

УДК 656.212.5

А. А. САФРОНЕНКО, начальник Гомельского производственного отдела, Государственное предприятие «Институт “Белжелдорпроект”», г. Гомель

РАЗРАБОТКА ОНТОЛОГИИ ИНФРАСТРУКТУРЫ ЖЕЛЕЗНЫХ ДОРОГ КАК ОСНОВЫ ЭФФЕКТИВНОЙ ИНФОРМАТИЗАЦИИ ПОДРАЗДЕЛЕНИЙ

Автоматизация работы отдельных инфраструктурных хозяйств железных дорог проводится в основном на базе электронного документооборота отраслевых форм отчетности без учета нужд смежных хозяйств дороги. В данной статье рассмотрены вопросы создания онтологии предметной области инфраструктуры железных дорог как основы формирования объектной модели инфраструктуры. Это позволит провести масштабную информатизацию инфраструктурных подразделений железных дорог и достичь высокого качества управления инфраструктурой.

В настоящее время информатизация на железнодорожном транспорте проходит внутри отдельных хозяйств. Все вновь создаваемые информационные системы направлены на автоматизацию конкретного хозяйства. Это обусловлено тем, что внутри отдельного подразделения всегда имеется стремление снизить трудоемкость рутинных операций (как правило, задач учета) имеющегося штата инженерных работников технических отделов, выполняющих эти операции.

Особое место при рассмотрении вопросов автоматизации можно отнести подразделениям, осуществляющим эксплуатацию объектов инфраструктуры железных дорог, – инфраструктурным подразделениям: дистанциям пути, электроснабжения, сигнализации и связи, промышленных и гражданских сооружений, участкам водоснабжения. В настоящее время вся техническая документация строится из необходимости учета объектов инфраструктуры и их существенных характеристик. Документация статически фиксирует состояние объектов, но не содержит связей между ними и механизмов их взаимодействия, то есть не обладает признаками системности. Отчасти это обусловлено тем, что все утвержденные формы отчетности в хозяйствах ориентированы на бумажное хранение, а предпринимаемые попытки автоматизации ведения этих форм отчетности сводятся к повторению бумажных форм, фактически упрощающих лишь внесение изменений, оформление и тиражирование. Поэтому несмотря на неоднократные попытки приступить к автоматизации некоторых процессов внутри этих подразделений, сейчас нет ни одной отраслевой информационной системы инфраструктурного хозяйства.

Уникальность инфраструктурной информации и, как следствие, перспективных информационных систем ее обработки состоит в том, что для формирования этих систем необходима формализация не только данных об обслуживаемых в отдельном хозяйстве устройствах, но и о смежной инфраструктуре. Так, для формирования схематического плана централизации в хозяйстве сигнализации и связи необходимы данные о путевом развитии станции (обслуживается хозяйством пути) и системе электроснабжения поста централизации (обслуживается хозяйством электроснабжения). Как следствие, невозможно изолированное развертывание информационной системы одного инфраструктурного хозяйства без информатизации другого. В противном случае, сократив трудоемкость процессов внутри хозяйства, увеличим затраты на увязку данных со смежными хозяйствами.

Основным препятствием на пути информатизации инфраструктурных хозяйств является рассогласованность нотации информации в различных утвержденных формах отчетности каждого из хозяйств. Это не позволяет в настоящее время с минимальными трудозатратами производить сверку документов. При каждом согласовании между смежными подразделениями данных об объектах инфраструктуры требуется ручная конвертация данных из одной формы отчетности в другую. Поэтому в большинстве случаев при ведении технической документации принимается значительное число отступлений от нормализованной информации с целью сокращения трудоемкости поддержания системы обращения этой документации:

1) не пересматривается ординатная привязка объектов инфраструктуры при изменении геометрии пути, жестко координируются мосты, переезды, оси станций и парков, что порождает необходимость наличия «неправильных пикетов»;

2) с целью сокращения объема сверяемой информации в документации одного хозяйства не указывается существенная информация об объектах, обслуживаемых смежным хозяйством, что усложняет сбор исходных данных для принятия решений об обслуживании инфраструктуры;

3) снижаются требования к содержанию установленных форм отчетности, когда форма документа устанавливается исходя не из потребного информационного наполнения, а от возможности его наполнения с минимальными затратами труда.

1 Направление совершенствования системы хранения инфраструктурной информации.

