# An Introduction to Small-World and Scale-Free Networks

DEREK SOLLBERGER

Department of Mathematics and Statistics
California State University, Long Beach
Derek.Sollberger@gmail.com

May 7, 2009

# Outline

# Examples

Before we looked into network patterns, dynamical systems have been used to study networks such as

- Biological oscillation
- Excitable media (unbounded growth, e.g. forest fire)
- Neural networks
- Spatial games (e.g. chess, checkers)

[2]

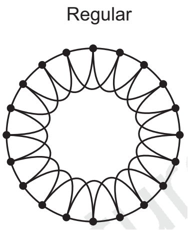
# Examples

However, networks tend to be wired as highly-clustered systems, hence small-world networks can include

- Brains of tiny animals
- US power grid
- Film actors
- Disease studies

[2]

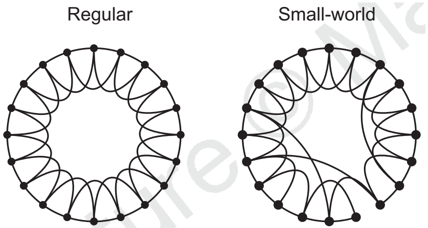
# Heuristic Model




- *n* vertices
- *k* edges per vertex

[2]

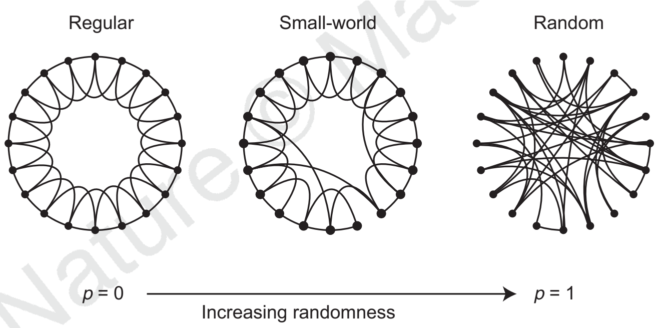
# Heuristic Model




- We rewire each edge at random with probability $p$

[2]

# Heuristic Model




- We can 'tune' the graph between regularity ($p = 0$) and disorder ($p = 1$)

[2]

# Definitions

To talk about network structure, let ...

- $L(p)$ be the characteristic path length
- $C(p)$ be the clustering coefficient

Loosely put, ...

- $L(p)$ is the most common distance
- $C(p)$ is the "cliquishness"

[2]

# Definitions

To talk about network structure, let ...

- $L(p)$ be the characteristic path length
- $C(p)$ be the clustering coefficient

Loosely put, ...

- $L(p)$ is the most common distance
- $C(p)$ is the "cliquishness"

[2]

# Model Behavior

To utilize one, connected graph, we require that

$$n > k > \ln n > 1$$

With those bounds on $k$, as $p \to 0$ (regular lattice):

- $L(p) \approx L_0 \sim \frac{n}{2k}$
- $C(p) \approx C_0 \sim \frac{3}{4}$

As $p \to 1$ (toward disorder):

- $L(p) \approx L_{random} \sim \frac{\ln n}{\ln k}$
- $C(p) \approx C_{random} \sim \frac{k}{n}$

[2]

# Model Behavior





To utilize one, connected graph, we require that

$$n > k > \ln n > 1$$

With those bounds on $k$, as $p \to 0$ (regular lattice):

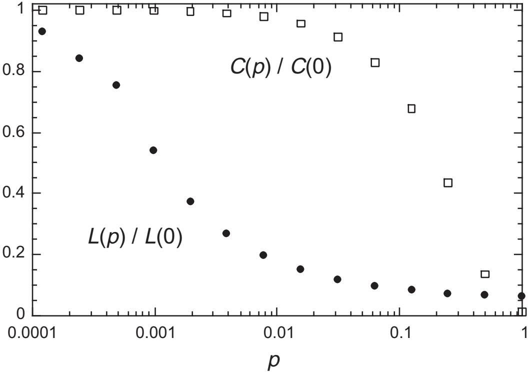
- $L(p) \approx L_0 \sim \frac{n}{2k}$
- $C(p) \approx C_0 \sim \frac{3}{4}$

