CREDIBILITY-BASED BINARY FEEDBACK MODEL FOR GRID RESOURCE PLANNING

Parasak Chokesatean and Martin B.H. Weiss

School of Information Sciences, University of Pittsburgh, Pittsburgh, PA 15260
{pac3@pitt.edu and mbw.pitt.edu}

Abstract

In commercial grids, Grid Service Providers (GSPs) can improve their profitability by maintaining the lowest possible amount of resources to meet client demand. Their goal is to maximize profits by optimizing resource planning. In order to achieve this goal, they require an estimate of the demand for their service, but collecting demand data is costly and difficult. In this paper we develop an approach to building a proxy for demand, which we call a *value profile*. To construct a value profile, we use binary feedback from a collection of heterogeneous clients. We show that this can be used as a proxy for a demand function that represents a client's willingness-to-pay for grid resources. As with all binary feedback systems, clients may require incentives to provide feedback and deterrents to selfish behavior, such as misrepresenting their true preferences to obtain superior services at lower costs. We use credibility mechanisms to detect untruthful feedback and penalize insincere or biased clients. Finally, we use game theory to study how cooperation can emerge in this community of clients and GSPs.

Keywords

Binary feedback, credibility mechanism, cooperation, value profile, and Grid Service Providers, resource planning.

1. Introduction

Grid computing systems are an ecosystem consisting of Grid Service Providers (GSPs) and clients. GSPs are agents who control resources while clients are users of those resources. Generally, resources refer to CPU cycles, memory space, disk space, and network bandwidth. Fig. 1 illustrates a typical commercial grid environment[1] (as opposed to grids based on voluntary resources) [1]. Many studies have modeled grid resource brokers, as summarized in Buyya [2]. In this paper, we only look at the interaction between clients and GSPs and ignore the role of brokers in order to simplify our analysis. We assume that a client sends a job based on his/her constraints (such as budget and preferred duration) directly to a GSP, who promises to provide a service to the client according to a service level agreement (SLA). After the job is completed, the GSP returns the result to the client along with the service charge.

---

[1] A grid resource broker is an agent who finds GSPs for clients while a grid market directory contains a list of all GSPs within a market.

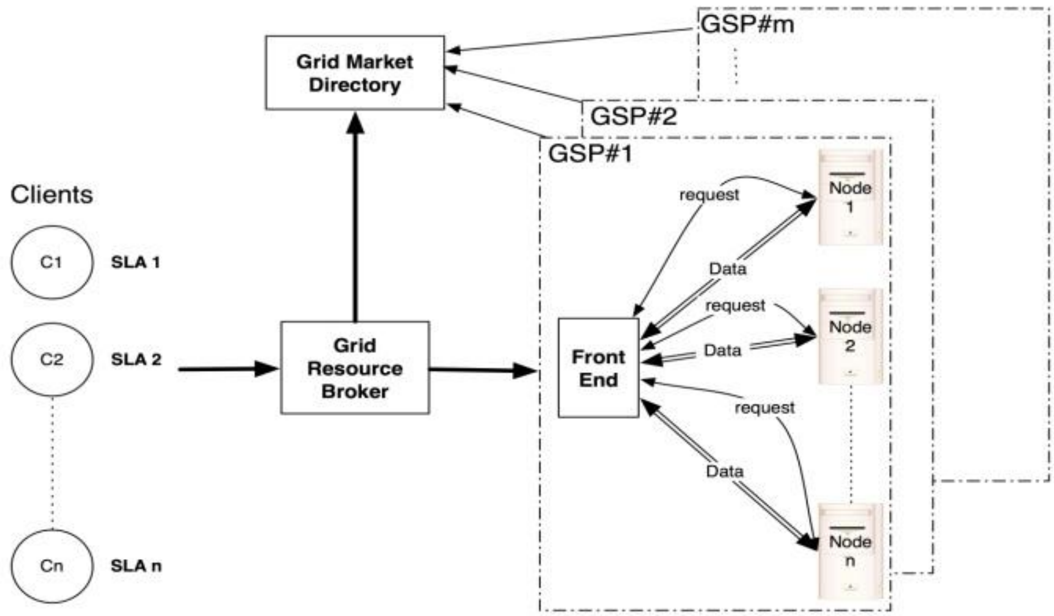


Figure 1 -- Typical Commercial Grid environment

GSPs can increase their profits by either increasing revenues or reducing costs. Increased revenues can be achieved through pricing strategies or by upgrading current services. We presume that an upgraded service attracts both new and current clients, which can boost revenues. This case usually occurs in busy GSPs that have high resource utilization. However, GSPs have to be careful not to overprovision, since resources might be idle for long periods of time. This raises the question of how many extra resources GSPs should acquire.

Alternatively, GSPs improve their profitability by maintaining the lowest amount of resources to meet client demand. In this scenario, a GSP might decide to downgrade a service by reducing their computing resources. This case commonly happens in GSPs that have low resource utilization. The challenge is finding the lowest resource level that is required to satisfy clients. If GSPs downgrade their quality of service (QoS) too much, they might end up losing their existing clients.

Since QoS is an essential attribute of demand and cannot be measured until after it is consumed, it is difficult for GSPs to estimate client demand. To improve this situation, feedback models, which work as signaling tools, may be used [3]. With feedback models, clients are asked to rate the received services based on their level of satisfaction. Since clients have to use resources to provide feedback, they might not do so if they cannot clearly see a benefit in their contribution. Furthermore, clients have incentives to provide untruthful feedback[2] since they (rationally) want to consume as many resources as possible at the lowest price possible. That is, they have an incentive to lie in order to receive better service even though they may already be satisfied with the current service. This leads to questions on how useful client feedback is and how truthful client feedback can be obtained.

Most previous work concentrates on market-based resource allocation without the use of service value to grid clients [2,4,5,6,7,8,9,10]. The main theme of this paper is to develop an approach to build a *value profile* using feedback models for a group of heterogeneous grid clients that GSPs can use as a proxy for demand to plan their resources economically. The goals of this paper are to study how client demand can be estimated for services and how cooperation can emerge. By understanding the conditions that allow it to emerge, it may be possible to suggest the development of cooperation in a particular set of conditions and provide the development of a credibility-based binary feedback model for grid resource planning.

This paper is organized as follows. In Section 2, we review theoretical frameworks used in this paper and we describe an approach to building a value profile. Sections 3-4 discuss clients' incentives to cooperate. Section 5 presents simulation results and discussions. Finally, Section 6 offers our conclusions.

[2] Untruthful feedback is a term used to describe a feedback received from insincere or biased clients. Insincere clients are clients who have an incentive to cheat. Biased clients are clients who have an incentive to provide an unfair rating.

2. Value Profiles

In this paper, a *value profile* is a proxy for a demand function that represents a client's willingness-to-pay for grid resources. To construct a value profile that is useful to GSPs, we use binary feedback, which is a form of on-line reputation. We examine the prior research on binary feedback in this section and demonstrate a way in which binary feedback can be used to construct a value profile.

2.1 Online-Reputation

Online reputation mechanisms, also known as feedback systems, have emerged as a significant quality signal and control mechanisms in private e-markets such as eBay.com and Amazon.com [3,11]. The objective of reputation mechanisms is to encourage trust and cooperation in online trading communities. Reputation systems are designed to collect feedback information from individual traders' prior behavior and publish it to communities as an individual feedback profile. The success of future transactions depends on how people behave today. Most studies [12,13,14] indicate that positive feedback increases the price and probability of sale while negative feedback decreases the price and probability of sale.

2.2 Binary Feedback

The binary feedback model is a mechanism where clients (buyers) can only rate past transactions as either "positive" (1) or "negative" (0). Positive ratings indicate that clients received high quality or satisfactory services (or goods), and negative ratings indicate that clients received low quality or unsatisfactory services. The summary of ratings is publicly available to all clients. As a result, clients know the quality of sellers (GSPs in our case) based on the summary of their most recent ratings [15].