В ходе анализа утвержденных для подразделений инфраструктурных хозяйств форм отчетности установлено, что все они сформированы исходя из «документоцентричного» подхода. То есть существенным для информации в отрасли является форма ее представления, что создает описанные выше сложности в ее актуализации. Поэтому все попытки автоматизация системы ведения технической документации «как есть» не окончились успехом.

Однако в каждом из хозяйств используется информация из смежного хозяйства, но в другой нотации. Для недопущения дублирования информации, во избежание рассогласования данных между смежными подразделениями требуется создать единую модель инфраструктурной информации дороги, которая позволяла бы:

1) упростить внедрение отраслевых информационных систем за счет закрепления структуры основных данных об объектах инфраструктуры, необходимых для формирования объектных моделей этих систем;

2) обеспечивать достоверность информации за счет единого хранилища данных;

3) повысить качество информации за счет строгого закрепления источников на уровне модели данных.

Для решения этих задач необходимо разработать структуру данных, инвариантную к отраслевым нотациям, что позволит перейти с «документоцентричного» к «датацентричному» подходу в хранении информации. Это позволит построить информационные системы ведения технической документации, структура данных которых будет сформирована исходя не из соответствия установленным формам отчетности, а из потребности простого доступа к данным. Таким образом, дальнейшее совершенствование автоматизации процессов, связанных с инфраструктурной информацией, должно вестись в два этапа:

1) разделение хранения инфраструктурной информации от внешней формы ее представления («датацентричный» подход);

2) формализация отношений и механизмов взаимодействия объектов инфраструктуры.

2 «Датацентричный» подход. Анализ форм отчетности в хозяйствах показал, что всю информацию i хозяйства можно разделить на 2 части:

$$A_i = A_i^S + A_i^O, \quad (1)$$

где A_i^S – часть информации, источником которой являются смежные хозяйства, A_i^O – часть информации, порождаемая и необходимая исключительно в рассматриваемом хозяйстве.

Задача формирования модели инфраструктурной информации сводится к поиску области, которая будет являться общеинфраструктурной или системной информацией об инфраструктуре и должна содержаться централизованным банком данных:

$$S = \cup A_i^S = \cap A_i^O. \quad (2)$$

Оставшаяся часть является отраслевой информацией и содержится отраслевыми системами отдельных хозяйств:

$$O = \cup A_i^O \setminus \cap A_i^O = \emptyset. \quad (3)$$

При этом формирование технической документации будет осуществляться как

$$A_i = f_i(S + A_i^O), \quad (4)$$

где f_i – функция преобразования нотации информации для i -го хозяйства, которая позволит избежать рассогласования данных в утвержденных формах отчетности.

Исходя из вышесказанного, необходимо разработать онтологию предметной области «инфраструктура железных дорог», что позволит:

– установить однозначное понимание терминов для упрощения информационного обмена между спе-

циалистами отрасли, а также между отраслевыми информационными системами;

– обеспечить повторное использование знаний.

Для разработки онтологии $S = \langle C, H, R, D \rangle$ необходимо выполнить следующие шаги:

– определить множество классов C ;

– установить иерархию между классами, которая определяет, что класс c_i является подклассом c_j при $c_i, c_j \in C$:

$$H = \{ \langle c_i, c_j \rangle \} \subset C \times C; \quad (5)$$

– определить отношения между классами, которые характеризуют технологическую взаимосвязь типа r_k между основным c_i и второстепенным c_j объектами инфраструктуры:

$$R = \{ \langle c_i, c_j, r_k \rangle \}; \quad (6)$$

– установить свойства классов d_i и ограничений на их значения e_i , при этом все свойства класса являются частью множества свойств подкласса:

$$\forall c \in C : \exists \bar{d} = \{ d_i, \{ e_i \}_i \}, \quad (7)$$

$$\forall h = \langle c_i, c_j \rangle \in H : \bar{d}_j \in \bar{d}_i. \quad (8)$$

Разработка онтологии $S = \langle C, H, R, D \rangle$ позволит сформировать базу знаний об объектах инфраструктуры и решать не только учетные задачи подразделений. Становится возможным создание качественно новых аналитических систем управления поездной и маневровой работы, разработка которых сковывалась отсутствием общей инфраструктурной модели железной дороги.

