



Coordination and Efficiency in Decentralized Collaboration

Daniel M. Romero
School of Information
University of Michigan

In collaboration with Dan Huttenlocher and Jon Kleinberg

Coordination in Decentralized Collaboration Environments

Explicit coordination: Requires interaction. Costly & time consuming.



Coordination in Decentralized Collaboration Environments

Explicit coordination: Requires interaction. Costly & time consuming.

Implicit coordination: Little interaction. **Requires a mutual mental model.** [Cannon-Bower 90]

Trade-off: Coordination benefits vs. cost



Coordination in Decentralized Collaboration Environments

Explicit coordination: Requires interaction. Costly & time consuming.

Implicit coordination: Little interaction. **Requires a mutual mental model.** [Cannon-Bower 90]



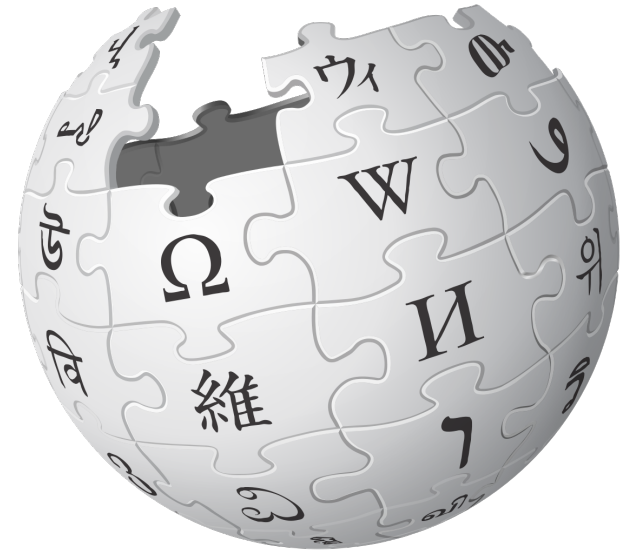
Trade-off: Coordination benefits vs. cost

In this paper: How coordination levels vary depending on:

- 1) quality of output
- 2) team composition

Coordination in Wikipedia

Data: All of Wikipedia edit history up to April 1, 2007. A total of 3.4 million articles edited by 500K users.



Coordination Tools:

- 1. Discussion Pages:** Discussion of any issues with the article.
- 2. Edit comments:** Comments that explain the nature of each edit.

Discussion and Comment Topics

Discussion Pages	
Justify edit	18
Suggest edit	33
Provide reference	20
Question	13
Copyright Issue	8
Dispute claim in article	12
Future direction	8
Other	6

Edit Comments	
Mentions section	52
Reverted edit	14
Minor edit	19
Added content	14
Removed content	7
Correction	2
Mentions other users	14
Other	11

Featured Wikipedia Articles

- One feature article chosen each day by Wikipedia community through a peer review process (2,563).
- They are reviewed for **accuracy, neutrality, completeness, and style.**



Featured Wikipedia Articles

- One feature article chosen each day by Wikipedia community through a peer review process (2,563).
- They are reviewed for **accuracy, neutrality, completeness, and style.**



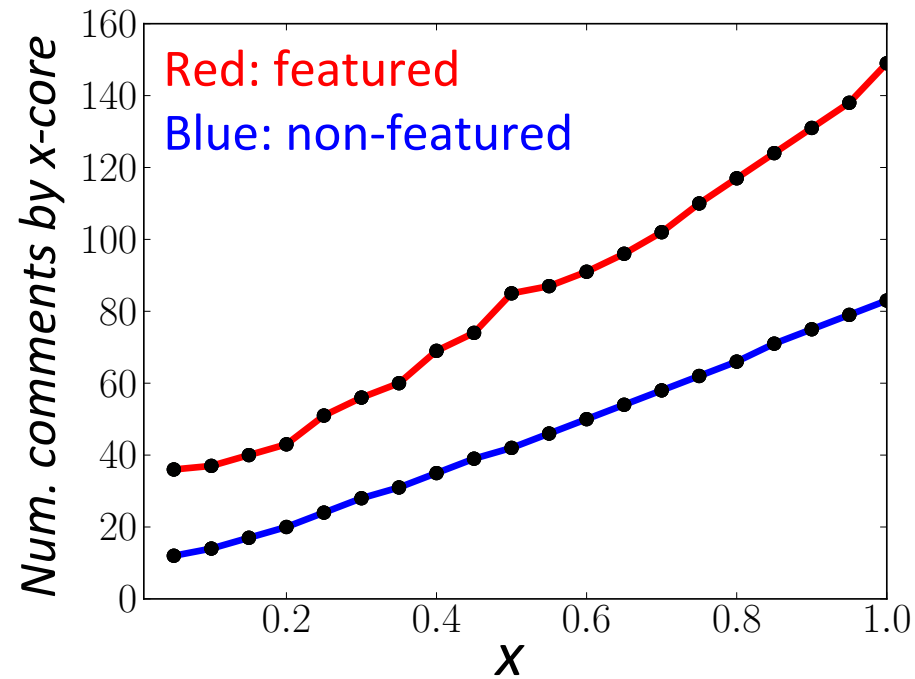
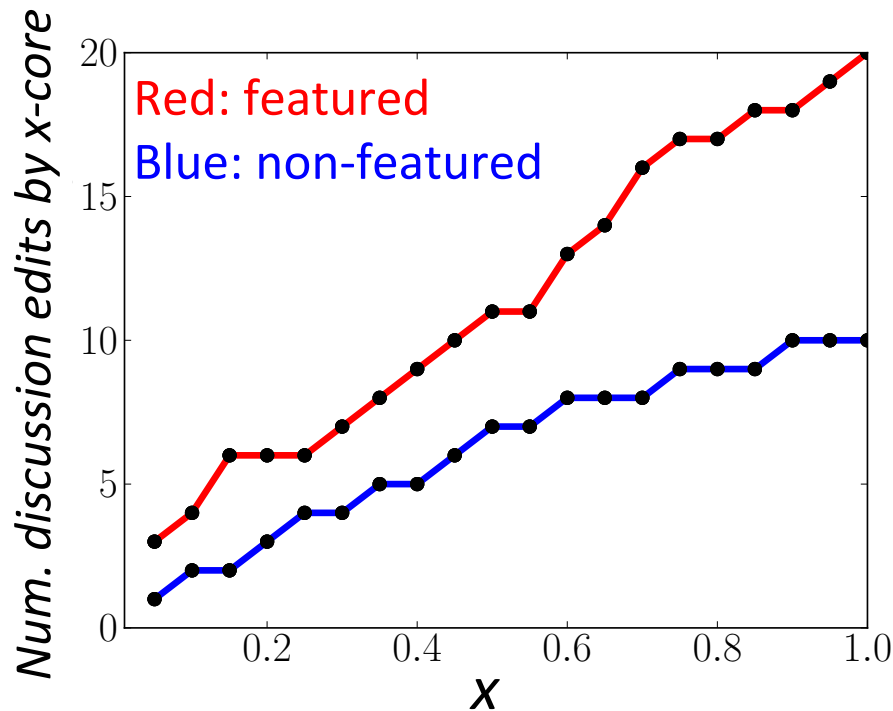
Do editors of featured articles coordinate more or less than editors of non-featured articles?

