

XV Международная научно-практическая конференция студентов аспирантов и молодых учёных  
«Молодёжь и современные информационные технологии»

## ВЫЯВЛЕНИЕ НЕДОБРОСОВЕСТНЫХ РАБОТНИКОВ В КРАУДСОРСИНГЕ ЭМОЦИЙ

Olga Korovina

Berestneva O.G.

Tomsk Polytechnic University, [olgakorovina92@gmail.com](mailto:olgakorovina92@gmail.com)

### Введение

Краудсорсинг стал важным инструментом для улучшения доступа к онлайн-ресурсам, особенно к нетекстовой информации. В связи с активным ростом пользовательского контента, важность тэгов растёт. Тэги это слова или короткие фразы, которые пользователь добавляет как ключевые слова к картинкам, видео и другим медиа контентам. [Golbeck].

Расставление тэгов - одно из самых популярных действий в наши дни, что включает в себя тэги объектов, людей или эмоций. Таггинг помогает не только подписать контент, но и также обеспечить поиск по ключевым словам. Хотя бы одни хэштег в Инстаграме увеличивает популярность фото на 12,6%. Самые популярные хэштеги в Инстаграм - название брендов. Инстаграм также позволяет искать картинки используя Эмоджи.

Тэги для картинок могут быть сгенерированы автоматически [Wu] или полуавтоматически, а также с помощью краудсорсинга [Aljanaki, Goto]. Проведенные исследования тэгов по контенту [<https://everypixel.com>, Bar-Ilan, Golbeck], эмоциям [Chafale] и качеству [<https://everypixel.com>]. В целом, генерация тэгов с помощью краудсорсинга более точная и практичная, чем автоматическая аннотация [Wu].

В большинстве исследовательских работ на сегодняшний день основное внимание уделяется инструменту для сбора эмоций по картины [Runge, Siegert], и меньше внимания уделено процессу анализу информации и выбора нужной эмоции.

Краудсорсинг широко используется в наши дни, потому что этот метод дешевле и даёт быстрый результат {ribeiro2011}. Но при краудсорсинге сложно проверить качество работников и достоверность данных, поскольку краудсорсинг обеспечивает небольшой контроль над средой в процессе выполнения задания. Более сложной задачей является проверка качества краудсорсинга субъективных задач. Под субъективными задачами понимается задание не имеющее одного правильного ответа. Как правило субъективные задачи нацелены на своб мнений.

Одной из проблем, связанных с краудсорсингом субъективных задач, является обнаружение некорректных работников, поскольку онлайн-работники выполняют задачи без надзора. В крупных исследованиях есть по крайней мере один работник, который представляет случайные ответы, не читая задание {ribeiro2011}.

### Способы обнаружения недобросовестных работников

Один из способов обнаружения недобросовестных работников - наблюдение за их поведением во время выполнения задания. Например, Mok et al. создали методологию для скрининга низкокачественных работников путем анализа траектории курсора мыши (mok2015). Или Kim et al. определяют работников, которые нажимают случайно {kim2012}. Другой способом выявить недобросовестных работников- задания, в которых правильный ответ известен {hossfeld2011}.

Проверка ответов после выполнения задач работниками и устранения “лживых” ответов поможет увеличить точность полученных данных. Предоставление бонусов и выплата вознаграждений также способствуют повышению качества ответов {ribeiro2011}.

### Существующие платформы для проверки недобросовестных работников краудсорсинга

Существующие краудсорсинговые платформы для оценки качества субъективных задач:

*CrowdMOS* - это экономичная и удобная платформа для оценки качества субъективных задач, в которой интернет-работники участвуют в стандартных тестах, таких как MOS (mean opinion score). Для оценки надёжности пользователей *CrowdMOS* использует ACR (absolute category rating)- коэффициент корреляции выборки между средним пользовательским рейтингом работника и глобальным средним рейтингом {ribeiro2011crowdmos}, {ribeiro2011}.

*Quadrant of Euphoria* - краудтест платформа для онлайн-оценки мультимедии. Данный метод использует парное сравнение TSR (transitivity satisfaction rate) \ cite {wu2013}, \ cite {chen2010}.

*WESP* (веб-платформа субъективной оценки) {rainer2013} на основе рекомендаций ITU (international communication union) для субъективных оценок качества мультимедии {itu1999}. *WESP* позволяет настраивать субъективную оценку качества в соответствии с требованиями и целями исследования. Это очень гибкая платформа, которая использует широкий спектр различных методологий, таких как единый стимул, двойной стимул, парное сравнение или непрерывная оценка качества, что позволяет добавлять предварительные вопросы, механизм голосования, рейтинговую шкалу и контрольные вопросы.

*QualityCrowd* - это краудсорсинг платформа для QoE-оценки (Quality of experience) {keimel2012}. *QualityCrowd* также является гибкой платформой и использует различные методологии тестирования (один стимул или двойной стимул, а также различные шкалы).

*In-momento crowdsourcing* использует QoE (Quality of experience) и оценку поведения пользователей. Данный метод позволяет контролировать время, полноэкранный режим или перемещение между окнами браузера {gardlo2014}

#### **Заключение**

Существует множество методологий для распознавания работников, выполняющих краудсорсинговые задания недобросовестно. Некоторые из них очень {ribeiro2011crowdmos}, в то время как другие направлены на выполнение определенного вида задания {valeri2016}. Некоторые платформы не предоставляют возможность настроить проверку работников {wu2013,chen2010}, другое очень гибкие и позволяют менять настройки под каждое задание {keimel2012,rainer2013}, многие из них позволяют добавить вопросы для определения недобросовестных работников.

#### **Список использованных источников**

1. Bar-Ilan, Judit, et al. "Structured versus unstructured tagging: a case study." *Online Information Review* 32.5 (2008): 635-647.
2. Chafale, Dhanashri, and Amit Pimpalkar. "Review on Developing Corpora for Sentiment Analysis Using Plutchik's Wheel of Emotions with Fuzzy Logic." *International Journal of Computer Sciences and Engineering (IJCSE)* 2 (2014): 14-18.
3. Dan-Glauser, Elise S., and Klaus R. Scherer. "The Geneva affective picture database (GAPED): a new 730-picture database focusing on valence and normative significance." *Behavior research methods* 43.2 (2011): 468-477.
4. Eleta, Irene, and Jennifer Golbeck. "A study of multilingual social tagging of art images: cultural bridges and diversity." *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*. ACM, 2012.
5. Peng, Kuan-Chuan, et al. "A framework of changing image emotion using emotion prediction." 2014 IEEE International Conference on Image Processing (ICIP). IEEE, 2014.
6. Renaningtyas, Luri, Intan Rizky Mutiaz, and Achmad Syarief. *NOSTALGIA AND AUTHENTICITY OF FAUX VINTAGE PHOTO IN INSTAGRAM THROUGH DIGITAL FILTERS*. Diss. Petra Christian University, 2014.
7. Runge, Nina, et al. "Tag your emotions: a novel mobile user interface for annotating images with emotions." *Proceedings of the 18th International Conference on Human-Computer Interaction with*

*Mobile Devices and Services Adjunct*. ACM, 2016.

8. Wu, Lei, et al. "Learning to tag." *Proceedings of the 18th international conference on World wide web*. ACM, 2009.