Вычислительные технологии **Том 11, Специальный выпуск, 2006**

РАСПРЕДЕЛЕНИЕ ПРОЦЕССОРНОГО ВРЕМЕНИ И ПАМЯТИ С ПОМОЩЬЮ ЦЕН

С. В. Бредихин
Институт вычислительной математики и математической геофизики СО РАН, Новосибирск, Россия
И. А. Вялков
Новосибирский государственный университет, Россия
А. Б. Хуторецкий
Новосибирский государственный педагогический университет, Россия
e-mail: bred@nsc.ru, hab@dus.nsc.ru

The distributed information-computing system is considered as a market of two commodities, namely processor time and the disc memory. Key players are then the suppliers and the consumers (users). A model of this market is proposed. The consumer's problem, that is the utility maximization under the given prices subject to the budget restriction, has been solved. A mechanism for the price adjustment of the supply and demand is described.

Введение

Перечень услуг, предоставляемых распределенными информационно-вычислительными системами (РИВС), постоянно расширяется. Соответственно растет число потенциальных пользователей РИВС и увеличивается полезность, которую мог бы получить каждый пользователь. С выходом РИВС за корпоративные рамки возникает проблема финансирования эксплуатации и развития системы.

Мы будем говорить о платных РИВС (ПРИВС), финансируемых за счет пользователей. Платность услуг регулирует предложение (стимулируя владельцев предоставлять свободные ресурсы для сетевого использования) и спрос (стимулируя пользователей откладывать решение задач с низким приоритетом и высокой стоимостью). Цены ресурсов порождают честные правила доступа к сетевым ресурсам для всех, устраняют необходимость централизованной координации при переговорах, позволяют поставщикам и пользователям принимать решения, максимизирующие полезность [1].

Система управления ресурсами ПРИВС должна руководствоваться в первую очередь не техническими критериями (как среднее время пребывания задания в системе), а стремлением каждого поставщика к максимизации прибыли и каждого пользователя — к максимизации полезности. Другими словами, ПРИВС является субрынком рынка информационно-вычислительных ресурсов, а система управления ресурсами (СУР) обеспечивает ценовое согласование спроса и предложения на этом субрынке, т. е. возможность (хотя бы принципиальную) достижения равновесия, установления таких цен, при которых спрос на ресурсы близок к предложению и равен предложению, если равновесие существует. Следовательно, минимальная “рыночная” роль СУР состоит в создании инфраструктуры, позволяющей агентам спроса и агентам предложения “находить друг друга” и договариваться об объемах и ценах сделок. Такой подход предложен, например, в [2 – 4].

Многие исследователи отводят СУР более активную роль. Они исходят из того, что каждого агента рынка представляет программа-брокер, которая “знает” бюджетное ограничение агента и модель его предпочтений на множестве ресурсных наборов. Взаимодействуя с программами-брокерами, СУР может вычислить цены равновесия (если оно существует). Затем программы, представляющие агентов рынка, определяют спрос и предложение при этих ценах. Если удастся (децентрализованно или с участием СУР) выбрать оптимальные ресурсные наборы пользователей так, чтобы они попарно не пересекались, то получим равновесное распределение ресурсов. Мы рассмотрим модель распределения процессорного времени и памяти [5, 6], модифицируем эту модель, чтобы устранить некоторые неточности, предложим подход к вычислению цен статического равновесия, основанный на работе [7], и обсудим проблемы, связанные с применением указанного подхода.

1. Исходная модель

Кратко опишем модель рынка, предложенную в [5, 6]. Потребитель формулирует спрос, указывая “работы”. Для каждой работы фиксированы объемы и длительности использования требуемых ресурсов. Потребитель имеет бюджет, который периодически пополняется. Однако деньги, не потраченные в данном периоде, не могут быть использованы для оплаты ресурсов в следующем периоде. В результате стимулируются не экономия средств и первоочередное выполнение самых важных работ, а полное расходование текущего бюджета и использование системных ресурсов для выполнения малозначимых работ в конце бюджетного периода.

Пользователь, желающий приобрести ресурс, указывает только потребный объем, но не длительность использования. Поставщик предоставляет купленное количество ресурса по текущей цене на весь период решения задачи [5, p. 757]. Это значит, что заключается неправдоподобный форвардный контракт с неопределенным сроком на поставку ресурса по фиксированной цене.

Авторы не описывают функцию полезности потребителя, заменяя ее следующим алгоритмом вычисления индивидуальной функции спроса при текущих ценах $\mathbf{p}$ (прямым полужирным шрифтом мы указываем векторные величины).

В каждый момент времени потребитель имеет очередь невыполненных работ. В начале единичного периода $[t; t+1)$ он вычисляет

$$C_1 = \frac{\sum\limits_i w_i p_i}{t} \quad \text{и} \quad C_2 = \frac{I(t)}{T_b - t}\ ,$$

где w_i — общее количество ресурса i, использованное потребителем с некоторого начального момента 0; $p_i(t)$ — текущая цена ресурса i; $I(t)$ — текущий бюджет потребителя; T_b — ближайший момент пополнения бюджета.

