

## МОДЕЛЬ И МЕТРИЧНАЯ ОЦЕНКА КАЧЕСТВА НАУЧНЫХ РАБОТ

д.т.н., проф. В.С. Харченко, к.т.н. О.М. Тарасюк, Л.Д. Харченко

*Проведен анализ существующих подходов к оценке качества научных работ. Предложена обобщенная модель качества, ориентированная на оценку диссертационных работ. Описан метричный метод оценки качества, базирующийся на операциях свертки значений метрик, определяющих степень выполнения требований и критериев.*

**Введение. Проблема оценки качества научных работ.** Успешное развитие науки характеризуется объемом и качеством публикуемых научных работ. Только после публикации научной статьи, получения патента, защиты диссертации, научный результат становится общим достоянием и фиксирует конкретный вклад автора. В условиях тщательного роста числа издаваемых научных журналов, сборников, трудов конференций, научно-технических средств особую важность приобретает качество публикаций.

Это касается, в первую очередь, диссертационных работ, к уровню которых сейчас обращено пристальное внимание государственных органов и научной общественности. Требования к ним детально описаны в документах ВАК и многих работах, которые разъясняют их конкретные положения [1, 2]. В тоже время важно, на наш взгляд, обратиться к проблеме оценки качества диссертации и научных работ в целом, учитывая отечественный и зарубежный опыт в этой области. Объективная и полная оценка качества должна на наш взгляд, опираться на четкие хорошо формируемые и легко верифицируемые критерии, которые могут быть детализированны иерархией более детальных субкритериев и требований.

Решение данной задачи может быть выполнено путем создания модели качества научных работ, которая бы задавала некоторый обобщенный профиль оценки, детализируемый для различных видов публикаций. При этом может использоваться подход к оценке, основанный на применении метрик, значения которых задаются в нечетной или булевой шкале (например, соответствует полностью СП, частично СЧ, не соответствует НС), а затем «сворачиваются» в соответствии с иерархией критериев и требований и правилами предикативных сверток. Такой подход показал свою эффективность при решении ряда задач, связанных с оценкой качества программируемого обеспечения [3], динамическим мониторингом результатов деятельности научных работников [4] и др.

**Цель статьи** – разработка модели и метода метричной оценки качества научных работ. Для получения обобщенного профиля оценки выбран наиболее сложный их вид – диссертации. При написании статьи учтен опыт оценки качества диссертаций в ряде европейских стран и практика рецензирования научных работ (статей-докладов) в международных журналах и при подготовке трудов конференций. Речь далее идет о работах научно-технического профиля, хотя, на наш взгляд, некоторые положения могут носить и более общий характер.

**Структура модели качества.** В основу методики оценки качества предлагается положить совокупность базовых критериев  $\{K_i\}_i^7 = 1$ , определяющих верхний уровень ее иерархии. Далее каждый из критериев  $K_i$  декомпозируется на подмножества субкритериев  $\{C_{ij}\}_{j=1}^{N_i}$ , которые, в свою очередь, могут представляться набором требований  $\{T_{ijk}\}_{k=1}^{R_{ij}}$ . В рамках данной статьи ограничимся двух-трехуровневой иерархией и покажем ее соответствие современным требованиям ВАК Украины. Дальнейшее развитие модели качества может быть осуществлено путем детализации требований (субкритериев) нижнего уровня.

При формировании множества критериев (как и множеств субкритериев и требований – нижних уровней иерархии) учитывались следующие требования, которым они должны соответствовать [5]: полнота, равнозначность, верифицируемость, непротиворечивость. Требования непротиворечивости заключаются в данном случае в том, что критерии должны быть максимально ортогональными, т.е. непересекающимися по «зоне действия».

Множество базовых критериев включает нижеперечисленные.

1. **Критерий мотивированности** ( $K_1$ ) определяет, насколько необходимой в практическом и научном контексте является работа, как она связана с другими более масштабными работами такого направления и насколько четко и аргументированно поставлена задача. С учетом этого критерий  $K_1$  объединяет *субкритерии актуальности* ( $C_{11}$ ), *спланированности* ( $C_{12}$ ) и *постановочности* ( $C_{13}$ ). В свою очередь, субкритерий  $C_{12}$  содержит требования практической потребности ( $T_{111}$ ) теоретической необходимости ( $T_{112}$ ), субкритериев  $C_{12}$  – требования к работе как части государственных или иных научно-технических программ, НИР, хозяйственных работ и д.т. ( $T_{121}$ ), а также требования к непосредственному участию автора в выполнении этих работ ( $T_{122}$ ), субкритерий  $C_{13}$  – требования к формулировке темы ( $T_{131}$ ), задач (общих и частных) ( $T_{132}$ ), определению объекта и предмета исследований ( $T_{133}$ ).

