

Predicting Build Failures using Social Network Analysis on Developer Communication



Timo Wolf

Adrian Schröter

Daniela Damian

Thanh Nguyen

SIEMENS AG

Corporate Technology
Software & Engineering

Software Engineering Global interAction Lab – SEGAL

Computer Science Department
University of Victoria, BC Canada

SEGAL
Software Engineering Global interAction Labs

Introduction

Introduction

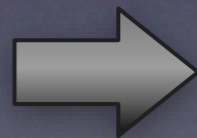


Introduction



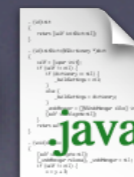
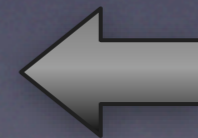
The damn build
broke again! WHY ?

Production



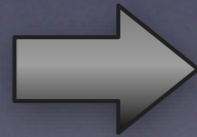
Build Failure

- Compile Error
- Test Failure



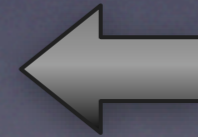
The damn build
broke again! WHY ?

DB changed!
But that's transparent.
Analyses takes time!
We should have talked earlier!



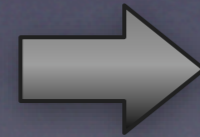
Build Failure

- Compile Error
- Test Failure



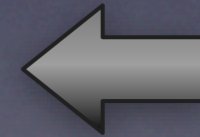
The damn build
broke again! WHY ?

DB changed!
But that's transparent.
Analyses takes time!
We should have talked earlier!



Build Failure

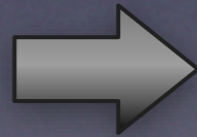
- Compile Error
- Test Failure



- Collaboration essential in software development
- Communication problems lead to coordination and integration failures

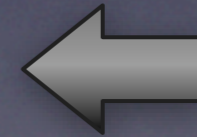
The damn build
broke again! WHY ?

DB changed!
But that's transparent.
Analyses takes time!
We should have talked earlier!



Build Failure

- Compile Error
- Test Failure



- Collaboration essential in software development
- Communication problems lead to coordination and integration failures

Do specific communication behaviors enable effective and successful coordination?

Communication for collaboration



Communication for collaboration

**Task related
Content &
Information**

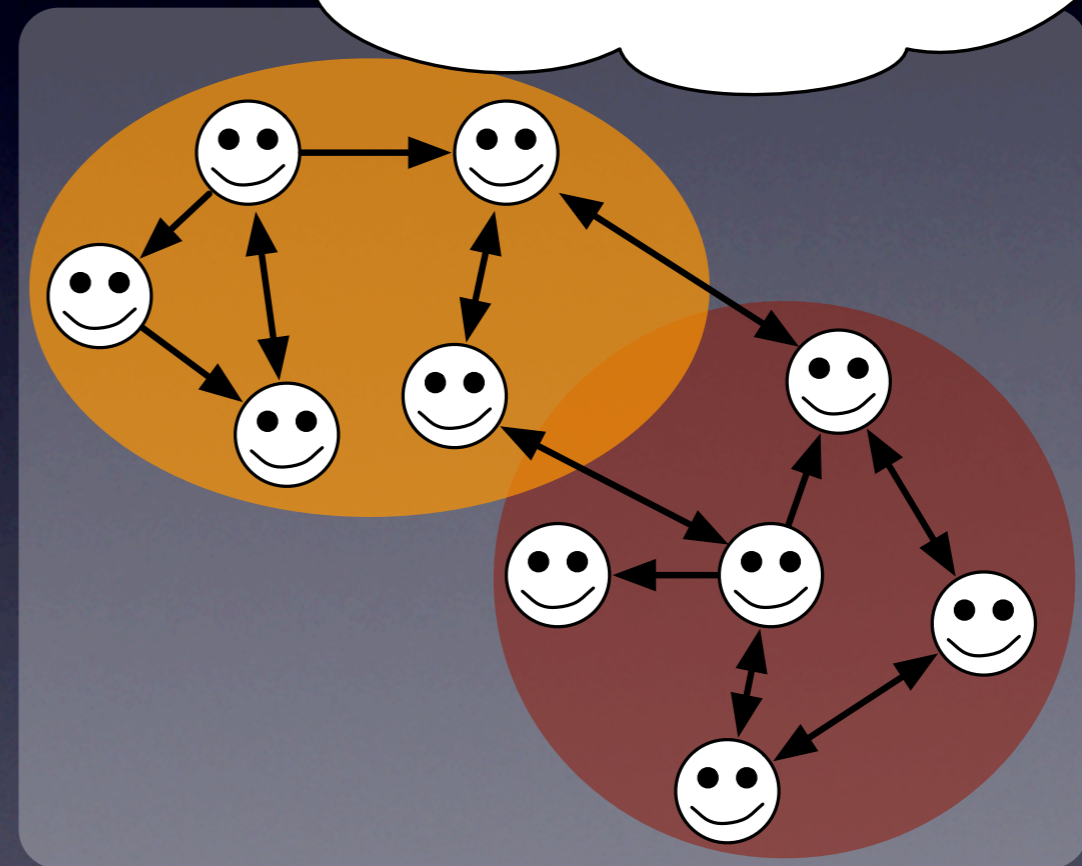
- Communication:
 - Task related content



Communication for collaboration

**Task related
Content &
Information**

- Communication:
 - Task related content
 - Structure

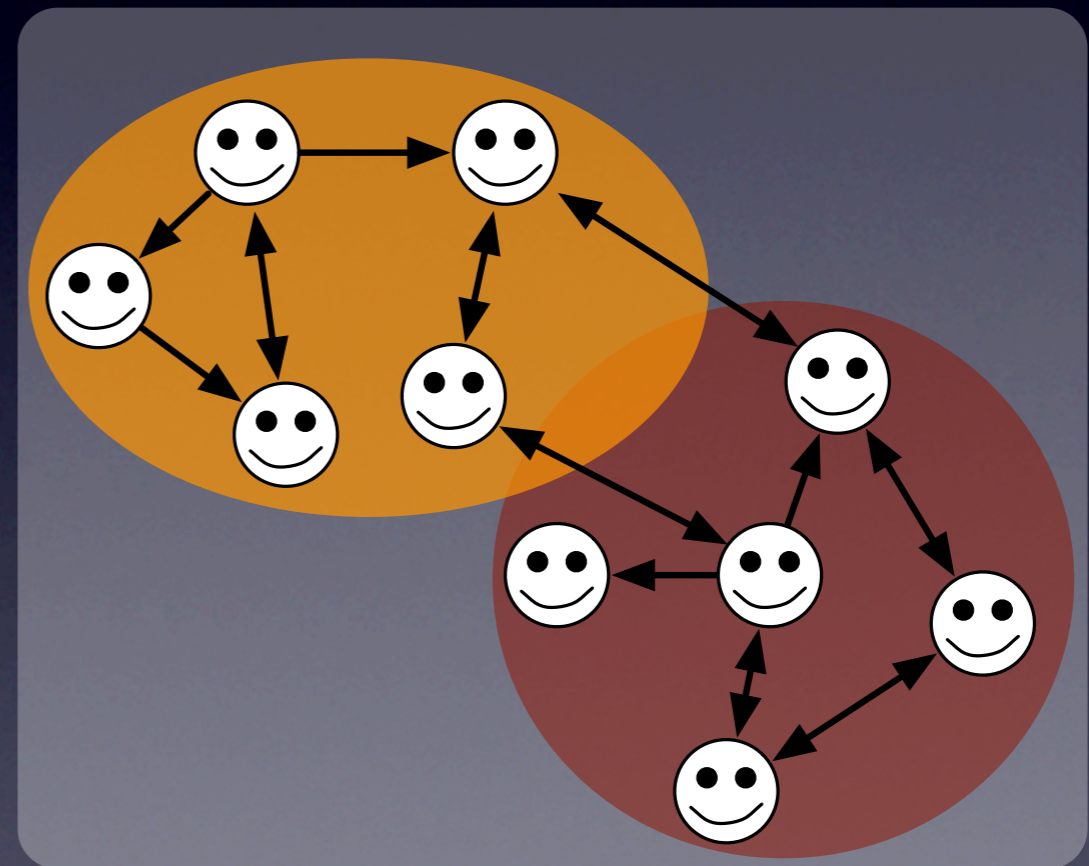


Communication for collaboration

- Communication:
 - Task related content
 - Structure

Our focus:

**Communication Structure and
relationship to coordination outcome**



Study Settings: IBM Jazz

Study Settings: IBM Jazz



& IBM Research

Study Settings: IBM Jazz



151 developers



47 functional teams

Study Settings: IBM Jazz

7 development sites
in USA, Canada, Europe

Study Settings: IBM Jazz

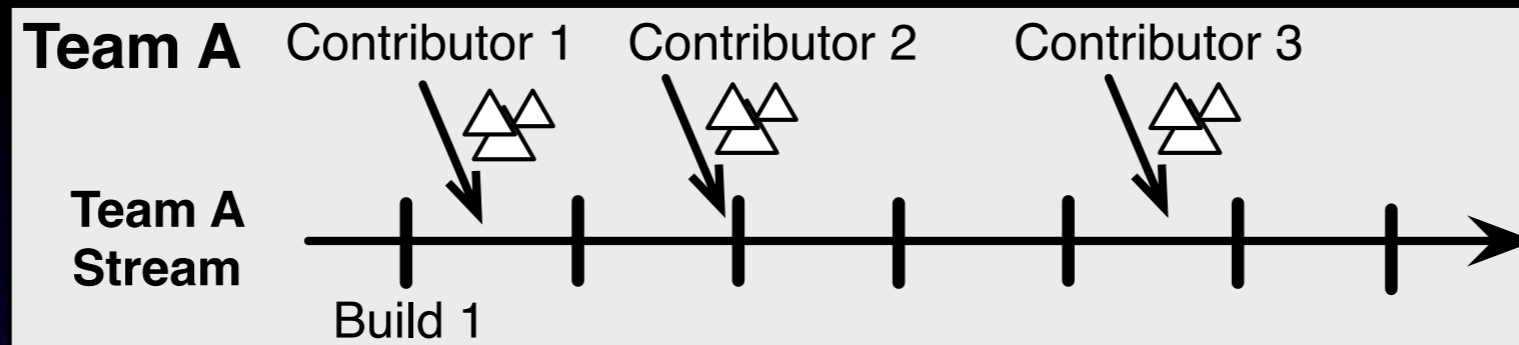


Release 1.0

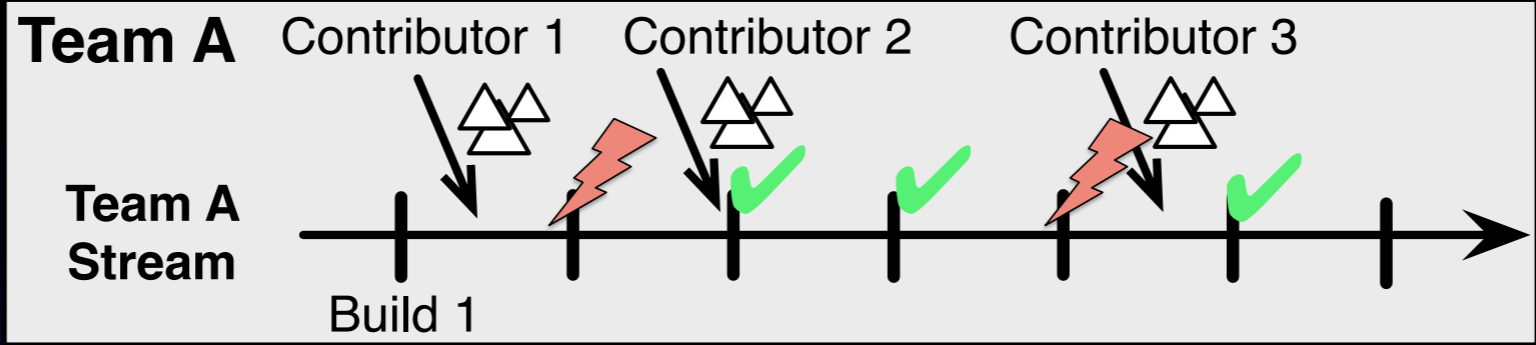


Coordination & Integration in Jazz

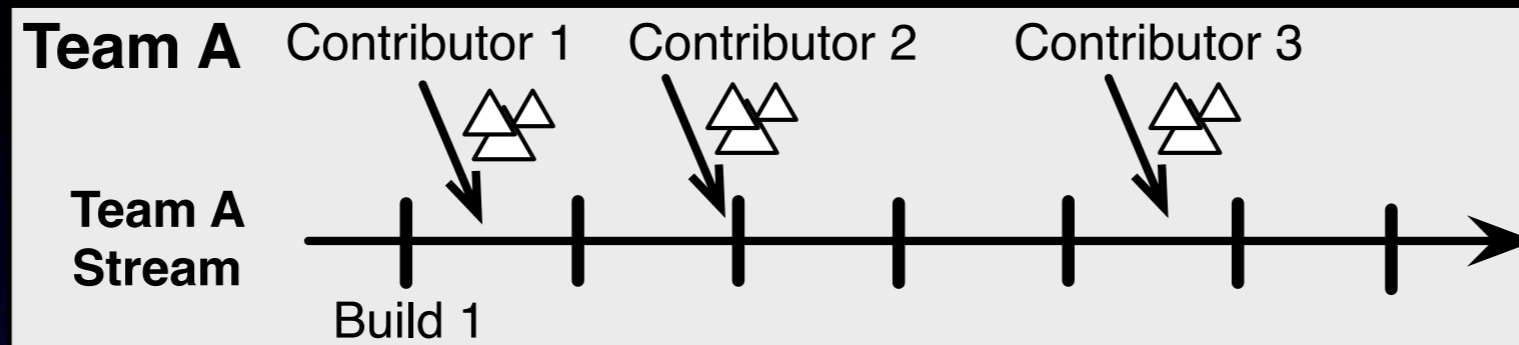
Coordination & Integration in Jazz



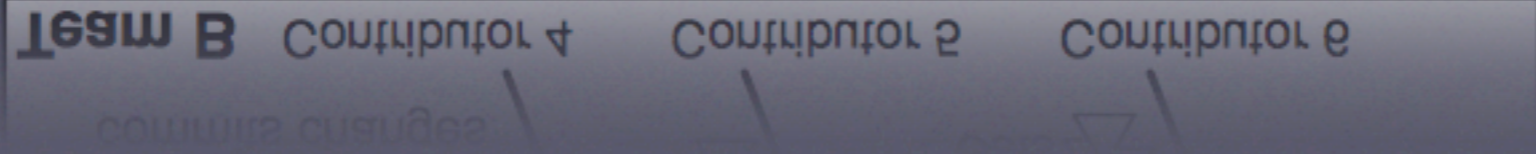
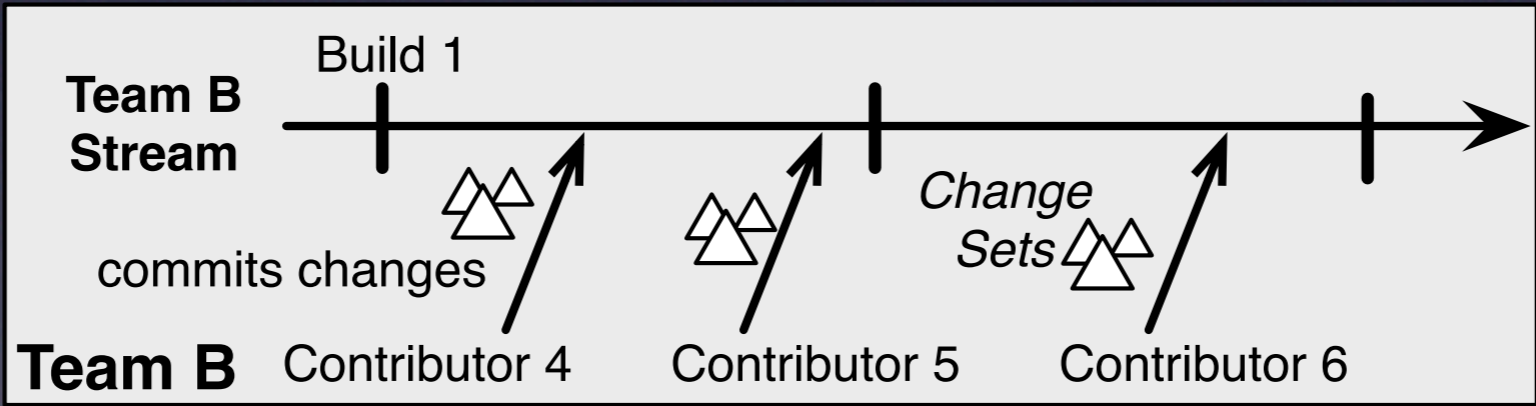
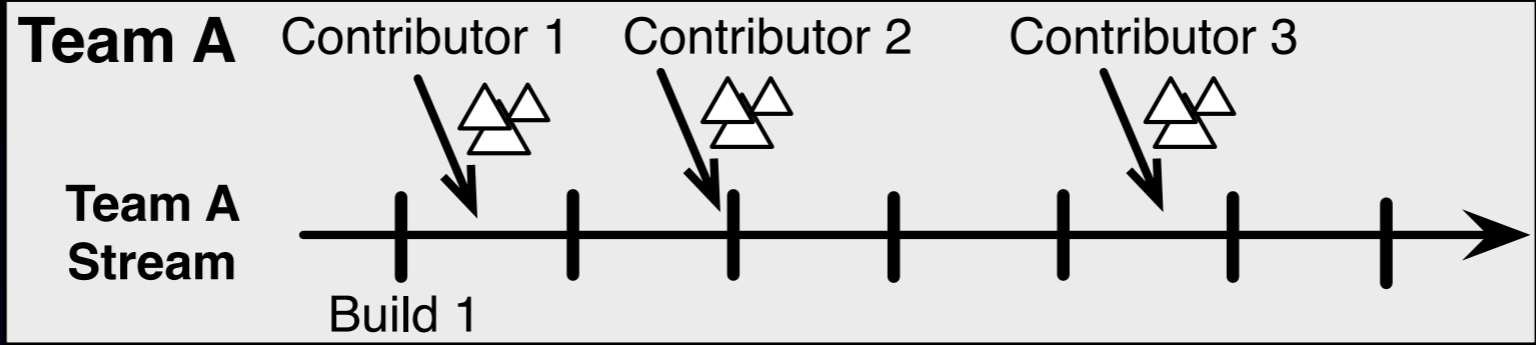
Coordination & Integration in Jazz