Coordination and Quality

x-core: Smallest set of editors responsible for x% of all edits.

Coordination and Quality

x-core: Smallest set of editors responsible for x% of all edits.

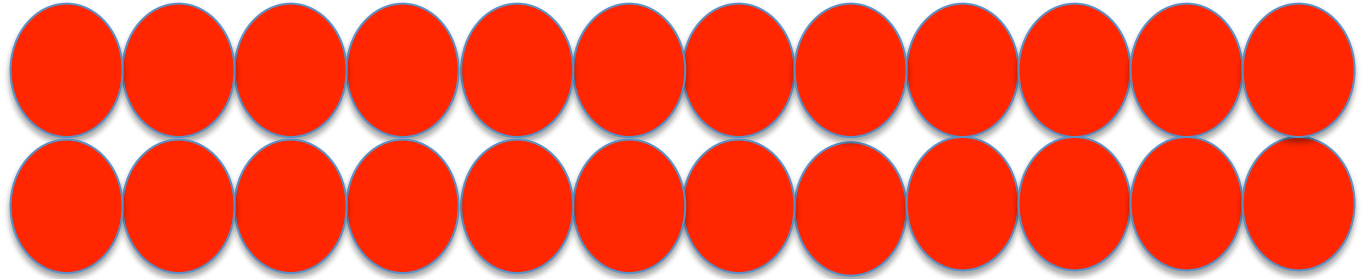


Higher levels of coordination associated with higher quality

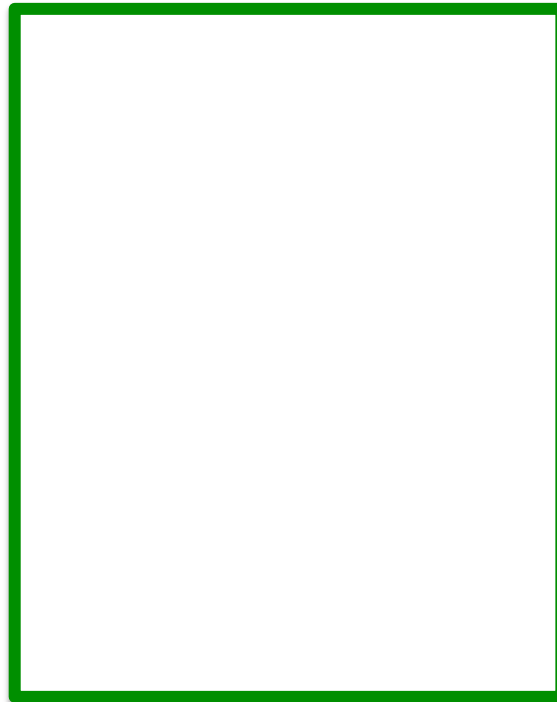
Coordination and Team Composition

Coordination and Team Composition

Num. Editors

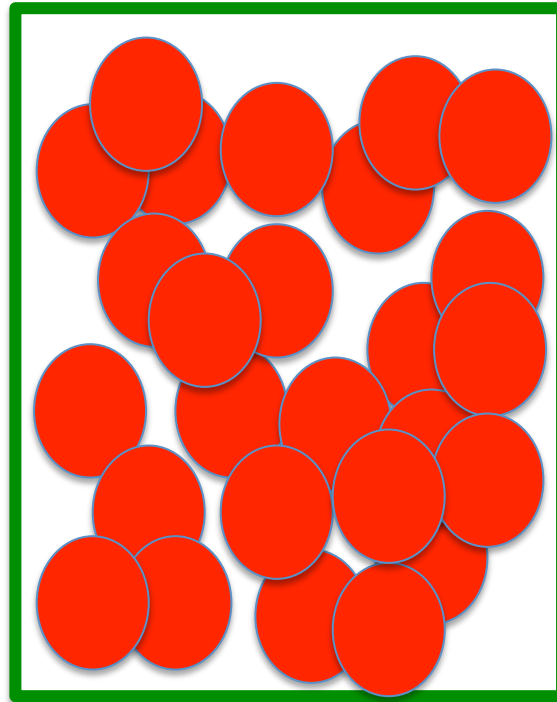


Amount of work



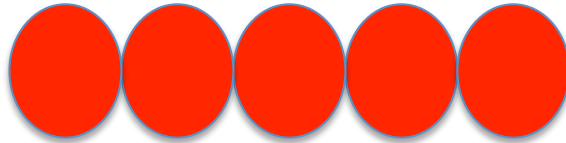
Coordination and Team Composition

Crowded Environment



