
TEORETYCZNE MODELE SIECIOWE

Karol Działowski
Zachodniopomorski Uniwersytet Technologiczny

14 października 2020

1 Zapoznanie się z podstawowymi funkcjami biblioteki iGraph w pakiecie R

1.1 Instalacja biblioteki

Zainstalowano bibliotekę przy użyciu komendy `install.packages("igraph")`.

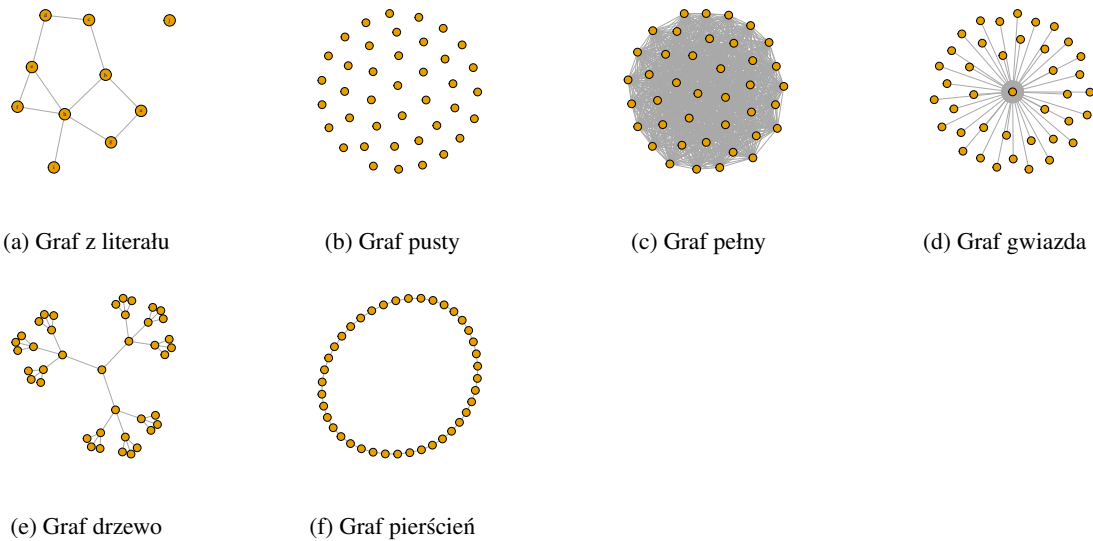
```
1 > install.packages("igraph")
2 WARNING: Rtools is required to build R packages but is not currently installed. Please download and
  install the appropriate version of Rtools before proceeding:
3
4 https://cran.rstudio.com/bin/windows/Rtools/
5 Installing package into C:/Users/karol/Documents/R/win-library/4.0
6 (as lib is unspecified)
7 prbowanie adresu URL https://cran.rstudio.com/bin/windows/contrib/4.0/igraph_1.2.6.zip
8 Content type application/zip length 9345285 bytes (8.9 MB)
9 downloaded 8.9 MB
10
11 package igraph successfully unpacked and MD5 sums checked
12
13 The downloaded binary packages are in
14 C:\Users\karol\AppData\Local\Temp\RtmpEnKdHo\downloaded_packages
```

1.2 Zapis i odczyt sieci z plików

```
1 library(igraph)
2 net <- read_graph("./data/facebook_combined.txt", format = "edgelist")
3 write_graph(net, "./data/out/file2g.txt", "edgelist")
```

1.3 Generowanie i wizualizacja prostych grafów

```
1 gl <- graph_from_literal(a-b-c-d-e-f, a-g-h-b, h-e:f:i, j)
2 plot(gl)
3 eg <- make_empty_graph(40)
4 plot(eg, vertex.size=10, vertex.label=NA)
5 fg <- make_full_graph(40)
6 plot(fg, vertex.size=10, vertex.label=NA)
7 st <- make_star(40)
8 plot(st, vertex.size=10, vertex.label=NA)
9 tr <- make_tree(40, children = 3, mode = "undirected")
10 plot(tr, vertex.size=10, vertex.label=NA)
11 rn <- make_ring(40)
12 plot(rn, vertex.size=10, vertex.label=NA)
13 er <- sample_gnm(n=100, m=40)
14 plot(er, vertex.size=6, vertex.label=NA)
```



