

Чумичкин А.А., Иванцов Д.В., Хрусталев Е.Ю.

Формирование сложноструктурированных документов средствами автоматизированных систем управления

Согласно предложенному подходу подготовка документа заключается в последовательном построении его фрагментов в соответствии с алгоритмами, заложенными в графических редакторах. Реализация разработанной технологии позволит оперативно и с минимальными затратами ресурсов разрабатывать новые формы представления данных.

Анализ состояния проблемы¹. В настоящее время в системе управления развитием продукции военного назначения (ПВН) существует множество автоматизированных систем управления, реализующих функции поддержки принятия решений. В ходе решения функциональных задач должностными лицами одной из важнейших задач, решаемых такими системами, является представление информации в форме, отражающей наиболее важные аспекты анализируемой информации и удобной для анализа. Большинство таких систем предназначены для реализации аналитических задач, имеющих высокую динамику применяемых показателей и исходных данных. Одними из таких систем являются автоматизированная система управления развитием вооружения и военной техники и программно-технический комплекс начальника вооружения Вооруженных Сил Российской Федерации. Помимо этого существуют аналогичные системы в заказывающих управлениях, видах и родах войск.

В системах поддержки принятия решений (СППР) при определении способа и формы представления информации пользователю, возникает необходимость решения следующих основных задач: определение необходимого для принятия решения массива данных; выбор аспектов, интересующих пользователя; разработка алгоритма подготовки массива и подготовка формы представления. Помимо этого, практика показывает, что отмеченная ранее сложность обрабатываемой в

СППР информации делает подготовку представления данных по всем интересующим аспектам в единой форме трудоемким, анализ затруднительным, а саму форму представления не сбалансированной с точки зрения эргономики. В связи с этим на практике при решении одной задачи, как правило, формируется несколько форм представления, отражающих те или иные параметры анализируемой информации – так называемых «разрезов» массива данных.

Анализ существующих и разрабатываемых, в органах военного управления автоматизированных систем управления и поддержки принятия решений показал, что все формы представления информации в этих системах по характеру способов представления можно с определенной долей условности разделить на две основные группы (рис. 1):

1) информационно-справочного характера – как правило, формы представления информации об элементарных единицах данных в словарях и справочниках систем: образцах ПВН, отдельных мероприятиях;

2) аналитического характера – большинство из них предусматривают предварительную обработку данных (агрегирование, группировку и т.д.).

Формы представления, относящиеся к первой группе, достаточно подробно изучены и существует множество методов их реализации, которые успешно используются в рассматриваемых автоматизированных системах управления. В связи с этим, а также с низкой динамикой таких форм, в данной статье рассматриваются формы представления данных второго из перечисленных типов, представляющие наибольший интерес. Вместе с тем, необходимо отметить, что некоторые формы являются выходными докумен-

¹ Статья подготовлена при поддержке гранта РФФИ (проект № 09-07-13516) и гранта президента по поддержке молодых ученых кандидатов наук № 02.120.11.8–МК

тами и бывают как первого, так и второго типов, но обладают достаточно высокой динамикой структуры и состава содержащейся информации, поэтому такие формы тоже будут исследоваться. Рассматриваемые формы, как правило, представляют собой набор таблиц, диаграмм, рисунков и текстовых фрагментов, определенным образом расположенных на ограниченном пространстве. Иногда в таких формах информация представляется иначе, например, как визуализация данных на картах или других объектах и т.д. Однако такие способы визуализации используются лишь в определенных автоматизированных системах управления, например, в тех, которые используют геоинформационные системы, и в разрабатываемой технологии не будут применяться.

Таблицы данных форм бывают различной сложности, информация в них представляется в виде строк и столбцов, а в качестве дополнительных измерений используются па-

раметры шрифтов и заливок. В настоящее время для представления информации в автоматизированных системах управления используется широкий спектр типов диаграмм, и выбор наиболее приемлемых осуществляется в зависимости от задач и представляемых данных.

Вторым способом реализации форм представления является вывод данных в различного рода графические редакторы и подготовка документов или форм анализа и контроля специалистами. Данный способ имеет как ряд достоинств, так и ряд недостатков. К достоинствам можно отнести возможности индивидуализации документов и широкие возможности по их оформлению. На практике наиболее целесообразно использовать данный способ при подготовке таких документов как презентации, доклады, аналитические записки и т.д.

