

В. Н. Костин, канд. тех. наук, доц., e-mail: vladimirkostin57@mail.ru, **Д. В. Даньшин**, студент, Оренбургский государственный университет

Метод оценки глубины прогноза развития (эволюции) характеристик сложных систем на основе энтропийного подхода

Предлагается на основе заданных соотношений энтропии (ошибок первого и второго рода) определять допустимое время прогнозирования развития характеристик любых сложных систем. Результаты прогнозирования могут использоваться для задания характеристик перспективных систем при их проектировании. Апробация метода проведена для оценки глубины прогнозирования развития социальной ситуации в целях формирования параметров модели нарушителя при проектировании систем физической защиты.

Ключевые слова: оценка глубины прогнозирования развития характеристик сложных систем, энтропия, ошибки первого и второго рода

Введение

При разработке любой новой системы (технической, экономической, экологической, социальной и т. д.), рассматривая процесс совершенствования как эволюционное развитие системы, необходимо задать перспективные (оптимальные) характеристики будущей системы на определенный момент времени. В качестве меры эволюции системы выступает информационная энтропия. Очевидно, что чем дальше по времени осуществляется прогноз развития, тем больше неопределенность и тем меньше достоверность и надежность получаемых характеристик, т. е. имеется оптимально приемлемое время прогнозирования.

Существует множество математических методов, которые позволяют прогнозировать поведение параметров системы во времени [1]. Однако методов оценки глубины прогнозирования развития систем крайне мало. Среди них марковские модели, которые позволяют определять время наступления предельных вероятностей переходов состояний в системе для однородной марковской модели.

Особенность информационного метода заключается в определении изменения прогнозируемых параметров развития системы в виде порции приращения соотношений энтропии. При достижении заданного значения соотношений энтропии (ошибок прогнозирования) проводится оценка характеристик системы и глубины прогнозирования по времени. Полученные характеристики являются оптимальными, так как система обладает наилучшим соотношением энтропии как меры степени упорядоченности и организованности структуры системы. Если порция энтропии от предыдущей системы будет велика, то система будет недоразвитой, т. е. это модернизация старой системы (не новая система). Если порция энтропии от предыдущей системы мала, то возникает опасность, что новая система будет не приспособлена (не адаптирована) к существующей

среде, т. е. новая система обладает малой преемственностью от предыдущей системы. При построении метода оценки глубины прогнозирования использовался принцип максимума энтропии.

Данные вопросы рассматривались в зарубежных источниках [2], а также в работах [3, 4] только с позиции теории принятия решений как выбор наилучшей системы из совокупности существующих систем. Результатом же применения данного метода являются прогнозируемые характеристики развития перспективной системы на определенный момент времени, которые являются входными данными для проектирования.

Постановка задачи

Необходимо путем прогноза развития параметров сложной системы определить ее характеристики и глубину прогнозирования по времени через заданные ошибки первого и второго рода в виде соотношений энтропии.

Известно, что степень упорядоченности и организованности структуры любой системы можно описать с помощью энтропии [3, 4], поэтому каждый период развития системы определялся в виде оценочного потенциала — энтропии. Математический аппарат метода опирается на материал, изложенный в работе [3].

Формализацию задачи осуществим на основе концепции метода комитетов, рассматривая ее с позиции правомерного существования схемы нахождения сложного предпочтения на множестве временных интервалов развития системы. Комитетом оценки эффективности принимаемых решений назовем такое конечное множество векторов класса R^n : $K = (x_1, x_2, \dots, x_n) \in R^n$, для которого принятие нулевой гипотезы H_0 при заданном уровне значимости обеспечивает определенное число компонент вектора X_m (характеристик системы).

Решение сформулированной задачи связано, во-первых, с определением вектора X_m , количественно отображающего потенциал системы; во-вторых, с оценкой уровня значимости $\alpha_0^{\text{эф}}$ и мощности критерия $\beta_i^{\text{эф}}$ стратегического прогноза. Информационное отображение конкретной ситуации при этом укладывается в следующую схему: имеется n сравниваемых между собой периодов развития системы; каждому периоду существования системы поставлена в соответствие совокупность m характеристик, определяющих потенциал системы в данное время. В этом случае ситуация прогнозирования в развернутой форме характеризуется табл. 1, у которой один вход образован множеством временной шкалой развития системы $\{A_j\}$, а другой — множеством ее признаков.