According to Dellarocas [11], quality can be divided into three categories: *real quality* (q_r), *advertised quality* (q_a), and *estimated quality* (q_e). Real quality is unknown to clients in advance and can only be determined after consumption. Generally, clients prefer higher quality to lower quality, although their willingness to pay for extra quality varies. Advertised quality, controlled by GSPs, informs clients through advertising, and it may or may not reflect real quality. Estimated quality is based on the information that is collected from clients. Basically, clients assess the quality based upon the advertised quality and GSPs' rating profile.

This quality information can be used to calculate client satisfaction, which is the difference between real quality and estimated quality. A client decides whether to rate a transaction based on satisfaction (S_j). If the real utility exceeds the expected utility ($S_j>0$), the client should rate that transaction as positive. On the other hand, if the real utility falls below some threshold ($-\lambda_j$) of the expected utility ($S_j \leq -\lambda_j$), the client should rate that transaction as negative. Furthermore, if the client receives slightly poor but not terribly poor service ($-\lambda_j<S_j\leq 0$), the client may not provide any feedback [11].

2.3 Value Profiles

Since the demand function either for a client or for a group of clients is hard to obtain, most studies assume a utility function (which is also difficult to empirically determine) and use it to calculate the price that clients are willing to pay. In this paper, we assert that a feasible way to obtain the willingness to pay of clients is to use feedback mechanisms like the binary feedback model. As discussed in Section 2.2, positive feedback implies that a client is satisfied with the received service in term of cost and time, so that (presumably) s/he is willing to pay for such service at a future time. Negative feedback implies that a client is unsatisfied with the received service, and s/he may not be willing to pay for such service again. As a result, we propose the use of binary feedback to deduce the willingness to pay of a collection of heterogeneous clients.

We will show that this value profile represents the willingness to pay of clients at different prices so it can function as a proxy for a market demand function. To do this, we must first build a simple model of GSPs and clients. Using this model, we can show that a value profile behaves like a demand function and is useful for grid resource planning. Since they are economic actors, we refer to them as GSP agents and client agents.

2.4 GSP Agents

GSP agents provide services to clients based on a SLA. In each period, the agents announce an estimated service price to clients. The price varies based on the utilization of system, which we express as [16]

$$p_i^{est} = a_i + b_i * \left(\rho_i \big/ (1 - \rho_i) \right)$$

The agents also announce an estimated computing time, which is the mean response time of the system. Table 1 provides notations and parameters of GSP agents.

Table 1 -- Parameters and definitions

	Symbol	Description
GSP	i	GSP index
	p_i^{est}	Estimated price for computing a job in that period, which equals an estimated cost to the client
	a_i	Constant for adjusting pricing range
	b_i	Constant for adjusting pricing range
	ρ_i	Average utilization of system *i* in that period
Client	j	Client index
	S_j	Service satisfaction of client j after his/her job is completed
	$\propto_j$	Constraint sensitivity of client j
	$t_j^{preferred\_durat}$	Preferred duration for completing a job of client j
	$t_j^{total\_computing}$	Total computing time for completing a job of client j
	$\%\Delta t_j^{computing}$	Percentage of the difference between the preferred duration and the total computing time for completing the job of client j
	$c_j^{\exp}$	Expected cost (job budget) for completing job of client j
	c_j^{real}	Real cost for completing job of client j
	$\%\Delta C_j^{computing}$	Percentage of the difference between the expected cost (job budget) and the real cost for completing the job of client j
	FB_j	Binary feedback from client j
	λ_j	QoS threshold of client j

2.5 Client Agents

Clients are usually satisfied when a job is completed within the budget and the preferred job duration. Thus, client satisfaction (S_j) can be calculated as the difference between the change in computing time and the change in computing cost, which we express as (see Table 1)

$$S_j = \alpha_j * \left(\% \Delta t_j^{computing}\right) - \left(1 - \alpha_j\right) * \left(\% \Delta C_j^{computing}\right)$$

where

$$\%\Delta t_j^{computing} = (t_j^{preferred\_duration} - t_j^{total\_computing}) \Big/ t_j^{preferred\_duration}$$

$$\%\ \Delta\ C_j^{computing} = \left(c_j^{\exp} - c_j^{real}\right) \Big/ c_j^{\exp}$$

Unless otherwise specified, we assume that clients are truthful and cooperative. As we described in Section 2.2, after a service, each client gives binary feedback FB_j based on his/her satisfaction, which is expressed as (see Table 1)

$$FB_j = \begin{cases} "1" & if \quad S_j > 0 \\ "0" & if \quad S_j \le -\lambda_j \\ no\ rate & if \quad -\lambda_j < S_j \le 0 \end{cases}$$

2.6 Assumptions and Justifications

We make a number of simplifying assumptions in this paper so that we can focus on the proof of the concept. We will relax some of these assumptions in future research. First, we offer jobs to the grid from heterogeneous clients that have different constraints (budget and preferred duration). Clients will decide whether to submit a job based on the GSPs' service announcements subject to their constraints, as summarized in Fig. 2. Then, we assume that each client has a different number of jobs to process and is able to submit only one job at a time. We also assume that all jobs are of the same size.

We use a M/M/1 queuing system with First Come First Serve (FCFS) policy. Thus, the price is directly proportional to system utilization. This helps regulate demand by encouraging clients who have low budgets and long preferred durations to wait for an off-peak period. Like purchasing airline tickets[3], we assume that prices will be fixed after GSPs and clients have an agreement; however, the completion time is uncertain. Thus, the estimated service cost is the same as the final cost ($\%\Delta C = 0$). We also assume that there is no discount rate for clients when GSPs miss a preferred deadline. Moreover, we assume that the relationship between the GSP and the client exists in the context of a competitive market.

To justify the concept of value profiles in Section 2.3, we ran preliminary tests[4] to determine whether value profiles are consistent with demand theory. Table 2 summarizes the parameters used in the preliminary tests. The algorithm for this experiment is presented in Fig. 2.

[3] When purchasing airline tickets, we pay a certain price but we do not know whether a flight will be delayed.
[4] We use Csim as the simulation tool.

```
Algorithm Design for Justifying a Value Profile

SET each client's constraint  (budget, preferred_duration, and no_of_jobs)
FOR all clients' job
  WHILE a client still has jobs to submit
      server announces server_estimate_price and server_estimate_computing_time
      IF client_remain_budget ≥ server_estimate_price THEN
        CALCULATE possible_job_done if a client submit his/her job at this price
          IF client_remain_jobs ≤ possible_job_done THEN
            IF  preferred_duration ≥ server_estimate_computing_time THEN
              SUBMIT job to queue
              WAIT until job is done
              DECREASE no_of_jobs
              RECORD total_service_price and total_computing_time
              CALCULATE satisfaction
              PROVIDE feedback based on the satisfaction
                IF satisfaction  > 0 THEN
                  feedback = 1 (positive)
                  COUNT no_of_jobs_with_positive_feedback
                ELSE IF satisfaction ≤ -λj THEN
                  feedback = 0 (negative)
                  COUNT no_of_jobs_with_negative_feedback
                ELSE IF -λj < satisfaction  ≤ 0 THEN
                  NO RATE
                END IF
              CALCULATE client_remain_budget
                IF client_remain_budget ≥ 0
                  Repeat submitting job process
                END IF
            ELSE
                WAIT for other period
            END IF
          ELSE
            WAIT for other period
          END IF
        ELSE
          WAIT for other period
      END IF
    END WHILE
END FOR
```

Figure 2 -- Algorithm design to evaluate value profile concept

Table 2 -- Parameter values used to test the value profile concept

	Parameter	Value
Environment	*NUM_CLIENTS*	100
	Interarrival rate	0.6 jobs/min
	Service rate	1.0 jobs/min
Client Agents	*Budget*	Uniform (1, {10,20})
	Preferred duration	Uniform (1, {5,20})
	No_of_jobs	Uniform (1, 50)
	λ_j	0
	$\propto_j$	1.0
	$t^{retransmit}$	1000 min
GSP Agents	a_i	0
	b_i	0.1
	p_i *(service price)*	Vary based on system utilization

