

Crowdsourcing and Its Applications in Scientific Research

Sheng-Wei CHEN

Institute of Information Science, Academia Sinica, Taiwan

swc@iis.sinica.edu.tw

Crowdsourcing is a term that combines “Crowd” and “Outsourcing”. It is a distributed model that assigns tasks traditionally undertaken by employees or contractors to an undefined crowd. It achieves the goal of mass collaboration via Web 2.0 technologies. The main difference between crowdsourcing and ordinary outsourcing is that a task is carried out by an unspecific Internet crowd rather than a specific group of people.

Crowdsourcing services, such as Amazon Mechanical Turk (MTurk), extend the interactivity of crowdsourcing tasks by more comprehensive user interfaces and micro-payment mechanisms. MTurk is a popular crowdsourcing service that provides a marketplace for a variety of tasks, and anyone who wishes to seek help from the Internet crowd can post their tasks on the website. Tasks can involve any kind of effort, such as participating in surveys, performing experiments, or answering certain specialized questions. Researchers have adopted MTurk to conduct user studies on image annotation, document relevance, and document evaluation. Since an Internet crowd can be quite large, crowdsourcing enables researchers to conduct experiments with a more diverse set of participants at a lower economic cost than would be possible under laboratory conditions. However, because participants carry out experiments without supervision, they may give erroneous feedback perfunctorily, carelessly, or dishonestly, even if they receive a reward for each experiment.

In this talk, I will first introduce crowdsourcing and its applications in various areas. Then, I will present several examples how researchers apply crowdsourcing to their research studies. Finally, I will discuss the trustworthiness issue in crowdsourcing and how it can be resolved with a paired-comparison-based technique.