Вес параметров в формировании оценочного потенциала характеризуется количественной мерой степени уверенности в ситуации объективно существующей неопределенности и отождествляется с распределением вероятностей p_{ji} . Компоненты $\{X_{ji}\}$ задаются в единых или различных физических шкалах. Поэтому для приведения компонент $\{X_{ji}\}$ к единой общей шкале воспользуемся естественной нормализацией, осуществляемой относительно экстремальных значений компонент $\{X_{ji}\}$ как без смены ингредиента на противоположный

$$r_{ji} = \frac{x_{ji}}{x_{\max j}}, \quad (1)$$

так и со сменой ингредиента на противоположный

$$r_{ji} = \frac{x_{\min j}}{x_{ji}} \quad (2)$$

с отображением в $x_{ji} \rightarrow r [0, 1]$. Зависимости (1) и (2) обеспечили отображение выборочного пространства, приведенного в табл. 1, в другое, имеющее мощность континуума, приведенного в табл. 2.

Элементы r_{ji} (пространства мощности континуума) в единой шкале будем идентифицировать с элементарными событиями. При этом определенная на r_{ji} нормированная мера соответствует вероятности $p(r)$, которая отождествляется с понятием интегрального потенциала заданного комплекса элементарных событий. Смысл данной меры состоит в том, чтобы соответствующим образом интерпретировать понятие вероятность. При этом вероятность как категория диалектики совмещает в себе и меру объективной возможности события и степень субъективной уверенности в появлении событий. В целях формализации задачи выбора решения идентифицируем множества систем с пространством событий $\{A\}$, а множество признаков — с со-

Морфологическая матрица

Характеристика системы (признак)	Временная шкала развития системы				
	$\{A_1\}$...	$\{A_j\}$...	$\{A_n\}$
X_1	X_{11}	...	X_{1j}	...	X_{1n}
X_j	X_{j1}	...	X_{ji}	...	X_{jn}
...
X_m	X_{m1}	...	X_{mi}	...	X_{mn}

Таблица 2

Матрица принятия решений

Характеристика системы (признак)	Временная шкала развития системы				
	$\{A_1\}$...	$\{A_j\}$...	$\{A_n\}$
X_1	r_{11}	...	r_{1j}	...	r_{1n}
X_j	r_{j1}	...	r_{ji}	...	r_{jn}
...
X_m	r_{m1}	...	r_{mi}	...	r_{mn}

бытиями $\{x\}$. Тогда связь между всеми компонентами, формирующими оценочный потенциал ситуации, осуществляется через определенную на этих компонентах нормированную меру, которая отождествляется с вероятностью $p(a)$. Распределение вероятностей $p(a)$ обеспечивает оценку нулевой гипотезы H_0 при заданном уровне значимости $\alpha_0^{\text{эф}}$ и мощности критерия $\beta_0^{\text{эф}}$ прогноза развития системы.

Информация в вероятностно-статистической теории выступает в качестве снимаемой уменьшаемой неопределенности, а ее количество измеряется с помощью энтропии через вероятность.

Поэтому дальнейшее построение метода связано с исследованием законов преобразования информации поля декартова произведения двух множеств (систем и признаков) в количественные составляющие информации. С этой целью в логическую схему введем такие понятия, как априорные, апостериорные и условные вероятности, применим теорему Байеса и формулу полной вероятности, а также введем понятие условной вероятности p проявления j -й характеристики в формировании оценочного потенциала при условии, что события, формирующие оценочный потенциал, произошли.

Для получения зависимости определения величины $p(r)$, являющейся нормированной мерой на элементарных событиях $\{r\}$, воспользуемся тем, что понятие оценочного потенциала заданного комплекса элементарных событий можно отождествить с функцией принадлежности, которая ставит в соответствие каждому r действительное число в интервале $[0, 1]$. При этом, не нарушая общности рас-