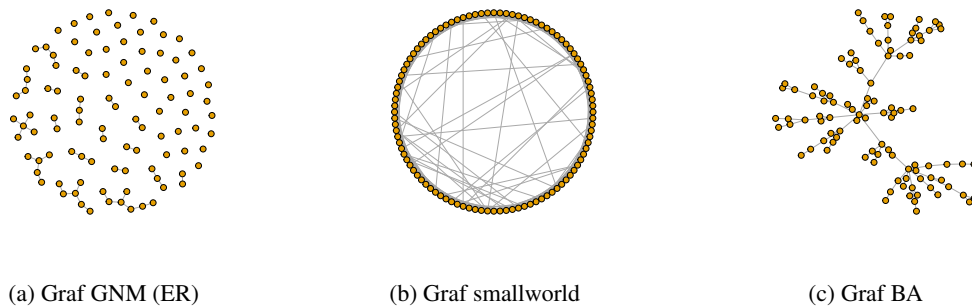
Rysunek 1: Wygenerowane proste grafy

1.4 Generowanie i wizualizacja sieci syntetycznych ER BA WS

```

1 er <- sample_gnm(n=100, m=40)
2 plot(er, vertex.size=6, vertex.label=NA)
3 sw <- sample_smallworld(dim=2, size=10, nei=1, p=0.1)
4 plot(sw, vertex.size=6, vertex.label=NA, layout=layout_in_circle)
5 ba <- sample_pa(n=100, power=1, m=1, directed=F)
6 plot(ba, vertex.size=6, vertex.label=NA)

```



Rysunek 2: Sieci syntetyczne

1.5 Układy grafów

Pakiet *igraph* dostarcza różne metody wzajemnego ułożenia wierzchołków i krawędzi. Poniżej przedstawiono różne układy dla grafu wygenerowanego modelem Barabasi-Albert.

```

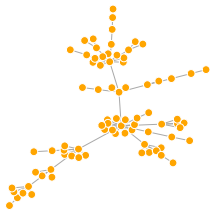
1 net.bg <- sample_pa(80)
2 V(net.bg)$size <- 8
3 V(net.bg)$frame.color <- "white"
4 V(net.bg)$color <- "orange"
5 V(net.bg)$label <- ""
6 E(net.bg)$arrow.mode <- 0
7 plot(net.bg)
8 l <- layout_in_circle(net.bg)

```

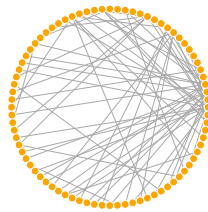
```

9 plot(net.bg, layout=1)
10 l <- layout_on_sphere(net.bg)
11 plot(net.bg, layout=1)
12 l <- layout_with_fr(net.bg)
13 plot(net.bg, layout=1)
14 l <- layout_with_kk(net.bg)
15 plot(net.bg, layout=1)

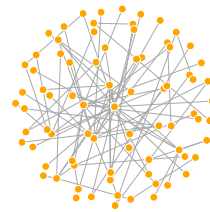
```



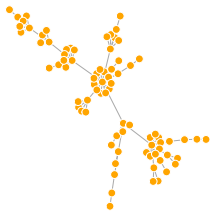
(a) Układ domyślny



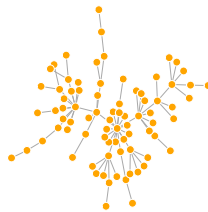
(b) Układ kołowy



(c) Układ sfery 3D



(d) Układ Fruchterman-Reingold



(e) Układ Kamada Kawai

Rysunek 3: Układy grafów

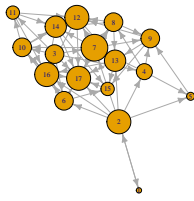
1.6 Stopnie wierzchołków i rozkłady

Stopień wężła to liczba połączeń wężła do innych wężłów. Analiza stopni wierzchołków pozwala nam wyekstrahować wiedzę z grafu.