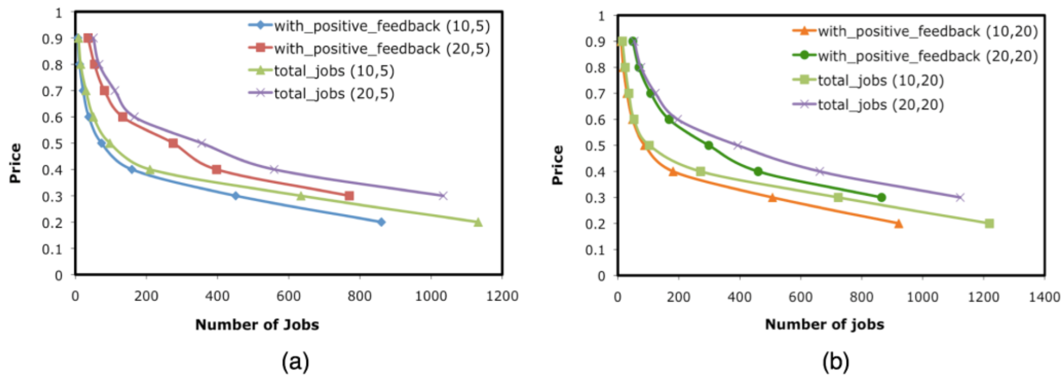


Figure 3 -- Demand curves based on value profiles (a) max_budget = {10,20} and max_preferred duration = 5 (b) max_budget = {10,20} and max_preferred duration = 20.

Fig. 3 shows the resulting value profiles. The curves clearly have the shape of typical demand functions since the number of jobs with positive feedback and the number of total jobs submitted to the system are inversely proportional to the service prices. Moreover, these curves shift up and to the right when clients increase their maximum budget or preferred duration (and vice versa). For example, at a price of 0.3, if they increase their budget from 10 to 20, the number of jobs with positive feedback and the number of total submitted jobs will increase from 451 to 770 and from 634 to 1035, respectively, as shown in Fig. 3(a). If they also increase their preferred duration from 5 to 20, the number of jobs with positive feedback and the number of submitted jobs will increase to 865 and 1123, respectively, as shown in Fig. 3(b). If they fix their budget at 10 and increase their preferred duration from 5 to 20, the number of jobs with positive feedback will increase from 451 to 510, as shown in Fig. 3(b). These results indicate that when clients' constraints change, value profiles can capture clients' willingness-to-pay at different prices. This is consistent with demand theory.

The results suggest that value profiles have the shape and characteristics of demand functions. Therefore, we conclude that binary feedback can be used to construct value profiles that represent clients' willingness-to-pay for grid resources. For the rest of this paper, we will use value profiles as proxies for demand functions. We believe that the use of value profiles can assist GSPs in finding an economic equilibrium point to plan their resource base.

3. Using Binary Feedback for Grid Resource Planning

The objective of this section is to explore clients' incentives related to the use of binary feedback. According to Dellarocas et al. [3], feedback contributions do not directly benefit the feedback providers, while other entities benefit more directly. Thus, individuals might have less of an economic incentive to provide feedback even if it is socially optimal for them to do so.

3.1 Incentives to Cooperate

Clients do not cooperate unless they receive some benefit from their contribution. In reputations on the popular Internet auction site eBay, eBayers receive benefits in their future trading. Even there, eBay persistently prompts transacting parties to leave feedback. In our study, GSPs ask clients to spend resources to provide feedback after they receive a result. If clients cannot clearly see benefits from providing feedback, they do not have any economic incentives to cooperate with GSPs. As a result, we require a mechanism to create an incentive for clients to cooperate.

People work hard when they think their effort will help them achieve outcomes that they value [17]. After providing feedback, clients will continue their contribution if they obtain noticeable benefits, which can be achieved through providing "selective incentives [18]" and by publishing "community activities

[19]". For selective incentives, GSPs have to treat feedback-providing clients better than non-feedback-providing clients; for example, in terms of higher priority in queuing and better rate. For community activity, like eBay, GSPs can show their feedback profile to represent how they manage jobs according to the SLA negotiated with clients.

3.2 Incentives to be Truthful

Rationally, self-interested individuals want to maximize their own payoff without any concern for another's payoff. In our case, clients prefer to consume as many resources as possible within their budgets. In the case of good service, although they may already be satisfied with the current service, they may still want a better service by lying to GSPs. If GSPs believe them, they will receive an upgraded service without any extra costs. On the other hand, in the case of bad service, clients can receive more benefits by claiming that the service is a lot worse than it actually is. Accordingly, they have an economic incentive to provide untruthful feedback. We assert that a feasible way to promote honesty in feedback models is to use credibility mechanisms.

The use of credibility mechanisms enables GSPs to detect and penalize insincere or biased clients. The idea is to ensure that sincere clients always receive more benefits than insincere or biased clients. Therefore, clients will fear penalties and will provide truthful feedback because of higher payoffs.

3.3 Promoting a Truthful Feedback

In this paper, we adopt the credibility mechanism that is proposed by Papaioannou and Stamoulis [20]. We assume that when clients decide to submit their job, it means that they are willing to pay that price and their preferred duration is greater than or equal to the estimated computing time ($t_j^{pref\_duration} \ge t_j^{estimated\_computing}$). Then, we assume that the estimated computing time can be used as the reference of clients' preferred duration.

With this assumption, GSPs can detect and penalize insincere or biased clients by comparing the client's preferred duration ($t_j^{preferred\_durat}$) with the total computing time ($t_j^{total\_computing}$). GSPs usually expect to receive positive feedback if they can finish jobs within preferred durations, and vice versa. Thus, the expected feedback[5] ($FB_j^{\exp}$) can be expressed as

$$FB_j^{\exp} = \begin{cases} "1" & if\ t_j^{total\_computing} \le t_j^{preferred\_duration} \\ "0" & if\ t_j^{total\_computing} > t_j^{preferred\_duration} \end{cases}$$

If the received feedback does not meet the expectations of the GSP ($FB_j^{\exp} \neq FB_j^{recei}$), that feedback is considered to be untruthful and will be discarded. Moreover, that client's job will be held for some periodof time ($t_j^{penalty}$) before receiving next service. The client will not be offered the opportunity to provide feedback after that job since it will certainly be negative because of the longer computing time ($t_j^{greedy\_computing} = t_j^{total\_computing} + t_j^{penalty}$).Feedback that matches expectations ($FB_j^{\exp} = FB_j^{recei}$) will be counted. This process continues until there are no new jobs. We believe that the use of a credibility mechanism such as the one described can promote truthful feedback. The algorithm of this mechanism is provided in [22].

---

[5] Note that GSPs' expected feedback function ($FB_j^{\exp}$) is not equivalent to clients' feedback function (FB_j) since clients may not provide any feedback when receiving slightly bad service (see Section 2.5).

4. Analysis of Clients' Incentives

Given the clients' incentives (discussed in Section 3), we now consider the interaction between clients and GSPs using game theory. In this paper, we use strategic games[6], which are often used for non-cooperative games. The objectives of this section are to analyze the cooperation between a GSP (player I) and a client (player II), and determine whether the use of credibility mechanism can help the cooperation emerge.

4.1 Individual Payoff

Before starting game analysis, we have to define a payoff for each individual player. The key players are a GSP and a client. The profit (payoff) for a GSP is known as

$$P = R - C$$

where R is a revenue from providing a service to a client, and C is the cost of providing that service. We assume that a GSP always sets its revenue to cover its cost ($R > C > 0$).

The payoff for a client is defined as

$$P = Q - R$$

where R is the service cost that the client has to pay, which is equivalent to the revenue R of the GSP. Q is the client's perception of value, which decays exponentially with time; this represents the QoS that the client received from the GSP. Therefore, in this function, the value of Q is proportional to the difference between preferred duration and total computing time, which we express as

$$Q = Re^{(t^{preferred\_duration} - t^{total\_computing})}$$

For each strategic combination, a payoff for each player is calculated from the difference between his/her payoff both before and after making such decision, which we express as

$$P_{GSP} = \Delta R - \Delta C$$
$$P_{client} = \Delta Q - \Delta R$$

4.2 Feedback Games

We play feedback games by assuming that all players are rational and selfish, and that a client always provides feedback. Then, we assume that a GSP plans resources based on the received feedback. Therefore, a GSP will upgrade a service if a client reports a poor service, and vice versa. Table 3 summarizes parameters used in these feedback games.