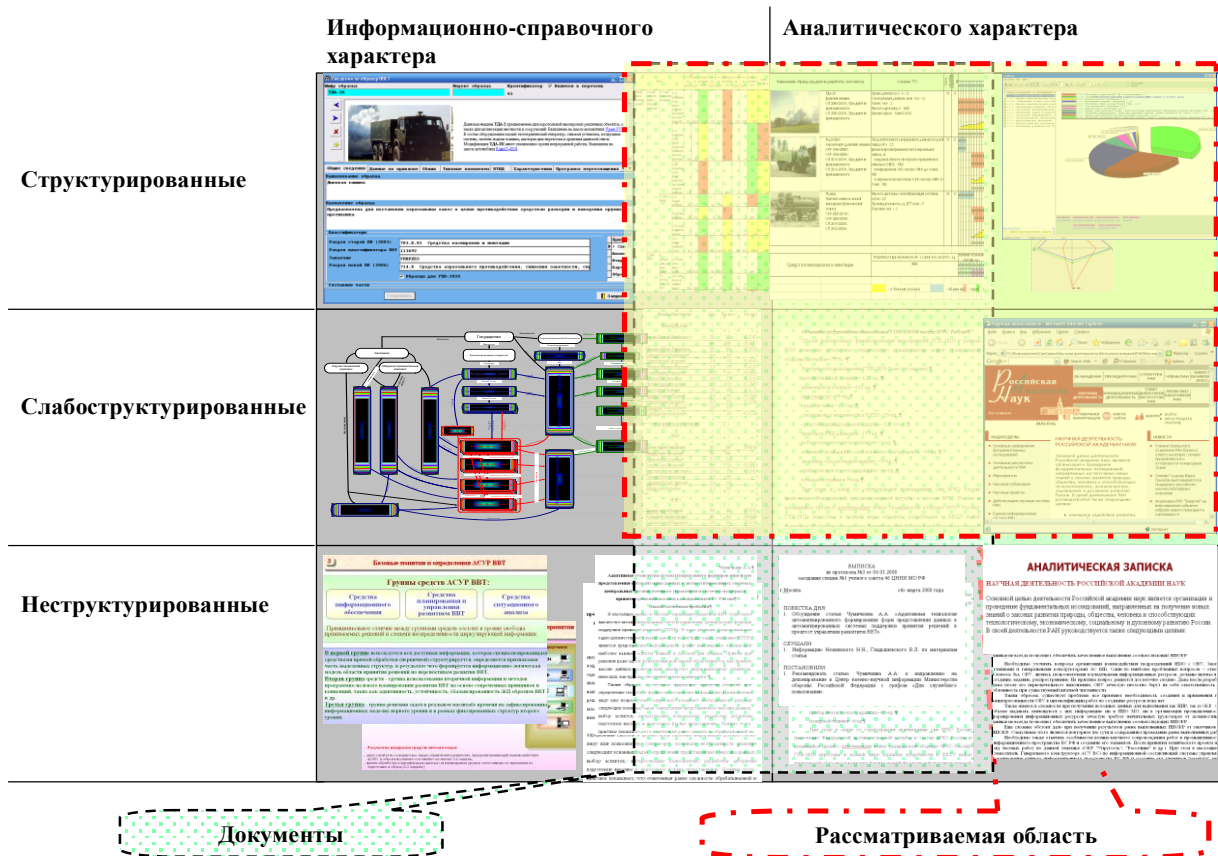


Рисунок 1. Группы форм представления данных

В качестве недостатков можно выделить высокую трудоемкость оформления, которая сильно зависит от объема обрабатываемых данных. Но основным недостатком данного подхода все-таки является отсутствие системы хранения шаблонов и структур документов, что приводит к интеллектуальным потерям, так как в процессе их разработки проводится большой объем интеллектуальной работы, результаты которой фиксируются в этих шаблонах. Так, например, в интересах решения какой-либо из функциональных задач разрабатывается система форм. Процесс ее разработки связан с большим количеством расчетов, экспериментов и согласований.

Однако, через некоторое время (два-четыре года) после ее разработки, при рассматриваемом способе реализации, для решения той же самой (или аналогичной) задачи разработка системы форм производится заново. Это распространенная в организациях проблема, которая связана с динамикой кадров и является частью общей проблемы управления знаниями.

