

# Social Network Analysis of NFL Coaches, 1980-2013

Nathan E Dire

November 2014

## 1 Introduction

Coaching in the United States' National Football League is a high stakes business. NFL coaches dominate the top paid coaches in U.S. sports [3]. The best coaches are paid in excess of \$7 million per year.

The group of NFL coaches is a relatively tight circle. Head coaches are often hired from the pool of offensive and defensive coordinators.

There is general perception that influential coaches pass on their knowledge to assistants, who then become successful head coaches themselves.

This report considers a few main questions about NFL coaches who have worked together. First, does centrality in the social network align with overall coaching success? Are there communities in the network? Small world characteristics?

## 2 Methodology

### 2.1 Data Acquisition

Data on NFL coaching staffs was gathered from the site <http://www.pro-football-reference.org> by reading the pages and parsing the listed coaching staff. For example, this page <http://www.pro-football-reference.com/teams/det/2010.htm> lists the coaches for the 2010 Detroit Lions. The team page generally lists the head coach, offensive coordinator, and defensive coordinator. Only the years 1980-2013 were considered. The code used to produce this data is available at [4].

### 2.2 Network Construction

The relationship between members of the coaching staff is subject to interpretation. Within this data, the coach may be both a coordinator and head coach.

For this analysis, I consider two constructions of the social network:

1. The first model attempts to follow the “coaching family tree” perception, and only has directed edges from head coach to coordinator. This approach attempts to model the influence of successful coaches through the success of their assistants. Edges are between the head coach and the coordinators; coordinators have no connection to each other. This model will be referred to as the *tree* model. The file is `coaches_tree.gml`.

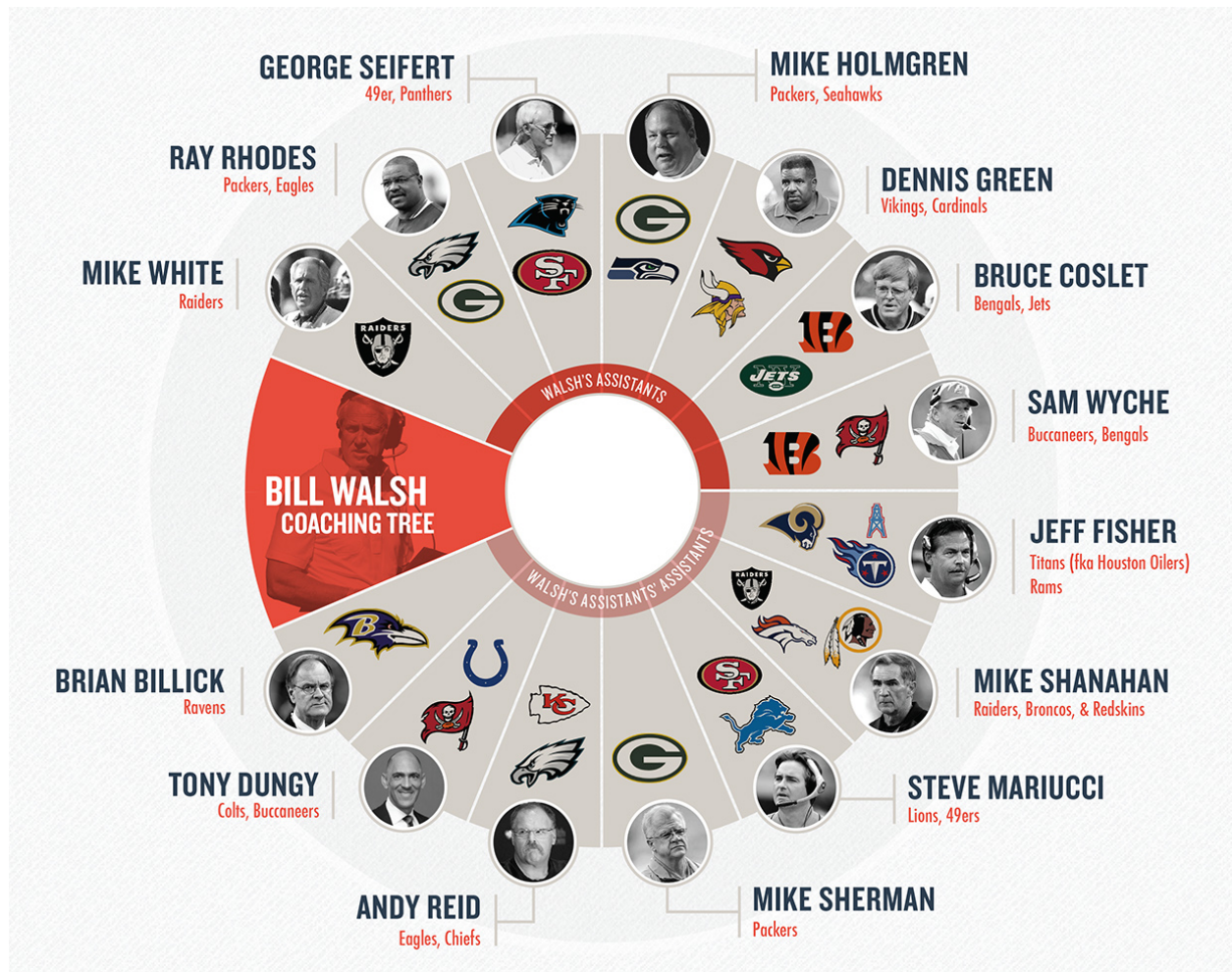


Figure 1: Image: The Bill Walsh Coaching Tree. Source: <http://blog.hubspot.com/marketing/paypal-mafia-bill-walsh-nfl-infographic>

2. The second model considers each member of the staff to be connected with every coach they served with, with increasing weight for each year served. This approach attempts to identify coaching “communities” who may tend serve together. This model will be referred to as the *peer* model. The file is `coaches_peer.gml`.