C_1 — это средние (за единицу времени) затраты, которые понес бы потребитель в периоде $[0; t)$, если бы всегда приобретал ресурсы по текущим ценам; C_2 — максимальные затраты за единицу времени, возможные в периоде $[t; T_b)$ при равномерном расходовании средств. Если $C_1 > C_2$, то спрос равен нулю, в противном случае пользователь включает в свой спрос периода $[t, t+1)$ “столько работ из очереди, сколько может себе позволить” [5]. Возникает вопрос: отражает ли так определенная функция спроса рациональные предпочтения пользователя?

Предположим, что на рынке только один ресурс и первая в очереди работа требует 10 единиц этого ресурса. Допустим, что $I(1) = 30$, $T_b = 3$ и за период $[0; 1)$ пользователь приобрел 10 единиц ресурса. Тогда в момент 1 при цене 2 имеем $C_1 = 20 > C_2 = 15$, спрос равен нулю. После этого $I(2) = 30$ и за период $[0; 2)$ пользователь приобрел те же 10 единиц ресурса. В момент 2 при цене 3 имеем $C_1 = 15 < C_2 = 30$, и работа включается в спрос. Пользователь, следовательно, предпочитает выполнить работу позже и заплатить за нее больше. Такую систему предпочтений трудно назвать разумной.

На рынке два ресурса — процессорное время и дисковая память. Поставщик памяти предназначает для продажи некоторое фиксированное количество S *файлов* определенного размера. Его функция полезности не указана, а функция предложения вычисляется следующим образом. В момент t поставщик вычисляет свой средний доход от одного файла в единицу времени за период $[0; t)$:

$$C = \frac{R}{tS}, \tag{1}$$

где R — суммарный доход, полученный этим поставщиком за тот же период. Если текущая цена меньше C, то предложение равно нулю, в противном случае поставщик в периоде $[t; t+1)$ предлагает ПРИВС все предназначенные для продажи файлы.

Единица измерения процессорного времени *слот* — это порция, какое-то количество α циклов в единицу времени, доля скорости процессора. Поставщик предназначает для продажи фиксированное число N слотов. Приобретая один слот, пользователь получает право на α циклов процессора в каждую единицу времени за период выполнения работы. Функция предложения описана следующим образом. Поставщик в момент t вычисляет

$$C = \frac{R}{tN},$$

где R определено, как в (1). Если текущая цена меньше C, то предложение равно нулю, в противном случае поставщик в периоде $[t; t+1)$ предлагает для продажи N слотов.

Легко привести пример нерационального поведения поставщика (например, поставщика памяти), имеющего описанную функцию предложения. Пусть $t = 1$, $S = 10$. Допустим, что выручка R поставщика за период $[0; 1)$ равна 100. Тогда $C = 10$ и при цене 9 предложение равно нулю. В момент 2 предшествующий доход по-прежнему равен 100 и $C = 5$, поэтому поставщик предоставит системе свой ресурс при цене, скажем, 6. Другими словами, он предпочитает продавать не раньше и дороже, а позже и дешевле.

2. Уточненная модель

Модифицируем исходную модель рынка, устраняя отмеченные выше ее недостатки. Рассматриваем временной период $[1; T]$, где T достаточно велико. Единичный период $[\tau - 1; \tau]$ будем называть “моментом τ”.

2.1. Агенты и товары

На рынке $I = IU + IC + IM$ агентов: пользователи с номерами $1, \ldots, IU$; поставщики процессорного времени с номерами $IU + 1, \ldots, IU + IC$; поставщики памяти с номерами $IU + IC + 1, \ldots, I$. Ресурсам присвоим номера $1, \ldots, J = IC + IM$, ресурс j принадлежит поставщику $j + IU$. Поставщик s, $IU < s \le I$, в каждый момент владеет $V(s)$ единицами соответствующего ресурса. Применительно к поставщикам машинного времени это означает, что процессоры, включенные в ПРИВС, имеют, вообще говоря, разные тактовые частоты и, соответственно, обеспечивают разные количества слотов в единицу времени.

Для дальнейшего нам удобно "индивидуализировать" ресурсные единицы: присвоим единицам ресурса, принадлежащего поставщику s, номера $k \in \{1, \ldots, V(s)\}$; k-ю единицу ресурса j в момент τ будем называть "единичным ресурсом (j, k, τ)". Любой распределенный во времени ресурсный набор A можно теперь представить *характеристическим вектором* $\mathbf{x} = \mathbf{x}(A) = (x_{jk}(\tau) \mid 1 \le j \le J,\ 1 \le k \le V(j + IU), 1 \le \tau \le T)$, где $x_{jk}(\tau) = 1$, если единичный ресурс (j, k, τ) входит в A, иначе $x_{jk}(\tau) = 0$. Пусть $\mathbf{x}$ представляет в указанном смысле набор A и $DR_{j\tau}(\mathbf{x})$ — количество ресурса j на момент τ, включенное в A, тогда

$$DR_{j\tau}(\mathbf{x}) = \sum_{k=1}^{V(j+IU)} x_{jk}(\tau).$$

2.2. Поставщики

Предоставление пользователю единицы ресурса на единичный период сопряжено для поставщика s с затратами $c_s > 0$. Он выбирает вектор $\mathbf{q}^s = (q^s_\tau)_{\tau=1}^T$ (q^s_τ — объем предложения в момент τ) и число m_s (спрос на деньги), $q^s_\tau \le V(s)$. *Функция полезности поставщика s* описывает его прибыль за рассматриваемый период:

$$U_s(\mathbf{q}^s, m_s) = m_s - c_s \sum_\tau q^s_\tau. \tag{2}$$

При ценах $\mathbf{p} = (p_j(\tau) \mid 1 \le j \le J,\ 1 \le \tau \le T)$ поставщик s имеет *бюджетное ограничение*

$$m_s \le \sum_\tau p_j(\tau) q^s_\tau, \tag{3}$$

где $j = s - IU$, т. е. полученная сумма не может превышать выручку от продажи ресурса.

