Statistical Analysis of Demand and Supply in a Distributed Problem Solving Platform of Software Development

Ta-Hsin Li
IBM Research, Yorktown Heights, NY 10598-2018, USA - thl@us.ibm.com

Yi-Min Chee
IBM Research, Yorktown Heights, NY 10598-2018, USA - ymchee@us.ibm.com

David Hoffman
IBM Global Business Services, Research Triangle Park, NC 27709-2195, USA - dah@us.ibm.com

Blain Dillard
IBM Software Group - Watson, Research Triangle Park, NC 27709-2195, USA - blaind@us.ibm.com

Abstract

An emerging business model in application software development in large enterprises is to employ a flexible workforce, or a resource pool, which consists of vetted freelancers, to support the application development process, including software design, coding, and application testing. This so-called crowdsourcing model is facilitated by a service platform where work items are posted with detailed requirements and where proposals are solicited from members of the resource pool. After evaluating the proposals and previous work history of the participants, the best candidate is selected to perform the task. The success of this model depends crucially on having the right participants at the right time when their skills are needed. However, the need for each set of skills fluctuates over time, depending on the software development activities of the business; the number of participants also fluctuates because participation is entirely voluntary and performed via self-selection of work. Therefore, maintaining the appropriate capacity, or supply, of the resource pool is an important and challenging problem for the service provider who utilizes this type of delivery platform. Undersupply of talent can impact project deliveries and cause work to go unstaffed, and oversupply reduces the effectiveness and commitment of the participants, causing them to lose interest when the work is not plentiful enough. In this paper, we present some results of a statistical analysis of the demand and supply in a resource pool operation in IBM's Global Business Services business unit. The analysis enables a data-driven strategy for capacity planning and management.

Keywords: capacity planning; freelancing; staffing; crowdsourcing; survival analysis; logistic regression.

1. Introduction

In large enterprises such as IBM, software development for business applications is a complicated undertaking that requires a workforce with diverse skill sets dictated by the ever-increasing complexity of technological requirements (e.g., languages, tools, platforms, APIs, frameworks, databases, etc.).The conventional delivery model is entirely project-based, where a team with required skills is formed and dedicated to a given project until its completion. The project-based delivery model is not very cost-efficient because not all skills are needed all the time during the project's life cycle. The resource-pool-based delivery model, also known as crowdsourcing (Howe, 2008; Vuković, 2009; Ranade & Varshney, 2012), is designed to improve the efficiency. A resource pool consists of vetted freelancers who are contracted and paid to perform software development tasks only when needed. To better utilize the resources, work items should be self-contained and short-cycle; following the principles of good project management, work items in a work breakdown structure should be about one week or less. Participating projects must break down the entire work effort into small pieces, or components, that can be developed and tested independently in short time by the resource pool and then integrated by the project team to produce the final product. This leads to the distributed problem solving

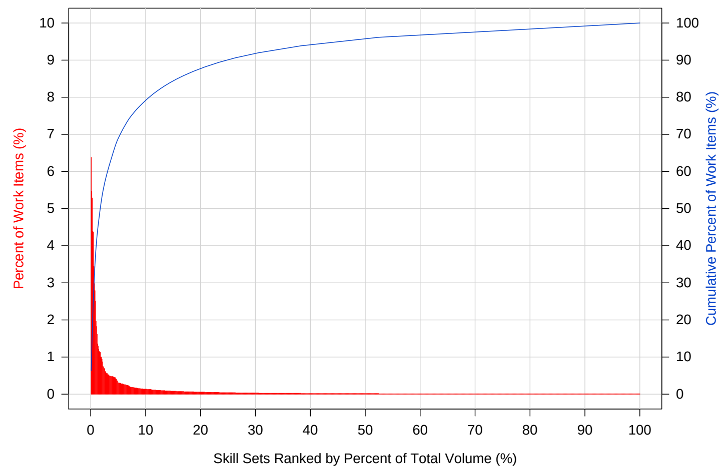


Figure 1. Fragmented distribution of work items over different sets of skills.

business model for software development. This business model needs to be supported by a shared service that maintains the resource pool and provides a platform (e.g. web portal) to facilitate the preparation of work items, the solicitation of proposals from the resource pool members, the selection of a winner, and the submission and review of the finished work. See Peng, Ali Babar & Ebert (2014) for a recent survey of crowdsourcing platforms.

