How Transitive Are Real-World Group Interactions? – Measurement and Reproduction

KDD2023 LONG BEACH, CA AUGUST 6 - 10

Sunwoo Kim, KAIST, kswoo97@kaist.ac.kr
Fanchen Bu, KAIST, boqvezen97@kaist.ac.kr
Minyoung Choe, KAIST, minyoung.choe@kaist.ac.kr
Jaemin Yoo, CMU, jaeminyoo@cmu.edu
Kijung Shin, KAIST, kijungs@kaist.ac.kr

https://github.com/kswoo97/hypertrans
https://doi.org/10.1145/3580305.3599382

Summary

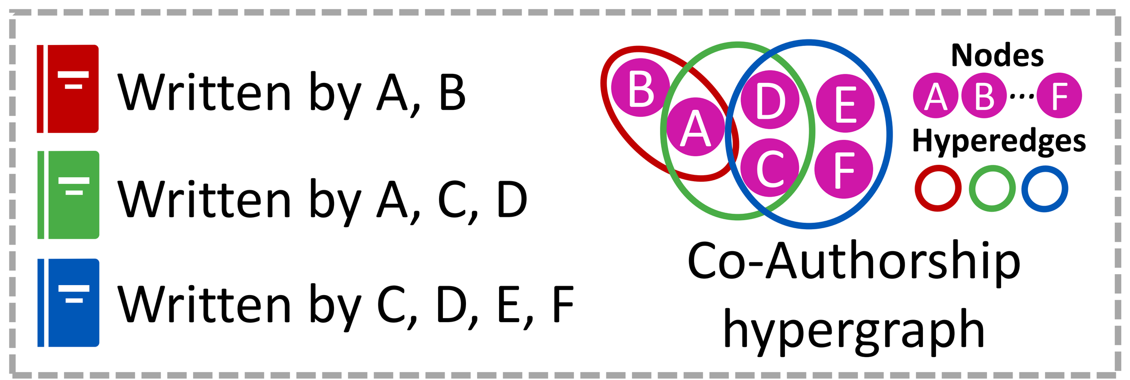
- **Goal:** To analyze **transitivity of real-world group interactions** and to develop a realistic and scalable synthetic hypergraph generator.
- **Previous Work:** Focus on transitivity of pairwise interactions but not on group interactions.
- **Contribution 1) Axioms:** Criteria that any well-defined transitivity measures should meet.
- **Contribution 2) Principled Measure:** A measure (**HyperTrans**) that meets all the criteria.
- **Contribution 3) Observations:** Transitivity patterns of real-world hypergraphs.
- **Contribution 4) Generator:** A scalable hypergraph generator (**THera**) that can reproduce the transitivity patterns of real-world hypergraphs.
  - **Reproduction**: More accurately reproduce transitivity patterns than competitors.
  - **Scalability:** Can generate hypergraphs of $10^{7.5}$ hyperedges within a minute.

** **Oral Presentation Info**: **Tuesday, August 8, 10:00 AM - 12:00 AM, Room 201B.**

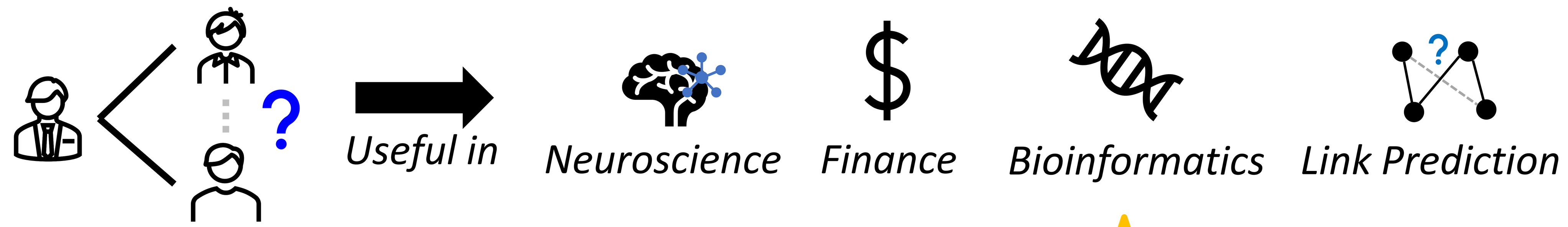
Motivation

- **Hypergraph:** Group interactions are prevalent around us, and a hypergraph is a data structure that is widely used to represent such group interactions.

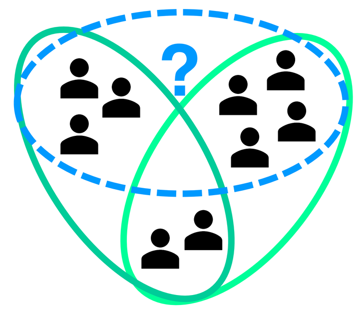
Examples of real-world group interactions



- **Graph Transitivity:** Probability of two neighbors of a node being also adjacent [1].

