ba-networks

Robert Jankowski

15 04 2020

```
load_file <- function(filename) {
  df <- read.csv(filename, header = FALSE)
  colnames(df) <- c("degree")
  df
}

log_scale <- function(x, n) exp(seq(log(x[1]), log(x[length(x)]), length.out = n))

fit_powerlaw <- function(ba, bins = 20) {
  log_breaks <- seq(min(ba$degree), max(ba$degree), length.out = bins)
  log_breaks <- log_scale(log_breaks, bins)
  h_degree <- hist(ba$degree, breaks = log_breaks, plot = FALSE)

  degree_fit <- lm(log(h_degree$density)~log(h_degree$mids))

  A <- degree_fit$coefficients[1]
  alpha <- degree_fit$coefficients[2]
  list(x = h_degree$mids,
       y = h_degree$density,
       log_breaks = log_breaks,
       A = A,
       alpha = alpha)
}
```

Degree distribution

```
plot_degree_distribution_m <- function(fit_ba_models, size) {
  ba_m_1 <- fit_ba_models[[1]]
  ba_m_2 <- fit_ba_models[[2]]
  ba_m_5 <- fit_ba_models[[3]]

  y_max <- max(ba_m_1$y, ba_m_2$y, ba_m_5$y)
  y_min <- min(ba_m_1$y, ba_m_2$y, ba_m_5$y)
