# Развивающиеся многомодульные технологии и системы на основе регуляризирующего байесовского подхода. Методологические аспекты интеграции в многомодульных системах

С. В. Прокопчина

*Финансовый университет при Правительстве РФ*

***Аннотация.* В статье определяется и рассматривается один из основных трендов развития информационных систем, состоящий в интеграции различных функциональных модулей в единую систему. Показана актуальность этого направления. Сформулированы методологические и метрологические требования к интеграции модульных систем. Предложена методология регуляризирующего подхода для создания развивающихся интегрированных систем на основе многомодульной архитектуры.**

***Ключевые слова: многомодульные системы, интеграция, регуляризирующий байесовский подход***

## I. Введение

В современном развитии информационных средств на основе методов искусственного интеллекта явно прослеживается тренд интеграции различных функциональных модулей и систем в единые комплексы, обеспечивающие эффективное решение прикладных задач.

С одной стороны это связано с усложнением прикладных задач, с другой стороны обусловлено стремлением разработчиков значительно расширить функциональности систем.

Примерами такой интеграции могут служить различные системы IIoT, вычислительные сети, объединяющие участников производственных процессов от инвесторов до потребителей, от маркетинга до реализации продукции; информационные системы в энергетике и водоснабжении, состоящие из модулей мониторинга и управления генерацией, распределением и потреблением энергетических и водных ресурсов; различные приложения, интегрированные с большими языковыми моделями.

В связи с этим возникает необходимость разработки сервисов для интеграции модулей в автоматизированном или автоматическом режимах.

Одним из таких сервисов является протокол MCP.

Model Context Protocol (MCP) – это открытый стандарт, разработанный компанией Anthropic в конце 2024 года, предназначенный для стандартизации взаимодействия между искусственным интеллектом (ИИ) и внешними источниками данных и инструментами. (nerdit.ru – https://nerdit.ru/model-context-protocol/). MCP позволяет ИИ-моделям, таким как крупные языковые модели и ИИ-агенты, безопасно и эффективно подключаться к различным сервисам, включая базы данных, API, файловые системы и облачные сервисы.

## II. Актуальные вопросы интеграции в многомодульных системах

Создание подобных сервисов безусловно значительно упрощает разработку комплексов со сложной архитектурой.

Однако остается ряд важных вопросов, которые необходимо решать до использования вышеуказанных средств. К их числу относятся, прежде всего, математический базис и метрологическое сопровождение процессов интеграции модулей.

Математический аппарат интеграции необходим при наличии мультимодальной архитектуры информационных систем.

В многомодульных системах использование мульмодальной архитектуры обусловлено разнотипностью данных и знаний исходной информации, их специфическими свойствами и свойствами алгоритмов их обработки.

В этой ситуации для реализации целостного многоступенчатого процесса обработки и интеграции информации различными модулями необходимо последовательная интеграция их решений.

Для этого можно использовать различные виды свертки решений отдельных этапов, выполняемых различными модулями.

Далее в статье будет рассмотрен регуляризирующий байесовский подход, основанный на вероятностной байесовской свертке [1, 3].

В связи с необходимостью реализации интеграции решений особенно актуальным является обеспечение метрологического сопровождения, определяющего качество и устойчивость получаемых решений.

В методологическом аспекте вопрос контроля качества решений связан с вопросом выбора и согласования их метрологических характеристик, то есть используемых метрик качества.

Так, при обработке измерительной информации от сенсорных устройств традиционным является применение мер неопределенности в виде мер смещения и дисперсии результатов, характеризующих точность полученных решений. А при использовании нейросетей используются метрики, характеризующие надежность решений, в том или ином виде определяющие уровни ошибок 1 и 2 рода. Кроме того, в условиях информационной неопределенности, вызванной неточностью, неполнотой, нечеткостью исходной информации, обязателен контроль показателей энтропии и количества информации, добавляемой на каждом интегрируемом этапе.

Это необходимо для обеспечения сходимости и устойчивости решений при интеграции.

Таким образом, комплексы метрологических характеристик при интеграции решений разнотипных модулей должны включать совокупность всех показателей качества их решений.