Можно считать, что, предлагая пользователям q^s_τ единиц ресурса в момент τ, поставщик тем самым предъявляет спрос на остальные $V(s) - q^s_\tau$ единиц. Двоичный вектор $\mathbf{x}$ подходящей размерности является характеристическим вектором для одного из определенных таким образом ресурсных наборов поставщика s, если и только если для всех τ:

$$DR_{j\tau}(\mathbf{x}) \in [0; V(s)] \ \text{ при } j = s - IU, \ \text{ иначе } DR_{j\tau}(\mathbf{x}) = 0 \tag{4}$$

(поставщик предъявляет спрос только на свой ресурс). Пусть X_s — множество всех векторов, удовлетворяющих (4).

2.3. Потребители (пользователи)

Каждый пользователь i хочет выполнить одно задание, которое требует M_i единиц памяти (файлов) в каждый момент от запуска до завершения и K_i единиц процессорного времени (слотов). Набор файлов и слотов, используемых заданием, не меняется в течение единичного периода, но может различаться в разные периоды. Смена ресурсов, если она нужна, происходит на стыке единичных периодов мгновенно и бесплатно. В каждый момент задание может использовать файлы, принадлежащие разным поставщикам, но только один процессор.

Пользователь i выделяет на выполнение своего задания сумму B_i (бюджет); он выбирает набор ресурсов, соответствующий некоторому характеристическому вектору $\mathbf{x}^i$, и число m_i (спрос на деньги). Нуль-вектор $\mathbf{x}^i = \mathbf{0}$ означает отказ от выполнения задания. Вектор $\mathbf{x}^i$ задает спрос $DC_\tau(\mathbf{x}^i)$ на процессорное время и спрос $DM_\tau(\mathbf{x}^i)$ на память в каждый момент τ:

$$DC_\tau(\mathbf{x}^i) = \sum_{j=1}^{IC} DR_{j\tau}(\mathbf{x}^i), \quad DM_\tau(\mathbf{x}^i) = \sum_{j=IC+1}^{IM} DR_{j\tau}(\mathbf{x}^i).$$

Задание i выполняется в момент τ по плану $\mathbf{x}^i$, если оно использует процессор, $DC_\tau(\mathbf{x}^i) > 0$. Опишем множество X_i векторов $\mathbf{x}^i$, соответствующих допустимым для пользователя i ресурсным наборам. Единичные периоды, в которые выполнение задания i по плану $\mathbf{x}^i$ начинается и кончается, будем обозначать $t^0(\mathbf{x}^i)$ и $t^1(\mathbf{x}^i)$ соответственно; $t^1(\mathbf{x}^i) = +\infty$ означает, что выполнение задания не завершается в периоде $[0;T]$.

1. Если $\sum_\tau DC_\tau(\mathbf{x}^i) = 0$ (задание не выполняется в периоде $[0;T]$), то $\mathbf{x}^i \in X_i$; положим $t^1(\mathbf{x}^i) = +\infty$.

2. Пусть $\sum_\tau DC_\tau(\mathbf{x}^i) > 0$. Тогда определено значение $t^0(\mathbf{x}^i) = \min\{\tau \mid DC_\tau(\mathbf{x}^i) > 0\}$.

2.1. Если не существует $t \in [t^0(\mathbf{x}^i);T]$, для которого выполнены условия

$$\sum_{\tau=t^0(\mathbf{x}^i)}^{t} DC_\tau(\mathbf{x}^i) \geq K_i, \tag{5}$$

$$DM_\tau(\mathbf{x}^i) \geq M_i \qquad \text{для всех } \tau \in [t^0(\mathbf{x}^i);t], \tag{6}$$

$$DR_{j\tau}(\mathbf{x}^i) > 0 \to DR_{k\tau}(\mathbf{x}^i) = 0 \qquad \text{для всех } j \in \{1,\dots,IC\},\ k \in \{1,\dots,IC\} \setminus \{j\}, \tau \in [t^0(\mathbf{x}^i),t], \tag{7}$$

то $\mathbf{x}^i \notin X_i$; положим $t^1(\mathbf{x}^i) = +\infty$. Условие (5) означает, что задание получает достаточное процессорное время в период $(t^0(\mathbf{x}^i)-1;t]$; условие (6) требует, чтобы задание было обеспечено памятью в указанном периоде; (7) запрещает заданию использовать более одного процессора в каждый момент $\tau \in [t^0(\mathbf{x}^i);t]$.

2.2. Существуют $t \in [t^0(\mathbf{x}^i);T]$, удовлетворяющие (5)–(7). Тогда $\mathbf{x}^i \in X_i$; положим $t^1(\mathbf{x}^i)$ равным наименьшему из таких t.