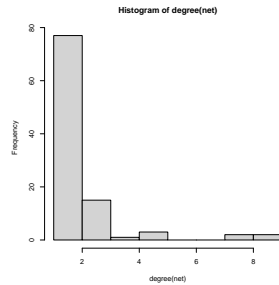
```

1 net <- read.graph("./data/test_16.txt", format = "edgelist")
2 deg <- degree(net, mode="all")
3 plot(net, vertex.size=deg*3)
4
5 net <- sample_pa(n=100, power=1, m=1, directed=F)
6 hist(degree(net))
7
8 net <- sample_pa(n=100, power=1, m=1, directed=F)
9 deg <- degree(net, mode="all")
10 deg.dist <- degree.distribution(net, cumulative=T)
11 plot(degree.distribution(net, cumulative=T), pch=20, xlab="degree", ylab="cumulative frequency",
      type="l")

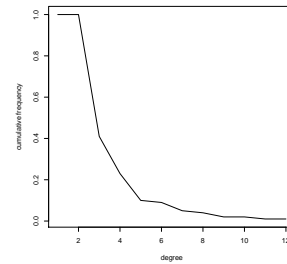
```



(a) Stopnie węzłów przedstawienie na grafie



(b) Histogram stopni węzłów



(c) Rozkład stopni wierzchołków

Rysunek 4: Stopnie węzłów i rozkłady

2 Badania sieci rzeczywistej

Celem zadania było wybranie z repozytoriów dowolnej sieci o rozmiarze około 100-200 węzłów i zaimportowanie jej przy użyciu *igraph*. Następnie należało przygotować wizualizację z wykorzystaniem wybranych algorytmów, np. Fruchterman-Reingold i Kamada Kawai oraz wygenerować wykres rozkładu stopni wierzchołków. Wykres i histogram zapisać do plików graficznych w formacie png.

W tym celu wybrano bazę *MAMMALIA-VOLES-BHP-TRAPPING-26*[1]. Reprezentuje ona sieć karczowników, małych ssaków podobnych do chomików. Zdjęcie karczownika przedstawiono na rysunku 5. Krawędzie w sieci były tworzone za każdym razem gdy dwa karczowniki były złapane w co najmniej jednej wspólnej klatce podczas jednej sesji łapania karczowników. Sieć przedstawia relacje „społeczne” pomiędzy karczownikami.



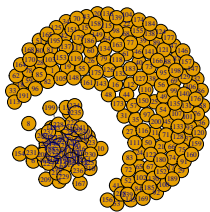
Rysunek 5: Karczownik.

```
1 library(igraph)
2
3 net <- read.graph("./data/voles/mammalia-voles-rob-trapping-11.edges", format = "edgelist")
4 plot(net)
5
6 l <- layout_with_fr(net)
7 plot(net, layout=l)
```

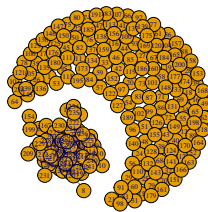
```

8
9 l <- layout_with_kk(net)
10 plot(net, layout=l)
11
12 deg <- degree(net, mode="all")
13 plot(net, vertex.size=deg*3)
14
15 hist(degree(net))
16
17 deg <- degree(net, mode="all")
18 deg.dist <- degree.distribution(net, cumulative=T)
19 plot(degree.distribution(net, cumulative=T), pch=20,xlab="degree", ylab="cumulative frequency",
      type="l")

