

ОПТИМИЗАЦИЯ СТОИМОСТИ ПРОГРАММНОГО ПРОЕКТА В УСЛОВИЯХ РЕАКТИВНОГО ПЕРЕПЛАНИРОВАНИЯ: ФОРМАЛЬНАЯ ПОСТАНОВКА ЗАДАЧИ

А.Б. Клименко

Дана постановка задачи, представлен краткий обзор ее аналогов и исследована возможность их применения к решению поставленной задачи с учетом особенностей процесса разработки программного обеспечения. Формализована задача оптимизации стоимости программного проекта.

Ключевые слова: планирование, неопределенность, расписание, комплектация, вычислительная система, оптимизация стоимости.

ВВЕДЕНИЕ

Реактивное перепланирование [1] представляет собой один из наиболее широко используемых подходов к планированию в условиях неопределенности. Реактивное перепланирование позволяет корректировать расписания выполнения работ в ответ на внешнее или внутреннее событие. Принятие решения относительно формирования дальнейшего расписания называется перепланированием, моменты времени, в которые происходит перепланирование — точками перепланирования.

В свою очередь, способ упорядочения решаемых задач проекта зачастую определяет эффективность функционирования системы: «... даже в простейших случаях существует большая разница в качестве функционирования системы, обусловленная только выбором очередности...» [2]. Можно утверждать, что, сводя стоимость проекта к стоимости эксплуатации исполнителей, можно достигнуть если не оптимальной стоимости проекта, то, по крайней мере, получить решение, приемлемое в аспекте финансовых вложений.

Сформулируем задачу оптимизации стоимости проекта в условиях реактивного перепланирования следующим образом: в точке перепланирования необходимо составить такое распределение нерешенных еще задач проекта с ограничением на следование и предварительно оцененной трудоемкостью по некоторому доступному числу исполнителей, чтобы проект был завершен не позднее ра-

нее запланированного срока, и при этом затраты на эксплуатацию исполнителей были минимальны. Иными словами, в точке перепланирования необходимо принять решение о том, сколько исполнителей потребуется для завершения проекта в указанный срок, а также составить план распределения задач проекта по исполнителям с учетом ограничения на их следование.

1. КРАТКИЙ ОБЗОР АНАЛОГОВ

Наиболее близкой по смыслу задачей можно считать задачу о комплектации вычислительной системы (ВС) минимальной стоимости, сформулированную и решенную А.Б. Барским [3]. Однако, прежде чем перейти к анализу возможности применения предложенного метода для решения задачи оптимизации стоимости программного проекта [3], перечислим некоторые значимые особенности процесса разработки программного обеспечения, отмеченные в работах [4—6].

- Увеличение числа исполнителей далеко не всегда ведет к росту общей производительности коллектива. В качестве основных причин приводятся рост накладных расходов на коммуникацию и существование ограничений на следование задач [5]. Широкую известность получило высказывание о том, что «добавление новых людей к запаздывающему проекту равносильно тушению огня керосином». Однако позже это высказывание было скорректировано следующим образом: добавле-



ние новых людей к проекту всегда ведет к его удорожанию, но не всегда к провалу — в зависимости от того, сколько новых людей и когда добавлено.

- Для новых исполнителей существует период так называемой ассимиляции в проекте [4, 6], на протяжении которого производительность исполнителя возрастает до определенного уровня.

- Как правило, после подключения к проекту новые исполнители обучаются. Для обучения принято выделять группу исполнителей из числа давно работающих над проектом (далее — обучающих экспертов). При этом производительность обучающих экспертов падает, так как каждый из них проводит некоторую долю рабочего времени за обучением новичков, и достигает прежнего уровня по мере перехода новичков в группу экспертов.

Трудоемкость задач, представленных к решению, предварительно оценена посредством одной из известных методик (например, СОСОМО [7]). Средняя производительность исполнителя также может быть оценена, что, в общем, не противоречит возможности ее изменения в ходе проекта.

Формализация задачи оптимизации стоимости проекта должна тем или иным образом отражать перечисленные особенности, а именно:

- существует связь «число исполнителей — производительность исполнителя»;
- имеется связь между составом подгрупп исполнителей и их производительностью;
- существует зависимость производительности исполнителей от времени их работы в проекте;
- задача характеризуется не временем ее решения, а трудоемкостью.

Проанализируем постановку задачи о комплектации ВС минимальной стоимости в аспекте выделенных особенностей процесса разработки программного обеспечения (см. таблицу).

