



Информационные процессы в пространстве «больших данных»



Борис ЛЁВИН
Boris A. LYOVIN

Виктор ЦВЕТКОВ
Victor Ya. TSVETKOV



Лёвин Борис Алексеевич – доктор технических наук, профессор, ректор Российского университета транспорта (МИИТ), Москва, Россия.

Цветков Виктор Яковлевич – доктор технических наук, профессор, заместитель руководителя центра стратегического анализа и развития НИИАС, Москва, Россия.

Information Processes in Big Data Environment

(текст статьи на англ. яз. –
English text of the article – p. 27)

Статья описывает результаты исследования объектных и ситуационных моделей при управлении транспортом. Раскрываются смысл и содержание такого типа информационных моделей. Показано сходство и различие между ними. Рассматриваются особенности статических и динамических информационных ситуаций, их дифференциация по критериям взаимодействия и критериям состояния. Предложен метод использования модели информационной ситуации для оценки информационной неопределённости при принятии решений.

Ключевые слова: информация, транспорт, управление, моделирование, информационная модель объекта, информационная модель ситуации, информационная неопределённость, функция доверия, функция правдоподобия.

Современное управление транспортом характеризуется возрастанием объёмов информации, ростом сложности управленческих ситуаций и сокращением допустимого времени принятия решений, особенно для высокоскоростного транспорта. В определённой степени эти факторы являются следствием проблемы «больших данных» [1], характерной для разных направлений. Дополнительным фактором управления стала проблема уменьшения или устранения информационной неопределённости [2], которая имеет две причины. Первая причина неопределённости состоит в нехватке необходимой информации для принятия решения. Вторая – в чрезмерном избытке информации, которая требует значительного времени для анализа и нахождения полезного контента, что может быть сопоставимо с допустимым временем принятия решения.

Одним из радикальных для управления транспортом зарекомендовал себя информационный подход [3], включающий применение различных информационных моделей. Модели транспортных систем [4, 5] предназначены для выявле-

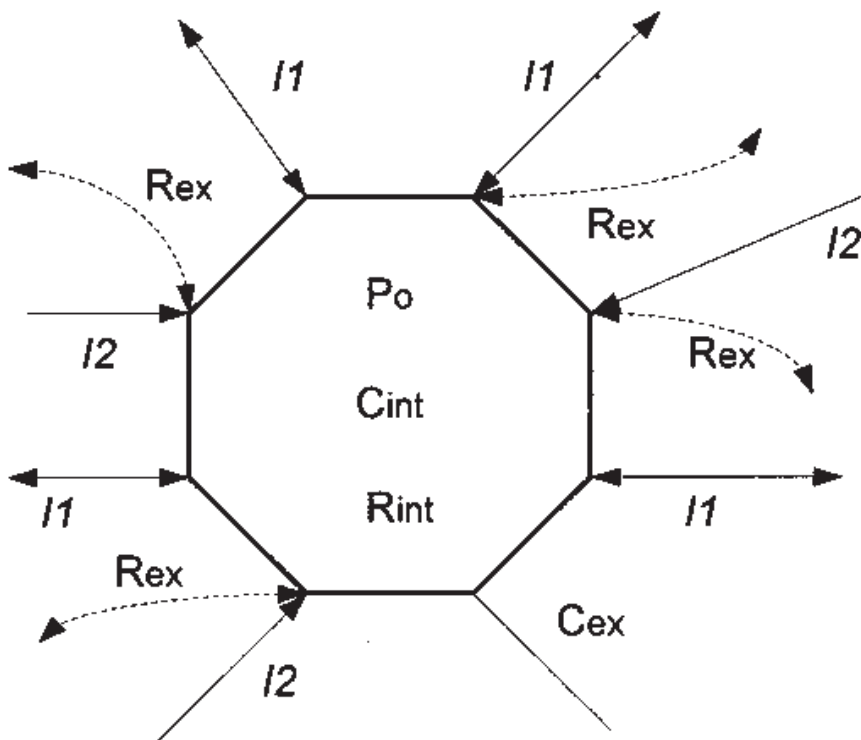


Рис. 1. Информационная модель объекта как совокупность связей и отношений с внешней средой.

ния резервов развития сферы транспортных услуг, повышения качества транспортной работы и прогнозирования будущего транспорта. Информационное моделирование не является простым переносом методов информатики и моделей в сферу транспорта. Оно требует новой организации информационных и электронных ресурсов и применения инновационных решений.

ВИДЫ ИНФОРМАЦИОННЫХ МОДЕЛЕЙ

Объектные модели описывают объекты. Наиболее ярким представителем в информационной области является информационная модель объекта (ИМО). Ситуационные модели описывают ситуацию, в которой находятся объекты. Одной из таких стала информационная модель ситуации (ИМС), или информационной ситуации. Обе эти информационные модели близки между собой, но имеют и различия. Хотя не являются и единственными для описания объектов и ситуаций. Существуют обобщающие модели, например, модель информаци-

онной конструкции [6], которая при необходимости может описывать как объект, так и ситуацию.

Основное различие между ИМО и ИМС – масштаб действия. Общее у них то, что они производное от понятия «информационная модель».