As $p \to 1$ (toward disorder):

- $L(p) \approx L_{random} \sim \frac{\ln n}{\ln k}$
- $C(p) \approx C_{random} \sim \frac{k}{n}$

[2]

# Tracking Disorder




Upon increasing disorder ($p$), we see that

- characteristic path length $L$ drops dramatically
- clustering ($C$) remains virtually unchanged

[2]

# Testing the Model

**Table 1 Empirical examples of small-world networks**

| | $L_{actual}$ | $L_{random}$ | $C_{actual}$ | $C_{random}$ |
|---|---|---|---|---|
| Film actors | 3.65 | 2.99 | 0.79 | 0.00027 |
| Power grid | 18.7 | 12.4 | 0.080 | 0.005 |
| *C. elegans* | 2.65 | 2.25 | 0.28 | 0.05 |

Characteristic path length $L$ and clustering coefficient $C$ for three real networks, compared to random graphs with the same number of vertices ($n$) and average number of edges per vertex ($k$). (Actors: $n = 225{,}226$, $k = 61$. Power grid: $n = 4{,}941$, $k = 2.67$. *C. elegans*: $n = 282$, $k = 14$.) The graphs are defined as follows. Two actors are joined by an edge if they have

In all three scenarios,

- $L_{actual} > L_{random}$ and $C_{actual} > C_{random}$

[2]

# Swine Flu!

http://www.thenetworkthinker.com/2009/04/network-structure-of-swine-flu-pandemic.html

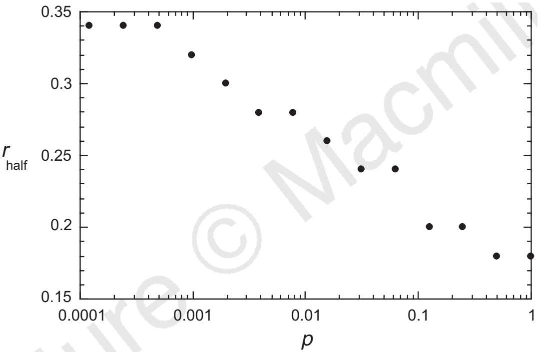
A person can infect another person with infectiousness probability *r*. [2]

# Disease Power

The critical infectiousness $r_{half}$ of a disease is the ability to infect half of the population based on infectiousness probability $r$. [2]

# Disease Power vs. Randomness

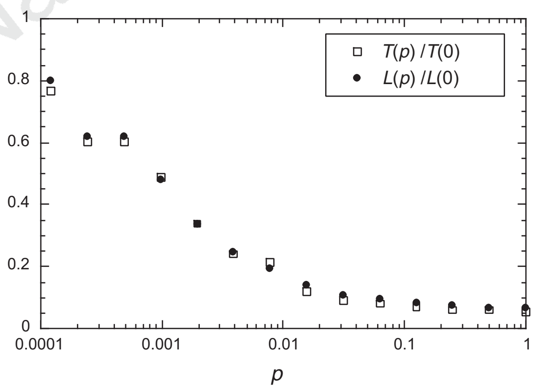
The critical infectiousness $r_{half}$ of a disease is the ability to infect half of the population based on infectiousness probability $r$.




[2]

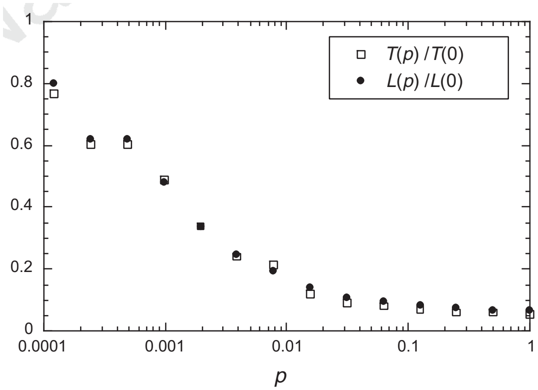
# Vindication

For an “ideal” disease ($r = 1$), the time it takes for the disease to infect the entire population $T(p)$ is represented by a curve that is similar to the $L(p)$ curve.




