Aggregating Crowd Intelligence over Open Source Information: An Inference Rule Centric Approach

Kai Kida, Hiroyoshi Ito, Masaki Matsubara, Nobutaka Suzuki, Atsuyuki Morishima

University of Tsukuba, Japan

kai.kida.2021b@mlab.info, {ito,masaki,nsuzuki,mori}@slis.tsukuba.ac.jp

Abstract

Approaches for aggregating duplicate task results from crowd workers often fail when the prerequisite for collective intelligence does not hold or when there are no results related to the task to predict the skills of workers. As an alternative approach, we explore an inference rule-centric approach where crowd workers derive inference rules towards conclusions over open source information. The result of our preliminary experiment shows its potential to derive correct answers when vote aggregation methods are ineffective.

Introduction

Aggregating crowd intelligence is one of the important topics in crowdsourcing. The most popular approach is to aggregate duplicate task results, such as majority voting and other sophisticated aggregation techniques to identify correct results, although it is well-known that such an approach often fails when the prerequisite for collective intelligence such as the diversity and independence of crowd workers does not hold. For example, as we show later, we asked crowd workers whether a photo was taken in the morning or in the evening. The majority of the workers told us that it was taken in the evening, which is incorrect.

We observe that crowd intelligence often does surprisingly great work with open source information; on SNS, the crowd sometimes identifies the picture's locations and the person's name if the topic is of great interest to the public. So we are interested in whether we can develop a systematic method that leverages the power of the crowd for any topic with general-purpose crowdsourcing platforms.

To answer the question, we explore an inference rule-centric approach because inference rules and logic are known as established means to derive conclusions from known facts. In the approach, crowd workers derive logical arguments while referencing open source information.

Figure 1 gives the overview of the framework. Given a set of conflicting hypotheses (such as (1) the photo was taken in the morning and (2) it was taken in the evening), we submit a set of tasks to ask crowd workers to complete with open information sources. The task results will eventually generate an inference tree with values representing the certainty of components. We conclude that the hypothesis is true if the value associated with the node corresponding to it is high.

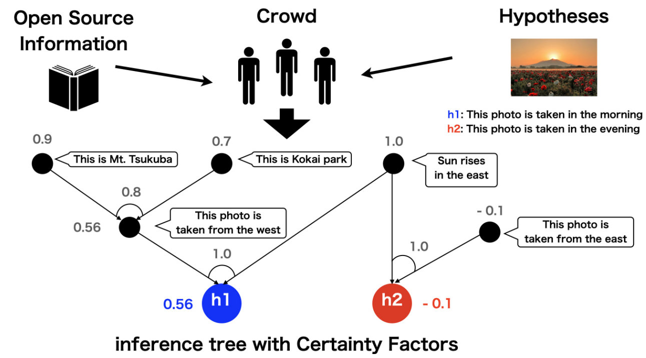

Figure 1: Our framework generates inference trees with certainty factors for given hypotheses, based on the results of different sets of microtasks. In our experiment, the framework strongly suggested that $h1$ is true, while $h2$ won in a simple majority vote. The correct hypothesis is $h1$.

This paper overviews the framework and reports the result of our preliminary experiments with a workflow we developed as the first step. The result shows that this approach is promising and can exploit the crowd's power of deriving a variety of ways to reach conclusions.

Related Work. There are a lot of works that leverage the power of the crowd to collect evidence for claims in a variety of contexts such as fake news (Tschiatschek et al. 2017), the location of photos (Popoola et al. 2013)(Venkatagiri et al. 2019), and general claims (Wijerathna et al. 2018). There is also a line of works that involve crowd to evaluate given inference rules (Zeichner, Berant, and Dagan 2012). In contrast, our focus is on a structured framework for deriving inference rules to make human-in-the-loop decisions with the help of algorithms. Therefore our method can be combined with many existing methods by adopting them as components of our framework. Finding the best combinations of such components will be one of our future works.