Table 3 -- Feedback game parameters (R > C > 0 and 0 < ω < k < 1)

	Upgrade	*Downgrade*
Q	Q = (1+ω)*Q	Q = (1-ω)*Q
R	R = (1+k)*R	R =R
C	C = (1+ω)*C	C = (1-ω)*C

[6] These games usually assume that both players move simultaneously, or that later movers do not have any information about the earlier players' moves.

In these games, we assume that a GSP has two resource planning strategies after receiving a client's feedback: *Trust* or *Don't trust* the feedback. Believing the feedback might result in services that are more costly to provide. A client also has to choose between providing *Truthful* and *Untruthful* feedback. Untruthful feedback gives an insincere client more benefits than truthful feedback. On the other hand, truthful feedback is more valuable than untruthful feedback to the GSP. A resource manager will disregard the client's feedback if s/he knew that the client lied. Table 4 analyzes the game given that the client already receives good service; Table 5 offers an analysis based on the given that the client receives poor service.

Table 4 -- Feedback game given good service

II (Client) / *I (GSP)*	*Truthful*	*Untruthful*
Trust	-ωQ ωC	ωQ-kR kR- ωC
Don't trust	0 0	0 0

Table 5 -- Feedback game given poor service

II (Client) / *I (GSP)*	*Truthful*	*Untruthful* *(badly poor, b>k>ω)*
Trust	ωQ-kR kR- ωC	bQ-kR kR-bC
Don't trust	0 0	0 0

In Tables 4-5, "Trust" strategy dominates[7] "Don't trust" strategy. Since both players are rational, the client realizes that the GSP always prefers to trust the received feedback[8]. Then, the client will provide untruthful feedback because of the higher payoff[9]. As a result, the rationality of both players leads to the conclusion that the client will provide untruthful feedback and the GSP will trust it. At this point, cooperation cannot emerge and the GSP requires a mechanism to detect an insincere client.

4.3 Credibility-based Feedback Game

What makes it possible for the cooperation to emerge is that both players have to meet each other again, recognize each other from the previous transaction, and recall how the other behaved last time [21]. The decision of players not only affects the outcome of the current move, but also influences future decisions of the players. This is called an iterated game.

Using the credibility mechanism, the GSP checks whether the received feedback is truthful by comparing it with their performance record. This mechanism allows the GSP to penalize the insincere client for previous non-cooperative play. This is similar to "Tit for tat" strategy[10] in game theory. In such games, cooperation might arise as an equilibrium outcome. The incentive for client to defect is overcome by the threat of penalties, which leads to the possibility of a cooperative outcome. Consequently, we integrate the credibility mechanism into the feedback games.

---

[7] Strategic dominance only occurs when one strategy gives higher payoff than another strategy for individual player, no matter how that player's opponents would play.

[8] Given good and poor service, the resulting payoffs of ωC, kR-ωC, and kR-bC when cooperating (trust) are greater than zero when defecting (don't trust), respectively.

[9] Given good service, the resulting payoff of –ωQ when cooperating (truthful) is less than ωQ-kR when defecting (untruthful). Likewise, given poor service, the resulting payoff of ωQ-kR when cooperating (truthful) is less than bQ-kR when defecting (untruthful).

[10] A player using this strategy will initially cooperate, and then respond based on another player's previous decision.

Table 6 -- Credibility-based feedback game with good service

I (GSP) \ *II (Client)*	*Truthful*	*Untruthful*
Trust	-ωQ ωC	-Q 0
Don't trust	0 0	-Q 0

Table 7 -- Credibility-based feedback game with poor service

I (GSP) \ *II (Client)*	*Truthful*	*Untruthful*
Trust	ωQ-kR kR- ωC	-Q 0
Don't trust	0 0	-Q 0

In Tables 6-7, the "Truthful" strategy dominates the "Untruthful" strategy because of the credibility mechanism. Since both players are rational, the GSP realizes that the client is afraid of the penalty so the client will provide truthful feedback[11]. As a result, the rationality of both players leads to the conclusion that the client will provide truthful feedback and the GSP will trust it[12]. Cooperation succeeds at this point.

In conclusion, the results show that the cooperation cannot emerge through the use of feedback alone; credibility mechanisms are required for cooperation to emerge. In this paper, we use credibility-based binary feedback [22] to build value profiles, which GSPs can use to optimally plan their resources.

4.4 Effect of Untruthful Feedback Game

According to our game analyses, clients are better off cooperating with GSPs when a credibility mechanism is used. However, the model cannot filter out all untruthful feedback because they receive limited information from clients. Thus, non-cooperative and untruthful clients still occur in two cases, as summarized in Table 8.

Table 8 -- The effect of untruthful feedback

Client Satisfaction	*Expected Feedback*	*Received Feedback*	*Analysis*
(1) $S_j \geq 0$	"+"	"No rate"	Non-cooperative / Untruthful
(2) $\lambda_j \leq S_j < 0$	"No rate"	"-"	Untruthful

In the first case, clients might not respond because of noncooperation or being untruthful even though the GSP finished the job within the preferred duration and budget. In the second case, untruthful clients will provide negative feedback instead of no response when GSPs slightly fail to meet clients' requirement. These two cases will affect value profiles and client satisfaction rates (or percentage of positive feedback) and might cause GSPs' investment decisions to change.

[11] Given good and poor service, the resulting payoffs of -ωQ and ωQ-kR when cooperating (truthful) are greater than –Q when defecting (untruthful), respectively.
[12] Given good and poor service, the resulting payoffs of ωC and kR-ωC when cooperating (trust) are higher than zero when defecting (don't trust), respectively.

5. Applying Value Profiles

This section presents the results and discussions of the use of a credibility-based binary feedback model for grid resource planning. An economic equilibrium is required to test these questions. This equilibrium point can be determined by the intersection of cost function and market demand. Thus, we must construct a resource cost function and a proxy for market demand.

5.1 Resource Cost Function

Since the lifetime of investments in computing resources is substantially longer than demand fluctuations, we must consider the long-run production cost function. To generate a cost function, we begin by selecting six server systems from the TPC-C Benchmark [23]. In this paper, we assume that these six server systems represent six different scales of GSPs measured by the number of processors per GSP. Then, we determine a long-run average total cost (LATC) for one year, which is the sum of the short-run average total cost (SATC) of each GSP. The $SATC_i$ can be calculated as

$$SATC_i = \frac{TCO_i + network\ cost_i + operation\ cost_i}{no.\ of\ \ job\ production_i}$$

where the total cost of ownership (TCO_i) is

$$TCO_i = hardware\ cost_i + storage\ cost_i + software\ cost_i$$

Table 9 -- One-year cost analysis of each GSP

GSP_i	(1) *tpmC* (Transaction Rate)	(2) = (1)/10^6 Service Rate (jobs/min)	(3) No. of Job Production ($x10^3$/yr)	(4) TCO_i ($x10^3$)
GSP_1	236,271	0.24	122	\$190
GSP_2	404,462	0.40	210	\$445
GSP_3	841,809	0.84	436	\$939
GSP_4	1,245,516	1.25	646	\$1,625
GSP_5	1,616,162	1.62	838	\$1,795
GSP_6	2,196,268	2.20	1,139	\$3,138

By assuming that each job has 10^6 transactions, Table 9 shows the one-year cost function for each GSP[13]. The quadratic LATC function in Fig.4 is determined by performing a regression[14], which can be expressed as

$$LATC = 3.59 + (-0.003)(no.of\ \ jobs) + \left(1.817 * 10^{-6}\right)(no.of\ \ jobs)^2$$

[13] Due to the TPC-C Benchmark in Table 9, column (1) shows the computing capacity of each system. By assuming that each job has 10^6 transactions, column (2) shows the service rate of each system. With these values, we can calculate the number of job production in one year, as shown in column (3).

[14] We run the regression model based on the values in columns (3) and (4). The statistical summary indicates that the cost function has a high coefficient of determination (R^2=0.811) and a low significance level (Sig.=0.082).