Таким образом, невзирая на то, что решение задачи о комплектации ВС минимальной стоимости дает ответ на вопрос как о составе множества ис-

полнителей, так и о порядке распределения упорядоченных задач, оно не может использоваться для решения задачи оптимизации стоимости проекта в силу того, что не позволяет учесть выделенные особенности процесса разработки программного обеспечения. Это, в свою очередь, делает актуальным вопрос о формализации задачи оптимизации стоимости программного проекта.

2. ФОРМАЛИЗАЦИЯ ЗАДАЧИ ОПТИМИЗАЦИИ СТОИМОСТИ ПРОГРАММНОГО ПРОЕКТА

2.1. Принимаемые допущения

Для формализации и постановки задачи оптимизации стоимости программного проекта будем опираться на следующие допущения:

- с увеличением числа исполнителей фактическая производительность каждого из них, по крайней мере, не возрастает;

- множество всех исполнителей, работающих над проектом, можно разбить на три группы: группу «экспертов», группу «обучающих экспертов», группу «не-экспертов»;

- по прошествии периода ассимиляции в проекте «не-эксперт» обретает номинальную производительность, значение которой в дальнейшем уже не меняется на относительно больших промежутках времени;

- на протяжении периода ассимиляции в проекте «не-эксперт» отнимает время на свое обучение у исполнителей из группы «обучающих экспертов», тем самым снижая их номинальную производительность;

- в любой точке перепланирования исполнитель из группы «не-экспертов» должен быть переведен в группу «экспертов» или «обучающих экспертов» посредством принятия управленческого решения, если время его работы в проекте превышает период ассимиляции.

2.2. Формализация исходных данных

Формализуем исходные данные для задачи минимизации стоимости проекта. Назовем момент времени t_q точкой перепланирования. Пусть в текущей точке перепланирования известно следующее.

- Информационный граф, полученный путем декомпозиции главной задачи проекта на подзадачи: $G = (H, X, \Gamma, B)$, где $H = \{i\} = \{1, \dots, n\}$ — вершины графа, $X = \{x_i\}$ — предварительно оцененные трудоемкости задач, Γ — множество дуг, определяющих связи между задачами по информации, $B = \{b_i\}$ — отношение выполненной части задачи к общему объему задачи в момент времени t_q . Если данная задача еще не назначалась на выполнение,

Анализ задачи о комплектации ВС минимальной стоимости

| Признак | Наличие |
|--|---|
| Связь «число исполнителей — производительность исполнителя» | Не отражена |
| Связи между составом подгрупп исполнителей и их производительностью | Не отражены |
| Изменение производительности исполнителей от времени их работы в проекте | Не отражена |
| Задачи характеризуются трудоемкостью | Задачи характеризуются временем их решения на процессорах |

то $b_i = 0$. В случае, если задача уже назначена на выполнение и решается, значение b_i может быть равно, к примеру, 50 %. Если задача выполнена, то, соответственно, $b_i = 100$ %. Вершины, для которых значения $b_i = 100$ %, не рассматриваются в дальнейшем. Полагаем, что информационный граф не имеет контуров.

Употребление термина «информационный граф» подразумевает, что каждая из задач должна быть выполнена один раз и задачи связаны по информации.

• Множество исполнителей $M = \{m_j\}, j = 1, \dots, m$, принимающих участие в решении задач проекта: $m_j = \langle j^{exp}, pnom_j(t), s_j, t_j^{join}, h_j, g_j \rangle$, где j — порядковый номер исполнителя, подключенного к проекту, $pnom_j(t)$ — номинальная производительность исполнителя, s_j — стоимость эксплуатации исполнителя, t_j^{join} — момент времени, в который произошло присоединение исполнителя к множеству исполнителей, работающих над проектом, h_j — номер задачи, которая может находиться на решении у j -го исполнителя в точке перепланирования, g_j — переменная, определяющая принадлежность исполнителя к одной из групп исполнителей: $g_j = (\langle exp \rangle | \langle n_exp \rangle | \langle learn_exp \rangle)$, где exp — строка, наличие которой означает принадлежность исполнителя к группе исполнителей-экспертов; n_exp — строка, наличие которой означает принадлежность исполнителя к группе исполнителей-неэкспертов; $learn_exp$ — строка, наличие которой означает принадлежность исполнителя к группе исполнителей-экспертов, в чьи обязанности входит консультирование не-экспертов.

