

COCS 6323: Statistical Methods in Research  
Group Project

Group <number>  
Department of Computer Science  
University of Houston  
March 7, 2019

## Contents

<b>1 Contribution</b>	<b>3</b>
<b>2 Figure 2A</b>	<b>4</b>
<b>3 Figure 2B</b>	<b>5</b>
<b>4 Figure 3A</b>	<b>6</b>
<b>5 Figure 3B</b>	<b>7</b>
<b>6 Figure 3C</b>	<b>8</b>
<b>7 Figure 3D</b>	<b>9</b>
<b>8 Figure 3E</b>	<b>10</b>
<b>9 Figure 3F</b>	<b>11</b>
<b>10 Supplementary S1</b>	<b>12</b>
<b>11 Supplementary S2</b>	<b>13</b>
<b>12 Supplementary S3</b>	<b>14</b>
<b>13 Supplementary S4</b>	<b>15</b>

## List of Tables

## List of Figures

1 Growth of cross-disciplinary social capital . . . . .	4
2 Growth of cross-disciplinary social capital . . . . .	5
3 Network Distributions for Direct and Mediated Associations of Biologists . . . . .	13
4 Network Distributions for Direct and Mediated Associations of Computer Scientists . . . . .	13
5 Three perspectives on the centrality of $F_i$ in the direct collaboration network . . . . .	14
6 Growth of cross-disciplinary social capital . . . . .	15

## 1 Contribution

Add your content here

## 2 Figure 2A

Giant Component

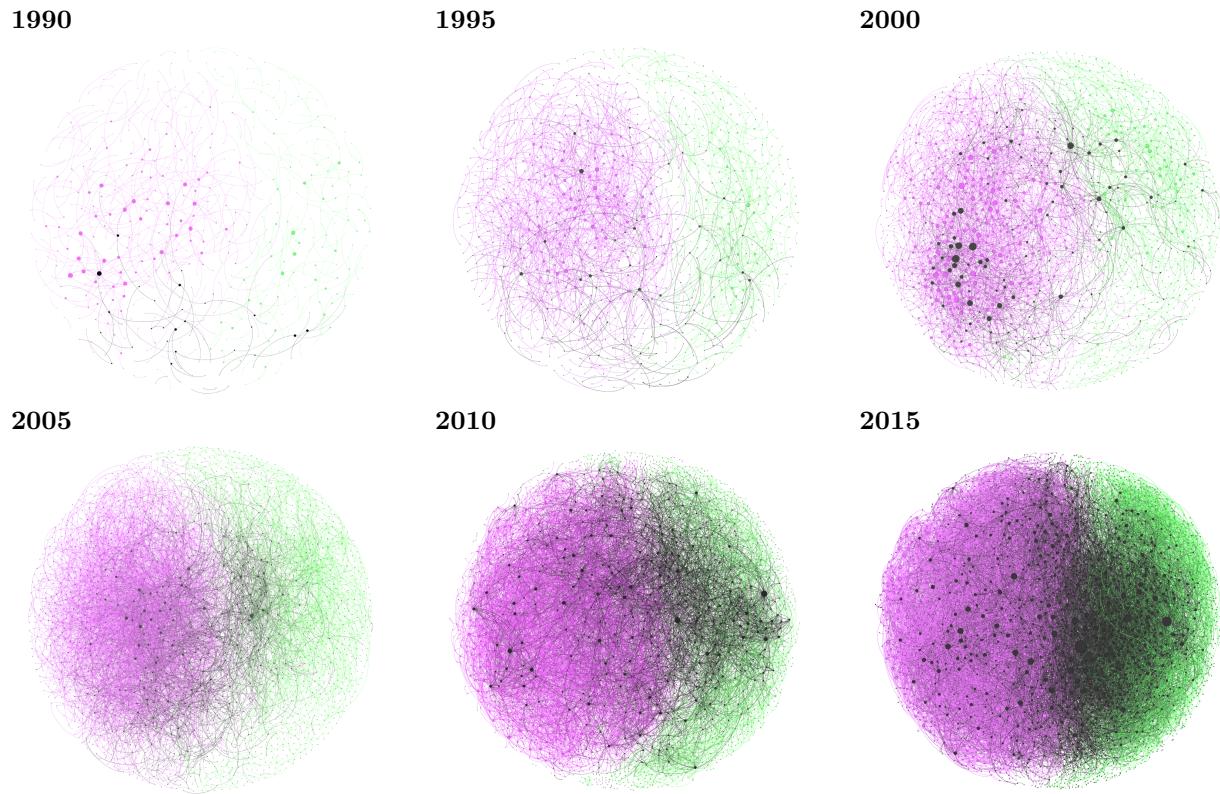


Figure 1: Growth of cross-disciplinary social capital

Evolution of the giant component in the U.S. biology-computing network. Green and magenta nodes represent faculty  $F_i$  with BIOF and CSF affiliation, respectively; black nodes represent faculty  $F_i$  that, by time  $t$ , published at least one cross-disciplinary publication and joined the XDF group; node size is proportional to the logarithm of the degree centrality,  $\ln C_{Di}$ , of  $F_i$  at time  $t$ .

