

© 2022 г. А.Н. ТЫРСИН, д-р техн. наук (at2001@yandex.ru)
(Уральский федеральный университет, Екатеринбург;
Научно-инженерный центр
«Надежность и ресурс больших систем и машин»
УрО РАН, Екатеринбург)

ЭНТРОПИЙНОЕ МОДЕЛИРОВАНИЕ СЕТЕВЫХ СТРУКТУР¹

В настоящее время довольно часто используется энтропия для описания сложных систем в различных областях. Рассмотрены вопросы использования дифференциальной энтропии для сетевых структур, представленных в виде связанных графов с корреляционными связями. Известно, что энтропию непрерывного случайного вектора можно разложить на две составляющие: энтропию случайности и энтропию самоорганизации. Для сетевых структур наряду с оценкой самой энтропии предложены другие полезные характеристики — энтропийная мера взаимосвязи между несколькими подсистемами и энтропия системы в отдельной вершине, которые расширяют возможности энтропийного моделирования для исследования сетевых структур: позволяют оценить взаимосвязанность разных участков между собой и определить, как изменяется энтропия внутри таких систем. Рассмотрены примеры на модельных данных.

Ключевые слова: сетевая структура, случайный вектор, дифференциальная энтропия, граф, система, подсистема, взаимозависимость.

DOI: 10.31857/S0005231022100130, EDN: ALLDVU

1. Введение

Структура — это совокупность устойчивых связей между элементами системы, обеспечивающих воспроизводимость при изменяющихся условиях [1]. В холистическом понимании структура вместе с элементами образует систему.

Сетевая структура рассматривается как децентрализованный комплекс взаимосвязанных элементов, способный расширяться путем включения новых звеньев, что придает сети гибкость и динамичность [2]. В сетевых структурах потенциально могут существовать связи между всеми элементами, причем эти связи могут быть разнонаправленными, т.е. могут быть и двойное подчинение, и межуровневое взаимодействие [3]. Также в системе могут быть подсистемы, что тоже должно отражаться в сетевой структуре как взаимосвязи на уровне подсистем. Поведение реальных систем часто обладает стохастичностью, а связи между их элементами можно адекватно описывать как корреляционные. Модели таких систем обычно акцентируют внимание на явном количественном описании вероятностных характеристик тех или иных ситуаций в жизненном цикле системы [4–7]. Однако такие модели не

¹ Исследование выполнено при финансовой поддержке Российского фонда фундаментальных исследований (проект № 20-51-00001).

позволяют учесть системные характеристики сетевых структур, что может ограничить возможности выработки эффективных управленческих решений.

Энтропия является универсальным параметром, позволяющим объединять различные проявления физического мира в единое целое, т.е. может служить системной характеристикой. В настоящее время достаточно распространено использование энтропии для описания поведения открытых стохастических систем в различных областях [8–11]. Общим в этих работах является использование введенной К. Шенноном информационной энтропии [12]

$$(1) \quad H(\mathbf{D}) = - \sum_{i=1}^L p_i \ln p_i,$$

где p_1, \dots, p_L — вероятности того, что система принимает конечное число соответствующих состояний D_i , т.е. $p_i = P(S \in D_i)$.

Согласно (1) модель системы представляется как функция от множества ее состояний \mathbf{D} . Однако использование информационной энтропии в качестве модели такой системы имеет существенные недостатки:

1. Требуется оценить вероятности p_i . Это требует больших выборок, для некоторых состояний статистику получить практически невозможно.
2. Некоторые состояния систем заранее могут быть вообще не известны.
3. Затруднено моделирование взаимосвязей между элементами многомерных систем.
4. Не учитывается изменение дисперсии.
5. Основные системные закономерности не учитываются.
6. Адекватные энтропийные модели разработаны только для частных задач.

Поэтому актуальна задача поиска более информативных энтропийных характеристик, описывающих поведение сетевых структур.

2. Постановка задачи энтропийного моделирования сетевых структур

Более адекватным подходом к описанию стохастических систем является выделение не отдельных состояний, а ее элементов, которые всегда можно задать для конкретной системы. Представим систему в виде непрерывного случайного вектора $\mathbf{Y} = (Y_1, \dots, Y_m)$ с взаимосвязанными компонентами. Плотность вероятности $p_{\mathbf{Y}}(x_1, \dots, x_m)$ полностью описывает распределение вероятностей многомерной случайной величины \mathbf{Y} . Поэтому вместо информационной энтропии будем использовать дифференциальную энтропию [12], являющуюся функционалом от плотности вероятности случайного вектора \mathbf{Y} ,

$$(2) \quad H(\mathbf{Y}) = - \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} p_{\mathbf{Y}}(x_1, \dots, x_m) \ln p_{\mathbf{Y}}(x_1, \dots, x_m) dx_1 \dots dx_m.$$

Формула (2) была предложена К. Шенноном как формальный аналог понятия информационной энтропии (энтропия непрерывного распределения)

для m -мерного непрерывного случайного вектора \mathbf{Y} . Эта величина впоследствии А.Н. Колмогоровым совместно с И.М. Гельфандом и А.М. Ягломом была исследована и названа дифференциальной энтропией [13].

