
Creative Story Writing through Crowdsourcing Empowerment

Silvana Castano

Alfio Ferrara

Stefano Montanelli

silvana.castano@unimi.it

alfio.ferrara@unimi.it

stefano.montanelli@unimi.it

Department of Computer Science

Università degli Studi di Milano

Milan, Italy

ABSTRACT

In this position paper, we envisage a possible crowdsourcing model for enforcing creative story writing in a fully crowd-assisted way. In particular, we discuss open issues regarding the capability to actively involve crowd workers in both the composition and the review/editing steps that constitute a comprehensive story-writing process. Considerations based on a preliminary experimentation in a real crowdsourcing system are also presented.

KEYWORDS

Social computing; online creative writing; crowdsourcing coordination

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

ACM CHI '19, May 04–09, 2019, Glasgow, UK

© 2019 Copyright held by the owner/author(s). Publication rights licensed to ACM.

INTRODUCTION

The crowdsourcing paradigm recently gained a lot of attention in a lot of situations where the introduction of human-supported computing (a.k.a. *social computing*) has been recognized as a valuable solution to provide an effective contribution for improving the quality of conventional tool-supported computing (e.g., resource labeling [2], item classification [6], entity linking [3]). In this context, the benefits of crowdsourcing are mostly based on the execution of the so-called *decide-question* tasks, where crowd workers are asked to choose the preferred option among a set of given candidates. Examples of human skills that are usually relevant for such a kind of crowdsourcing task are *perceptual speed*, *deductive reasoning*, and *flexibility of closure* [4]. On the opposite, it is uncommon that creative human skills like *fluency of ideas*, *problem sensitivity*, and *originality* represent crucial features of a worker profile to be involved in a crowdsourcing campaign. Moreover, also when a campaign requires to address *create-question* tasks where creativity-oriented abilities are useful, it is very hard to assess whether the self-declared worker skills are really owned by the worker and whether they are successfully employed in task resolution. Collaborative writing represents an actual application field where creative human skills can be concretely employed not only for copyediting nor proofreading [5, 9, 10]. However, the integration between the needs of collaborative writers (e.g., enforce story coherence during writing) and the process constraints of existing crowdsourcing systems (e.g., independent task assignment without storyline awareness) is not straightforward [11]. In the recent literature, some ideas are being proposed to enforce creative activities with the support of crowdsourcing contributions (see also online systems like <http://foldingstory.com/>, <http://www.crowdstories.com/>, and <http://www.crowdstory.com/>). On the one side, the focus is on how to enforce intuitive and easy-to-execute crowdsourcing microtasks [12]. On the other side, the discussion is on how to effectively involve the crowd in both writing story contributions and evaluating the generated stories [7, 8]. However, the crucial point is on how to go beyond the use of prefixed worker roles, so that the human abilities of involved crowdworkers can naturally emerge in all the executed tasks.

In this position paper, we envisage a possible crowdsourcing model for enforcing creative story writing in a fully crowd-assisted way. The model is characterized by the active involvement of crowd workers in both the composition and the review/editing steps that constitute a comprehensive story-writing process. In particular, we discuss some open issues and we present our considerations based on a preliminary experimentation in a real crowdsourcing platform.

MODELING CROWD-ASSISTED CREATIVE WRITING

Consider a text defined as a tree $\mathcal{T} = (T, h)$, where T is a set of text snippets (or story snippets) and $h(T_i) \rightarrow T_j$ is a link function which maps a snippet T_i on its parent snippet T_j (which is empty when T_i is the initial text snippet). A text may be read as an hypertext in which a snippet T_j has multiple

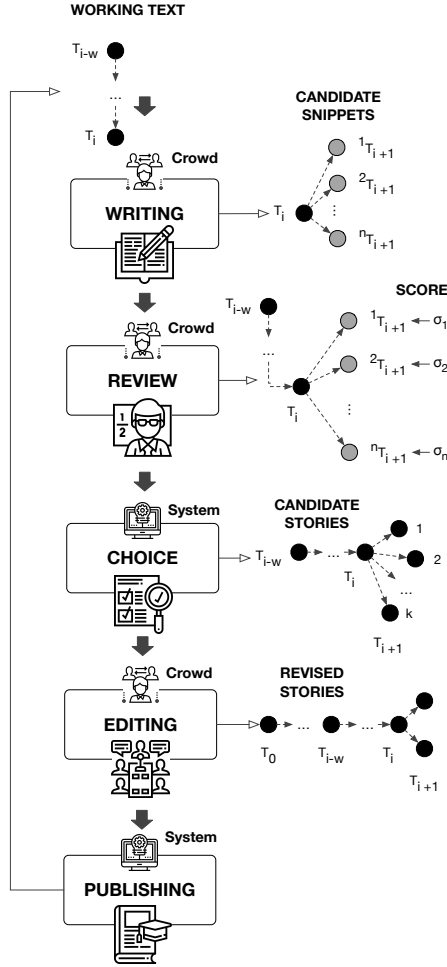


Figure 1: Our proposed model for crowd-assisted creative writing

child nodes in $\mathcal{T}$ representing possible continuations T_i . In particular, a text $\mathcal{T}$ contains a number of stories $\mathcal{S}$, each one defined as a path $\mathcal{S} = \langle T_1, T_2, \dots, T_n \rangle$, where $h(T_{i+1}) \rightarrow T_i, \forall T_i, T_{i+1} \in \mathcal{S}$, T_1 is the root node (i.e., the starting point of the story), and T_n is a leaf node (i.e., the story conclusion).