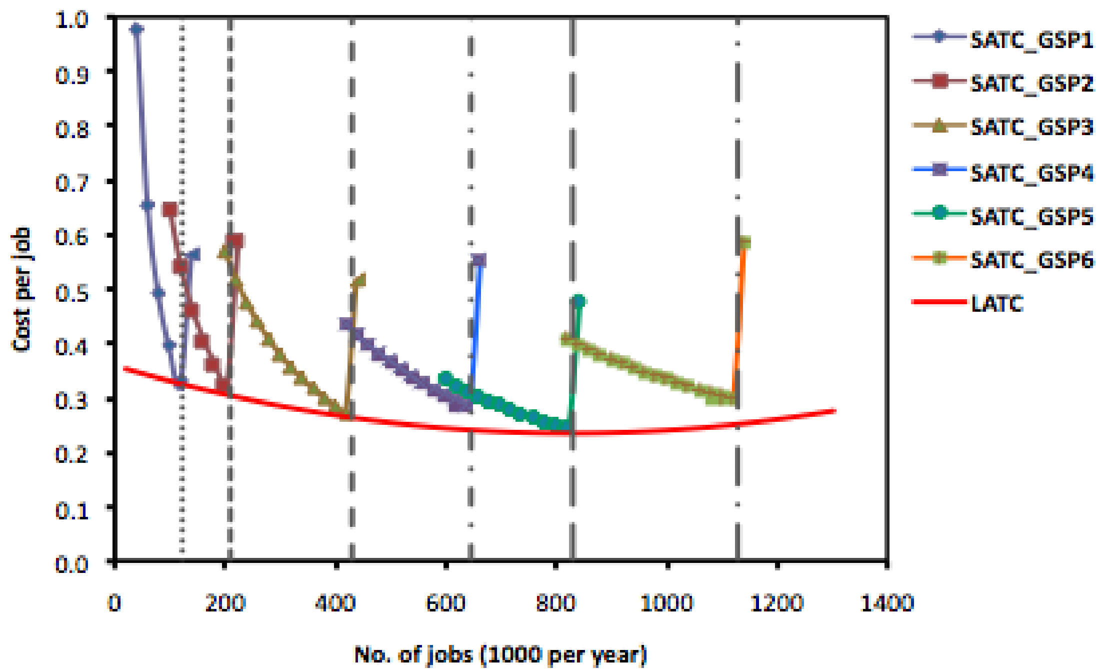


Figure 4 -- LATC curve for GSPs

5.2 Market Demand

We argued earlier that a value profile could function as a proxy for the demand function of an individual GSP, so the sum of each value profile can be used as a proxy for market demand[15]. By assuming that all value profiles are published, market demand can be expressed as

$$Market\ Demand = \sum_{i=1}^{\#GSPs} (value\_profile_i)$$

In this paper, we use this market demand to represent the willingness to pay of all grid clients. Furthermore, since we assume that the market is a perfectly competitive market, normal profits[16] only arise when the GSP's long-run economic equilibrium is reached. Then, with the use of value profiles, the economic equilibrium can be determined by the intersection of the market demand curve and the LATC curve.

5.3 Experimental Approach

In these experiments, we assume that each client always provides feedback FB_j based on their satisfaction level S_j which can be expressed as (see Table 1 for the summary of notations and parameters of client agents)

[15] At any price, market demand is the sum of the quantities demanded by each individual's demand.
[16] Normal profit is a zero economic profit (where revenue equals cost) since economic profit does not occur in a perfectly competitive market in long-run equilibrium regarding to the free entry and exit of GSPs.

$$FB_j = \begin{cases} "1" & if \quad S_j > 0 & with\ P(contribution) \\ "no\ rate" & if \quad S_j > 0 & with\ (1-P(contribution)) \\ "0" & if \quad S_j \le -\lambda_j & with\ P(contribution) \\ "no\ rate" & if \quad S_j \le -\lambda_j & with\ (1-P(contribution)) \\ "no\ rate" & if \quad -\lambda_j < S_j \le 0 & with\ (1-P(untruthful)) \\ "0" & if \quad -\lambda_j < S_j \le 0 & with\ P(untruthful) \end{cases}$$

We believe that GSPs can increase clients' incentives to cooperate, as discussed in Section 3, so in this section we assume that clients always provide feedback. We consider two cases: *a low demand case* and *a high demand case*. In the first case, a high-capacity GSP might want to downsize because of low demand. In the second case, a low-capacity GSP might want to upgrade because of high demand. These two cases frequently arise when making capital investment decisions. Table 10 summarizes the default values of parameters for the experiments.

Table 10 -- Simulation parameters

	Parameter	Value
Environment	*TIME*	518400 minutes (1 year)
	Job load	• Low demand (market case 1) ○ Num_clients = 5,000 ○ Interarrival rate = 0.3 jobs/min • High demand (market case 2) ○ Num_clients = 50,000 ○ Interarrival rate = 0.5 jobs/min
	Service rate	{0.24, 0.40, 0.84, 1.25, 1.62, 2.20}
Client Agents	*Budget*	Uniform (5, 100)
	Preferred duration	Uniform (5, 30)
	No_of_jobs	Uniform (10, 100)
	QoS threshold (λ_j)	Uniform (-2,0)
	$\propto_j$	1.0
	$t^{retransmit}$	1000 min
	P(contribution)	1
	P(untruthful)	{0, 1.0}
GSP Agents	a_i	0
	b_i	0.1
	p_i *(service price)*	Vary based on system utilization

5.4 Case I: Low Demand with Truthful Feedback

In this case, we study whether a GSP_6 should downgrade its computing capacity because of low demand. Without value profiles, GSP_6 produces 274,711 jobs/year and its revenue is below the LATC curve. This means that GSP_6 loses approximately \$5,494, as shown in the small shaded rectangle in Fig. 5.

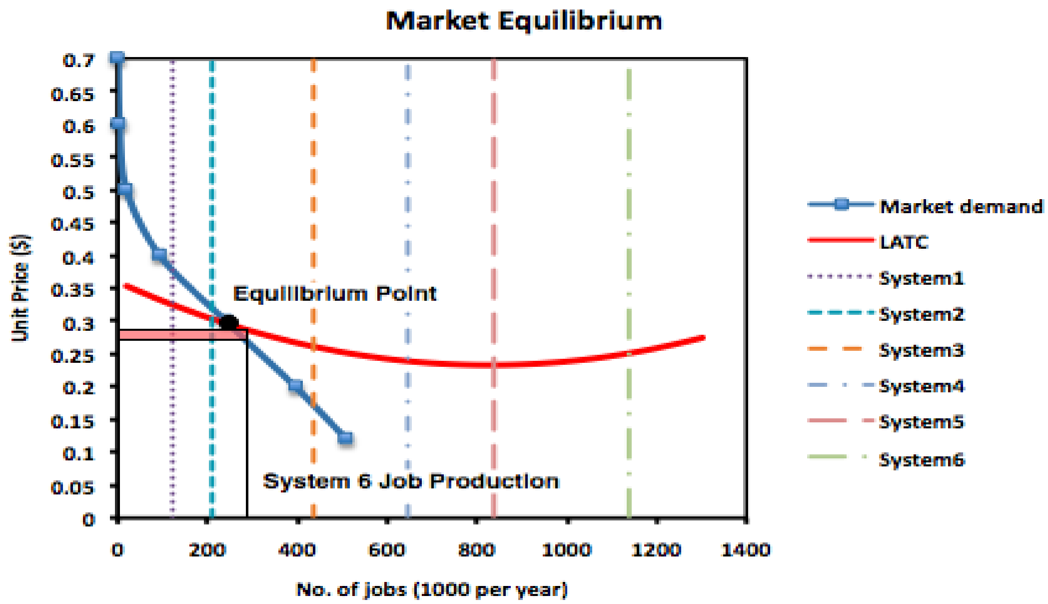


Figure 5 -- Market equilibrium in Case I

With value profiles, as shown in Fig. 5, GSP_6 will profit by reducing its service rate from 2.2 to 0.4 jobs/min since the equilibrium point is close to GSP_2. Fig. 6 summarizes the tradeoffs after downgrading to the same capacity as GSP_2. The results clearly show that system utilization significantly increases from 45.7% to 82.5% and the percentage of positive feedback decreases from 99.4% to 94.8%. The clients' job success rate also decreases from 99.3% to 92.5%. Figure 6(d) clearly shows the tradeoffs when using value profiles,. While system utilization increases by 36.8% percentage points, percentage of positive feedback and clients' job success rate decreases by 4.6% and 6.8% percentage points, respectively. In summary, the use of value profiles benefits GSP_6 in resource cost savings even though client satisfaction rate decreases slightly (but remains over 90%).