[2]

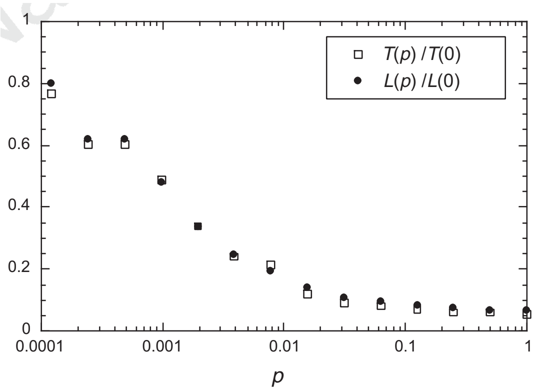
# Punch Line




- An infectious disease spreads faster in a small world.
- Even with a small amount of rewiring ("short cuts"), the world becomes quite small.

[2]

# Punch Line




- An infectious disease spreads faster in a small world.
- Even with a small amount of rewiring (“short cuts”), the world becomes quite small.

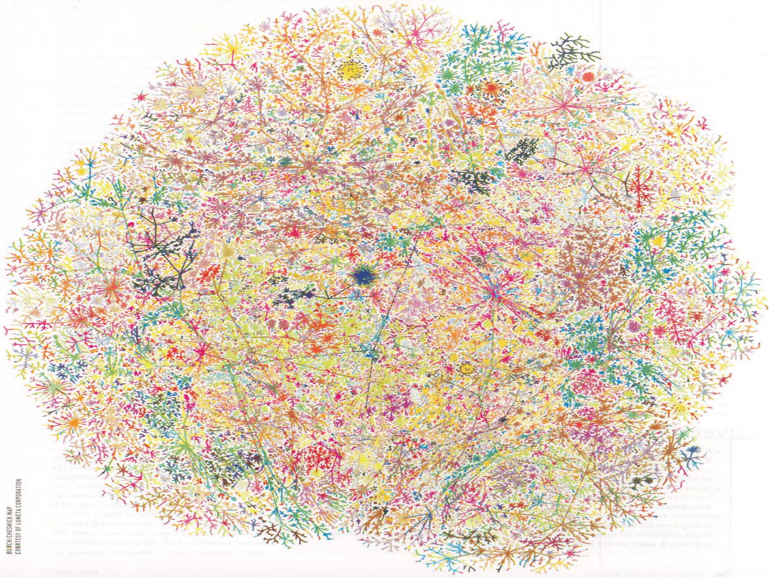
[2]

# Moving Forward

Some assumptions may have to be relaxed in other models, such as

- Not all groups have connected individuals
- The random rewiring may have an underlying distribution, such as ‘majority-rule’.

[2]


BURCH/CHESWICK MAP
COURTESY OF LUMETA CORPORATION

# More Examples

- neural networks
- societies
- internet
- power grids
- transportation systems

[1]

# For Your Consideration

- How do malfunctioning nodes in genetics lead to cancer?
- How does diffusion occur so rapidly in social networks?
- How can some networks still function with failed nodes?

[1]

# The Pattern

The World Wide Web:

- Many web pages have a few links (e.g. my website)
- A few webpages have many links (e.g. `http://www.yahoo.com/`)

Hollywood Actors:

- Many actors have been in a few films (e.g. Tim Hohn, "Dying Man #2" in *Vlad & Antoinette*)
- A few actors have been in many films (e.g. Kevin Bacon)

We say that this pattern represents scale-free networks. [1]

# The Pattern

The World Wide Web:

- Many web pages have a few links (e.g. my website)
- A few webpages have many links (e.g. `http://www.yahoo.com/`)

Hollywood Actors:

- Many actors have been in a few films (e.g. Tim Hohn, “Dying Man #2” in *Vlad & Antoinette*)
- A few actors have been in many films (e.g. Kevin Bacon)

We say that this pattern represents scale-free networks. [1]

# Assumptions

- No node is typical of another one
- Scale-free networks are resistant to accidental failures, but vulnerable to coordinated attacks
- There are underlying patterns for making connections and clustering

[1]

# Early Hypotheses

In 1959, Paul Erdós and Alfréd Renyi suggested modeling networks with random links.

