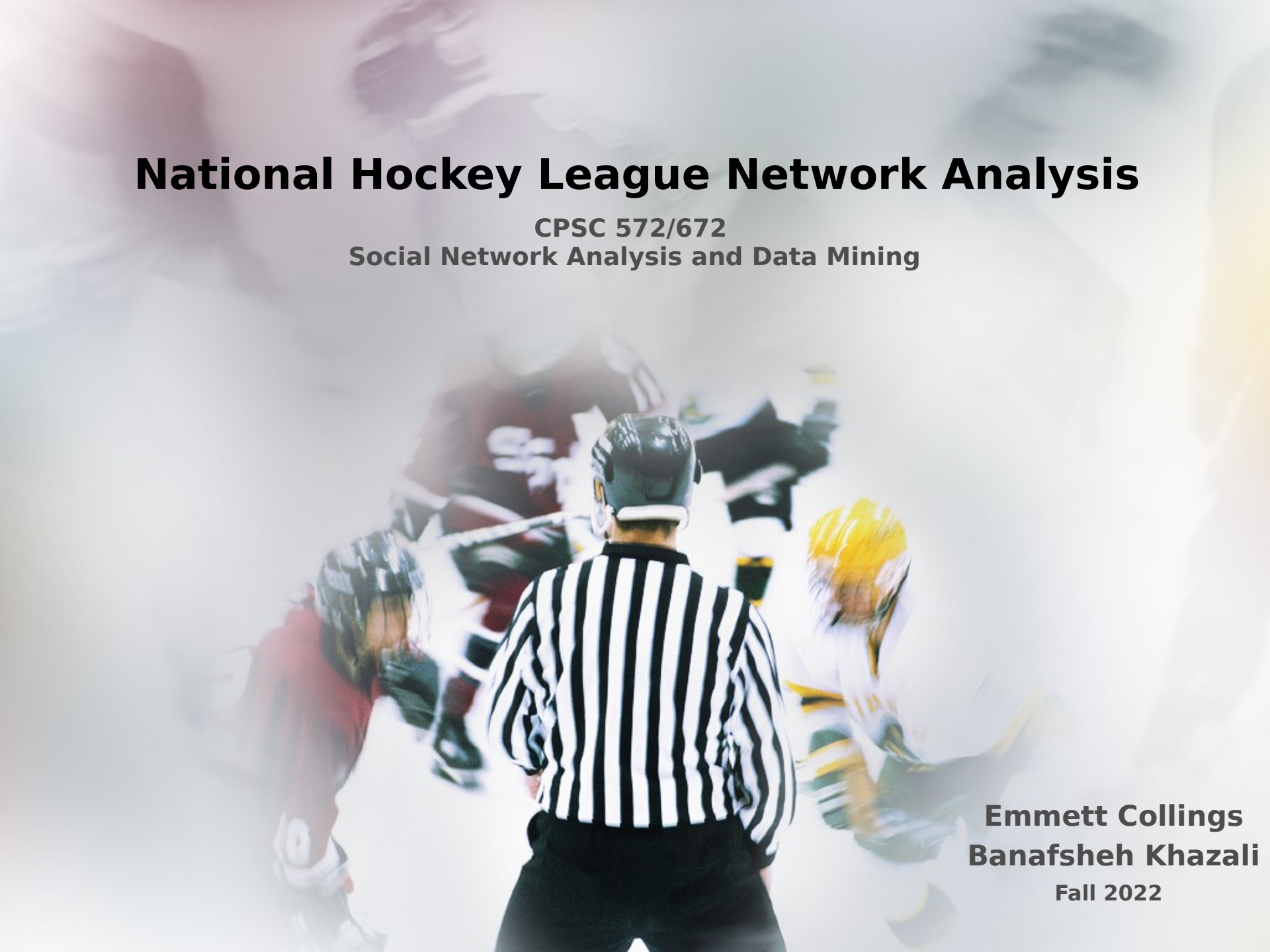
National Hockey League Network Analysis

CPSC 572/672
Social Network Analysis and Data Mining



Emmett Collings
Banafsheh Khazali

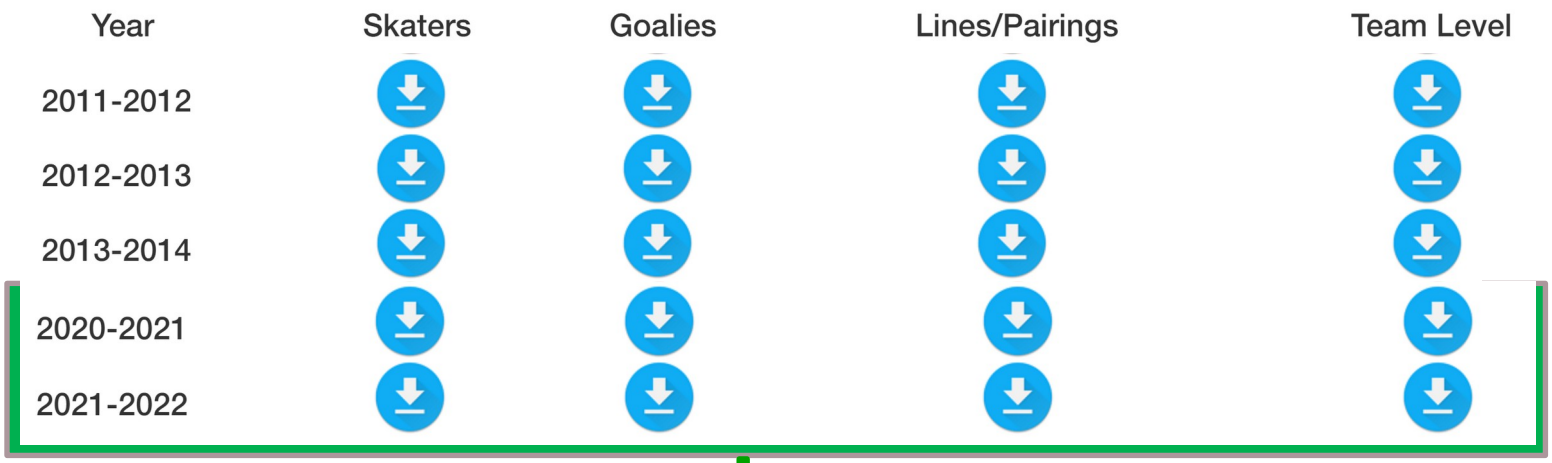
Fall 2022

Data Acquisition

Dataset name: National Hockey League Dataset

Source: https://moneypuck.com/data.htm

Data: player and team raw data CSV network format



H	G	F	E	D	C	B	A	
icetime	games_pla	situation	position	team	name	season	lineId	1
787	5	5on5	line	PHI	Van Riemsdyk-Laughton-Atkinson	2021	8.47E+20	2
762	3	5on5	pairing	NJD	Bahl-Geertsen	2021	8.48E+13	3
16662	72	5on5	pairing	DAL	Lindell-Hakanpaa	2021	8.48E+13	4
1113	5	5on5	line	COL	Lehkonen-Compher-Newhook	2021	8.48E+20	5

Network Description

Nodes: Individual players 

C
name
Van Riemsdyk-Laughton-Atkinson
Bahl-Geertsen
Lindell-Hakanpaa

Edges: The link between players if they have been deployed together for more than 600 minutes in a season at 5-on-5. 

Network Size

Nodes: ~2000, over a 10-year period.

Edges: 10000+, each player has at least one partner.

Questions we plan to ask and why we care?

Questions

- What players are the most/least effective at increasing the effectiveness of their linemates?
- Are there players that notably affect linemates in specific metrics such as defensive or offensive play?
- Are there potentially more effective player deployments that synergize more effectively?
- Can a player's synergistic effects with other players be inflating/ deflating their contract value?

Why do we care?

- Sports leagues are hyper competitive.
- more successful teams have immediate financial successes.