Функция полезности пользователя i квазилинейна и имеет вид

$$U_i(\mathbf{x}^i, m_i) = u_i(\mathbf{x}^i) + m_i, \tag{8}$$

где $u_i(\mathbf{x}^i) = g_i(t^1(\mathbf{x}^i))$, а $g_i(\tau)$ — денежная оценка полезности для пользователя завершения задания в момент τ. Будем считать, что функция $g_i(\tau)$ монотонно убывает и стремится к нулю при $\tau \to +\infty$.

При ценах $\mathbf{p} = (p_j(\tau) \mid 1 \leq j \leq J,\ 1 \leq \tau \leq T)$ пользователь i имеет *бюджетное ограничение*

$$m_i + \sum_{j,\tau} DR_{j\tau}(\mathbf{x}^i) p_j(\tau) \leq B_i. \tag{9}$$

2.4. Равновесия

Поставщик s при ценах $\mathbf{p}$ решает задачу максимизации функции (2) при условиях (3) и $0 \le q_\tau^s \le V(s)$ для всех τ. Пусть $j = s - IU$. В оптимуме ограничение (3) активно, $m_s = \sum\limits_\tau p_j(\tau) q_\tau^s$. Задача принимает вид

$$\sum_\tau (p_j(\tau) - c_s) q_\tau^s \to \max \quad \text{при } 0 \le q_\tau^s \le V(s) \ \forall \tau. \tag{10}$$

Каждому вектору $\mathbf{q}^s$ можно сопоставить ресурсный набор поставщика, которому, в свою очередь, соответствует характеристический вектор $\mathbf{x}^s \in X_s$. Тогда

$$\sum_\tau (p_j(\tau) - c_s) q_\tau^s = \sum_\tau (p_j(\tau) - c_s)(V(s) - DR_{j\tau}(\mathbf{x}^s)) = C + u_s(\mathbf{x}^s) - \sum_\tau p_j(\tau) DR_{j\tau}(\mathbf{x}^s),$$

где

$$C = \sum_\tau (p_j(\tau) - c_s) V(s), \quad \text{а} \quad u_s(\mathbf{x}^s) = c_s \sum_\tau DR_{j\tau}(\mathbf{x}^s).$$

Поэтому задача (10) эквивалентна следующей:

$$v_s(\mathbf{x}^s, \mathbf{p}) = u_s(\mathbf{x}^s) - \sum_\tau p_j(\tau) DR_{j\tau}(\mathbf{x}^s) \to \max \quad \text{при } 0 \le DR_{j\tau}(\mathbf{x}^s) \le V(s) \ \forall \tau. \tag{11}$$

Множество всех решений этой задачи обозначим $D_s(\mathbf{p})$. Тогда $\mathbf{p} \mapsto D_s(\mathbf{p})$ есть *отображение спроса для поставщика* $s \in \{IU + 1, \dots, I\}$.

Пользователь i при ценах $\mathbf{p}$ решает задачу максимизации функции (8) при условиях (3) и $\mathbf{x}^i \in X_i$. В оптимуме должно быть $m_i = B_i - \sum\limits_{j,\tau} DR_{j\tau}(\mathbf{x}^i) p_j(\tau)$. После подстановки m_i в (8) и исключения константы B_i получим следующую задачу:

$$v_i(\mathbf{x}^i, \mathbf{p}) = u_i(\mathbf{x}^i) - \sum_{j,\tau} DR_{j\tau}(\mathbf{x}^i) p_j(\tau) \to \max \quad \text{при } \mathbf{x}^i \in X_i. \tag{12}$$

Множество всех решений этой задачи обозначим $D_i(\mathbf{p})$. Тогда $\mathbf{p} \mapsto D_i(\mathbf{p})$ есть *отображение спроса для пользователя* $i \in \{1, \dots, IU\}$.

Пара $(\mathbf{x}, \mathbf{p})$ является *равновесием*, если $\mathbf{x} = (\mathbf{x}^i)_1^I \in D_1(\mathbf{p}) \times \dots \times D_I(\mathbf{p})$ и $\sum_i \mathbf{x}^i = \mathbf{e}$, где $\mathbf{e}$ — вектор, все компоненты которого равны единице. В этом случае $\mathbf{x}$ будем называть *равновесным распределением* ресурсов, а компоненты вектора $\mathbf{p}$ — *ценами равновесия*.

Другими словами, равновесное распределение есть набор векторов спроса $\mathbf{x}^i$, по одному для каждого агента рынка, причем $\mathbf{x}^i$ при ценах равновесия доставляет агенту i максимум полезности в рамках бюджетного ограничения, а каждая единица ресурса в каждый момент предназначена ровно одному агенту (пользователю или владельцу этого ресурса).

3. Механизм ценовой адаптации рынка

Выделяют два типа механизмов перехода рынка к равновесию: *нащупывающие* (*tâtonnement*) и *ненащупывающие*. Нащупывающий механизм состоит в следующем. Агенты рынка непосредственно или через некоторое центральное звено, “вальрасова аукциониста”, обмениваются информацией о спросе, предложении и ценах [8, p. 3 и 9, p. 746]. Ситуация, в которой ни один агент рынка не хочет изменить свое сообщение, является равновесием [8, p. 4]. Сделки происходят только в равновесии. В частности, процесс Вальраса — Самуэльсона предполагает, что агенты рынка, зная вектор текущих цен $\mathbf{p}$, сообщают аукционисту свои векторы спроса/предложения. Аукционист вычисляет вектор $\mathbf{E}(\mathbf{p})$ агрегированного избыточного спроса (спрос минус предложение); он повышает цену товара i, если $E_i(\mathbf{p}) > 0$ (спрос больше предложения), и понижает, если $E_i(\mathbf{p}) < 0$. Один из вариантов такого процесса в непрерывном времени описывается векторным дифференциальным уравнением Эйлера