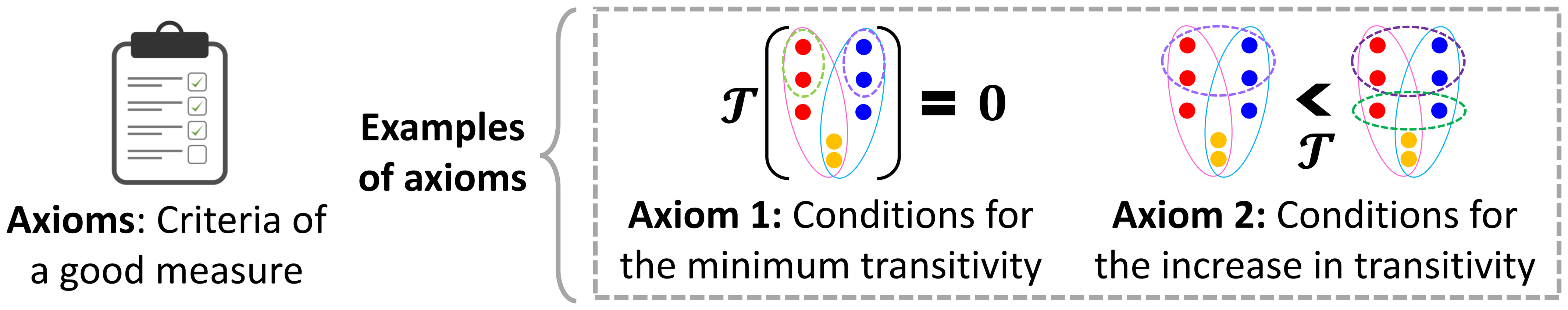


- **Research Interest: Transitivity of real-world group interactions.**
  - No existing studies focus on the transitivity of group interactions.
  - Q1) How should we measure transitivity of group interactions?
  - Q2) How transitive are real-world group interactions?
  - Q3) How can we reproduce realistic transitivity patterns?

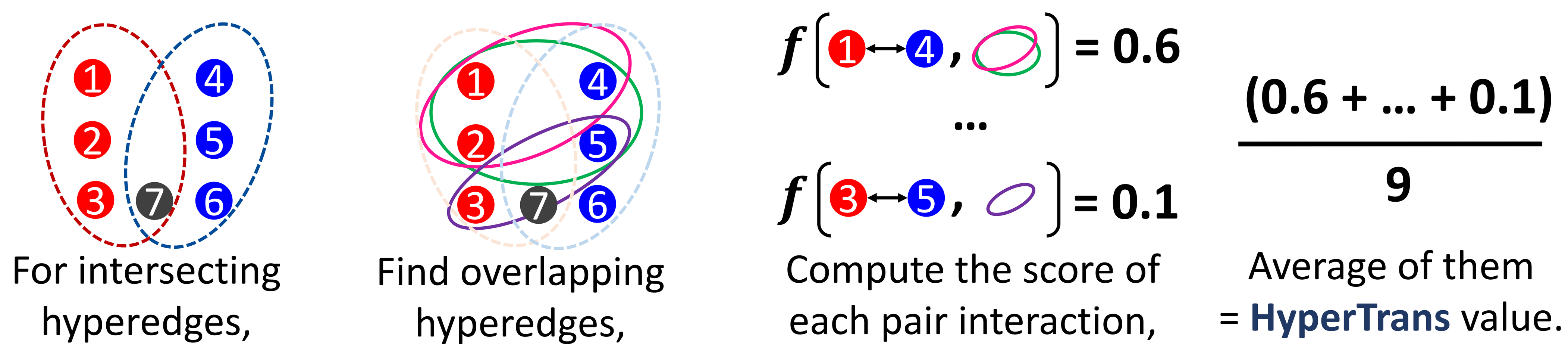


Axiom & Measure

- **Q1-1)** What are the **criteria that any well-defined transitivity measure should meet?**
- **A1-1) Axioms:** We propose 8 axioms that formalize characteristics that a well-defined transitivity measure ($\mathcal{T}$) should satisfy.