2. **Критерий новизны** ( $K_2$ ) определяет, насколько весомыми и достаточными по объему (*субкритерий весомости*  $K_{21}$ ), отличающимися от

известных (*субкритерий отличительности*  $C_{22}$ ), ценными в теоретическом плане (*субкритерий научной ценности*  $C_{23}$ ), существенными по степени новизны (*субкритерий новизны*  $C_{24}$ ) и вклада в теорию (*субкритерий значимости*  $C_{25}$ ) являются полученные результаты. Весомость ( $C_{21}$ ) определяется исходя из полученного множества результатов (методов, моделей, аналитических зависимостей и т.д.). При этом для из результатов должны быть определены их сущность и отличительные признаки («метод оценки надежности программного обеспечения, отличающийся от известных тем, что базируется на метрико-вероятностных процедурах ...») в соответствии с субкритерием  $C_{22}$ , достигаемый эффект («метод ..., отличающийся тем, что базируется ... и позволяет учесть дополнительные характеристики и таким образом повысить достоверность оценки») в соответствии с субкритерием  $C_{23}$ . Кроме того, для каждого из результатов должна быть степень их новизны (впервые, усовершенствован, получил дальнейшее развитие – субкритерий  $C_{24}$ ) и сформулирован вывод о вкладе в соответствующую область теоретических знаний («метод ... развивает теорию надежности компьютерных систем и их программного обеспечения в части ...») – субкритерий  $C_{25}$ .

3. **Критерий доказательности** ( $K_3$ ) определяет насколько обоснованными (*субкритерий обоснованности*  $C_{31}$ ) и достоверными (*субкритерий достоверности*  $C_{32}$ ) являются результаты. Первый из субкритериев включает требования к методике исследования, ее обоснованности, логичности ( $T_{311}$ ), корректности используемого математического аппарата ( $T_{312}$ ), трассируемости (соответственно) поставленных задач и полученных результатов ( $T_{313}$ ). Субкритерий достоверности  $C_{32}$  базируется на требованиях к сводимости полученных результатов к общеизвестным и общепризнанным (например, при граничных значениях новых ранее не учитывавшихся параметров) –  $T_{321}$ , экспериментальной или практической подтверждаемости научных результатов –  $T_{322}$ , сходимости результатов при использовании различных (диверсных) методиках моделирования, расчета и т.д. –  $T_{323}$ .

4. **Критерий практичности** ( $K_4$ ) определяет практическую значимость полученных результатов. Он включает: *субкритерий готовности* к практическому использованию результатов ( $C_{41}$ ), т.е. доведенности их до уровня инженерных методик, алгоритмов, программных продуктов; *субкритерий практической ценности* ( $C_{42}$ ), указывающий конкретный эффект от их возможного внедрения; *субкритерий степени внедрения результатов* ( $C_{43}$ ), определяющий объем и реальный эффект от их критического использования; *субкритерий наличия и обоснованности рекомендаций по расширенному внедрению результатов* ( $C_{44}$ ).

5. **Критерий известности** ( $K_5$ ) указывает, насколько известными и самостоятельными являются полученные результаты. Этот критерий

включает *субкритерии опубликованности* ( $C_{51}$ ), *апробированности* ( $C_{52}$ ) и личного вклада – *самостоятельности* ( $C_{53}$ ). Субкритерий  $C_{51}$  и  $C_{52}$  базируются на идентичных по структуре требованиях к объему – количеству публикаций  $T_{511}$  (докладов  $T_{521}$ ); уровню публикаций  $T_{512}$  (апробаций  $T_{521}$ ), исходя из рейтинга журналов, сборников, научных работ (конференций); полноте публикаций  $T_{513}$  (апробаций  $T_{523}$ ), т.е. степени «покрытия» результатов публикациями (докладами на конференциях).

