User Reputation in Community Question Answering

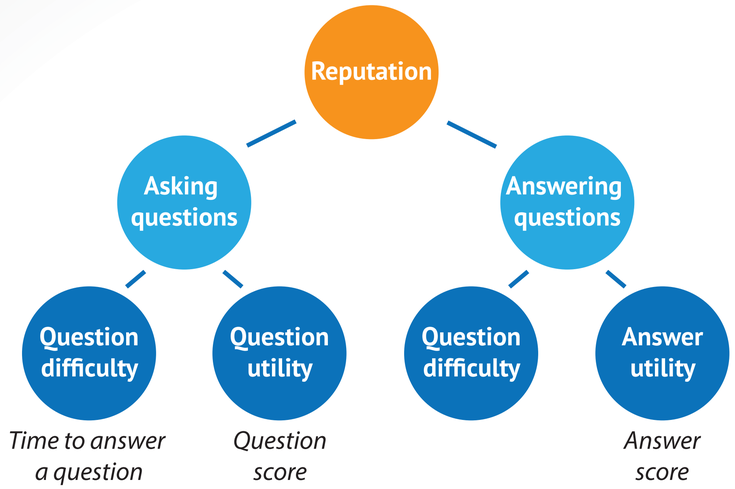
Author: **Adrián Huňa** Supervisor: **Ivan Srba**

MOTIVATION

"Reputation is a combination of user activity and expertise"

Existing methods:
- Focus on user activity.
- Do not reflect bias in topics in Community Question Answering (CQA) systems.
- Do not measure level of expertise.

REPUTATION CALCULATION



- Long tail distribution of question/answer scores and times to answer questions.
  - Solved by logarithm of values.
- Bias in topics solved by values **normalization** by maximum value for each tag a question belongs to.

- **Two variants** of our method:
  - Sum of partial reputations.
  - Average of partial reputations.
    - In order to completely eliminate influence of activity.

EVALUATION

- Problem with golden standard.
- Evaluated using experimental infrastructure in Askalot.
- Programmers.**StackExchange**.com dataset.
- Relative ranking of users for one question.

	P@1	P@2	MRR	nDCG	Questions
Full variant *(sum)*	40,093	38,074	65,538	83,162	20552
Full variant *(average)*	**43,971**	**41,154**	**66,278**	**84,511**	20552
Answers only *(sum)*	40,179	38,267	63,632	83,223	20324
Answers only *(average)*	43,623	40,926	66,182	84,521	20324
StackOverflow Reputation	40,699	38,474	63,943	83,370	20558
Best Answer Ratio	**41,881**	**40,078**	**64,585**	**83,728**	20324
Z-score	38,388	37,022	62,322	82,534	20558
Number of Answers	38,570	37,308	62,481	82,647	20324

APPLICATION IN ASKALOT

- Askalot.fiit.stuba.sk
- Reputation visualisation.

Adrian
Adrian Huna

CONCLUSION

- Reputation focused on expertise.
- Reflection of topics' bias.

- User expertise is important for user reputation estimation.
- Our proposed method outperformed all baseline methods.

STU FIIT

PeWe@FIIT
personalized web group