

РАЗМЕТКА ИЗОБРАЖЕНИЙ МАССОВОГО МЕРОПРИЯТИЯ ЕГО УЧАСТНИКАМИ НА ОСНОВЕ НЕМОНЕТАРНОГО СТИМУЛИРОВАНИЯ

А. В. Пономарев^а, канд. техн. наук, старший научный сотрудник

^аСанкт-Петербургский институт информатики и автоматизации РАН, Санкт-Петербург, РФ

Введение: растущая информатизация и распространение глобальных сетей позволяют использовать естественные когнитивные способности людей, координируемых с помощью современных информационно-коммуникационных технологий, для ряда задач, автоматическое решение которых на данный момент не позволяет достигнуть необходимого качества результатов. **Цель:** анализ возможности использования скоординированных усилий сообщества людей, действующих на основе немонетарного стимулирования, для разметки изображений на примере разметки фотографий легкоатлетических пробегов номерами изображенных на них участников в целях облегчения поиска фотографий. **Результаты:** разработан интернет-сервис для совместной разметки изображений. Проведенные эксперименты, в которых приняли участие несколько сотен человек, показали, что «внутреннее» немонетарное стимулирование, связанное с облегчением поиска фотографий, оказывается достаточным для привлечения необходимого количества участников сообщества к разметке, а качество получаемой разметки сопоставимо с качеством, достигаемым с помощью платформы краудсорсинга Amazon Mechanical Turk при использовании монетарного стимулирования. **Практическая значимость:** полученные результаты могут быть использованы при разработке различных систем, основанных на скоординированных усилиях сообщества на базе немонетарного стимулирования.

Ключевые слова — краудсорсинг, крауд-вычисления, обработка изображений, коллективный интеллект, социальные сети.

Введение

Несмотря на успехи компьютерных наук, по-прежнему существуют задачи обработки информации, полностью автоматические решения которых не всегда возможны или не обеспечивают требуемый уровень качества получаемых результатов. Одной из таких задач является распознавание изображений. И хотя современные системы распознавания изображений проделали длинный путь и во многих важных с практической точки зрения приложениях прекрасно справляются со своей задачей (например, при распознавании напечатанного текста), в ряде сложных ситуаций (неравномерная освещенность, наличие искажений) качество распознавания может существенно снижаться.

Одним из способов решения этой проблемы, продиктованным растущей информатизацией и распространением глобальных сетей, является использование естественных когнитивных способностей людей, координируемых посредством современных информационно-коммуникационных технологий. Данный способ получил название крауд-вычисления (или, в более общем смысле, краудсорсинг).

Однако вовлечение человека в процесс обработки информации неизбежно требует учета ряда факторов: ограниченной его производительности, необходимости мотивации и возможности ошибки или даже целенаправленного искажения информации. Это все предъявляет специфиче-

ские требования к проектированию и построению систем крауд-вычислений.

Одной из практических задач, демонстрирующих подобные сложные для современных систем распознавания условия, является распознавание на фотографиях номеров участников легкоатлетических пробегов. Любительские легкоатлетические пробеги являются популярным видом досуга, причем количество как самих пробегов, так и их участников растет с каждым годом и в России, и за рубежом [1–4]. Учитывая широкое распространение цифровых фотокамер, пробеги, как правило, фиксируются как любителями, так и профессиональными фотографами, многие из которых публикуют фотографии в открытом доступе. Участники пробега, в свою очередь, часто бывают заинтересованы в получении собственных фотографий — либо в качестве сувенира, либо для анализа техники бега. Однако поиск фотографий является трудоемкой задачей, поскольку количество опубликованных в социальных медиа фотографий с мероприятия среднего уровня может исчисляться тысячами. К счастью, организация большинства пробегов предполагает, что участники прикрепляют нагрудные номера (являющиеся их уникальными идентификаторами в рамках пробега), что теоретически позволяет передать задачу поиска от самого заинтересованного участника любому, кто в состоянии распознавать номера на фотографиях. Существует несколько работ, в которых для решения этой задачи предлагаются системы, использующие технологии компью-

терного зрения [5, 6], однако в силу того, что нагрудные номера являются гибкими (следовательно, могут быть искажены) и на фотографии могут быть частично закрыты (руками участника или другими участниками пробега), полностью автоматические решения зачастую не обеспечивают желаемого качества распознавания.

Настоящая статья посвящена исследованию возможности использовать скоординированные усилия сообщества для решения задачи разметки фотографий пробега нагрудными номерами изображенных на них участников в целях последующего облегчения поиска фотографий заданного участника. Сообщество это формируется (преимущественно) из самих участников пробега, а основным предполагаемым стимулом к участию в разметке является сокращение времени, затрачиваемого на поиск фотографий.

Разметка (и аннотирование) изображений является одним из наиболее используемых и хорошо изученных применений краудсорсинга [7–10]. Однако решаемая в рамках данной статьи задача имеет существенные отличия от рассмотренных ранее в литературе:

1) участники пробега, как правило, более заинтересованы в просмотре фотографий (с эмоциональной точки зрения), чем случайные исполнители (*workers*) коммерческой платформы краудсорсинга (например, Amazon Mechanical Turk — АМТ [11]);

2) использование желания участников найти свои фотографии с наименьшими усилиями создает немонетарный стимул, позволяющий при определенных условиях выполнить трудоемкую задачу без финансовых затрат (что может быть полезно администрации пробега).