Информационная модель объекта (ИМО) [7] определяется как взаимосвязанная совокупность параметров, наиболее важных связей и отношений. Термин «наиболее важных» означает, что в модель включаются существенные связи и отношения, а несущественные исключаются. Это общее свойство любых моделей, в том числе и неинформационных. Формальное описание ИМО приведено в выражении:

$$\text{ИМО} = F(I(P_o, C_{int}, C_{ex}, R_{int}, \text{Rex}, I_1, I_2, SO)), \quad (1)$$

где P_o – параметры объекта, C_{int} – внутренние связи между частями объекта, C_{ex} – внешние связи с другими объектами и со средой, R_{int} – внутренние информационные отношения между частями объекта, R_{ex} – внешние информационные отношения, I_1 – информационные



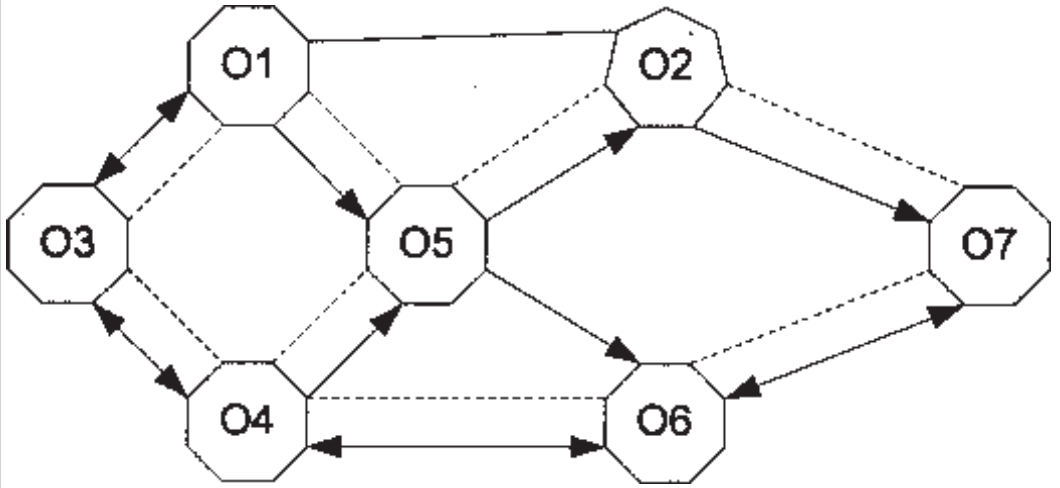


Рис. 2. Пример закрытой информационной ситуации по взаимодействиям.

взаимодействия объекта с другими объектами и со средой, I2 – информационные воздействия на объект, SO – системность объекта (необязательное свойство). Объект может быть частью другого объекта как более сложной системы или может быть самостоятельной системой, обладающей целостностью и системностью. В последнем случае появляется свойство системности объекта SO.

На рис. 1 приведена информационная модель объекта. Внутренняя структура не показана, а обозначена через Po, Cint, Rint. Связи и взаимодействия показаны сплошными линиями, отношения – пунктирными. Взаимодействия – двойными стрелками, воздействия – односторонними стрелками. Связи без взаимодействия – прямыми линиями без стрелок.

Область существенного влияния на объект выделена. Остальные объекты в этой модели отсутствуют и показаны через внешние связи и отношения. Целевое назначение ИМО – описание индивидуального объекта.

Информационная модель ситуации (ИМС) – взаимосвязанная совокупность параметров, наиболее важных связей и отношений для данной ситуации. Целевое назначение ИМС – описание качественно разных ситуаций: взаимодействия объектов, поведения одного или совокупности объектов в данной ситуации, динамики ситуации безотносительно к объектам. Ситуация имеет всегда больший масштаб, чем модель объекта. Информационная ситуация более разнообразна, чем ИМО. У неё предметная

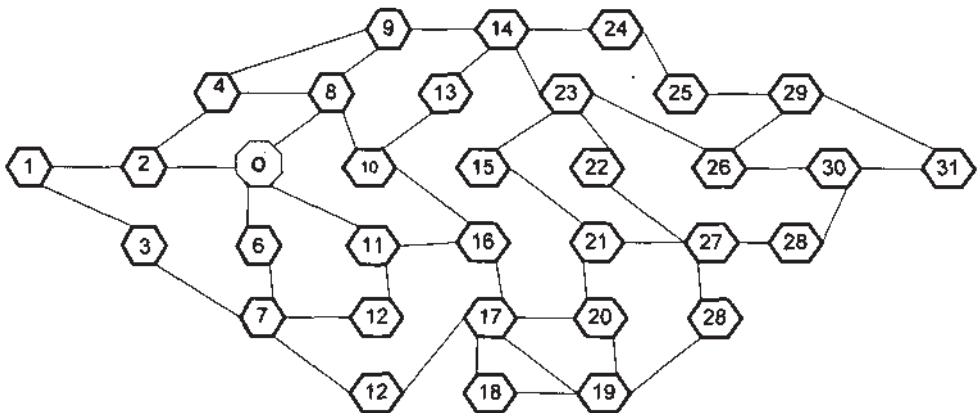


Рис. 3. Информационная ситуация по состояниям.

ориентация. Например, информационная ситуация взаимодействия объектов, информационная ситуация движения объекта, информационная ситуация состояния объекта. Формально модель ИМС приведена в выражении:

$ИМС = F_2(Ps, Co, Cp, Ros, Rps, IS1, IS2, S)$, (2) где Ps – параметры ситуации, Co – связи между объектами, Cp – связи между параметрами объекта, Ros – отношения между объектами, Rps – отношения между параметрами, $IS1$ – информационные взаимодействия в ситуации между объектами, $IS2$ – информационные воздействия в ситуации, S – системность ситуации. На рис. 2 приведён пример информационной ситуации по взаимодействиям. На нём условно показаны семь объектов.

Объекты обозначены через O_i , отношения показаны пунктиром, связи и взаимодействия – сплошными линиями, отношения – пунктирными. Взаимодействия – двойными стрелками, воздействия – односторонними стрелками. Закрытая информационная ситуация чаще всего обладает свойством системности и может быть рассмотрена как сложная система со всеми системными свойствами. Отношения дополняют состояния. Это могут быть отношения иерархии, отношения эквивалентности и др.

На рис. 3 приведена информационная ситуация по состояниям. Связи между состояниями показаны непрерывными линиями. Состояния обозначены шестиугольниками и не заштрихованы. Объект (O) – восьмиугольником и заштрихован.

Рисунок может предполагать множество ситуаций. Например: поезд находится на станции отправления (условно состояние 1); поезд – на промежуточной станции (условно состояние 5 или O); поезд – на станции прибытия (условно состояние 31).