Coordination and Team Composition

Num. Editors

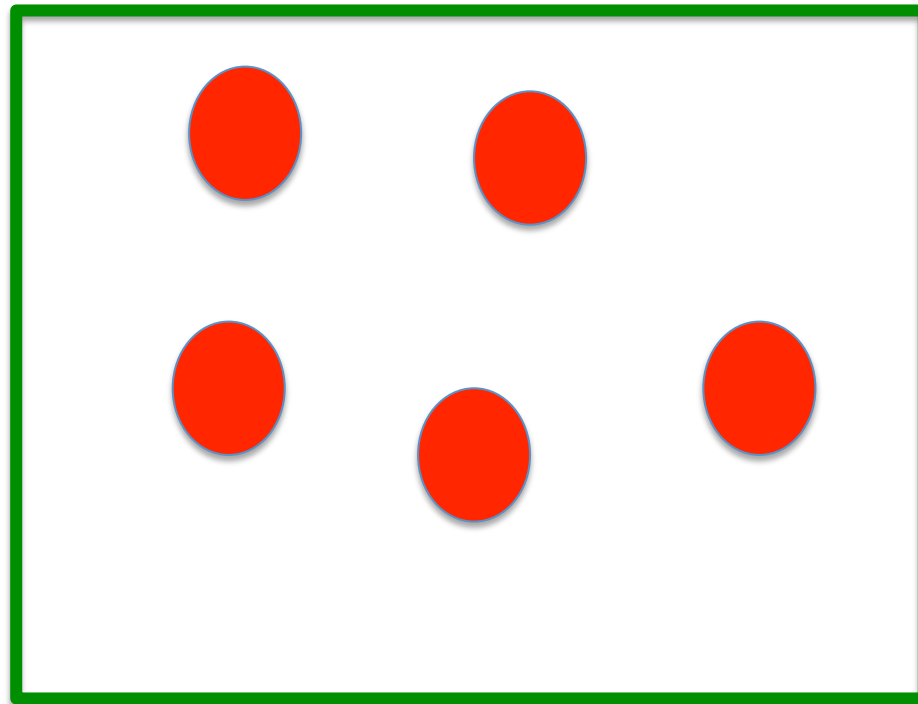


Amount of work



Coordination and Team Composition

Less Crowded Environment



Coordination and Team Composition

Consider the first 100 edits to each article

Num. editors

Eventual size of article in bytes (or word count)

Num. discussion edits by the 100th article edit

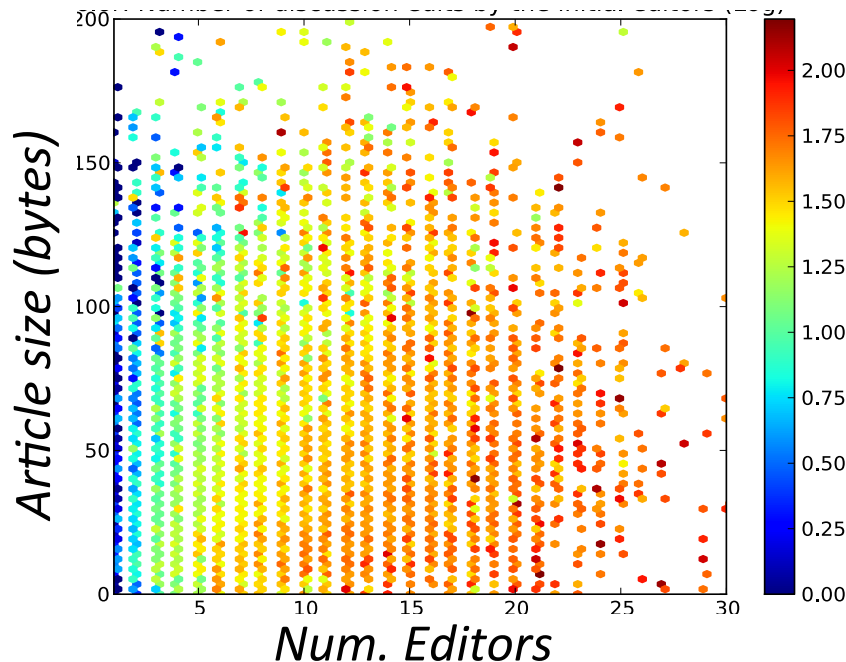
Coordination and Team Composition

Consider the first 100 edits to each article

Num. editors

Eventual size of article in bytes (or word count)