Проведенный анализ существующих подходов к формированию форм представления информации в автоматизированных системах органов управления выявил противоречие между необходимостью поддержания форм представления информации в СППР в состоянии, адекватном изменяющимся условиям, и затратностью доработки соответствующего СПО, приводящей порой к необходимости создания отдельных его элементов заново.

Постановка задачи и разработка замысла ее решения. В каждой из рассмотренных автоматизированных систем необходима технология автоматизированного формирования форм представления данных, которая будет соответствовать следующим, на первый взгляд, противоречивым требованиям:

– должна обеспечиваться возможность автоматизированного формирования форм представления данных;

– должно обеспечиваться сохранение форм, алгоритмов подготовки исходных данных и способов визуализации;

– должна обеспечиваться возможность создание новых и корректировки всех параметров существующих форм представления;

– должна обеспечиваться независимость форм представления от изменения данных;

– должна обеспечиваться возможность использования элементов дизайнерского искусства и основных возможностей современных графических редакторов;

– и, при всем этом, *не должны требоваться изменения* никаких элементов автоматизированной системы (в том числе СПО) кроме информационного обеспечения.

То есть с одной стороны существует необходимость непрерывного изменения существующей системы форм в интересах ее адаптации меняющимся условиям, а с другой – отсутствует возможность такой динамики программного обеспечения АСУ.

В интересах разработки такой технологии был проведен анализ существующих форм представления данных и по степени структуризации выявлены следующие их группы (рис. 1):

– структурированные формы – формы, имеющие четко определенную структуру, массив данных и правила визуализации (сводные таблицы, диаграммы, альбомы, справочники и т.д.);

– слабо структурированные – формы, имеющие определенную структуру, но имеющие текстовые элементы, содержание которых зависит от ситуационно ориентированных структурированных данных (пояснительные записки к структурированным документам, справки и т.п.)

– неструктурированные – текстовые или графические документы, в которых невозможно однозначно определить как структуру информации, получаемой из базы данных, так и форму изложения.

Наибольший интерес, с точки зрения разработки автоматизированной технологии формирования форм представления информации, является первая и вторая группы, которые и являются предметом исследования.

Рассмотрим подробнее структурированные формы представления. К этой группе, согласно приведенному ранее определению, можно отнести подавляющее большинство форм представления, начиная от простейших таблиц и заканчивая сложными документами, включающими сложноструктурированные таблицы, диаграммы различных типов, а также рисунки

и фотографии. Несмотря на все разнообразие таких форм, все они состоят из трех основных элементов:

- массива исходных данных, на основе которого строится форма;
- структуры формы, определяющей расположение и компоновку ее элементов;
- правила построения, визуализации формы (например, выделения цветами или шрифтом тех или иных элементов в зависимости от значений данных).

Исходя из того, что массив данных необходимый для построения форм, должен выбираться из базы данных той или иной автоматизированной системы, вместо него будем рассматривать правило получения из базы данных такого массива.

Таким образом, каждую из форм можно описать как:

$$f(\Omega, \theta, \beta), \quad (1)$$

где f – форма представления данных, Ω – правило получения данных, θ – структура (шаблон) формы, β – правило (алгоритм) построения и визуализации.

На практике одним из видов модификации форм является изменение параметров массивов исходных данных, в связи с этим введем для Ω множество параметров $X = \{x_1, x_2 \dots x_n\}$, где x – параметры-ограничения базы данных. Таким образом, правило получения массива данных примет вид: $\Omega(X)$.

Помимо массивов данных у форм, имеющих одну и ту же структуру, могут изменяться параметры правил построения и визуализации. Аналогично массивам данных введем функцию параметров $\lambda(y)$ и правило визуализации примет следующий вид: $\beta(\lambda(y))$.

Если же изменениям подвергается структура формы, будем считать, что это уже другая форма представления данных.

Теперь множество форм представления данных, использующихся в автоматизированной системе можно описать как:

$$F = f_i(\Omega_j(X), \theta_i, \beta_k(\lambda(y))), \quad (2)$$

где $i \in I$ – номер формы представления, $j \in J$ – номер правила выборки исходных данных, $k \in K$ – номер правила визуализации.