Для исследования такой возможности автором статьи был разработан интернет-сервис *Bibtaggers* [12], позволяющий координировать разметку изображений участниками сообщества, а также проведен ряд экспериментов, целью которых был поиск ответа на следующие вопросы.

1. Является ли разметка фотографий нагрудными номерами участников разумным подходом к решению более общей задачи — разметки фотографий идентификаторами участников?

2. Как влияет на качество результата использование при разметке *контекста изображения*, т. е. дополнительной информации, не содержащейся непосредственно на изображении. В частности, насколько полезным оказывается исследование соседних фотографий (сделанных до и после данной)?

3. Является ли немонетарное стимулирование, основанное на облегчении поиска, достаточным для привлечения к разметке необходимого количества участников сообщества?

4. Каково ожидаемое качество разметки, осуществленной участниками сообщества?

5. Как соотносится качество разметки, осуществленной участниками сообщества на основе «внутреннего», немонетарного стимулирования, и разметки, полученной с помощью АМТ при использовании монетарного стимулирования?

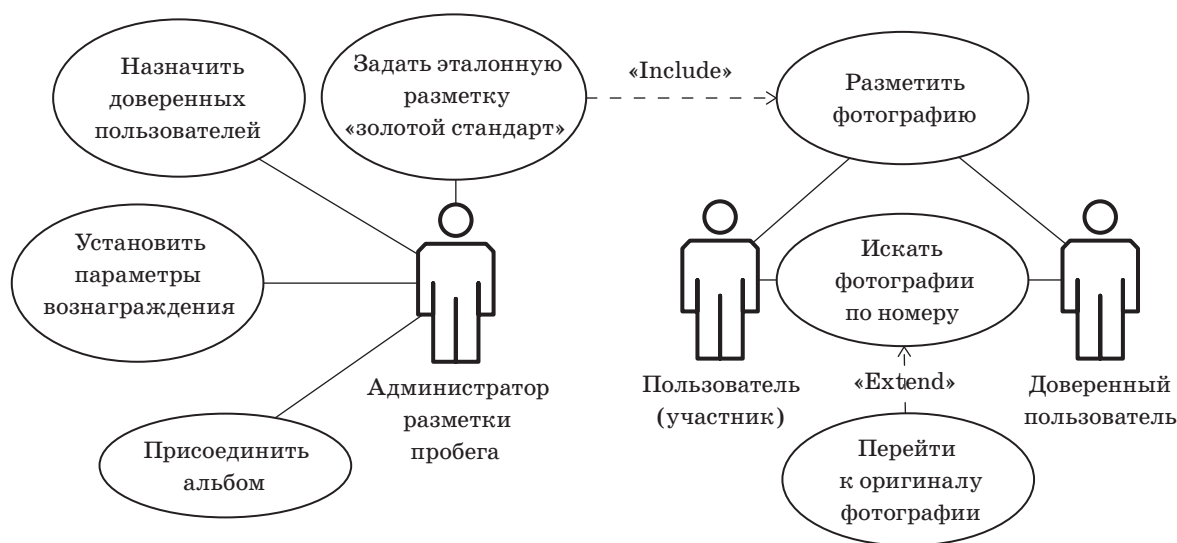
Структура и принцип работы интернет-сервиса

Основные положения

Необходимым условием функционирования любой системы обработки информации, включающей операции, выполняемые человеком, является создание заинтересованности человека в выполнении этих операций. Применительно к крауд-вычислениям подобное создание заинтересованности получило в англоязычной литературе название *incentivization* (или стимулирование). При разработке интернет-сервиса предполагалось, что основным мотивом к участию является снижение времени и трудоемкости поиска собственных фотографий, поэтому выбрана функциональная стратегия стимулирования, ставящая доступность нужных пользователю функций (в частности, поиска фотографий по номеру) в зависимость от вклада, который пользователь внес в разметку фотографий. Выделено два вида «вознаграждения» пользователя: возможность поиска фотографий по номеру участника (при этом найденные фотографии отображаются в небольшом разрешении, достаточном для того, чтобы проверить, что участник действительно присутствует на фотографии, но недостаточном для того, чтобы распечатать фотографию или опубликовать в социальных сетях) и возможность перехода к полноразмерным оригиналам найденных фотографий. Каждой из таких «функций-наград» поставлен в соответствие порог, задающий, при каком количестве обработанных пользователем фотографий должна активироваться соответствующая функция. Значения порога выбираются для каждого пробега отдельно в зависимости от количества фотографий и предполагаемого количества участников разметки.

Упрощенная диаграмма вариантов использования интернет-сервиса представлена на рис. 1. Выделены три типа акторов: администратор разметки пробега, пользователь (участвующий в разметке) и доверенный пользователь («эксперт»).

Администратор разметки пробега осуществляет конфигурирование различных параметров, связанных с разметкой. В частности, он может присоединять к пробегу альбомы, опубликованные различными фотографами на внешних сервисах обмена фотографиями (например, Яндекс.Фотки [13]); устанавливать значения параметров, влияющих на распределение вознаграждения; назначать доверенных пользователей и задавать



■ Рис. 1. Диаграмма вариантов использования веб-приложения

эталонную разметку фотографий («золотой стандарт»), используемую для контроля качества.