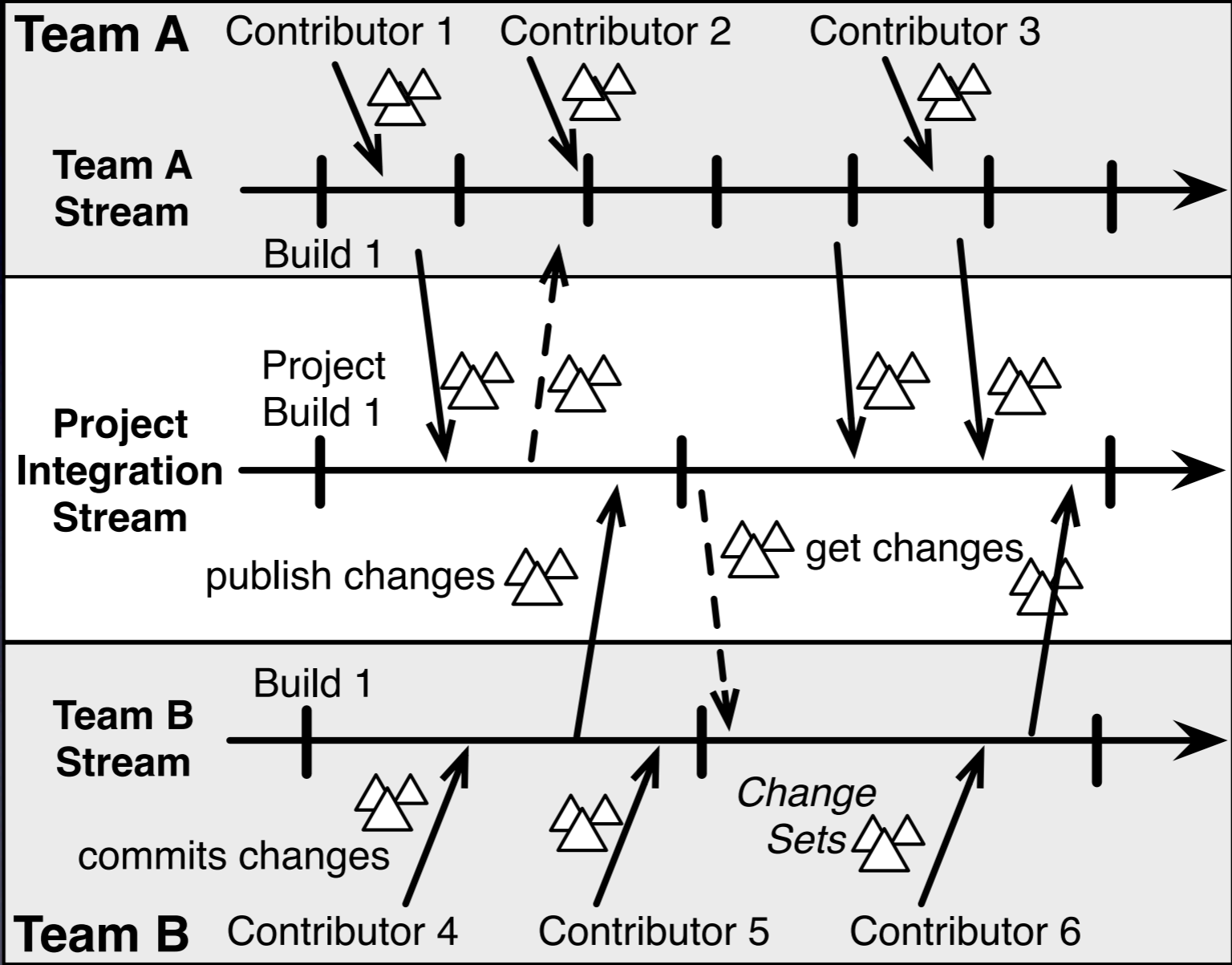
Coordination & Integration in Jazz



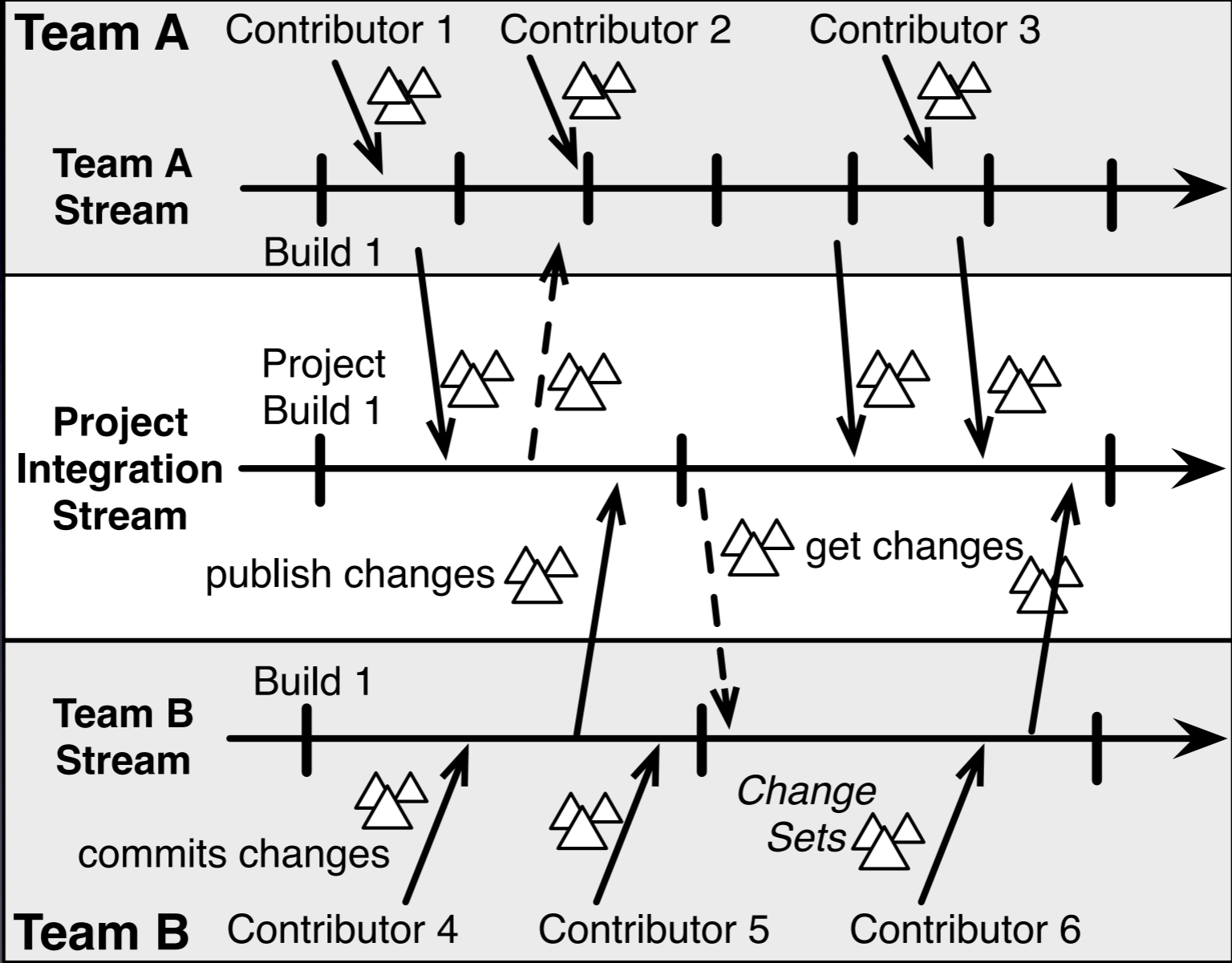
Coordination & Integration in Jazz



Coordination & Integration in Jazz



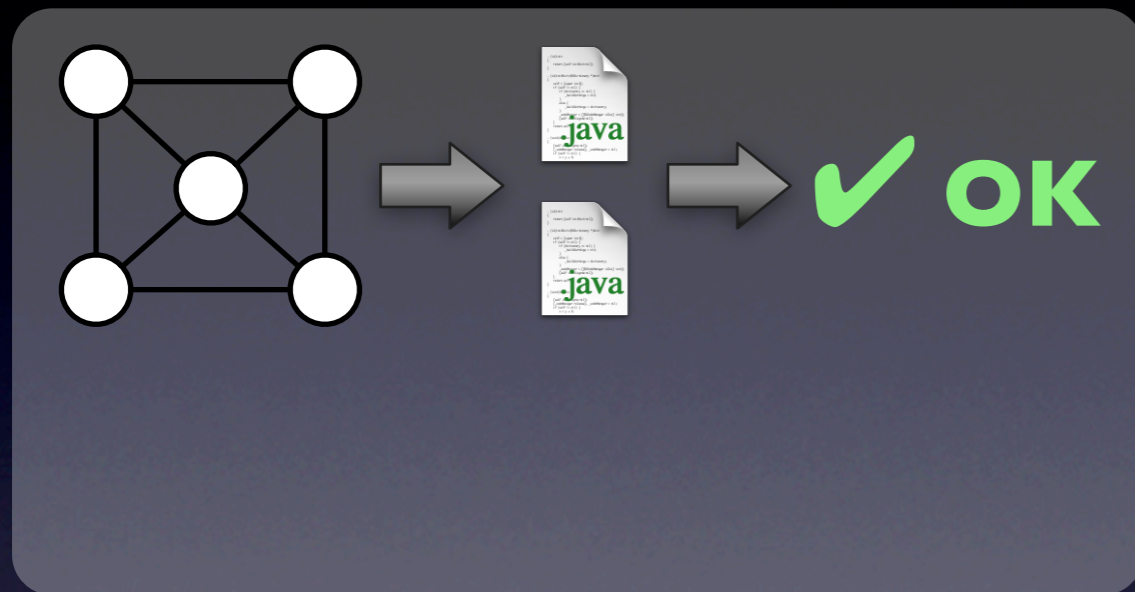
Coordination & Integration in Jazz



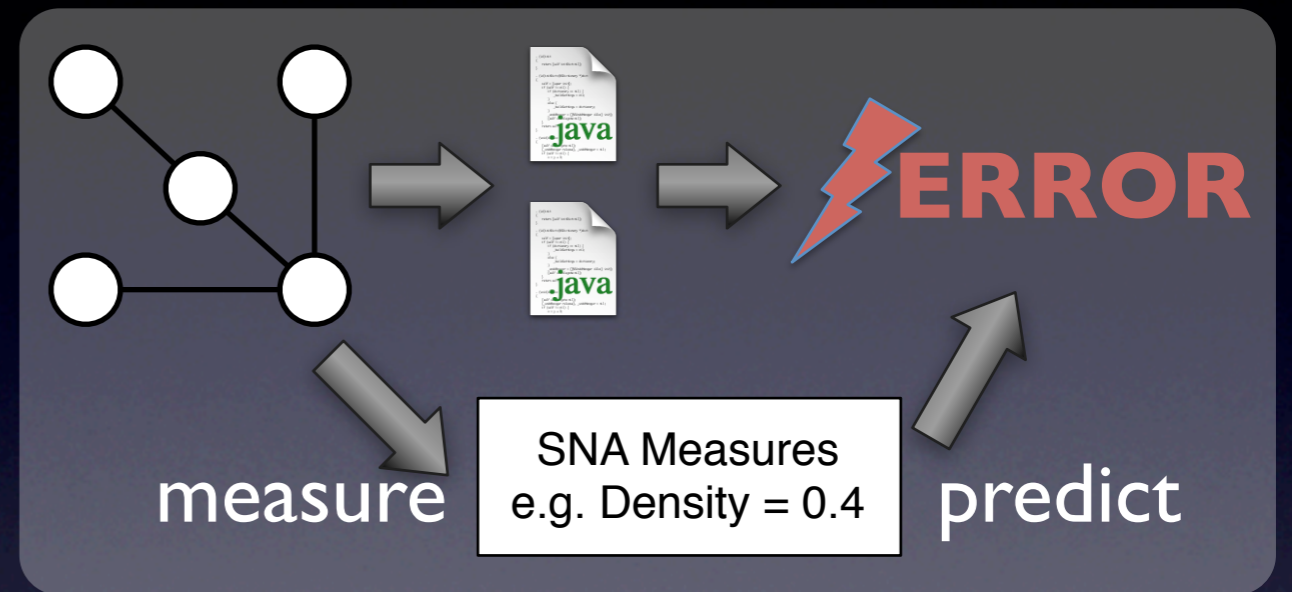