Num. discussion edits by the 100th article edit



Higher levels of coordination associated with crowdedness

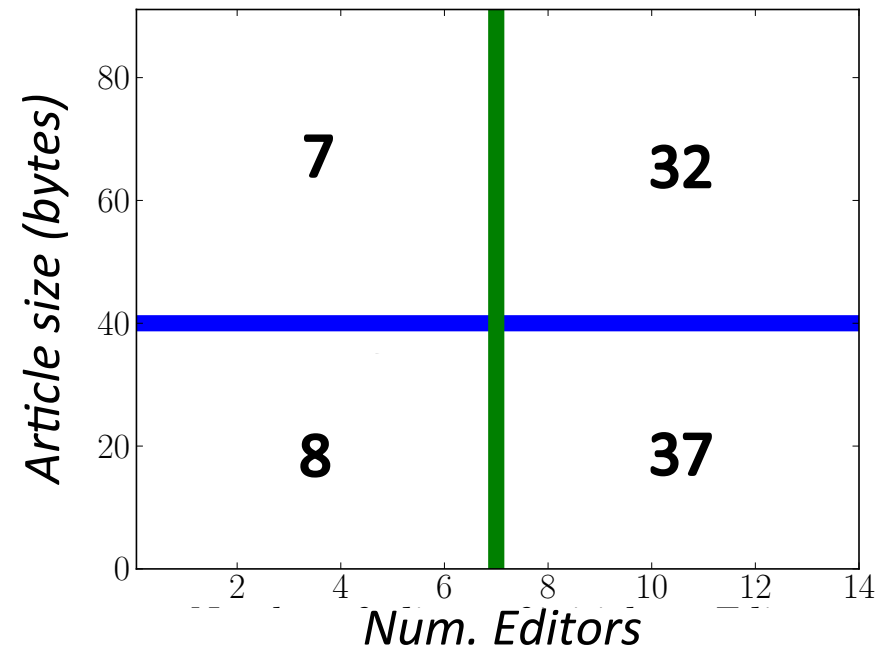
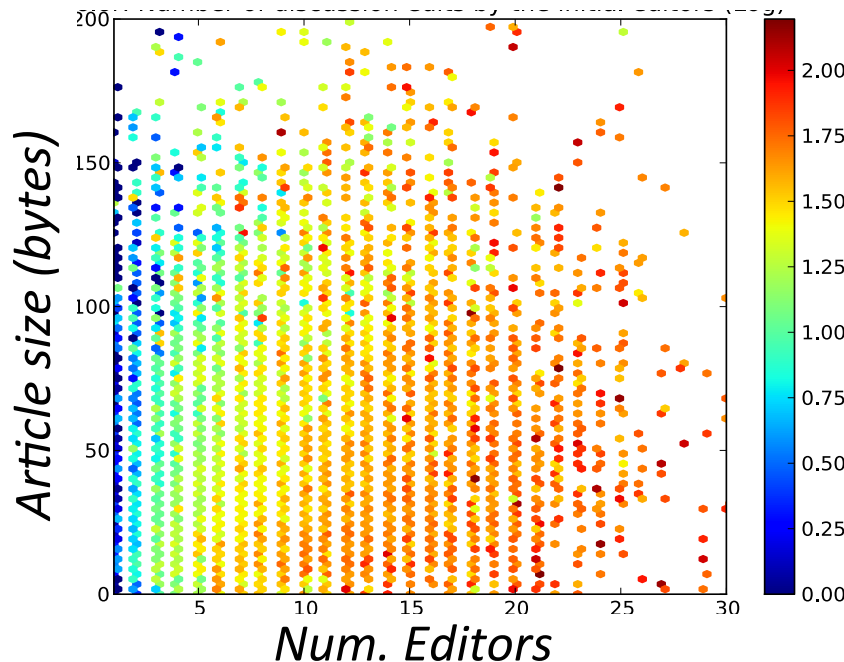
Coordination and Team Composition

Consider the first 100 edits to each article

Num. editors

Eventual size of article in bytes (or word count)