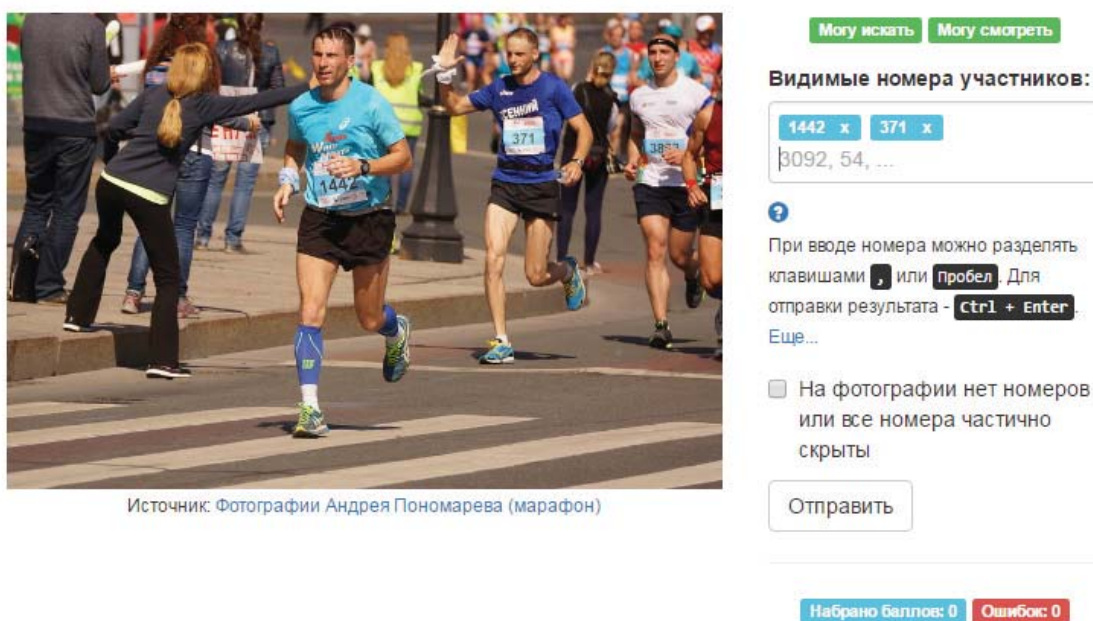
Пользователь (участник) может действовать по двум основным направлениям: просмотреть предложенные интернет-сервисом фотографии и указать видимые на них номера участников; искать фотографии по номеру участника. Первый сценарий доступен любому пользователю (зарегистрированному в системе или анонимному), а второй, как правило, требует внесения определенного вклада в разметку, размер которого устанавливается администратором пробега. Вклад вычисляется как разница между количеством обработанных пользователем фотографий и ко-

личеством ошибок на эталонных фотографиях (см. п. «Защита от мошенничества»).

Доверенный пользователь с точки зрения основных сценариев использования идентичен обычному пользователю, однако разметка, полученная от такого пользователя, считается надежной и не требует проверки, в отличие от разметки обычного пользователя. Перечень доверенных пользователей назначается администратором пробега.

Пользовательский интерфейс

Пользовательский интерфейс сервиса состоит из нескольких страниц, основной из которых является страница разметки (рис. 2), содержащая фото-



Источник: Фотографии Андрея Пономарева (марафон)

■ Рис. 2. Интерфейс страницы разметки

графию пробега, поле ввода списка видимых номеров участников, отдельное (бинарное) поле ввода для идентификации того, что на демонстрируемой фотографии не видно ни одного номера, и кнопку для отправки формы. Помимо перечисленных обязательных элементов, страница разметки может содержать ссылки на предыдущую и последующую фотографии в серии (это может быть полезно для распознавания частично заслоненного номера) и бинарное поле ввода для указания того, что данный список номеров должен восприниматься как эталонный (поле отображается только администратору пробега). Роль отдельного бинарного поля ввода для выделения фотографий, не содержащих номеров, заключается в дополнительной проверке целостности ввода пользователя. Наличие этого бинарного поля делает невозможным отправку незаполненной формы; участник должен обязательно совершить какое-либо действие: либо перечислить видимые номера, либо явно указать, что их нет.

Задания

В рамках сервиса принято следующее разграничение понятий «задача» и «задание». Под «задачей» понимается потребность получения множества номеров участников, видимых на некоторой фотографии. Решение такой задачи, как правило, подразумевает выполнение нескольких «заданий», каждое из которых предъявляется одному участнику разметки и требует от него выполнения определенных действий. Задание может иметь определенные требования к участникам, которым оно может быть предъявлено. Так, некоторые задания могут предъявляться всем участникам, а некоторые — только доверенным участникам (как правило, такие задания используются для согласования противоречивых результатов).

Стратегия планирования

Под планированием понимается процесс определения результата задачи на основе имеющихся результатов соответствующих заданий и, если это оказывается необходимым, потребность в выполнении дополнительных заданий. Реализованная стратегия планирования использует основанный на избыточности контроль качества с ослабленным консенсусом. В первую очередь определяются результаты выполнения двух заданий; если они оказываются схожими (подробнее о формальном определении сходства ниже), то результат задачи определяется как объединение результатов заданий. В противном случае создается третье задание. Если среди результатов выполнения трех заданий присутствуют два схожих, то результат задачи вычисляется как объединение двух наиболее схожих, иначе создается доверенное задание, и итоговым решением задачи станет результат, предоставленный доверенным пользователем. Степень сходства

между результатами выполнения заданий определяется как индекс Жаккара [14]:

$$J = |R_i \cap R_j| / |R_i \cup R_j|,$$

где R_i и R_j — это множества номеров, указанные в результатах заданий i и j соответственно. Результаты считаются схожими, если индекс превышает некоторый порог (обычно 0,6).