- Poisson distribution for the number of links for a node

[1]

# Mapping the Web

In 1998, Hawoong Jeong and Réka Albert set out to map the internet.

- Web crawlers collected data on links

Some of the results included:

80% of webpages had fewer than 4 links, while

$<$ 0.01% of webpages had more than 1000 links.

[1]

# Mapping the Web

In 1998, Hawoong Jeong and Réka Albert set out to map the internet.

- Web crawlers collected data on links

Some of the results included:

80% of webpages had fewer than 4 links, while
$<$ 0.01% of webpages had more than 1000 links.

[1]

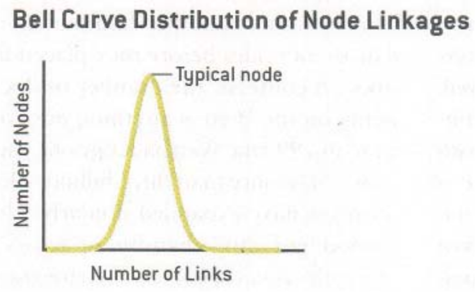
# Example: Transportation Networks

The US highway system resembles a random network.

Most of the nodes (cities) have the same number of outgoing links (highways) [1]

# Example: Transportation Networks

Thus we can conjecture that a network's links per node follow a bell curve.




[1]

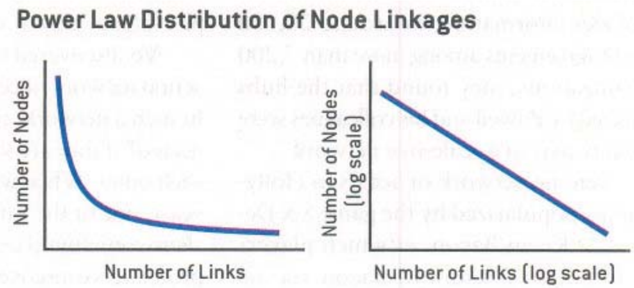
# Example: Transportation Networks

Airline routes resemble a scale-free network.

A few hubs have many outgoing routes. [1]

# Example: Transportation Networks

It seems that the best trend to fit the data would be a power function.




[1]

## Examples of Scale-Free Networks

| NETWORK | NODES | LINKS |
|---|---|---|
| **Cellular metabolism** | Molecules involved in burning food for energy | Participation in the same biochemical reaction |
| **Hollywood** | Actors | Appearance in the same movie |
| **Internet** | Routers | Optical and other physical connections |
| **Protein regulatory network** | Proteins that help to regulate a cell's activities | Interactions among proteins |
| **Research collaborations** | Scientists | Co-authorship of papers |
| **Sexual relationships** | People | Sexual contact |
| **World Wide Web** | Web pages | URLs |

# Revising the Random-Network Model

We still need to explain why hubs exist and grow.

Assuming a power function model $f(x) = x^n$, what should the exponent be?

- Empirical tests have found results between 2 and 3.

[1]

# Revising the Random-Network Model

We still need to explain why hubs exist and grow.

Assuming a power function model $f(x) = x^n$, what should the exponent be?

- Empirical tests have found results between 2 and 3.

[1]

# Revising the Random-Network Model

We will relax a couple of the assumptions from the old Erdós and Rényi model. [1]

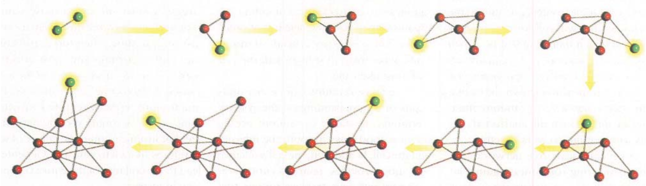
# Dynamical Systems

Networks grow over time. [1]

# Not All Nodes Are Equal

Linking toward more nodes can be affected by "preferential attachment".