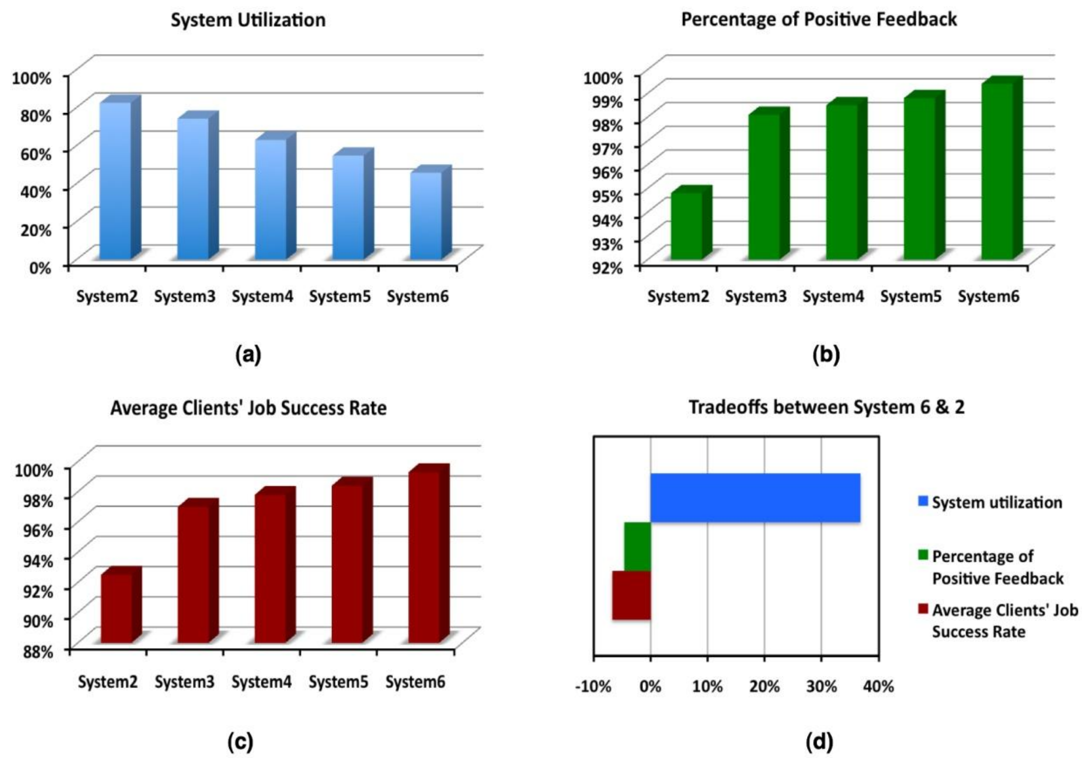


Figure 6 -- Summary of tradeoffs using value profiles in Case I

Since clients are worse off in this case, they might not want to cooperate with GSP_6 or will even provide untruthful feedback. This will cause the market demand curve and the equilibrium point to change. To investigate this, we changed the values of *P(contribution)* and *P(untruthful)* in Table 10. In this experiment, each client is randomly assigned with a different probability and observe the outcome.

As explained in Section 3.3, GSPs expect to receive positive feedback if they can finish a job within a preferred duration. Thus, this expected positive feedback can be used as the reference market demand curve when clients are always cooperative, *P(contribution)* = 1, and truthful, *P(untruthful)* = 0.

With non-cooperative and untruthful clients, the received market demand curve shifts down and to the left, as shown in Fig. 7. The result shows that the received and reference equilibrium points are close to GSP_1 and GSP_2, respectively. With the received demand information, GSP_6 will downsize to the same capacity as GSP_1 instead of GSP_2. In other words, GSP_6 will reduce its service rate from 2.2 to 0.24 jobs/min, which is a reduction of 1.96 jobs/min. If clients are cooperative, GSP_6 will reduce its capacity rate to 0.4 jobs/min, which is a reduction of only 1.8 jobs/min. So if clients are non-cooperative and untruthful, clients will experience more delays in the system, so they are better off cooperating with GSP_6.

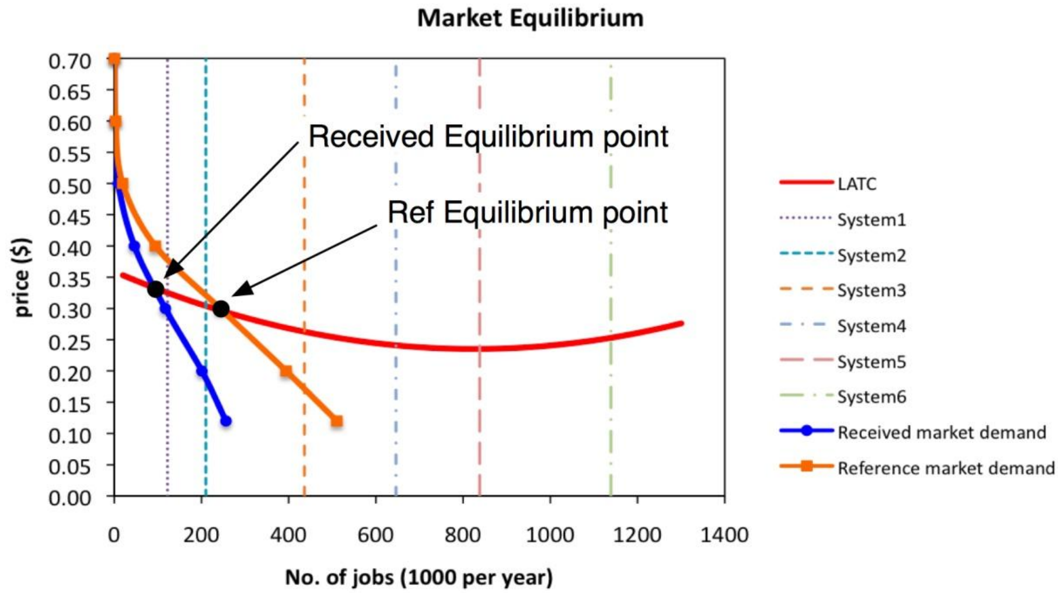


Figure 7 -- Change in demand curve and equilibrium for Case I

To further investigate the downgrade scenario[17], we play a non-cooperative game between a GSP and a client; this will enable us to study the development of cooperation for this situation. Both the GSP and the client have two strategies: *Cooperate* or *Defect*. Table 1 shows that the resulting payoff of $-\omega Q$ when cooperating is larger than $-kQ$ when defecting. As a result, clients will cooperate with the GSP_6 (the cooperation will emerge).

Table 11 -- Client's incentive game

II (Client) / *I (GSP)*	*Cooperate*	*Defect*
Cooperate	ωC, $-\omega Q$	kC, $-kQ$
Defect	0, 0	0, 0

5.5 Case II: Low Demand with Untruthful Feedback

Untruthful feedback might cause the accuracy of percentage of positive feedback to decrease. In this case, we examine whether untruthful feedback affects GSP_6's decision in case I.

[17] See Table 4.