Вопрос принадлежности исполнителя к той или иной группе будет более подробно рассмотрен в дальнейшем.

• Множество исполнителей $M_{avail} = \{m_j\}, j = 1, \dots, k$, определяющее наличие исполнителей, для которых существует возможность подключения к проекту. Для любого $m_j \in M_{avail}$ полагаем $t_j^{join} = \infty$, $h_j = 0$, $g_j = \langle n_exp \rangle$. Значение $pnom_j(t)$ для элементов множества M_{avail} является ожидаемой номинальной производительностью после ассимиляции.

• Декларируемое время завершения проекта T^{plan} , т. е. момент времени, в который проект должен быть завершён по плану.

• Время ассимиляции нового исполнителя в проекте t_{assim} , по истечении которого производительность нового исполнителя достигает некоторого номинального значения, которое далее остается постоянным на относительно большом промежутке времени.

2.3. Формализация процессов качественных и количественных изменений в множестве исполнителей

Формализуем наиболее значимые в аспекте производительности процессы, протекающие внутри множества исполнителей, работающих над проектом. Разобьем множество исполнителей, принимающих участие в проекте, на три основные подгруппы:

- эксперты M^{exp} ;
- обучающие эксперты M^{learn_exp} , выделенные для обучения новых исполнителей;
- не-эксперты M^{n_exp} , которые были относительно недавно подключены к проекту (или подключаются в текущей точке перепланирования).

Такое разбиение исполнителей на подгруппы принято в работах [4, 6] и позволяет формализовать процессы, протекающие в множестве исполнителей с течением времени.

Схематично эти процессы представлены на рис. 1.

Сформулируем функциональные зависимости номинальной производительности исполнителя от времени $pnom_j(t)$ в зависимости от принадлежности исполнителя к той или иной группе.

Производительность вновь принятого на работу исполнителя возрастает с течением времени в соответствии с кривой обучения (например, Stanford-B, S-curve) [6]. Однако для упрощения исследования на начальных этапах примем в качестве допущения, что в среднем номинальная производительность не-эксперта возрастает линейно, и если представить графически изменение номинальной производительности не-эксперта на промежутке времени $t > t_q$, где t_q — момент вре-

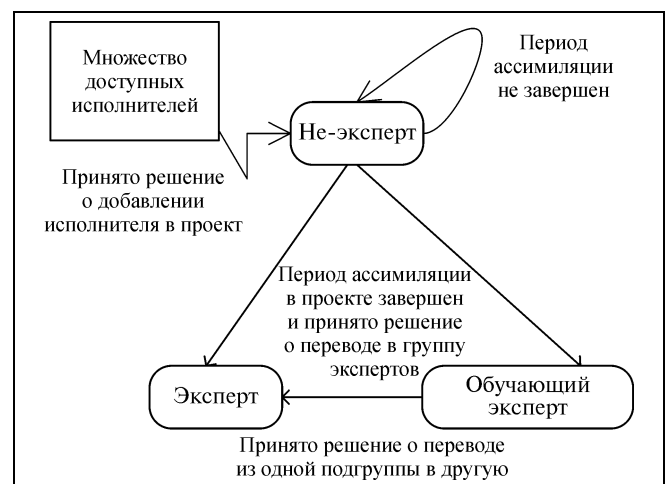


Рис. 1. Схема процессов изменения состава групп исполнителей

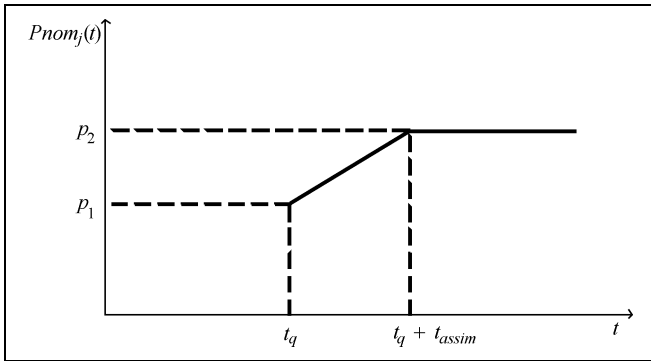


Рис. 2. Зависимость номинальной производительности новичка от времени:

p_1 — исходная производительность исполнителя, p_2 — производительность, достигаемая по окончании периода ассимиляции в проекте

мени, в который происходит перепланирование и данный не-эксперт подключается к проекту, то график зависимости производительности подключенного к проекту не-эксперта от времени примет вид, показанный на рис. 2.