Num. discussion edits by the 100th article edit



Higher levels of coordination associated with crowdedness

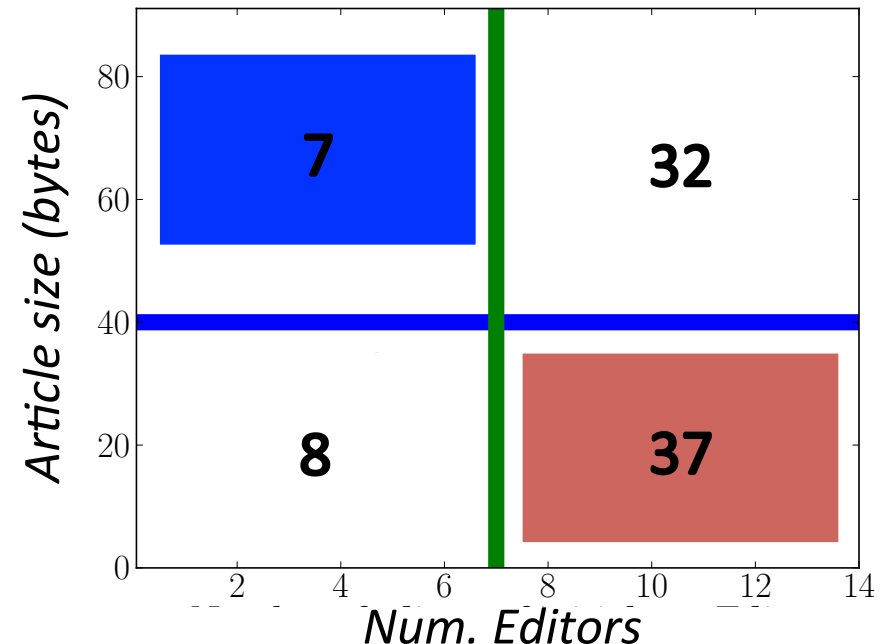
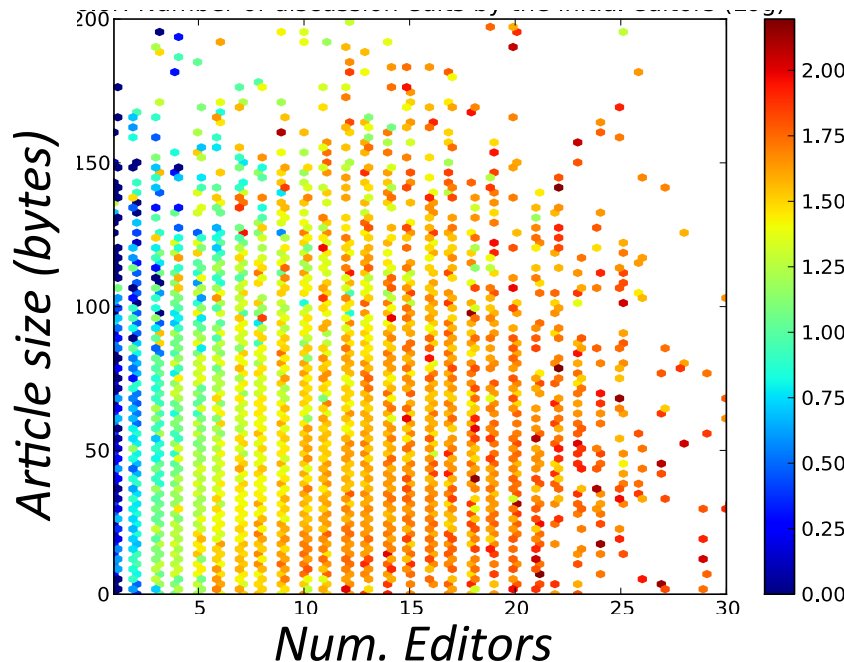
Coordination and Team Composition

Consider the first 100 edits to each article

Num. editors

Eventual size of article in bytes (or word count)

Num. discussion edits by the 100th article edit



Higher levels of coordination associated with crowdedness

Coordination Trade-Off Model

- Each articles has N “parts” and E editors.
- Each part is either “empty” or “full”
- The goal is to fill in as many parts as possible

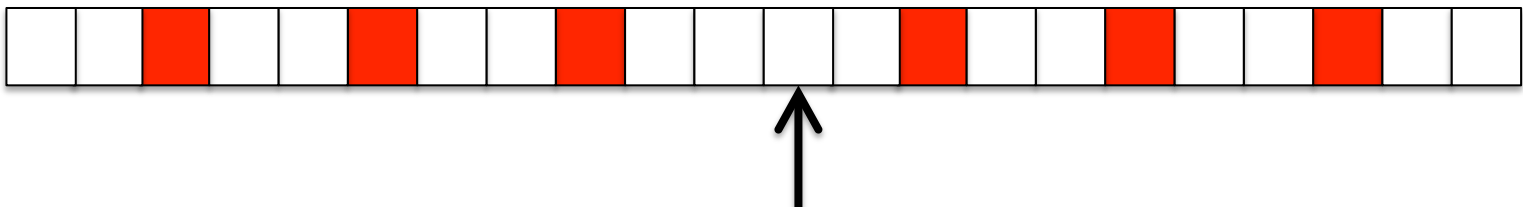


Coordination Trade-Off Model

- Each articles has N “parts” and E editors.
- Each part is either “empty” or “full”
- The goal is to fill in as many parts as possible

When editors arrive:

- If they choose an empty part, the part will become full (net gain to article)

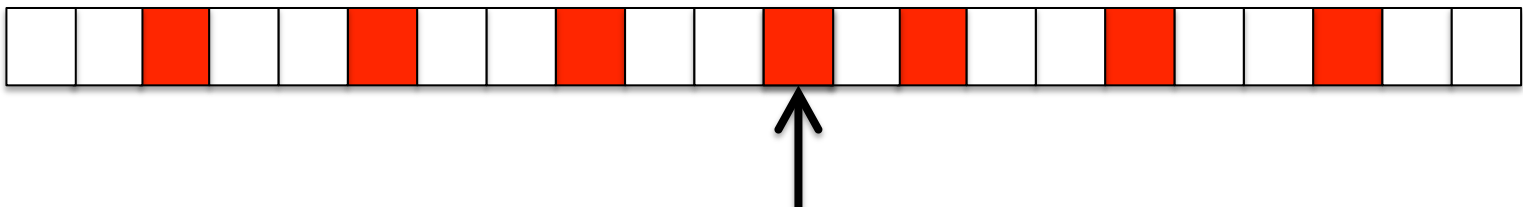


Coordination Trade-Off Model

- Each articles has N “parts” and E editors.
- Each part is either “empty” or “full”
- The goal is to fill in as many parts as possible

When editors arrive:

- If they choose an empty part, the part will become full (net gain to article)

