Labor in the cloud: Designing incentives for sharing performance data

Panos Ipeirotis

New York University

Going beyond the provision of "traditional" computing services through the cloud, there is another cloud service that is changing everyday life: the emergence of crowdsourcing (or "labor as a service") as a way to get immediate access to human intelligence, from within a computer system. This idea has changed the way that computer systems are designed and implemented. While this immediate access to human intelligence allows for innovative applications, at the same time we see questions that are not typically encountered when designing purely computational systems.

Separation of task execution and recruiting

For example, a crucial question is how should we abstract and separate the different functionalities provided by crowdsourcing platforms? What are the basic services that a platform should provide? Today, we see a separation of the recruitment functionality from the user-interface and task-handling functionality: Most services abstract away from the "labor channel" and then build their own "application logic" on top. The major example of a recruiting service is Mechanical Turk, and other companies (oDesk, Samasource, etc) are also providing similar labor channels. On the other hand, task handling is done by the employer under a separate interface, almost always ignoring any task handling functionality provided by the labor platform.

The tragedy of the commons in crowdsourcing systems

While this is a natural separation of specialties and focus, this separation of labor channel from task execution also creates the following problem: Each task-handling participant evaluates the recruited employers and knows their performance. However, there is no incentive to share back this information with the labor channel. Providing feedback is effectively a public good. Feedback benefits others but not the one who provides it. Combining this public good properties with a competitive environment, where each service is competing for access to the best workers, creates an environment where effectively nobody has the incentives to share private knowledge about the worker performance. Who wants to tell competitors who are the most trusted and reliable employees?

This is a highly suboptimal solution. First, when there is no public information about the performance of each worker, all employers need to devise their own tests and measurements, and learn by trial-and-error who the workers are that provide high quality services. To make the parallel with computing services, we would have to run benchmarks on every cloud service before even starting executing anything of interest. Combining that with the fact that workers have limited capacity, it is understandable why an employer does not want to share this information with others. Second, even two competitors may end up being better off by sharing information: If two employers have information about, say, 50% of the workforce, they could share information with each other and have information about all workers, saving each other the cost of testing.

Unfortunately, sharing such valuable information generates a prisoner's dilemma situation. While sharing is a better solution for both parties, it is even better for someone to back off and wait for others to share. Or even worse, a malicious employer may give incorrect information to others, in order to feed false information to competitors and lead them to hire incompetent or malicious workers.

While this setting is currently limited to crowdsourcing services, we can see a similar problem emerging in distributed cloud systems, where resources are contributed by a variety of participants and are not controlled strictly by a single provider (Amazon, Google, etc.)

The research challenge

What are the structures that can encourage and incentivize truthful revealing of reputation by the employers who have evaluated workers? Clearly, the worker that provides a public reputation feedback should expect something in return, and not just see others free-ride on this information.

A potential approach is to get employers to post feedback in an "escrow" system which others can probe/search but not browse. In other words, you need the identifying information about the worker whose profile is requested, you cannot just query for "give me the best workers". This ensures that someone can get information only for workers for whom there was some interaction, not for the global pool of workers. Incentive-wise, this means that you need to be participating in the system to get access to performance data. To ensure that there are no abuses of the probing privilege, a tit-for-tat mechanism (e.g., 10 queries, after posting information for a single worker) can mitigate such concerns.

Of course, this does not ensure that the employers will be truthful. In fact, the incentives are to post incorrect information, in order to fool the competitors and lead them to hire incorrect workers. One potential avenue towards fixing this problem is to examine the level of agreement when evaluating a single worker: If an employer provides the same feedback as other good employers, then the information should be trusted. If not, then the information is not reliable and the employer should not be rewarded with access to the pool of information about the workers.

We can augment this simple model to account for factors such as worker reliability (reliable workers get same scores from everyone, unreliable are difficult to label correct), worker focus (scores depend on area of focus), time, and other factors: We can simply leverage all the recent work on managing noisy workers in crowdsourced environments. This can generate a reputation system in which there is a "tit for tat" system of contribution and employers actually benefit by contributing to this common pool of semi-public feedback.

Alternative approaches can also be explored, trying to answer a broader question: How can we share valuable information only with others that contribute back valuable information?