В [14] доказано, что если все компоненты Y_i имеют дисперсии $\sigma_{Y_i}^2$, то дифференциальная энтропия (далее, энтропия) $H(\mathbf{Y})$ случайного вектора \mathbf{Y} равна

$$(3) \quad H(\mathbf{Y}) = \sum_{i=1}^m \ln \sigma_{Y_i} + \sum_{i=1}^m \kappa_i + \frac{1}{2} \sum_{k=2}^m \ln \left(1 - R_{Y_k|Y_1 \dots Y_{k-1}}^2 \right),$$

где $\kappa_i = H(Y_i/\sigma_{Y_i}) = H(\widehat{Y}_i) = -\int_{-\infty}^{+\infty} p_{\widehat{Y}_i}(x) \ln p_{\widehat{Y}_i}(x) dx$ — дифференциальные энтропии по плотностям с единичными дисперсиями; $R_{Y_k|Y_1 \dots Y_{k-1}}^2 = \frac{\sigma_{Y_k|Y_1 \dots Y_{k-1}}^2}{\sigma_{Y_k}^2}$ — индекс детерминации в общем случае нелинейной регрессионной зависимости случайной величины Y_k , от случайных величин Y_1, \dots, Y_{k-1} , (доля дисперсии Y_k , объясняемая изменением переменных Y_1, \dots, Y_{k-1}). При неизвестном виде зависимости для определения $R_{Y_k|Y_1 \dots Y_{k-1}}^2$ можно воспользоваться методами непараметрического регрессионного анализа [15].

Первые два слагаемых $H_V(\mathbf{Y}) = \sum_{i=1}^m \ln \sigma_{Y_i} + \sum_{i=1}^m \kappa_i = \sum_{i=1}^m H(Y_i)$ равны сумме энтропий элементов $H(Y_i)$, что соответствует случаю, когда элементы Y_i системы \mathbf{Y} взаимно независимы. Величина $H_V(\mathbf{Y})$ равна предельной энтропии, соответствующей взаимной независимости элементов системы, и характеризует рассмотрение целостного объекта как состоящего из частей (аддитивность системы). Поэтому $H_V(\mathbf{Y})$ назовем «энтропией хаотичности».

Возьмем третье слагаемое в (3) со знаком «-»:

$$-G_R(\mathbf{Y}) = \frac{1}{2} \sum_{k=2}^m \ln \left(1 - R_{Y_k|Y_1 \dots Y_{k-1}}^2 \right), \text{ т.е. } G_R(\mathbf{Y}) = -\frac{1}{2} \sum_{k=2}^m \ln \left(1 - R_{Y_k|Y_1 \dots Y_{k-1}}^2 \right).$$

Назовем $G_R(\mathbf{Y})$ энтропийной «мерой самоорганизации». Она отражает степень взаимосвязи между элементами системы \mathbf{Y} , характеризуя свойства системы как целого (целостность системы): при полной взаимной независимости между элементами системы $G_R(\mathbf{Y}) = 0$; при строгой функциональной зависимости между хотя бы двумя элементами системы $G_R(\mathbf{Y}) \rightarrow +\infty$. Отметим, что в случае двухкомпонентного вектора $\mathbf{Y} = (Y_1, Y_2)$ величина $G_R(\mathbf{Y})$ определяется средней взаимной информацией между Y_1 и Y_2 , которая всегда неотрицательна [16].

Если \mathbf{Y}° — гауссов случайный вектор, тогда

$$(4) \quad H_V(\mathbf{Y}^\circ) = \sum_{i=1}^m \ln \sigma_{Y_i} + m \ln \sqrt{2\pi e}, \quad G_R(\mathbf{Y}^\circ) = -\frac{1}{2} \ln |\mathbf{R}_Y|,$$

где $\mathbf{R}_Y = \left\{ \rho_{Y_i^\circ Y_j^\circ} \right\}_{m \times m}$ — корреляционная матрица случайного вектора \mathbf{Y}° .

В рамках энтропийного моделирования сетевыми структурами будем называть системы, каждый из элементов которой связан хотя бы с одним из других элементов системы. Они могут быть представлены в виде связанных графов, в которых связь между элементами (вершинами) задается в виде тесноты корреляционной связи.

Физические компоненты инфраструктуры моделируются как связанный граф, где узлы представляют собой такие компоненты, как районы, развязки автодорог, железнодорожные депо, генераторы, телефонные коммутаторы, резервуары воды и т.д. Дуги графа характеризуют взаимосвязи между элементами. В рамках энтропийного моделирования эти взаимосвязи будем описывать теснотой корреляционной связи между элементами системы (узлами графа). Для сетевых структур недостаточно ограничиться только моделью (2)–(3), так как наряду с оценкой самой энтропии (2) и ее составляющими $H_V(\mathbf{Y})$ и $G_R(\mathbf{Y})$ необходимо оценивать энтропийные характеристики как взаимодействие между собой подсистем, так и для каждого элемента сетевой структуры.