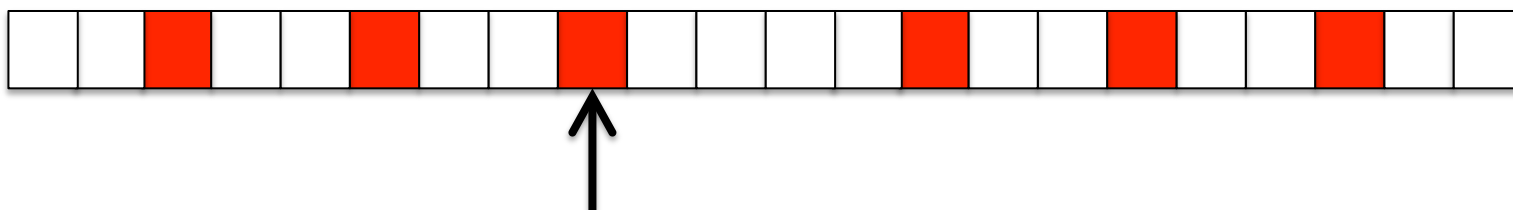


Coordination Trade-Off Model

- Each articles has N “parts” and E editors.
- Each part is either “empty” or “full”
- The goal is to fill in as many parts as possible

When editors arrive:

- If they choose an empty part, the part will become full (net gain to article)
- If they choose a full part, the part will become empty (net loss to the article) with probability α .

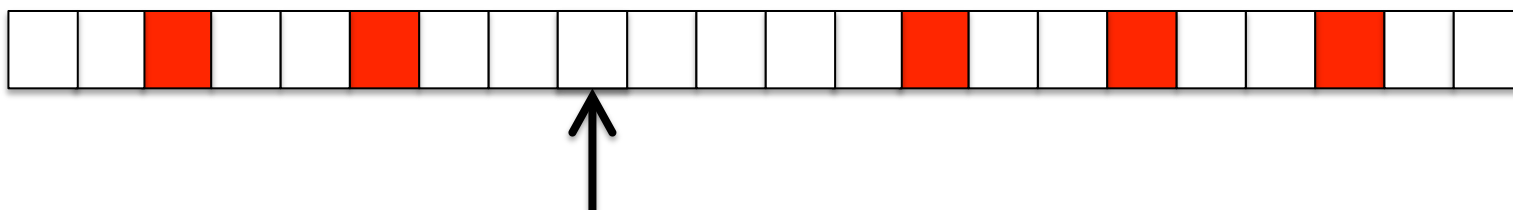


Coordination Trade-Off Model

- Each articles has N “parts” and E editors.
- Each part is either “empty” or “full”
- The goal is to fill in as many parts as possible

When editors arrive:

- If they choose an empty part, the part will become full (net gain to article)
- If they choose a full part, the part will become empty (net loss to the article) with probability α .



Coordination Trade-Off Model

Each editor has 2 options:

- Not coordinate: Choose 2 **random** parts to edit: contribute -2, to 2 parts to the article

Coordination Trade-Off Model

Each editor has 2 options:

- Not coordinate: Choose 2 **random** parts to edit: contribute -2, to 2 parts to the article
- Coordinate: Choose 1 **empty** part to edit: Contribute **exactly** 1 part to the article.

Coordination Trade-Off Model

Each editor has 2 options:

- Not coordinate: Choose 2 **random** parts to edit: contribute -2, to 2 parts to the article
- Coordinate: Choose 1 **empty** part to edit: Contribute **exactly** 1 part to the article.

Each editor coordinates with probability β (fixed)

Find β that maximizes the number of finished parts (in terms of N and E)

Optimal Coordination

- Fix N and β and let P_i be the expected number of parts filled after the first i editors.