According to Figure 1, the writing process consists in the creation of $\mathcal{T}$ and it is enforced through crowdsourcing by iteratively generating a new stage $i + 1$ with k story continuations to append to the leaf nodes of the stage i . At a considered stage $i + 1$, the crowd-assisted steps are:

Writing. The crowd workers are involved to formulate new snippets, namely story continuations, through the execution of create-question tasks. In writing, a worker has a local story visibility, meaning that the task context provides a preview of story snippets created at stage $i, i - 1, \dots, i - w$. The parameter w is set to determine the tradeoff between the need to keep the task as simple as possible (low value of w) and the need to give a meaningful story preview to the worker for the sake of plot coherence (high value of w).

Review. The crowd workers are involved in the evaluation of candidate story snippets through the execution of decide-question tasks. Receiving a task, a worker has to rate a link $h(T_{i+1}^j) \rightarrow T_i$ with a score σ_j in terms of quality of T_{i+1}^j with respect to the previous T_i . According to the crowd review results, conventional crowdsourcing techniques (e.g., majority voting) are employed in the *choice* step to select the story snippets to preserve among the set of snippets T_{i+1}^j created at stage $i + 1$.

Editing. The crowd workers are involved in the evaluation of story paths. In editing, the worker has a global story visibility, meaning that the task context provides a preview of the entire story path from root to the last story continuation including a considered T_{i+1}^j snippet. The goal is to evaluate the story with a decide-question task where the crowd can decide to approve/reject a story continuation. This step differs from the review step in terms of task request and complexity. In the review step, the worker has a local visibility of the story and can rate the story snippet quality. In the editing step, the worker has a global story visibility but the choice is just approve/reject. Finally, the *publishing* step is executed to assimilate the approved story continuations and to trigger a new writing cycle.

CHALLENGES IN CROWD-ASSISTED CREATIVE WRITING

For implementation of the proposed story-writing model in a real crowdsourcing platform, the issues to be considered can be distinguished in *task-related*, *worker-related*, and *budget-related* issues.

Task-related issues. The proposed model is characterized by the use of three different kinds of tasks, namely writing (create-question), review (decide-question), and editing (decide-question) tasks. A specific issue to consider for task management in story writing is due to the need of interleaving different task typologies within a single stage. Thus, advanced scheduling algorithms are required to dynamically address *constraint-based task assignment* and to avoid possible bias such as for example the involvement of a worker in both writing and review/editing tasks related to a certain story snippet.

Table 1: Examples of consistent and inconsistent story snippets

- (1) *But now Buck did not know what to do: call for help or read the paper? Mary was getting up: "I call the neighbors". Now, one thought was in the head of Buck...*
- (2.a) **(consistent)** *As Mary left the room, Buck picked up the paper. Unfortunately, it was all stained with blood and could no longer read what was written. He approached the paper to the lamp and stood petrified in reading the only sentence visible at the bottom of the page.*
- (2.b) **(inconsistent)** *A code. A combination of numbers and letters. Someone trying to tell him something... his task was to understand what, maybe it concerned the disappearance of Mary. He concentrated and began to try to decipher the mysterious code.*

For the sake of readability, the story has been translated from Italian to English.

Worker-related issues. In the proposed model, three different worker roles can be distinguished, namely *writer*, *reviewer*, and *editor*. On this point, a possible innovative aspect to manage is concerned with the capability to recognize the worker skills as long as tasks are executed. The idea is that a worker W is encouraged to execute further writing tasks when the story snippets created by W are appreciated by the other workers. Similarly, review and editing tasks can be assigned to those workers that show the ability to choose popular and appropriate story snippets. A possible idea is to integrate *learning and association-rule mechanisms* in crowdsourcing platforms to enable the dynamic recognition of emergent worker skills.

Budget-related issues. The specification of *cost strategies* represents an open issue to exploit for enabling the configuration of a creative writing campaign so that the number of stories to generate are put in relation with the number of tasks to execute (and the budget/resources to invest).

