ba-networks

Robert Jankowski

23 kwiecień, 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)
}
```

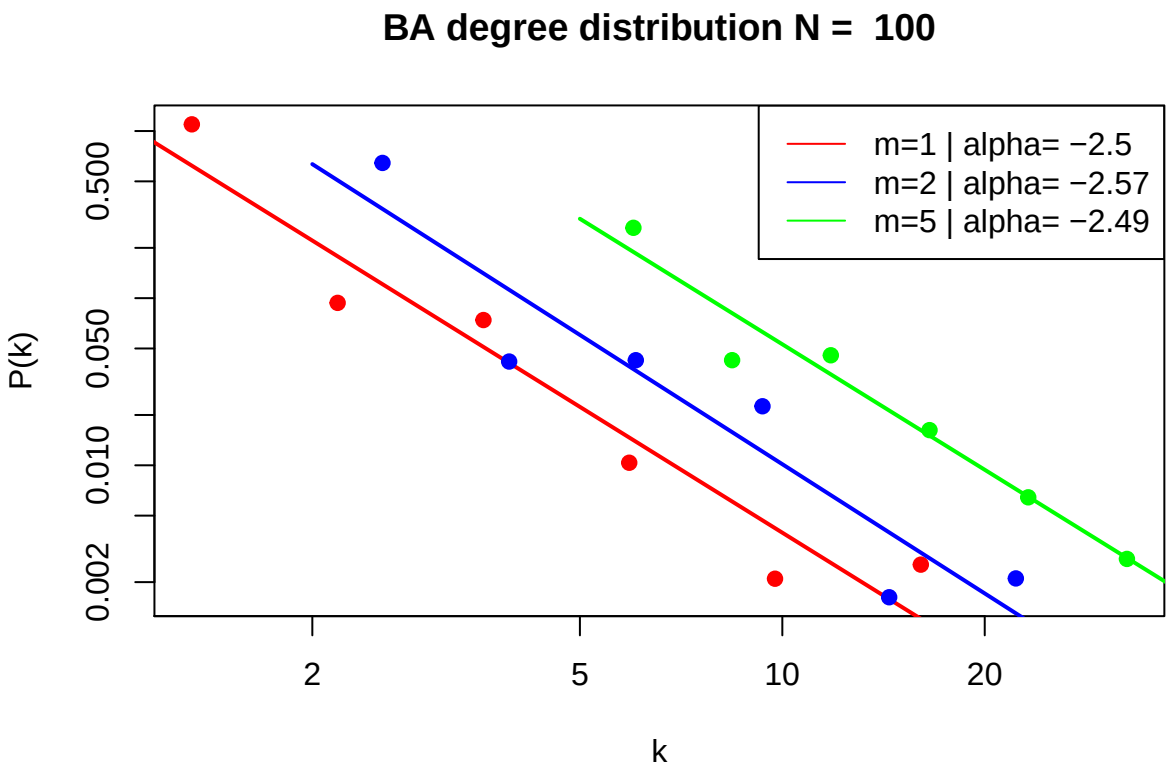
Degree distribution

```
library(zeallot)

min_max <- function(a1, a2, a3) {
    a_min <- min(a1, a2, a3)
    a_max <- max(a1, a2, a3)
    c(a_min, a_max)
}

plot_by_m <- function(ba_m, colour) {
  points(ba_m$x, ba_m$y, pch = 19,  col = colour)
  lines(ba_m$log_breaks, exp(ba_m$A) * ba_m$log_breaks ** ba_m$alpha,
        col = colour, lty = 1, lwd = 2)
}

legend_description <- function(m, alpha) {
  paste("m=", m, " | alpha= ", round(alpha, 2), sep = "")
}

plot_degree_distribution_m <- function(fit_ba_models, size) {
  ba_m_1 <- fit_ba_models$ba_m_1
```