Защита от мошенничества

Под мошенничеством в данном контексте понимается получение вознаграждения через предоставление заведомо низкокачественных результатов (пустых или сфурмированных с минимальными усилиями, случайных, не имеющих отношения к изображению фотографии). Для ограничения возможностей мошенничества используется два метода: эталонная разметка («золотой стандарт») и контроль времени.

«Золотой стандарт» представляет собой несколько фотографий, размеченных администратором пробега. Такие фотографии показываются участнику обязательно в начале каждой сессии и затем в ее ходе с некоторой вероятностью, зависящей от количества правильно размеченных фотографий «золотого стандарта». Неправильная разметка таких фотографий отодвигает момент активации вознаграждения, в чем участник не заинтересован. Продолжить процесс разметки можно, только дав правильный (эталонный) ответ; об ошибке участник оповещается.

Контроль времени позволяет выявить слишком быстрые ответы, как правило, отправленные пользователем, не предпринявшим должных усилий по анализу представленного ему изображения и заполнению полей формы. При выявлении таких действий пользователю отображается соответствующее предупреждение.

Эти меры могут быть полезны не только для борьбы с мошенничеством в обозначенном выше смысле, но также помогают ограничить вред от умышленного искажения разметки. Если пользователь систематически заявляет, что на фотографии изображены номера, которые на ней не изображены (или наоборот), и исправляет ввод при попадании на фотографию «золотого стандарта», то через несколько фотографий вероятность показа ему только фотографий «золотого стандарта» будет близка к 1, что существенно уменьшит его влияние на общее качество разметки.

Определение качества разметки

Разметка фотографий номерами участников, основанная на визуальном анализе самой фотографии, — это лишь один из способов достичь более общей цели, а именно разметки фотографий

идентификаторами людей, на ней изображенных. Визуальная разметка номерами участников имеет несколько принципиальных ограничений, в частности, участник может быть виден на фотографии, однако его номер может быть полностью или частично заслонен. Другие подходы к разметке, например, основанные на распознавании лиц или на применении радиомаячков, могут в этой ситуации оказаться успешными. Таким образом, необходимы способы оценки качества решения задачи, которые могут быть использованы как для сравнения различных схем стимулирования при разметке фотографий усилиями сообщества, так и для оценки ограничений, связанных в принципе с разметкой, основанной на визуальном анализе номеров участников. Использование существующих наборов данных («бенчмарков») осложняется тем, что при разметке усилиями сообщества его участники не заинтересованы в разметке фотографий таких публичных наборов данных, в отличие от исполнителей краудсорсинговой платформы, действующих на базе монетарного стимулирования, поэтому для оценки качества разметки подготовлено несколько наборов данных [15].

DS0 — набор данных, содержащий разметку фотографий номерами участников, причем единственной информацией, доступной эксперту-составителю набора, была сама размечаемая фотография. Другими словами, фотография в *DS0* имеет ассоциированный с ней номер, если эксперт видит участника с этим номером на фотографии. Если эксперт не в состоянии непротиворечиво определить номер, то он не должен указывать его вообще (т. е. он не должен указывать все возможные варианты). Целью этого набора данных является измерение качества разметки, полученной от участников сообщества, по сравнению с качеством разметки, полученной от эксперта.

DS1 — набор данных, в котором эксперту разрешается при разметке использовать минимальный контекст, т. е. дополнительную информацию — помимо размечаемой фотографии, просматривать одну фотографию, сделанную непосредственно до, и одну фотографию, сделанную непосредственно после данной, для выяснения нечетко видных фрагментов номера. Целью данного набора является измерение качества разметки участниками сообщества по сравнению с разметкой экспертом при возможности использования минимального контекста.

*DS** — набор данных, содержащий идеальную разметку фотографий идентификаторами участников (в действительности, содержащий разметку, лишь близкую к идеальной). В строгом смысле, идеальная разметка фотографии включала бы идентификаторы (в простейшем случае — номера) всех участников, видимых на фотографии. Очевидно, для ряда фотографий (на-

пример, панорамных) такую разметку построить невозможно. Поэтому при построении *DS** критерий был смягчен, и фотография в *DS** размечена номером участника, если номер видим или если изображение участника хорошо различимо на фотографии. При этом, если номер не видим, то для участника случайным образом генерировался уникальный идентификатор, и фотография помечалась этим идентификатором. Целью этого набора данных было проанализировать, как разметка, основанная на анализе изображений, соотносится с эталонной разметкой.

Все три набора данных построены на одном и том же множестве из 100 фотографий, случайным образом выбранных из 4447 фотографий одного пробега, сделанных независимо двумя фотографами.

Пример того, как одна и та же фотография размечена в разных наборах, приведен на рис. 3. Номер 1010 хорошо различим на фотографии, поэтому содержится в *DS0*. Анализ предыдущего изображения позволяет восстановить номер крайней слева участницы (это 1004). Наконец, третий и четвертый участники хорошо видимы, но их номера не различимы, поэтому в *DS** им сопоставлены уникальные метки.

Частичные ответы на некоторые вопросы, касающиеся сопоставления различных способов разметки, могут быть получены без экспериментов с сообществом только за счет сопоставления подготовленных наборов.