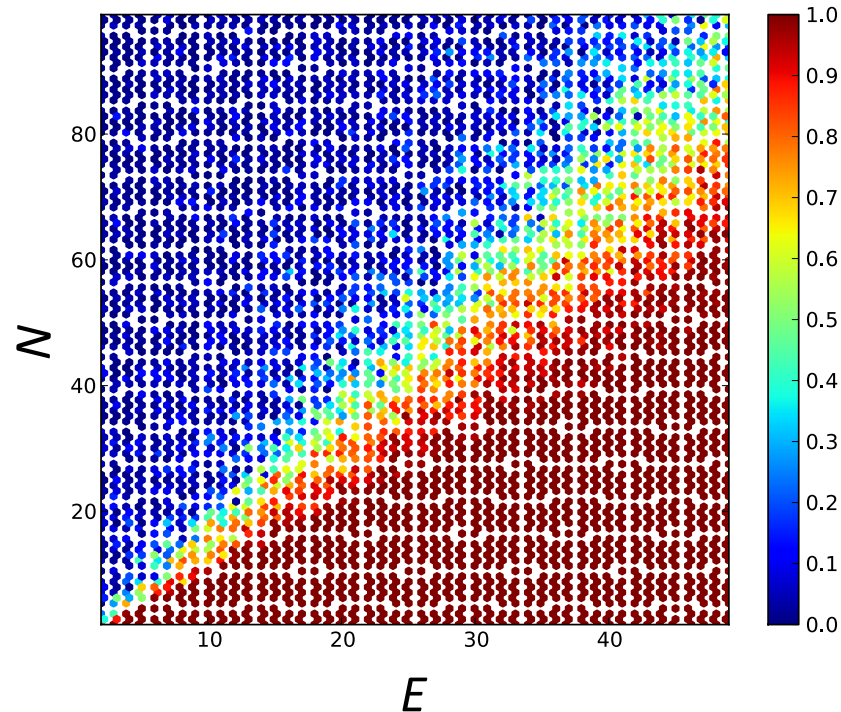
$$P_{i+1} = AP_i + P_0$$

$$A = \frac{4(1-\beta)}{N^2} - \frac{2(1-\beta)}{N} + 1$$

$$P_0 = -\frac{1-\beta}{N} - \beta + 2$$

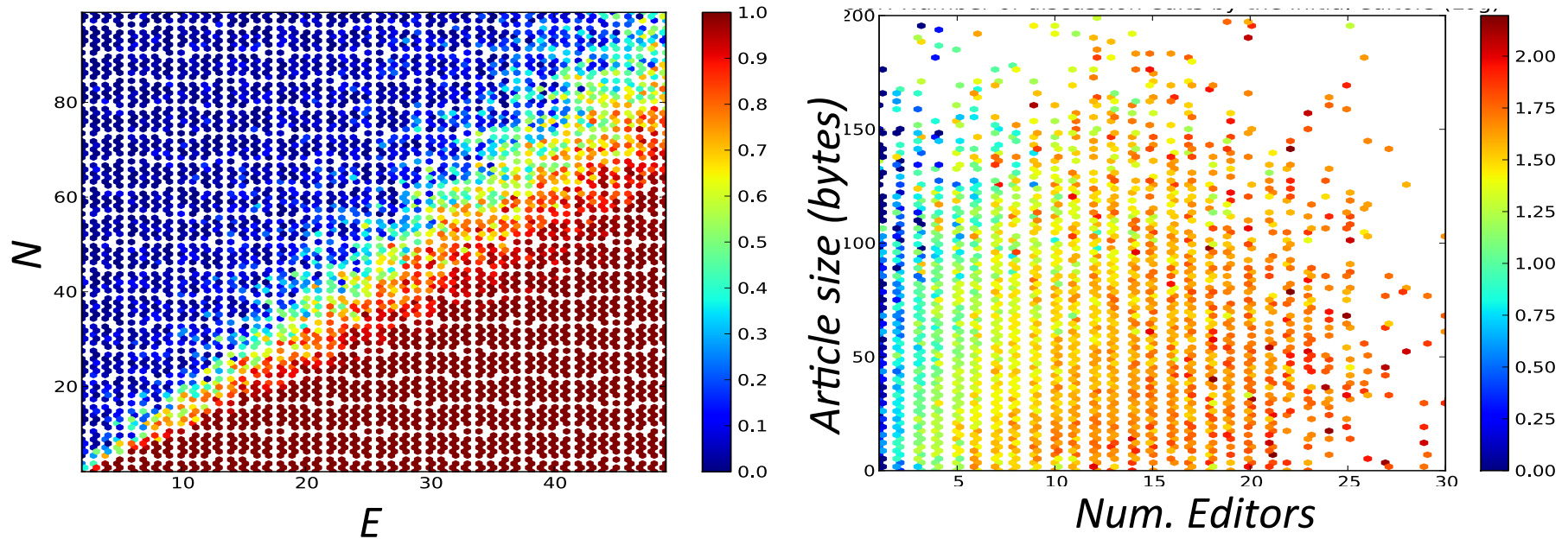
$$P_i = P_0 \frac{A^i - 1}{A - 1}$$

Optimal Coordination



- When $E > N$, $\beta = 1$ (fill all parts by coordinating)
- When E is small enough, $\beta = 0$ (colliding is unlikely)
- In between, the best β lies away from both 0 and 1.

Comparison with Wikipedia



Higher levels of coordination in crowded articles

Discussion

- **Coordination trade-off:** Projects with high performance and in crowded environments exhibit higher levels of coordination.

Discussion

- **Coordination trade-off:** Projects with high performance and in crowded environments exhibit higher levels of coordination.
- **Implication for design:** Coordination mechanisms should be emphasized more strongly on crowded projects.