```
  ba_m_2 <- fit_ba_models$ba_m_2
  ba_m_5 <- fit_ba_models$ba_m_5

  c(y_min, y_max) %<-% min_max(ba_m_1$y, ba_m_2$y, ba_m_5$y)
  c(x_min, x_max) %<-% min_max(ba_m_1$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)

  plot_by_m(ba_m_2, "blue")
  plot_by_m(ba_m_5, "green")

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

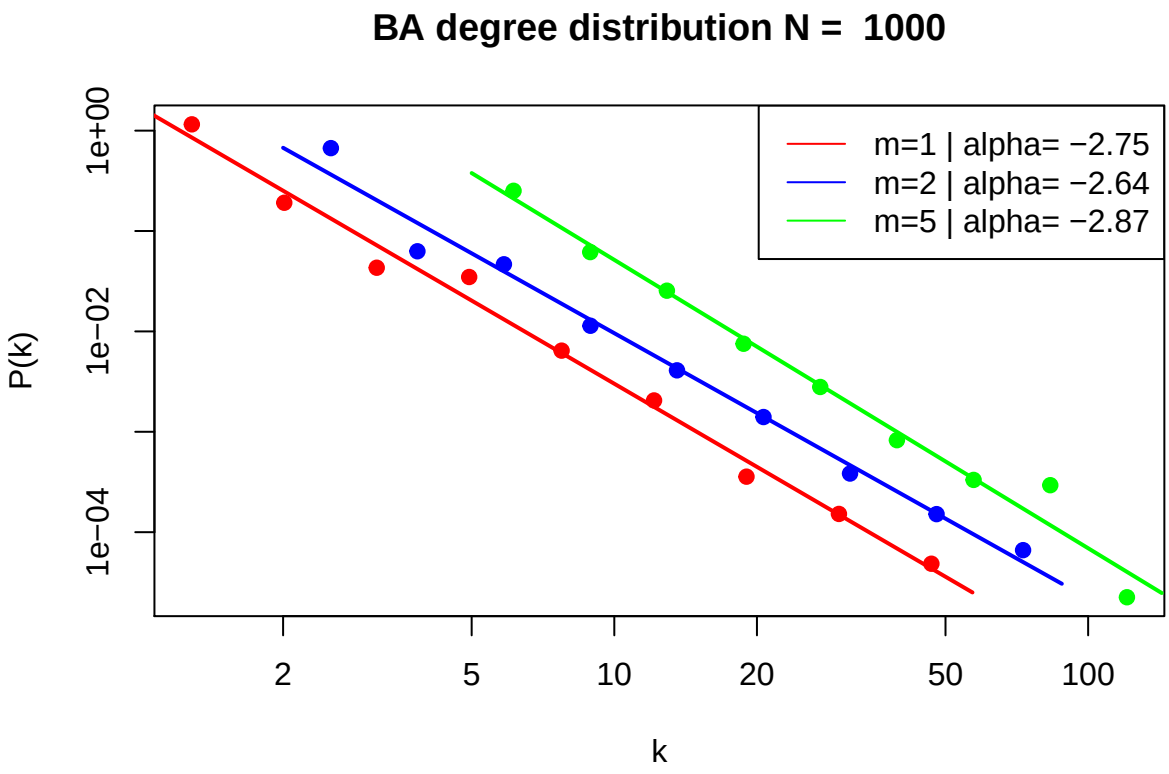
```
get_degree_and_fit <- function(filenames, bins) {
  ba_m_1 <- load_file(filenames$m_1)
  ba_m_2 <- load_file(filenames$m_2)
  ba_m_5 <- load_file(filenames$m_5)

  fit_ba_m_1 <- fit_powerlaw(ba_m_1, bins = bins)
  fit_ba_m_2 <- fit_powerlaw(ba_m_2, bins = bins)
  fit_ba_m_5 <- fit_powerlaw(ba_m_5, bins = bins)

  list(ba_m_1 = fit_ba_m_1,
       ba_m_2 = fit_ba_m_2,
       ba_m_5 = fit_ba_m_5)
}
```

```
# png("ba_n=100.png", width = 600, height = 400)
filenames_n_100 <- list(m_1 = "output/ba_degree_n=100_m=1.txt",
                        m_2 = "output/ba_degree_n=100_m=2.txt",
                        m_5 = "output/ba_degree_n=100_m=5.txt")
fit_ba_n_100 <- get_degree_and_fit(filenames_n_100, 7)

plot_degree_distribution_m(fit_ba_n_100, 100)
```