A SCALE-FREE NETWORK grows incrementally from two to 11 nodes in this example. When deciding where to establish a link, a new node (*green*) prefers to attach to an existing node (*red*) that already has many other connections. These two basic mechanisms—growth and preferential attachment—will eventually lead to the system's being dominated by hubs, nodes having an enormous number of links.



[1]

Network growth and preferential attachment $\Rightarrow$ hubs.

The authors' work with computer simulations corroborated this trend. [1]

Network growth and preferential attachment $\Rightarrow$ hubs.

The authors' work with computer simulations corroborated this trend. [1]

# Winner Take All?

It appears that the preferential attachment tends to be linear.

- A new node is twice as likely to link to an existing node that has twice as many connections as its neighbor

If that mechanism is faster, then one hub would gain links from all the new nodes, hence a “winner take all” scenario. [1]

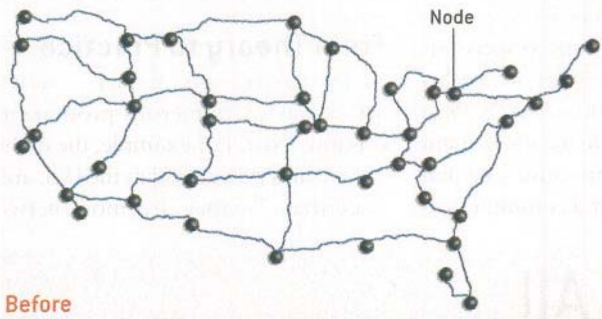
# What Makes a Network Robust?

Complex networks (and their nodes), such as

- power grids (power plants)
- communication webs (routers)
- living systems (proteins)

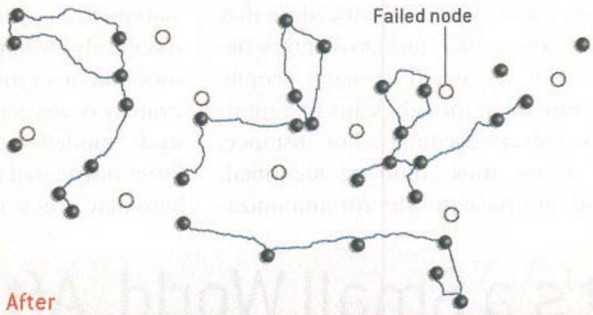
seem to be resilient versus accidental failures. [1]

# Random Network: Accidental Node Failure



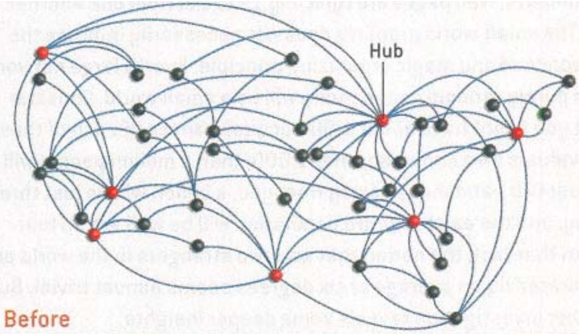

Nearly homogeneous topology [1]

# Random Network: Accidental Node Failure



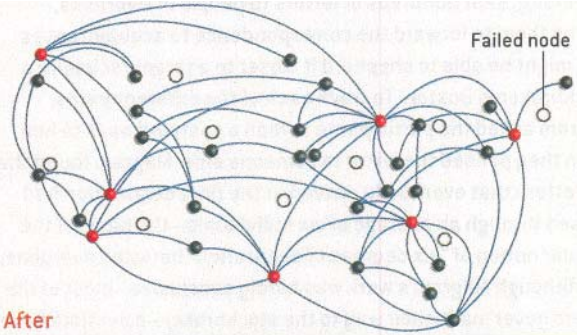

Quick deterioration into disconnected, noncommunicating islands [1]