$$\dot{\mathbf{p}} = \mathbf{E}(\mathbf{p}). \tag{13}$$

Пусть $\mathbf{p}(t)$ — решение уравнения (13). Предел $\mathbf{p}(t)$ при $t \to +\infty$ существует и является вектором цен равновесия [10, 11], если функция $\mathbf{E}$ удовлетворяет условию “валовой заменимости” (*gross substitutes*):

$$\frac{\partial E_i}{\partial p_i} < 0 \ \text{ для всех } i, \quad \frac{\partial E_i}{\partial p_j} > 0 \ \text{ при } i \neq j. \tag{14}$$

Условие (14) означает, что рынки разных товаров сильно взаимосвязаны: повышение цены на рынке товара i вызывает мгновенный отток денег с этого рынка, причем какая-то часть этих денег попадает на рынок каждого товара $j \neq i$. Если (14) не выполнено, то процесс (13), вообще говоря, не имеет предела [12]. В рассматриваемой модели ресурсы не заменяемы, а дополнительны, (14) наверняка не выполняется. Поэтому в [5, 6] предложено использовать более универсальную схему Смэйла (S. Smale):

$$\dot{\mathbf{p}} \cdot \mathbf{J_E}(\mathbf{p}) = -\lambda \mathbf{E}(\mathbf{p}), \tag{15}$$

где $\mathbf{J_E}(\mathbf{p})$ — якобиан вектор-функции $\mathbf{E}(\mathbf{p})$. Если функция избыточного спроса $\mathbf{E}$ дважды непрерывно дифференцируема и удовлетворяет еще некоторым техническим условиям, то процесс (15) сходится к вектору цен равновесия [13].

Процессы (13) и (15) являются нащупывающими: агенты рынка заключают и перезаключают контракты и обмениваются сигналами, но не производят и не потребляют, пока не установятся цены, уравнивающие спрос и предложение. Только после этого осуществляются сделки, приносящие каждому агенту рынка максимальную полезность в рамках его бюджетного ограничения.

Ненащупывающий механизм, напротив, допускает сделки при ненулевом избыточном спросе по текущим ценам при условии, что рынок упорядочен (*orderly*), “короткая” сторона рынка удовлетворена. Это значит, что после завершения всех сделок (возможных и желательных для участников при текущих ценах) для каждого товара либо спрос равен нулю, либо предложение равно нулю. Избыточный спрос (ненулевой, если цены не уравновешивают рынок) при этом, конечно, не изменяется, а неудовлетворенность одной из сторон рынка стимулирует изменения цен. Такой механизм не требует участия аукциониста, если ценами управляют владельцы товаров. Ненащупывающий механизм более реалистичен, но сделки, осуществленные вне равновесия, могут оказаться не оптимальными *ex post*, когда цены равновесия будут достигнуты. Различие между нащупывающими и ненащупывающими процессами адаптации цен можно сформулировать следующим образом: в первом случае по ценам равновесия осуществляются все сделки, а во втором, вообще говоря, только заключительные, “очищающие” рынок.

3.1. О применении схемы Смэйла в работах [5, 6]

Нестабильность спроса вынуждает авторов указанных работ строить динамическое равновесие. В таком случае вектор цен $\mathbf{p}$ должен иметь структуру $(\mathbf{p}^1, \ldots, \mathbf{p}^T)$, где $\mathbf{p}^\tau$ — вектор цен в момент τ (мы рассматривали такой вектор цен в подразд. 2.4). Этого требует и применяемый авторами механизм (15) (см. [14]). Агенты рынка при фиксированном векторе $\mathbf{p}$ должны формулировать спрос на ресурсы и предложение ресурсов для каждого $\tau \in \{1, \ldots, T\}$. И адаптация цен происходит одновременно для всех единичных периодов. Это особенно важно для рынка рассматриваемого типа, где спрос и предложение следующего периода зависят от распределения ресурсов в предыдущих периодах. Нам кажется (не можем утверждать с уверенностью, так как в указанных работах процедура описана недостаточно подробно), что авторы отождествляют реальное время τ с параметром t процесса (15), когда применяют его дискретный вариант

$$\Delta\mathbf{p}(t) \cdot \mathbf{J}_\mathbf{E}(\mathbf{p}(t)) = \mathbf{E}(\mathbf{p}(t)). \tag{16}$$

Это косвенно подтверждается тем, что в [5, с. 760] сказано: "…модель Smale'a предполагает функции спроса и предложения независимыми от времени …". Это неверно, в динамической модели рынка товар в разных периодах рассматривается как разные товары, и цены на все такие товары входят в вектор $\mathbf{p}$. По-видимому, авторы по ценам момента t вычисляют избыточный спрос в этот момент и находят цены момента $t+1$, решая систему уравнений (16). Однако избыточный спрос момента t зависит от цен и объемов сделок, осуществленных в предшествующие моменты (см. разд. 1), когда цены равновесия не были достигнуты (в противном случае избыточный спрос был бы равен нулю). Следовательно, ресурсы распределяются и используются вне равновесия. Это значит, что реализуется некий ненащупывающий механизм, но не механизм Смэйла.