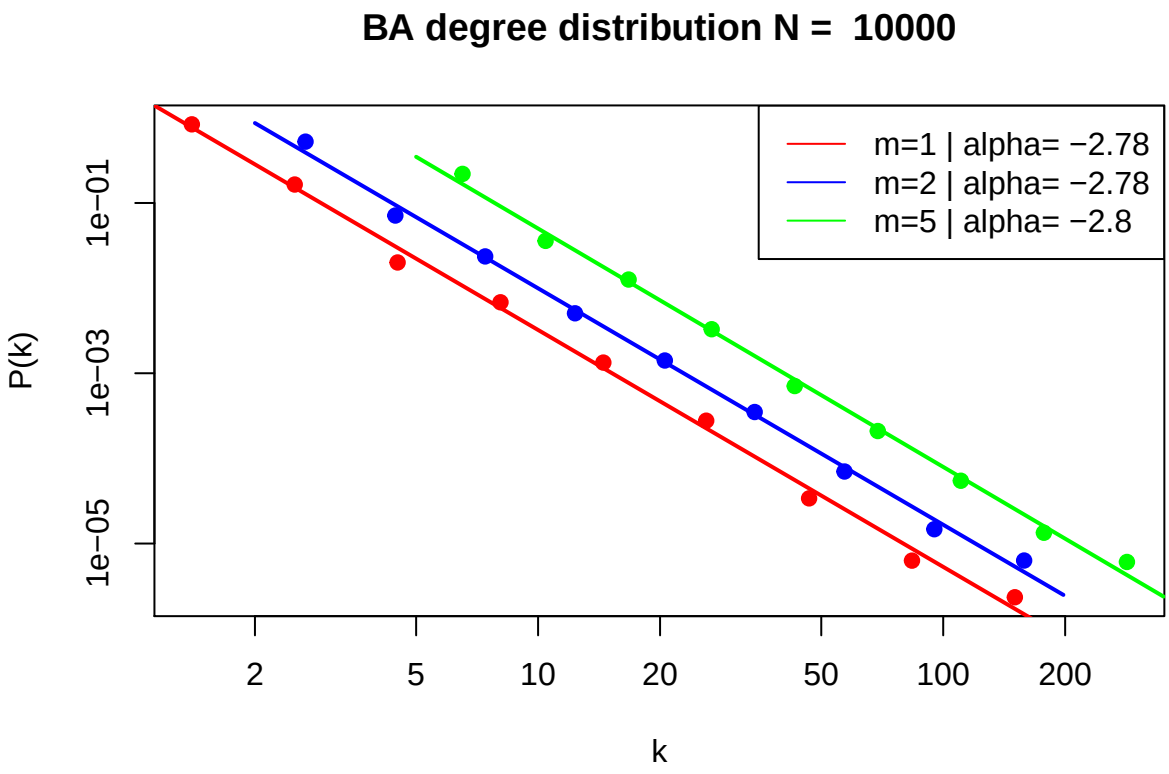

```
# dev.off()
```

```
# png("ba_n=1000.png", width = 600, height = 400)
filenames_n_1000 <- list(m_1 = "output/ba_degree_n=1000_m=1.txt",
                         m_2 = "output/ba_degree_n=1000_m=2.txt",
                         m_5 = "output/ba_degree_n=1000_m=5.txt")
fit_ba_n_1000 <- get_degree_and_fit(filenames_n_1000, 10)

plot_degree_distribution_m(fit_ba_n_1000, 1000)
```