```



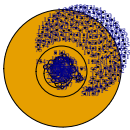
(a) Układ domyślny



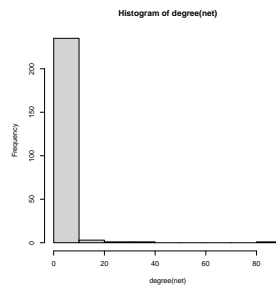
(b) Układ Fruchterman-Reingold



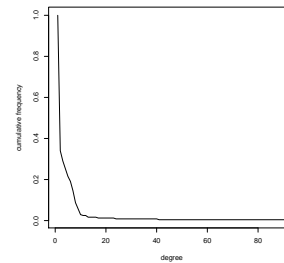
(c) Układ Kamada Kawai



(d) Stopień węzłów.



(e) Histogram stopni.



(f) Rozkład stopni wierzchołków.

Rysunek 6: Analiza sieci karczowników.

Z danych można zaobserwować, że:

- większość karczowników nigdy nie została złapana z innym karczownikiem (159 z 241 karczowników)
- tylko 6 karczowników miało więcej niż 10 powiązań (komenda: `sum(degree(net)>10, na.rm = TRUE)`)
- część karczowników (społecznych) jest odseparowana od karczowników nieśmiały
- najbardziej popularny karczownik miał 90 „partnerów”

3 Badania na sieciach syntetycznych

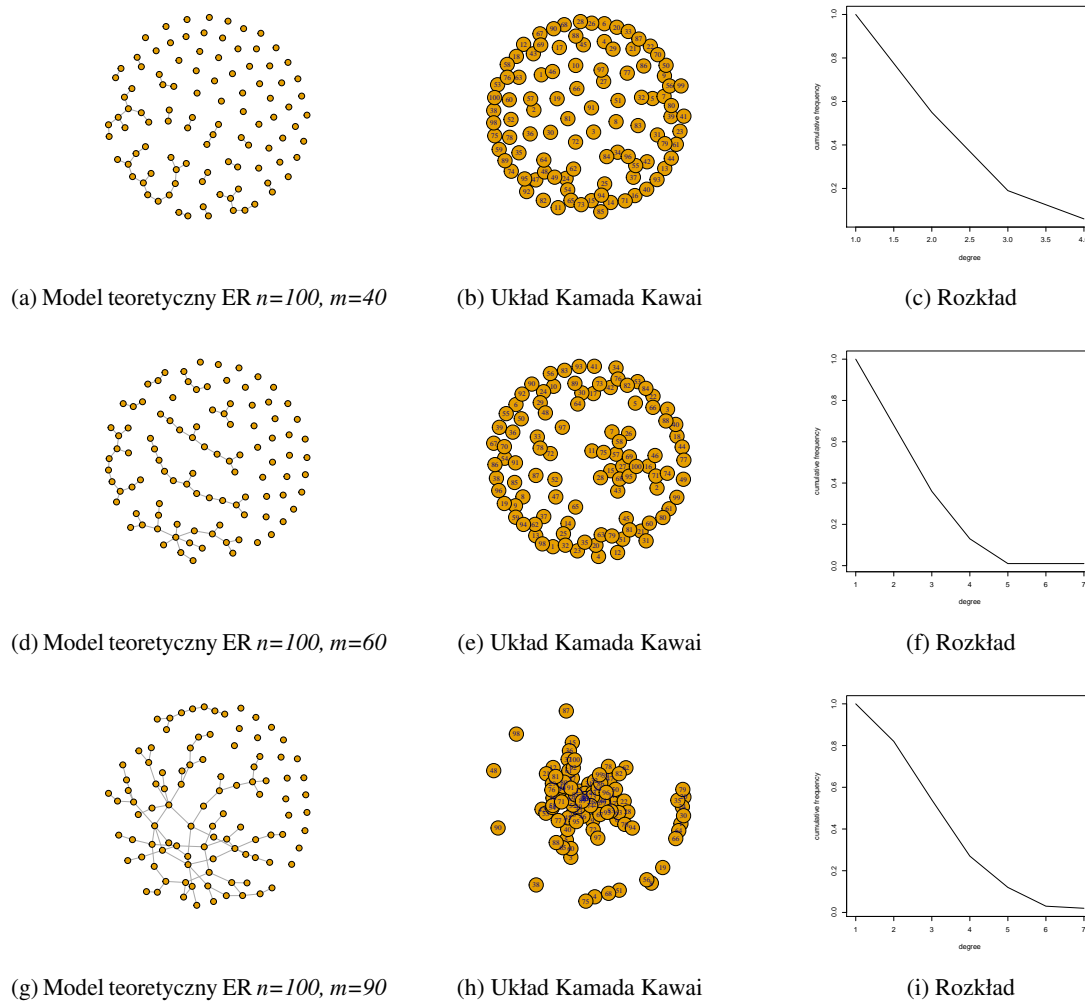
Trzecie zadanie składało się z trzech kroków:

1. Wygenerować po trzy sieci syntetyczne dla każdego z modeli teoretycznego ER BA WS z wykorzystaniem dowolnie wybranych zestawów parametrów dla każdego typu sieci i zapisać do plików np. ba1.txt, ba2.txt, ba3.txt
2. Zwizualizować każdą z sieci i zastosować wybrany algorytm do optymalizacji układu grafu. Zapisać wykresy do plików.

3. Dla każdej z 9 sieci wygenerować wykresy rozkładu stopni wierzchołków.

Należało też podsumować to zadanie opisem i interpretacją wyników. Przykładowo jak parametry sieci syntetycznych wpływają na charakterystyki uzyskanych grafów i na rozkłady stopni wierzchołków. Kod z rozwiązaniem znajduje się w appendiksie w listingu 1.

3.1 Model ER

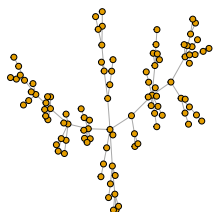


Rysunek 7: Model teoretyczny Erdos-Renyi

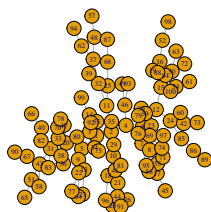
3.1.1 Obserwacje

Parametr m odpowiada za liczbę krawędzi w grafie. Wraz z wzrostem liczby krawędzi układ Kamada-Kawai przesuwa węzły o większym stopniu do centrum grafu. Na rozkładzie widoczny jest trend zwiększania liczby połączeń pomiędzy węzłami. W grafie wygenerowanym z pierwszymi parametrami mamy do czynienia z maksymalnym węzłem 4 stopnia, w drugim grafie z maksymalnym węzłem 5 stopnia, w ostatnim grafie z maksymalnym węzłem 7 stopnia.

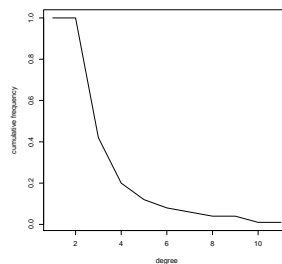
3.2 Model BA



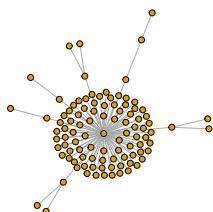
(a) Model teoretyczny BA $n=100$, $power=1$, $m=1$



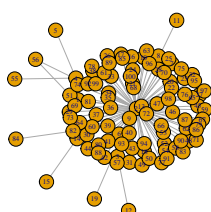
(b) Układ Kamada Kawai



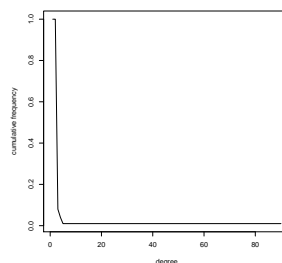
(c) Rozkład



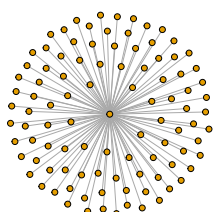
(d) Model teoretyczny BA $n=100$, $power=2$, $m=1$