  x_max <- max(ba_m_1$x, ba_m_2$x, ba_m_5$x)
  x_min <- min(ba_m_1$x, ba_m_2$x, ba_m_5$x)

  plot(ba_m_1$x, ba_m_1$y, log = "xy", pch = 19, xlab = "k", ylab = "P(k)",
       col = "red", ylim = c(y_min, y_max), xlim = c(x_min, x_max),
       main = paste("BA degree distribution N = ", size))
  lines(ba_m_1$log_breaks, exp(ba_m_1$A) * ba_m_1$log_breaks ** ba_m_1$alpha,
        col = "red", lty = 1, lwd = 2)

  points(ba_m_2$x, ba_m_2$y, pch = 19,  col = "blue")
  lines(ba_m_2$log_breaks, exp(ba_m_2$A) * ba_m_2$log_breaks ** ba_m_2$alpha,
        col = "blue", lty = 1, lwd = 2)
```

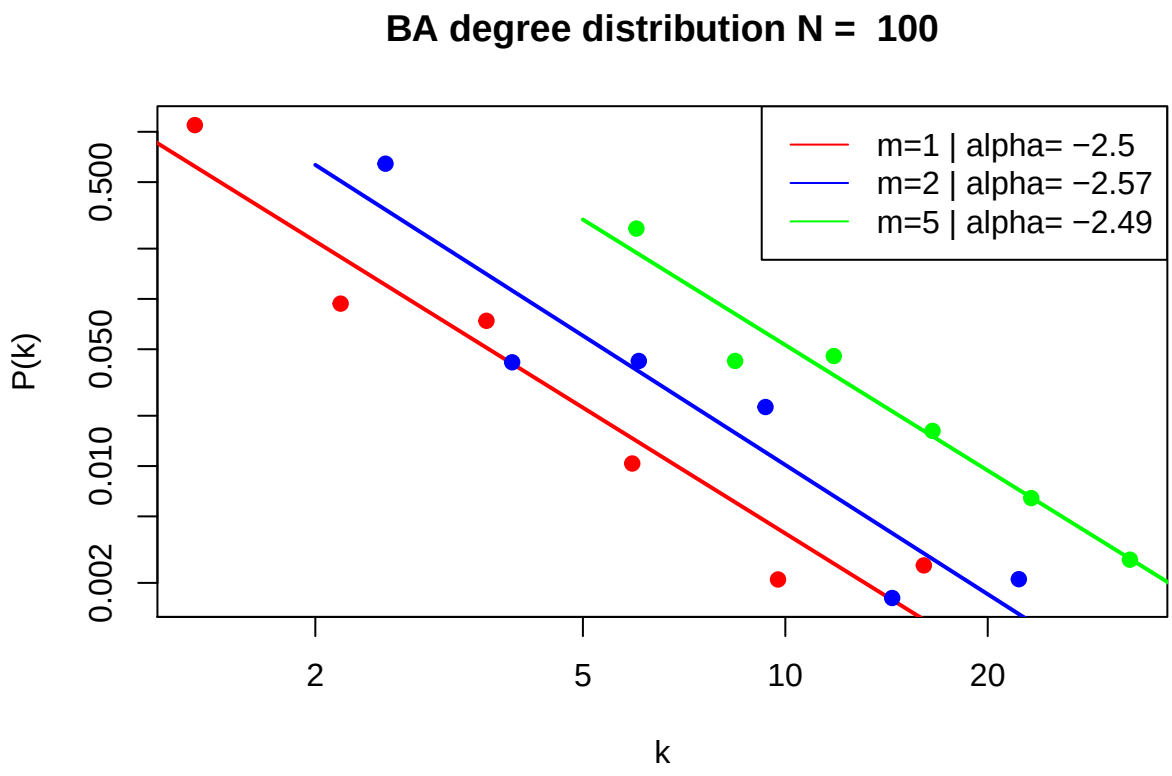
```
  points(ba_m_5$x, ba_m_5$y, pch = 19, col = "green")
  lines(ba_m_5$log_breaks, exp(ba_m_5$A) * ba_m_5$log_breaks ** ba_m_5$alpha,
        col = "green", lty = 1, lwd = 2)

  m_1 <- paste("m=1 | alpha=", round(ba_m_1$alpha, 2))
  m_2 <- paste("m=2 | alpha=", round(ba_m_2$alpha, 2))
  m_5 <- paste("m=5 | alpha=", round(ba_m_5$alpha, 2))
  legend("topright", legend = c(m_1, m_2, m_5), col = c("red", "blue", "green"),
