What makes a successful rebuttal in computer science conferences? : A perspective on social interaction

Junjie Huang[1,2] Win-bin Huang[3] Yi Bu[3] Qi Cao[1,2] Huawei Shen[1,2] Xueqi Cheng[1,2]

[1]Institute of Computing Technology, Chinese Academy of Sciences [2]University of Chinese Academy of Sciences
[3]Peking University

Abstract

With an exponential increase in submissions to top-tier Computer Science (CS) conferences, more and more conferences have introduced a rebuttal stage to the conference peer review process. The rebuttal stage can be modeled as social interactions between authors and reviewers. A successful rebuttal often results in an increased review score after the rebuttal stage. In this paper, we conduct an empirical study to determine the factors contributing to a successful rebuttal using over 3,000 papers and 13,000 reviews from ICLR2022, one of the most prestigious computer science conferences. **First**, we observe a significant difference in review scores before and after the rebuttal stage, which is crucial for paper acceptance. **Furthermore**, we investigate factors from the reviewer's perspective using signed social network analysis. A notable finding is the increase in balanced network structure after the rebuttal stage. **Subsequently**, we evaluate several quantifiable author rebuttal strategies and their effects on review scores. These strategies can help authors in improving their review scores. **Finally**, we used machine learning models to predict rebuttal success and validated the impact of potential factors analyzed in this paper. Our experiments demonstrate that the utilization of all features proposed in this study can aid in predicting the success of the rebuttal. **In summary, this work presents a study on the impact factors of successful rebuttals from both reviewers' and authors' perspectives and lays the foundation for analyzing rebuttals with social network analysis.**

Introduction

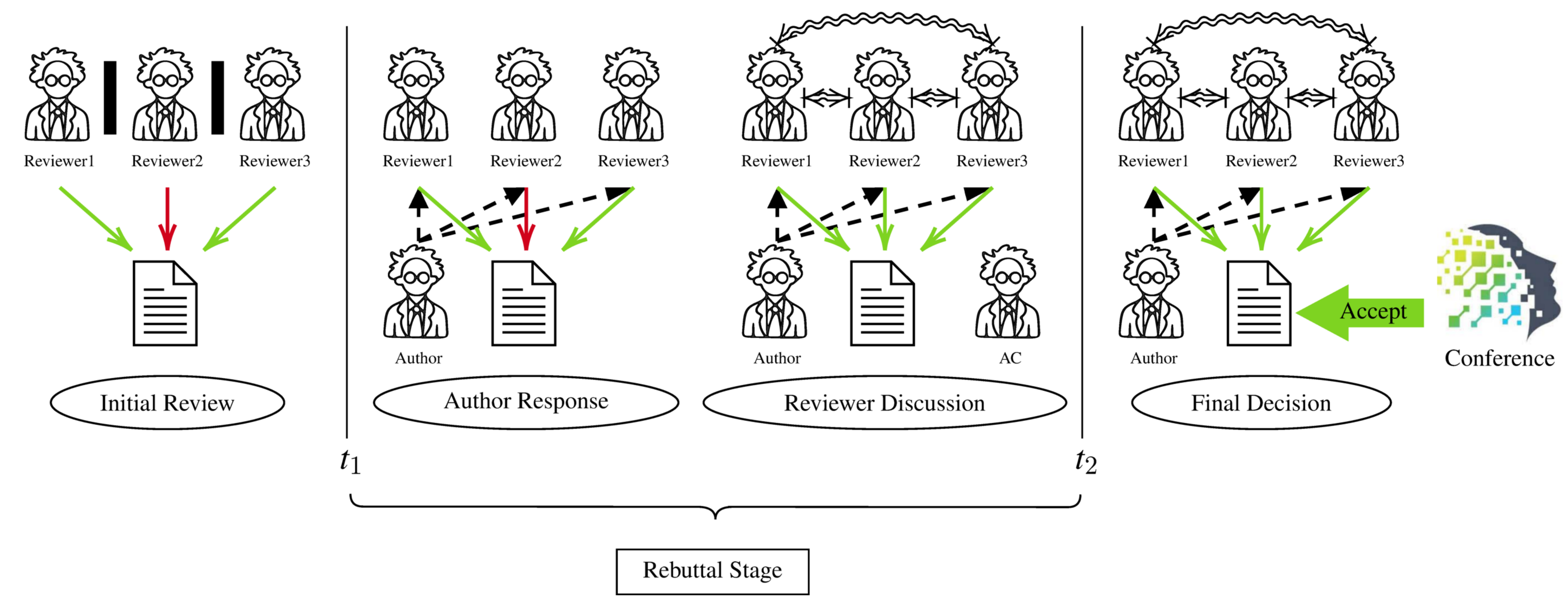

Figure. 1. Illustration of peer review process in ICLR2022. ICLR2022 mainly includes four stages: initial review, author response, reviewer discussion and final decision. In this paper, we mainly focus on the rebuttal stage between t_1 and t_2.