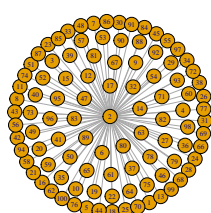
(e) Układ Kamada Kawai



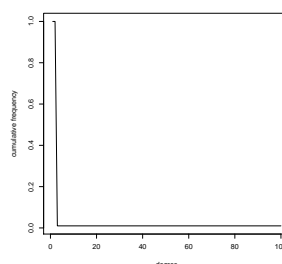
(f) Rozkład



(g) Model teoretyczny BA $n=100$, $power=3$, $m=1$



(h) Układ Kamada Kawai



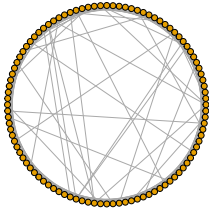
(i) Rozkład

Rysunek 8: Model teoretyczny Barabasi-Albert Model

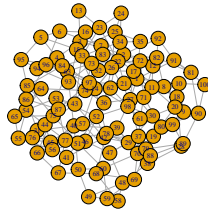
3.2.1 Obserwacje

Parametr *power* odpowiada za moc przywiązania preferencyjnego. Można to zinterpretować jako skupienie grafu. Wraz z zwiększaniem mocy wyłaniają się węzły, będące punktem centralnym. W takim przypadku rozkład ulega degradacji i sieć przybiera postać gwiazdy.

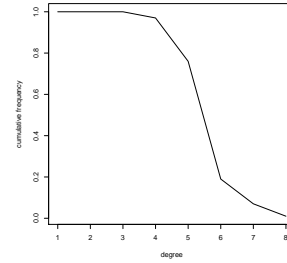
3.3 Model WS



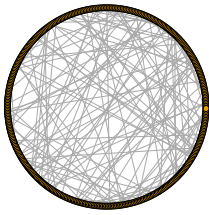
(a) Model teoretyczny WS $dim=2$, $size=10$, $nei=1$, $p=0.1$



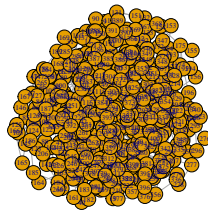
(b) Układ Kamada Kawai



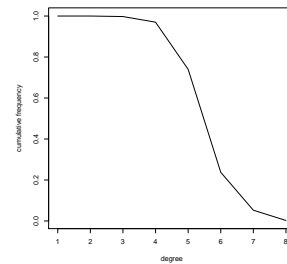
(c) Rozkład



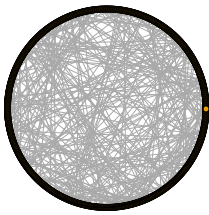
(d) Model teoretyczny WS $dim=2$, $size=20$, $nei=1$, $p=0.1$



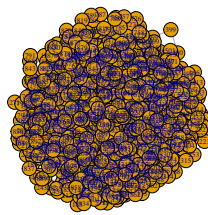
(e) Układ Kamada Kawai



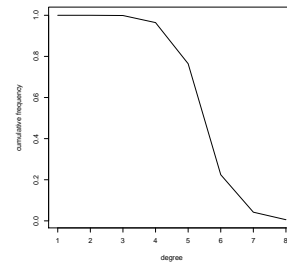
(f) Rozkład



(g) Model teoretyczny WS $dim=2$, $size=30$, $nei=1$, $p=0.1$



(h) Układ Kamada Kawai



(i) Rozkład

Rysunek 9: Model teoretyczny Watts-Strogatz (smallworld)

3.3.1 Obserwacje

W modelu smallworld parametr *size* odpowiada za wielkość siatki wzdłuż każdego wymiaru. To znaczy, że jest więcej połączeń pomiędzy skupiskami (małymi światami). Układ Kamada-Kawai w przypadku tego modelu jest nieczytelny. Wraz z zwiększaniem parametru *size* rozkład stopni wierzchołków nie zmienia się.

