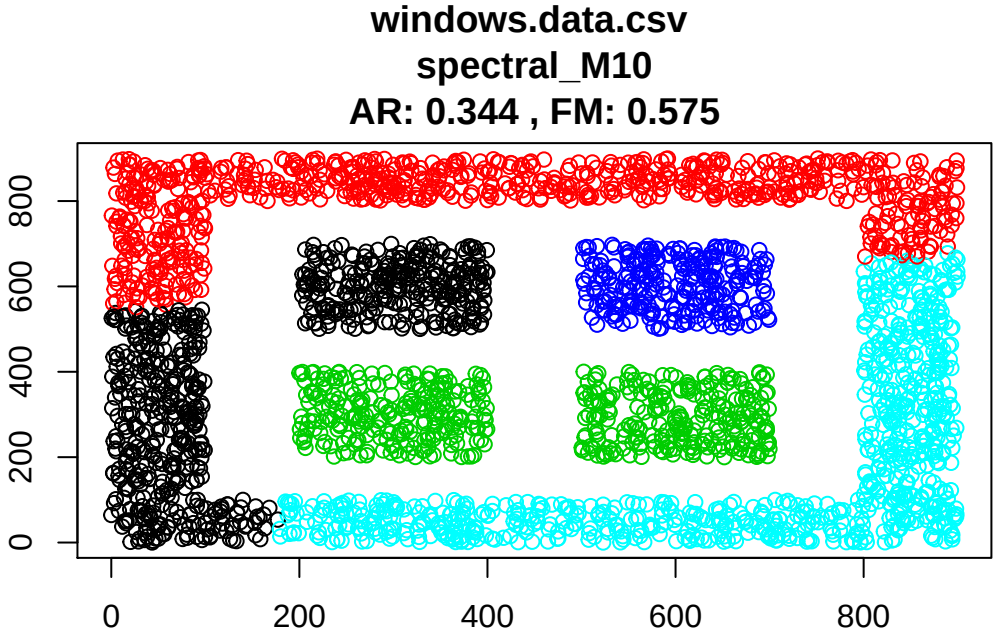

labirynth.data.csv
spectral_Mk
AR: 0.719 , FM: 0.794
0
200
400
600
0
100
200
300
400
500
600
700