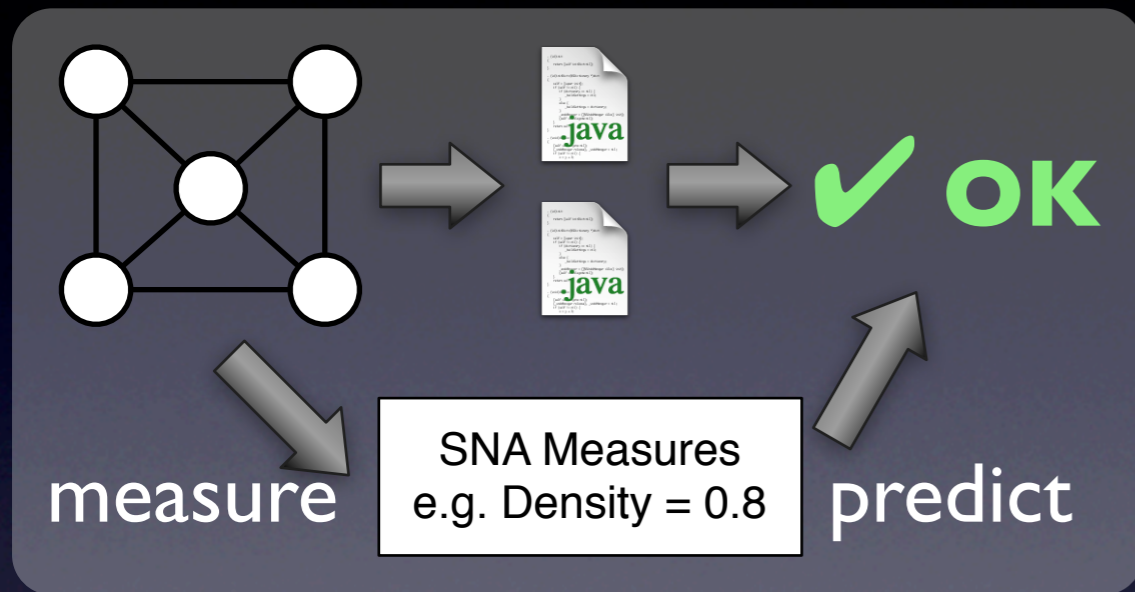
Coordination Outcome: Build Result

Research Questions

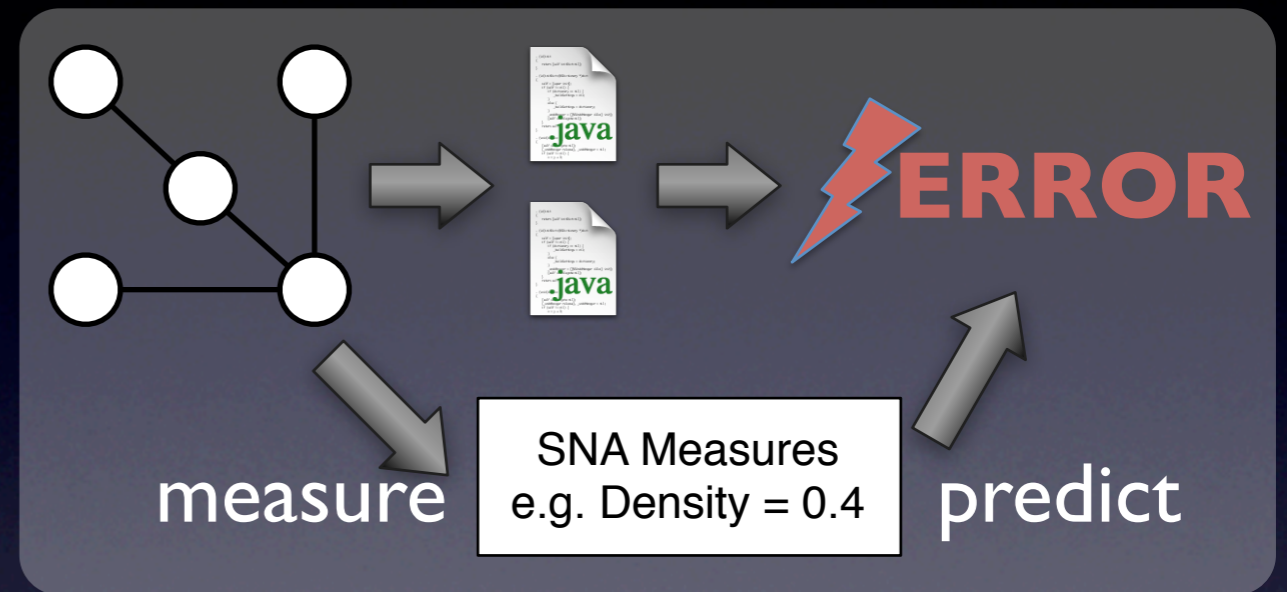
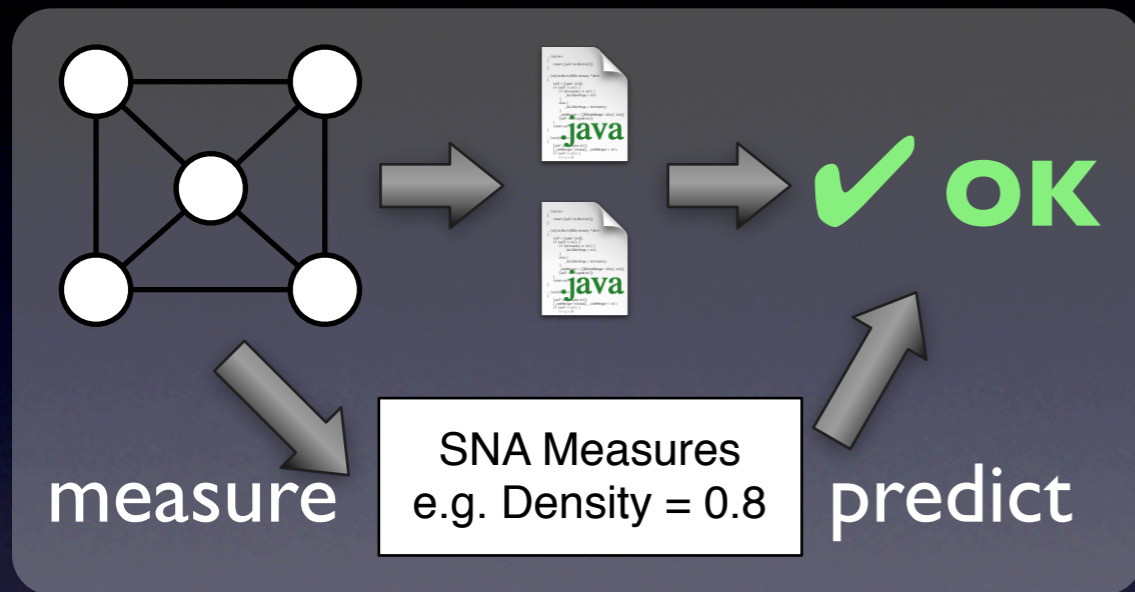
Research Questions



Research Questions

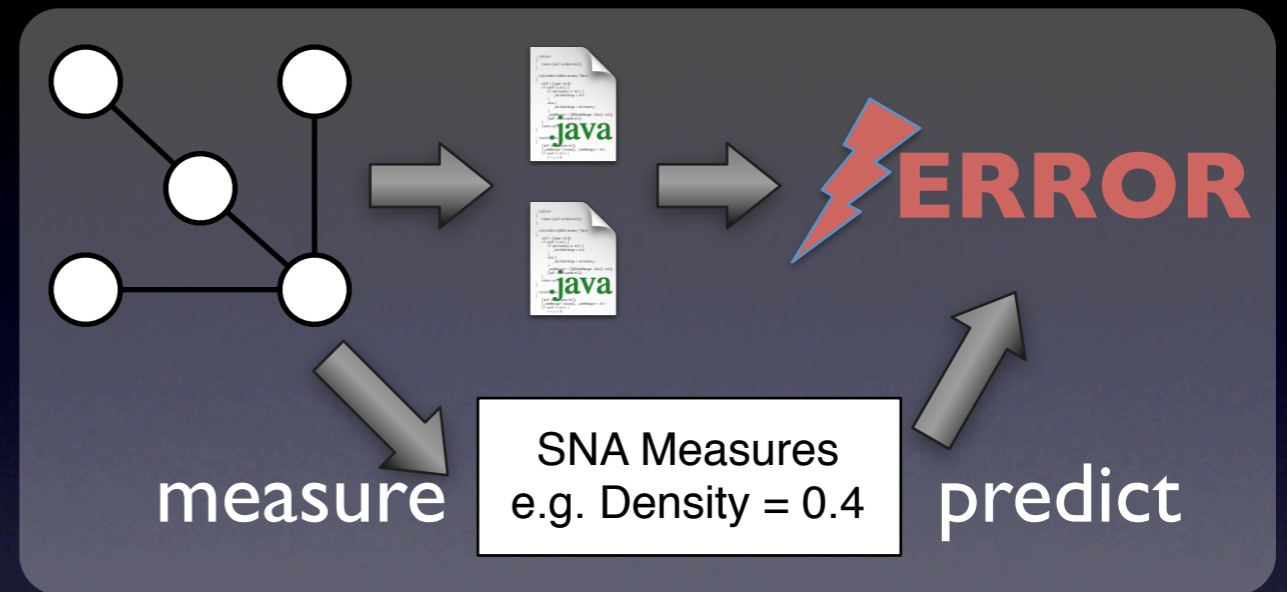
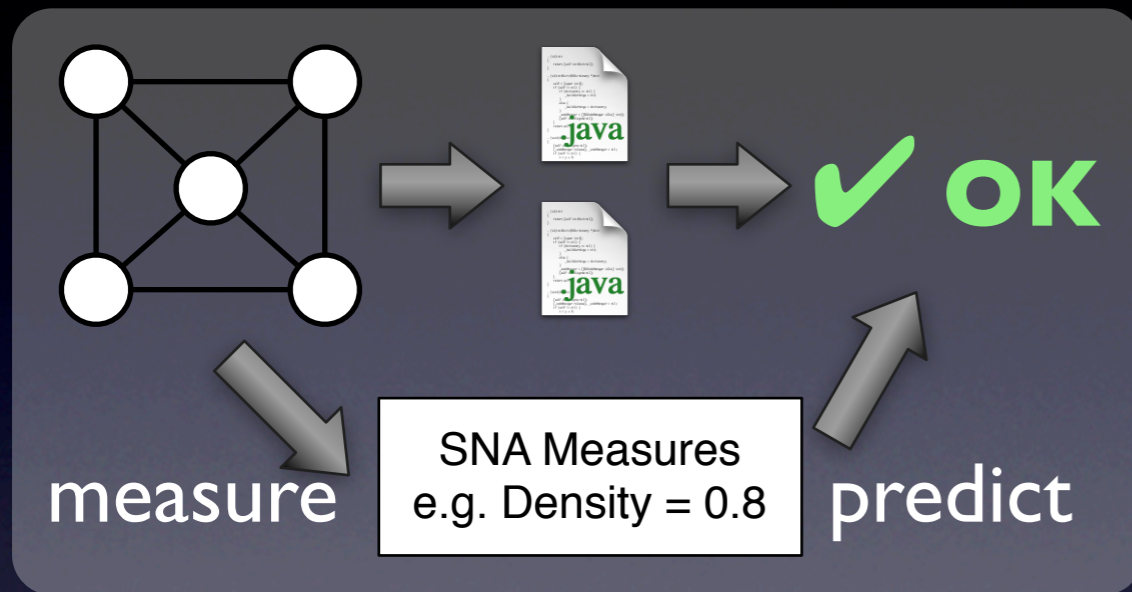