Информационная ситуация по состояниям описывает один объект, перемещающийся по возможным фиксированным состояниям. Она строится, когда есть начальное и целевое состояние. Если цель не определена чётко, а имеется некая целевая парадигма, то от управления по состояниям переходят к управлению по позициям [8]. Пример такой ситуации связан с рынком, когда одной из парадигм

управления является обеспечение конкурентоспособности. В этом случае проводят сравнительный анализ и определяют позицию объекта в информационной ситуации. Сравнивая позицию объекта с позицией других объектов, вырабатывают стратегию улучшения позиции объекта «O» с учётом изменения остальных позиций. Такое управление будет динамическим и требует применения динамической модели ситуации, термин «позиция» употребляется в двух значениях. Пространственная позиция, которая описывает перемещение объекта в пространстве, и параметрическая позиция, которая характеризует позицию объекта по выбранному критерию, например конкурентоспособность или надёжность.

Для полноты рассмотрения транспортных объектов необходимо применять динамические модели. Перемещение транспорта осуществляется по транспортной сети, аналог которой приведён на рис. 3.

Динамические модели включают фактор времени, по этой причине именно они служат основой управления как временного процесса [9, 10]. При перемещении ИМО, как правило, стационарна по внутренним характеристикам. Динамика проявляется в первую очередь на модели информационной ситуации:

$$ИМС(t) = F_3 [Ps(t), Co(t), Cp(t), Ros(t), Rps(t)]. \quad (3)$$

В процессе динамики информационная ситуация рис. 3 характеризует перемещение объекта, изменение его состояния и позиции. Часто при таком анализе применяют топологические модели [11], которые решают не только задачи выбора пути, но и оценивают риск перемещения, текущую и итоговую стоимость перевозки.

При моделировании следует принимать во внимание, что объекты транспорта и транспортная инфраструктура находятся в реальном пространстве. Поэтому для построения информационных моделей необходимо использовать пространственную информацию и пространственные модели [12]. В отдельных случаях для моделирования управления нужны пространственные знания [13]. Включение пространственных факторов в моделирование делает востребованными-



ми методы геоинформатики и геоинформационных моделей [14]. Кроме того, современное пространственное моделирование широко использует космические технологии. В силу этого применяют интеграцию технологий дистанционного зондирования и геоинформатики.

ИНСТРУМЕНТ ОЦЕНКИ НЕОПРЕДЕЛЁННОСТИ

Причинами информационной неопределённости могут быть разные факторы, но из них можно выделить три, из которых два диаметрально противоположные: недостаток информации; избыточность информации, которая в когнитивных терминах проявляется как «необозримость» и «невосприимчивость». Третий фактор обусловлен развитием высокоскоростного транспорта и состоит в требовании сокращения времени на принятие решений.

Эти все причины можно назвать объективными. Дополнительно к ним появилась ещё одна — повышение информационной нагрузки на руководителей, рост риска, обусловленного «человеческим фактором». Информационная неопределённость мотивирована разными «не факторами» [16]: неосведомлённость, неинформированность, неистинность информации (заблуждение), неадекватность моделирования.

Неосведомлённость и неинформированность близкие понятия, но между ними есть различие. Неосведомлённость больше носит субъективный оттенок. Она может возникать при наличии осведомляющей информации, но не полученной принимающими решения по субъективным причинам. Неинформированность же обусловлена лишь отсутствием информации.

Неистинность информации (заблуждение) связана с наличием «правдоподобной», но не до конца точной информации. Целью информационного взаимодействия и диагностики, как и любого научного исследования, является познание истины. Однако вследствие неправильных исходных посылок, неправильной трактовки условий, логических ошибок и т.п. — результатом информационного процесса может быть заблуждение. Под заблуждением [17] обычно понимают определённый

вид ложных высказываний, отличающихся от прочих тем, что такое ложное высказывание принимается за истинное.

Проблемы моделирования, нацеленного на устранение неопределённости, проистекают в основном от того, что информационная неопределённость обусловлена неадекватностью моделей, отражающих реальную ситуацию. В настоящее время информационное моделирование стало более точным инструментом, позволяющим создавать качественно разные модели: ситуаций, процессов, объектов, явлений. Это помогает эффективно устранять неопределённость.

Рассмотрим ситуационные модели, содержащие неопределённость. Простейшая: «Поезд опаздывает на станцию прибытия». Это констатирующая модель, которая содержит неопределённость без какой-либо количественной меры.

На практике возможно применение моделей, содержащих количественные оценки неопределённости. Например, поезд прибывает вовремя на станцию с вероятностью 0,5; другая ситуация — поезд опаздывает на станцию прибытия с вероятностью 0,15. Эти варианты разнородны и вместе создают сложную информационную ситуацию.

Сложная информационная ситуация содержит истинную и не истинную информацию. Для обработки такой информации применяют теорию Демпстера—Шафера (ТДШ) [18]. Это математическая теория свидетельств, основанная на функции доверия (belief functions — Bel) и функции правдоподобия (plausible reasoning — Pl), которые используют, чтобы скомбинировать части разнородной информационной ситуации для вычисления вероятности событий.

Функция доверия представляет собой вероятность события: $Bel: P(X) \rightarrow [0, 1]$.

Пусть X — универсум: набор всех состояний системы (рассматриваемых утверждений). Мощность множества 2^X содержит все подмножества множества X , включая пустое множество \emptyset . Например, если $X = \{a, b\}$, то $2^X = \{\emptyset, \{a\}, \{b\}, X\}$.