         lty = 1, cex = 1)
}
```

```
# png("ba_n=100.png", width = 600, height = 400)
ba_m_1_n_100 <- load_file("output/ba_degree_n=100_m=1.txt")
ba_m_2_n_100 <- load_file("output/ba_degree_n=100_m=2.txt")
ba_m_5_n_100 <- load_file("output/ba_degree_n=100_m=5.txt")

fit_ba_m_1_n_100 <- fit_powerlaw(ba_m_1_n_100, bins = 7)
fit_ba_m_2_n_100 <- fit_powerlaw(ba_m_2_n_100, bins = 7)
fit_ba_m_5_n_100 <- fit_powerlaw(ba_m_5_n_100, bins = 7)

plot_degree_distribution_m(list(fit_ba_m_1_n_100, fit_ba_m_2_n_100, fit_ba_m_5_n_100), 100)
```



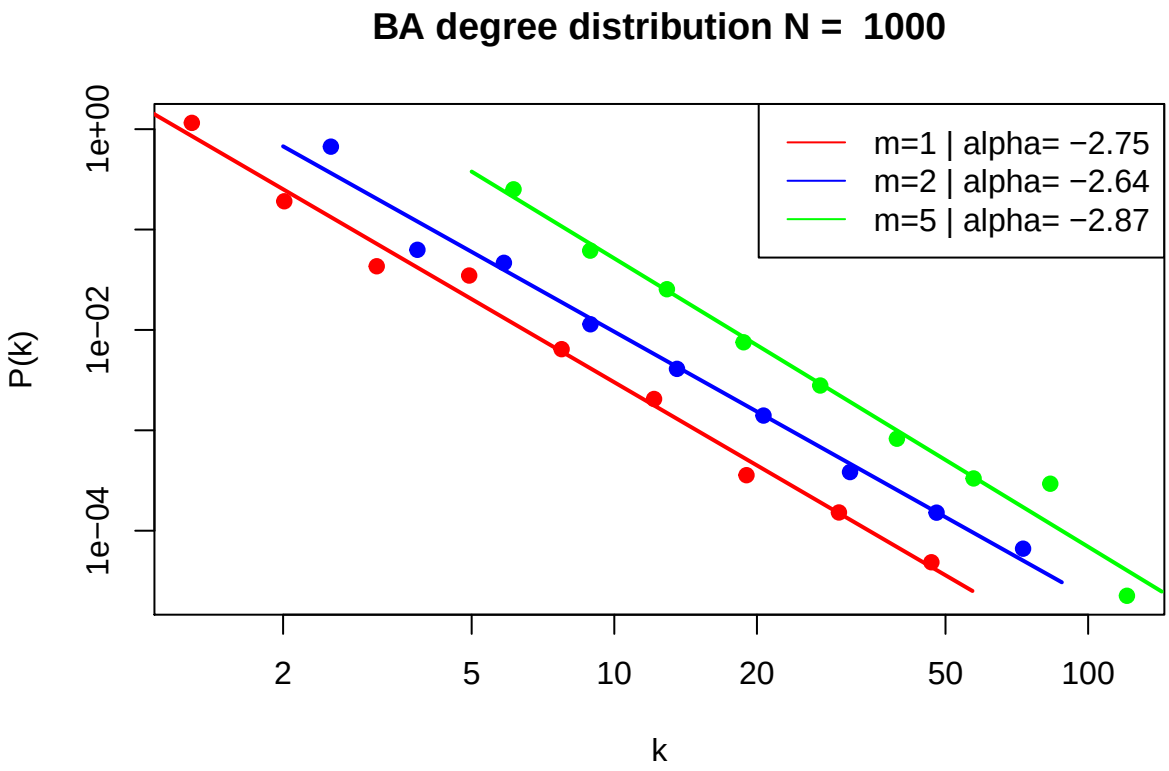

```
# dev.off()
```

```
# png("ba_n=1000.png", width = 600, height = 400)
ba_m_1_n_1000 <- load_file("output/ba_degree_n=1000_m=1.txt")
```

```r
ba_m_2_n_1000 <- load_file("output/ba_degree_n=1000_m=2.txt")
ba_m_5_n_1000 <- load_file("output/ba_degree_n=1000_m=5.txt")

fit_ba_m_1_n_1000 <- fit_powerlaw(ba_m_1_n_1000, bins = 10)
fit_ba_m_2_n_1000 <- fit_powerlaw(ba_m_2_n_1000, bins = 10)