Первой и основной задачей является формирование множества C , составляющего онтологию, при этом необходимо произвести разграничения:

1) между классами и подклассами. Предварительный анализ инфраструктурной информации указывает на необходимость восходящей разработки множества классов: первоначальное определение самых конкретных классов с последующей группировкой этих классов в более общие понятия. При формировании классов у них проявляются наборы одинаковых свойств, свидетельствующие о необходимости группировки их как свойства класса более высокого уровня иерархии:

$$c_i, c_j \mid \bar{d}_i \cap \bar{d}_j \neq \emptyset. \quad (9)$$

$$\exists c_k : \{ \langle c_i, c_k \rangle, \langle c_j, c_k \rangle \} \in H, \quad (10)$$

$$\bar{d}_k = \bar{d}_i \cap \bar{d}_j. \quad (11)$$

Решение о выделении надкласса необходимо принимать в зависимости от технологических особенностей построения инфраструктуры;

2) между классами и их свойствами, так как при контекстном рассмотрении свойства d_j некоторого класса c_i оно может рассматриваться как самостоя-

тельный класс. В таком случае данное свойство необходимо выделять в класс

$$c_k : \langle c_k, c_i \rangle \in H, \quad (12)$$

$$\bar{d}_k = \langle d_j, \{e_i\}_j \rangle. \quad (13)$$

3 Формализация отношений. Описанное выше разделение данных от формы их представления помимо прямых эффектов совершенствования существующей системы обращения технической документации порождает сопутствующий эффект – возможность расширения перечня решаемых задач за счет дополнения модели хранения данных формальным описанием возможных отношений между ними. Модель данных строится на перечне классов предметной области, имеющих (исходя из технических и технологических особенностей объектов, описываемых этими классами) некоторые отношения между собой. Формализация отношений поможет дополнить модель данных поведенческими характеристиками объектов. Самой простой из задач, возможных к решению при помощи отношений, является проверка сценария «что-если», отображающего, какой объем изменений объектов инфраструктуры произойдет при принятии решения об изменении отдельного ее элемента. Такие задачи являются самыми многочисленными из решаемых в подразделениях дороги на основании инфраструктурной информации:

- 1) выработка технических заданий и условий на проектирование объектов;
- 2) разработка указаний по системе эксплуатации объектов;
- 3) оценка ущерба при чрезвычайных ситуациях;
- 4) расчет издержек на техническое перевооружение.

Исходя из сказанного, необходимо установление отношений между классами онтологии инфраструктуры. Отношение в общем виде описывается как кортеж $r = \langle c_i, c_j, r_k \rangle$, характеризующий технологическую взаимосвязь типа r_k между второстепенным c_i и основным c_j объектами инфраструктуры.

Иерархия. Отношения иерархии, которое характеризует подчиненность одного класса другому, выражается как «является подклассом». Устанавливается между классами, имеющими связь «частное–общее».

Стрелочный перевод характеризует элемент верхнего строения пути, предназначенный для перевода подвижного состава с одного пути на другой. По конструкции может быть обыкновенным и имеет свойства, отличные от других конструкций стрелочных переводов. Следовательно, можно выделить отношения между классами: «обыкновенный стрелочный перевод» (c_1) «является подклассом» (r_0) «стрелочный перевод» (c_0) – $\langle c_1, c_0, r_0 \rangle$.

Иерархия объектов в банке данных реализуется посредством связи таблиц базы данных дочернего и родительского классов.

Расположение. Отношение расположения, которое характеризует размещение одного объекта на другом, выражается как «расположен на». Устанавливаются для объектов, имеющих общие геометрические координа-

ты. Совпадение координат легко установить в геометрической нотации информации, которую обеспечивает масштабный план или геоинформационная система.

Стрелочный перевод укладывается в путь и имеет общую с ним геометрию. Поэтому может быть установлено отношение: «стрелочный перевод» (c_0) «расположен на» (r_1) «путь» (c_2) – $\langle c_0, c_2, r_1 \rangle$.

Для геометрического установления множества отношений между объектами необходимо для каждого класса онтологии разработать условное графическое обозначение им описываемых объектов на масштабных планах. Условные обозначения могут быть разделены на точечные, линейные и контурные.