Первый вопрос такого рода заключается в оценке того, насколько эффективна разметка фотографий номерами участников на основе анализа изображения в качестве решения более общей задачи — разметки фотографий идентификаторами участников. Для ответа на этот вопрос следует сопоставить наборы данных, содержащие только номера участников (*DS0*, *DS1*) и набор данных *DS**, содержащий помимо номеров уникальные идентификаторы для тех участников пробега, номера которых оказались не видны.

В целях сопоставления каждый набор данных представлен в виде множества пар, состоящих из идентификатора фотографии и номера (идентификатора) участника. Если в наборе данных фотография размечена несколькими номерами, то такой фотографии будет соответствовать несколько пар. Для измерения степени соответствия наборов данных использованы меры точность (*precision*) и полнота (*recall*), широко используемые в информационном поиске (например, см. [16]). Точность соответствует доле элементов «испытываемого» набора, присутствующих в эталонном наборе, а полнота, наоборот, доле элементов эталонного набора, вошедших в испытываемый.

Результаты вычисления точности и полноты *DS0* и *DS1* по отношению к *DS**, рассматриваемо-



DS0:
1010
DS1:
1010, 1004
DS:*
1010, 1004,
_GREY,
_BLUE

Ленинградская область, г. Сестрорецк
Сестрорецкий полумарафон

15 мая 2016 г.
Фото: Шалимов Владимир, <http://nabegu.spb.ru> & bibtaggers.ru

■ **Рис. 3.** Иллюстрация соотношения тестовых наборов

му в качестве эталона, представлены в табл. 1. Из таблицы видно, что показатель точности для обоих наборов равен 1,0 — это следствие способа конструирования DS^* и соотношения $DS0 \subseteq DS1 \subseteq DS^*$ между содержимым наборов. Полнота же показывает, что для случайным образом выбранного изображения участника на фотографии вероятность того, что этот факт будет зафиксирован экспертом, исследующим только изображение, оценивается как 0,65 ($DS0$). Если же эксперту разрешается обращаться к соседним фотографиям, то вероятность повышается до 0,803 ($DS1$), поскольку соседние фотографии могут быть использованы для установления неразличимых на данной фотографии номеров.

Позволяет табл. 1 оценить также эффект анализа соседних фотографий в серии. Видно, что на исследуемом наборе фотографий прирост полноты составил около 23,5 %.

В определенном смысле полнота $DS1$ устанавливает теоретический максимум, который может быть достигнут при разметке силами сообщества, когда участникам доступен минимальный

контекст, т. е. разрешается «заглядывать» на соседние изображения. Существуют и другие типы контекста, которые могут оказаться полезными для увеличения полноты разметки (например, с привлечением информации протоколов пробега), но увеличение количества исследуемых типов контекста увеличивает и количество времени и усилий, которые потенциально будут затрачиваться на их анализ, одновременно снижая шансы того, что участники сообщества, действующие на основе немонетарного стимулирования, будут активно использовать эти возможности.

Экспериментальное исследование

Целью проведенного экспериментального исследования был анализ поведения участников сообщества и оценка качества разметки, выполненной с их помощью.

Статистика использования интернет-сервиса

Сервис был использован для разметки фотографий пяти спортивных мероприятий (обозначенных E1–E5). Некоторые данные об использовании сервиса в ходе этих мероприятий приведены в табл. 2.

Фотографии трех пробегов (E1, E4 и E5) были размечены полностью. Видно также, что в разметке фотографий этих трех пробегов приняли участие несколько сотен человек. Принципиальное отличие E1, E4 и E5 от E2 и E3 в том, что в первую группу входят более крупные мероприятия, собравшие значительно большее число участни-

■ **Таблица 1.** Сопоставление наборов данных

Набор данных	Точность	Полнота	Прирост полноты
$DS0$	1,0	0,65	—
$DS1$	1,0	0,803	23,5%

■ Таблица 2. Данные об использовании сервиса

Показатель	E1	E2	E3	E4	E5
Количество участников пробега	~1500	~400	~700	~1300	~7300
Количество фотографий	2770	821	563	4447	7126
Количество фотографов	1	1	1	2	3
Количество пользователей, принявших участие в разметке фотографий	~228	~7	~15	~155	~379
Все фотографии обработаны	+	-	-	+	+
Приблизительное время обработки, ч	32	-	-	65	34
Количество найденных номеров (подтвержденных)	1223	201	118	1088	4902
Количество номеров, по которым были поисковые запросы	639	27	32	529	3197
Настройки порогов активации вознаграждения					
Порог активации поиска	25	30	20	30	20
Порог активации перехода к оригиналам	30	35	25	40	30
Статистика по выполнению заданий					
Выполнено заданий:	7181	902	595	9861	16 476
анонимными пользователями	3622	450	378	5195	7784
зарегистрированными пользователями	3559	452	217	4666	8692
Зарегистрированных пользователей	~72	4	4	62	160
Анонимных сессий	156	3	11	93	219

ков и хорошо представленные в социальных сетях. Ссылки на страницы соответствующих пробега были опубликованы в соответствующих тематических группах, и большинство участников разметки пришло именно по этим ссылкам. События E2 и E4 также были достаточно многочисленными, но значительно слабее освещены в социальных сетях, соответственно, значительно меньше участников пробега знали о возможности разметки и последующего поиска фотографий по номеру. Это объяснение подтверждается и статистикой поисков. В пробегах E1, E4 и E5 поисковые запросы в течение первой недели после мероприятия были осуществлены для 42, 40 и 44 % номеров участников соответственно. Таким образом, приблизительно 40 % участников пробега заинтересованы в поиске своих фотографий (этот процент может быть еще больше, поскольку не все участники используют социальные сети). Для E2 и E3 только 5 % участников пробега искали свои фотографии, что говорит либо о совершенно другом сообществе (с другими приоритетами), либо о недостатке информированности о такой возможности. Положительный результат для пробега E1, E4 и E5 позволяет сделать вывод о достаточности немонетарного стимулирования для привлечения участников к решению задачи.