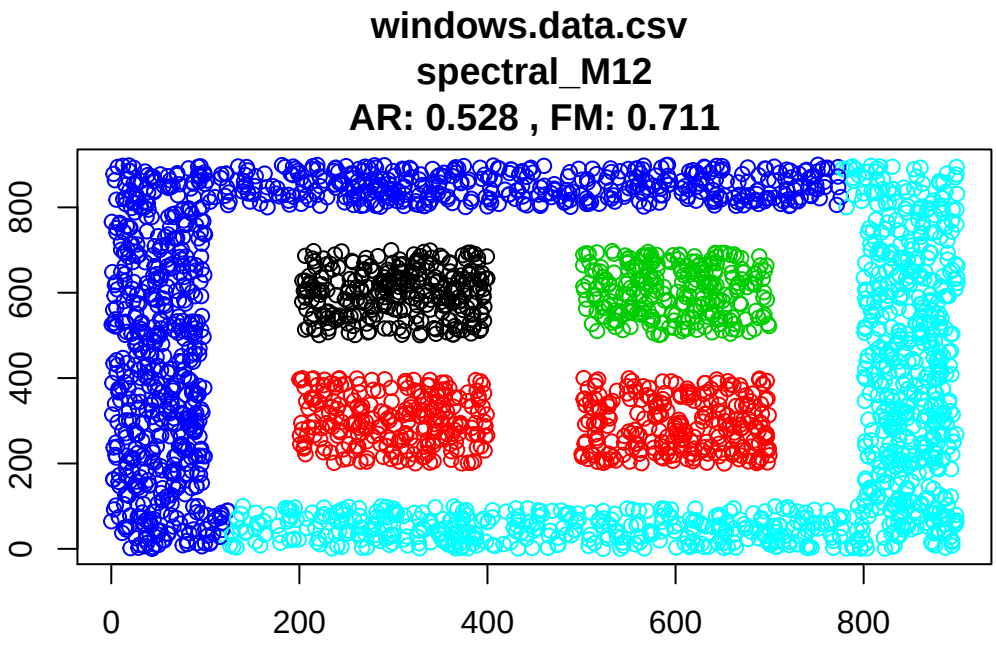
Zbiór windows

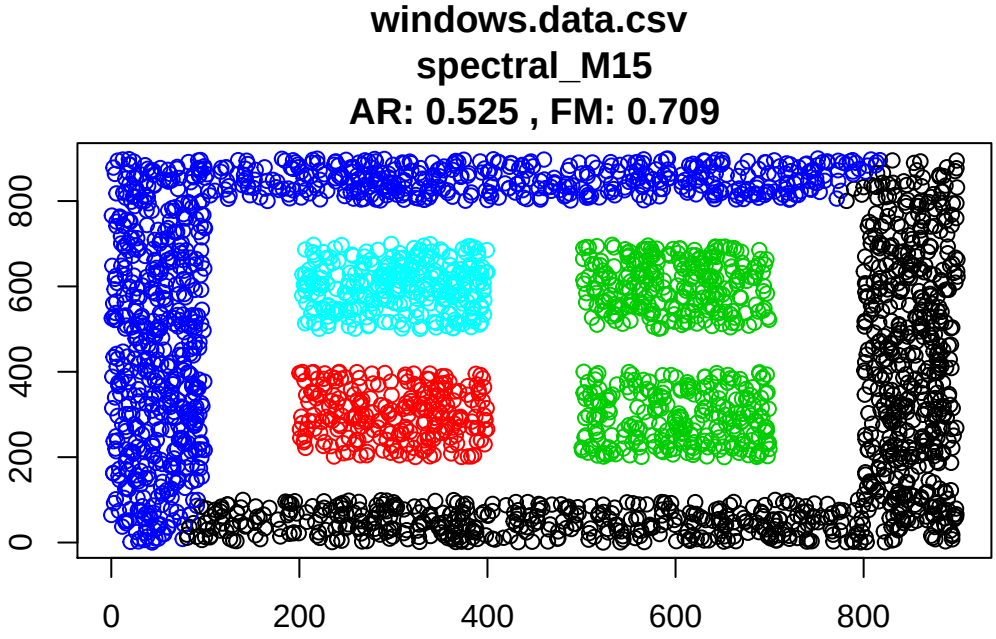
Algorithm	RA	FM
genie_0.5	1.000	1.000
genie_0.8	1.000	1.000
hclust_single	1.000	1.000
spectral_M12	0.528	0.711
spectral_M8	0.527	0.711
spectral_M15	0.525	0.709
spectral_M20	0.435	0.654
spectral_Mk	0.364	0.589
spectral_M10	0.344	0.575
spectral_M30	0.311	0.550
genie_0.2	0.201	0.498
genie_0.3	0.201	0.498
spectral_M5	0.188	0.454
hclust_wardD2	0.103	0.395
fuzzy_10	0.099	0.404
hclust_average	0.099	0.386
hclust_median	0.099	0.400
hclust_wardD	0.097	0.387
hclust_complete	0.090	0.393
hclust_mcquitty	0.089	0.383
hclust_centroid	0.083	0.380
hclust_centroid2	0.082	0.380
kmeans	0.081	0.371
fuzzy_5	0.080	0.372
spectral_M2	0.078	0.366
fuzzy_2	0.074	0.365
fuzzy_default	0.074	0.366