fit_ba_m_5_n_1000 <- fit_powerlaw(ba_m_5_n_1000, bins = 10)

plot_degree_distribution_m(list(fit_ba_m_1_n_1000, fit_ba_m_2_n_1000, fit_ba_m_5_n_1000), 1000)
```

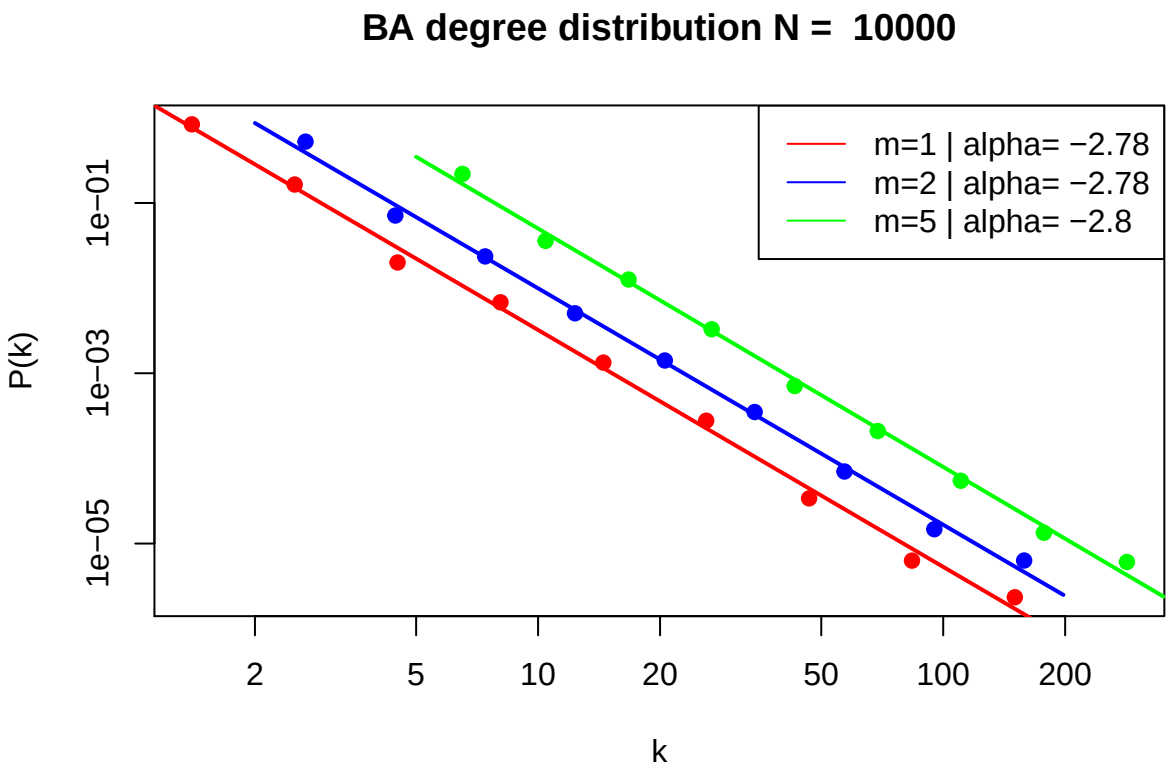



```r
# dev.off()
```

```r
# png("ba_n=10000.png", width = 600, height = 400)
ba_m_1_n_10000 <- load_file("output/ba_degree_n=10000_m=1.txt")
ba_m_2_n_10000 <- load_file("output/ba_degree_n=10000_m=2.txt")
ba_m_5_n_10000 <- load_file("output/ba_degree_n=10000_m=5.txt")

fit_ba_m_1_n_10000 <- fit_powerlaw(ba_m_1_n_10000, bins = 10)
fit_ba_m_2_n_10000 <- fit_powerlaw(ba_m_2_n_10000, bins = 10)
fit_ba_m_5_n_10000 <- fit_powerlaw(ba_m_5_n_10000, bins = 10)

plot_degree_distribution_m(list(fit_ba_m_1_n_10000, fit_ba_m_2_n_10000, fit_ba_m_5_n_10000), 10000)
```

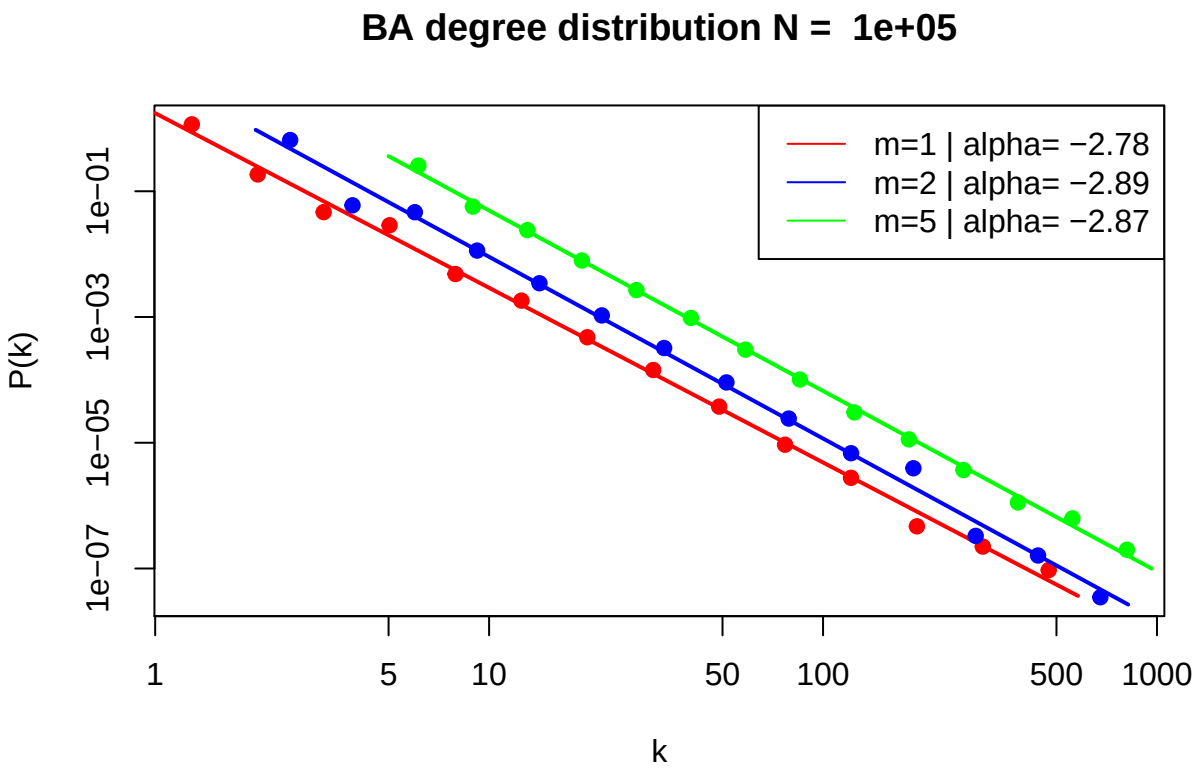



```
# dev.off()
```

```
# png("ba_n=100000.png", width = 600, height = 400)
ba_m_1_n_100000 <- load_file("output/ba_degree_n=100000_m=1.txt")
ba_m_2_n_100000 <- load_file("output/ba_degree_n=100000_m=2.txt")