By subscribing to the resource pool service platform, the projects can tap into the talent in the resource pool when needed without having to maintain an underutilized workforce of their own. The platform provider must manage the capacity of the resource pool carefully so that the rights skills are available at the right time when they are needed. Although the economies of scale make the capacity planning easier, it does not eliminate uncertainty in the demand or the supply of skills. The uncertainty on the demand side comes from the volatility of project activities, although it can be reduced considerably as compared to that of a single project team, due to the aggregation effect of the resource pool model. The uncertainty on the supply side depends on the behavior of individual members of the resource pool and their personal preferences for choosing work to participate in. Understanding these uncertainties is a crucial step toward the development of an effective strategy for capacity planning and management of resource pool services.

In this paper, we present some results of a statistical analysis of the demand and supply in a resource pool operation at IBM Global Business Services. On the demand side, we show the phenomenon of demand fragmentation by skill sets. i.e., sparse demand for a large number of skill combinations. We also show the erroneous projections made by the project teams for the timing of work items and their idiosyncratic characteristics that occur across different project managers. On the supply side, we show the diverse behavior of participation for the resource pool members in bidding for eligible work items. We also show the effect of tenure in the resource pool on the participation rate of resource pool members. All these, and other findings from the analysis, provide a deeper understanding of the demand and supply dynamics of the resource pool that could be utilized to design an effective strategy of demand supply modeling and forecasting for capacity planning and management.

2. Analysis of Demand

Work demand is categorized by the sets of required skills in each work item needed by project teams and determined based on project plans. Due to the complexity of business applications and their technological requirements, the number of such skill sets can be very large. Some of these sets are commonly used by the participating projects, and some other sets are required only occasionally. While the common skill sets naturally deserve a lot of attention, the rare skill sets cannot be totally ignored because their numbers can also scale to a significant volume.

Figure 1 depicts the percentage of work items for different sets of skills. It shows that while the largest 10%

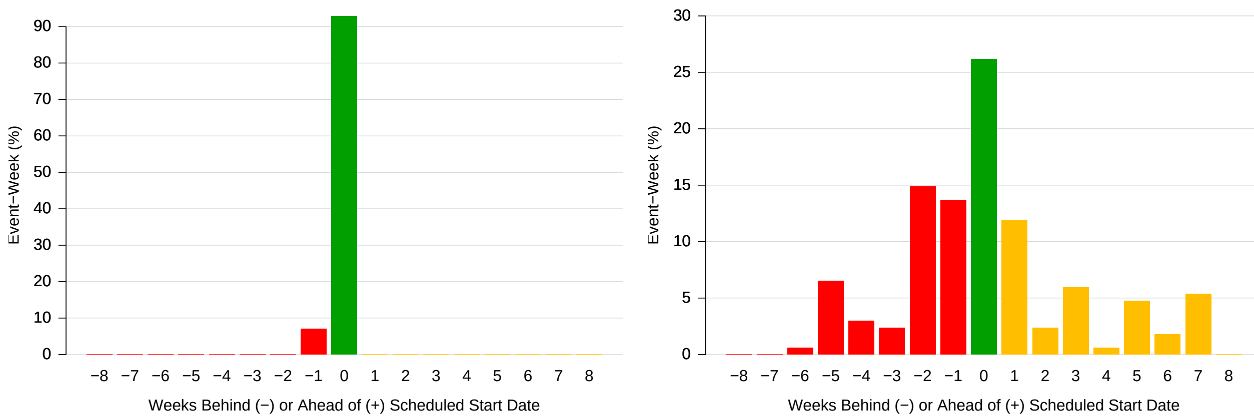


Figure 2. Distribution of discrepancy in scheduled start date from two project managers.

of the sets of skills counts for 80% of the total work items, the remaining 20% of work items spread out thinly over the other 90% of the sets of skills. The fragmented demand makes statistical modeling and forecasting very difficult at the granularity of skill sets. It calls for a judiciously designed aggregation scheme in which fragmented sets of skills are combined to form larger classes without fundamentally changing the requirements of the work items affected by the aggregation. A rare set of skills tends to contain a large number of skill components, and many of them may not be critical to the successful execution of the work item. Rooting out these unnecessary components offers an opportunity to achieve the objective of aggregation.