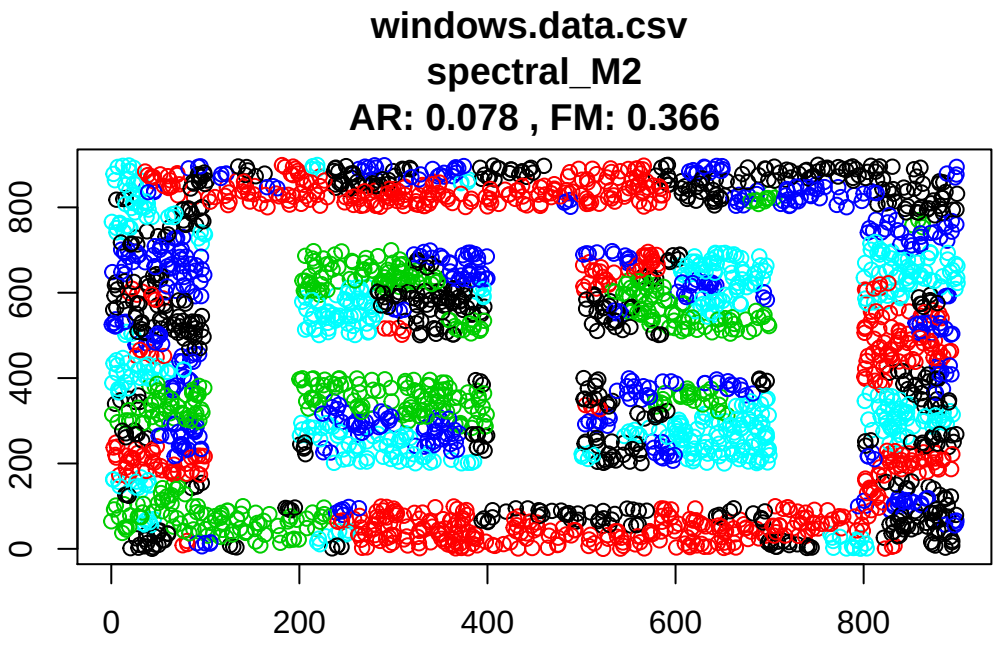
Poza grupą trzech algorytmów, które znalazły poprawne rozwiązanie, większość algorytmów poradziła sobie z tym zbiorem bardzo słabo. Algorytmy spektralne są jednak wyraźnie najlepsze sposród tej grupy.