Таким образом, разработка технологии, реализующей хранение и реализацию множества форм, представленных соотношением (2), позволит обеспечить ее соответствие рассмотренным ранее требованиям.

Технологические решения. Как отмечено ранее, одним из основных требований к разрабатываемой технологии является то, что она не должна требовать изменения элементов автоматизированной системы (в том числе и СПО) кроме информационного обеспечения. То есть специальное программное обеспечение, входящее в состав АСУ должно быть разработано однажды, а изменяться может только информация, которая может храниться как в базе данных под управлением системы управления базами данных (СУБД), так и файловой системе. Рассмотрим реализацию основных элементов форм.

Так как все операции в базе данных под управлением СУБД, использующихся в рассматриваемых АСУ, осуществляются посредством выполнения запросов на языке SQL, правила выборки информации из базы данных $\Omega_j(X(x))$ будут представлять собой SQL-запросы с параметрами. Множество этих запросов необходимо организовать в библиотеку запросов, которая будет храниться в базе данных и хранить следующую информацию: идентификатор запроса – j , неформальное (текстовое) описание запроса, текст запроса на языке SQL, описание параметров запроса $X(x)$.

Следующий элемент – структуру форм представления информации θ_i наиболее целесообразно реализовывать в виде шаблонов документов графических редакторов, например ПС-Таблиц (MS Excell для систем, использующих ОС Windows). В некоторых случаях можно использовать и шаблоны других форматов таких как DBML, XML и др., однако преимуществом предложенного способа является возможность использования инструментария графических редакторов на этапе автоматизированного построения, а также в процессе изменения формы после его окончания.

Правила формирования и визуализации форм отображения данных будем описывать в виде алгоритмов, реализованных на языке программирования, выбираемом в соответствии с

шаблоном формы, рассмотренным выше. Хранение правил построения и их параметров $\beta_k(\lambda(y))$ необходимо организовать в соответствующей библиотеке.

Множество форм представления информации будет храниться в библиотеке форм. Каждая форма определяется:

- идентификатором;
- наименованием;
- неформальным описанием (комментарием);
- алгоритмом;
- идентификатором правила выбора исходного массива информации;
- ссылкой на шаблон формы;
- идентификатором правила построения и визуализации.

Следующая задача – разработка специального программного обеспечения, реализующего разработанную технологию. Данное СПО должно быть одним из модулей АСУ и обеспечивать решение следующих основных задач:

- ведение библиотеки форм представления данных;
- осуществление поиска и выбора необходимых для построения форм;
- изменение параметров элементов форм;
- выполнение алгоритма построения выбранных форм с заданными параметрами их элементов.

Разработанная технология обладает следующими характеристиками:

- за счет алгоритмизации всех процессов обеспечивает автоматизированное формирование форм представления данных;
- обеспечивает сохранение форм, алгоритмов подготовки исходных данных и способов визуализации;
- обеспечивает возможность создания новых и корректировки параметров существующих форм представления;
- обеспечивает независимость форм представления от изменения данных;
- за счет внедрения соответствующих шаблонов обеспечивает возможность использования элементов дизайнерского искусства и основных возможностей графических редакторов;
- требует изменения только элементов информационного обеспечения.

Следовательно, созданная адаптивная технология удовлетворяет предъявляемым к ней требованиям. Поэтому закономерен вывод о том, что ее реализация позволит разрешить существующее противоречие между необходимостью поддержания форм представления информации систем поддержки принятия решений в состоянии, адекватном изменяющимся условиям программного управления развитием ПВН, и невозможностью непрерывной доработки соответствующих элементов АСУ.

Формирование сложноструктурированных документов

Рассматриваемые формы представления информации, как правило, представляют собой набор таблиц, диаграмм, рисунков и текстовых блоков, определенным образом расположенных на ограниченном пространстве. Ввиду этого, такие формы представления данных различаются по сложности, составу информации и способам представления информации. В качестве дополнительных измерений используются параметры шрифтов и заливок, в зависимости от задач и представляемых данных используется широкий спектр типов диаграмм и графических объектов. Таким образом, любую форму представления данных можно разделить на несколько простейших, но в тоже время цельных фрагментов, выстраивание которых в различные структуры позволяет получать различные формы представления данных.