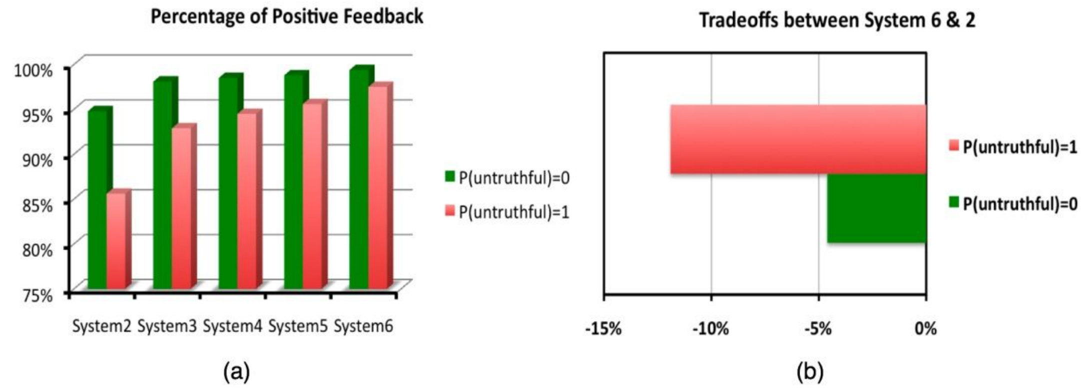


Figure 8 -- Percent positive feedback, Case II vs. Case I

With truthful feedback in case I, positive feedback decreases slightly by 4.6% percentage points after GSP_6 downsizes its computing capacity to level of GSP_2. However, with untruthful feedback, the positive feedback decreases from 97.5% to 85.6%, which is a 11.9% decline in percentage points as presented in Fig. 8(b). Therefore, untruthful feedback reduces the accuracy of positive feedback by 7.36% percentage points.

If GSP_6 can live with this untruthful client satisfaction rate, it will continue using this computing capacity (GSP_2). On the other hand, if GSP_6 prefers to keep its client satisfaction rate over 90%, it might decide to upgrade its computing capacity to be the same as GSP_3. However, GSP_6 will end up with losses since revenue of GSP_3 is below the LATC curve, as presented in Fig 5. As a consequence, this would not happen because GSP_6 has to have a profit. Thus, GSP_6 will not change its decision for this case.

5.6 Case III: High Demand with Truthful Feedback

In the high demand case, we examine whether GSP_2 should upgrade its computing capacity to generate more revenue. Without value profiles, GSP_2 produces 186,758 jobs/year and its revenue is clearly over the LATC curve. This means that GSP_2 has a super normal profit, as shown in the green square in Fig. 9. Note that when a firm has a super normal profit, it attracts new entrants to the industry. By increasing its capacity to produce more jobs at lower cost, a GSP would achieve economies of scale and reduce the entry incentive. Thus, GSP_2 must increase its service rate from 0.4 to 1.62 jobs/min since the equilibrium point is close to GSP_5, as shown in Fig. 9.

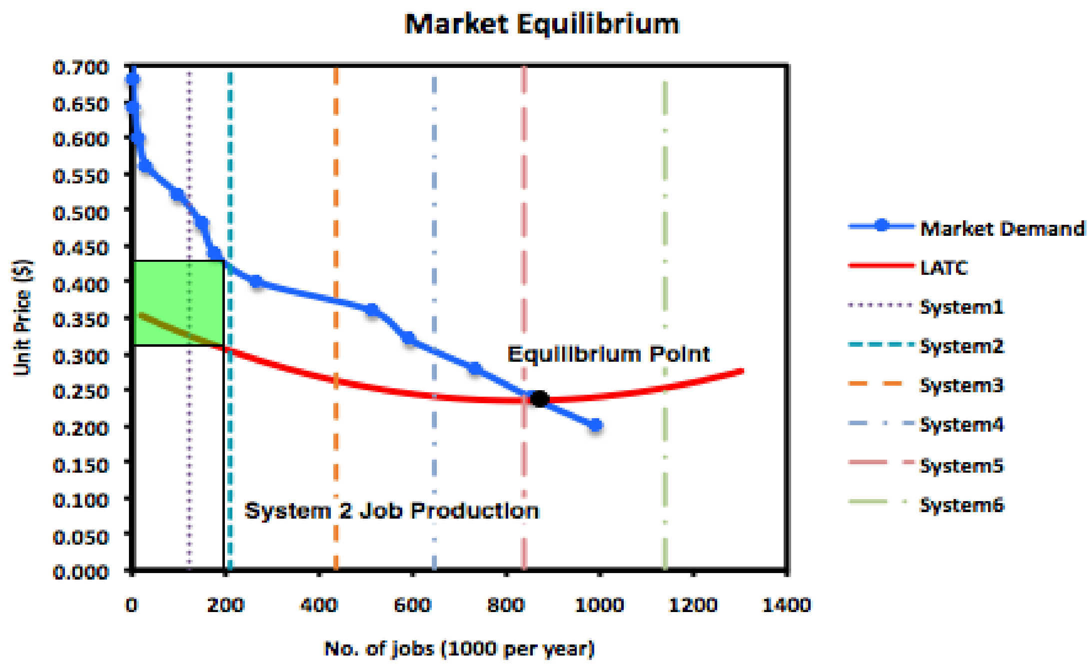


Figure 9 -- Market equilibrium in Case III

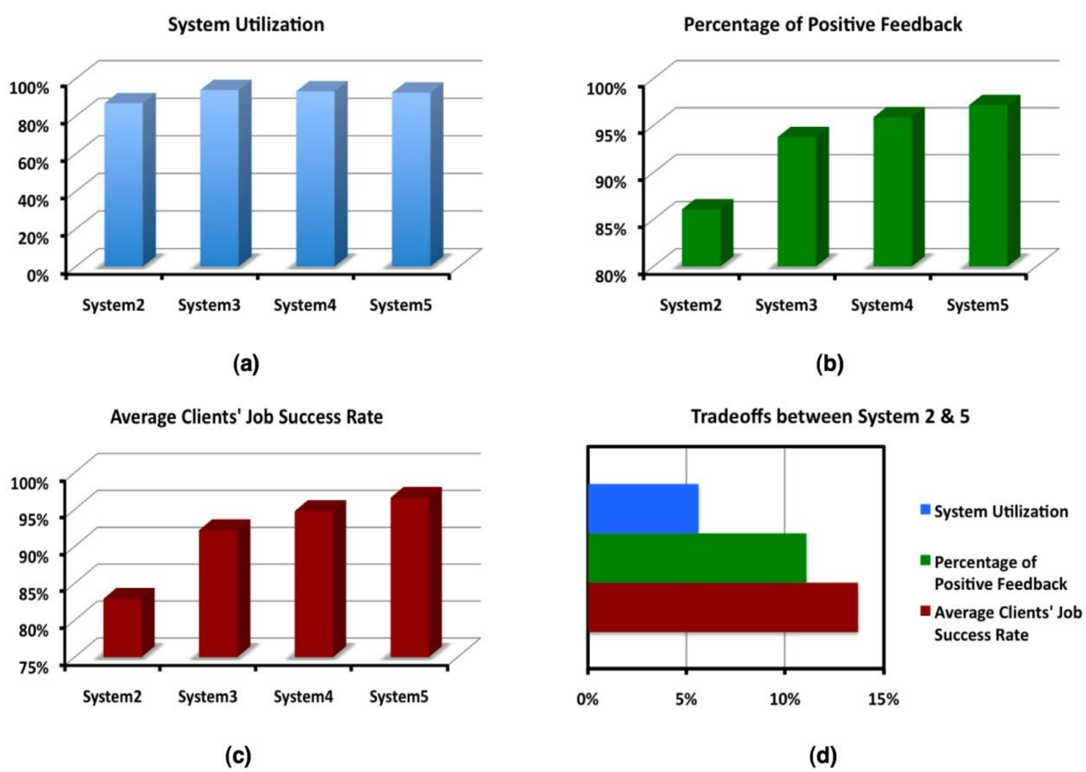


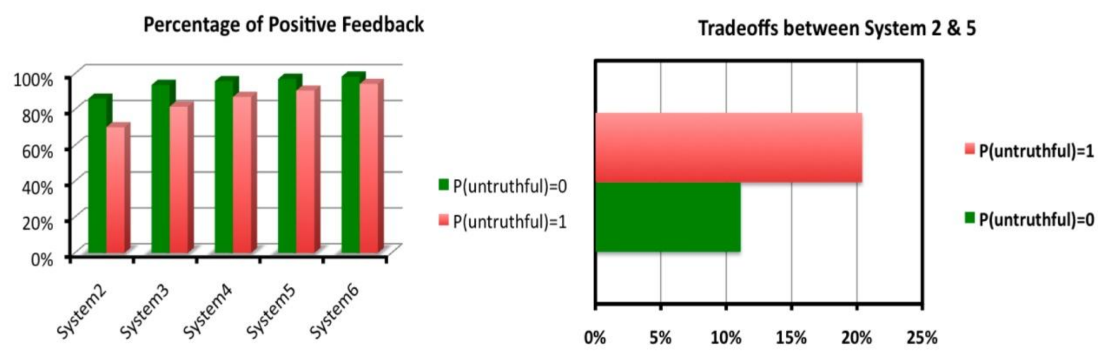
Figure 10 -- Summary of the tradeoffs when using value profiles in Case III