Для восстановления отношения расположения необходимо разработать алгоритмы обработки графических представлений объектов, например:

- попадание точечного объекта на линейный: «стрелочный перевод» на «путь», «жесткая поперечина» на «опора»;
- попадание точечного объекта на точечный: «светильник» на «опора»;
- попадание линейного объекта на линейный: «ферма» на «участок теплотрассы», «ферма моста» на «путь».

Вхождение. Отношения вхождения, которые характеризуют расположение одного объекта в другом, выражается как «входит в». Устанавливается между объектами в случае, когда один объект попадает в другой контурный объект, отражающий технологическую взаимосвязь входящих в него объектов. Попадание объектов в контур легко установить в геометрической нотации информации, которую обеспечивает масштабный план или геоинформационная система.

Группа путей с одинаковой специализацией объединяется в парк, стрелочные переводы – в горловины. Поэтому может быть установлено отношение:

- «путь» (c_2) «входит в» (r_1) «парк» (c_3) – $\langle c_2, c_3, r_2 \rangle$;
- «стрелочный перевод» (c_0) «входит в» (r_2) «горловина» (c_4) – $\langle c_0, c_4, r_2 \rangle$.

Для геометрического установления отношения вхождения необходимо разработать алгоритмы обработки графических представлений объектов, например:

- попадание точечного объекта в контурный: «стрелочный перевод» в «горловина», «горловина» в «парк», «парк» в «станция»;
- попадание (пересечение) линейного объекта в контурный: «путь» в «парк».

Связь. Отношение связи устанавливается для объектов, связанных между собой либо технически, либо технологически, но не отношением расположения, выражается как «связан».

Каждому стрелочному переводу должен соответствовать предельный столбик, перекрестному стрелочному переводу – два предельных столбика. Координатно в явном виде они не связаны, но предельный столбик устанавливается относительно стрелочного перевода согласно правилу: после стрелочного перевода в середине междупутья в том месте, где расстояние между осями расходящихся путей равно нормальному междупутному расстоянию на перегонах. Между классами

«стрелочный перевод» и «пределный столбик» может быть установлено отношение: «пределный столбик» (c_3) «связан» (r_3) «стрелочный перевод» (c_0) – $\langle c_3, c_0, r_3 \rangle$.

Так как технические и технологические связи заранее известны, то необходимо алгоритмизировать процедуру установления отношений связи объектов, таких как:

– «пределный столбик» с «стрелочный перевод», «светофор» с «граница рельсовой цепи»;

– наличие общей границы контурных объектов: «узел» с «участок», «станция» с «перегон», «станция» с «подъездной путь».

Заключение. Для построения «датацентрической» системы обращения инфраструктурной информации обоснована необходимость разработки онтологии системной информации S , которая должна быть универсальной для построения любой инфраструктурной информационной системы. Вся отраслевая информация соответственно будет накапливаться в отраслевой онтологии A_i^o , формируемой и развивающейся для нужд решения задач конкретной службы. Для обеспечения соответствия форм отчетности необходимо из данных онтологий генерировать утвержденную нотацию посредством функции преобразования f_i .

Получено 12.09.2013

A. A. Safronenko. Forming of railway infrastructure ontology as a basis of railway departments informatization.

Automation of individual railway infrastructure departments is carried out mainly on the basis of the electronic document management of the department's reporting forms without the needs of the others departments. This article describes how to create a domain ontology of the railway infrastructure as the basis of the infrastructure objects model. This would allow for a large-scale computerization of railway infrastructure departments and achieve a high quality infrastructure management.

Для формирования предложенной онтологии в статье определен состав онтологии предметной области инфраструктуры железных дорог:

1) для универсальности онтологии определены правила формирования перечня классов предметной области, задания перечня их свойств и ограничений на свойства;

2) для расширения аналитических возможностей систем хранения технической документации за счет возможности учета связанности объектов инфраструктуры, выделены четыре типа отношений объектов инфраструктуры.

Список литературы

1 ГОСТ 3.1109-82. Единая система технологической документации. Термины и определения основных понятий.

2 Большая энциклопедия транспорта : в 8 т. / Н. С. Конарев [и др.]; под общ. ред. Н. С. Конарева. – М. : Большая Российская энциклопедия, 2003. – Т. 4: Железнодорожный транспорт. – 1039 с.

3 Правдин, Н. В. Железнодорожные станции и узлы / Н. В. Правдин, В. Г. Шубко. – М. : Транспорт, 2005. – 502 с.