Пример такого комплекса метрологических характеристик для многомодульных систем. включающих сенсорные блоки и модули на основе искусственного интеллекта (нейросети, системы с нечеткой логикой, аналитические вероятностно-статистические системы), разработан в рамках регуляризирующего байесовского подхода [3].

Вопрос интеграции, гарантирующей требуемое качество решений, решается на основе принципов метрологического синтеза систем метрологического сопровождения. Методологические вопросы и практические решения этого подхода приведены в [3–5].

## III. Методологические аспекты интеграции на основе регуляризирующего байесовского подхода

В информационном аспекте интеграция модулей должна производиться на основе интеграции решений каждого из модулей. Решения модулей H получаются путем трансформации исходной информации X посредством вычислительного фильтра F, то есть модели или иной аналитической структуры, реализованной в модуле

Формализованная запись решения i-модуля имеет вид:

$$H_i = F_i\,(X_i) \qquad (1)$$

Обобщенная модель интегрального решения $G^{(o)}$ при такой интеграции модулей Gi и Gi $_{+1}$ может быть записана в виде следующего формульного выражения.

$$H^{(o)} = F_i\,(X_i) * F_{i+1}\,(X_{i+1}), \qquad (2)$$

где символ * обозначает свертку решений.

Если модули функционируют в многомодальном режиме, то исходная информация X состоит из ряда информационных потоков $\{x_i\}$, и формула (2) принимает вид:

$$H^{(o)} = F_i\ (\{x_i\}) * F_{i+1}\,(\{x_{i+1}\}) \qquad (3)$$

Если модуль представляет собой систему моделей G, имеет сложную иерархическую структуру и состоит из N модульных подсистем Gn, каждая из которых имеет решение Hn, то выражение (3) для интегрального решения имеет следующий вид:

$$Hn^{(n)} = F_1\ (\{x_{i1}\}) *\ldots* F_n\,(\{x_{i,n}\}) \qquad (4)$$

Если и эти модульные подсистемы имеют также иерархическую структуру, состоящую из $m_n$ подсистем для каждой n-той подсистемы, то общее интегрированное решение может быть получено по формуле:

$$Hn^{(n)} = (F_{1,1}\ (\{x_{i1,1}\}) *\ldots* F_{1,m}\,(\{x_{i,1,m}\})*\ldots * (F_{n,1}\ (\{x_{in,1}\}) *\ldots* (\,F_{n,m}\,(\{x_{i,\,n,m}\}) \qquad (5)$$

В развивающихся многомодульных системах номенклатура модулей имеет динамический характер. В зависимости от условий функционирования состав ее модулей может меняться во времени t. Тогда выражение (5) для интегрированного решения может быть сформулировано следующим образом:

$$H_{n,t}^{(n)} = (F_{1,1,t}\ (\{x_{i1,1,t}\}) *\ldots* F_{1,m,t}\,(\{x_{i,1,m,t}\})*\ldots * (F_{n,1,t}\,(\{x_{in,1,t}\}) *\ldots*(\,F_{n,m,t}\,(\{x_{i,\,n,m,t}\}) \qquad (6)$$

Интеграция модулей на основе РБП происходит на базе шкал с динамическими ограничениями. Вид такой шкалы представлен на рис. 1.

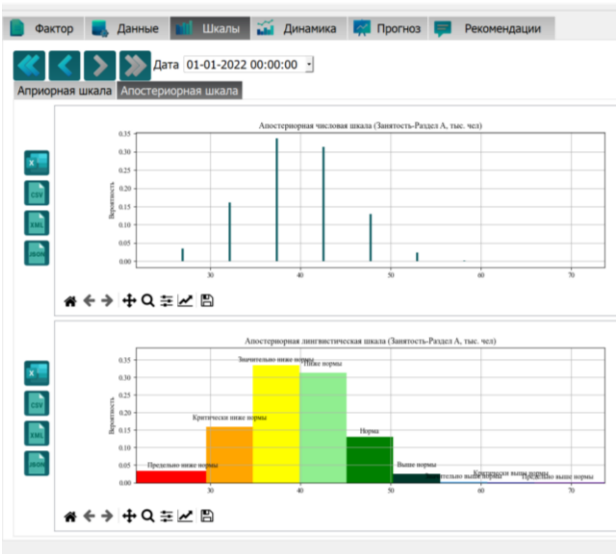


Рис. 1. Шкала с динамическими ограничениями

Шкала имеет двузвенную структуру с верхней шкалой для обработки числовых значений и нижней шкалой для обработки информации в лингвистической форме. Теория и применения таких шкал подробно рассмотрены и под обозначениями приведены в [1–3].