Элементы множества мощности могут быть взяты для представления предложений относительно фактического состояния системы, содержащего все и только те

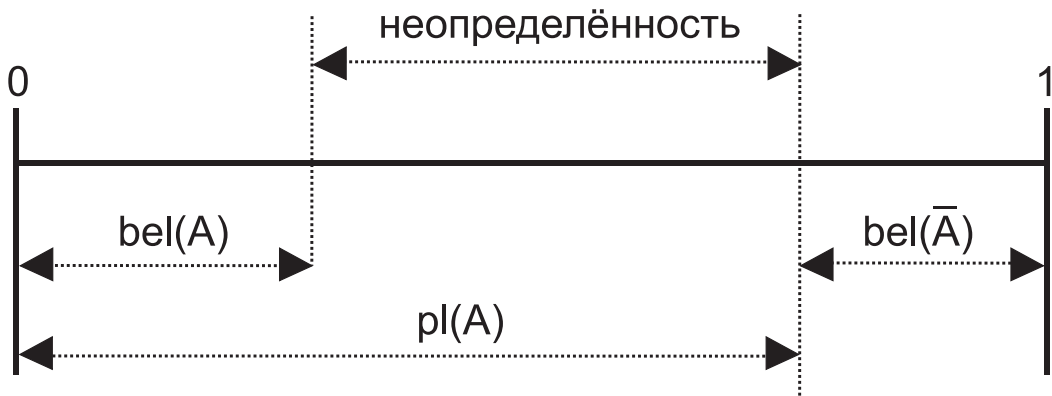


Рис. 4. Отношения между функцией доверия и функцией правдоподобия.

состояния, в которых утверждение верно. Теория доказательства присваивает массу убеждения каждому элементу набора мощности. Вводится понятие массы (m), заимствованное частично из физики, частично из математики, из теории вероятности. Формально функция $m: 2^x \rightarrow [0, 1]$ называется базовым убеждением (basic belief assignment – ВВА), когда оно имеет два свойства:

1. Масса пустого множества равна нулю: $m(\emptyset) = 0$.

2. Массы оставшихся элементов показательного множества нормированы на единичную сумму:

$$\sum_{A \in P(x)} m(A) = 1.$$

Масса $m(A)$ элемента A , заданного члена набора, выражает долю всех релевантных и доступных доказательств, которые подтверждают утверждение о том, что фактическое состояние принадлежит A , но не подмножеству A . Значение $m(A)$ относится только к множеству A и не создаёт никаких дополнительных утверждений о других подмножествах A , каждое из которых, по определению, имеет свою собственную массу. Исходя из приписанных масс, можно определить верхнюю и нижнюю границы возможностей вероятностного интервала.

Этот интервал содержит точную величину вероятности рассматриваемого подмножества $P(A)$ (в классическом смысле) и ограничен двумя неаддитивными непрерывными мерами, называемыми доверием $\{belief\}$ (или поддержка $\{support\}$) и правдоподобием $\{plausibility\}$: $Bel(A) \leq P(A) \leq Pl(A)$.

Доверие $Bel(A)$ к множеству A определяется как сумма всех масс собственных подмножеств рассматриваемого множества:

$$Bel(A) = \sum_{B \vee B \subseteq A} m(B).$$

Правдоподобие $Pl(A)$ – сумма масс всех множеств B , пересекающихся с рассматриваемым множеством A :

$$Pl(A) = \sum_{B, A \cap B \neq \emptyset} m(B).$$

$$Pl(A) = 1 - Bel(\bar{A}) \text{ – для всех } A \in P(X).$$

Отношения между функцией доверия и функцией правдоподобия иллюстрирует рис. 4.

ТДШ позволяет интерпретировать доверие и правдоподобие как границы интервала возможного значения истинности гипотезы: доверие \leq какая-то мера истинности \leq правдоподобие.

При этом постулируется, что:

- доверие к гипотезе определяется как сумма масс свидетельств, однозначно поддерживающих гипотезу;
- правдоподобие определяется как единица – сумма масс всех свидетельств, противоречащих или опровергающих гипотезу.

По существу, ТДШ использует нечёткий оппозиционный подход. Различие между утверждением и оппозиционным утверждением определяет область неопределённости.

Например, пусть у нас есть гипотеза «поезд прибывает по расписанию». Если для неё доверие 0,5 и правдоподобие 0,85,





то это значит, что у нас есть свидетельства (общей массой 0,5), однозначно указывающие, что поезд прибывает по расписанию; но имеются и свидетельства (общей массой 0,15), однозначно указывающие, что поезд не прибывает по расписанию (доверие «поезд не прибывает» = $1 - 0,85 = 0,15$). Оставшаяся масса (дополняющая 0,5 и 0,15 до $1,0 = 0,35$), она же – зазор между правдоподобием 0,85 и доверием 0,5, соответствует «неопределённости» или наличию свидетельства, что поезд точно есть, но не говорящих ничего о том, прибывает он вовремя или не прибывает. Интервал $[0,5; 0,85]$ характеризует неопределённость истинности гипотезы, исходя из имеющихся свидетельств.

Таким образом, модель информационной ситуации с включением вероятностных характеристик позволяет оценивать неопределённость не только качественно, но и количественно. Это дает возможность использовать такие оценки в автоматизированных и интеллектуальных системах для повышения качества и надежности управления.

ЗАКЛЮЧЕНИЕ

Применение объектных и ситуационных моделей для управления транспортом является одним из направлений повышения эффективности деятельности транспорта и основой развития управления. Современное управление на транспорте всё больше переходит от эвристического к автоматизированному и интеллектуальному. Трудность такого перехода связана с трудностью формализации ситуаций, содержащих неопределённость.

Предложенный метод даёт формализовать неопределённость и проводить автоматизированную обработку ситуаций. Объектные и ситуационные модели позволяют на порядки уменьшить объём информации, анализируемый человеком, и использовать такие модели для принятия решений в ситуационных комнатах. В то же время теория Демпстера–Шафера остаётся только одним из подходов к оцен-

ке неопределённости и нуждается в дальнейшем развитии.