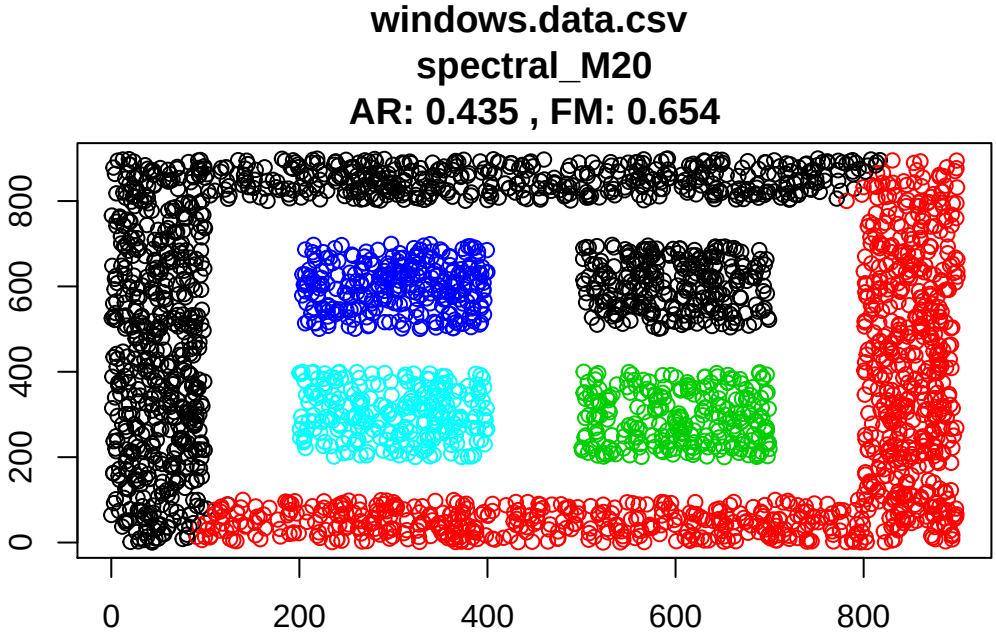

windows.data.csv
expert
AR: 1 , FM: 1
0
200
400
600
800

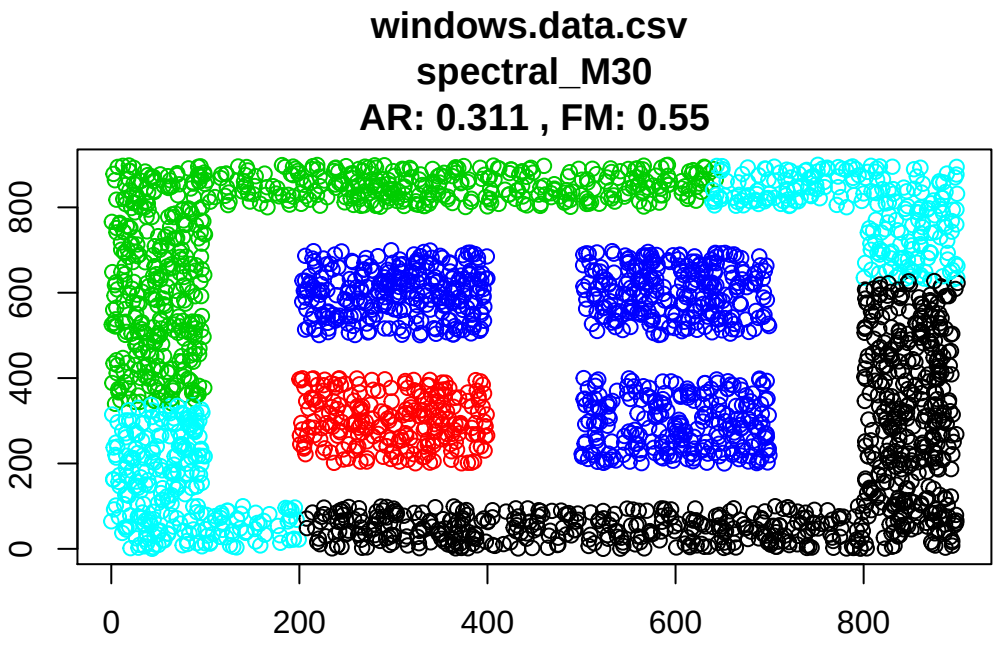

windows.data.csv
spectral_M10
AR: 0.344 , FM: 0.575
0
200
400
600
800

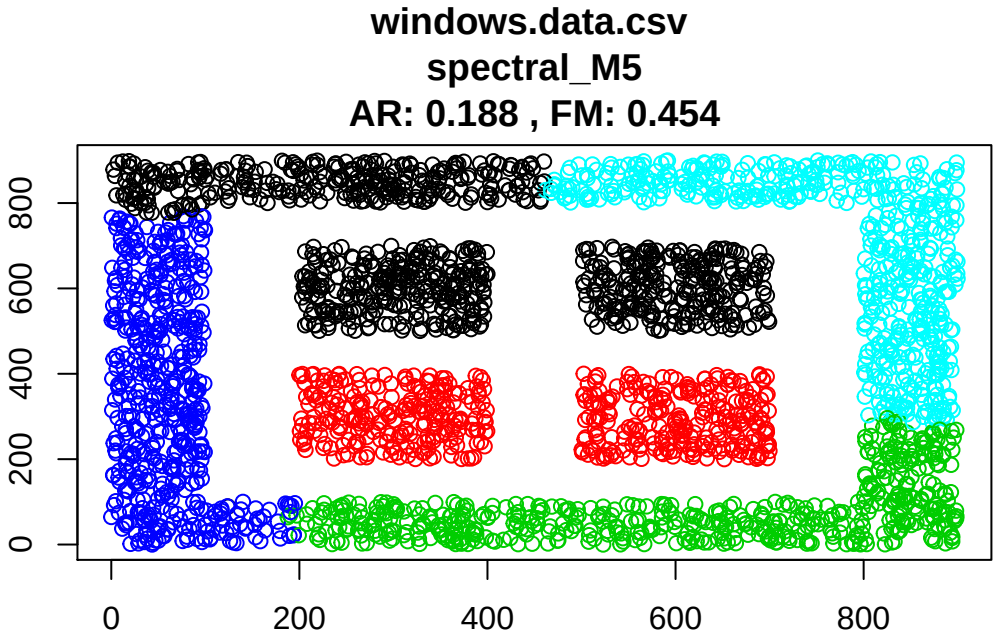

windows.data.csv
spectral_M12
AR: 0.528 , FM: 0.711
0
200
400
600
800

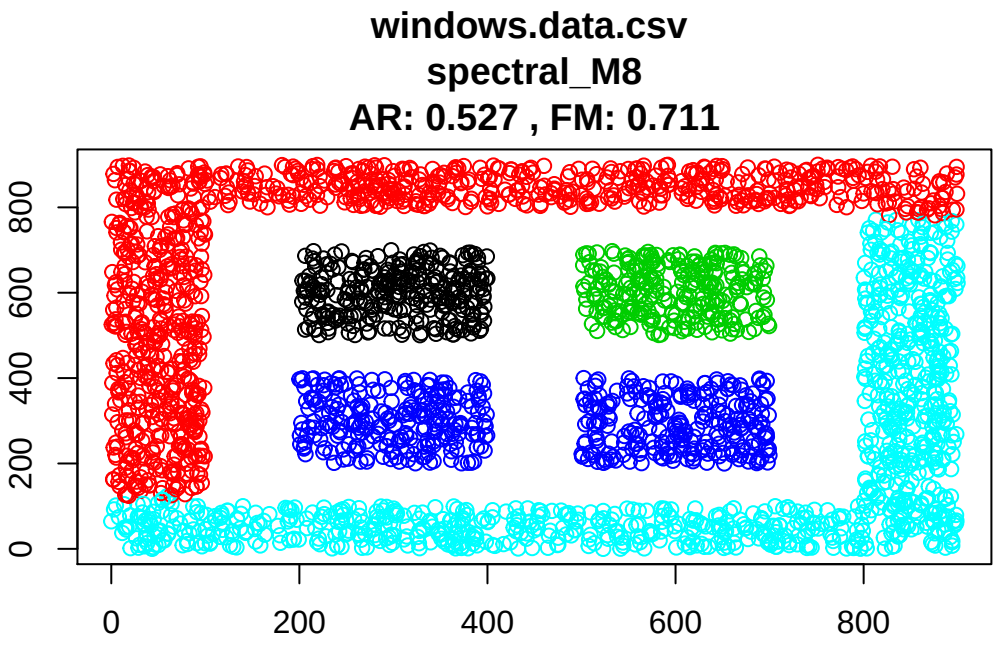

windows.data.csv
spectral_M15
AR: 0.525 , FM: 0.709
0
200
400
600
800