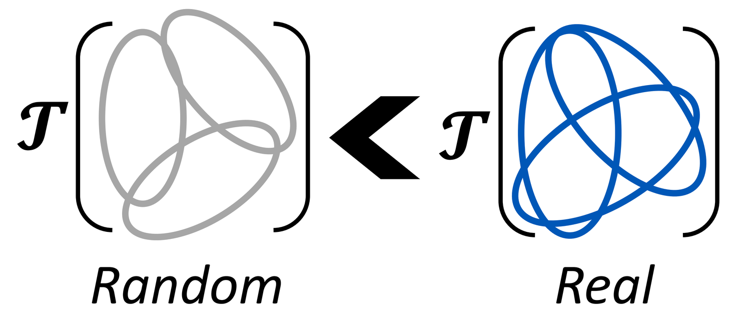


- **Q1-2)** How can **we systematically quantify** the transitivity of group interactions?
- **A1-2) HyperTrans:** Our proposed principled transitivity measure that satisfies all the proposed axioms (*No baseline measures satisfy all the proposed axioms!*).



Observation

- **Q3)** What are the **transitivity-related patterns of real-world group interactions**?
- **A3) Observations:** We analyze transitivity of 12 real-world hypergraphs.
  - Email & Drug & Student Communication & Co-Authorship & QnA
  - **Null Model**: To verify that the patterns are not random, we generate random counterparts of real-world hypergraphs by using HyperCL [2].
  - **Observation 1:** Real-world Hypergraphs are transitive.

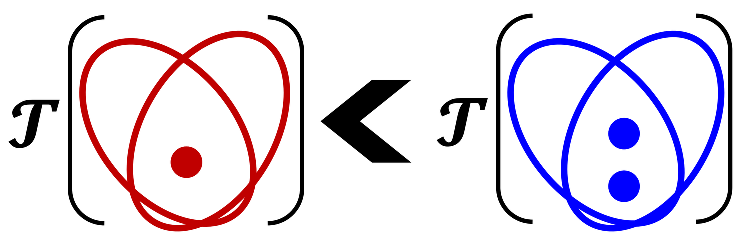


Measured hypergraph transitivity values (x100).

Data	E1	E2	D1	D2	S1	S2	C1	C2	C3	Q1	Q2	Q3
Real	19.5	12.5	5.2	1.9	34.5	33.6	0.7	0.5	0.2	0.5	0.5	0.3
Null	7.8	5.3	0.8	0.5	11.9	22.3	0.0*	0.0*	0.0*	1.4	1.7	4.0

0.0*: Value $\leq 10^{-4}$

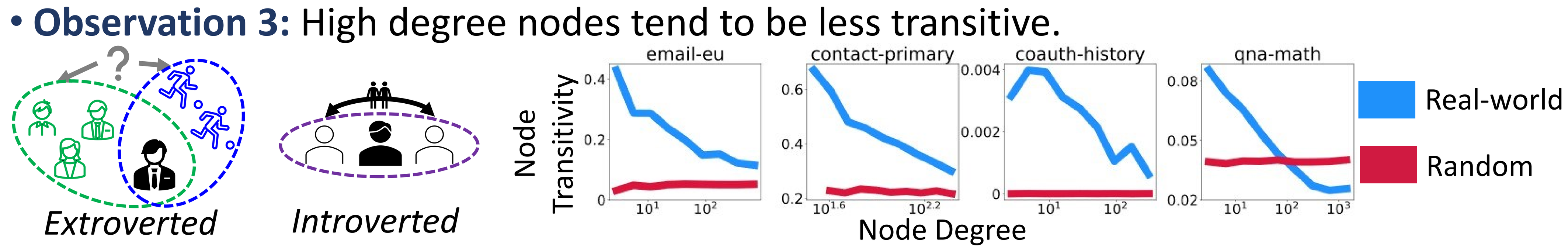
- **Observation 2:** Interactions with larger intersection tend to be more transitive.



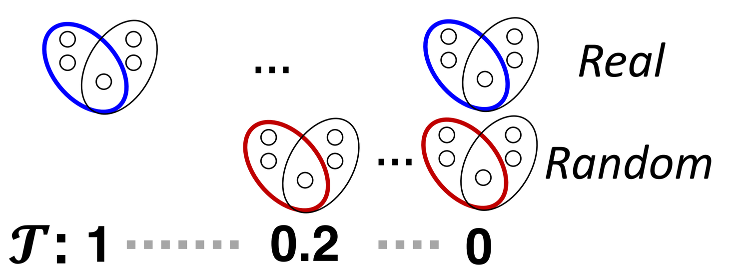
Correlation between intersection and transitivity (x100).

Data	E1	E2	D1	D2	S1	S2	C1	C2	C3	Q1	Q2	Q3
Real	9	12	32	14	13	13	12	14	12	4	4	4
Null	-9	-14	-1	-1	0.0*	0.0*	0.0*	0.0*	5	0.0*	0.0*	1

- **Observation 3:** High degree nodes tend to be less transitive.



- **Observation 4:** Real-world hyperedges have a wide transitivity range.