Цель статьи — предложить дополнительный инструментарий, который бы позволил в рамках энтропийного моделирования учесть системные свойства сетевых структур для задач принятия решений.

3. Энтропийная мера взаимосвязи между несколькими подсистемами сетевой структуры

Пусть заданы n подсистем $\mathbf{Y}^{(j)}$ системы $\mathbf{Y} = (Y_1, \dots, Y_m)$, таких что $\mathbf{Y}^{(j)} = (Y_{j,1}, \dots, Y_{j,m_j}) \subset \mathbf{Y}$, $j = 1, \dots, n$, $n \in \{2, 3, \dots, m\}$, любая компонента Y_i входит в состав только одной подсистемы (случайного вектора) $\mathbf{Y}^{(j)}$. Также считаем, что все компоненты \mathbf{Y}_i имеют дисперсии. Определим «энтропийную меру взаимосвязи» между подсистемами (случайными векторами) $\mathbf{Y}^{(1)}, \dots, \mathbf{Y}^{(n)}$ как разность между суммой энтропий подсистем и энтропией системы (объединения подсистем) $\mathbf{Y} = \bigcup_{j=1}^n \mathbf{Y}^{(j)} = (Y_1, \dots, Y_m)$

$$(5) \quad G \left(\bigcap_{j=1}^n \mathbf{Y}^{(j)} \right) = \sum_{j=1}^n H(\mathbf{Y}^{(j)}) - H \left(\bigcup_{j=1}^n \mathbf{Y}^{(j)} \right).$$

Теорема 1. Энтропийная мера взаимосвязи между несколькими подсистемами $\mathbf{Y}^{(1)}, \dots, \mathbf{Y}^{(n)}$ сетевой структуры $\mathbf{Y} = \bigcup_{j=1}^n \mathbf{Y}^{(j)}$ равна

$$(6) \quad G \left(\bigcap_{j=1}^n \mathbf{Y}^{(j)} \right) = -\frac{1}{2} \ln \frac{1 - d_e \left(\bigcup_{j=1}^n \mathbf{Y}^{(j)} \right)}{\prod_{j=1}^n (1 - d_e(\mathbf{Y}^{(j)}))},$$

где $d_e(\mathbf{Y}) = 1 - \prod_{k=2}^m (1 - R_{Y_k|Y_1 \dots Y_{k-1}}^2)$, $d_e(\mathbf{Y}^{(j)}) = 1 - \prod_{k=2}^{m_j} (1 - R_{Y_{j,k}|Y_{j,1} \dots Y_{j,k-1}}^2)$ — коэффициенты тесноты корреляционной связи между компонентами случайных векторов \mathbf{Y} , $\mathbf{Y}^{(j)}$ соответственно.

Доказательство теоремы 1 приведено в Приложении.

Замечание 1. Энтропийная мера взаимосвязи (6), как видно из (5), не содержит компоненты, характеризующей хаотичность подсистем. Если все элементы подсистем между собой взаимно независимы, то $G\left(\bigcap_{j=1}^n \mathbf{Y}^{(j)}\right) = 0$. В случае, когда хотя бы два элемента разных подсистем между собой строго функционально зависимы, то $G\left(\bigcap_{j=1}^n \mathbf{Y}^{(j)}\right) \rightarrow +\infty$.

Замечание 2. Рассмотрим частные случаи:

$$1. \text{ Для } n = 2 \quad G(\mathbf{Y}^{(1)} \cap \mathbf{Y}^{(2)}) = -\frac{1}{2} \ln \left(1 - \frac{1 - d_e(\mathbf{Y}^{(1)} \cup \mathbf{Y}^{(2)})}{(1 - d_e(\mathbf{Y}^{(1)}))(1 - d_e(\mathbf{Y}^{(2)}))} \right).$$

2. Пусть U и V — непрерывные случайные величины, у которых существуют дисперсии. Пусть \mathbf{X} — случайный вектор, у всех компонент которого существуют дисперсии. Тогда энтропийные меры взаимосвязи между \mathbf{X} и U и между U и V равны

$$(7) \quad G(\mathbf{X} \cap U) = 1 - \frac{1 - d_e(\mathbf{X} \cup U)}{1 - d_e(\mathbf{X})},$$

$$(8) \quad G(U \cap V) = -\frac{1}{2} \ln \left(1 - R_{V|U}^2 \right) = -\frac{1}{2} \ln \left(1 - R_{U|V}^2 \right),$$

где $R_{U|V}$, $R_{V|U}$ — теоретические корреляционные отношения между U и V .

