ISPY: Automatic Issue-Solution Pair Extraction from Community Live Chats
面向群智的问题及解决方案自动提取技术
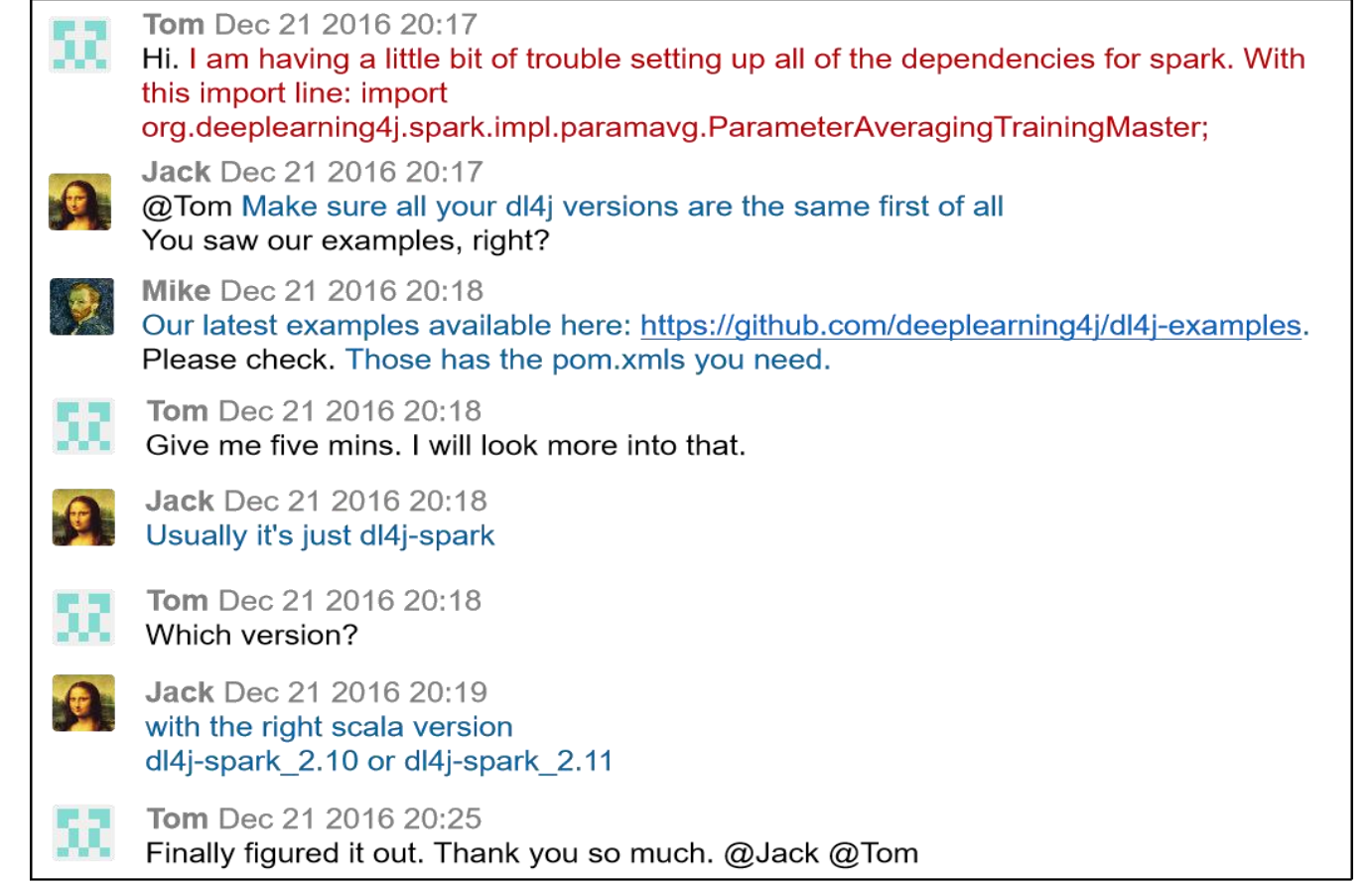
石琳 江子攸 杨叶 陈啸 张昱旻 沐方文 蒋涵芝 王青