Кроме того, вычисление производных функции избыточного спроса на дискретные ресурсы [5, с. 760] выглядит неубедительно, и вряд ли существует удовлетворительное решение этой проблемы. Поэтому мы предлагаем использовать не требующую производных процедуру адаптации цен, которая описана ниже.

3.2. Механизм J. Ma и F. Nie

Нетрудно проверить, что модель рынка, сформулированная в разд. 2, является частным случаем модели рынка неделимых товаров, описанной в [7]. В этой статье обоснован механизм ценового согласования спроса и предложения (*MN-механизм*), который обобщает дискретную форму схемы Эйлера (13) и состоит в следующем.

Зафиксируем произвольно число $b > 0$ и разобьем отрезок $[0, b]$ на n равных промежутков длины $\Delta_n = b/n$ точками $t_0 = 0, t_1, \ldots, t_n = b$ (равномерность разбиения несущественна, мы предполагаем ее для упрощения). Зафиксируем также произвольный вектор цен $\mathbf{p}_0$ и положим $\mathbf{p}(0) = \mathbf{p}_0$. Для всех $l \in \{1, \ldots, n\}$ вычисляем

$$\mathbf{p}(t_l) = \mathbf{p}(t_{l-1}) + \Delta_n[\mathbf{f}(\mathbf{p}(t_{l-1})) - \mathbf{e}], \tag{17}$$

где $\mathbf{f}(\mathbf{p})$ — любая выпуклая линейная комбинация оптимальных ресурсных наборов агента i при ценах $\mathbf{p}$ и все компоненты вектора $\mathbf{e}$ равны единице. Для рынка рассматриваемого типа функция $\mathbf{f}(\mathbf{p})$ обобщает функцию избыточного спроса.

Завершив вычисления по формуле (17), получим вектор-функцию $\mathbf{p}_n(t)$, определенную в точках t_l. В [7] доказаны следующие результаты.

1. При $n \to \infty$ последовательность функций $\mathbf{p}_n(t)$ равномерно сходится к некоторой функции $\mathbf{p}(t)$, определенной на промежутке $[0; b]$.

2. При любом выборе $\mathbf{p}_0$ и $\mathbf{f}(\cdot)$ в (17) вектор $\mathbf{p}(b)$ в определенном смысле близок к вектору цен равновесия настолько, насколько возможно.

3. Если равновесие существует, то $\mathbf{p}^*$ — вектор цен равновесия.

Таким образом, последовательность $\mathbf{p}_n(b)$ в любом случае сходится к "хорошей" системе цен $\mathbf{p}(b)$. Однако использование MN-механизма для распределения ресурсов сопряжено с серьезными техническими трудностями, которые мы обсудим применительно к модели рынка, описанной в разд. 2 (назовем ее "модель G1").

4. О применении MN-механизма к модели G1

В MN-механизме каждая единица ресурса рассматривается как самостоятельный "единичный ресурс". Поэтому вектор цен должен иметь вид $\mathbf{p} = (p^\tau_{jk})_{j,k,\tau}$, где p^τ_{jk} — цена единичного ресурса (j, k, τ). Допуская, что в модели G1 разные единицы ресурса j в момент τ могут иметь разные цены, получим "дезагрегированную" модель, которую обозначим DG1. Задачи (11) и (12) для модели DG1 имеют вид

$$v_s(\mathbf{x}^s, \mathbf{p}) = u_s(\mathbf{x}^s) - \sum_{k,\tau} p^\tau_{jk} x^s_{jk}(\tau) \to \max \quad \text{при } 0 \le \sum_k x^s_{jk}(\tau) \le V(s)\ \forall \tau,\ j = s - IU \tag{18}$$

и, соответственно,

$$v_i(\mathbf{x}^i, \mathbf{p}) = u_i(\mathbf{x}^i) - \sum_{j,k,\tau} x^i_{jk}(\tau) p^\tau_{jk} \to \max \quad \text{при } \mathbf{x}^i \in X_i. \tag{19}$$

Теорема. *Если $(\mathbf{x}, \mathbf{p})$ — равновесие в модели DG1, то существует уравновешивающий распределение $\mathbf{x}$ вектор цен $\hat{\mathbf{p}}$, для которого $\hat{p}^\tau_{jk}$ не зависит от k (цены всех единиц ресурса j в момент τ равны).*

Доказательство теоремы [15] дает конструктивный способ построения системы цен $\hat{\mathbf{p}}$. Понятно, что это цены равновесия для модели G1. Иными словами, если DG1 имеет равновесие, то и G1 имеет равновесие. Обратное, конечно, тоже верно. Однако не всякое равновесие для DG1 является равновесием для G1.

Уравнения системы (17) для модели DG1 имеют вид

$$p^\tau_{jk}(t_l) = p^\tau_{jk}(t_{l-1}) + \Delta_n[f^\tau_{jk}(\mathbf{p}(t_{l-1})) - \mathbf{e}], \tag{20}$$

где $f^\tau_{jk}(\mathbf{p})$ — соответствующая компонента вектор-функции $\mathbf{f}(\mathbf{p})$.