ЛИТЕРАТУРА

1. Чехарин Е. Е. Большие данные: большие проблемы // Перспективы науки и образования. – 2016. – № 3. – С. 7–11.
2. Цветков В. Я. Информационная неопределённость и определённость в науках об информации // Информационные технологии. – 2015. – № 1. – С. 3–7.
3. Коваленко Н. И. Информационный подход при построении картины мира // Перспективы науки и образования. – 2015. – № 6. – С. 7–11.
4. Лёвин Б. А., Мамаев Э. А., Багинова В. В. О концепции построения моделей производственно-транспортных систем // Наука и техника транспорта. – 2003. – № 4. – С. 8–17.
5. Мамаев Э. А. Особенности транспортных систем и специфика их моделирования // Системное моделирование социально-экономических процессов: Тезисы докл. и сообщ. XXV межд. науч. школы-семинара им. акад. С. Шаталина. – Часть 2. – М.: ЦЭМИ РАН, 2002. – С. 92.
6. Дешко И. П. Информационное конструирование: Монография. – М.: МаксПресс, 2016. – 64 с.
7. Цветков В. Я. Информационные модели объектов, процессов и ситуаций // Дистанционное и виртуальное обучение. – 2014. – № 5. – С. 4–11.
8. Цветков В. Я. Когнитивное управление: Монография – М.: МаксПресс, 2017. – 72 с.
9. Неймарк Ю. И., Коган Н. Я., Савельев В. П. Динамические модели теории управления. – М.: Наука, 1985. – 400 с.
10. Барабанов И. Н. и др. Динамические модели информационного управления в социальных сетях // Автоматика и телемеханика. – 2010. – № 11. – С. 172–182.
11. Маркелов В. М. Применение топологических моделей геоанализа для оптимизации транспортных маршрутов // Славянский форум. – 2012. – № 2. – С. 56–61.
12. Павлов А. И. Цифровое моделирование пространственных объектов // Славянский форум. – 2015. – № 4. – С. 275–282.
13. Кужелев П. Д. Пространственные знания для управления транспортом // Государственный советник. – 2016. – № 2. – С. 17–22.
14. Бутко Е. Я. Геоинформатика как метод построения картины мира // Славянский форум. – 2017. – № 1. – С. 34–41.
15. Цветков В. Я. Ресурсность и интегративность сложной организационно-технической системы // Международный журнал прикладных и фундаментальных исследований. – 2016. – № 5 (часть 4). – С. 676.
16. Нариньяни А. С. НЕ-факторы: краткое введение // Новости искусственного интеллекта. Вып. 2. – 2004. – М.: Комкнига, 2006. – С. 52–63.
17. Тихонов А. Н., Цветков В. Я. Методы и системы поддержки принятия решений. – М.: МаксПресс, 2001. – 312 с.
18. Shafer Glenn. Dempster–Shafer theory, 1992. [Электронный ресурс]: <http://www.glennshafer.com/assets/downloads/articles/article48.pdf>. Доступ 30.05.2017. ●

Координаты авторов: Лёвин Б. А. – tu@miit.ru,
Цветков В. Я. – cvj2@mail.ru.

Статья поступила в редакцию 30.05.2017, принята к публикации 23.08.2017.

INFORMATION PROCESSES IN BIG DATA ENVIRONMENT

Lyovin, Boris A., Russian University of Transport, Moscow, Russia.

Tsvetkov, Victor Ya., Research, Design and Project Institute of Railway Informatics, Automatics and Communications (NIIAS), Moscow, Russia.

ABSTRACT

The article describes the results of the study on object and situation oriented modelling of transport control, reveals the sense and contents of relevant information models, similarities and differences between described models, considers

features of static and dynamic data situations, their differentiation by criteria of interaction and state. The authors suggest method of the use of the model of information situation for assessment of information uncertainty for decision-making.

Keywords: information, data, transport, management, control, modelling, information model of an object, situational information model, information uncertainty, belief function, likelihood function.

Today, transportation management is characterized by growing volumes of information, increasingly complex situation management challenges, and ever-shrinking periods of time available for decision-making, especially in the case of high-speed transport systems. To an extent, these factors are corollaries following from the more general problem of «big data» [1] that pervades many different fields. An additional factor affecting management is the problem of reducing or eliminating information uncertainty [2] that has a two-fold cause. Uncertainty is caused, firstly, by the lack of information necessary for making a decision. Secondly, it stems from excessive, and overwhelming, amounts of information that require a long time for their analysis and identification of useful content, which time is comparable to the acceptable decision-making time.

A radical method of transport management known as the information-based approach was suggested [3] that involves the application of various information models. Models of transportation systems [4, 5] are designed to reveal opportunities for the development of transportation services, improving their quality, and predicting the future of the transportation industry. Information modeling is not a simple transfer of methods and models used by information science to the field of transportation. It requires a new array of information and electronic resources and a set of innovative information solutions.

Types of information model

Object models describe objects. The OIM, i.e. object information model, is the best example of concepts used by the information science. Situation models describe situations of objects. The situation information model (SIM) is an example, sometimes referred to as the information situation. The two information models are close to each other but distinct. They are not the only ones that describe objects and situations. Generalized models have been developed, such as the information construction model [6] that can be used to describe either objects or situations as needed.

The key distinction between the OIM and SIM lies in their scopes of applicability. Their shared area is that both are derivatives of the notion of 'information model.'

Object information model (OIM) [7] is defined as an interrelated set of parameters, the most important links and relations. The term «most important» means that included in the model are essential links and relations, while non-essential ones are left out. This is the general feature of any models, including non-information ones. A formal description of the OIM is provided by the formula:

$$OIM = f(P_o, C_{int}, C_{ex}, R_{int}, R_{ex}, I_1, I_2, SO), \quad (1)$$

where P_o is the object's parameters, C_{int} is the internal links between parts of the object, C_{ex} is the external links with other objects and the environment, R_{int} is the internal information relations between parts