В настоящее время в автоматизированных системах управления существует практика реализации представления данных тремя основными способами [1]. Первый заключается во внедрении соответствующих модулей, форм и алгоритмов непосредственно в специальное программное обеспечение [2,3]. Данный способ наиболее приемлем для построения простых форм представления данных. Но, не смотря на это, он применяется и для реализации более сложных форм представления данных, что не всегда является рациональным. Это обусловлено тем, что процесс разработки новых программных модулей, форм и алгоритмов может существенно отстать от высокой динамики изменения как структуры, так и содержания форм представления данных и тем более от непрерывного изменения форм выходных документов.

Второй способ заключается в выводе данных в различного рода графические редакторы [4] и подготовке документов или форм анализа и контроля специалистами вручную [5]. К достоинствам данного способа можно отнести возможности индивидуализации документов и широкие возможности по их оформлению. Основными недостатками данного способа является высокая трудоемкость оформления и отсутствие системы хранения шаблонов и структур документов. На практике наиболее целесообразно использовать данный способ при подготовке таких документов как презентации, доклады и аналитические записки [6].

При третьем [7] подходе осуществляется вывод данных в различного рода графические редакторы и подготовка документов или форм анализа и контроля правилами (алгоритмами) построения и визуализации заложенными в графических редакторах. Данный подход лишен недостатков первых двух способов и в настоящее время является передовым и широко применяется при разработке форм представления в автоматизированных системах управления. Несмотря на это, данный способ также имеет ряд недостатков, основным из которых является жесткое соответствие шаблона представления данных правилам (алгоритмам) построения и визуализации, сложность которых пропорциональна сложности форм представления данных. Так, например, для внедрения новой формы представления данных необходимо разрабатывать новое правило (алгоритм) построения и визуализации, что требует определенных трудовых и временных ресурсов. Поэтому применение данного способа ограничивается только определенным типом форм. Актуальность совершенствования данного метода обусловлена тем, что в настоящее время формы представления информации приобретают более сложный характер, что влияет на состав содержащейся в них информации, которую невозможно передать в графический редактор в рамках одного обращения к системе управления базами данных (СУБД).

Проведенный анализ существующих методов формирования форм представления информации в автоматизированных системах управления показал, что на сегодняшний день самым совершенным является третий метод, но при его использовании, в процессе разра-

ботки новых форм представления информации, необходимо разрабатывать новые уникальные правила (алгоритмы) построения и визуализации, которые зачастую имеют высокую степень сложности, что требует затрат определенных ресурсов. Также использование этого метода ограничивается невозможностью представления некоторых сложноструктурированных форм.

Обобщая вышеизложенное, можно сделать вывод, что в автоматизированных системах управления существует определенный класс форм представления данных, для которых в настоящее время не существует методов автоматизированного формирования удовлетворяющим современным требованиям. Все это обуславливает актуальность разработки технологии автоматизированного формирования форм представления данных, которая будет соответствовать следующим требованиям:

- обеспечивать возможность автоматизированного формирования форм представления данных;
- обеспечивать возможность автоматизированной разработки и корректировки форм представления данных;
- обеспечивать сохранение структур форм представления информации и алгоритмов подготовки исходных данных для них;
- основываться на использовании библиотек базовых простейших правил (алгоритмах) построения и визуализации;
- обеспечивать независимость форм представления от изменения данных;
- обеспечивать возможность использования элементов дизайнерского искусства и основных возможностей современных графических редакторов;
- не должна требовать изменения никаких элементов автоматизированной системы (в том числе СПО) кроме информационного обеспечения.

В интересах разработки такой технологии был проведен анализ существующих форм представления данных. На основе результатов проведенного анализа формы представления информации, с определенной долей условности, можно разделить на три основные группы:

- структурированные формы – формы, имеющие четко определенную структуру, массив данных и правила визуализации (сводные

таблицы, диаграммы, альбомы, справочники и т.д.);

– слабо структурированные – формы, имеющие определенную структуру, но имеющие текстовые элементы, содержание которых зависит от ситуационно ориентированных структурированных данных (пояснительные записки к структурированным документам, справки и т.п.)

– неструктурированные – текстовые или графические документы, в которых невозможно однозначно определить как структуру информации, так и форму изложения.

Наибольший интерес, с точки зрения разработки автоматизированной технологии формирования форм представления информации, является первая и вторая группы, третья группа в данной работе рассматриваться не будет.