Research Questions



- RQ1: Can individual measures of communication structure predict integration failure?

Research Questions



- RQ1: Can individual measures of communication structure predict integration failure?
- RQ2: Can the combination of communication structure measures predict integration failure?

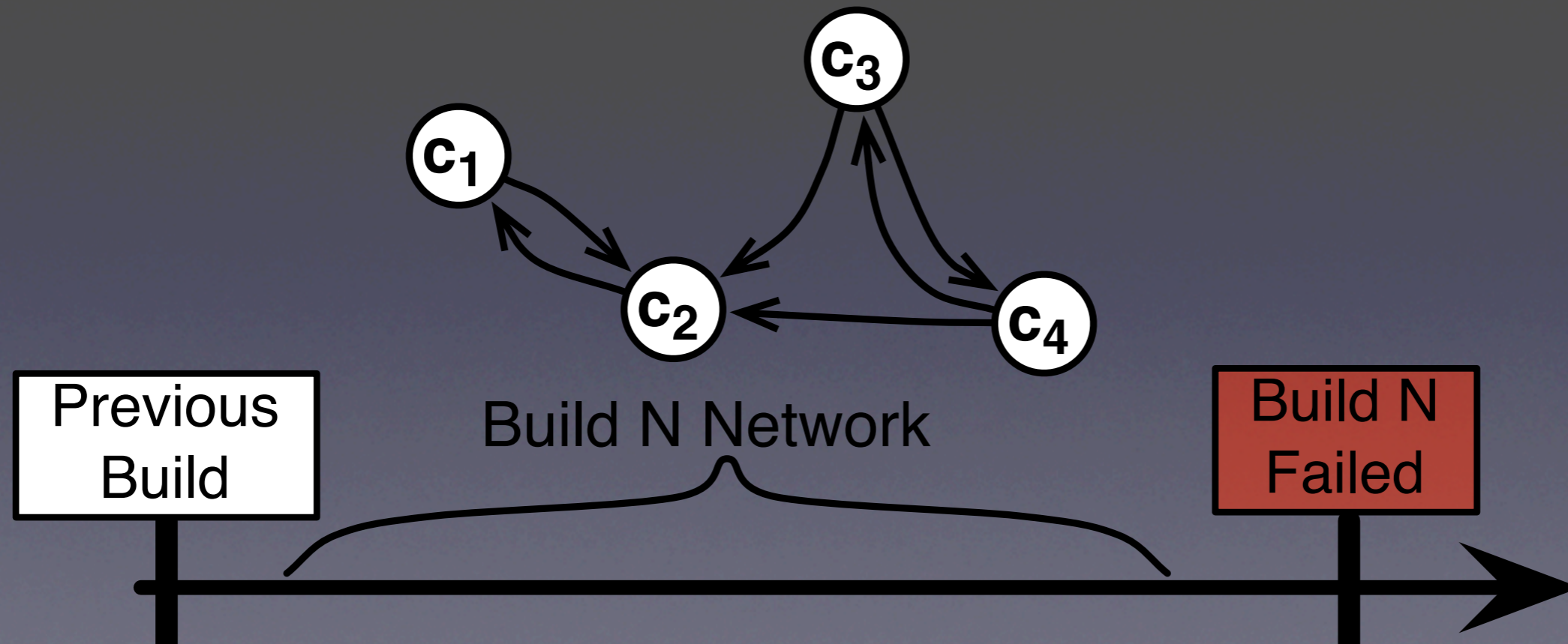
Constructing Social Networks in Jazz



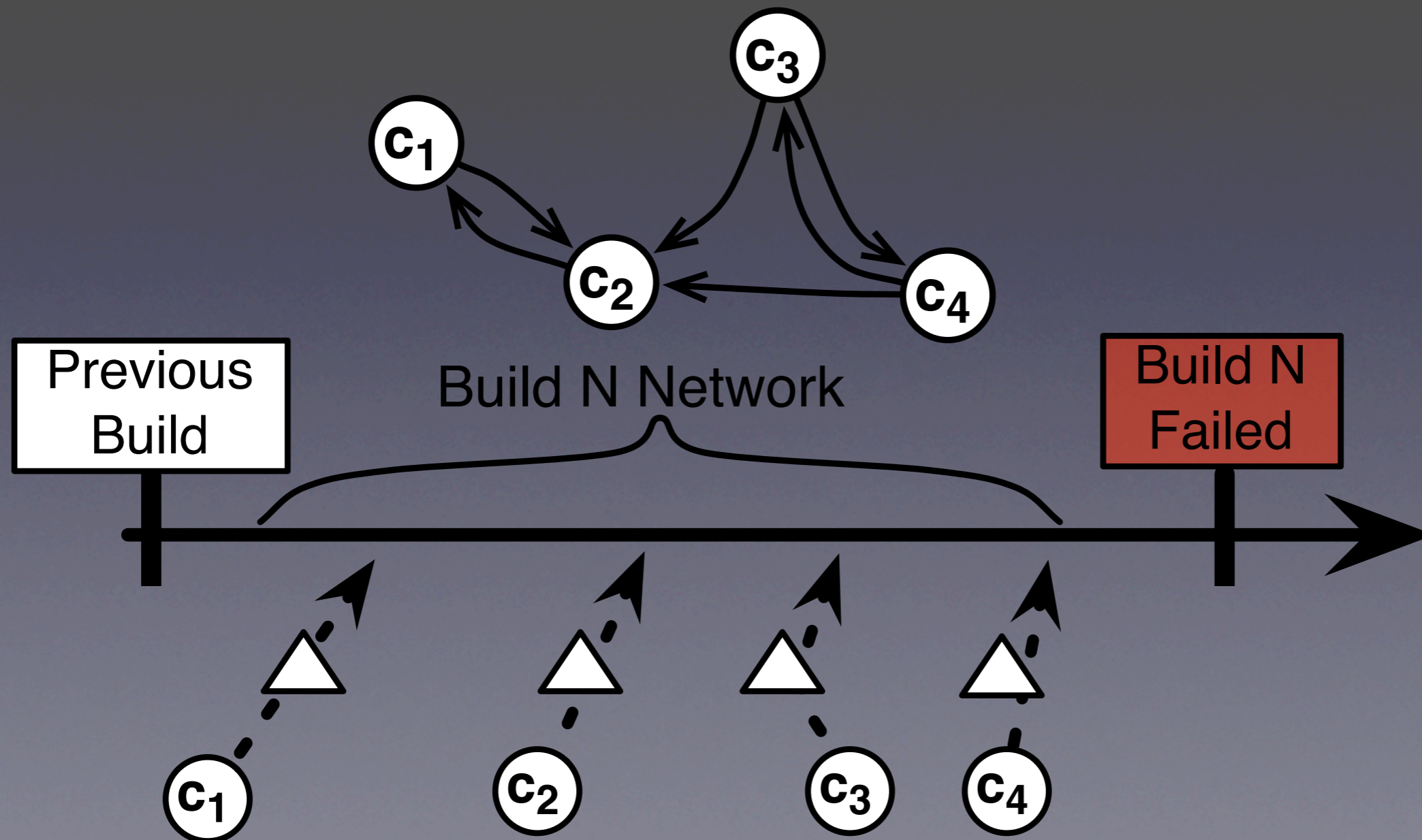
Constructing Social Networks in Jazz



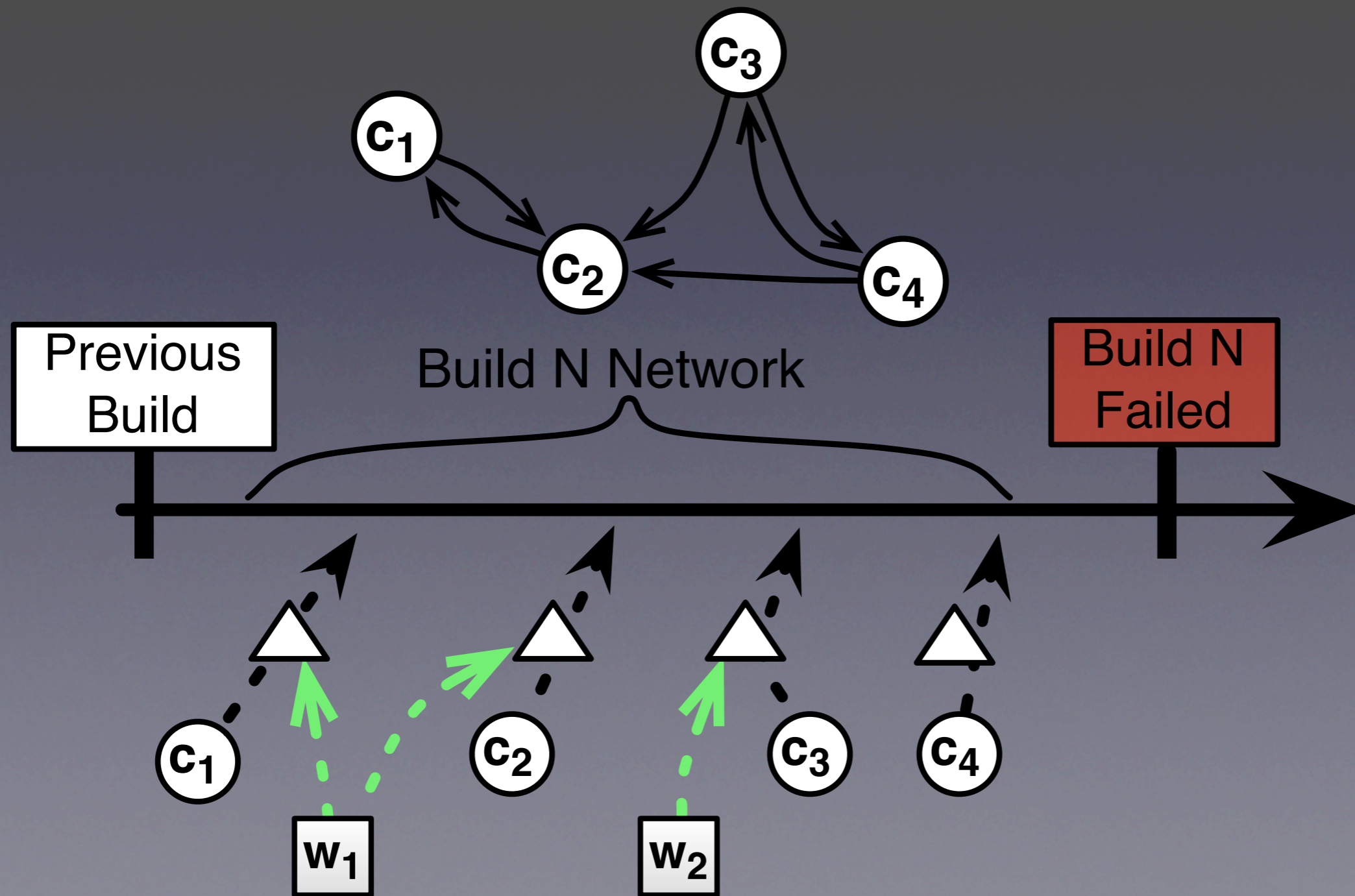
Constructing Social Networks in Jazz



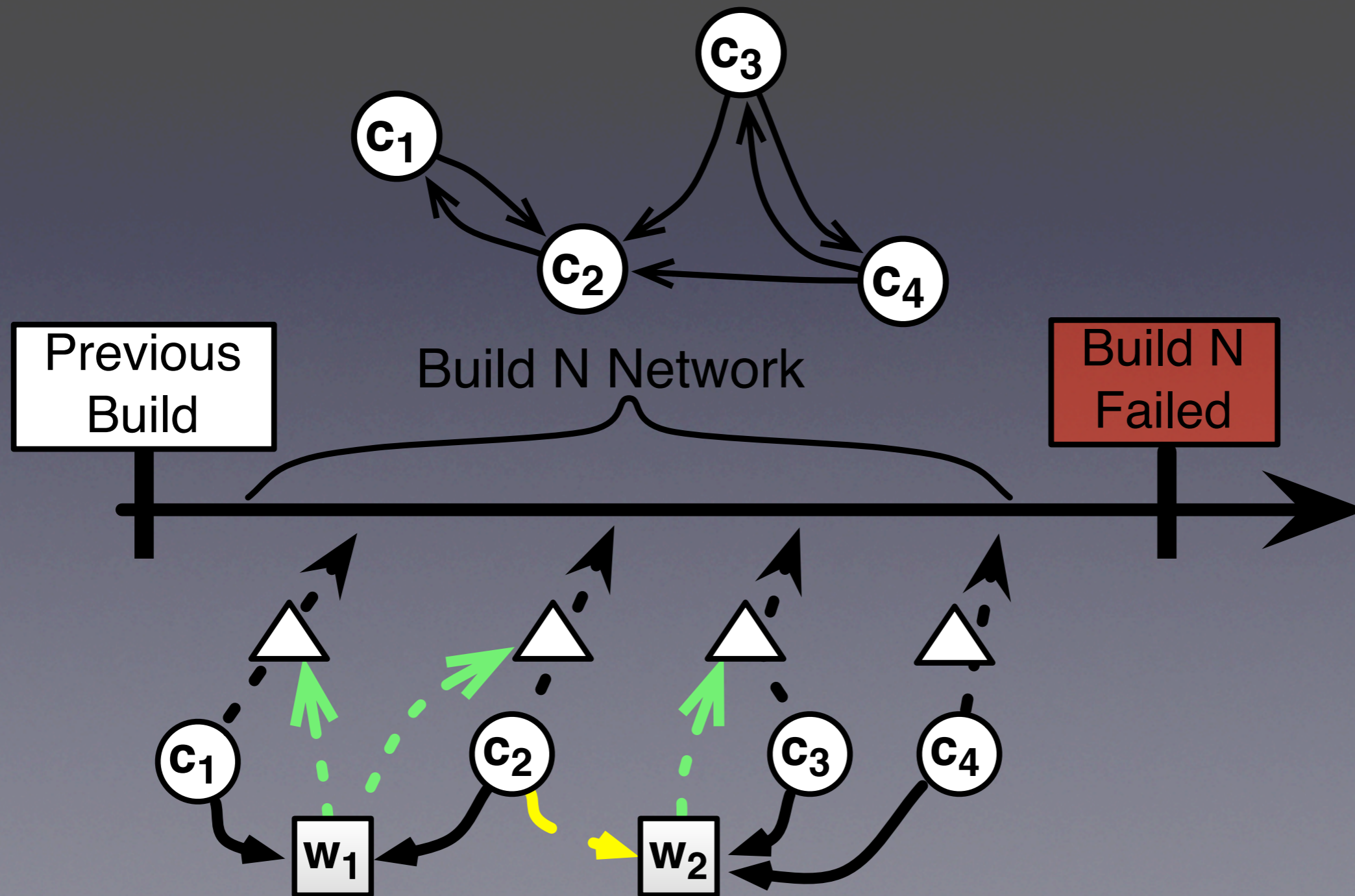
Constructing Social Networks in Jazz



Constructing Social Networks in Jazz

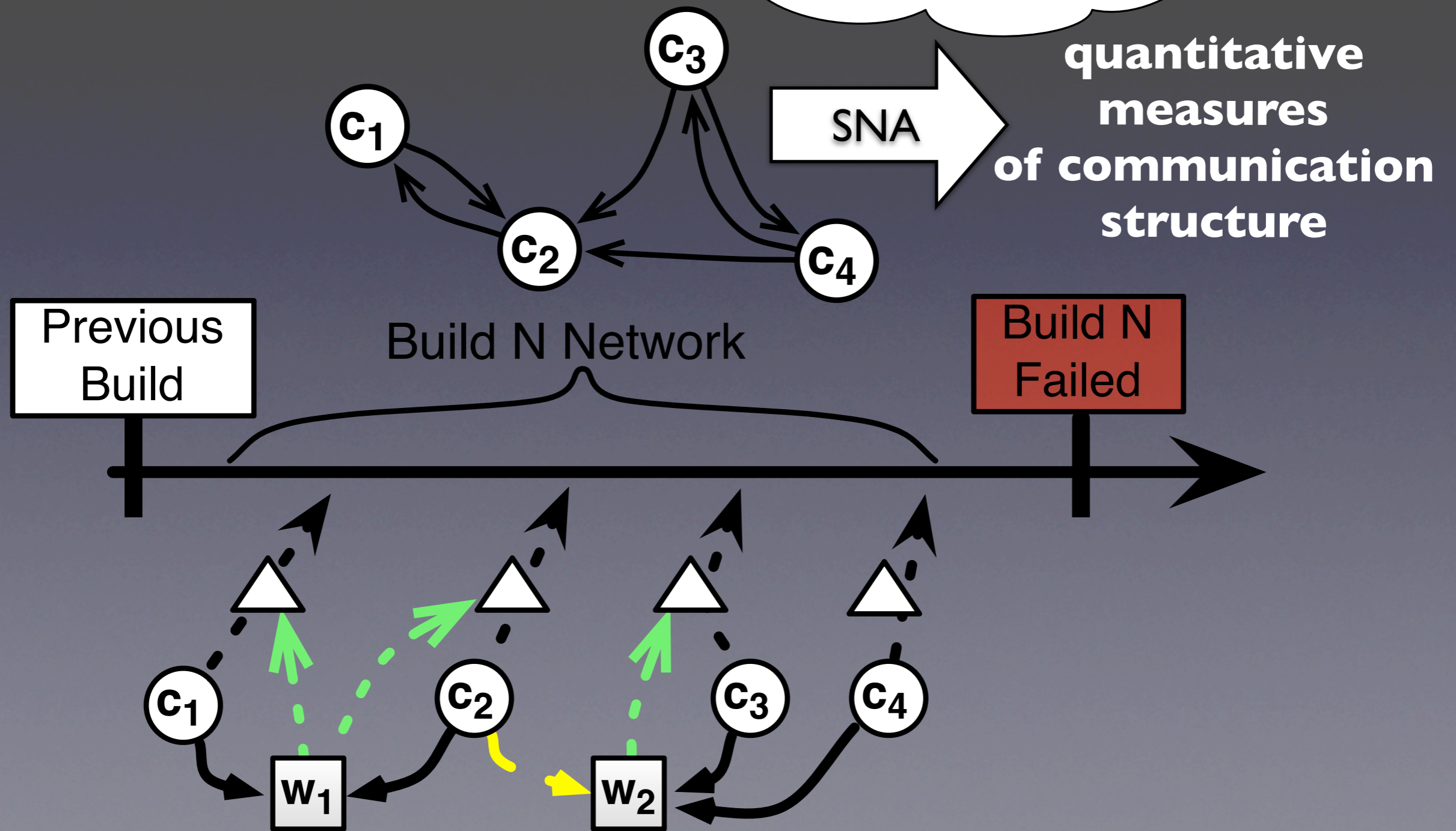


Constructing Social Networks in Jazz



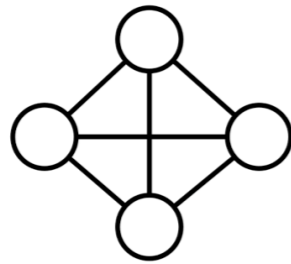
Constructing Social

Social Network Analysis (SNA) in Jazz

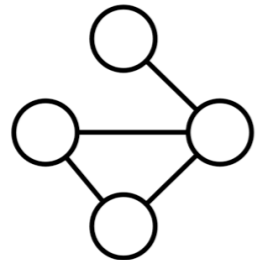


SNA Measures

Density

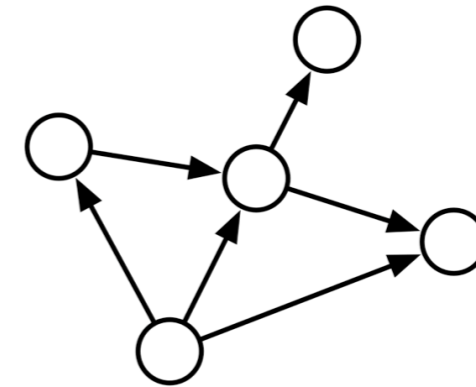


$$6/6 = 1$$

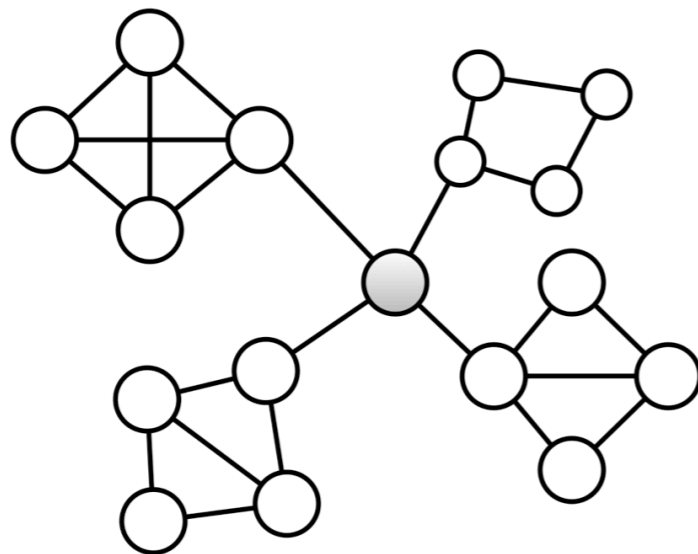


$$4/6 = 0.66$$

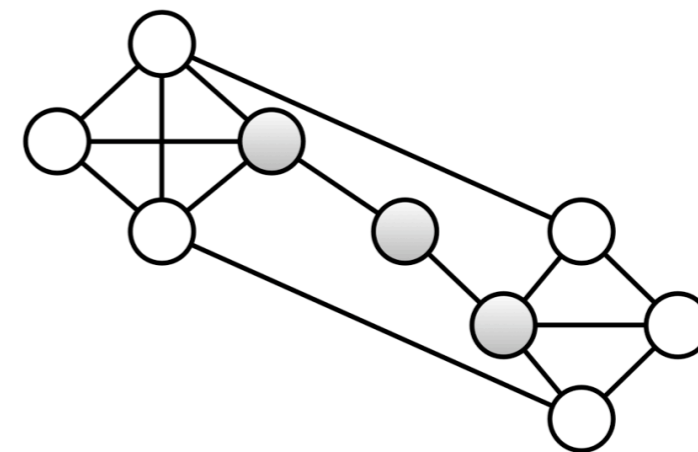
Degree Centrality



Structural Holes

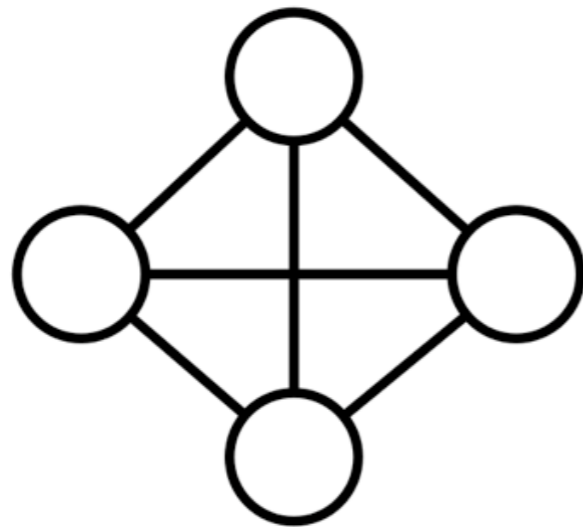


Betweenness Centrality

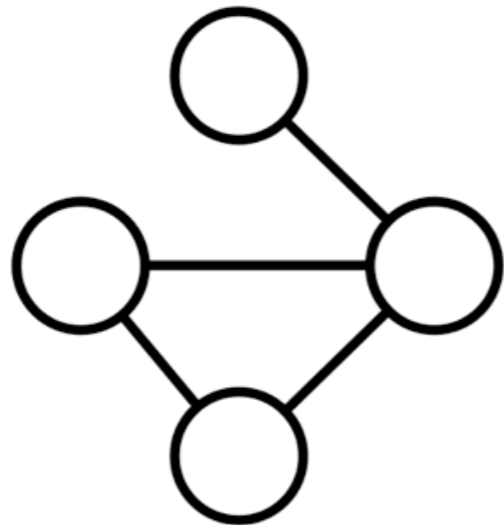


SNA Measures

Density

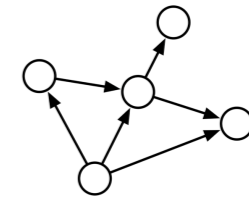