## 2.3 Analysis

Network analysis is done in the R programming language with the `igraph` package. The calculations are included in this document using the `knitr` package.

## 2.4 Limitations

There are several important limitations of this analysis:

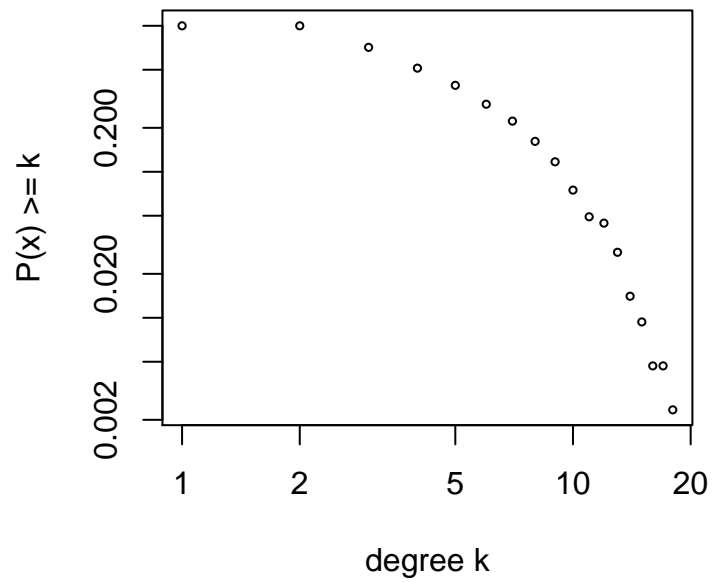
- Only the head coach and coordinators are considered. This represents the only clean data available from <http://pro-football-reference.com>. It also avoids the complexities resulting from teams having different names for the various position coaches. This has the downside of not recognizing cases where head coaches and coordinators take jobs as position coaches.
- Only the years 1980-2013 are considered. The consistent reporting of offensive and defensive coordinators appears to start around 1980. This results in the network being truncated for coaches employed before and after 1980.
- Coaches are considered to serve the entire year. In reality, teams may occasionally fire coaches mid-year, and coaches may take a leave of absence.

## 3 Network Properties

```
tree_g = read.graph("coaches_tree.gml",format="gml")
summary(tree_g)

## IGRAPH U--- 427 777 --
## attr: href (v/c), guid (v/c), label (v/c), id (v/n), value (e/n)

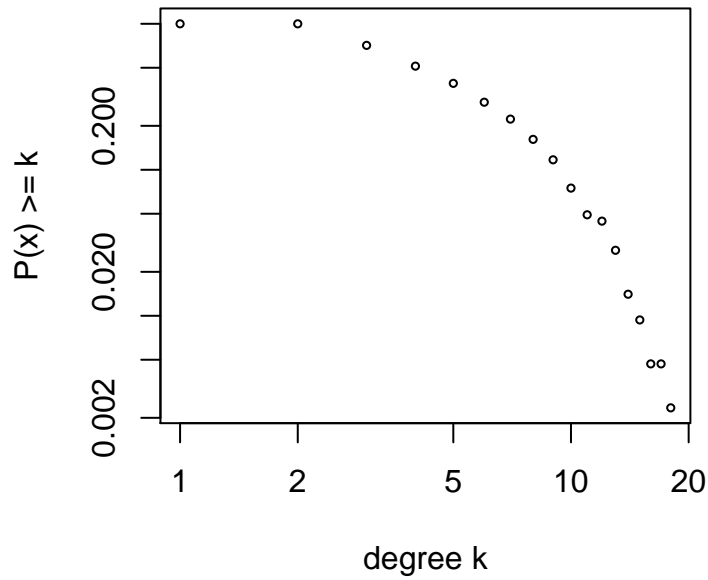
tree_deg = degree.distribution(tree_g, cumulative = TRUE)
plot(tree_deg,log="xy",xlab="degree k",ylab="P(x) >= k",cex=0.5)
```



```
peer_g = read.graph("coaches_peer.gml",format="gml")
summary(peer_g)

## IGRAPH U--- 427 1244 --
## attr: href (v/c), guid (v/c), label (v/c), id (v/n), value (e/n)

peer_deg = degree.distribution(tree_g, cumulative = TRUE)
plot(peer_deg,log="xy",xlab="degree k",ylab="P(x) >= k",cex=0.5)
```



## 4 Centrality

### 4.1 Degree

### 4.2 Betweenness

```
bb = betweenness.estimate(tree_g, v=V(tree_g), directed=FALSE, cutoff=16)
```

```
bb = betweenness.estimate(peer_g, v=V(peer_g), directed=FALSE, cutoff=16)
```

### 4.3 Closeness

## 5 Community Structure

## 6 Small World Network Properties

## 7 Conclusion

This report analyzed the social network of NFL coaches serving as head coach or coordinator in the years 1980-2013.

The content used to generate this report can be found at [?].

## References

- [1] Harrison, C.K. & Associates (2013). Coaching Mobility (Volume I in the Good Business Series). A Report for the NFL Diversity and Inclusion Series.
- [2] Harrison, C.K. & Bukstein, S. (2014). NFL Occupational Mobility Patterns (Volume III). A report for the NFL Diversity and Inclusion â€”Good Businessâ€” Series.
- [3] <http://www.forbes.com/sites/chris-smith/2013/05/22/the-highest-paid-coaches-in-us-sports/>
- [4] <https://github.com/ndire/pfr-scraper>
- [5] <https://github.com/ndire/sna-nfl-coaches>