The 36th IEEE/ACM International Conference on
Automated Software Engineering (ASE 2021)

主要联系人：石琳（15001193593，shilin@iscas.ac.cn）

Introduction

- Community live chats contain rich sets of information for potential improvement on software quality and productivity. One of the important applications is to mine knowledge on issues and their potential solutions.
- We first formulate the problem of issue-solution pair extraction from developer live chat data, and propose an automated approach, named ISPY, based on natural language processing and deep learning techniques with customized enhancements, to address the problem.
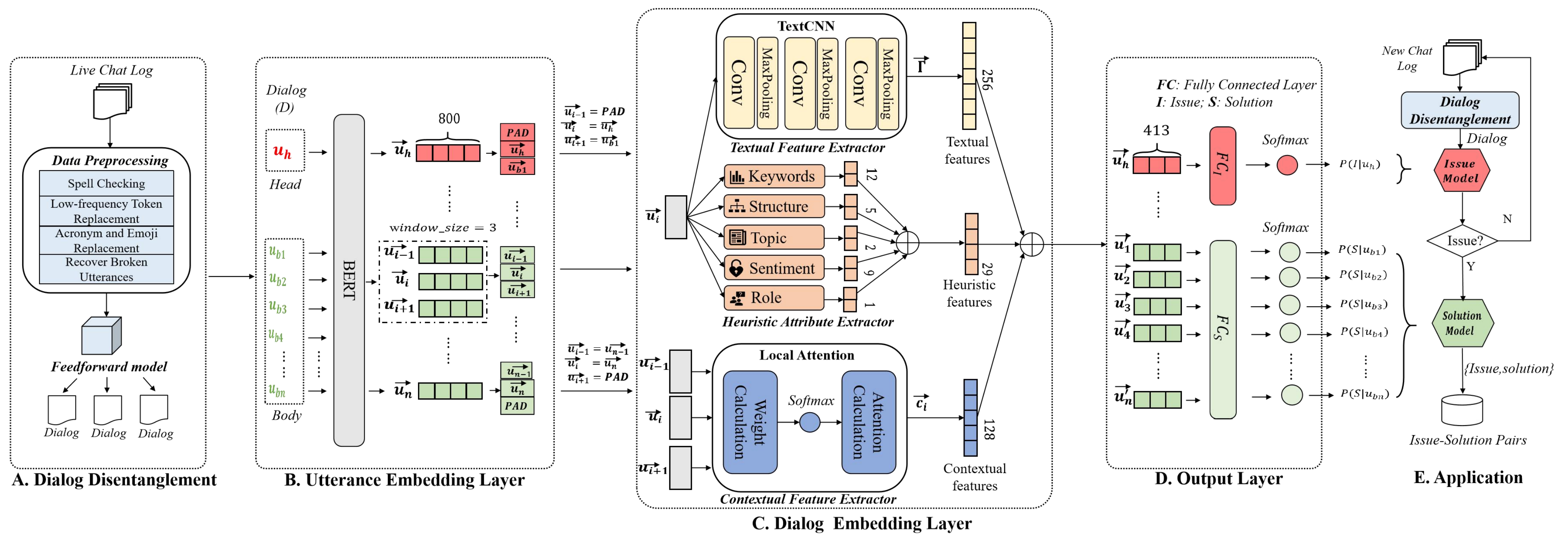