суждений, искомая зависимость функции принадлежности представляется в форме

$$p_{ji}(r) = r_{ji} / \sum_{i=1}^n r_{ji}. \quad (3)$$

Один из методов расчета вероятности проявления j -го признака сравниваемых интервалов времени на формирование оценочного потенциала основан на введенном В. В. Хоменюком понятии потенциального распределения вероятности [3, стр. 106]. Формализованный расчет этих вероятностных оценок связан с принципом максимума неопределенности, сами оценки получаются в результате решения следующей задачи на условный экстремум:

$$H(P) = - \sum_{j=1}^m p_j \log p_j \rightarrow \max; \quad (4)$$

$$\sum_{j=1}^m p_j = 1; \quad (5)$$

$$\prod_{j=1}^m r^{p_j} = \text{const}. \quad (6)$$

Формула (4) — это информационная энтропия Больцмана—Шеннона, выступающая в качестве меры неопределенности, где p_j — вероятность проявления j -го признака системы. Формула (5) является условием нормировки, а (6) постулирует постоянство среднегеометрического показателя:

$$r_{ji} = \sum_{i=1}^n r_{ji}^{p_j}. \quad (7)$$

Отыскивая условный экстремум (4)—(6) по правилу неопределенного множителя Лагранжа, получаем зависимость для определения оценок вектора потенциального распределения вероятностей, доставляющей максимум целевой функции (4):

$$\hat{p}_j(r) = \sum_{i=1}^n r_{ji} / \sum_{j=1}^m \sum_{i=1}^n r_{ji}. \quad (8)$$

Принцип потенциального распределения вероятностей основан на том, что предпочитается выбор с большей вероятностью тех признаков системы, свойства которых имеют больший вклад в суммарное значение оценочного потенциала. При этом отметим, что для принципа потенциального распределения вероятностей (8) априорная информация о состоянии характеристик основана на принципе недостаточности знаний.

Однако не вызывает сомнения тот факт, что вес различных признаков в формировании оценочного потенциала различен. Получение оценок априорного распределения p_j связано с отношением порядка предпочтения, которое подробно исследовано в тру-

дах Фишборна [5]. Для простого линейного отношения порядка оценки Фишборна априорных вероятностей образуют убывающую арифметическую прогрессию:

$$\check{p}_j = 2(m - j + 1)/(m(m + 1)). \quad (9)$$

Вводя на основе оценок Фишборна априорную вероятность, мы, задавая "вход" в модель, учитываем различный вес характеристики в формировании оценочного потенциала системы. Затем, используя принцип потенциального распределения (8) и положения теоремы Байеса, получаем логически обоснованный "выход" из модели в виде апостериорных условных вероятностей в форме

$$p_j = \sum_{i=1}^n r_{ji} \check{p}_j / \sum_{j=1}^m \sum_{i=1}^n r_{ji} \check{p}_j. \quad (10)$$

Таким образом, польза от введения априорной вероятности, с одной стороны, состоит в том, что это дает присоединение необходимой для анализа информации, а с другой стороны, это выигрыш с точки зрения логической ясности.

После введения априорной вероятности в информационно-вероятностный метод и вычисления апостериорного значения условной вероятности p_j перейдем к следующему этапу моделирования, связанному с получением вероятностных оценок проявления j -го признака i -го варианта на формирование оценочного потенциала. С этой целью воспользуемся теоремой Байеса, в которой речь идет об обращении порядка утверждений в условной вероятности, т. е. в принятых нами обозначениях связываются $p_{ji}(r)$ и p_j . Тогда вероятность $p(a)$ в рассматриваемой информационной ситуации определяется зависимостью

$$p_{ji}(a) = \frac{p_{ji}(r)p_j}{\sum_{i=1}^n \sum_{j=1}^m p_{ji}(r)p_j}. \quad (11)$$

Для построения решающего алгоритма воспользуемся аналогией статистических понятий, а именно понятиями ошибок первого и второго рода. *Ошибка первого рода* заключается в непринятии проверяемой гипотезы H_0 , когда она верна. Вероятность совершения такой ошибки $\alpha_0^{\text{эф}}$ называется уровнем

значимости. Уровень значимости $\alpha_0^{\text{эф}}$ характеризует риск разработчика новой системы. При этом очевидно, что риск разработчика тем выше, чем меньше степень упорядоченности и организации рассматриваемой системы, характерной для того или иного периода времени. В нашем случае это вероятность оказаться в старой системе, т. е. принять старую систему (параметры новой системы незна-

чимо отличаются от существующей системы — модернизация).