Качество разметки

Для оценки качества разметки был применен тот же подход, что и при сравнении наборов данных, составленных экспертами. А именно, собранный набор был представлен в виде множества пар, и затем были оценены точность и полнота этого набора относительно эталонных наборов. Помимо точности и полноты, для простоты сопоставления различных вариантов разметки вычислялась также интегральная мера $F1$ [16]:

$$F1 = 2PR / (P + R),$$

где P — точность, а R — полнота.

Результаты сведены в табл. 3. Можно заметить, что разметка, полученная силами сообщества, ближе всего к $DS0$ и существенно отличается

■ Таблица 3. Данные об использовании сервиса

Эталонный набор данных	Точность	Полнота	$F1$
$DS0$	0,925	0,912	0,919
$DS1$	0,957	0,763	0,849
DS^*	0,957	0,613	0,747

■ Таблица 4. Качество разметки участниками АМТ

Эталонный набор данных	\$ 0,05			\$ 0,02		
	Точность	Полнота	F1	Точность	Полнота	F1
DS0	0,858	0,958	0,905	0,906	0,923	0,914
DS1	0,879	0,794	0,835	0,925	0,763	0,836
DS*	0,879	0,639	0,740	0,925	0,614	0,738

ся от *DS1* и *DS**. Поскольку точность достаточно высока, участники сообщества в подавляющем большинстве случаев указывали только те номера, которые действительно присутствовали на изображении. Однако полнота оказалась не очень высокой, что свидетельствует о том, что члены сообщества приложили в целом меньше усилий к указанию всех видимых на фотографиях участников пробега, нежели составители эталонного набора. Низкая полнота на наборе *DS** может быть объяснена частично внутренними ограничениями разметки, основанной только на анализе изображений, а частично следует из низкой полноты на *DS1*. Последняя, в свою очередь, показывает, что очень немногие члены сообщества в действительности пытались использовать соседние снимки для выяснения (частично) скрытых номеров. Практическим следствием этого является необходимость модификации используемой системы стимулирования таким образом, чтобы повысить привлекательность использования соседних снимков (и, как следствие, повысить общее качество разметки).

Сравнение с монетарным стимулированием на платформе АМТ

Важным вопросом, возникающим в связи с разметкой фотографий с помощью участников сообщества, является то, как соотносится ожидаемое качество разметки с качеством разметки, получаемом при решении аналогичной задачи людьми за вознаграждение. Для ответа на этот вопрос были проведены два эксперимента, в ходе которых исполнителям, сотрудничающим с коммерческой краудсорсинговой платформой АМТ, были предложены фотографии из эталонного набора данных (следует подчеркнуть, что данные исполнители не были частью сообщества участников пробега, они назначены с помощью внутренних механизмов платформы АМТ и выполняли задания только за счет монетарного стимулирования). По каждой фотографии было собрано три ответа, которые были обобщены с помощью того же алгоритма, что использован в разработанном интернет-сервисе *Bibtaggers*. Разница заключалась в том, что в первом эксперименте за выполнение задания исполнитель получал \$ 0,05, а во втором — \$ 0,02. Размер вознаграждения выбран

в соответствии с исследованием [17]. Результаты оценки качества представлены в табл. 4.

Как видно из табл. 4, разница в размере вознаграждения почти не оказала влияния на качество получаемого результата (в смысле точности, полноты и значения меры *F1*), однако время разметки (не указано в таблице) при стоимости \$ 0,02 было несколько больше, в то же время количество намеренно некачественных результатов было больше при стоимости \$ 0,05. Это согласуется с существующими исследованиями АМТ [18, 19]. Сопоставление данных табл. 3 и 4 позволяет сделать вывод, что разметка силами сообщества на основе немонетарного стимулирования по интегральному показателю качества не уступает разметке на основе монетарного стимулирования. При этом суммарная стоимость разметки составила \$ 15 (для эксперимента с вознаграждением \$ 0,05 за фотографию) и \$ 6 (для эксперимента с вознаграждением \$ 0,02 за фотографию). Полная разметка фотографий, полученных, например, на мероприятии *E5*, стоила бы \$ 1068 и 427 соответственно, не отличаясь существенно по качеству от результатов, полученных усилиями сообщества, действующего на основе немонетарного стимулирования.

Заключение

В ходе проведения исследования разработан интернет-сервис для разметки фотографий номерами участников с помощью сообщества самих участников пробега. Практическая оценка сервиса показала, что немонетарное стимулирование, связанное с облегчением поиска размеченных таким образом фотографий, оказывается адекватным, и участники с готовностью включаются в разметку, будучи информированы о такой возможности. Эксперименты позволили выявить еще один источник мотивации, не очевидный в ходе проектирования и реализации системы. Этот источник сводится к ценности ощущения включенности в цепочку общественно-полезных действий, выполняемых на некоммерческой основе, своего рода «благодарность» за труд фотографа, который осуществлял съемку и бесплатно сделал фотографии доступными для сообщества. За счет этого вклад некоторых членов сообщества

значительно превысил необходимый для активации возможности поиска.