4.1. Варианты выбора функции $\mathbf{f}(\mathbf{p})$

Самый простой способ фиксации функции $\mathbf{f}$ состоит в следующем: каждому $\mathbf{p}$ сопоставляем $\mathbf{x}^i(\mathbf{p}) \in D_i(\mathbf{p})$ для всех $i \in \{1, \ldots, I\}$ и полагаем $\mathbf{f}(\mathbf{p}) = \mathbf{f}_1(\mathbf{p}) = \sum\limits_i \mathbf{x}^i(\mathbf{p})$. Тогда для реализации процедуры (17) нужно уметь при любом $\mathbf{p}$ находить хотя бы одно решение задач (18) и (19).

При таком выборе функции $\mathbf{f}$ цены разных единиц одного ресурса в один и тот же момент могут различаться как в ходе процесса, так и в пределе. Если равновесие существует и соответствующие цены будут построены, то мы сможем их "подправить" согласно теореме. Если равновесия нет или мы не нашли хорошее приближение к вектору цен равновесия, то все равно две единицы одного ресурса в один момент эквивалентны для всех агентов рынка и желательно, чтобы их цены совпадали. Мы предлагаем обеспечить такое совпадение на всех итерациях MN-механизма, выбрав функцию $\mathbf{f}$ следующим образом.

Для любого $\mathbf{p}$ положим $\mathbf{x}^i(\mathbf{p})$ равным среднему арифметическому всех элементов из $D_i(\mathbf{p})$ (это конечное множество) и $\mathbf{f}(\mathbf{p}) = \mathbf{f}_2(\mathbf{p}) = \sum_i \mathbf{x}^i(\mathbf{p})$. Тогда $f^\tau_{jk}(\mathbf{p})$ в (20) равно частоте ν^τ_{jk} вхождений единичного ресурса (j, k, τ) в наборы из $D_i(\mathbf{p})$. Следовательно, для реализации MN-механизма достаточно на каждой итерации знать число оптимальных наборов $n(i)$ для любого агента рынка и число $m(i, j, k, \tau)$ оптимальных наборов, включающих единичный ресурс (j, k, τ) (для всех i, j, k, τ).

Предположим, что вектор цен $\mathbf{p}(t_{l-1})$ обеспечивает *единую цену* ресурса j в момент τ ($p^\tau_{jk}(t_{l-1})$ не зависит от k) для всех j, τ. Тогда замена единичного ресурса (j, k_1, τ) единичным ресурсом (j, k_2, τ) не изменяет полезность набора для всех агентов рынка. Это значит, что величины $m(i, j, k, \tau)$ не зависят от k; тогда и $f^\tau_{jk}(\mathbf{p})$, и $\mathbf{p}(t_l)$ в (20) не зависят от k. Другими словами, если $\mathbf{f}(\cdot) = \mathbf{f}_2(\cdot)$ и исходный вектор цен $\mathbf{p}(0)$ обеспечивает единые цены, то они поддерживаются на всех итерациях. В частности, можно положить $p^\tau_{jk}(0) = c_s$ для $s = j + IU$ (цена каждой единицы ресурса равна отправной цене поставщика). В этом случае все уравнения системы (20), соответствующие фиксированным j и τ, совпадут, достаточно сохранить только одно из них.

4.2. Об определении оптимальных ресурсных наборов для агентов рынка

Речь идет о решении задач (18) и (19). Из предыдущего раздела следует, что при реализации MN-механизма нас может интересовать один элемент из $D_i(\mathbf{p})$, $1 \le i \le I$, или частота вхождений каждого единичного ресурса в наборы из $D_i(\mathbf{p})$. Кроме того, когда цены, по которым будут распределяться ресурсы, найдены, желательно иметь полные описания множеств $D_i(\mathbf{p})$.

Сначала рассмотрим задачу (18), которая легко решается в общем виде. Зафиксируем s и $\mathbf{p}$, положим $j = s - IU$. Подставив в (18) выражение для u_s (см. подразд. 1.2.1), приведем ее к виду

$$\sum_{k,\tau} (c_s - p^\tau_{jk}) x^s_{jk}(\tau) \to \max \quad \text{при } 0 \le \sum_k x^s_{jk}(\tau) \le V(s)\ \forall \tau.$$

Пусть K^+, K^- и K^0 — множества единичных ресурсов (j, k, τ), таких, что $p^\tau_{jk} > c_s$, $p^\tau_{jk} < c_s$ и $p^\tau_{jk} = c_s$ соответственно. Тогда каждый ресурсный набор из $D_s(\mathbf{p})$ не включает единичные ресурсы из K^+, а включает все единичные ресурсы из K^- и некоторое количество ресурсов из K^0. Число таких наборов $2^{|K^0|}$, причем каждый единичный ресурс из K^0 входит в половину из них. Следовательно, ν^τ_{jk} равно 0, 1 или 0.5, если единичный ресурс (j, k, τ) принадлежит K^+, K^- или K^0 соответственно.

Перейдем к задаче (19). В общем случае это, по-видимому, громоздкая комбинаторная задача. Однако она существенно упрощается, если предположить, например, что каждое задание либо не выполняется, либо выполняется с постоянной скоростью (и без перерывов). Мы сделаем другое упрощающее предположение: все процессоры одинаковы и каждый из них может предоставить только один слот в единицу времени. Это значит, что $V(s) = 1$ для $IU < s \le IU + IC$.