3. Если \mathbf{Y}° — гауссов случайный вектор, то

$$(9) \quad G\left(\bigcap_{j=1}^n \mathbf{Y}^{(j)\circ}\right) = -\frac{1}{2} \ln \frac{|\mathbf{R}_{\bigcup_{j=1}^n \mathbf{Y}^{(j)\circ}}|}{\prod_{j=1}^n |\mathbf{R}_{\mathbf{Y}^{(j)\circ}}|}.$$

4. Энтропия в отдельном узле сетевой структуры

Пусть $\mathbf{Y} = (Y_1, \dots, Y_m)$ — сетевая структура. Определим «энтропию сетевой структуры в узле Y_l » как разницу между энтропией всей \mathbf{Y} системы и энтропией системы без элемента Y_l

$$(10) \quad H(Y_l(\mathbf{Y})) = H(\mathbf{Y}) - H(\mathbf{Y} \setminus Y_l).$$

Поскольку нумерация элементов в системе не влияет на ее энтропию, то без потери общности считаем, что $l = m$. Тогда выражение (10) примет вид

$$\begin{aligned} H(\mathbf{Y}) - H(\mathbf{Y} \setminus Y_m) &= \sum_{i=1}^m \ln \sigma_{Y_i} + \sum_{i=1}^m \kappa_i + \frac{1}{2} \sum_{k=2}^m \ln \left(1 - R_{Y_k|Y_1 \dots Y_{k-1}}^2 \right) - \\ &- \sum_{i=1}^{m-1} \ln \sigma_{Y_i} - \sum_{i=1}^{m-1} \kappa_i - \frac{1}{2} \sum_{k=2}^{m-1} \ln \left(1 - R_{Y_k|Y_1 \dots Y_{k-1}}^2 \right) = \\ &= \underbrace{\ln \sigma_{Y_m} + \kappa_m}_{H(Y_m)} + \frac{1}{2} \ln \left(1 - R_{Y_m|Y_1 \dots Y_{m-1}}^2 \right). \end{aligned}$$

Поэтому для произвольного номера l элемента в сетевой структуре имеем

$$(11) \quad H(Y_l(\mathbf{Y})) = H(Y_l) + \frac{1}{2} \ln \left(1 - R_{Y_l|Y_1 \dots Y_{l-1} Y_{l+1} \dots Y_m}^2 \right).$$

С учетом (11) можно указать на теоретико-информационную интерпретацию формулы (10) как энтропию в узле сети в виде условной энтропии $H(\mathbf{Y}) - I(Y_l; \mathbf{Y} \setminus Y_l)$, где $I(Y_l; \mathbf{Y} \setminus Y_l)$ — неотрицательная средняя взаимная информация между вершиной Y_l и сетью $\mathbf{Y} \setminus Y_l$, выраженная через коэффициенты корреляции.

Таким образом, справедлива следующая

Лемма 1. Энтропия сетевой структуры $\mathbf{Y} = (Y_1, \dots, Y_m)$ в узле Y_l определяется по формуле (11).

Замечание 3. Энтропию сетевой структуры в узле $H(Y_l(\mathbf{Y}))$ можно разложить на «энтропию хаотичности» $H(Y_l(\mathbf{Y}))_R$ и «энтропийную меру самоорганизации» $H(Y_l(\mathbf{Y}))_V$ в узле Y_l как

$$(12) \quad H(Y_l(\mathbf{Y})) = H_V(Y_l(\mathbf{Y})) - G_R(Y_l(\mathbf{Y})),$$

где $H_V(Y_l(\mathbf{Y})) = H(Y_l) = \ln \sigma_{Y_l} + \kappa_l$ — дифференциальная энтропия элемента (случайной величины) Y_l , $G_R(Y_l(\mathbf{Y})) = -\frac{1}{2} \ln \left(1 - R_{Y_l|Y_1 \dots Y_{l-1} Y_{l+1} \dots Y_m}^2 \right)$ — энтропийная мера взаимосвязи между случайной величиной Y_l и случайным вектором $\mathbf{Y} \setminus Y_l$ (при отсутствии корреляционной связи между Y_l и случайным вектором $\mathbf{Y} \setminus Y_l$ $G_R(Y_l(\mathbf{Y})) = 0$, при строгой функциональной зависимости между Y_l и хотя бы двумя элементами случайного вектора $\mathbf{Y} \setminus Y_l$ $G_R(Y_l(\mathbf{Y})) \rightarrow +\infty$).

Замечание 4. Если \mathbf{Y}° — гауссов случайный вектор, то

$$(13) \quad H(Y_l^\circ(\mathbf{Y}^\circ)) = \ln \sigma_{Y_l^\circ} + \ln \sqrt{2\pi e} + \frac{1}{2} \ln \frac{|\mathbf{R}_{\mathbf{Y}^\circ}|}{|\mathbf{R}_{\mathbf{Y}^\circ \setminus Y_l^\circ}|}.$$

5. Примеры реализации. Обсуждение результатов

Выражения (5)–(13) позволяют исследовать сетевые структуры: оценивать взаимосвязанность между собой различных участков, а также оценивать, как меняется энтропия внутри таких систем. Рассмотрим примеры. Для упрощения интерпретации результатов рассмотрим гауссовы сетевые структуры.