Сравнение результатов разметки со специальным образом подготовленными эталонными наборами показало, что качество разметки, проведенной усилиями членов сообщества, сопоставимо с «поверхностной» разметкой, при которой определяются только явно видимые номера и не делаются попытки проанализировать контекст в виде хронологически соседних фотографий (хотя реализованный сервис допускает такую возможность). Вместе с тем при активизации этой возможности полнота разметки может вырасти достаточно существенно (около 20 %).

Сопоставление результатов разметки, полученных с применением различных видов стимулирования, показало, что качество разметки от-

личается несущественно. При значительно меньших финансовых затратах разметка усилиями сообщества является более привлекательной.

Среди дальнейших планов — исследование влияния других немонетарных механизмов стимулирования (например, геймификации) на характер участия пользователей в системе и качество получаемых результатов, разработка механизмов стимулирования для использования дополнительной контекстной информации при разметке и, как следствие, повышение ее качества, а также применение проблемно ориентированных мер качества, фиксирующих тот факт, что с точки зрения конечного пользователя полнота поисковой выдачи (в данной задаче) оказывается важнее точности.

Работа выполнена при финансовой поддержке РФФИ (грант № 16-37-60107).

Литература

1. Московский марафон. https://ru.wikipedia.org/wiki/Московский_марафон (дата обращения: 09.02.2017).
2. Белые ночи (марафон). [https://ru.wikipedia.org/wiki/Белые_ночи_\(марафон\)](https://ru.wikipedia.org/wiki/Белые_ночи_(марафон)) (дата обращения: 09.02.2017).
3. RunRepeat.com Marathon Performance Across Nations. <http://runrepeat.com/research-marathon-performance-across-nations> (дата обращения: 09.02.2017).
4. Scheerder J., Breedveld K. Running Across Europe: The Rise and Size of One of the Largest Sport Markets. — Palgrave Macmillan, 2015. — 297 p. doi:10.1057/9781137446374
5. Roy S., Shivakumara P., Mondal P., Raghavendra R., Pal U., Lu T. A New Multi-modal Technique for Bib Number/Text Detection in Natural Images // Advances in Multimedia Information Processing: Lecture Notes in Computer Science. 2015. Vol. 9314. P. 483–494. doi:10.1007/978-3-319-24075-6_47
6. Ben-Ami I., Basha T., Avidan S. Racing Bib Numbers Recognition // British Machine Vision Conf. (BMVC). 2012. <http://people.csail.mit.edu/talidekel/papers/RBNR.pdf> (дата обращения: 09.02.2017).
7. Muhammadi J., Rabiee H.R., Hosseini A. Crowd Labeling a Survey // ArXiv. 2013. <http://arxiv.org/abs/1301.2774> (дата обращения: 09.02.2017).
8. Kara Y., Genc G., Aran O., Akarun L. Modeling Annotator Behaviors for Crowd Labeling // Neurocomputing. 2015. Vol. 160. P. 141–156. doi:10.1016/j.neucom.2014.10.082
9. Yang Y., Zhu B., Guo R., Yang L., Li S., Yu N. A Comprehensive Human Computation Framework — with Application to Image Labeling // Proc. of the 16th ACM Intern. Conf. on Multimedia Pages. 2008. P. 479–488. doi:10.1145/1459359.1459423
10. Ahn L. von, Blum M., Langford J. Telling Humans and Computers Apart Automatically // Communications of the ACM. 2004. Vol. 47. N 2. P. 57–60. doi:10.1145/966389.966390
11. Платформа краудсорсинга Amazon Mechanical Turk. <https://www.mturk.com/mturk/welcome> (дата обращения: 07.03.2017).
12. Интернет-сервис разметки фотографий пробегов Bibtaggers. <http://bibtaggers.ru> (дата обращения: 07.03.2017).
13. Интернет-сервис для хранения фотографий Яндекс.Фотки. <https://fotki.yandex.ru/> (дата обращения: 07.03.2017).
14. Levandowsky M., Winter D. Distance Between Sets // Nature. 1970. Vol. 234. P. 34–35. doi:10.1038/234034a0
15. Наборы данных для оценки качества разметки. <http://bibtaggers.ru/static/datasets.html> (дата обращения: 07.03.2017).
16. Manning C., Raghavan P., Schütze H. Introduction to Information Retrieval. — Cambridge University Press, 2008. — 506 p.
17. Mason W., Suri S. Conducting Behavioral Research on Amazon's Mechanical Turk // Behavior Research Methods. 2012. N 1 (44). P. 1–23. doi:10.3758/s13428-011-0124-6
18. Marge M., Banerjee S., Rudnicky A. I. Using the Amazon Mechanical Turk for Transcription of Spoken Language // Proc. of the 2010 IEEE Conf. on Acoustics, Speech and Signal Processing. 2010. P. 5270–5273.
19. Mason W. A., Watts D. J. Financial Incentives and the Performance of Crowds // Proc. of the ACM SIGKDD Workshop on Human Computation. 2009. P. 77–85.