Model

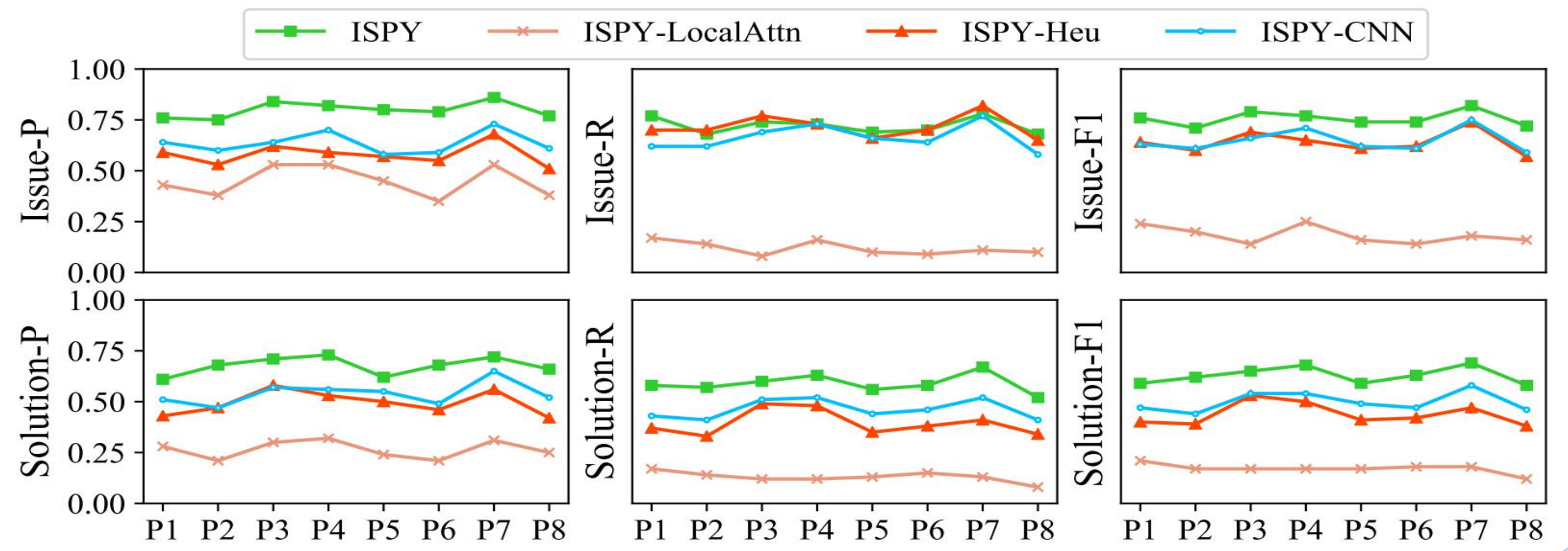A. Dialog Disentanglement　B. Utterance Embedding Layer　C. Dialog Embedding Layer　D. Output Layer　E. Application

Result

Performance in detecting issue dialogs (RQ1): ISPY outperforms the six baselines in detecting issue dialogs across most of the studied projects.