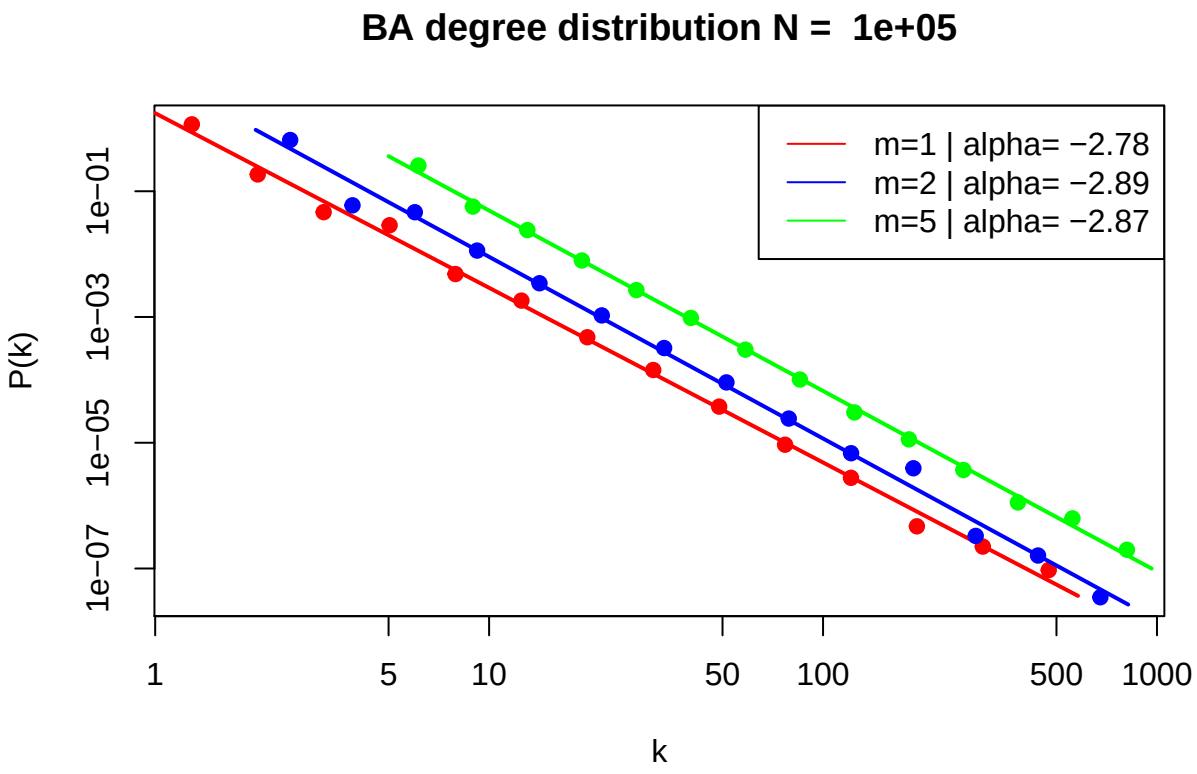

```
# dev.off()
```

```
# png("ba_n=10000.png", width = 600, height = 400)
filenames_n_10000 <- list(m_1 = "output/ba_degree_n=10000_m=1.txt",
                          m_2 = "output/ba_degree_n=10000_m=2.txt",
                          m_5 = "output/ba_degree_n=10000_m=5.txt")
fit_ba_n_10000 <- get_degree_and_fit(filenames_n_10000, 10)

plot_degree_distribution_m(fit_ba_n_10000, 10000)
```