4 Podsumowanie

Podczas laboratoriów przećwiczono posługiwanie się biblioteką *igraph*. Przeprowadzono analizę na sieci rzeczywistej oraz na zestawie sieci wygenerowanych.

Literatura

- [1] Ryan A. Rossi and Nesreen K. Ahmed. The network data repository with interactive graph analytics and visualization. In *AAAI*, 2015.

A Kod do zadania 3

```

1 library(igraph)
2
3 pdf(file="./data/out/3/er1.pdf")
4 net <- sample_gnm(n=100, m=40)
5 plot(net, vertex.size=6, vertex.label=NA)
6 write_graph(net, "./data/out/3/er1.txt", "edgelist")
7 dev.off()
8 pdf(file="./data/out/3/er1kk.pdf")
9 l <- layout_with_kk(net)
10 plot(net, layout=l)
11 dev.off()
12 pdf(file="./data/out/3/er1dist.pdf")
13 deg <- degree(net, mode="all")
14 deg.dist <- degree.distribution(net, cumulative=T)
15 plot(degree.distribution(net, cumulative=T), pch=20,xlab="degree", ylab="cumulative frequency",
16     type="l")
17 dev.off()
18
19 pdf(file="./data/out/3/er2.pdf")
20 net <- sample_gnm(n=100, m=60)
21 plot(net, vertex.size=6, vertex.label=NA)
22 write_graph(net, "./data/out/3/er2.txt", "edgelist")
23 dev.off()
24 pdf(file="./data/out/3/er2kk.pdf")
25 l <- layout_with_kk(net)
26 plot(net, layout=l)
27 dev.off()
28 pdf(file="./data/out/3/er2dist.pdf")
29 deg <- degree(net, mode="all")
30 deg.dist <- degree.distribution(net, cumulative=T)
31 plot(degree.distribution(net, cumulative=T), pch=20,xlab="degree", ylab="cumulative frequency",
32     type="l")
33 dev.off()
34
35 pdf(file="./data/out/3/er3.pdf")
36 net <- sample_gnm(n=100, m=90)
37 plot(net, vertex.size=6, vertex.label=NA)
38 write_graph(net, "./data/out/3/er3.txt", "edgelist")
39 dev.off()
40 pdf(file="./data/out/3/er3kk.pdf")
41 l <- layout_with_kk(net)
42 plot(net, layout=l)
43 dev.off()
44 pdf(file="./data/out/3/er3dist.pdf")
45 deg <- degree(net, mode="all")
46 deg.dist <- degree.distribution(net, cumulative=T)
47 plot(degree.distribution(net, cumulative=T), pch=20,xlab="degree", ylab="cumulative frequency",
48     type="l")
49 dev.off()
50
51
52 pdf(file="./data/out/3/sw1.pdf")
53 net <- sample_smallworld(dim=2, size=10, nei=1, p=0.1)
54 write_graph(net, "./data/out/3/sw1.txt", "edgelist")
55 plot(net, vertex.size=6, vertex.label=NA, layout=layout_in_circle)
56 dev.off()
57 pdf(file="./data/out/3/sw1kk.pdf")
58 l <- layout_with_kk(net)
59 plot(net, layout=l)