EMV Workflow

We devised an iterative workflow, named the EMV (Expansion-Merge-Verification) workflow, in which each iteration consists of Expansion, Merge, and Verification

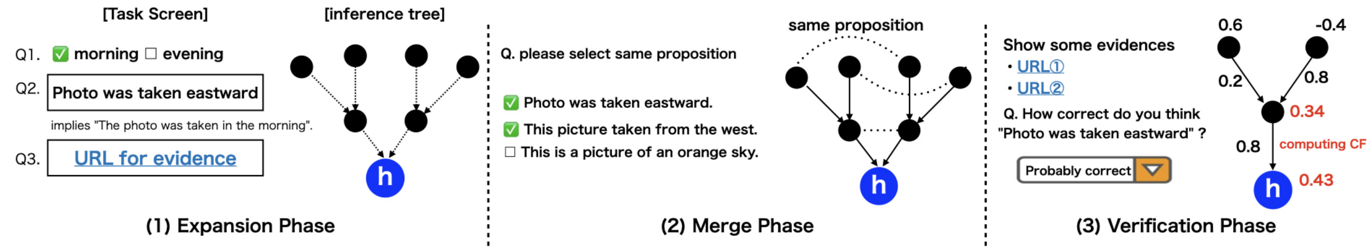

Figure 2: One iteration of Expansion, Merge, and Verification to derive the inference tree. In the expansion phase, we add logical expressions that reach the conclusion that some of the hypotheses are correct. The merge phase merges equivalent propositions. The verification phase gives confidence scores to each proposition and derivation.

phases (Figure 2). The three-phase iteration recursively expands the tree and ends when the *certainty factor* attached to every leaf node (proposition) is highly positive or negative. Here, we adopt Mycin's certainty factors (Buchanan and Shortliffe 1984) because it is a simple method that can deal with multiple evidences to derive the same conclusion.

Phase 1: Expansion. We submit tasks to collect logical expressions that reach the conclusion that some of the hypotheses are correct and add the expressions to the current inference tree (in the beginning, it has nodes that correspond to the given hypothesis only). In the task, we first ask the crowd worker which hypothesis H_i she thinks is true and then enter antecedents of the rule $X_1, X_2, X_3 \rightarrow H_i$, with the evidences to support X_j. For example, if H_i is "the photo is taken in the evening", a worker answers "There is a photo on the web that has the same shape as the mountain, with the description that this photo was taken from the east" as X_1, and gives the URI of the page as its evidence. The antecedent of the rule is added as a node to the inference tree and connected to the conclusion (hypothesis) by a direct edge.

Phase 2: Merge. We submit tasks to merge equivalent propositions into one, to avoid duplicate propositions in the inference tree. For each rule R_i we obtained in the expansion phase, we generate a task for the marge phase as follows. We choose $R_j (j \neq i)$ that shares at least one term to be included in the candidate list to be merged with R_i. Then, the task asks workers whether there are rules that have the same meaning with R_i.

Phase 3: Verification. For each rule $R : X \rightarrow A$, we submit tasks to ask the crowd give confidence scores to each antecedent X and rule R, which will be used to compute the certainty factors we explain below. We generate a task for each leaf node and edge in the inference tree after Phase 2.

Inference Rules and Certainty Factors

We adopt *certainty factors* introduced by Mycin, as a means for evidence combination. The certainty factor is a measure of an expert's belief in a fact or a rule. Given a rule $R : X \Rightarrow Y$, $CF(R)$ denotes a certainty factor that represents the strength of the belief that $X \Rightarrow Y$ holds. Likewise, we use $CF(X)$ to denote a certainty factor that represents the strength of the belief that X holds. The certainty factor ranges from -1 (definitely false) to 1 (definitely true).

We can derive the certainty by the following process: (1) calculating the certainty for the antecedent X and the rules R, (2) calculating the certainty for the consequent Y, and (3) propagating the certainty of them.