EXPERIMENTAL EXPERIENCE AND CONSIDERATIONS

For a preliminary evaluation, we performed a real crowdsourcing experience of story writing with our Argo crowdsourcing platform [1]. In Argo, the solutions envisaged in this paper are not implemented, however the experiment allowed us to better understand the possible impacts of our ideas. In particular, we run a one-month campaign involving 332 students from the courses of Arts and Humanities of the University of Milan. The experiment produced 94 stories written in Italian, of which a sample of 32 has been evaluated by a literature expert. The focus of our experiment was to assess the overall thematic and narrative consistency of the crowd-generated stories. In Table 1, we show an example of a two-snippet excerpt, the former evaluated as consistent, and the latter as inconsistent since a snippet mentions the disappearance of a character (Mary) who is present in the previous snippet. We asked the expert to evaluate the consistency of the stories with a quantitative score k in the range from 1 to 5, where 1 denotes a totally inconsistent plot and 5 denotes a totally consistent plot. In particular, the value 3 has been used to denote the stories that are sufficiently consistent to be published for reading. The result is that about 30% of the stories are inconsistent (i.e., $k < 3$), but the 70% of them are readable (i.e., $k \geq 3$) and 20% of them are considered good (i.e., $k \geq 4$). Furthermore, we exploited the Pearson correlation coefficient for evaluating the degree of correlation between the consistency of the stories, their length, and the trustworthiness of workers, that is a measure provided by Argo to assess the worker reliability in writing, review, and editing. In the results, we observe a negative correlation (-0.872) between consistency and length and a positive correlation (0.523) between consistency and worker trustworthiness. Such a positive correlation is a very interesting result since the quality of the crowdsourcing work is conventionally measured at the level of a single task, while the consistency of stories represent a global quality evaluation of crowd work involving multiple tasks. We argue that this result depends on the capability to assign tasks to workers according to their profiles. For this reason, emergent worker profiling is our ongoing research activity on this topic.

REFERENCES

- [1] Silvana Castano, Alfio Ferrara, Lorenzo Genta, and Stefano Montanelli. 2016. Combining Crowd Consensus and User Trustworthiness for Managing Collective Tasks. *Future Generation Computer Systems* 54 (2016).
- [2] Silvana Castano, Alfio Ferrara, Lorenzo Genta, and Stefano Montanelli. 2018. Leveraging Crowd Skills and Consensus for Collaborative Web-resource Labeling. *Future Generation Computer Systems* (2018).
- [3] Gianluca Demartini, Djellel Eddine Difallah, and Philippe Cudré-Mauroux. 2012. ZenCrowd: Leveraging Probabilistic Reasoning and Crowdsourcing Techniques for Large-Scale Entity Linking. In *Proc. of the 21st WWW Conference*. ACM, Lyon, France, 469–478.
- [4] Edwin A. Fleishman. 1975. Toward a Taxonomy of Human Performance. *American Psychologist* 30, 12 (1975), 1127–1149.
- [5] Linda Flower and John R. Hayes. 1981. A Cognitive Process Theory of Writing. *College Composition and Communication* 32, 4 (1981), 365–387.
- [6] David R. Karger, Sewoong Oh, and Devavrat Shah. 2013. Efficient Crowdsourcing for Multi-class Labeling. In *Proc. of the Int. Conference on Measurement and Modeling of Computer Systems (SIGMETRICS 2013)*. ACM, Pittsburgh, PA, USA, 81–92.
- [7] Joy Kim, Justin Cheng, and Michael S. Bernstein. 2014. Ensemble: Exploring Complementary Strengths of Leaders and Crowds in Creative Collaboration. In *Proc of the 17th ACM Conference on Computer Supported Cooperative Work and Social Computing*. Baltimore, MD, USA.
- [8] Michael Nebeling, Alexandra To, Anhong Guo, Adrian A. de Freitas, Jaime Teevan, Steven Dow, and Jeffrey P. Bigham. 2016. WearWrite: Crowd-Assisted Writing from Smartwatches. In *Proc. of the 34th Conference on Human Factors in Computing Systems*. San Jose, CA, USA.
- [9] Christine M. Neuwirth, David S. Kaufer, Ravinder Chandhok, and James H. Morris. 1994. Computer Support for Distributed Collaborative Writing: Defining Parameters of Interaction. In *Proc. of the ACM Conference on Computer Supported Cooperative Work*. ACM, Chapel Hill, North Carolina, USA, 145–152.
- [10] Sylvie Noël and Jean-Marc Robert. 2004. Empirical Study on Collaborative Writing: What Do Co-authors Do, Use, and Like? *Computer Supported Cooperative Work (CSCW)* 13, 1 (2004), 63–89.
- [11] Judith S. Olson, Dakuo Wang, Gary M. Olson, and Jingwen Zhang. 2017. How People Write Together Now: Beginning the Investigation with Advanced Undergraduates in a Project Course. *ACM Trans. Comput.-Hum. Interact.* 24, 1 (2017), 4:1–4:40.
- [12] Jaime Teevan, Shamsi T. Iqbal, and Curtis von Veh. 2016. Supporting Collaborative Writing with Microtasks. In *Proc. of the 34th Conference on Human Factors in Computing Systems*. ACM, San Jose, CA, USA, 2657–2668.