Рассмотрим подробнее структурированные формы представления. К этой группе, согласно приведенному ранее определению, можно отнести подавляющее большинство форм представления данных, начиная от простейших

Каждый из этих фрагментов содержит данные, относящиеся к одному предприятию, поэтому каждому фрагменту необходим свой, относящийся к конкретному предприятию массив исходных данных, а правило построения и визуализации однотипных фрагментов остается постоянным для любого из этих предприятий. Ввиду этого, данный перечень предприятий можно представить в виде структуры повторяющихся фрагментов (**рисунок 2**).

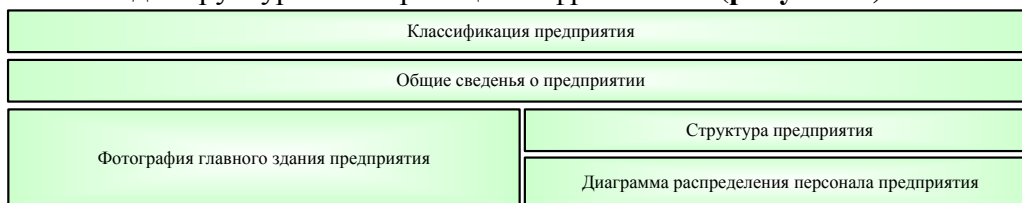


Рисунок 2 – Структура перечня предприятий

Таким образом, на основе рассмотренного примера можно сделать вывод, что любой сложноструктурированный документ можно представить в виде структуры его повторяющихся фрагментов, что позволяет формализовать любую структурированную форму представления данных, ввиду того, что все они состоят из четырех основных элементов:

– набор фрагментов, из которых строится форма (любую форму можно разделить на несколько фрагментов, которые повторяются в ней с некоторой периодичностью и в совокупности образуют форму представления информации в целом);

таблиц и заканчивая документами, которые состоят из различных комбинаций таблиц, диаграмм, рисунков, фотографий, — сложноструктурированные документы. Пример сложноструктурированного документа представлен на **рисунке 1**.

Форма представления данных о перечне предприятий, который включает в себя их классификацию согласно некому классификатору, а также основную информацию о них. Данные о предприятиях состоят из следующих фрагментов:

- классификация предприятия — представляет собой текстовый блок;
- общие сведения о предприятии — представляет собой сводную таблицу;
- изображение главного здания предприятия — представляет собой фотографию;
- структура предприятия — представляет собой иерархическую структуру;
- распределение персонала предприятия — представляет собой круговую диаграмму.

– структура формы (определяет расположение и компоновку фрагментов, из которых состоит форма);

– массив исходных данных, на основе которого строятся фрагменты (информация, которая является содержанием фрагмента формы);

– правила построения и визуализации фрагментов (алгоритм определяющий способ представления содержания фрагмента).

Исходя из того, что массив данных необходимый для построения фрагментов, должен выбираться из базы данных той или иной автоматизированной системы, вместо него будем рассматривать правило получения из базы данных такого массива.

Так как любой документ рассматриваемого типа можно разделить на несколько фрагментов, совокупность которых образует структуру документа, то в общем случае документ можно описать как:

$$f(\alpha, \beta) \quad (1)$$

где α — матрица параметров документа, β — матрица описания структуры документа.

Матрица параметров документа представляет собой таблицу значений различных параметров размерностью $l \times k$ и описывается следующим образом:

$$\alpha = \begin{bmatrix} p_{1,1} & \dots & p_{1,l} \\ \vdots & \ddots & \vdots \\ p_{k,1} & \dots & p_{k,l} \end{bmatrix} \quad (2)$$

где p — значение параметра.