Ошибка второго рода — это принятие проверяемой гипотезы H_0 , когда она неверна. Вероятность совершения ошибки второго рода $\beta_1^{\text{эф}}$ характеризует риск заказчика (потребителя). Отметим, что заказчик рискует не только тогда, когда альтернативный вариант обладает низкой степенью упорядоченности и организации, но и когда он обладает низкой степенью приспособляемости к изменению различных внешних условий. Ошибка второго рода характеризует степень порции энтропии преемственности от предыдущей ситуации. Таким образом, это вероятность того, что прогнозируемые параметры не будут обладать преемственностью от исходных данных начальной системы, т. е. прогнозируемые параметры не будут адаптированы к новым условиям.

В нашем случае ошибка второго рода означает, что дальнейшее прогнозирование не целесообразно, так как возникает большая неопределенность использования системы, т. е. характеристики прогнозируемой системы обладают низкой приспособленностью к существующей среде (не связаны с реальностью).

Измерение степени упорядоченности и степени организации системы осуществляется через количество энтропии. Тогда уровень значимости оценим через функцию неопределенности:

$$H_i(p) = -\sum p_{ji}(a) \log p_{ji}(a).$$

При этом не опровергается та система, которой соответствует меньшее значение величины:

$$\alpha_i^{\text{эф}} = H_{\text{best}}(p) - H_i(p),$$

где $H_{\text{best}}(p)$ — значение энтропии гипотетической прогнозируемой системы, обладающей наилучшими характеристиками для данной информационной ситуации; $H_i(p)$ — значение энтропии для i -го варианта существующей системы.

Известно, что количество накопленной и сохраняемой в структуре систем информации I_i в точности равно уменьшению их энтропии $\Delta H_i(p)$. При этом разность между энтропией $H_i(p)$, объективно существующей в оценке влияния каждой конкретной характеристики на формирование оценочного потенциала, и максимальной энтропией H_{max} в рассмотренной информационной ситуации и есть количество информации I_i , накопленной в данной области исследований:

$$I_i = H_{\text{max}} - H_i(p). \quad (12)$$

Чтобы любая система в процессе своей эволюции не достигла предела "приспособленности" (в результате которого системы существуют только в определенных жестко детерминированных условиях), она должна сохранять в себе непредсказуемость, характеризующую определенной порцией энтро-

пии. Удельный вес этой порции определяется зависимостью

$$G_H^i = H_i(p)/I_i.$$

Следующий шаг конструктивистского метода связан с решением задачи минимизации меры неопределенности. Решение данной задачи обеспечит определение количественного значения удельного веса энтропии.

В результате анализа большого числа работ из различных научных областей доказано существование оптимального значения величины G_H , характеризующей удельный вес порции энтропии [3, 4]. Получено оптимальное значение порции энтропии $G_H^{\text{опт}} = 0,27$, это наилучшее соотношение непредсказуемости и детерминированности. Следовательно, чем больше расчетная величина G_H^i отличается от оптимального значения, тем выше вероятность принятия гипотезы G_0 , когда она не верна. Поэтому ошибка второго рода, определяемая мощностью критерия (удельный вес порции энтропии), представляется зависимостью в форме

$$\beta_i^{\text{эф}} = G_H^i - G_H^{\text{опт}}. \quad (13)$$

Величины $\alpha_0^{\text{эф}}$ и $\beta_0^{\text{эф}}$ количественно характеризуют систему договоренностей, т. е. для принятия гипотезы H_0 необходимо представителям заказчика и разработчика договориться о числовом значении ошибок первого и второго рода. При этом если выполняется неравенство вида

$$\alpha_i^{\text{эф}} \leq \alpha_0^{\text{эф}} \text{ и } \beta_i^{\text{эф}} \leq \beta_0^{\text{эф}},$$

то нулевая гипотеза H_0 (вариант i -й системы) принимается, и процесс прогнозирования заканчивается.