Учитывая, что начальная производительность p_1 нового исполнителя составляет α , $\alpha \leq 1$, от номинальной производительности в конце периода обучения p_2 ; $p_1 = \alpha p_2$, зависимость номинальной производительности нового исполнителя от времени можно записать следующим образом:

$$pnom_j^{n-exp}(t) = \begin{cases} p_{1j}, & t < t_q, \\ \frac{p_{2j}(1-\alpha)}{t_{assim}}(t-t_q) + \alpha p_2, & t_q \leq t \leq t_q + t_{assim}, \\ p_{2j}, & t > t_q + t_{assim}. \end{cases} \quad (1)$$

Каждый новый j -й исполнитель, обучаясь, отнимает определенную долю рабочего времени обучающегося эксперта на свое обучение. Считается, что обучение — один из способов социализации знаний [8]. Одним из лучших и наиболее эффективных методов обучения является межличностное общение [9], которое и положено в основу метода обучения «специалист в области слышимости».

В среднем значение доли «потерянного» обучающимися экспертами времени на одного нового j -го исполнителя примем равным $\mu_j(t)$, учитывая, что на протяжении периода t_{assim} для не-эксперта значение $\mu_j(t)$ по крайней мере, не возрастает.

Поскольку кривая, демонстрирующая зависимость времени общения не-экспертов с обучающимися экспертами от времени в литературе не опи-

сана, предположим следующее: $\mu_j(t)$ описывается линейной функцией, убывающей с течением времени таким образом, что ее значения изменяются от L_j до 0, где L_j — исходное значение времени, отнимаемого j -м не-экспертом на свое обучение в долях от нормы рабочего времени. Предполагается, что к моменту окончания обучения значение $\mu_j(t)$ становится равным нулю:

$$\mu_j(t) = \begin{cases} L_j, & t \leq t_q, \\ -\frac{L_j}{t_{assim}}(t-t_q) + L_j, & t_q < t \leq t_q + t_{assim}, \\ 0, & t > t_q + t_{assim}, \end{cases} \quad (2)$$

где L_j — доля времени, отнимаемая j -м исполнителем у множества экспертов в момент начала обучения, $0 \leq L < 1$.

Функция (2) описывает зависимость изменения доли рабочего времени, отнимаемое не-экспертом у обучающих экспертов от времени.

Далее рассмотрим, как меняется номинальная производительность обучающих экспертов $pnom_j^{learn-exp}(t)$ в период времени $(t_q, t_q + t_{assim})$ в том случае, если в момент времени t_q было произведено добавление новых исполнителей. Пусть:

$$pnom_j^{learn-exp}(t) = pnom_j^{learn-exp}(0) \cdot \left(1 - \frac{1}{m^{learn-exp}} \sum_{k=1}^{m^{n-exp}} \mu_k(t) \right), \quad (3)$$

где $\mu_k(t)$ — доля времени, отнимаемая у экспертов k -м не-экспертом на собственное обучение; m^{n-exp} — число не-экспертов; $m^{learn-exp}$ — число обучающих экспертов.

Для исполнителей, принадлежащих к группе экспертов, будем полагать, что их номинальная производительность не подвержена изменениям в связи с обучением не-экспертов.

Кроме того, необходимо для всех исполнителей, подключенных к проекту, учитывать факт влияния общего числа исполнителей на их номинальную производительность [5].

Номинальная производительность исполнителей уменьшается с возрастанием их числа в силу накладных расходов на коммуникацию. Таким образом, зависимость фактической производительности исполнителей от их числа будет иметь вид:

$$pfact_j(t) = pnom_j(t) \left(1 - \frac{0,06m^2}{100} \right), \quad (4)$$

где $pfact_j(t)$ — фактическая производительность исполнителя, $pnom_j(t)$ — исходная номинальная производительность исполнителя, m — общее число исполнителей.

2.4. Целевая функция задачи минимизации стоимости проекта

В качестве целевой функции будем рассматривать суммарную стоимость эксплуатации исполнителей, принявших участие в проекте, т. е.:

$$S = \min(S_{fact}), \quad (5)$$

где S_{fact} — стоимость эксплуатации исполнителей, принявших участие в проекте на временном интервале $[0; T^{fact}]$, где T^{fact} — момент времени фактического завершения проекта, причем допускается $T^{fact} < T^{plan}$.