Discussion

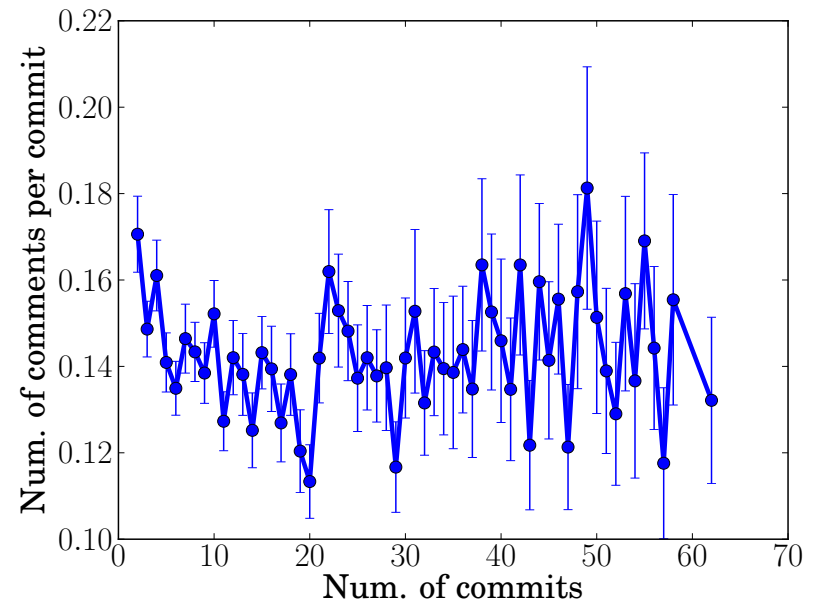
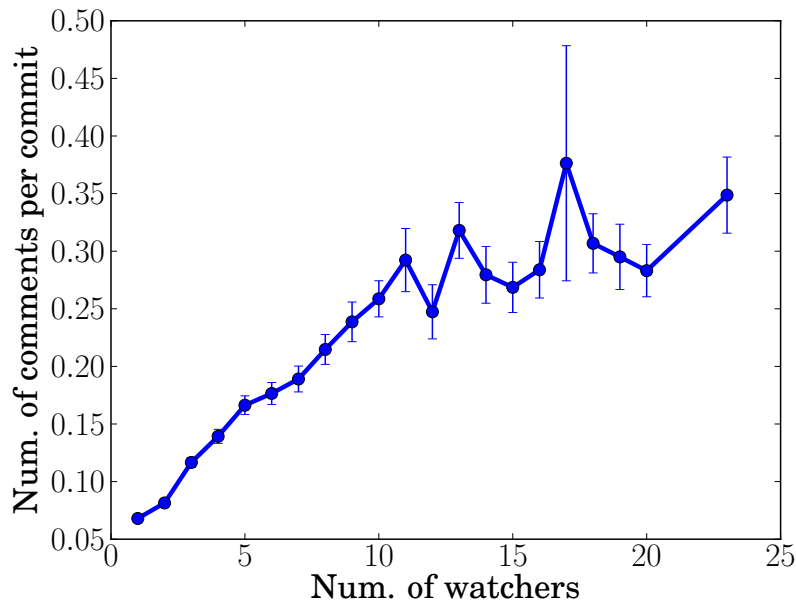
- **Coordination trade-off:** Projects with high performance and in crowded environments exhibit higher levels of coordination.
- **Implication for design:** Coordination mechanisms should be emphasized more strongly on crowded projects.
- **Generalizability:**
 - Findings hold in two very different domains: Wikipedia and GitHub (See paper).
 - Proposed framework (model and measures) can be directly adapted to new settings, where a group produces a primary work product and a separate channel for coordination.

Discussion

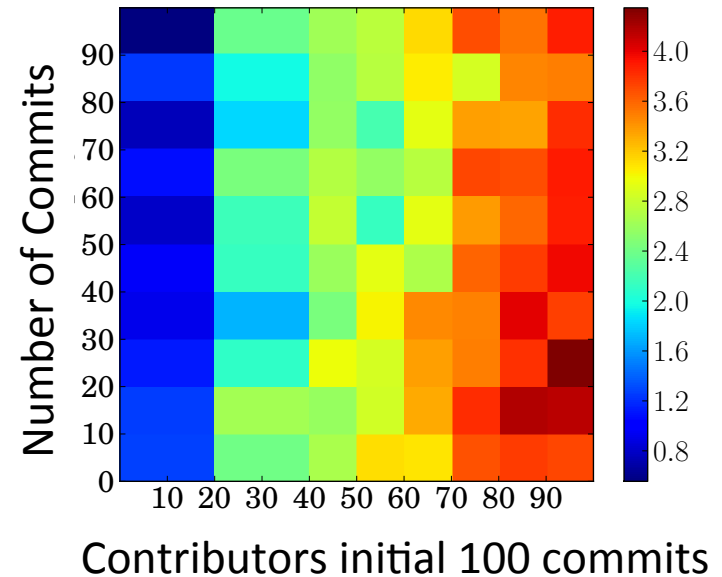
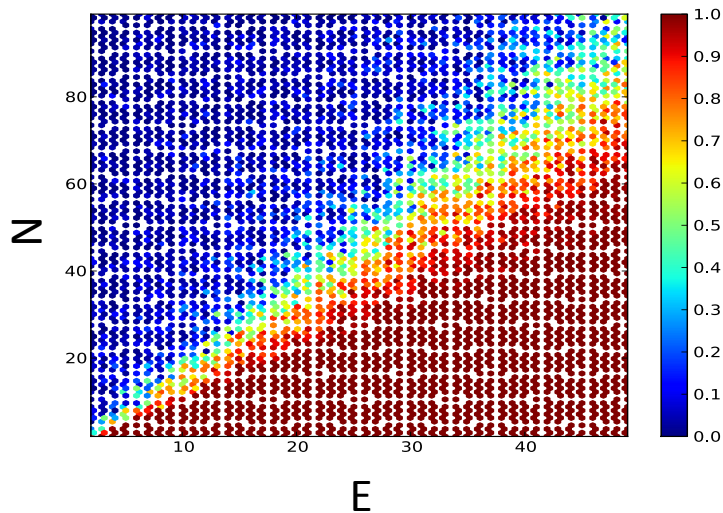
- **Coordination trade-off:** Projects with high performance and in crowded environments exhibit higher levels of coordination.
- **Implication for design:** Coordination mechanisms should be emphasized more strongly on crowded projects.
- **Generalizability:**
 - Findings hold in two very different domains: Wikipedia and GitHub (See paper).
 - Proposed framework (model and measures) can be directly adapted to new settings, where a group produces a primary work product and a separate channel for coordination.

Coordination in GitHub

- Coordination: Number of comments
- Size of project: Number of commits
- Status: Number of watchers



Coordination in GitHub



High coordination increases when article becomes crowded -- small size and many participants