$$6/6 = 1$$

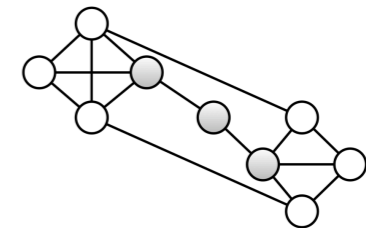


$$4/6 = 0.66$$

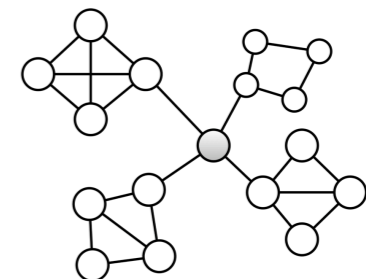
Degree Centrality



Betweenness Centrality

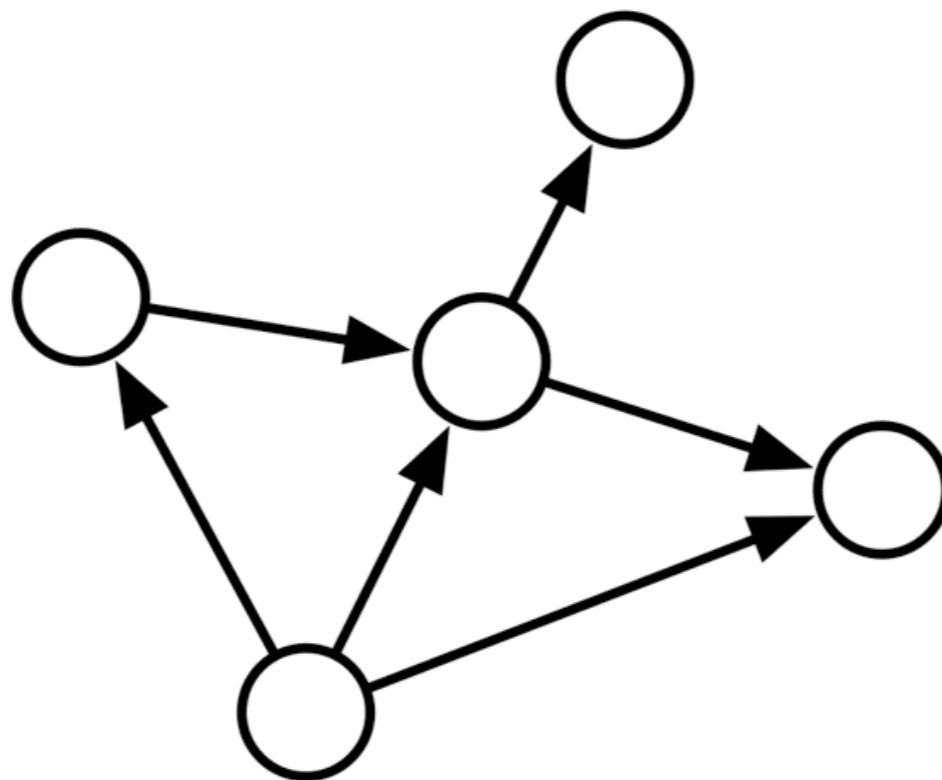


Structural Holes

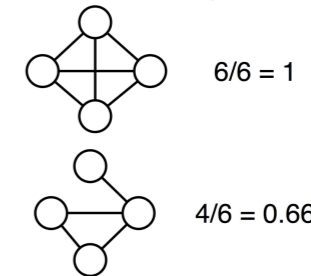


SNA Measures

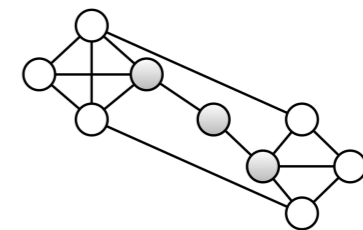
Degree Centrality



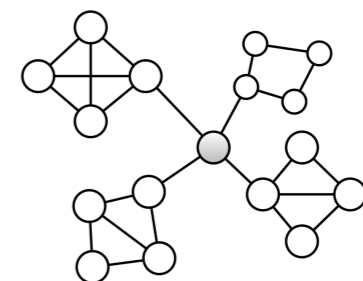
Density



Betweenness Centrality

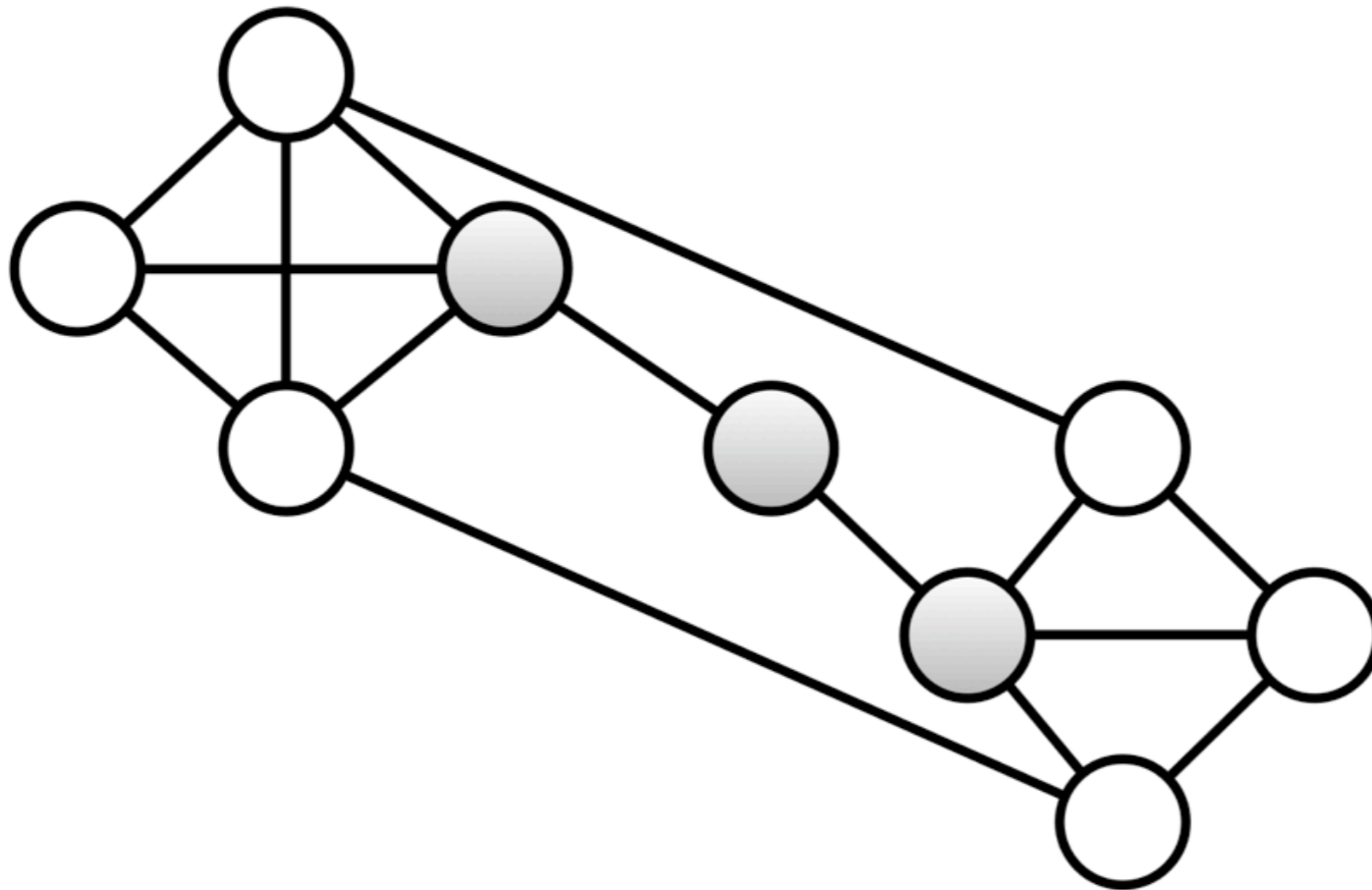


Structural Holes

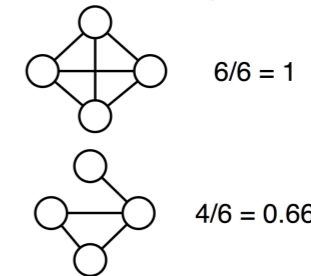


SNA Measures

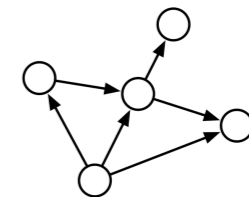
Betweenness Centrality



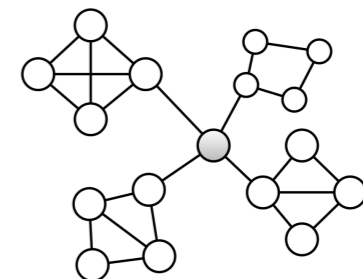
Density



Degree Centrality

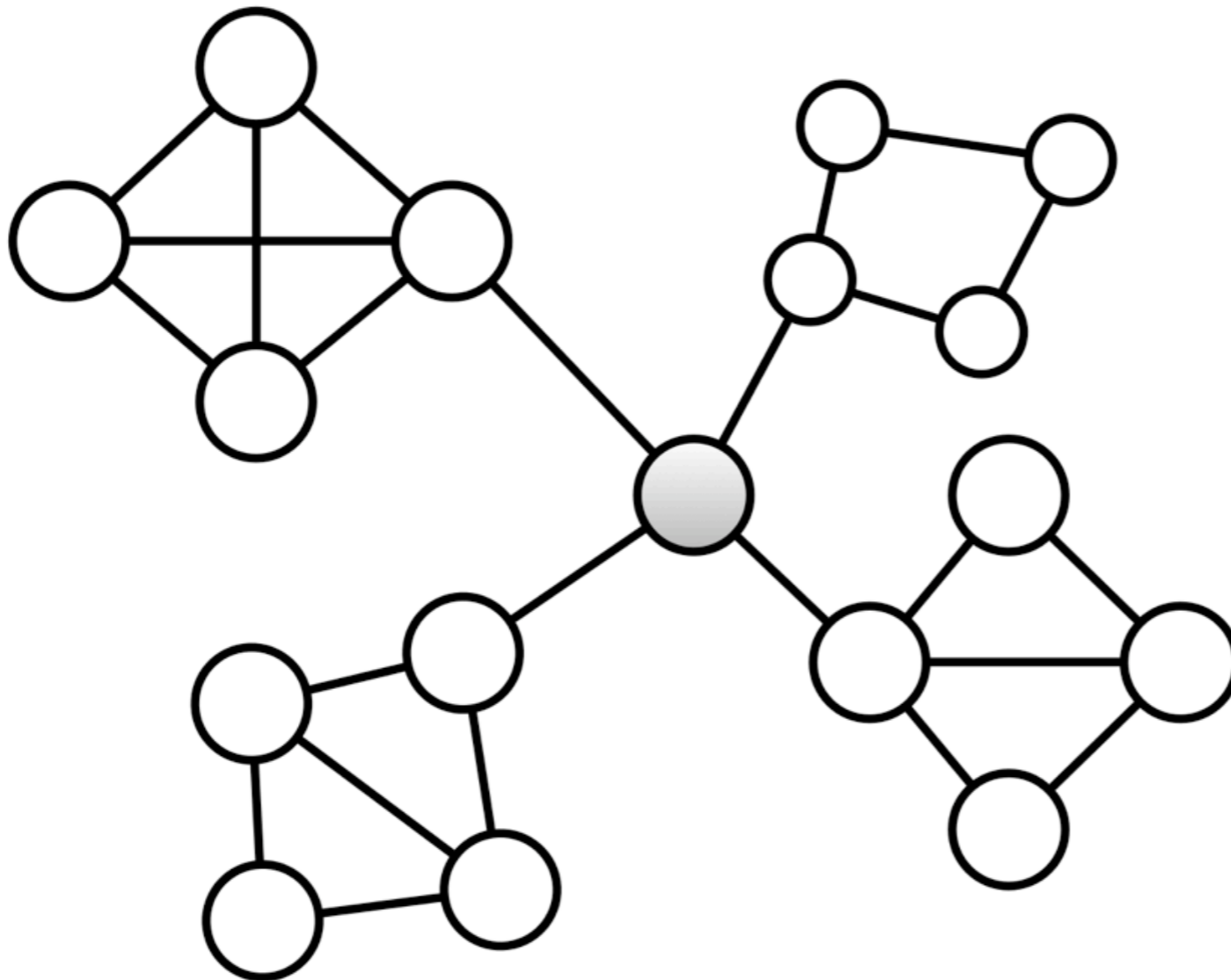


Structural Holes

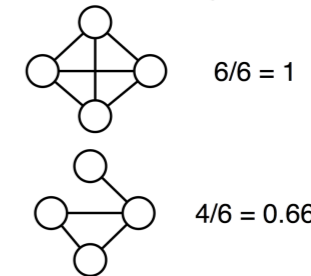


SNA Measures

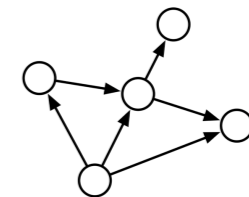
Structural Holes



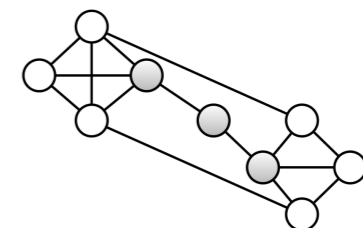
Density



Degree Centrality



Betweenness Centrality

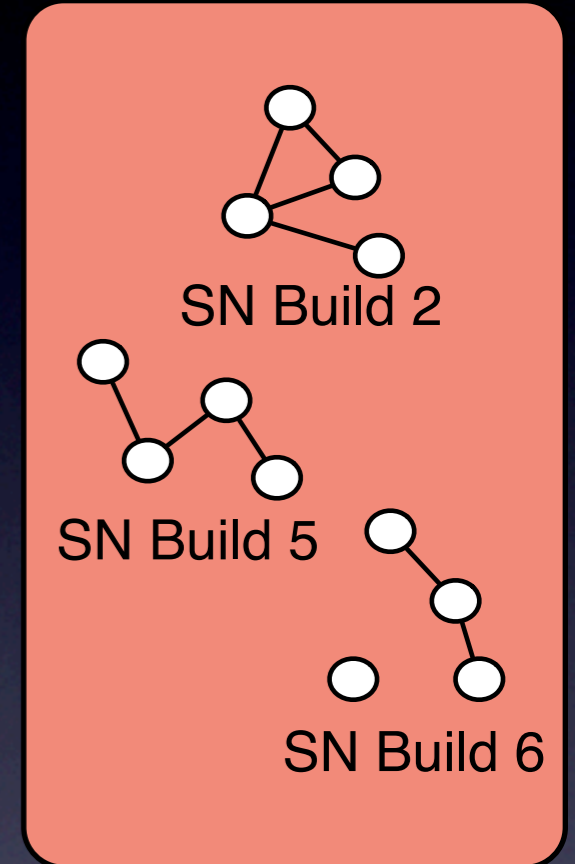
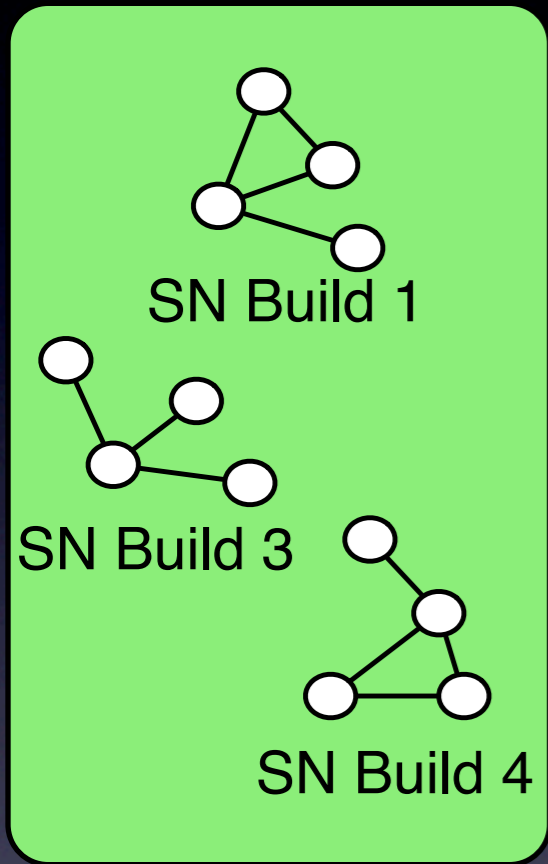


Data Collection & Selection

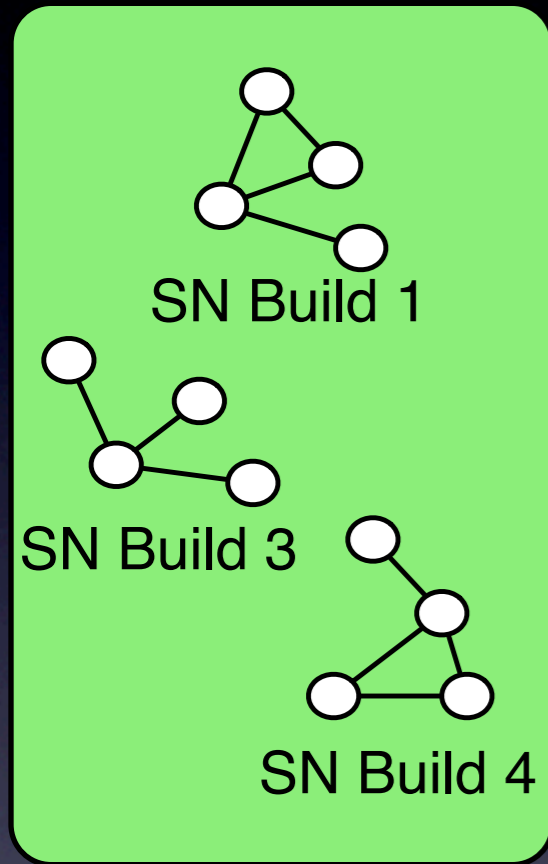
- Time Range: Nov. 5, 2007 - Feb. 26, 2008
 - 1288 Build Results
 - 25713 Work Items
 - 71019 Comments
- Selected Build Results of 5 Teams
- Selected 3 Integration Builds (nightly, weekly, beta)