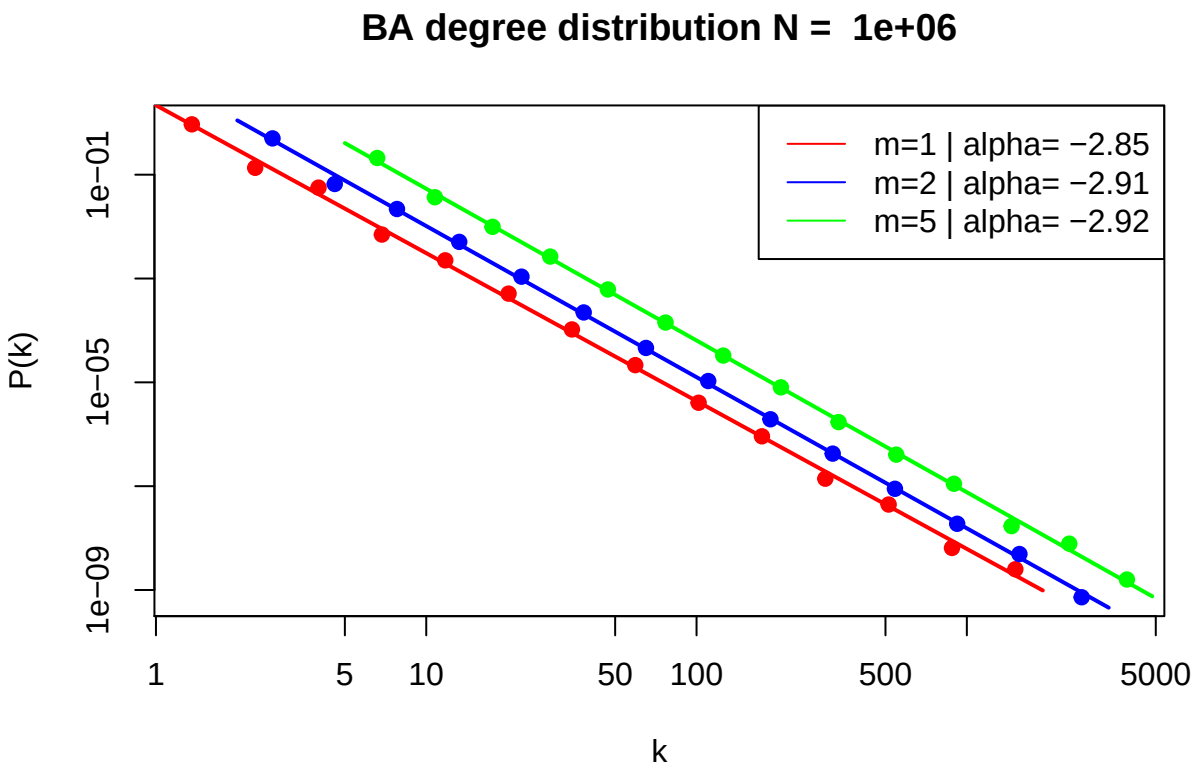

```
# dev.off()
```

```
# png("ba_n=100000.png", width = 600, height = 400)
filenames_n_100000 <- list(m_1 = "output/ba_degree_n=100000_m=1.txt",
                           m_2 = "output/ba_degree_n=100000_m=2.txt",
                           m_5 = "output/ba_degree_n=100000_m=5.txt")
fit_ba_n_100000 <- get_degree_and_fit(filenames_n_100000, 15)

plot_degree_distribution_m(fit_ba_n_100000, 100000)
```