Пример 1. Пусть дана гауссова система \mathbf{X} , граф которой приведен на рис. 1. Она состоит из двух подсистем $\mathbf{X}^{(1)} = (X_1, X_2, X_3, X_4, X_5)$ и $\mathbf{X}^{(2)} = (X_6, X_7, X_8, X_9, X_{10})$.

В табл. 1, 2 приведены значения среднеквадратических отклонений и корреляционной матрицы случайного вектора \mathbf{X} .

Выполним вспомогательные расчеты: $|\mathbf{R}_{\mathbf{X}^{(1)}}| = 0,30920$, $|\mathbf{R}_{\mathbf{X}^{(2)}}| = 0,017723$, $|\mathbf{R}_{\mathbf{X}^{(1)} \cup \mathbf{X}^{(2)}}| = |\mathbf{R}_{\mathbf{X}}| = 2,6472 \cdot 10^{-6}$. Отсюда энтропийная мера взаимосвязи $H(\mathbf{X}^{(1)} \cap \mathbf{X}^{(2)}) = 3,82$.

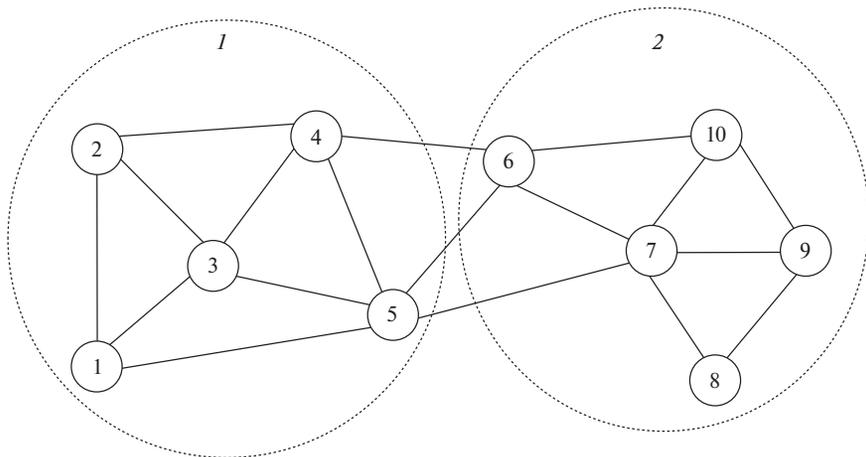


Рис. 1. Граф сетевой структуры \mathbf{X} , состоящей из двух подсистем.

Выполнив расчеты по формулам (12), (13), получим значения энтропий в узлах сетевой структуры \mathbf{X} , которые приведены в табл. 3. Видим, что наибольшие значения энтропийной меры самоорганизации находятся в вершинах X_5 , X_6 , X_7 , которые расположены на границе между подсистемами. Самые высокие значения энтропии хаотичности наблюдаются в вершинах, имеющих наибольшие дисперсии.

Таблица 1. Среднеквадратические отклонения элементов случайного вектора \mathbf{X}

X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}
1,53	1,08	2,00	1,23	1,21	1,84	1,34	1,78	1,00	1,54

Таблица 2. Корреляционная матрица случайного вектора \mathbf{X}

	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}
X_1	1,00	0,41	-0,13	0,21	0,04	0,05	-0,01	-0,04	-0,05	0,07
X_2	0,41	1,00	0,08	0,35	0,00	-0,04	-0,01	0,03	0,00	-0,01
X_3	-0,13	0,08	1,00	0,60	0,41	0,23	-0,12	0,04	0,00	-0,01
X_4	0,21	0,35	0,60	1,00	0,46	0,58	0,02	0,05	0,01	0,14
X_5	0,04	0,00	0,41	0,46	1,00	0,40	-0,34	0,01	0,02	-0,02
X_6	0,05	-0,04	0,23	0,58	0,40	1,00	0,59	0,18	0,02	0,63
X_7	-0,01	-0,01	-0,12	0,02	-0,34	0,59	1,00	0,59	0,33	0,45
X_8	-0,04	0,03	0,04	0,05	0,01	0,18	0,59	1,00	0,74	-0,28
X_9	-0,05	0,00	0,00	0,01	0,02	0,02	0,33	0,74	1,00	-0,58
X_{10}	0,07	-0,01	-0,01	0,14	-0,02	0,63	0,45	-0,28	-0,58	1,00

Таблица 3. Значения энтропий в вершинах сетевой структуры \mathbf{X}

X_l	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}
$H(X_l(\mathbf{X}))$	1,17	-0,35	0,21	-1,34	-1,72	-1,74	-2,18	-1,02	-0,11	0,08
$H_V(X_l(\mathbf{X}))$	1,84	1,50	2,11	1,63	1,61	2,03	1,71	2,00	1,42	1,85
$G_R(X_l(\mathbf{X}))$	0,67	1,85	1,90	2,97	3,33	3,76	3,89	3,02	1,53	1,77