**6. Критерий согласованности** ( $K_6$ ) определяет, насколько соответствует диссертационная работа паспорту *специальности* (*субкритерий согласованности специальности*  $C_{61}$ ; если оценивается научная статья, доклад он определяет степень соответствия профилю журнала, конференции), содержание автореферата диссертации (или реферата, аннотации статьи, доклада их содержанию) (*субкритерий согласованности автореферата*  $C_{62}$ ), полученные результаты – поставленным задачам (*субкритерий согласованности результатов*  $C_{63}$ ), ключевые слова и УДК – содержанию работы (*субкритерий классификации*  $C_{64}$ ).

**7. Критерий документированности** ( $K_7$ ) определяет, насколько качественно оформлена работа. Он включает субкритерии, учитывающие требования к структуре – *субкритерий структуры*  $C_{71}$  (включая требования к наличию и содержанию основных элементов  $T_{711}$ , последовательности  $T_{712}$ , структурированности  $T_{713}$ ), *субкритерий стиля изложения*  $C_{72}$  (включая требования к лаконичности  $T_{721}$ , языку грамотности  $T_{722}$ , терминологии  $T_{723}$ ), *субкритерий текстового оформления*  $C_{73}$  (включая требования к шрифтам и интервалам  $T_{731}$ , заголовкам  $T_{732}$  и др.), *субкритерий оформления графической части*  $C_{74}$  (включая требования к рисункам  $T_{741}$  и таблицам  $T_{742}$ , в частности, к их наглядности, правильному оформлению, достаточности и т.д.), *субкритерий оформления литературы* –  $C_{75}$  (включая требования к достаточности списка источников и ссылок на них  $T_{751}$ , наличию необходимых современных  $T_{752}$  и англоязычных  $T_{753}$  ссылок, полноте информации по каждому источнику  $T_{754}$ , правильности оформления  $T_{755}$ ).

**Анализ соответствия модели качества и требований к диссертациям.** Результаты анализа соответствия рассмотренной модели и требований ВАК Украины [1] представлены в табл. 1. В левой части таблицы даны элементы модели качества, в правой части – указаны пункты инструкции по требованиям к диссертациям и автореферату. Звездочкой отмечены те субкритерии, выполнение которых в большей степени оценивается экспертом работы.

Таблица 1

Соответствие модели качества требованиям к диссертациям

Критерии	Модель качества		Требования к диссертациям (пункты [1])
	Субкритерии	Требования	
К <sub>1</sub>	С <sub>11</sub>	T <sub>111</sub>	3.4
		T <sub>112</sub>	3.4
	С <sub>12</sub>	T <sub>121</sub>	3.4
		T <sub>122</sub>	3.4
	С <sub>13</sub> *	T <sub>131</sub>	1.2, 3.4
		T <sub>132</sub>	3.4
T <sub>133</sub>		3.4	
К <sub>2</sub>	C <sub>21</sub> *		3.4
	C <sub>22</sub>		3.4
	C <sub>23</sub>		3.4
	C <sub>24</sub> *		3.4
	C <sub>25</sub> *		3.4
К <sub>3</sub>	С <sub>31</sub>	T <sub>311</sub>	3.4
		T <sub>312</sub>	3.4
		T <sub>313</sub>	3.4
	С <sub>32</sub>	T <sub>321</sub>	3.4
		T <sub>322</sub>	3.4
		T <sub>323</sub>	3.4
К <sub>4</sub>	C <sub>41</sub>		3.4
	C <sub>42</sub>		3.4
	C <sub>43</sub> *		3.4
	C <sub>44</sub> *		3.4
К <sub>5</sub>	С <sub>51</sub>	T <sub>511</sub>	3.4
		T <sub>512</sub>	3.4
		T <sub>513</sub>	3.4
	С <sub>52</sub>	T <sub>521</sub>	3.4
		T <sub>522</sub>	3.4
		T <sub>523</sub>	3.4
	C <sub>53</sub> *		1.3, 3.4
К <sub>6</sub>	C <sub>61</sub> *		3.1
	C <sub>62</sub> *		5
	C <sub>63</sub> *		3.6
	C <sub>64</sub> *		3.1
К <sub>7</sub>	С <sub>71</sub>	T <sub>711</sub>	2, 3
		T <sub>712</sub>	2
		T <sub>713</sub>	4.1
	С <sub>72</sub>	T <sub>721</sub>	1.4
		T <sub>722</sub>	1.4
		T <sub>723</sub>	1.4, 3.3