RQ I: individual measure prediction?

RQ I: individual measure prediction?

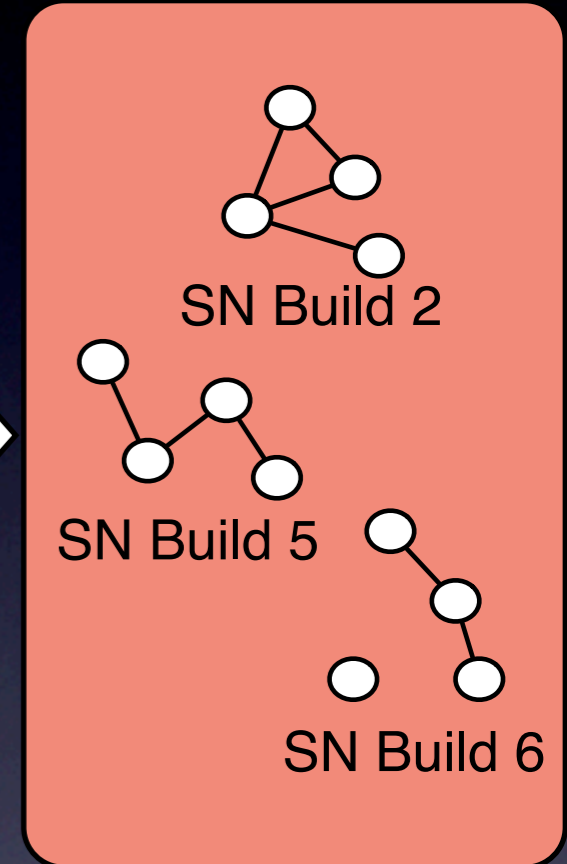


RQ1: individual measure prediction?



Social Network Measures

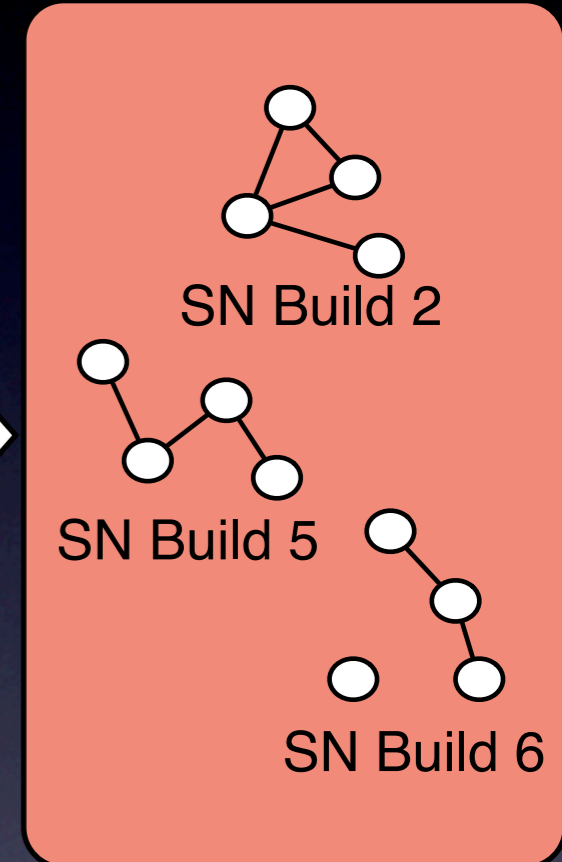
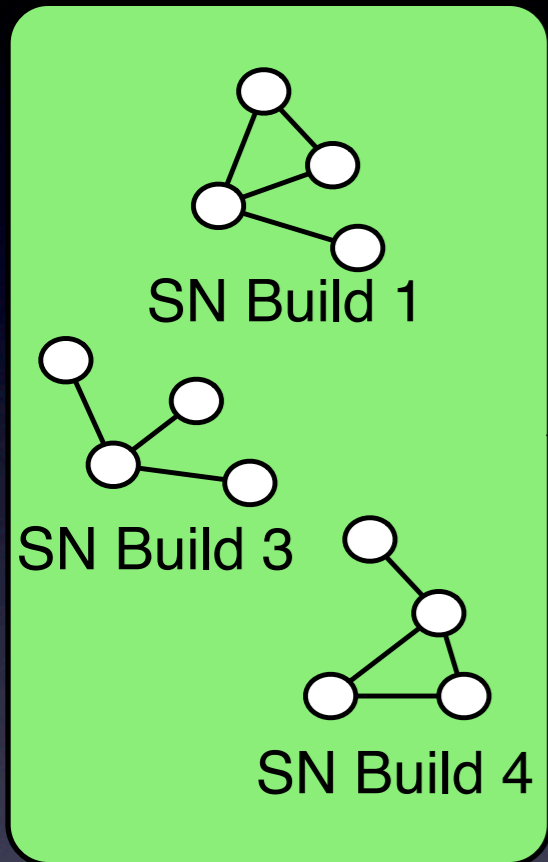
- **Density**
for Completeness
- **Degree Centrality**
for Activity
- **Betweenness Centrality**
for Control
- **Structural Holes**
for Redundancy



RQ1: individual measure prediction?

Social Network Measures

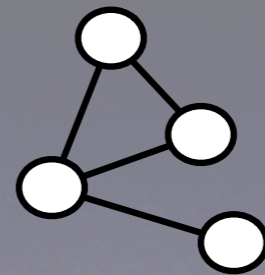
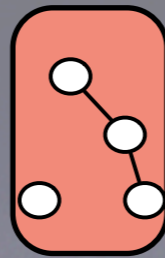
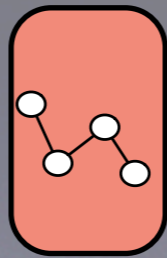
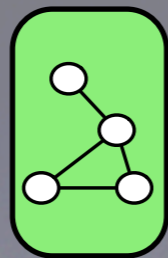
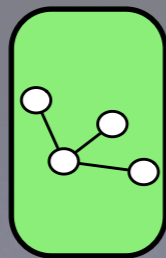
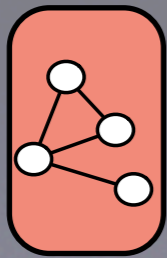
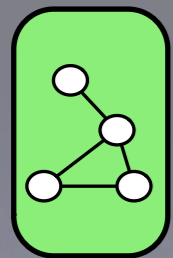
- Density for Completeness
- Degree Centrality for Activity
- Betweenness Centrality for Control
- Structural Holes for Redundancy



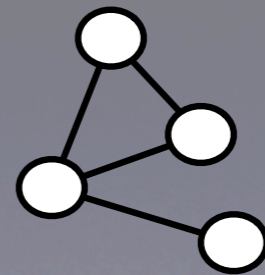
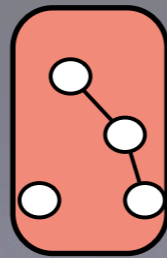
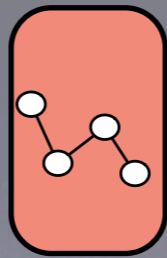
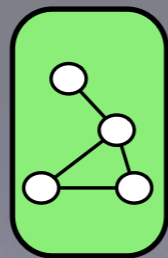
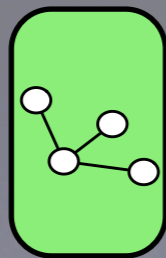
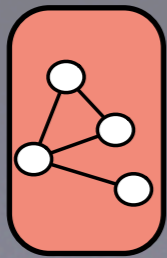
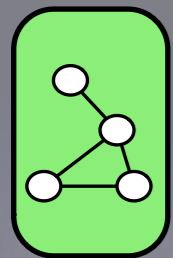
No statistically significant difference!

RQ2: multi measures prediction?

RQ2: multi measures prediction?