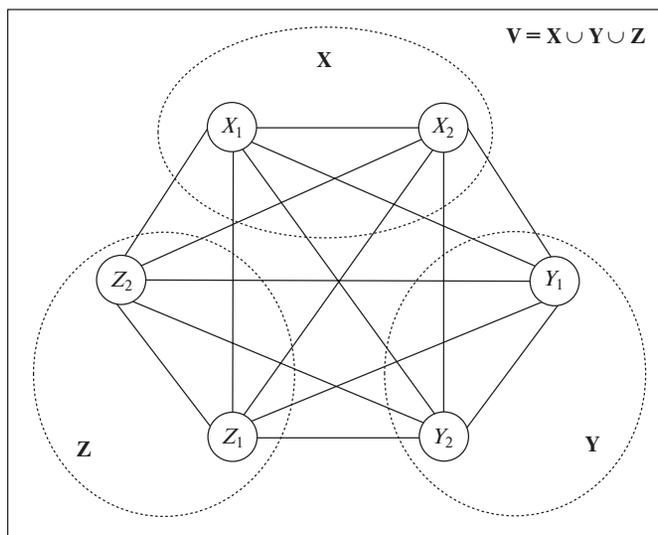


Рис. 2. Граф гауссовой системы \mathbf{V} , состоящей из трех подсистем.

Пример 2. Имеем гауссову систему \mathbf{V} , состоящую из подсистем $\mathbf{X} = (X_1, X_2)$, $\mathbf{Y} = (Y_1, Y_2)$, $\mathbf{Z} = (Z_1, Z_2)$, т.е.

$$\mathbf{V} = \mathbf{X} \cup \mathbf{Y} \cup \mathbf{Z} = (X_1, X_2, Y_1, Y_2, Z_1, Z_2).$$

Граф системы приведен на рис. 2.

Зададим дисперсии всех шести элементов равными 1. Корреляционная матрица системы \mathbf{V} приведена в табл. 4.

Найдем по формулам (12), (13) энтропию во всех шести узлах системы. Результаты расчетов приведены в табл. 5. Самая высокая энтропия наблюдается в узлах Y_1 и Z_1 , а самая низкая — в вершинах X_2 и Z_2 .

Изменим значения парных корреляций.

Случай 1. Коэффициенты парной линейной корреляции приведены в табл. 4.

Таблица 4. Исходная корреляционная матрица

Элемент	X_1	X_2	Y_1	Y_2	Z_1	Z_2
X_1	1	0,5	0,5	0,7	0,3	0,5
X_2	0,5	1	0,5	0,5	0,2	-0,3
Y_1	0,5	0,5	1	0,7	0,6	0,3
Y_2	0,7	0,5	0,7	1	0,5	0,5
Z_1	0,3	0,2	0,6	0,5	1	0,5
Z_2	0,5	-0,3	0,3	0,5	0,5	1

Таблица 5. Значения энтропии в вершинах системы \mathbf{V}

Вершина	X_1	X_2	Y_1	Y_2	Z_1	Z_2
Энтропия в узле	0,66	0,47	0,95	0,72	0,97	0,40

Таблица 6. Значения энтропийной меры взаимосвязи между подсистемами **X, Y, Z**

Случай	1	2	3
Энтропия взаимосвязи	1,52	1,76	1,68

Случай 2. Значение коэффициента парной линейной корреляции между элементами Y_1 и Z_2 изменим с 0,3 на 0,45.

Случай 3. Значение коэффициента парной линейной корреляции между элементами Z_1 и Z_2 изменим с 0,5 на 0,1.

Результаты расчета энтропийной меры взаимосвязи между подсистемами **X, Y, Z** для трех случаев приведены в табл. 6.

Результаты в табл. 6 говорят о следующем:

1. Увеличение тесноты корреляции между элементами разных подсистем приводит к росту тесноты взаимосвязи в целом между подсистемами.
2. Уменьшение тесноты корреляции между элементами в отдельной подсистеме приводит к росту тесноты взаимосвязи в целом между подсистемами.

6. Заключение

В рамках векторной энтропийной модели введены новые энтропийные характеристики — энтропийная мера взаимосвязи между несколькими подсистемами и энтропия системы в отдельном узле. Эти характеристики расширяют возможности исследования сетевых структур: позволяют оценивать взаимосвязанность между собой различных участков, а также определять, как меняется энтропия внутри таких систем.