### 3 Figure 2B

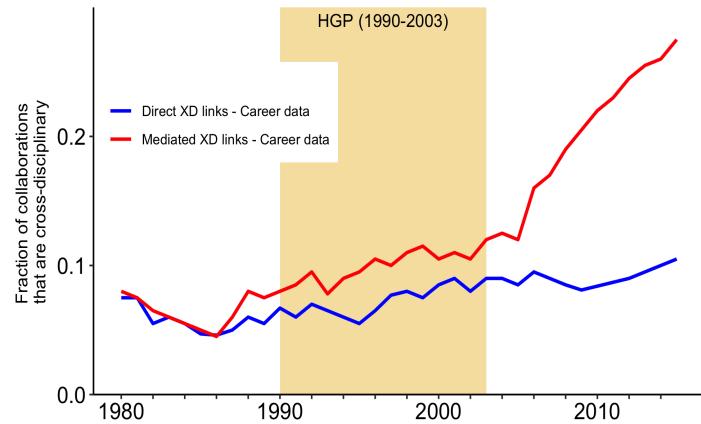


Figure 2: Growth of cross-disciplinary social capital

## **4   Figure 3A**

Add the content of this section

## **5    Figure 3B**

Add the content of this section

## **6    Figure 3C**

Add the content of this section

## **7   Figure 3D**

Add the content of this section

## **8    Figure 3E**

Add the content of this section

## **9    Figure 3F**

Add the content of this section

## **10 Supplementary S1**

Add your content here

## 11 Supplementary S2

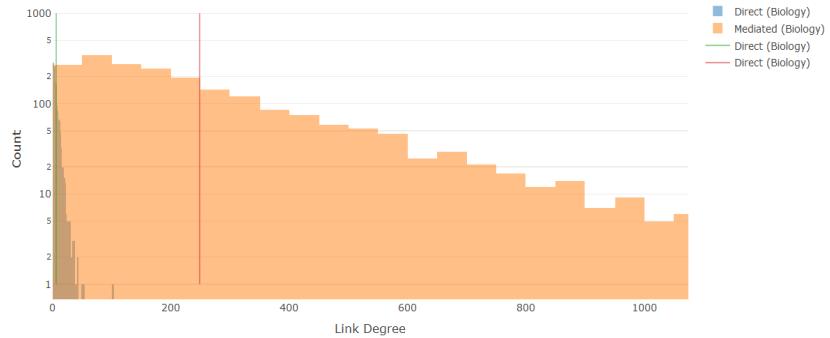


Figure 3: Network Distributions for Direct and Mediated Associations of Biologists

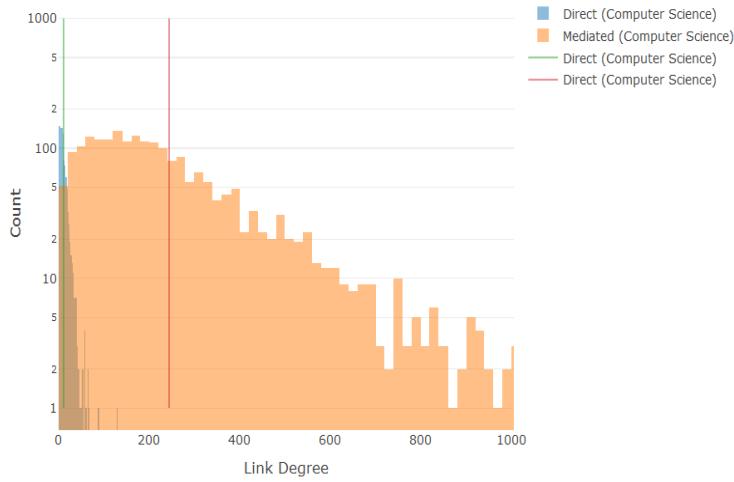


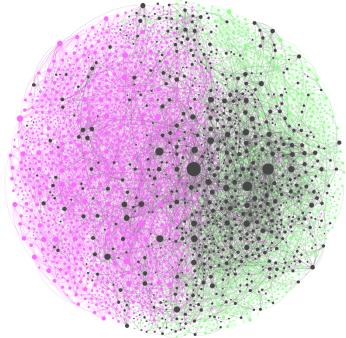
Figure 4: Network Distributions for Direct and Mediated Associations of Computer Scientists

Figure S2 depicts the number of associations between professors within biology departments (a), and computer science departments (b) with other researchers studying genomics. The dark area shows counts the direct connections between professors, while the lighter orange area shows counts of mediated connections. Vertical lines represent means of each distribution. Direct links are made when professors within the dataset collaborate to publish a paper. Mediated links are established when professors have a collaborator in common. The histograms above demonstrate the importance of the mediated connections, with regards to the robustness of the network.