In a typical resource pool operation, all work items must be prepared in the system. During the preparation stage, the project manager enters the description and requirement of the work item as well as the start date of the work into the service platform. The specified timing of these work items, together with their technological requirements, can be utilized to predict future demand. However, the specified timing, called scheduled start date, is not a reliable predictor of the actual launch time because any work item can be rescheduled to a later date or an earlier date by the project manager at any time before the work item is actually launched. As a result, the predicted future demand based on the scheduled start date can be erroneous.

The discrepancy between the scheduled start date and the actual launch time varies considerably, depending on the behavior of project managers. Figure 2 shows the distribution of discrepancy in scheduled start date for work items prepared by two different project managers. While one of the project managers is able to predict the actual launch time very well, the other poorly predicts in nearly 75% of the event weeks (an event week is defined as a work item spending one week in the preparation stage). The actual launch time can be later than the scheduled start date (negative discrepancy) or early than the scheduled start date (positive discrepancy). In the second example, the discrepancy ranges from -6 weeks to $+7$ weeks. To predict the demand accurately, one must take into account the erroneous nature of the scheduled start dates and differential the idiosyncratic behavior of project managers.

Table 1. Survival analysis for the lifetime of work items by proportional hazard regression

Class ID	Coefficient	Standard Error	Z-Score	p-Value
170	-0.1516	0.185	-0.821	0.41000
229	0.1263	1.135	0.636	0.52000
446	0.4238	0.187	2.272	0.02300
567	0.7618	0.208	3.662	0.00025
607	0.6744	0.204	3.305	0.00095

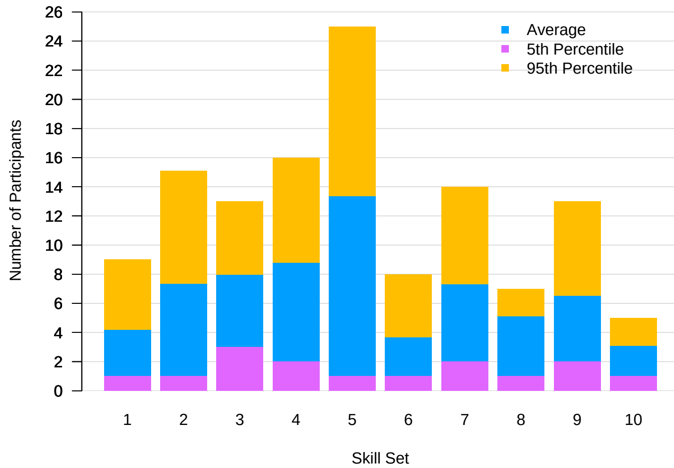


Figure 3. Number of participants by skill sets.

The lifetime of work items in the preparation stage offers another way to characterize the behavior of project managers. Some project managers tend to prepare their work items way ahead of the launch time and some others tend to do it on short notice. Table 1 shows the result of a survival analysis for the lifetime of work items in the preparation stage using Cox's proportional hazard model (Cox, 1972; Cox and Oakes, 1984). It contains the estimated regression coefficients in the model together with the corresponding standard errors, Z-scores, and p-values for five selected classes of work items. The work item classes are defined jointly by the sets of skills needed and the project managers. As can be seen, the first two classes do not differ significantly from the baseline behavior, whereas the last two classes exhibit considerably different characteristics as indicated by the very small p-values ($\ll 0.01$). The positivity of the coefficients for the last two classes implies that the work items in these classes tend to have a longer-than-average lifetime in the preparation stage. In other words, the corresponding project managers tend to prepare their work items earlier than average, thus offering a longer-than-average lead time for demand outlook.

3. Analysis of Supply

Because participation of resource pool events is entirely voluntary and self-selective, it is important for the service provider to ensure a sufficient number of participants for each work item. The level of participation is a key indicator of success in resource pool services. Suitable measures, such as offering better incentives (Mason & Watts, 2009) or recruiting more engaged resources with at-risk skill sets (Satzger, Psaier, Schall & Dustdar, 2013), have to be taken to boost the participation level when it is too low.