UDC 004.9

doi:10.15217/issn1684-8853.2017.3.105

Tagging Public Event Images based on Non-Monetary Incentives

Ponomarev A. V.^a, PhD, Senior Researcher, ponomarev@iiias.spb.su^aSaint-Petersburg Institute for Informatics and Automation of the RAS, 39, 14 Line, V. O., 199178, Saint-Petersburg, Russian Federation

Introduction: Informatization and the spread of global networks lead to using natural cognitive abilities of people coordinated by modern information and communication technologies (crowd computing) as a way to solve a number of information processing problems whose fully automated solution at the moment does not provide the required quality of the results. **Purpose:** The goal is to analyze the possibility of using coordinated community efforts driven by non-monetary incentives in order to solve the problem of tagging running race photos with bib numbers of the visible race participants. **Results:** An internet service has been developed for collaborative tagging of images. Experiments with several hundreds of participants have shown that "internal" non-monetary incentives given by the facilitation of the search for photos is sufficient to attract race participants to take part in the tagging. The tagging quality is comparable with the quality obtained with the help of the crowdsourcing platform Amazon Mechanical Turk by using monetary incentive. **Practical relevance:** The obtained results can be used for developing various systems based on coordinated efforts of community on the basis of non-monetary incentives.

Keywords — Crowdsourcing, Crowd Computing, Image Processing, Collective Intelligence, Social Networks.

References

1. *Moskovskii marafon* [Moscow Marathon]. Available at: https://en.wikipedia.org/wiki/Moscow_Marathon (accessed 09 February 2017).
2. *Belye nochi (marafon)* [White Nights (Marathon)]. Available at: [https://ru.wikipedia.org/wiki/Белые_ночи_\(марафон\)](https://ru.wikipedia.org/wiki/Белые_ночи_(марафон)) (accessed 09 February 2017).
3. *RunRepeat.com Marathon Performance Across Nations*. Available at: <http://runrepeat.com/research-marathon-performance-across-nations> (accessed 09 February 2017).
4. Scheerder J., Breedveld K. (eds.) *Running Across Europe: The Rise and Size of One of the Largest Sport Markets*. Palgrave Macmillan, 2015. doi:10.1057/9781137446374
5. Roy S., Shivakumara P., Mondal P., Raghavendra R., Pal U., Lu T. A New Multi-modal Technique for Bib Number/Text Detection in Natural Images. *Lecture Notes in Computer Science "Advances in Multimedia Information Processing"*, 2015, vol. 9314, pp. 483–494. doi:10.1007/978-3-319-24075-6_47
6. Ben-Ami I., Basha T., Avidan S. Racing Bib Numbers Recognition. *British Machine Vision Conf. (BMVC)*, 2012. Available at: <http://people.csail.mit.edu/talidekel/papers/RBNR.pdf> (accessed 09 February 2017).
7. Muhammadi J., Rabiee H.R., Hosseini A. Crowd Labeling a Survey. *ArXiv*, 2013. Available at: <http://arxiv.org/abs/1301.2774> (accessed 09 February 2017).
8. Kara Y., Genc G., Aran O., Akarun L. Modeling Annotator Behaviors for Crowd Labeling. *Neurocomputing*, 2015, vol. 160, pp.141–156. doi:10.1016/j.neucom.2014.10.082
9. Yang Y., Zhu B., Guo R., Yang L., Li S., Yu N. A Comprehensive Human Computation Framework — with Application to Image Labeling. *Proc. of the 16th ACM Intern. Conf. on Multimedia Pages*, 2008, pp. 479–488. doi:10.1145/1459359.1459423
10. Ahn L. von, Blum M., Langford J. Telling Humans and Computers Apart Automatically. *Communications of the ACM*, 2004, vol. 47, no. 2, pp. 57–60. doi:10.1145/966389.966390
11. *Platforma kraudsorsinga Amazon Mechanical Turk* [Crowdsourcing Platform Amazon Mechanical Turk]. Available at: <https://www.mturk.com/mturk/welcome> (accessed 07 March 2017).
12. *Internet-servis razmethi fotografii probegov Bibtaggers* [Bibtaggers — Internet Service for Race Bib Number Tagging]. Available at: <http://bibtaggers.ru> (accessed 07 March 2017).
13. *Internet-servis dlia khraneniia fotografii Iandeks.Fotki* [Photo Sharing Site Yandex.Fotki]. Available at: <https://fotki.yandex.ru/> (accessed 07 March 2017).
14. Levandowsky M., Winter D. Distance between Sets. *Nature*, 1970, vol. 234, pp. 34–35. doi:10.1038/234034a0
15. *Nabory dannykh dlia otsenki kachestva razmethi* [Datasets to Evaluate Tagging Quality]. Available at: <http://bibtaggers.ru/static/datasets.html> (accessed 07 March 2017).
16. Manning C., Raghavan P., Schütze H. *Introduction to Information Retrieval*. Cambridge University Press, 2008. 506 p.
17. Mason W., Suri S. Conducting Behavioral Research on Amazon's Mechanical Turk. *Behavior Research Methods*, 2012, no. 1(44), pp. 1–23. doi:10.3758/s13428-011-0124-6
18. Marge M., Banerjee S., Rudnicky A. I. Using the Amazon Mechanical Turk for Transcription of Spoken Language. *Proc. of the 2010 IEEE Conf. on Acoustics, Speech and Signal Processing*, 2010, pp. 5270–5273.
19. Mason W. A., Watts D. J. Financial Incentives and the Performance of Crowds. *Proc. of the ACM SIGKDD Workshop on Human Computation*, 2009, pp. 77–85.