Решение Ht на такой шкале представляется в виде набора решений $h_{k,t}$, соответствующих реперам шкалы в виде нечеткого числа или лингвистического выражения.

При регуляризирующем байесовском подходе для развивающихся многомодульных систем может быть использована вероятностная свертка по модернизированной байесовской формуле вида [1]:

$$P^{(ap)}(h_{k,t}\,|\,\{Mx_{i,t}\}\,|\,Y_{k,t}) = \frac{P^{a}_{k,t}(h_{k,t}\,|\,\{Mx_{i,t}\}\,|\,Y_{i-1,t})P(\tilde{h}_{k,t}\,|\,\{Mx_{i,t}\}\,|\,Y_{i,t})}{\sum_{j=1}^{K} P^{a}_{k,t}(h_{j,t}\,|\,\{Mx_{i,t}\}\,|\,Y_{i-1,t})P(\tilde{h}_{j,t}\,|\,\{Mx_{i,t}\}\,|\,Y_{i-1,t})} \quad (7)$$

где под обозначениями $P^a$ и $P^{ap}$ понимается вероятность предыдущего и текущего решения, а под обозначением Mx – комплекс метрологических характеристик решения.

При свертке решений двух модулей предлагается использовать формулу:

$$P^{(in)}_{k,i,t}(h_{k,t}\,|\;h_{i,t}) = \frac{P^{(ap)}(h_{k,t}\,|\,\{Mx_{k,t}\}\,|\,Y_{k-1,t})P^{(ap)}(h_{i,t}\,|\,\{Mx_{i,t}\}\,|\,Y_{i-1,t})}{\sum_{k=1}^{K}\sum_{i=1}^{I} P^{(ap)}(h_{k,t}\,|\,\{Mx_{k,t}\}\,|\,Y_{k-1,t})P^{(ap)}(h_{i,t}\,|\,\{Mx_{i,t}\}\,|\,Y_{i-1,t})} \quad (8)$$

Применяя последовательно формулу (8) для свертки решений, на всех уровнях иерархии многомодульной системы можно определять интегрированные решения на этих уровнях.

## IV. Метрологическое сопровождение интеграции в многомодульных системах на основе РБП

Для метрологической аттестации решений при интеграции измерительных и аналитических модуле на основе ИИ, в частности. нейросетей, как уже отмечалось, требуется совокупность их характеристик качества.

Комплекс метрологических характеристик РБП (формулы (7) и (8)) содержит совокупность показателей качества, традиционно используемых как в измерительных системах (точность), так и в системах искусственного интеллекта (метрики, определяемые на основе ошибок 1 и 2 рода).

В качестве показателей точности могут быть, вообще говоря, выбраны любые применяемые оценки точности результатов измерений, например, абсолютная, относительная или приведенная систематическая погрешность измерений. При этом расстояние $\rho(h_s, h_{s+1})$ характеризует абсолютную погрешность, равную для каждой пары элементов носителя шкалы РБП расстоянию между ними. При равномерной шкале эта величина будет постоянной и иметь ту же размерность, что и основное измеряемое свойство. При неравномерной шкале БИИ эта составляющая равна максимальному расстоянию между соседними элементами носителя шкалы:

$$\Delta_{max} = \max_{h_s \in H_K} \rho(h_s, h_{s+1})$$

Для шкалы РБП, а особенно универсальной (безразмерной), наиболее удобной формой представления точности результата БИИ является приведенная систематическая погрешность измерения:

$$\xi_s = \xi_{max} = \frac{\Delta_{max}}{\rho(h_K - h_1)}, \quad (9)$$

Отметим, что показатель $\xi_{max}$ является одновременно и параметром регуляризации алгоритмов РБП, которая обеспечивает устойчивость решений. Точность результата БИИ можно оценить по формуле:

$$\forall h_s \in H_K, \xi_s = \frac{\max[\rho(h_s, h_{s+1}); \rho(h_s; h_{s-1})]}{\rho(h_K; h_1)}, \quad (10)$$