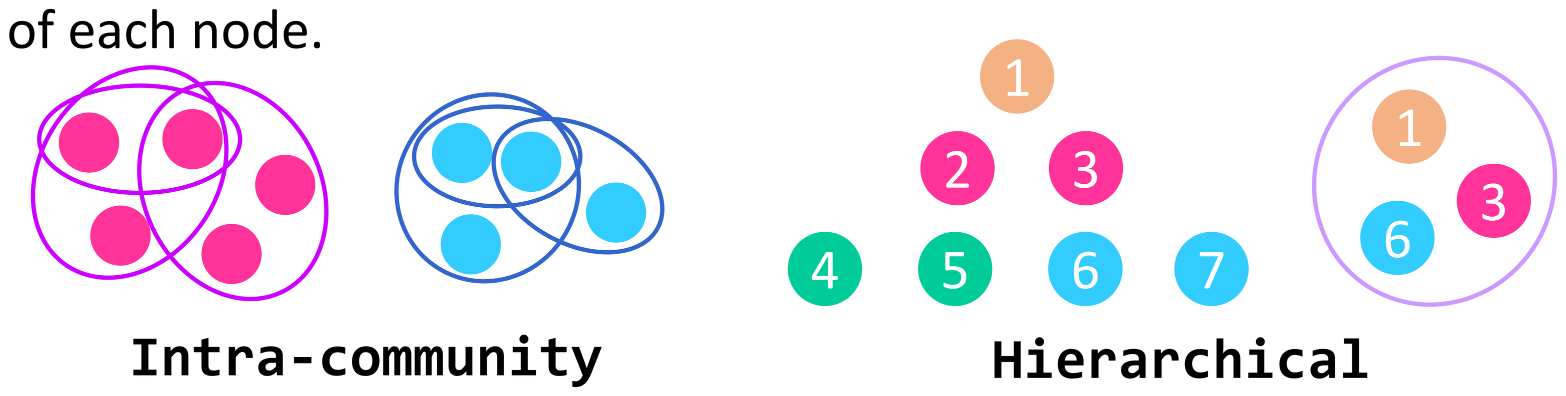


Range of hyperedge transitivity (x100).

Data	E1	E2	D1	D2	S1	S2	C1	C2	C3	Q1	Q2	Q3
Real	72.5	80.9	60	100	79.4	69.3	100	100	100	66.7	66.7	66.7
Null	27.9	24.8	7.5	3.2	31.6	39.5	10.5	6.9	33.3	50	33.3	100

Generator

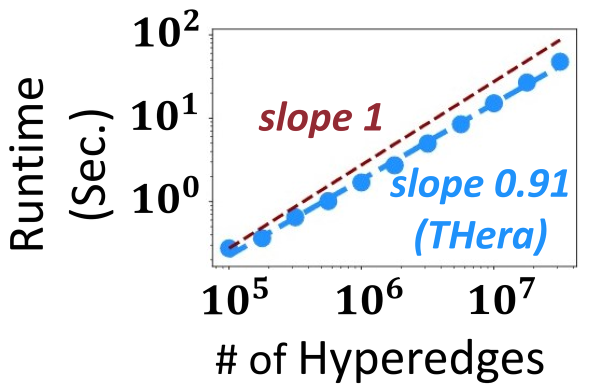
- **Q4)** How can we **reproduce transitivity patterns of real-world hypergraphs?**
- **A4) THera:** We propose a scalable hypergraph generator that can reproduce transitivity patterns of real-world hypergraphs.
  - **Description: THera** mainly utilizes (1) **community** and (2) **hierarchical** structure of nodes. It stochastically creates two types of hyperedges:
    - `Intra-community`: **THera** assigns a community to each node and create hyperedges within each community.
    - `Hierarchical`: **THera** creates tree of nodes and sample nodes proportional to the level of each node.



  - **Evaluation: THera outperforms all competitors** in terms of (1) pattern reproducibility, (2) speed, and (3) memory consumption.

Generators	THera	H-PA [3]	H-FF [4]	H-LAP [2]	H-LAP+ [2]
Pattern reproducibility	**1.1**	4.2	3.9	3.0	2.8
Runtime	**1.7**	5.0	2.7	1.8	3.7
Memory consumption	**2.0**	4.2	**2.0**	3.1	3.8

Avg. rank of generators (H-: Hyper).



Reference

[1]: Watts and Strogatz. Collective dynamics of 'small-world' network. Nature (393) 1998.
[2]: Lee et al. How do hyperedges overlap in real-world hypergraphs? WWW 2021.
[3]: Do et al. Structural patterns and generative models of real-world hypergraphs. KDD 2020.
[4]: Ko et al. Growth patterns and models of real-world hypergraphs. KAIS (64) 2022.