Предложенный подход рассчитан на сетевые структуры, представимые в виде корреляционных графов. Например, он может применяться для исследования городских систем, транспортных систем, систем связи, систем промышленной безопасности и т.д. Представляет интерес связать векторное энтропийное моделирование с методами риск-анализа. Из-за ограниченности объема статьи был рассмотрен частный случай гауссовой сетевой структуры. Для него определение введенных энтропийных характеристик сводится к определению средних квадратических отклонений компонент системы и определителей корреляционных матриц подсистем и всей системы (сетевой структуры). В дальнейшем планируется рассмотреть применение метода для практических задач.

ПРИЛОЖЕНИЕ

Доказательство теоремы 1.

Зададим n непрерывных случайных векторов произвольных размерностей, у которых нет совпадающих компонент. Обозначим их как $\mathbf{Y}^{(1)} = (Y_1^{(1)}, \dots, Y_{m_1}^{(1)})$, $\mathbf{Y}^{(2)} = (Y_1^{(2)}, \dots, Y_{m_2}^{(2)})$, ..., $\mathbf{Y}^{(n)} = (Y_1^{(n)}, \dots, Y_{m_n}^{(n)})$. Компоненты в любом из случайных векторов $\mathbf{Y}^{(l)}$ могут быть корреляционно взаимно зависимыми. Для этого множества случайных векторов введем их

объединение в виде вектора

$$(II.1) \quad \mathbf{Y} = \bigcup_{j=1}^n \mathbf{Y}^{(j)} = \\ = \underbrace{(Y_1, \dots, Y_{m_1})}_{\mathbf{Y}^{(1)}}, \underbrace{(Y_{m_1+1}, \dots, Y_{m_1+m_2})}_{\mathbf{Y}^{(2)}}, \dots, \underbrace{(Y_{m_1+\dots+m_{n-1}+1}, \dots, Y_{m_1+\dots+m_n})}_{\mathbf{Y}^{(n)}}.$$

Поскольку

$$H \left(\bigcup_{j=1}^n \mathbf{Y}^{(j)} \right) = H(\mathbf{Y}),$$

то

$$\sum_{j=1}^n H_V(\mathbf{Y}^{(j)}) = H_V \left(\bigcup_{j=1}^n \mathbf{Y}^{(j)} \right) = H_V(\mathbf{Y}).$$

Поэтому формула (5) с учетом (3) примет вид

$$(II.2) \quad G \left(\bigcap_{j=1}^n \mathbf{Y}^{(j)} \right) = G_R(\mathbf{Y}) - \sum_{j=1}^n G_R(\mathbf{Y}^{(j)}).$$

Пусть $M_j = \sum_{k=1}^j m_k$, $j = 1, \dots, n$, $M_0 = 0$, $M_n = m_1 + \dots + m_n = m$. В [17] показано, что

$$(II.3) \quad G_R(\mathbf{Y}^{(j)}) = -\frac{1}{2} \sum_{k=2}^{m_j} \ln \left(1 - R_{Y_k^{(j)}|Y_1^{(j)} \dots Y_{k-1}^{(j)}}^2 \right),$$

$$(II.4) \quad G_R(\mathbf{Y}) = -\frac{1}{2} \sum_{k=2}^{M_n} \ln \left(1 - R_{Y_k|Y_1 \dots Y_{k-1}}^2 \right).$$

Очевидно, что если $m_j = 1$, то $G_R(\mathbf{Y}^{(j)}) = 0$. С учетом (II.1)–(II.4) имеем

$$e^{2G(\bigcap_{j=1}^n \mathbf{Y}^{(j)})} = \frac{e^{-2 \sum_{j=1}^n G_R(\mathbf{Y}^{(j)})}}{e^{-2G_R(\mathbf{Y})}} = \frac{\prod_{j=1}^n \prod_{k=M_{j-1}+2}^{M_j} \left(1 - R_{Y_k^{(j)}|Y_1^{(j)} \dots Y_{k-1}^{(j)}}^2 \right)}{\prod_{k=2}^{M_n} \left(1 - R_{Y_k|Y_1 \dots Y_{k-1}}^2 \right)} = \\ = \prod_{j=1}^{n-1} \left(\left(1 - R_{Y_{M_{j+1}}|Y_1 \dots Y_{M_j}}^2 \right) \prod_{k=M_j+2}^{M_{j+1}} \frac{1 - R_{Y_k|Y_1 \dots Y_{k-1}}^2}{1 - R_{Y_k|Y_{M_j+1} Y_{M_j+2} \dots Y_{k-1}}^2} \right).$$

Далее умножим слева и справа последнее выражение на

$$\prod_{j=1}^{n-1} \prod_{k=M_j+2}^{M_{j+1}} \left(1 - R_{Y_k|Y_{M_j+1}Y_{M_j+2}\dots Y_{k-1}}^2\right)$$