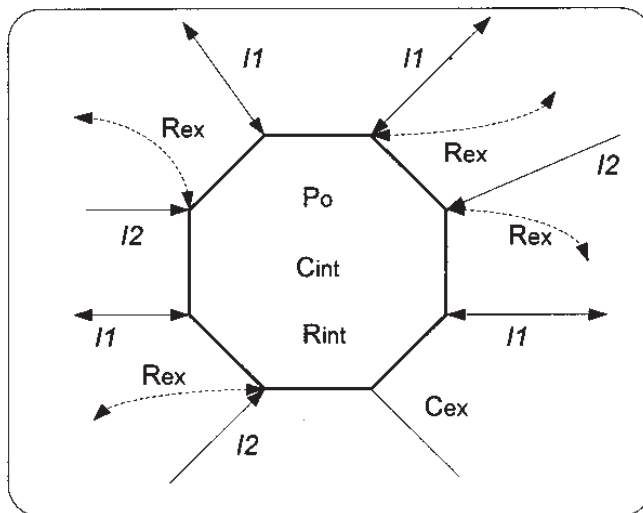
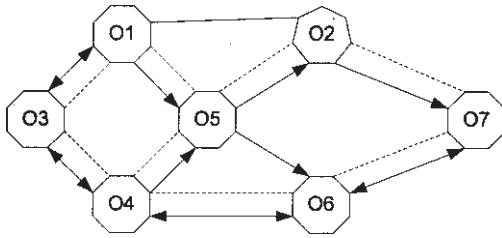
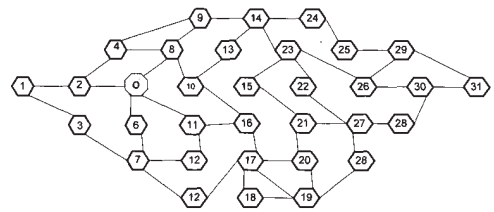


Fig. 1. An object information model as an aggregate of links and relations with external environment.



Pic. 2. Example of a closed information situation described with interactions.



Pic. 3. Information situation by the object's states.

of the object, *Rex* is the external information relations, *I1* is the information interactions between the object, other objects, and the environment, *I2* is the information impacts on the object, *SO* is the systemic characteristic of the object (an optional parameter). An object can be a part of another object that is a more complex system, or it can be a stand-alone system that has the attributes of completeness and systematicity. The latter is a manifestation of the object's systematicity (systematicity of the object, *SO*).

Pic. 1 provides an example of object information model. The internal structure is not shown but indicated with *Po*, *Cint*, *Rint*. Links and interactions are shown with solid lines; relations, with dashed lines. Double-tipped arrows show interactions; single-tipped arrows, impacts. Straight lines with no arrows indicate links without interactions.

The area of material impact on the object is shown with a bold line. Other objects are not included in the model and are referenced with external links and relations. The intended purpose of the OIM is to describe an individual object.

Situation information model (SIM) is an interrelated aggregate of parameters, links and relations that are the most relevant for a given situation. The intended purpose of the SIM is to describe qualitatively different situations: interactions between objects, behavior of a single object or a group of objects in a given situation, dynamics of the situation irrespective of the objects. A situation is always of a greater scale than an object model. An information situation is always more versatile than an OIM. It is object-oriented. Examples include information situation of interactions between objects, information situation of a moving object, information situation of an object's state. A formal description of the SIM is provided by the formula:

$$SIM = F2(Ps, Co, Cp, Ros, Rps, IS1, IS2, S), \quad (2)$$

where *Ps* is the situation's parameters; *Co* is the links between objects, *Cp* is the links between the object's parameters, *Ros* is the relations between objects, *Rps* is the relations between the object's parameters, *IS1* is information interactions between objects involved in the situation, *IS2* is the information impacts in the situation, *S* is the situations' systematicity. *Pic. 2* provides an example of an information situation described with interactions. It shows seven objects.

Objects are indicated with *O_i*, relations are shown with dashed lines; links and interactions, with solid lines. Interactions are shown with double-tipped arrows; impacts, with single-tipped arrows. In most cases, closed information situations are systemic and can be viewed as complex systems possessing the full range of systemic properties. Relations supplement states. The former can be relations of hierarchy, equivalency, etc.

Pic. 3 describes an information situation by the object's states. Links between states are shown with solid lines. States are shown with clear hexagons. The object (O) is represented with a shaded octagon.

The picture can describe multiple situations. For instance, a train is at the station of departure (conventionally, State 1); the train is at an interim station (State 5 or O); the train is at the station of destination (conventionally, State 31).

An information situation by states describes a single object that moves across possible fixed states. Such a model is built when we have an initial state and the target state. If the target is not clearly defined but a certain target paradigm is available, then a transition is made from managing by states to managing by positions [8]. An example of such a situation would be a market where ensuring competitiveness is one of management paradigms. In such a case, a comparative analysis is made to determine the object's position in the information situation. Comparing the object's position with the positions of the other objects, the analyst works out a strategy for improving the position of the object «O» in view of the changes in the other positions. Such management will be dynamic and requires a dynamic situation model; the term «position» has two meanings. The spatial position that describes the movement of the object in space, and the parametric position that characterizes the position of the object by a chosen criterion, such as market competitiveness or reliability.

To make the analysis of transport objects more comprehensive, dynamic models need to be used. Transportation vehicles move on a transportation network, a model of which is showed in *Pic. 3*.

Dynamic models incorporate a temporal dimension, and for this reason such models serve as the basis of management as a time process [9, 10]. When an OIM is moving, it is usually stationary in terms of its internal characteristics. The dynamics are primarily manifested in the situation information model:

$$HMC(t) = F3 [Ps(t), Co(t), Cp(t), Ros(t), Rps(t)]. \quad (3)$$

As the dynamics unfold, the information situation of *Pic. 3* characterizes the movement of the object, the change in its state and position. In many cases, such analysis uses topological models [11] that not only solve route selection problems but assess the movement's risks, the current and total cost of the transportation.

In modeling, it is important to keep in mind that transport objects and transport infrastructure exist in real space. For this reason, spatial information and spatial models should be used as building blocks in the construction of information models [12]. In some cases, management modeling needs spatial knowledge [13]. Incorporation of spatial factors into

modeling necessitates the use of geoinformatics methods and geoinformatics models [14]. In addition, modern spatial modeling relies widely on aerospace technologies. This requires integration of remote sensing and geoinformatics technologies.