Отметим, что в стоимость эксплуатации входит и стоимость простоев исполнителей. Иными словами, все исполнители, подключенные к проекту, оплачиваются до момента времени фактического завершения проекта с момента их подключения и не оплачиваются после отключения.

На рис. 3 схематически показано разбиение времени фактического завершения проекта точкой перепланирования. Рассмотрим процесс вычисления затрат для временного интервала $t \leq t_q$.

Диаграмма подключения исполнителей к проекту может иметь вид, представленный на рис. 4, т. е. в различных точках перепланирования к проекту могут подключаться различные исполнители, выбираемые из множества M_{avail} .

Обозначим через J множество номеров исполнителей, подключенных к проекту после точки перепланирования, через J' — соответственно множество номеров исполнителей, подключенных к проекту до точки перепланирования.

Тогда эксплуатационные расходы на l -м временном интервале

$$c_l = \sum_j s_j(t_l - t_{l-1}), \quad j \in J'. \quad (6)$$

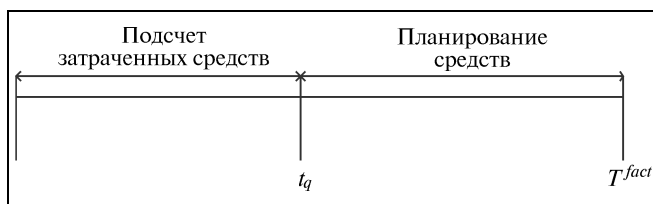


Рис. 3. Планирование суммарной стоимости проекта

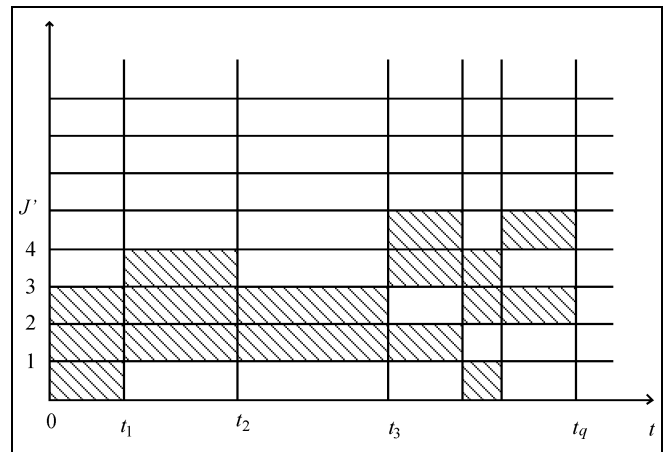


Рис. 4. Диаграмма подключений исполнителей к проекту

Средства, потраченные на эксплуатацию исполнителей до точки перепланирования, вычисляются по формуле

$$S_{spent} = \sum_{l=1}^q c_l. \quad (7)$$

Планирование стоимости эксплуатации исполнителей на оставшейся части проекта опирается на подсчет затрат в соответствии с составляемым расписанием работ. Здесь важно отметить, что производится «оптимистичная» оценка дальнейших работ по проекту, т. е. подразумевается, что никаких событий, продуцирующих точку перепланирования, не произойдет. Тогда, в соответствии с выражениями (6) и (7) получаем

$$S_{T^{fact}} = \sum_{l=1}^q c_l + (T^{fact} - t_q) \sum_j s_j, \quad j \in J. \quad (8)$$

Для получения значения T^{fact} необходимо составить расписание выполнения задач проекта.

Распределение работ по исполнителям опишем посредством матриц R и T :

$$R = \begin{bmatrix} R_{11} & R_{1j} & R_{1m} \\ \dots & R_{ij} & \dots \\ R_{n1} & \dots & R_{nm} \end{bmatrix},$$

где

$$R_{ij} = \begin{cases} 1, & \text{если } i\text{-я задача решается } j\text{-м исполнителем,} \\ 0 & \text{в противном случае,} \end{cases}$$

$T = \{T_i\}$, где T_i — момент времени начала решения i -й задачи.



Как уже было сказано, значение T^{fact} определяется вычислением момента окончания выполнения последней работы.

Вычислим значение

$$T^{fact} = t_q + \max_{i,j} (T_i + \tau_{ij}), \quad (9)$$

где τ_{ij} — время выполнения j -м исполнителем i -й задачи.

Отметим, что здесь вычисление τ_{ij} представляет собой отдельную задачу по причине изменения производительности исполнителей.