# Scale-Free Network: Accidental Node Failure



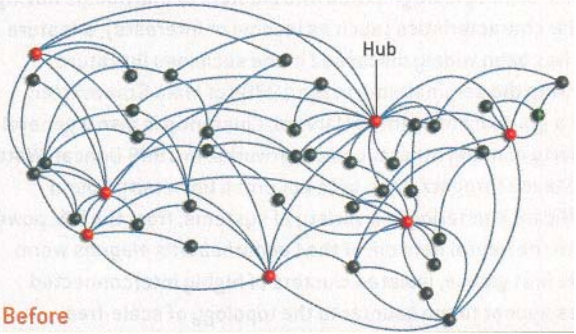

Inhomogeneous topology [1]

# Scale-Free Network: Accidental Node Failure



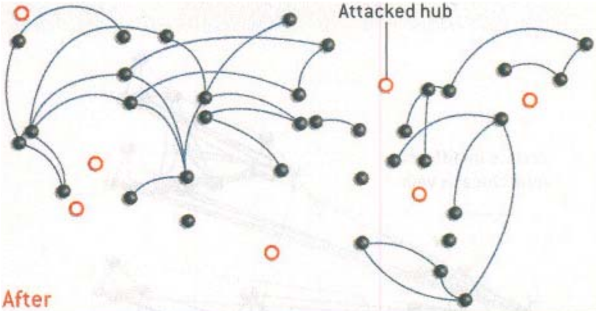

Smaller nodes are disrupted, but the larger ones remain [1]

# Scale-Free Network: Attack on Hubs




Prior knowledge of the hubs can lead to a coordinated attack [1]

# Scale-Free Network: Attack on Hubs




Even a robust, scale-free network can be disrupted [1]

# Coordinated Attacks

Good intentions ...

... protect hubs

Evil intentions ...

... attack hubs

[1]

# Coordinated Attacks

Good intentions ...

... protect hubs

Evil intentions ...

... attack hubs

[1]

# Viruses

In a scale-free network, the threshold is zero.

- Diseases can appear at any time
- Computer viruses cannot be completely eradicated

[1]

# Viruses

In a scale-free network, the threshold is zero.

- Diseases can appear at any time
- Computer viruses cannot be completely eradicated

[1]

# Viruses

In a scale-free network, the threshold is zero.

- Diseases can appear at any time
- Computer viruses cannot be completely eradicated

[1]

# Immunizations

Can knowledge of an underlying, network structure help us study medicine?

It makes sense to immunize the people act as hubs ...

... but who are the hubs?

[1]

# Example of a Coordinated Attack

In the 1950s, Pfizer studied the rate at which doctors prescribe new drugs.

Perhaps, more studies into scale-free networks can give a mathematical framework to such a process. [1]

# Counterexamples

Some examples of networks that are not scale-free (i.e. do not follow the power function model) include

- US highway system
- Power grid
- Crystal lattice
- Small food webs

The reasons include

- Man-made projects start off small, and are then scaled for larger projects
- With few choices, there could be less preferential attachment

[1]

# Counterexamples

Some examples of networks that are not scale-free (i.e. do not follow the power function model) include

- US highway system
- Power grid
- Crystal lattice
- Small food webs

The reasons include

- Man-made projects start off small, and are then scaled for larger projects
- With few choices, there could be less preferential attachment

[1]

# Future Model Updates

Relaxing more assumptions, we could attempt to get more accurate models

- Consider a network's diameter
- Measure speed/strength of links
- Node content is not homogeneous

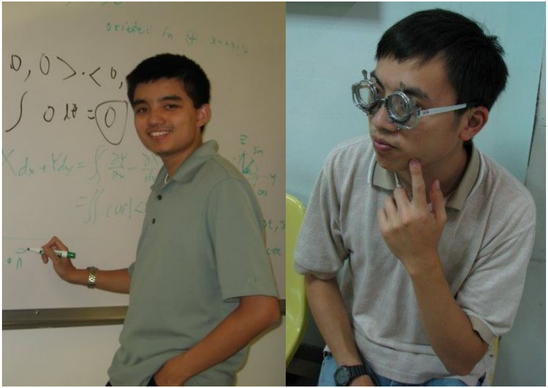
[1]

Albert-Laszlo Barabasi and Eric Bonabeau.
Scale-free networks.
*Scientific American*, 288:50–59, 2003.