A tool for assessing the uncertainty

Information uncertainty may be caused by a variety of factors, three of which can be singled out as the most salient. Out of the three, two are diametric opposites: deficit of information, and excess of information. In cognitive terms, these are described as «opacity» and «indiscernibility». The third factor is related to the advent of high-speed transport and is manifested in the shortening of the time available for decision-making.

All these factors can be regarded as objective. In addition to them, another one has emerged: information overload of managers and increased risks associated with the «human factor». Information uncertainty is also driven by various «non-factors» [16]: unawareness, uninformed-ness, untrue information (misunderstanding), inadequate modeling.

Unawareness and uninformed-ness are close but distinct notions. Unawareness is largely subjective. It can arise when the relevant information is available, but the decision-maker has never received it for subjective reasons. Uninformed-ness stems from just a lack of information.

Untrue information (causing fallacies) is related to the availability of information that is plausible but is not fully accurate. The purpose of information interaction and diagnostics, as well as any scientific research, is learning the truth. However, as a consequence of incorrect initial premises, incorrect interpretation of the conditions, errors in logic, etc., an information process may result in a fallacy. A fallacy [17] is normally construed as a type of false statement that is distinct from the other false statements in that it is taken to be true.

Problems in modeling that aims to eliminate uncertainty largely follow from the fact that information uncertainty results from the inadequacy of models simulating a real situation. By now, information modeling has become a more accurate tool that allows the creation of qualitatively different models: models of situations, processes, objects, phenomena. This tool is effective in helping eliminate uncertainty.

Let us review some situation models that contain uncertainty. An elementary one would be «The train is set to arrive at its destination late». This is a statement model that contains uncertainty without any quantitative measure.

In practice, models are possible that do contain quantitative assessments of uncertainty. For example, the train will be at the station on time with a probability of 0.5; another situation, the train will be at the station late with a probability of 0.15. These variants are disparate and together they create a complex information situation.

A complex information situation will contain both true and false information. To process such information, the Dempster-Shafer theory (DST) is used [18]. This is a mathematical theory of evidence that is based on the belief function (Bel) and the plausible reasoning function (Pl), which functions are used to combine parts of a disparate information situation for the purpose of calculating probabilities of events.

The belief function is the probability of an event: $Bel: P(X) \rightarrow [0, 1]$.

Let X be the universe, i.e. the set of all states of a system (statements under review). The power set 2^X contains all subsets of set X , including the empty set \emptyset . For example, if $X = \{a, b\}$, then $2^X = \{\emptyset, \{a\}, \{b\}, X\}$.

The elements of the power set can be taken to represent assumptions on the actual state of the system that contains all the states for which the statement is true, and only such states. The proof theory assigns a mass of belief to each element of the power set. The notion of mass (m) is introduced, borrowed in part from physics, and in part from the probability theory. Formally, a function $\tau: 2^X \rightarrow [0, 1]$ is called a basic belief assignment (BBA), when it has two properties:

1. The mass of the empty set equals zero: $\tau(\emptyset) = 0$.
2. The masses of the remaining members of the power set add up to a total of 1:

$$\sum_{A \in P(X)} m(A) = 1.$$

The mass $m(A)$ of member A , a given member of the power set, expresses the proportion of all relevant and available evidence that supports the claim that the actual state belongs to A but to no particular subset of A . The value of $m(A)$ pertains only to the set A and makes no additional claims about any subsets of A , each of which have, by definition, their own mass. From the mass assignments, the upper and lower bounds of a probability interval can be defined.

This interval contains the precise probability of a set of interest $P(A)$ (in the classical sense), and is bounded by two non-additive continuous measures called belief (or support) and plausibility: $Bel(A) \leq P(A) \leq Pl(A)$.

The belief $Bel(A)$ for a set A is defined as the sum of all the masses of subsets of the set of interest:

$$Bel(A) = \sum_{B \subseteq A} m(B).$$

The plausibility $Pl(A)$ is the sum of all the masses of the sets B that intersect the set of interest A :

$$Pl(A) = \sum_{B, A \cap B \neq \emptyset} m(B).$$

$$Pl(A) = 1 - Bel(\bar{A}) \text{ for all } A \in P(X).$$

Pic. 4 illustrates the relationship between the belief function and the plausibility function.

DST allows to interpret belief and plausibility as the bounds of the interval where the true value of a hypothesis is possible: belief \leq a measure of truth \leq plausibility.

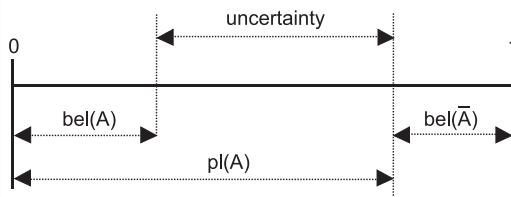
Here, it is postulated that:

- belief in a hypothesis is constituted by the sum of the masses of all sets enclosed by it;
- plausibility is defined as 1 minus the sum of the masses of all sets that contradict or reject the hypothesis.

In essence, DST uses an approach of fuzzy oppositions. The difference between a statement and the opposing statement defines the area of uncertainty.

For example, let us assume that we have the hypothesis «the train arrives on schedule». If, for this hypothesis, belief is 0.5 and plausibility 0.85, then it means that we have evidence (with a total mass of 0.5) that unequivocally indicate that the train arrives on schedule; but there also is evidence (with a total mass of 0.15) that unequivocally indicate that the train does not arrive on schedule (the belief «train does not arrive» = $1 - 0,85 = 0,15$). The remaining mass (complementing 0.5 and 0.15 to $1.0 = 0.35$), the gap





Pic. 4. Relationship between the belief function and the plausibility function.

between the plausibility of 0.85 and the belief 0.5, constitutes the «uncertainty» or evidence that the train definitely exists but asserting nothing about whether or not the train arrives on schedule. The interval $[0, 5; 0, 85]$ characterizes the uncertainty of the initial hypothesis's truthfulness based on the available evidence.