и учтем, что $\forall j \quad 1 - d_e(\mathbf{Y}^{(j)}) = \prod_{k=M_{j-1}+2}^{M_j} \left(1 - R_{Y_k|Y_{M_{j-1}+1}Y_{M_{j-1}+2}\dots Y_{k-1}}^2\right)$:

$$\begin{aligned} e^{2G(\bigcap_{j=1}^n \mathbf{Y}^{(j)})} \prod_{j=2}^n \left(1 - d_e(\mathbf{Y}^{(j)})\right) &= \\ &= \prod_{j=1}^{n-1} \left(1 - R_{Y_{M_j+1}|Y_1\dots Y_{M_j}}^2\right) \prod_{k=M_j+2}^{M_{j+1}} \left(1 - R_{Y_k|Y_{M_j+1}Y_{M_j+2}\dots Y_{k-1}}^2\right). \end{aligned}$$

Умножив и разделив последнее выражение справа на

$$\prod_{k=2}^{M_1} \left(1 - R_{Y_k|Y_1\dots Y_{k-1}}^2\right)$$

и учтя, что $1 - d_e(\mathbf{Y}^{(1)}) = \prod_{k=2}^{M_1} \left(1 - R_{Y_k|Y_1\dots Y_{k-1}}^2\right)$, получим

$$\begin{aligned} e^{2G(\bigcap_{j=1}^n \mathbf{Y}^{(j)})} \prod_{j=1}^n \left(1 - d_e(\mathbf{Y}^{(j)})\right) &= \prod_{k=1}^m \left(1 - R_{Y_k|Y_1\dots Y_{k-1}}^2\right) = \\ &= 1 - d_e\left(\bigcup_{j=1}^n \mathbf{Y}^{(j)}\right). \end{aligned}$$

Следовательно,

$$G\left(\bigcap_{j=1}^n \mathbf{Y}^{(j)}\right) = -\frac{1}{2} \ln \frac{1 - d_e\left(\bigcup_{j=1}^n \mathbf{Y}^{(j)}\right)}{\prod_{j=1}^n \left(1 - d_e(\mathbf{Y}^{(j)})\right)},$$

что и требовалось доказать.

СПИСОК ЛИТЕРАТУРЫ

1. Современный философский словарь: 2-е изд. / Под общей ред. проф. В.Е. Кемерова. Лондон, Франкфурт-на-Майне, Париж, Люксембург, Москва, Минск: ПАНПРИНТ, 1998.

2. *Лысак И.В., Косенчук Л.Ф.* Современное общество как общество сетевых структур // Информационное общество. 2015. № 2–3. С. 45–51.
3. *Новиков Д.А.* Сетевые структуры и организационные системы. М.: ИПУ РАН, 2003.
4. *Goldstein H.* Multilevel Statistical Models: 4th ed. Wiley, 2011.
5. *Pardoe I.* Applied Regression Modeling: 2nd ed. Wiley, 2012.
6. *Lanchier N.* Stochastic Modeling. Springer, 2017.
7. *Булатов В.В.* Введение в математические методы моделирования сложных систем. М.: ОнтоПринт, 2018.
8. *Frank S.A., Smith D.E.* Measurement Invariance, Entropy, and Probability // Entropy. 2010. V. 12. No. 3. P. 289–303.
9. *Wilson A.G.* Entropy in Urban and Regional Modelling: Retrospect and Prospect // Geographical Analysis. 2010. V. 42. No. 4. P. 364–394.
10. *Czyz T., Hauke J.* Entropy in Regional Analysis // Quaestiones Geographicae. 2015. V. 34. No. 4. P. 69–78.
11. *Попков Ю.С., Дубнов Ю.А., Попков А.Ю.* Энтропийная редукция размерности в задачах рандомизированного машинного обучения // АиТ. 2018. № 11. С. 106–122.
Popkov Y.S., Dubnov Y.A., Popkov A.Y. Entropy Dimension Reduction Method for Randomized Machine Learning Problems // Autom. Remote Control. 2018. V. 79. No. 11. P. 2038–2051.
12. *Shannon C.E.* A Mathematical Theory of Communication // The CityplaceBell System Technical Journal. 1948. V. 27. No. 3. P. 379–423; No. 4. P. 623–656.
13. *Гельфанд И.М., Колмогоров А.Н., Яглом А.М.* Количество информации и энтропия для непрерывных распределений / Труды III Всесоюзного математического съезда. М.: АН СССР, 1958. Т. 3. С. 300–320.
14. *Тырсин А.Н.* Энтропийное моделирование многомерных стохастических систем. Воронеж: Научная книга, 2016.
15. *Хардле В.* Прикладная непараметрическая регрессия: Пер. с англ. М.: Мир, 1993.
16. *Галлагер Р.* Теория информации и надежная связь: Пер. с англ. М.: Советское радио, 1974.
17. *Тырсин А.Н.* Скалярная мера взаимозависимости между случайными векторами // Зав. лаборатория. Диагностика материалов. 2018. Т. 84. № 7. С. 76–82. <https://doi.org/10.26896/1028-6861-2018-84-7-76-82>.

Статья представлена к публикации членом редколлегии А.А. Лазаревым.

Поступила в редакцию 23.01.2022

После доработки 11.06.2022

Принята к публикации 29.06.2022