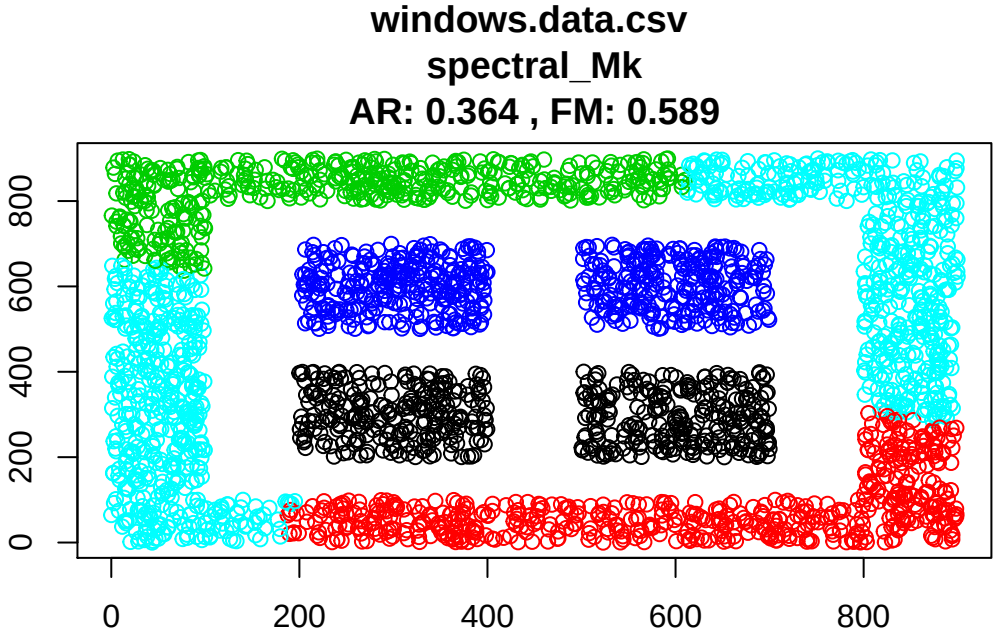

windows.data.csv
spectral_M2
AR: 0.078 , FM: 0.366
0
200
400
600
800


windows.data.csv
spectral_M20
AR: 0.435 , FM: 0.654
0
200
400
600
800


windows.data.csv
spectral_M30
AR: 0.311 , FM: 0.55
0
200
400
600
800


windows.data.csv
spectral_M5
AR: 0.188 , FM: 0.454
0
200
400
600
800


windows.data.csv
spectral_M8
AR: 0.527 , FM: 0.711
0
200
400
600
800


windows.data.csv
spectral_Mk
AR: 0.364 , FM: 0.589
0
200
400
600
800