```

```

60 dev.off()
61 pdf(file="./data/out/3/swldist.pdf")
62 deg <- degree(net, mode="all")
63 deg.dist <- degree.distribution(net, cumulative=T)
64 plot(degree.distribution(net, cumulative=T), pch=20,xlab="degree", ylab="cumulative frequency",
      type="l")
65 dev.off()
66
67
68 pdf(file="./data/out/3/sw2.pdf")
69 net <- sample_smallworld(dim=2, size=20, nei=1, p=0.1)
70 write_graph(net, "./data/out/3/sw2.txt", "edgelist")
71 plot(net, vertex.size=6, vertex.label=NA, layout=layout_in_circle)
72 dev.off()
73 pdf(file="./data/out/3/sw2kk.pdf")
74 l <- layout_with_kk(net)
75 plot(net, layout=l)
76 dev.off()
77 pdf(file="./data/out/3/sw2dist.pdf")
78 deg <- degree(net, mode="all")
79 deg.dist <- degree.distribution(net, cumulative=T)
80 plot(degree.distribution(net, cumulative=T), pch=20,xlab="degree", ylab="cumulative frequency",
      type="l")
81 dev.off()
82
83
84 pdf(file="./data/out/3/sw3.pdf")
85 net <- sample_smallworld(dim=2, size=30, nei=1, p=0.1)
86 write_graph(net, "./data/out/3/sw3.txt", "edgelist")
87 plot(net, vertex.size=6, vertex.label=NA, layout=layout_in_circle)
88 dev.off()
89 pdf(file="./data/out/3/sw3kk.pdf")
90 l <- layout_with_kk(net)
91 plot(net, layout=l)
92 dev.off()
93 pdf(file="./data/out/3/sw3dist.pdf")
94 deg <- degree(net, mode="all")
95 deg.dist <- degree.distribution(net, cumulative=T)
96 plot(degree.distribution(net, cumulative=T), pch=20,xlab="degree", ylab="cumulative frequency",
      type="l")
97 dev.off()
98
99
100
101
102
103 pdf(file="./data/out/3/ba1.pdf")
104 net <- sample_pa(n=100, power=1, m=1, directed=F)
105 write_graph(net, "./data/out/3/ba1.txt", "edgelist")
106 plot(net, vertex.size=6, vertex.label=NA)
107 dev.off()
108 pdf(file="./data/out/3/ba1kk.pdf")
109 l <- layout_with_kk(net)
110 plot(net, layout=l)
111 dev.off()
112 pdf(file="./data/out/3/ba1dist.pdf")
113 deg <- degree(net, mode="all")
114 deg.dist <- degree.distribution(net, cumulative=T)
115 plot(degree.distribution(net, cumulative=T), pch=20,xlab="degree", ylab="cumulative frequency",
      type="l")
116 dev.off()
117
118
119 pdf(file="./data/out/3/ba2.pdf")
120 net <- sample_pa(n=100, power=2, m=1, directed=F)

```

```

121 write_graph(net, "./data/out/3/ba2.txt", "edgelist")
122 plot(net, vertex.size=6, vertex.label=NA)
123 dev.off()
124 pdf(file="./data/out/3/ba2kk.pdf")
125 l <- layout_with_kk(net)
126 plot(net, layout=l)
127 dev.off()
128 pdf(file="./data/out/3/ba2dist.pdf")
129 deg <- degree(net, mode="all")
130 deg.dist <- degree.distribution(net, cumulative=T)
131 plot(degree.distribution(net, cumulative=T), pch=20,xlab="degree", ylab="cumulative frequency",
      type="l")
132 dev.off()
133
134 pdf(file="./data/out/3/ba3.pdf")
135 net <- sample_pa(n=100, power=3, m=1, directed=F)
136 write_graph(net, "./data/out/3/ba3.txt", "edgelist")
137 plot(net, vertex.size=6, vertex.label=NA)
138 dev.off()
139 pdf(file="./data/out/3/ba3kk.pdf")
140 l <- layout_with_kk(net)
141 plot(net, layout=l)
142 dev.off()
143 pdf(file="./data/out/3/ba3dist.pdf")
144 deg <- degree(net, mode="all")
145 deg.dist <- degree.distribution(net, cumulative=T)
146 plot(degree.distribution(net, cumulative=T), pch=20,xlab="degree", ylab="cumulative frequency",
      type="l")
147 dev.off()

```

Listing 1: Badania sieci syntetycznych.