ba_m_5_n_100000 <- load_file("output/ba_degree_n=100000_m=5.txt")

fit_ba_m_1_n_100000 <- fit_powerlaw(ba_m_1_n_100000, bins = 15)
fit_ba_m_2_n_100000 <- fit_powerlaw(ba_m_2_n_100000, bins = 15)
fit_ba_m_5_n_100000 <- fit_powerlaw(ba_m_5_n_100000, bins = 15)

plot_degree_distribution_m(list(fit_ba_m_1_n_100000, fit_ba_m_2_n_100000, fit_ba_m_5_n_100000), 100000)
```

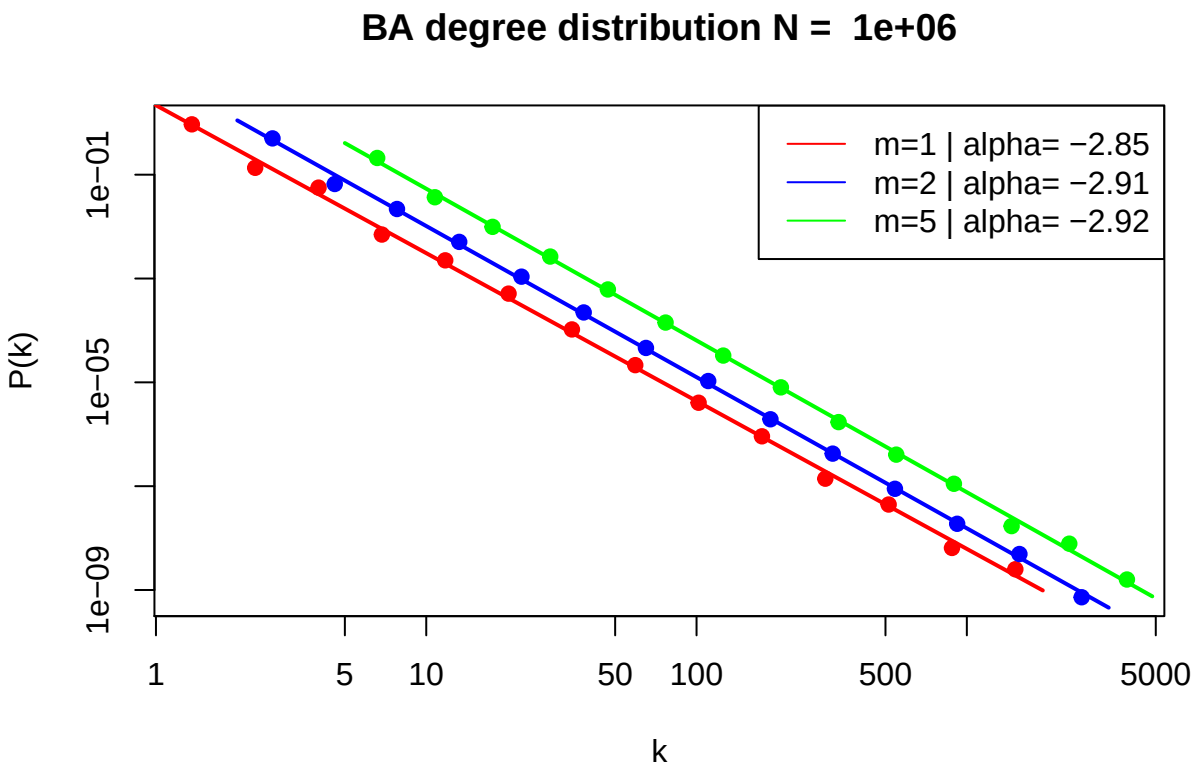



```
# dev.off()
```

```
# png("ba_n=1000000.png", width = 600, height = 400)
ba_m_1_n_1000000 <- load_file("output/ba_degree_n=1000000_m=1.txt")
ba_m_2_n_1000000 <- load_file("output/ba_degree_n=1000000_m=2.txt")
ba_m_5_n_1000000 <- load_file("output/ba_degree_n=1000000_m=5.txt")

fit_ba_m_1_n_1000000 <- fit_powerlaw(ba_m_1_n_1000000, bins = 15)
fit_ba_m_2_n_1000000 <- fit_powerlaw(ba_m_2_n_1000000, bins = 15)
fit_ba_m_5_n_1000000 <- fit_powerlaw(ba_m_5_n_1000000, bins = 15)

plot_degree_distribution_m(list(fit_ba_m_1_n_1000000, fit_ba_m_2_n_1000000, fit_ba_m_5_n_1000000), 1000
```