Duncan J. Watts and Steven H. Strogatz.
Collective dynamics of 'small-world' networks.
*Nature*, 393(6684):409–410, 1998.

# One Last Example

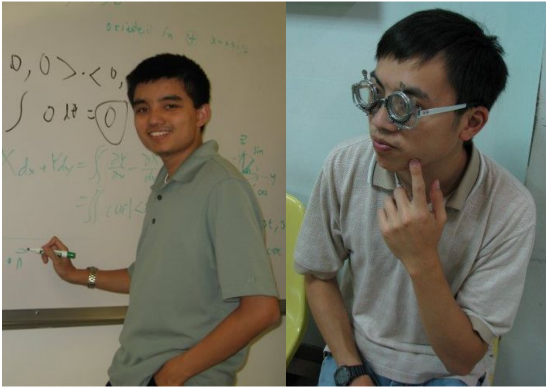
I was in a community service club



with Kevin Lin
(photo used with permission)

# One Last Example

I was in a community service club



with Kevin Lin
(photo used with permission)

# One Last Example

who was in Berkeley engineering

with Tiffany Shiau
(photos used with permission)

# One Last Example

who was in Berkeley engineering

with Tiffany Shiau
(photos used with permission)

# One Last Example

who played “Factory Worker #2” in *One Child* (2008)

with Joanne Chew
(from http://www.imdb.com/media/rm1190696448/nm2415065)

# One Last Example

who played "Factory Worker #2" in *One Child* (2008)

with Joanne Chew
(from `http://www.imdb.com/media/rm1190696448/nm2415065`)

# One Last Example

who played “Park Twin #1” in *Street Kings* (2008)

with Forest Whitaker
(from http://www.imdb.com/media/rm119069648/nm2415065 and http://www.imdb.com/media/rm2685243648/nm0001845)

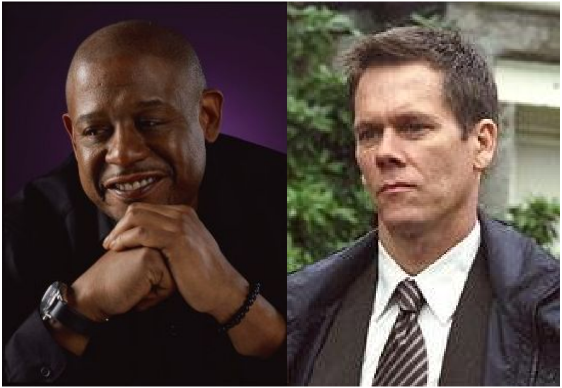
# One Last Example

who played "Park Twin #1" in *Street Kings* (2008)

with Forest Whitaker
(from http://www.imdb.com/media/rm1190696448/nm2415065 and http://www.imdb.com/media/rm2685243648/nm0001845)

# One Last Example

who played "Happiness" in *The Air I Breathe* (2007)

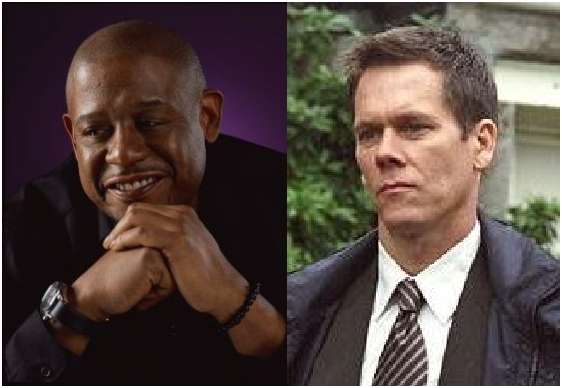


with Kevin Bacon!

(from http://www.imdb.com/media/rm2685243648/nm0001845 and http://www.imdb.com/media/rm1837931776/nm0000102)

# One Last Example

who played "Happiness" in *The Air I Breathe* (2007)



with Kevin Bacon!

(from `http://www.imdb.com/media/rm2685243648/nm0001845` and `http://www.imdb.com/media/rm1837931776/nm0000102`)

# The End