Task	Methods	Angular			Appium			Docker			DL4j			Ethereum			Gitter			Nodejs			Typescript			Average		
		P	R	F1	P	R	F1	P	R	F1	P	R	F1	P	R	F1	P	R	F1	P	R	F1	P	R	F1	P	R	F1
Issue	ISPY	0.76	0.77	0.76	0.75	0.68	0.71	0.84	0.74	0.79	0.77	0.68	0.72	0.82	0.73	0.77	0.8	0.69	0.74	0.79	0.70	0.74	0.86	0.78	0.82	0.8	0.72	0.76
	NB	0.36	0.40	0.38	0.41	0.30	0.35	0.47	0.36	0.41	0.70	0.56	0.62	0.08	0.25	0.13	0.22	0.42	0.29	0.30	0.50	0.37	0.15	0.40	0.22	0.34	0.40	0.36
	RF	0.56	0.25	0.34	0.69	0.30	0.42	0.75	0.23	0.35	0.84	0.44	0.58	1	0.17	0.29	0.50	0.25	0.33	0.33	0.13	0.18	0.23	0.30	0.26	0.61	0.26	0.36
	GBDT	0.27	0.75	0.40	0.40	0.7	0.51	0.59	0.79	0.61	0.73	0.44	0.55	0.21	0.76	0.33	0.19	0.67	0.29	0.30	0.88	0.44	0.18	0.9	0.30	0.35	0.65	0.46
	Casper	0.39	0.35	0.37	0.08	0.03	0.05	0.59	0.26	0.36	0.46	0.40	0.43	0.19	0.42	0.26	0.14	0.17	0.15	0.05	0.06	0.06	0.15	0.40	0.22	0.26	0.26	0.26
	CNC_PD	0.20	0.55	0.29	0.23	0.50	0.32	0.23	0.36	0.28	0.12	0.32	0.17	0.24	0.42	0.30	0.12	0.42	0.19	0.10	0.50	0.17	0.05	0.40	0.10	0.16	0.43	0.24
	DECA_PD	0.33	0.50	0.40	0.28	0.37	0.31	0.33	0.36	0.34	0.64	0.28	0.39	0.42	0.42	0.42	0.44	0.67	0.53	0.32	0.50	0.39	0.04	0.10	0.06	0.35	0.40	0.37
Solution	ISPY	0.61	0.58	0.69	0.68	0.57	0.62	0.71	0.60	0.65	0.66	0.62	0.68	0.73	0.63	0.68	0.62	0.56	0.59	0.68	0.58	0.63	0.72	0.67	0.69	0.68	0.59	0.63
	NB	0.21	0.58	0.30	0.24	0.48	0.32	0.31	0.59	0.40	0.58	0.49	0.53	0.33	0.80	0.47	0.10	0.67	0.17	0.37	0.55	0.44	0.08	0.09	0.09	0.28	0.53	0.37
	RF	0.26	0.83	0.39	0.27	0.52	0.36	0.31	0.56	0.40	0.80	0.13	0.23	0.15	0.40	0.22	0.10	0.33	0.15	0.50	0.65	0.57	0.54	0.64	0.58	0.37	0.51	0.43
	GBDT	0.26	0.92	0.40	0.26	0.74	0.38	0.22	0.62	0.32	0.71	0.16	0.26	0.11	0.40	0.17	0.10	0.67	0.17	0.42	0.75	0.54	0.42	0.73	0.53	0.31	0.62	0.41
	UIT	0.29	0.17	0.21	0.21	0.13	0.17	0.30	0.12	0.17	0.24	0.07	0.12	0.37	0.16	0.22	0.27	0.11	0.17	0.19	0.15	0.18	0.31	0.18	0.23	0.27	0.14	0.18
	CNC_SP	0.18	0.17	0.17	0.26	0.28	0.27	0.21	0.31	0.25	0.46	0.20	0.28	0.30	0.31	0.30	0.46	0.54	0.50	0.41	0.28	0.33	0.38	0.25	0.30	0.33	0.27	0.30
	DECA_SP	0.00	0.00	0.00	0.00	0.00	0.00	0.20	0.09	0.12	0.25	0.02	0.24	0.00	0.00	0.00	0.57	0.20	0.30	0.33	0.09	0.14	0.17	0.05	0.08			

Performance in extracting solutions (RQ2): ISPY outperforms the six baselines in extracting solution utterances in terms of Precision and F1.

Ablation experiment (RQ3): The contextual feature extractor provides a more significant contribution to the effectiveness of ISPY than others.

Application

ISPY helps with unanswered issues on Stack Overflow, and there are 6/26 solutions that have been accepted as best answers.

Conclusion

- We propose an approach, named ISPY, to automatically extract issue-solution pairs from development community live chats.
- We build a dataset with 750 dialogs including 171 issue-solution pairs and evaluate ISPY on it. The evaluation results show that our approach outperforms both issue-detection baselines and solution-extraction baselines by substantial margins.
- By applying ISPY, we automatically generate a dataset with over 30K issue-solution pairs extracted from 11 community live chats, and we utilize the dataset to provide solutions for 26 recent issues posted on Stack Overflow