Пример. Применение метода рассмотрено для прогнозирования развития социальной ситуации при решении модельного примера. Базовых нарушителей выберем согласно руководящему документу. Приказом министра промышленности и энергетики РФ от 04.05.2007 № 150 определены следующие типы нарушителей:

- x_1 — внешний нарушитель 1-го типа: террористическая группа 5—12 человек. Целью нарушителя является совершение террористического акта (ТА). Последствия несанкционированного действия нарушителя будут выходить за пределы федеральной, региональной или территориальной зон чрезвычайной ситуации (ЧС);
- x_2 — внешний нарушитель 2-го типа: малочисленная группа (2—4 человека). Целью нарушителя является совершение ТА. Последствия действий нарушителя будут выходить за пределы санитарной зоны объекта;

Таблица 3

Статистика криминальной ситуации

Тип нарушителя	Временная шкала ситуаций (год)				
	2008	2009	2010	2011	2012
x_1	2	3	2	2	3
x_2	4	5	4	5	6
x_3	7	6	8	7	8
x_4	2	3	1	4	3
x_5	3	2	3	4	4
x_6	2	1	3	1	2

- x_3 — внешний нарушитель 3-го типа: одиночный подготовленный нарушитель, не имеющий санкционированного доступа на объект. Данный тип нарушителя действует под принуждением или воздействием психотропных препаратов. Целью нарушителя является ТА;
 - x_4 — внешний нарушитель 4-го типа: одиночный нарушитель, не имеющий санкционированного доступа на объект, имеющий целью хищение материальных ценностей;
 - x_5 — внутренний нарушитель 1-го типа: работник объекта, имеющий санкционированный доступ на объект. Целью нарушителя является хищение ради собственной наживы, однако не исключается возможность совершения ТА;
 - x_6 — внутренний нарушитель 2-го типа: работник охраны объекта. Нарушитель данного типа может осуществить хищение материальных ценностей, а также вступить в сговор с внешним нарушителем 1-го и 2-го типа в целях наживы.
- Имеется, например, статистика по каждому типу угрозы в регионе (федерации) за несколько лет,

которая представлена в табл. 3. Типы нарушителей отсортированы в порядке убывания опасности.

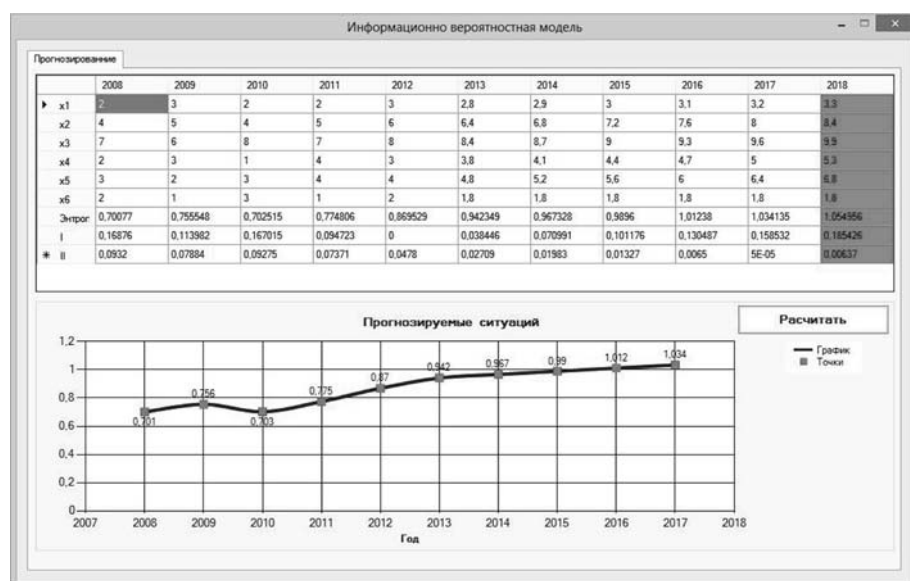
Прогнозирование поведения параметров ситуации на будущие года для табл. 3 осуществлялось на основе метода наименьших квадратов по критерию:

$$\min \sum_{i=0}^n [f(x_i) - P_3(x_i)]^2, \quad (14)$$

с использованием полинома третьей степени $P_3(x_i)$, так как данный полином описывает статистические результаты с наименьшей ошибкой согласования со статистическими данными $f(x_i)$ табл. 3 [1]. Последователно с помощью аппроксимирующей функции получали прогнозируемые результаты ситуаций и с использованием информационно вероятностного метода [формулы (1)–(3), (8)–(14)] оценивали ошибку первого и второго рода. Задача решается итерационно, пока ошибки не достигнут заданных значений. Процесс расчета автоматизирован при помощи разработанной программы на языке программирования C#, результаты приведены на рисунке.