In this paper, we select ICLR2022 as our research subject to examine the rebuttal stage in the CS conference peer review process. As shown in Figure 1, the core process encompasses the **initial review**, **author rebuttal**, **reviewer discussion** with reviewers and ACs, and the **final decision**. We examine the changes in review scores between t_1 and t_2 and investigate the factors that might influence these changes.

Research question

Based on the review process in ICLR2022, we focus on exploring the following research questions in this paper:

- **RQ1**: Does rebuttal stage matter? Is there a difference between the initial and final review results in ICLR2022?
- **RQ2**: Does "peer effect" influence the score changes for reviewers? How to model it with signed social network analysis?
- **RQ3**: Are there effective strategies that authors can employ for a successful rebuttal?
- **RQ4**: Can we build machine learning models to predict whether reviewers will revise their score after rebuttal?

Results and Discussion

Rebuttal Results

Table 1. Statistics of different types of reviews and papers.

	#Review	#Paper	%Accept	Δ
KEEP	10,374	1,727	13.26%	$4.52 \rightarrow 4.52$
INC	2,310	1,444	58.38%	$5.37 \rightarrow 6.10$
DEC	179	167	13.77%	$5.17 \rightarrow 4.74$
Total	12,863	3,338	32.80%	$4.92 \rightarrow 5.22$

- The average scores of 1,444 papers show an increase, resulting in an acceptance rate of approximately 58.38%, significantly higher than the acceptance rates of the 167 papers with decreased scores (13.77%) and the 1,727 papers with unchanged scores (13.26%).
- While 43.25% (1,444/3,338) of papers experienced an increase in scores, only 17.95% (2,310/12,863) of reviews displayed a similar increase.

Results and discussion

Signed Social Network Analysis

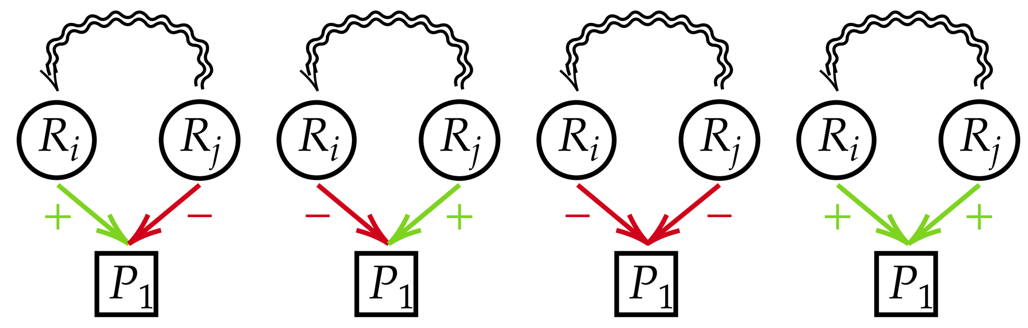

Figure. 2. The illustration of signed motifs in peer reviews. we define the first two motifs as unbalanced, and the last two motifs as balanced.

Table 2. The signed network analysis on three top computer science conference datasets. (* means that the p-value < 1e-3 for paired t-test)

	ICLR2022		ACL2018		TCSC	
	Before	After	Before	After	Before	After
# Links	12,863	13,021	3,875	4,054	1,170	1,170
% Positive Links	35.0	44.9(↑)	43.0	42.3(↓)	40.3	39.7(↓)
# Balanced motifs	11,720	13,329(↑)	2,324	2,679(↑)	1,002	1,134(↑)
# Unbalaned motifs	7,208	6,087(↓)	1,098	1,044(↓)	690	558(↓)
% Averaged Positive Ratio	35.2	45.1*(↑)	43.0	41.9(↓)	40.2	39.7(↓)
% Averaged Balanced Ratio	61.7	68.7*(↑)	61.8	69.7*(↑)	58.5	66.1*(↑)

- The ratio of positive links is below 50% across all three datasets.
- In all three datasets, the number of balanced motifs and the proportion of balanced motifs per paper increase after rebuttal and the unbalanced ones decrease (e.g., 11,720 → 13,329(↑) and 7,208 → 6,087(↓)).
- Another interesting observation is that, unlike ICLR2022, the negative link ratio of ACL2018 and TCSC increases after rebuttal.