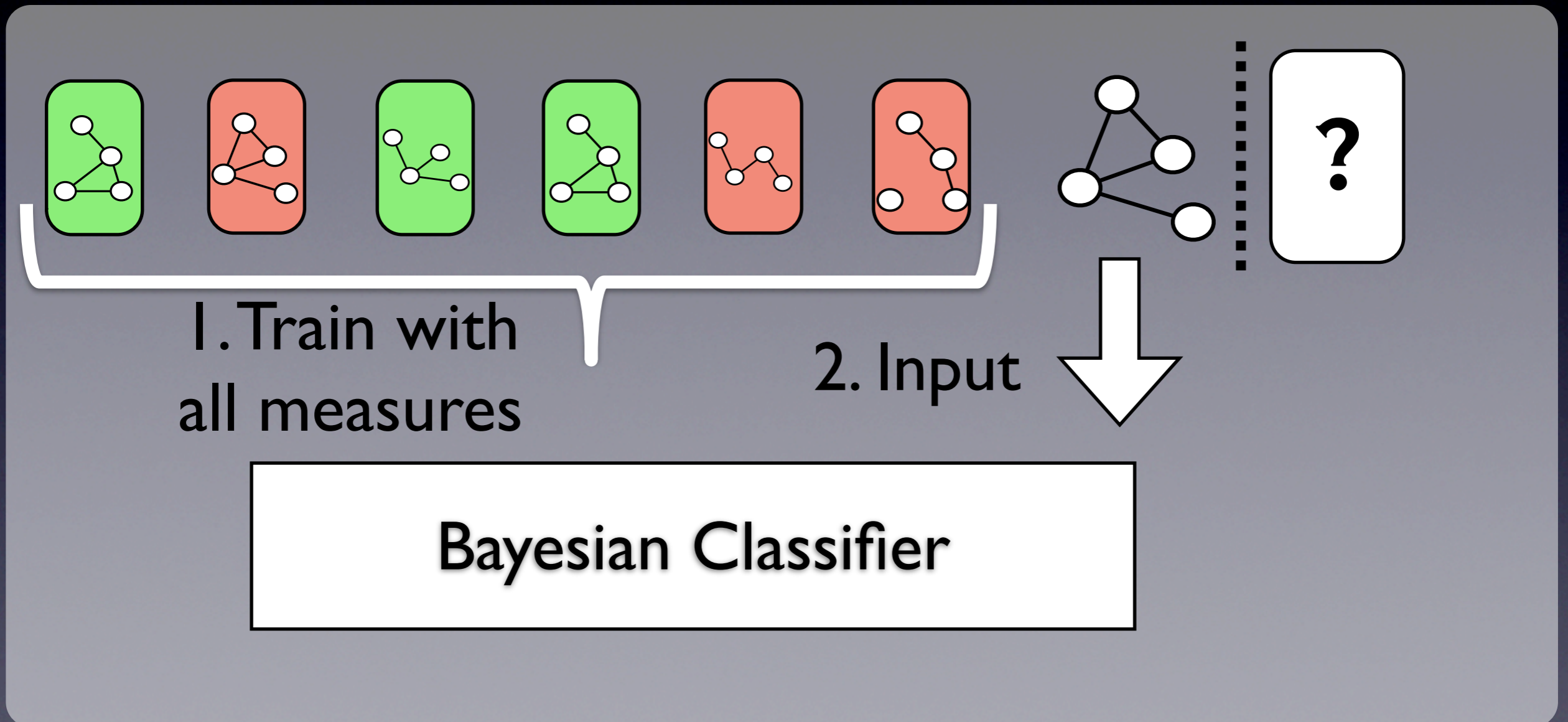
RQ2: multi measures prediction?



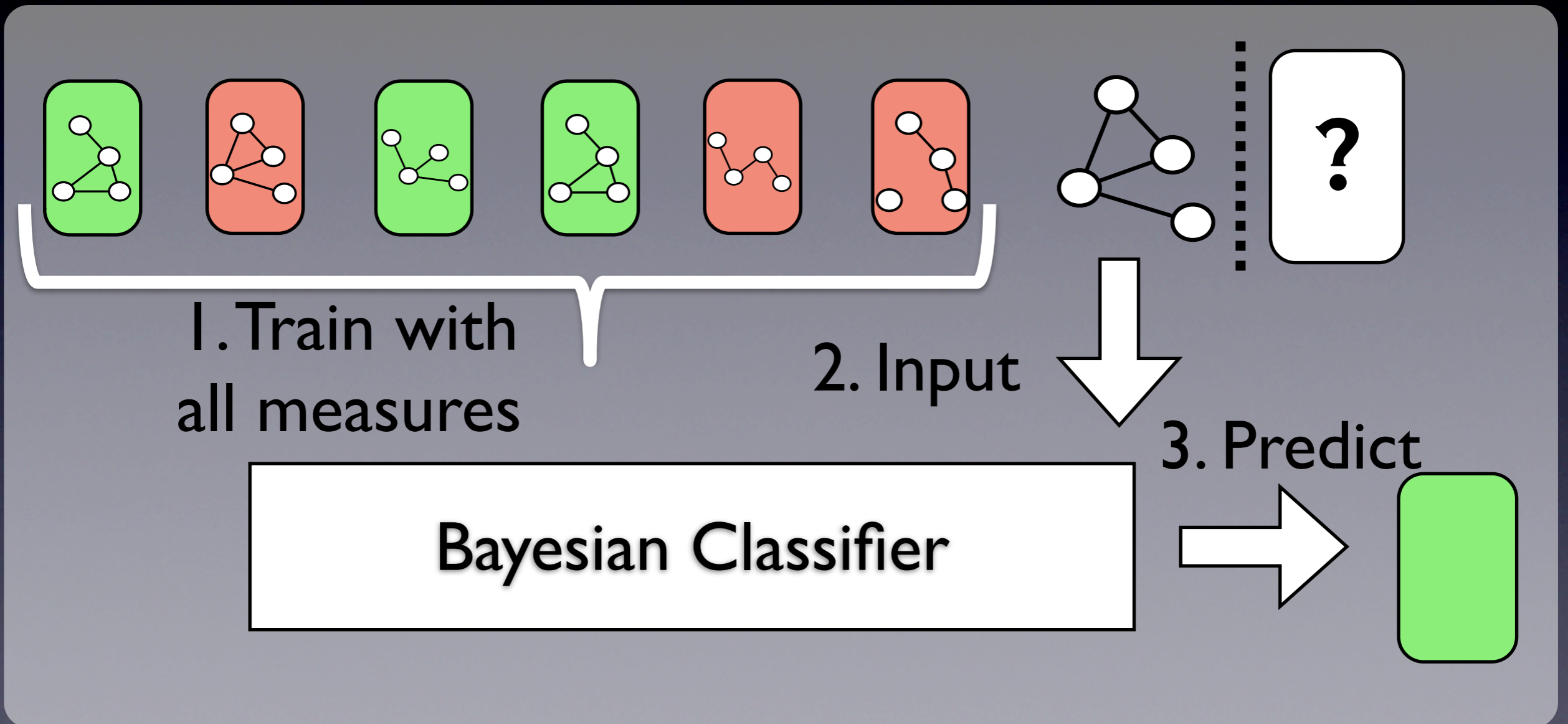
I. Train with
all measures

Bayesian Classifier

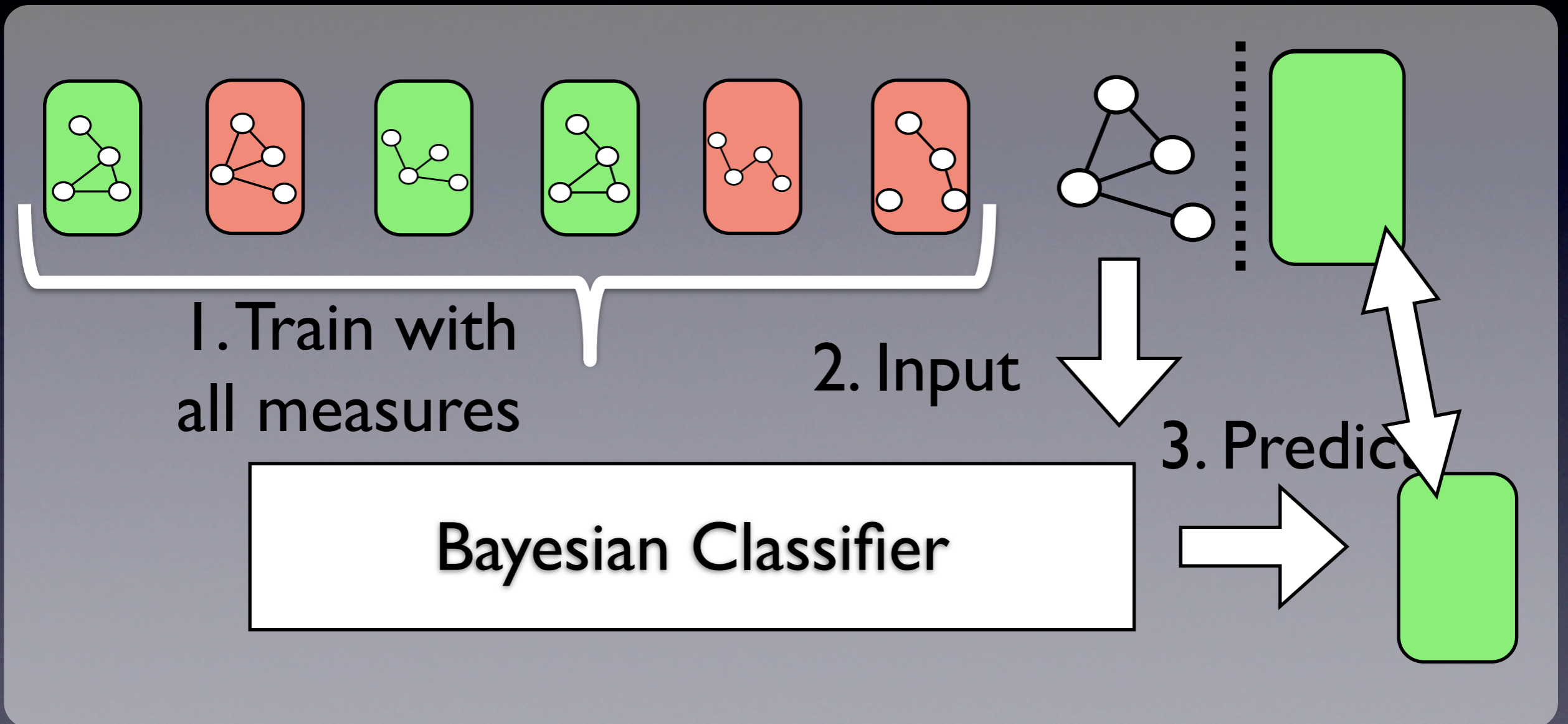
RQ2: multi measures prediction?



RQ2: multi measures prediction?



RQ2: multi measures prediction?



Prediction Results

Prediction Results

ERROR	Team Builds					Project Builds		
	B	C	F	P	W	nightly	weekly	beta
Recall	55%	75%	62%	66%	74%	89%	100%	92%
Precision	52%	50%	75%	76%	66%	73%	92%	92%

Prediction Results

ERROR	Team Builds					Project Builds		
	B	C	F	P	W	nightly	weekly	beta
Recall	55%	75%	62%	66%	74%	89%	100%	92%
Precision	52%	50%	75%	76%	66%	73%	92%	92%

Prediction Results

ERROR	Team Builds					Project Builds		
	B	C	F	P	W	nightly	weekly	beta
Recall	55%	75%	62%	66%	74%	89%	100%	92%
Precision	52%	50%	75%	76%	66%	73%	92%	92%

- Recall:
 - 62 of 100 failed **builds** are classified correctly

Prediction Results

ERROR	Team Builds					Project Builds		
	B	C	F	P	W	nightly	weekly	beta
Recall	55%	75%	62%	66%	74%	89%	100%	92%
Precision	52%	50%	75%	76%	66%	73%	92%	92%

- Recall:
 - 62 of 100 failed **builds** are classified correctly
- Precision:
 - 75 of 100 failure **predictions** are actual failures
 - 75% confidence that a prediction is correct

Summary

Summary

- Quantitative empirical evidence:
Communication structures make a difference!

Summary

- Quantitative empirical evidence:
Communication structures make a difference!
- Communication structures:
 - impact integration build results
 - are key for successful collaboration

Summary

- Quantitative empirical evidence:
Communication structures make a difference!
- Communication structures:
 - impact integration build results
 - are key for successful collaboration
- Single measures **do not** predict build results

Summary

- Quantitative empirical evidence:
Communication structures make a difference!
- Communication structures:
 - impact integration build results
 - are key for successful collaboration
- Single measures **do not** predict build results
- Combination measures **do** predict build results

Practical Implications

- Communication assessment for future builds
- Notification system for build failures
- Communication structure improvement

Future Work

- Factor analysis of combined measures
- Include technical information
- Prediction effectiveness case study

Questions?



Timo Wolf
timowolf@siemens.com



Daniela Damian
danielad@cs.uvic.ca



Adrian Schröter
schadr@uvic.ca



Thanh Nguyen
duythanh@uvic.ca

SIEMENS AG

Corporate Technology
Software & Engineering

SEGAL
Software Engineering Global interAction Labs

<http://segal.uvic.ca>

Backup Slides

Descriptive Statistics I

	Team Level Builds					Project Level Builds		
	B	C	F	P	W	nightly	weekly	beta
# Builds	60	48	55	59	55	15	15	16
# ERRORS	20	16	24	29	31	9	11	13
# OKs	40	32	31	30	24	6	4	3
<i># Contributors:</i>								
Min	3	9	6	5	13	43	37	55
Median	6	16.5	18	15	20	55	57	69.5
Mean	12.68	18.02	20.15	17.98	22.87	57.93	52.27	67.81
Max	58	31	64	61	52	75	75	79
<i># Directed Connections:</i>								
Min	0	1	2	0	11	81	56	144
Median	13	39.5	95	36	74	236	149	280
Mean	51.58	53.4	87.78	63	88.35	253.1	171.9	285.8
Max	361	139	355	401	300	434	496	446

Descriptive Statistics II

	B	Team Level Builds				Project Level Builds		
		C	F	P	W	nightly	weekly	beta
<i># Change Sets:</i>								
Min	1	15	8	32	83	80	62	82
Median	10	38	35	46	111	117	115	178.5
Mean	10.83	44.38	42.65	47.25	115.3	129	114.2	166.8
Max	33	101	91	75	156	199	173	196
<i># Work Items:</i>								
Min	0	2	1	1	10	11	5	31
Median	6.5	12	20	12	18	67	51	98
Mean	16.43	15.56	23.07	19.34	29.49	72.13	56.87	96.81
Max	131	50	100	107	119	132	202	170