Thus, an information situation model that includes a probability characteristic makes it possible to assess uncertainty not only qualitatively but also quantitatively. This, in turn, allows the use of such assessments in automated and smart systems for the purpose of improving the quality and reliability of management.

Conclusion

The application of object and situation models in transport management is a method to improve efficiency in operating the transport industry, and a basis for improving management theory and practices. In the transport industry, modern management is making an accelerated transition from heuristic methods to methods involving automatics and smart equipment and techniques. The difficulty of this transition is caused by the difficulty of formalizing situations that contain uncertainty.

The method suggested above makes it possible to formalize such uncertainty and process situations automatically. Object and situation models reduce, by orders of magnitude, the scope of information that must be analyzed by humans, and make it possible to use such models for decision-making in situation rooms. That said, Dempster–Shafer theory is just one of many approaches to uncertainty assessment, and needs to be further developed.

REFERENCES

1. Cheharin, E. E. Big data: big problems [*Bol'shie dannye: bol'shie problemy*]. *Perspektivy nauki i obrazovaniya*, 2016, Iss. 3, pp. 7–11.
2. Tsvetkov, V. Ya. Information uncertainty and certainty in information science [*Informatsionnaya neopredelennost' i opredelennost' v naukakh ob informatsii*]. *Informatsionnye tekhnologii*, 2015, Iss. 1, pp. 3–7.
3. Kovalenko, N. I. Information approach in building world view [*Informatsionnyj podkhod pri postroenii kartiny mira*]. *Perspektivy nauki i obrazovaniya*, 2015, Iss. 6, pp. 7–11.
4. Lyovin, B. A., Mamaev, E. A., Baginova, V. V. On the concept of construction of the models of transport industrial systems [*O kontseptsii postroeniya modelej*

proizvodstvenno-transportnykh sistem]. *Nauka i tekhnika transporta*, 2003, Iss. 4, pp. 8–17.

5. Mamaev, E. A. Features of transport systems and specific properties of their modelling. Theses of the report and brief report at 25th international school-workshop named after academician S. Shatalin [*Osobennosti transportnykh sistem i spetsifika ikh modelirovaniya*]. *Sistemnoe modelirovanie sotsial'no-ekonomicheskikh protsessov*: Tezisy dokl. i soobshh. XXV mezhd. nauch. shk. – seminar im. akad. S. Shatalina. Part II. Moscow, TsEMI [Central economic and mathematical institute of Russian academy of sciences], 2002, p. 92.
6. Dешko, I. P. Information design. Monograph [*Informatsionnoe konstruirovaniye: Monografiya*]. Moscow, MAX Press publ., 2016, 64 p.
7. Tsvetkov, V. Ya. Information models of objects, processes and situations [*Informatsionnye modeli ob'ektov, protsessov i situatsiy*]. *Distantsionnoe i virtual'noe obucheniye*, 2014, Iss. 5, pp. 4–11.
8. Tsvetkov, V. Ya. Cognitive management. Monograph [*Kognitivnoe upravleniye: Monografiya*]. Moscow, MAX Press publ., 2017, 72 p.
9. Neimark, Yu. I., Kogan, N. Ya., Saveliev, V. P. Dynamic models of management theory [*Dinamicheskije modeli teorii upravleniya*]. Moscow, Nauka publ., 1985, 400 p.
10. Barabanov, I. N. [et al]. Dynamic models of information administration in social networks [*Dinamicheskije modeli informatsionnogo upravleniya v sotsial'nykh setyakh*]. *Avtomatika i telemekhanika*, 2010, Iss. 11, pp. 172–182.
11. Markelov, V. M. Application of topological models of geodata for optimization of transportation routes [*Primeneniye topologicheskikh modelej geodannykh dlya optimizatsii transportnykh marshrutov*]. *Slavyanskiy forum*, 2012, Iss. 2, pp. 56–61.
12. Pavlov, A. I. Digital modelling of spatial objects [*Tsifrovoe modelirovanie prostranstvennykh ob'ektov*]. *Slavyanskiy forum*, 2015, Iss. 4, pp. 275–282.
13. Kuzhelev, P. D. Spatial knowledge for transport management [*Prostranstvennyye znaniya dlya upravleniya transportom*]. *Gosudarstvennyy sovetnik*, 2016, Iss. 2, pp. 17–22.
14. Butko, E. Ya. Geoinformatics as a method of building of world view [*Geoinformatika kak metod postroeniya kartiny mira*]. *Slavyanskiy forum*, 2017, Iss. 1, pp. 34–41.
15. Tsvetkov, V. Ya. Resource requirements and integrative character of complex organization and technical system [*Resursnyye i integrativnost' slozhnoj organizatsionno-tekhnicheskoy sistemy*]. *Mezhdunarodnyy zhurnal prikladnykh i fundamental'nykh issledovaniy*, 2016, Iss. 5 (part 4), p. 676.
16. Narigiani, A. S. Nonfactors: brief introduction [*NE-factory: kratkoe vvedeniye*]. *Novosti iskusstvennogo intellekta*, 2004, Iss. 2, Moscow, Komkniga publ., 2006, pp. 52–63.
17. Tikhonov, A. N., Tsvetkov, V. Ya. Methods and systems of support of decision-making [*Metody i sistemy podderzhki prinyatiya resheniy*]. Moscow, MAX Press publ., 2001, 312 p.
18. Shafer, Glenn. Dempster–Shafer theory, 1992. [Electronic resource]: <http://www.glennshafer.com/assets/downloads/articles/article48.pdf>. Last accessed 30.05.2017. ●

Information about the authors:

Lyovin, Boris A. – D.Sc. (Eng), professor, rector of Russian University of Transport, Moscow, Russia, tu@miit.ru.

Tsvetkov, Victor Ya. – D.Sc. (Eng), professor, deputy head of the centre of strategic analysis and development of Research, Design and Project Institute of Railway Informatics, Automatics and Communications (NIAS), Moscow, Russia, cvj2@mail.ru.

Article received 30.05.2017, accepted 23.08.2017.