Strategy Analysis

Table 3. Results of rebuttal strategy analysis (Mean±SD).

Strategy	Metrics	G_0	G_1	G_2	p-value
Work hard	Reply number		33.63%±47.25%	12.42%±32.98%	$< 1e^{-3}$
Work hard	Reply word count		36.37%±48.11%	9.29%±29.04%	$< 1e^{-3}$
Never miss	Text similarity (DL)		23.47%±42.39%	21.23%±40.90%	$< 1e^{-3}$
Never miss	Text similarity (TF-IDF)	2.82%±16.55%	25.84%±43.78%	16.59%±37.21%	$< 1e^{-3}$
Be polite	Politeness		33.52%±47.21%	13.54%±34.21%	$< 1e^{-3}$
Add references	Reference		28.49%±45.14%	19.45%±39.58%	$< 1e^{-3}$
Make consensus	Mention other reviewer		28.92%±45.35%	21.04%±40.76%	$< 1e^{-3}$

- The success rate of the group that adopted the rebuttal strategy (i.e. G_1) was significantly higher than that of the group that did not adopt the strategy.

Rebuttal Success Prediction

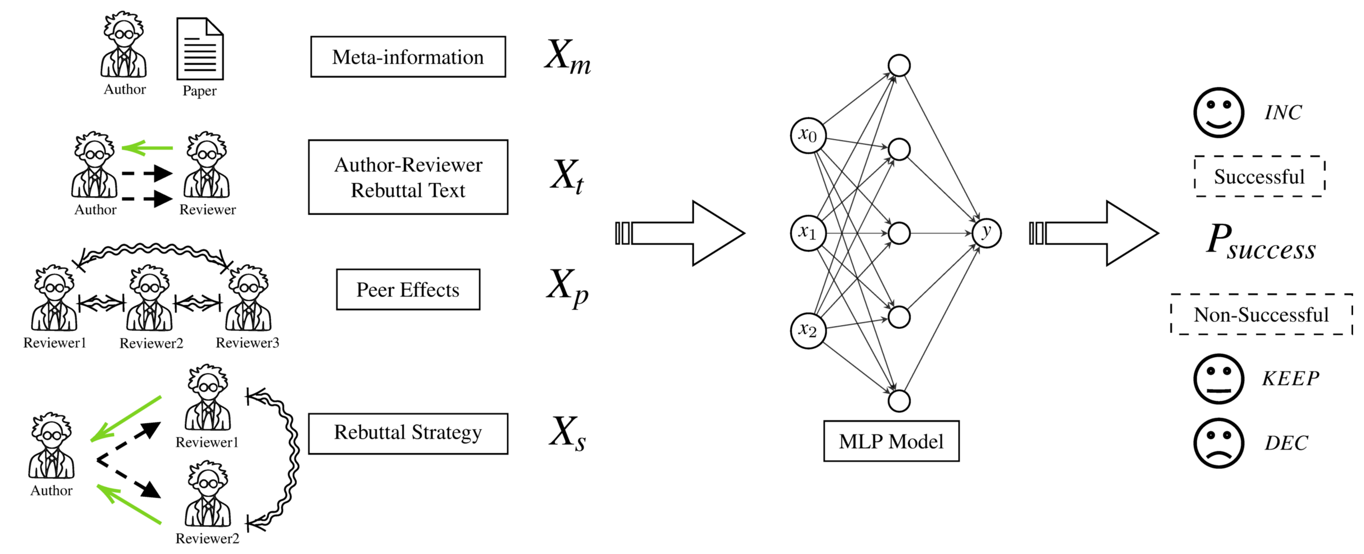

Figure. 3. The multi-factor prediction model for rebuttal success prediction.

Table 4. Results of rebuttal success prediction.

Model	Major Baseline	Random Baseline	MLP(X_m)	MLP(X_t)	MLP(X_s)	MLP(X_p)	MLP(X_p, X_m, X_s, X_t)
AUC (↑)	0.5000	0.4898	0.6477	0.6285	0.7143	0.7704	0.7739
Macro-F1 (↑)	0.4379	0.4484	0.4379	0.4886	0.4605	0.6175	0.6401

Conclusion and limitation

- In this paper, we conduct an empirical study on the impact of a successful rebuttal stage in CS conference peer reviews, including author response and reviewer discussion.
- We hope our research can illuminate strategies for crafting successful rebuttals for reviews and assist authors in getting their submissions accepted.
- This statistical approach may have certain limitations, and additional experiments using causal analysis could be applied to assess strategies more effectively.