Figure 3 depicts the average number of participants for ten selected sets of skills together with the 5th and 95th percentiles. It can be seen that the number of participants vary considerably over different sets of skills and across different work items that require the same set of skills. Therefore, it is important for work items with low participation to get quality participants in order to minimize the risk of failures.

The level of participation for work items that require a given set of skills depends on the individual behaviors of eligible members of the resource pool as well as the skills and capabilities that they each possess. It is possible that among a large number eligible members only a few participate with high probability. Figure 4 shows the estimated participation probability of eligible members for work items that require a particular set of skills. In this example, there are 12 eligible members who have the required skills, but only 7 of them can be regarded as likely participants because the others never or rarely participate. Among the 7 likely participants, the estimated probability of participation varies from 0.50 to 0.85. These personalized participation characteristics must be taken into account when evaluating the supply conditions and risks of the resource pool.

The tenure of a member in the resource pool could also play a role in characterizing the participation level.

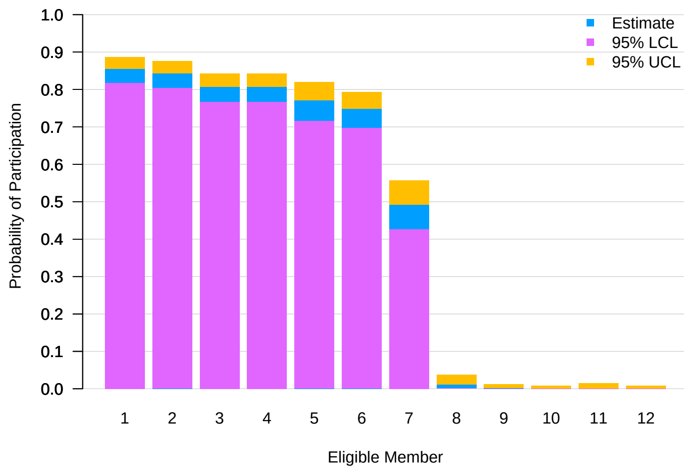
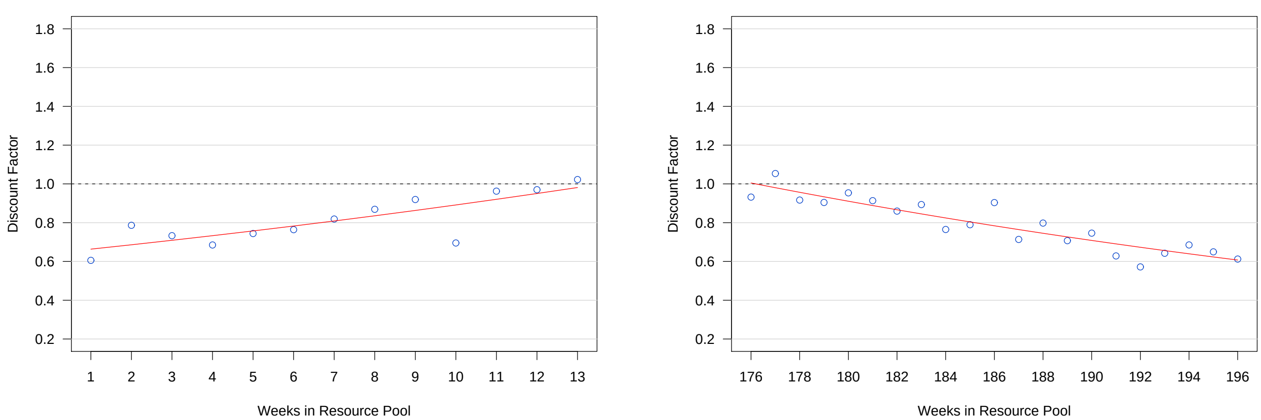


Figure 4. Probability of participation by eligible members.