```
# dev.off()
```

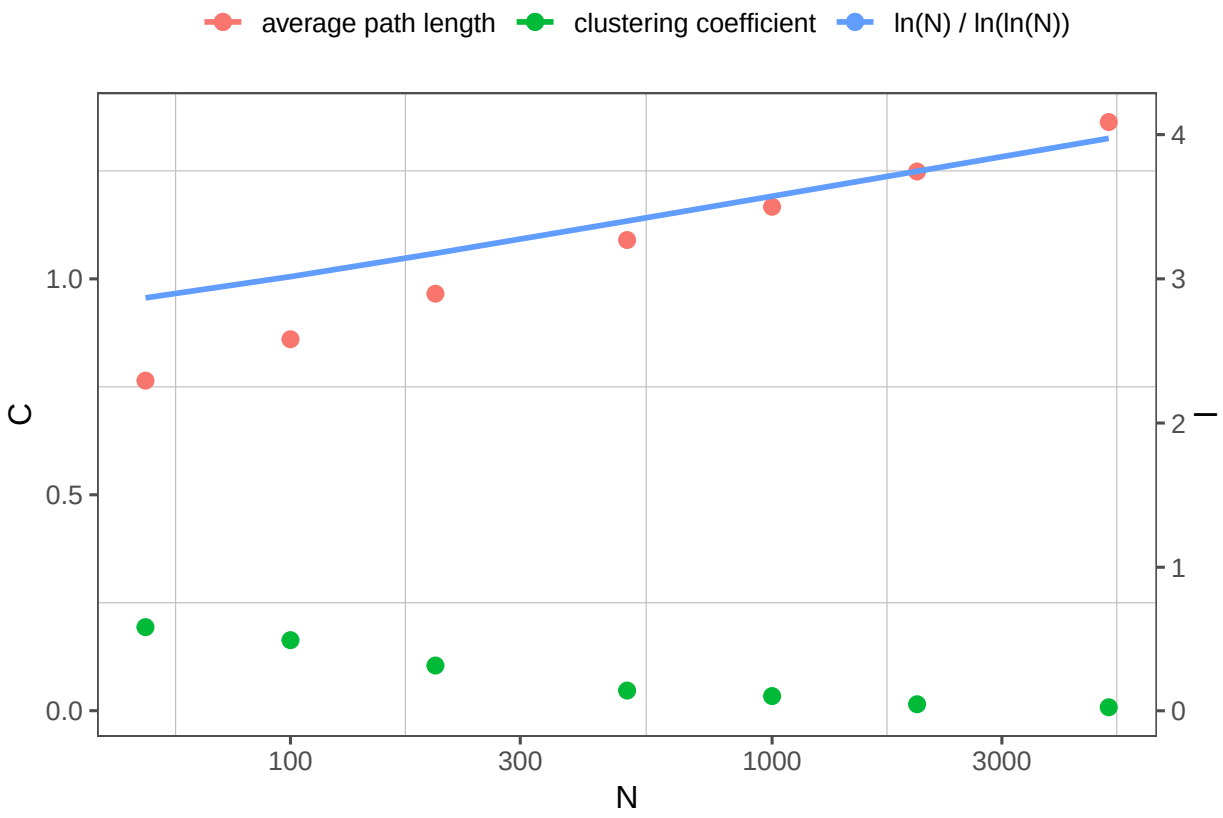
Clustering coefficient

```
library(ggthemes)

clustering <- read.csv("output/ba_simulation_clustering.txt", header = FALSE)
average_path <- read.csv("output/ba_simulation_avg_path.txt", header = FALSE)
colnames(clustering) <- c("C", "N")
colnames(average_path) <- c("l", "N")

path_theory <- data.frame(l_theory=log(average_path$N) / log(log(average_path$N)), N=average_path$N)

ggplot(clustering %>% left_join(average_path) %>% left_join(path_theory)) +
  geom_point(aes(N, C, color = "clustering coefficient"), size = 2.5) +
  geom_point(aes(N, l / 3, color = "average path length"), size = 2.5) +
  geom_line(aes(N, l_theory / 3, color = "ln(N) / ln(ln(N))"), size = 1) +
  scale_y_continuous("C", sec.axis = sec_axis(~ .*3, name = "l")) +
  scale_x_log10() +
  theme_few() +
  labs(color = "") +
  theme(legend.position = "top", panel.grid.minor = element_line(size = 0.2, colour = "grey"))
```




```
ggsave("ba_clustering_average_path.png")
```