Заключение

Полученные результаты показывают, что прогнозирование поведения ситуации по времени целесообразно осуществлять на шесть лет. При этом ошибка первого рода составит $\alpha_0 = 0,185$, а ошибка второго рода $\beta_0 = 0,0064$, т. е. G_H^i стремится к наилучшему соотношению непредсказуемости и детерминированности. Характеристики социальной ситуации на 2010 г. приведены на рисунке в последнем столбце и используются как входные данные в методике при задании требований к СФЗ при проектировании [6].



Прогнозирование развития ситуации

Список литературы

1. Вержбицкий В. М. Основы численных методов: учебник для вузов. М.: Высш. шк., 2002. 840 с.
2. Gutlin Libal Information Theory and the Living System. N. Y.: Columbia University Press, 1972.
3. Мушков А. Ю., Тихомиров В. А., Тихомиров А. В. Модели и методы стратегического управления сложными экономическими и технологическими системами. Тверь: ВУ ПВО, 2003. 244 с.
4. Седов Е. А. Эволюция и информация. М.: Наука, 1976.
5. Фишборн П. С. Теория полезности для принятия решений. М.: Наука 1978.
6. Костин В. Н., Шевченко С. Н. Методика формирования требований к системе физической защиты на основе концептуальной имитационной модели // Инфокоммуникационные технологии. 2013. № 2. С. 91–98.

Method of an Assessment of Depth of the Forecast of Development (Evolution) Difficult Systems on the Basis of Entropy Approach

In article it is offered on the basis of the set ratios of entropy (errors of the first and second sort) to define admissible time of forecasting of improvement (development) of parameters of any difficult systems. Results of forecasting can be used for a task of characteristics of perspective systems at their design. Approbation of a method is carried out for an assessment of depth of forecasting of development of a social situation for the purpose of formation of parameters of model of the violator at design of systems of physical protection of potentially dangerous objects.

Keywords: assessment of depth of forecasting of development of difficult systems, entropy, errors of the first and second sort

References

1. **Verzhbitskiy V. M.** *Osnovy chislennykh metodov: uchebnik dlya vuzov.* M.: Vyssh. Shk., 2002. 840 p.
2. **Gutlin Libal** *Information Theory and the Living System.* N. Y.: Columbia University Press, 1972.
3. **Mushkov A. Yu., Tikhomirov V. A., Tikhomirov A. V.** *Modeli i metody strategicheskogo upravleniya slozhnymi ekonomicheskimi i tekhnologicheskimi sistemami.* Tver': VU PVO, 2003. 244 p.

4. **Sedov E. A.** *Evolutsiya i informatsiya.* M.: Nauka, 1976.
5. **Fishborn P. S.** *Teoriya poleznosti dlya prinyatiya resheniy.* M.: Nauka, 1978.
6. **Kostin V. N., Shevchenko S. N.** Metodika formirovaniya trebovaniy k sisteme fizicheskoy zashchity na osnove kontseptual'noy imitatsionnoy modeli. *Infokommunikatsionnye tekhnologii.* 2013. N. 2. P. 91—98.

ИНФОРМАЦИЯ



IX Международная научная конференция "Параллельные вычислительные технологии" (ПаВТ 2015)

30 марта — 3 апреля 2015 г.
Уральский федеральный университет
имени первого Президента России Б.Н. Ельцина (г. Екатеринбург)

"Параллельные вычислительные технологии (ПаВТ) 2015" — международная научная конференция, девятая в серии ежегодных конференций, посвященных развитию и применению параллельных вычислительных технологий в различных областях науки и техники. **Главная цель конференции** — предоставить возможность для обсуждения перспектив развития параллельных вычислительных технологий и представления результатов, полученных ведущими научными группами в использовании суперкомпьютерных технологий для решения задач науки и техники в странах СНГ и всего Мира.

Организаторы конференции:

- Российская академия наук
- Суперкомпьютерный консорциум университетов России

Тематика конференции покрывает все аспекты применения высокопроизводительных вычислений в науке и технике, включая приложения, аппаратное и программное обеспечение, специализированные языки и пакеты.

В первый день работы конференции будет объявлена **22-я редакция списка Top50** самых мощных компьютеров СНГ.

Во все дни работы конференции будет действовать суперкомпьютерная выставка, на которой ведущие производители аппаратного и программного обеспечения представят свои новейшие разработки в области высокопроизводительных вычислений.

Языки конференции: русский, английский.

Официальный сайт конференции: <http://agora.guru.ru/pavt2015/>