В [15] описан алгоритм решения задачи (19) при указанных предположениях. Этот алгоритм фактически перебирает целочисленные пары (a, b), $1 \le a \le b \le T$, с некоторыми усечениями. Он позволяет за приемлемое время найти один из оптимальных наборов. Однако подсчет числа оптимальных наборов и тем более перечисление всех таких наборов приводят, вообще говоря, к экспоненциальным переборам.

Заключение

Предположим, что с помощью MN-механизма удалось построить предельную систему цен $\mathbf{p}$. Следующая проблема — распределение ресурсов по этим ценам. Для этого нужно описать множества $D_i(\mathbf{p})$. Если все они одноэлементные, то равновесие существует [7, Theorem 8], а соответствующее распределение получим, выделив каждому агенту рынка i ресурсы в соответствии с единственным вектором спроса, входящим в $D_i(\mathbf{p})$. При многоэлементных $D_i(\mathbf{p})$ распределение ресурсов становится трудной комбинаторной задачей, даже если равновесие существует.

В литературе есть много примеров отсутствия равновесия на рынке неделимых ресурсов. Можно предположить, что и в нашей модели неделимость файлов и слотов сделает точное равенство спроса и предложения редким явлением. Поэтому постановка задачи распределения ресурсов при ценах $\mathbf{p}$ должна учитывать возможность того, что они не уравновешивают систему.

Проблему можно сформулировать следующим образом. Дана система $(D_i \mid i \in I)$ семейств подмножеств некоторого конечного множества (ресурсных единиц) M. Надо построить набор попарно не пересекающихся подмножеств множества M, содержащий не более одного представителя из каждого D_i и в каком-то смысле максимальный (например, по числу подмножеств, по мощности объединения или суммарному приоритету "представленных" номеров i). Если равновесие существует, то такой набор определит равновесное распределение.

Отметим, что предложенная выше модель не в полной мере отражает динамику работы ПРИВС. Обсуждавшиеся выше распределения ресурсов динамические в том смысле, что они планируют выполнение заданий во времени. Однако мы предполагали, что наборы заданий и единичных ресурсов не изменяются. В реальности это не так и приходится перераспределять ресурсы по мере появления/исчезновения пользователей и/или поставщиков. Очевидное предложение: система собирает заявки на выполнение заданий в период $[t; \infty)$ до момента $t - \varepsilon$ и распределяет этим заданиям ресурсы в течение промежутка $[t - \varepsilon; t)$. Но решение задачи распределения ресурсов может потребовать столько ресурсов, что нечего будет распределять. Возможно, процедура, подобная аукциону, более адекватна моделируемой ситуации.

Список литературы

[1] Buyya R., Abramson D., Giddy J. An economy driven resource management architecture for global computational power grids // Intern. Conf. on Parallel and Distributed Proc. Techniques and Applications (PDPTA 2000). CSREA Press, 2000.

[2] Buyya R., Abramson D., Giddy J., Stockinger H. Economic models for resource management and scheduling in grid computing // Concurrency and Computation: Practice and Experience. 2002. Vol. 14. P. 1507–1542.

[3] Buyya R., Giddy J., Abramson D. An evaluation of economy-based resource trading and scheduling on computational power grids for parameter sweep applications // 2nd Intern. Workshop on Active Middleware Services (AMS 2000). N.Y.: Kluwer Acad. Press, 2000.

[4] Buyya R., Murshed M. GridSim: a toolkit for the modeling and simulation of distributed resource management and scheduling for grid computing // Concurrency and Computation: Practice and Experience. 2002. Vol. 14. P. 1175–1220.

[5] Wolski R., Brevik J., Plank J.S., Bryan T. Grid resource allocation and control using computational economies // Grid Computing. Making the Global Infrastructure a Reality, Chapter 32. Chichester, England: John Wiley & Sons, 2003.

[6] Wolski R., Plank J.S., Bryan T., Brevik J. Analyzing market-based resource allocation strategies for the computational grid // Intern. J. of High Performance Computing Applications. 2001. Vol. 15(3).

[7] Ma J., Nie F. Walrasian equilibrium in an exchange economy with indivisibilities // Math. Social Sci. 2003. Vol. 46. P. 159–192.

[8] van den Elzen A.H. Adjustment Processes for Exchange Economies and Noncooperative Games (Lecture Notes in Economics and Mathematical Systems, 402). Berlin: Springer-Verlag, 1993.

[9] Hahn F. Stability // Handbook of Mathematical Economics. Vol. II (5th edition). Amsterdam: North-Holland, 1993.

[10] Arrow K., Hurwicz L. The stability of the competitive equilibrium I // Econometrica. 1958. Vol. 26. P. 522–552.

[11] Arrow K., Hurwicz L. The stability of the competitive equilibrium II // Econometrica. 1959. Vol. 27. P. 82–109.

[12] Scarf H. Some examples of global instability of the competitive equilibrium // Intern. Econ. Review. 1960. Vol. 1. P. 157–172.

[13] Smale S. A convergent process of price adjustment and global newton methods // J. of Math. Econ. 1975. Vol. 3. P. 107–120.

[14] Smale S. Dynamics in general equilibrium theory // Amer. Econ. Review. 1975. Vol. 66(2). P. 288–294.

[15] Бредихин С.В., Вялков И.А., Хуторецкий А.Б. Две модели ценового согласования при распределении вычислительных ресурсов // Сиб. журн. индустр. мат. 2006. Т. 1.

Поступила в редакцию 26 сентября 2006 г.