Окончание таблицы 1

	C <sub>73</sub>	T <sub>731</sub>	1.1, 1.5, 4.1, 4.2, 4.5, 4.8
--	-----------------	------------------	------------------------------

	C <sub>74</sub>	T <sub>732</sub>	1.1, 1.5, 4.1, 4.2, 4.5, 4.8
		T <sub>741</sub>	1.1, 4.1, 4.2, 4.3
		T <sub>742</sub>	1.1, 4.1, 4.2, 4.4
	C <sub>75</sub> *	T <sub>751</sub>	1.3, 4.6
		T <sub>752</sub>	4.6
		T <sub>753</sub>	4.6
		T <sub>754</sub>	4.6
		T <sub>755</sub>	1.1, 3.7, 4.7

**Метричная оценка качества.** В результате оценки диссертационной работы (научной статьи, доклада) должен быть сделан вывод о ее соответствии предъявляемым требованиям (рассмотренной выше модели качества). Результат оценки работы может быть следующим: соответствует полностью (СП): соответствует частично, замечания могут быть устранены автором без повторного рассмотрения (рецензирования) (C<sub>41</sub>); соответствует частично, требуется существенная доработка и повторная оценка (C<sub>42</sub>); не соответствует (НС). Такая шкала оценок согласуется с практикой экспертизы диссертационных работ, рецензированием научных работ и докладов.

Метод метричной оценки качества при наличии ее модели включает две основные процедуры – оценку выполнения субкритериев (требований – П<sub>т</sub> и получение общей оценки П<sub>о</sub>).

Для реализации процедуры П<sub>т</sub> каждому требованию должна быть поставлена в соответствие метрика, определена шкала и методика измерения. Метрика M<sub>ijl</sub> (M<sub>ij</sub>, M<sub>i</sub>) – это количественная или качественная мера выполнения требования T<sub>ijl</sub> (субкритерия C<sub>ij</sub>, критерия K<sub>i</sub>). Она может вычисляться: как отношение позитивно выполненных элементов требования A<sub>ijl</sub> к общему числу таких элементов B<sub>ijl</sub> (порядковая шкала значений от 0 до 1. При этом метрика может вычисляться также как взвешенная сумма позитивно оцененных элементов); булева функция, принимающая как и переменные (значения выполнения элементов требований) два значения 0 («да» – требование выполнено) и 1 («нет» – требование не выполнено); нечетная или предикативная функция, принимающая несколько значений (например, СП, C<sub>41</sub>, C<sub>42</sub>, НС) в зависимости от ее документов, имеющих тоже или иное множество значений.

Кроме того, оценочные значения элементов требований могут выражаться и в абсолютных единицах (например, число публикаций – для требования T<sub>54</sub>), которые соотносятся к установленным порогам (существующим сейчас для этого требования по кандидатским диссертациям – 3, докторским – 20).

Процедура П<sub>о</sub> выполняется путем пошаговой свертки значений метрик (аддитивной, булевой, предикативной и др. [6]). В данном случае наи-

более предпочтительной является предикативная свертка. Если метрики принимают значения СП,  $C_{41}$ ,  $C_{42}$ , НС, то для каждого субкритерия, критерия и оценки в целом должны быть определены соответствующие предикаты. Они могут задаваться на основе операций маскируемых или немаскируемых конъюнкций под нечеткими переменными [7]. Для общей оценки  $\Sigma$  на основе оценок по критериям  $K_1 - \Sigma_i$  возможно следующее выражение

$$\Sigma = \begin{cases} \text{СП, если } \forall i : \Sigma_i = \text{СП} ; \\ C_{41}, \text{ если } (\exists i : \Sigma_i < \text{СП}) \& (\forall j : \Sigma_j \geq C_{41}) ; \\ C_{42}, \text{ если } (\exists i : \Sigma_i < C_{41}) \& (\forall j : \Sigma_j \geq C_{42}) ; \\ \text{НС, если } \exists i : \Sigma_i = \text{НС} . \end{cases}$$

Таким образом, результаты метричной оценки могут быть представлены в виде табл. 2.