In particular, there could be what we call the effect of learning for resource pool members: the probability of participation tends to be lower than the average for members who just joined the resource pool recently. The discount factor diminishes as the tenure of these members grow until a point is reached where it is no longer necessary. The effect of learning can be explained by the fact that new members need sufficient time to get acquainted with the system and expected participation patterns before they can fully participate in the resource pool activities. Similarly, there could be what we call the effect of experience where the participation level begins to decrease with the tenure when it is sufficiently long. This can be explained by the fact that experienced members become more selective in choosing the resource pool events to participate.

Figure 5 depicts the observed discount factor as a function of the tenure in the resource pool together with the logistic regression fit under two scenarios. The effect of learning and the effect of experience are clearly observed in Figure 5(a) and Figure 5(b), respectively. In this example, the participation probability for members in their first week of tenure is only 60% of the average participation probability. As the tenure increases, so does the participation probability, until 13 weeks in the pool when the participation probability reaches the average value. For members with a tenure longer than 176 weeks, the participation level begins to deteriorate consistently and drops back to 60% for those with 196 weeks of tenure.



(a) (b)

Figure 5. Discount factor of participation probability for members with different tenures in the resource pool. (a) The effect of learning. (b) The effect of experience.

Table 2. Logistic regression analysis of the participation probability

Effect	Covariate	Coefficient	Standard Error	Z-Score	p-Value
Learning	Intercept	-2.391369	0.027536	-86.84	$<$ 2e-16
	Tenure (1-13 Weeks)	0.036486	0.003413	10.69	$<$ 2e-16
Experience	Intercept	3.050933	0.773600	3.944	8.02e-05
	Tenure (176-196 Weeks)	-0.028073	0.004242	-6.618	3.64e-11

A formal statistical analysis for the impact of tenure on the participation probability is shown in Table 2 with the help of logistic regression (Hosmer, Lemeshow & Sturdivant, 2013). The table contains the estimated coefficients in the logistic regression model and their p-values in both early period of tenure and the late period of tenure. As we can see, the tenure variable is statistically significant in both cases, although to a lesser degree in the later case due to a smaller sample size.

4. Conclusions

In this paper, we have presented some statistical analysis results for the demand and supply of software development skills in a resource pool service. The analysis reveals a number of characteristics that are potentially beneficial to the design of capacity planning and management strategies. In particular, the analysis shows that the scheduled start date of work items in the preparation stage can be erroneous as a predictor for the timing of these work items and should be taken into account when it is used to construct demand outlook. The analysis also shows that the participation level of the resource pool is highly personal and can vary with the tenure at different stages of service. All these findings deserve serious consideration in demand and supply forecasting to ensure the success of resource pool services.

References

Cox, D. R (1972). "Regression models and life-tables," *Journal of the Royal Statistical Society, Series B*, 34 (2), 187–220.

Cox, D. R., & Oakes, D. (1984). *Analysis of Survival Data*. New York, Chapman & Hall.

Hosmer, D., Lemeshow, S., & Sturdivant, R. (2013). *Applied Logistic Regression*. 3rd Edition. Hoboken, New Jersey. Wiley & Sons.

Howe, J. (2008). *Crowdsourcing: Why the Power of the Crowd is Driving the Future of Business*. New York. Three Rivers Press.

Mason, W., & Watts, D. J. (2009), "Financial incentives and the 'performance of crowds'," *Proceedings of the 1st Human Computation Workshop (HCOMP 2009)*, 77–85.

Peng, X., Ali Babar, M., & Ebert, C. (2014). "Collaborative software development platforms for crowdsourcing," *IEEE Software*, 31(2), 30–36.

Ranade, G. V., & Varshney, L. R. (2012). "To crowdsource or not to crowdsource?" *Proceedings of the 4th Human Computation Workshop (HCOMP 2012)*, 150–156.

Satzger, B., Psaier, H., Schall, D., & Dustdar, S. (2013). "Auction-based crowdsourcing supporting skill management" *Information Systems*, 38, 347–560.

Varshney, L. R. (2012). "Participation in crowd systems," *Proceedings of the 50th Annual Allerton Conference on Communication, Control, and Computing*, 996–1001.

Vuković, M. (2009). "Crowdsourcing for enterprises," *Proceedings of 2009 World Conference on Services - I*, 686–692.