С учетом выражений (6)—(9), целевая функция (5) примет следующий вид:

$$S = \min \left(\sum_{l=1}^q c_l + (\max_{i,j} (T_i + \tau_{ij})) \sum_j s_j \right), \quad (10)$$

$$j \in J, \quad i = 1, \dots, n.$$

Перечислим ограничения, выполнение которых обязательно при оптимизации целевой функции (10).

1. В точке перепланирования имеется ограничение на ресурсы (исполнителей) такое, что $M^{exp} \cup M^{n_exp} \cup M^{learn_exp} \subseteq M \cup M^{avail}$.

2. Для каждого исполнителя, работающего в группе экспертов либо присоединяемого к ней, время работы в проекте должно превышать период ассимиляции в проекте: $\forall m_j \in M^{exp} : t_j^{join} + t_{assim} \leq t_q$.

3. Ограничение 2 справедливо и для исполнителей, работающих в группе обучающихся экспертов: $\forall m_j \in M^{learn_exp} : t_j^{join} + t_{assim} \leq t_q$.

4. Все задачи проекта должны быть выполнены за время, не превышающее запланированное: $\max_{i,j} (T_i + \tau_{ij}) \leq T^{plan}$.

5. Для всех задач время начала их решения положительно: $\forall i : T_i \geq 0$.

$$6. t_q \leq T^{plan}.$$

7. Каждая задача может быть решена только одним исполнителем, иными словами — решение любой задачи не может быть начато исполнителем до того, как он завершит решение предыдущей задачи: $\forall I, J : T_{Ik} - T_{Jk} \geq \tau_{kI}$ или $T_{Jk} - T_{Ik} \geq \tau_{kJ}$.

Сформулируем задачу минимизации стоимости программного проекта в окончательном виде: для исходных данных, подробно перечисленных и описанных в п. 2.2, при ограничениях 1—7, а также с учетом изменения производительности исполни-

телей в соответствии с функциями (1)—(4), необходимо найти такие R_{ij} , T_i , M^{exp} , M^{n_exp} и M^{learn_exp} , чтобы обеспечивалось

$$S = \min \left(\sum_{l=1}^q c_l + (\max_{i,j} (T_i + \tau_{ij})) \sum_j s_j \right), \quad j \in J.$$

ЗАКЛЮЧЕНИЕ

Новая формулировка задачи делает актуальным вопрос о поиске методов ее решения. В поставленной задаче совмещены признаки как структурной, так и параметрической оптимизации; кроме того, значительное отличие поставленной задачи состоит во введении функциональной зависимости производительности как от времени, так и от структуры множества исполнителей. Прогнозируемые изменения производительности, в свою очередь, оказывают влияние на время решения задач и, как следствие, на выбираемую структуру множества исполнителей. Поставленная задача комплексная и требует дальнейших исследований возможных методов ее решения.

ЛИТЕРАТУРА

1. *Vieira G., Herrmann W., Lin E.* Rescheduling manufacturing systems: a framework of strategies, policies, and methods. URL: <http://www.isr.umd.edu/Labs/CIM/projects/jos-rescheduling.pdf> (дата обращения: 20.01.2011).
2. *Conway R.W., Maxwell W.L.* Theory of Scheduling. — Massachusetts: Addison Wesley Publishing Company, Reading, 1967.
3. *Барский А.Б.* Параллельные процессы в вычислительных системах: планирование и организация. — М.: Радио и связь, 1990.
4. *Abdel-Hamid T., Madnick S.* Software Project Dynamics, Englewood Cliffs. — NJ: Prentice-Hall, 1991.
5. *Brooks F.* The mythical man-month (Essays on Software Engineering). — Massachusetts: Addison-Wesley Publishing Company Reading, 1975.
6. *Madachy R.* Software Process Dynamics. — IEEE Press, 2008.
7. *Boehm B.W.* Software Engineering Economics. — N-Y.: Prentice-Hall, 1981.
8. *Нонака И., Takeuchi X.* Компания — создатель знания. Зарождение и развитие инноваций в японских фирмах. — М.: Олимп-Бизнес, 2003.
9. *DeMarco T., Lister T.* Peopleware / 2nd Edition. — Dorset House, 1999.

Статья представлена к публикации членом редколлегии В.Н. Бурковым.

Клименко Анна Борисовна — аспирант, Технологический институт Южного федерального университета, г. Таганрог, ✉ anna_klimenko@mail.ru.