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

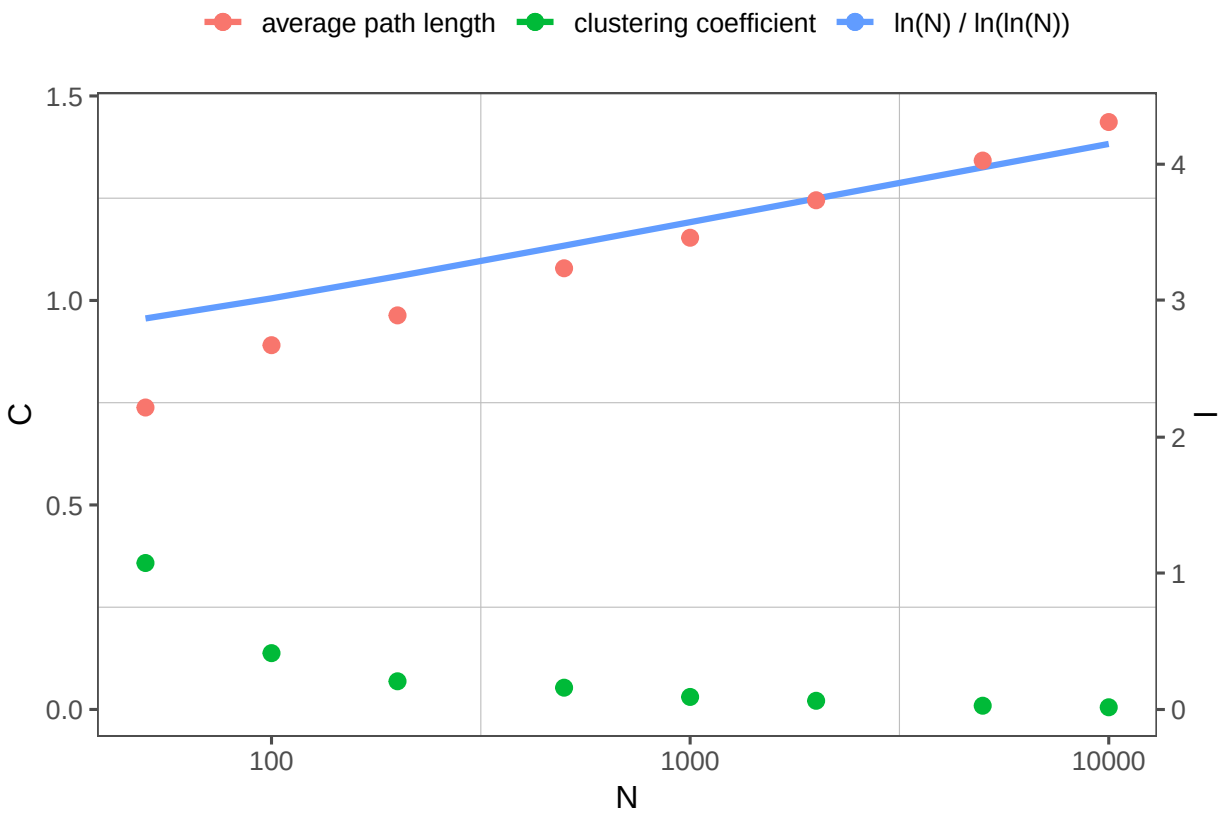
```
# png("ba_n=1000000.png", width = 600, height = 400)
filenames_n_1000000 <- list(m_1 = "output/ba_degree_n=1000000_m=1.txt",
                            m_2 = "output/ba_degree_n=1000000_m=2.txt",
                            m_5 = "output/ba_degree_n=1000000_m=5.txt")
fit_ba_n_1000000 <- get_degree_and_fit(filenames_n_1000000, 15)

plot_degree_distribution_m(fit_ba_n_1000000, 1000000)
```




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

Clustering coefficient

```
library(ggthemes)

clustering_avg_length <-
  read.csv("output/ba_simulation_par_2_clustering.txt", header = FALSE, col.names = c("C", "N")) %>%
  left_join(read.csv("output/ba_simulation_par_2_avg_path.txt", header = FALSE, col.names = c("l", "N"))
  mutate(l_theory = log(N) / log(log(N)))

ggplot(clustering_avg_length) +
  geom_line(aes(N, l_theory / 3, color = "ln(N) / ln(ln(N))"), size = 1.1) +
  geom_point(aes(N, C, color = "clustering coefficient"), size = 2.5) +
  geom_point(aes(N, l / 3, color = "average path length"), size = 2.5) +
  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")
```