Таблица 2

Результаты метричной оценки

Итоговая оценка	Критерии		Субкритерии		Требования	
	формулировка	оценка	формулировка	оценка	формулировка	оценка
$\Sigma$	$K_i$	$\Sigma_i$	$C_{il}$	$\Sigma_{il}$	$T_{il1}$	$\Sigma_{il1}$
					$\vdots$	
					$T_{ilm_i}$	$\Sigma_{ilm}$
			$\vdots$	$\vdots$	$\vdots$	
			$C_{in_i}$	$\Sigma_{in}$	$T_{in_1l}$	$\Sigma_{in1}$
					$\vdots$	
					$T_{in_im_i}$	$\Sigma_{inm}$
			$\vdots$	$\vdots$	$\vdots$	

**Выводы.** Предложенная модель и метричный метод оценки качества диссертационных и иных научных работ позволяет формализовать и сделать более компактными результаты оценки. Они могут быть представлены значениями метрик в соответствии с табл. 1, к которой даются короткие комментарии по частично выполненным или невыполненным требованиям, аргументирующим сделанные оценки.

Модель качества дает также укрупненную схему оценки промежуточных результатов подготовки научной работы. При этом в процессе ее подготовки могут даваться промежуточные оценки (для контроля или самоконтроля), которые удобно иллюстрировать с использованием динамических радиальных метрических диаграмм [8].

Кроме того, описанную модель и метод целесообразно развить для оценки циклов научных работ и динамического рейтингования деятель-

ности научных сотрудников или коллективов [9].

## ЛИТЕРАТУРА

1. Довідник здобувача наукового ступеня // Збірник нормативних документів та інформаційних матеріалів з питань атестації наукових кадрів вищої кваліфікації. Редакція ВАК України. – К.: Толока. – 2003. – С. 14 – 35.
2. Бажал Ю.М. Кількісний аналіз наукового апарату джерельної бази публікації (на матеріалі статей у фахових виданнях з економіки). // Бюлетень ВАК України. – 2003. – № 4 – С. 19 – 28.
3. Харченко В.С., Скляр В.В., Тарасюк О.М. Методы моделирования и оценки качества и надежности программного обеспечения: Учебное пособие. – Х.: НАКУ «ХАИ», 2004. – 188 с.
4. Харченко Л.Д. Элементы методологии и систем динамического рейтингования в управлении персоналом // Системы обробки інформації. – Х.: НАНУ, ПАНМ, ХВУ. – 2003. – Вып. 6 – С. 143 – 147.
5. Ларин А.А. Теоретические основы управления. Автоматизация управленческой деятельности: Учебн. пос. Ч. 4. – М.: РВСН, 2000. – 351 с.
6. Харченко В.С., Тарасюк О.М. Использование радиальных метрических диаграмм для оценки характеристик программного обеспечения // Открытые информационные и компьютерные интегрированные технологии: Сб. науч. трудов. – Х.: Нац. аэрокосмич. ун-т «ХАИ». – 2003. – Вып. 18. – С. 123 – 133.
7. Применение методов искусственного интеллекта в управлении проектами / Под. ред. А.Ю. Соколова – Х.: Нац. аэрокосмич. ун-т «ХАИ», 2002. – 474 с.
8. Тарасюк О.М., Харченко В.С. Динамические радиальные метрические диаграммы в задачах управления качеством программного обеспечения // Зб. наук. праць. ін-ту проблем моделювання в енергетиці ім. Г.Є. Пухова. – К.: НАНУ, ІПМЕ. – 2003. – Вып. 22. – С. 202 – 205.
9. Харченко Л.Д. Модели динамического рейтингования на основе многофазного представления результатов деятельности // Системы обробки інформації. – Х.: ХВУ. – 2004. – Вып. 2. – С. 112 – 116.

Поступила 19.01.2005

**ХАРЧЕНКО Вячеслав Сергеевич**, доктор техн. наук, профессор, заведующий кафедрой компьютерных систем и сетей Национального аэрокосмического университета «ХАИ», эксперт Государственного научно-технического центра (ГНТЦ) ядерной и радиационной безопасности (ЯРБ). Область научных исследований – надежность, живучесть и безопасность компьютерных систем для критического применения.

**ТАРАСЮК Ольга Михайловна**, канд. техн. наук, старший преподаватель кафедры компьютерных систем и сетей Национального аэрокосмического университета «ХАИ». Область научных исследований – методы и средства моделирования и оценки качества и надежности программного обеспечения и сложных организационно-технических систем.

**ХАРЧЕНКО Людмила Дмитриевна**, научный сотрудник научно-исследовательского отдела Харьковского университета Воздушных Сил. Область научных исследований – информационные технологии в системах управления, оценки качества и эффективности образования.