Fig. 10 summarizes the tradeoffs involved. After upgrading to the same capacity as GSP_5, GSP_2's system utilization increases from 87% to 92.6%, and percentage of positive feedback goes up from 86.1%

to 97.2%. Moreover, clients' job success rate significantly improves from 82.9% to 96.6%. As shown in Fig. 10(d), we do not observe the tradeoffs we saw in Case I. The results show that system utilization, percentage of positive feedback, and clients' job success rate increase by 5.6, 11.1, and 13.7 percentage points, respectively. In summary, the use of value profiles provides benefits to GSP_2 both in terms of revenue and client satisfaction rate. Likewise, clients receive the benefit from the use of value profiles since their job success rate goes up.

5.7 Case IV: High Demand with Untruthful Feedback

As in Section 5.5, the purpose of this case is to determine whether untruthful feedback affects GSP_2's investment decision in Case III. With truthful feedback in Case III, positive feedback increases by 11.1% percentage points after GSP_2 upgrades its computing capacity to the same as GSP_5. With untruthful feedback in this case, positive feedback increases from 70.2% to 90.6%, which is an increase of 20.4 percentage points as shown in Fig. 11(b). Therefore, untruthful feedback significantly inflates the accuracy of positive feedback by 9.18% percentage points.



(a) (b)

Figure 11 -- Percent positive feedback, Case IV vs. Case III

As this untruthful client satisfaction rate is around 90%, GSP_2 will continue to use this computing capacity (GSP_5). Unless GSP_2 requires client satisfaction rate of 95%, it might decide to upgrade its computing capacity to be the same as GSP_6. In doing so, GSP_2 will lose since the revenue of GSP_6 is below the LATC curve, as shown in Fig. 9.

6. Conclusion

As the goal of GSPs is to improve their economic viability by maintaining the lowest possible amount of resources to meet client demand, they have to know how clients value their services. In this paper, we describe how we constructed a value profile using binary feedback for a collection of heterogeneous grid clients, which GSPs can use to economically plan their resources. The results show that binary feedback can be used to construct a value profile that will serve as a proxy for a demand function to represent client's willingness-to-pay for grid resources at different prices.

Since clients have incentives to provide untruthful feedback, we use credibility mechanisms to detect untruthful feedback and our game analysis shows that a credibility mechanism can help cooperation merge. Using the mechanism of the value function, we show that a credibility-based binary feedback can assist GSPs in economically planning their resources.

References

[1] Weiss, Martin "Models of Grids: Cost and Standards Implications" (Invited) 1st IEEE Workshop on Grid Economics and Business Models, Seoul Korea, 23 April 2004.
[2] R. Buyya, Economic-based Distributed Resource Management and Scheduling for Grid Computing, Ph.D. dissertation, Monash University, Melbourne, Australia, April 2002.
[3] C. Dellarocas, F. Dini, and G. Spagnolo, Designing Reputation (Feedback) Mechanisms, Handbook of Procurement, Cambridge University Press, 2006.
[4] M. Siddiqui, A. Villazon, and T. Fahringer, Grid Capacity Planning with Negotiation-based Advance Reservation for Optimized QoS, in Proc. of the 2006 ACM/IEEE conference on Supercomputing (SC'06), Tampa, Florida, no. 103, 2006.
[5] D. Minoli, A Networking Approach to Grid Computing, J. Wiley, 2005.
[6] K. Krauter, R. Buyya, and M. Maheswaran, A taxonomy and survey of grid resource management systems for distributed computing, International Journal of Software: Practice and Experience (SPE), no. 32, 2002, pp.135-164.
[7] C. S. Yeo and R. Buyya, A Taxonomy of Market-based Resource Management Systems for Utility-driven Cluster Computing, Software-Practice & Experience, vol. 36, issue 13, Nov. 2006, pp. 1381-1419.
[8] L. ChunLin and L. Layuan, A Utility-based Two Level Market Solution For Optimal Resource Allocation In Computational Grid, in Proc. of the 2005 International Conference on Parallel Processing (ICPP'05), Jun. 2005, pp. 23-30.
[9] R. Wolski, J. S. Plank, J. Brevik, and T. Bryan, Analyzing market-based resource allocation strategies for the computational grid, International Journal of High Performance Computing Applications, vol. 15, issue 3, Aug. 2001, pp. 258-281.
[10] K. Subramoniam, M. Maheswaran, and M. Toulouse, Towards a Micro-Economic Model for Resource Allocation in Grid Computing Systems, in Proc. of the 2002 IEEE Canadian Conference on Electrical & Computer Engineering, 2002, pp. 783-785.
[11] C. Dellarocas, Analyzing the Economic Efficiency of eBay-like Online Reputation Reporting Mechanisms, in Proc. Of the 3rd ACM conference on Electronic Commerce, Tampa, Florida, 2001, pp. 171-179.
[12] P. Resnick, R. Zeckhauser, J. Swanson, and K. Lockwood, The Value of Reputation on eBay: A controlled experiment, (July 2002). KSG Working Paper Series No. RWP03-007. [Online]. Available: http://ssrn.com/abstract=385206
[13] M. I. Melnik and J. Alm, Does A Seller's Ecommerce Reputation Matter? Evidence From eBay Auction, The journal of Industrial Economics, vol. 50, Sep. 2002, pp.337-349.
[14] D. Lucking-Reiley, D. Bryan, N. Prasad, and D. Reeves, Pennies from eBay: the Determinants of Price in Online Auctions, in Proc. of the eleventh ACM SIGKDD International Conference on Knowledge Discovery in Data Mining, Chicago, IL, 2005, pp. 411-418.
[15] C. Dellarocas, Sanctioning Reputation Mechanisms in Online Trading Environments with Moral Hazard (July 2004). MIT Sloan Working Paper No. 4297-03. [Online]. Available: http://ssrn.com/abstract=393043
[16] S. Tareen, M. B. H. Weiss, and T. F. Znati, A Study of Pricing Strategies and Traffic Behavior in Packet Networks, PhD Dissertation, Telecommunications Program and Department of Computer Science, University of Pittsburgh, Nov. 7, 1998.
[17] G. Beenen, K. Ling, X. Wang, K. Chang, D. Frankowski, P. Resnick, and R. E. Kraut, Using Social Psychology to Motivate Contributions to Online Communities, in Proc. of the 2004 ACM conference on Computer Supported Cooperative Work, Chicago, IL, 2004, pp. 212-221.
[18] M. Olson, The Logic of Collective Action, Cambridge, MA: Harvard University Press, 1964.
[19] C. Nagele, Social Networks Research Report, Jul. 2005. [Online]. Available: http://tidbit.wildbit.com/2005/07/social_networks.html
[20] T. G. Papaioannou and G. D. Stamoulis, An Incentives' Mechanism Promoting Truthful Feedback in Peer-to-Peer Systems, in Proc. of the Fifth IEEE International Symposium on Cluster Computing and the Grid (CCGrid'05), vol. 1, 2005, pp. 275-283.
[21] R. Axelrod, The Evolution of Cooperation, New York: Basic Books, 1984.
[22] P. Chokesatean, Credibility-based Binary Feedback Model for Grid Resource Planning, Ph.D. dissertation, University of Pittsburgh, Pennsylvania, USA, June 2008.
[23] Transaction Processing Performance Council (TPC), TPC-C benchmark Version 5.9, (February 2008). [Online]. Available: http://www.tpc.org/tpcc/results/tpcc_results.asp?orderby=hardware**.**