Надежность результата БИИ характеризует устойчивость решения, то есть вероятность того, что принятое квазиустойчивое решение будет принадлежать интервалу, соответствующему выбранному элементу носителя шкалы. Этот интервал обозначен как $\left[h_{\gamma_s}, h_{\gamma_{s+1}}\right]$. Это событие можно рассматривать как совместное появление двух событий: нахождению результата БИИ в указанном интервале и отсутствия в нем оценки при истинности любого другого элемента шкалы БИИ.

Очевидно, что первое из этих событий будет характеризовать вероятность ошибки первого рода:

$$P_{\alpha_s} = \int_{h_{\gamma_s}}^{h_{\gamma_{s+1}}} f\left(\tilde{h}|h_s|Y_i\right) d\tilde{h}; \quad (11)$$

а второе – уровень ошибки 2 рода (мощность отвержения неверных решений):

$$\mathcal{D}_{s_j} = \int_{h_{\gamma_s}}^{h_{\gamma_{s+1}}} f\left(\tilde{h}|h_j|Y_i\right) d\tilde{h}; \quad (12)$$

При этом показатель надежности решения определяется по формуле:

$$V_s = \left(P_{\alpha_s} \cdot \mathcal{D}_s\right)|Y_i;\ \ \mathcal{D}_s = \min_{\substack{h_j \in H_K; \\ h_j \neq h_s}} \mathcal{D}_{s_j}|Y_i, \quad (13)$$

Показателем достоверности результата БИИ на данной шкале может быть апостериорная вероятность его появления. Таким образом комплекс метрологических характеристик, определяющий качество результата БИИ имеет вид [1]:

$$\{MX\}_s = \{\xi_s; V_s; P_s\}. \quad (14)$$

В комплекс метрологических характеристик входят также показатели риска применения решения, количества добавленной энтропии и количества полученной информации (по Фишеру) на каждом этапе получения интегрального решения.

Следует отметить, что до проведения этапа интеграции необходимо метрологически обосновать собственные решения каждого интегрируемого модуля.

При интеграции решений их комплексы метрологических характеристик также интегрируются в виде сверток собственных комплексов решений.

$$\{MX\}^{(0)}_{s,t,} = \{MX\}^{(1)}_{s,t} * \ldots * \{MX\}^{(n)}_{s,t}|Y^{(0)}_l; \quad (15)$$

где $Y^{(0)}_l$ – совокупность свойств информационных потоков

Методология интеграции их приведена в ряде работ автора [1, 3], в которых подробно рассмотрены разработанные методики интеграции комплексов.

Практические примеры интеграции моделей на основе РБП приведены в работах автора и его коллег [1–3].

Так методология и средства на основе РБП были применены для интеграции систем IIoT с нейросетевыми модулями обработки изображений и аналитическими системами. В настоящее время осуществлена интеграция средств на основе РБП (цифровая платформа «Инфоаналитик» с большими языковыми моделями. Интеграция таких моделей с платформой «Инфоаналитик» позволило автоматизировать реализацию дерева моделей, получить дополнительную информацию, а также повысить уровень объяснимости и интерпретируемости получаемых интегрированных решений.

## V. Заключение

Таким образом, интеграция модулей в многомодульных системах требует методологического и метрологического обоснования. Для реализации этого требования предлагается использовать методологию, технологии и средства регуляризирующего байесовского подхода.

## Список литературы

[1] Прокопчина С.В. Интеллектуальные измерения на основе регуляризирующего байесовского подхода. М.: Изд. Дом "Научная Библиотека", 2021. 479 с.

[2] Прокопчина С.В., Щербаков Г.А., Ефимов Ю.В. Моделирование социально-экономических систем в условиях неопределенности. М.: Изд. Дом "Научная Библиотека", 2018. 495 с.

[3] Прокопчина С.В. Основы теории шкалирования в экономике. М.: Изд. Дом "Научная Библиотека", 2021. 280 с.