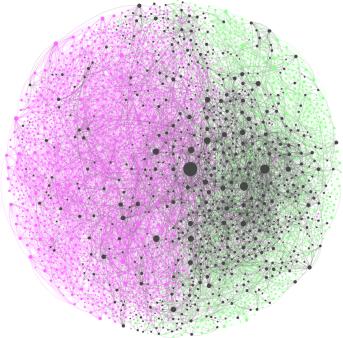
## 12 Supplementary S3

Centrality

Degree



PageRank



Betweenness

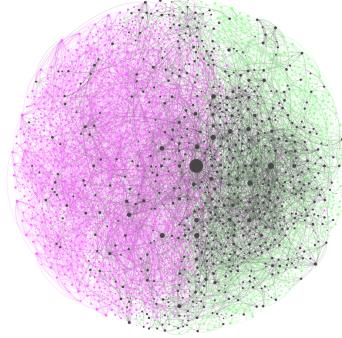


Figure 5: Three perspectives on the centrality of Fi in the direct collaboration network

Shown is the giant connected component of the faculty network  $F$  using all data up to 2015. The nodes and links across each network are fixed, only the node sizes vary according to the indicated centrality measure. Notably, the most central according to each of the three measures is Eric Lander, one of the leaders of the HGP.

## 13 Supplementary S4

Dismishing of Giant Component

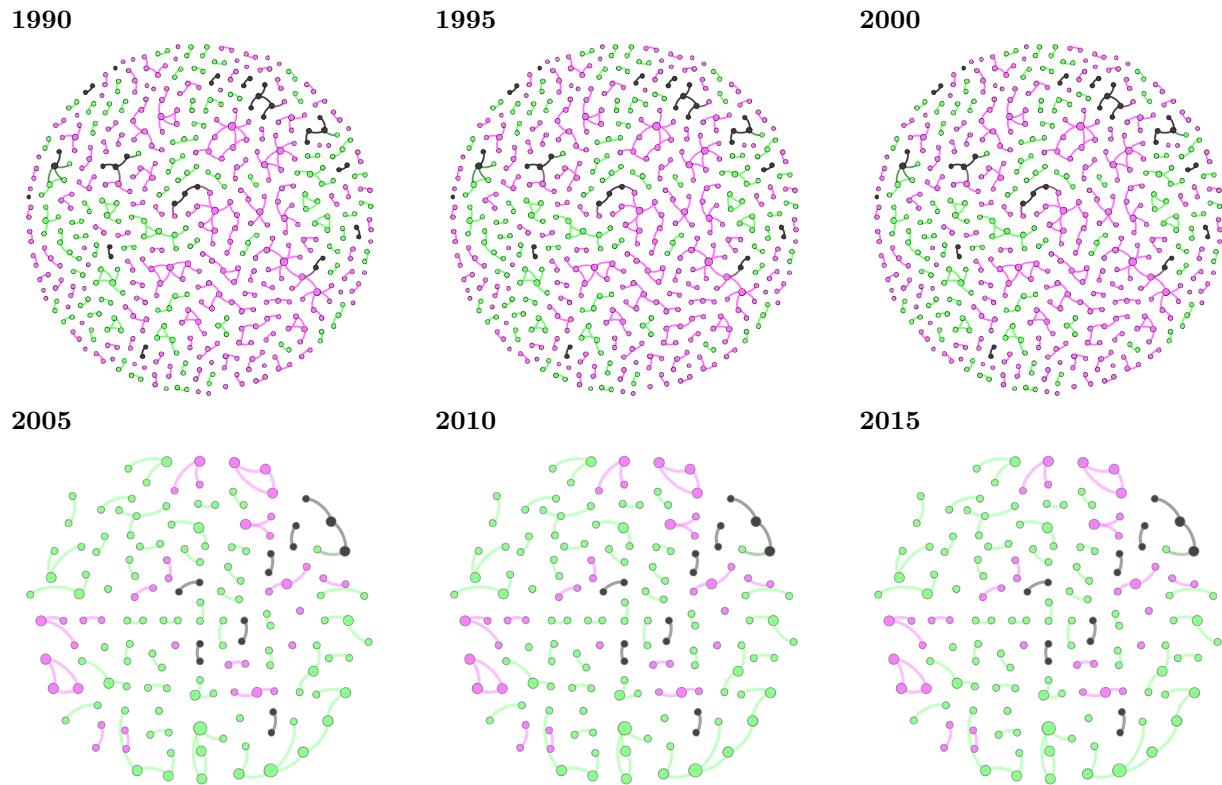


Figure 6: Growth of cross-disciplinary social capital

Green and magenta nodes represent faculty  $F_i$  with BIO and CS affiliation, respectively; black nodes represent faculty  $F_i$  that by time  $t$  collaborated with at least one faculty from the opposite department and thus joined the XD group.