For a single antecedent rule "$X \Rightarrow Y$", $CF(Y, \{X\})$ is defined as follows:

$$CF(Y, \{X\}) = CF(X) * CF(R) \tag{1}$$

For a conjunctive multiple antecedent rule "$X_1 \wedge X_2 ... \Rightarrow Y$", the certainty factor of the consequent is defined as:

$$CF(Y, \{X_1, X_2, \ldots\}) = CF(min(X_i)) * CF(R) \tag{2}$$

If we have more than one rule that has the same consequent, such as $X_1 \Rightarrow Y_1$ and $X_2 \Rightarrow Y_1$, the certainty factor is defined as:

$$CF_{COMB}(CF_1, CF_2) = \begin{cases} CF_1 + CF_2 * (1 - CF_1) (3.1) \\ \dfrac{CF_1 + CF_2}{1 - min(|CF_1|, |CF_2|)} (3.2) \\ CF_1 + CF_2 * (1 + CF_1) (3.3) \end{cases}$$

Preliminary Experiment

We applied the workflow to the question shown in Figure 1 on Lancers and Yahoo! Crowdsourcing. We paid 100 JPY (for each Phase 1 and 3 tasks) and 50 JPY (for each Phase 2 task) to workers. In Phase 1, 20 crowd workers on Lancers provided 17 (3) rules with 26 (4) antecedents for h_1 (h_2). Examples include "X_1: The photo shows fog." and "X_2 : Fog tends to occur in the morning." "$R : X_1 \wedge X_2 \Rightarrow h_1$". In Phase 2, we generated 30 tasks and 40 crowd workers on Yahoo merged 23 antecedents with others, resulting in 7 antecedents and 6 rules. In Phase 3, we asked 10 crowd workers on Yahoo to do the tasks. We obtained $CF(h_1) = 0.589$ and $CF(h_2) = -0.002$, which means that h_1 is probably correct. We also asked 160 crowd workers to complete a simple voting task independently. 83 workers voted for h_2.

Summary and Future Work

This paper overviewed an inference-rule-centric approach for aggregating crowd intelligence over open source information sources and showed a preliminary experimental result. We plan to explore the approach to identify better workflows and limitations of this approach.

Acknowledgment. This work was partially supported by JSPS KAKENHI Grant Number 22H00508 and 21H03552. This work was approved by the IRB of University of Tsukuba.

References

Buchanan, B. G.; and Shortliffe, E. H. 1984. *Rule-based expert systems : the MYCIN experiments of the Stanford Heuristic Programming Project / edited by Bruce G. Buchanan, Edward H. Shortliffe.* The Addison-Wesley series in artificial intelligence. Reading, Mass: Addison-Wesley. ISBN 0201101726.

Popoola, A.; Krasnoshtan, D.; Toth, A.-P.; Naroditskiy, V.; Castillo, C.; Meier, P.; and Rahwan, I. 2013. Information Verification during Natural Disasters. In *Proceedings of the 22nd International Conference on World Wide Web*, WWW '13 Companion, 1029–1032. New York, NY, USA: Association for Computing Machinery. ISBN 9781450320382.

Tschiatschek, S.; Singla, A.; Gomez Rodriguez, M.; Merchant, A.; and Krause, A. 2017. Detecting Fake News in Social Networks via Crowdsourcing.

Venkatagiri, S.; Thebault-Spieker, J.; Kohler, R.; Purviance, J.; Mansur, R. S.; and Luther, K. 2019. GroundTruth: Augmenting Expert Image Geolocation with Crowdsourcing and Shared Representations. *Proc. ACM Hum.-Comput. Interact.*, 3(CSCW).

Wijerathna, N.; Matsubara, M.; and Morishima, A. 2018. Finding Evidences by Crowdsourcing. In *2018 IEEE International Conference on Big Data (Big Data)*, 3560–3563.

Zeichner, N.; Berant, J.; and Dagan, I. 2012. Crowdsourcing Inference-Rule Evaluation. In *The 50th Annual Meeting of the Association for Computational Linguistics, Proceedings of the Conference, July 8-14, 2012, Jeju Island, Korea - Volume 2: Short Papers*, 156–160. The Association for Computer Linguistics.