В свою очередь матрица описания структуры документа представляет собой таблицу фрагментов документа размерностью $m \times n$ расположенных в определенном порядке и описывается следующим образом:

$$\beta = \begin{bmatrix} f_{1,1}(\Omega, \theta, \gamma) & \dots & f_{1,m}(\Omega, \theta, \gamma) \\ \vdots & \ddots & \vdots \\ f_{n,1}(\Omega, \theta, \gamma) & \dots & f_{n,m}(\Omega, \theta, \gamma) \end{bmatrix} \quad (3)$$

$$\beta = \begin{bmatrix} f_{1,1}(\Omega_j(M_i), \theta_g, \gamma_k(\varphi(y))) & \dots & f_{1,m}(\Omega_j(M_i), \theta_g, \gamma_k(\varphi(y))) \\ \vdots & \ddots & \vdots \\ f_{n,1}(\Omega_j(M_i), \theta_g, \gamma_k(\varphi(y))) & \dots & f_{n,m}(\Omega_j(M_i), \theta_g, \gamma_k(\varphi(y))) \end{bmatrix} \quad (4)$$

где $i \in I$ — номер строки матрицы параметров документа, $g \in G$ — номер формы представления фрагмента, $j \in J$ — номер правила выборки исходных данных для построения фрагмента документа, $k \in K$ — номер правила визуализации фрагмента документа.

Таким образом, разработка технологии, реализующей хранение и реализацию множества форм, описанных в (4), позволит обеспечить ее соответствие рассмотренным ранее требованиям.

где f — фрагмент документа, Ω — правило получения данных, θ — структура (шаблон) фрагмента, γ — правило (алгоритм) построения и визуализации.

Матрицы α и β связаны между собой посредством Ω , причем таким образом, что некоторые элементы строки матрицы α являются элементами множества параметров Ω . Это позволяет модифицировать фрагменты документа путем изменения массива исходных данных. Таким образом, для Ω множеством параметров является подмножество элементов i -ой строки матрицы — M_i , и правило получения массива данных примет вид — $\Omega(M_i)$, где $M_i \subset P_i$, а $P_i \in \{p_{i,1} \dots p_{i,l}\}$.

Помимо массивов данных у фрагментов, имеющих одну и ту же структуру, могут изменяться параметры правил построения и визуализации. Аналогично массивам данных введем функцию параметров $\varphi(y)$ и правило визуализации примет следующий вид: $\gamma(\varphi(y))$.

Теперь матрицу описания структуры документа можно описать как:

Технологические решения

Формирование документов, должно быть основано на взаимодействии следующих элементов:

1. Специального программного обеспечения (СПО);
2. Информационного обеспечения (ИО);
3. Шаблон документа;
4. Шаблон фрагментов.

Обобщенная схема формирования документа представлена на рисунке 3.

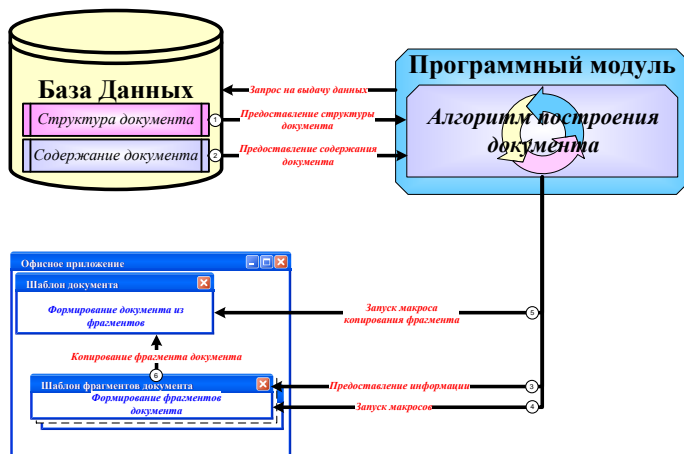


Рисунок 3 – Обобщенная схема формирования документа

Из рисунка 3 видно, что СПО содержит в себе алгоритм построения документа и является связующим звеном между базой данных и шаблоном документа. Этим достигается одно из основных требований к разрабатываемой технологии — она не должна требовать изменения СПО. То есть СПО должно быть разработано однажды.

В ИО в свою очередь, должна храниться информация, необходимая для построения документа, что обеспечит гибкость данной технологии, которая достигается путем изменении информации описывающей правила построения документа, а не изменением СПО. Также посредством ИО, а именно СУБД, осуществляется выборка информации содержащейся в документе по правилам $\Omega_j(M_i)$, которые представляют собой SQL-запросы с параметрами.

Экземпляры шаблона документа и шаблона фрагментов должны быть реализованы в виде шаблонов документов графических редакторов, например MS Excell для систем, использующих ОС Windows. Экземпляр шаблона фрагментов должен содержать в себе структуру (шаблон) форм представления информации — θ_g , а также правило (алгоритм) построения и визуализации — $\gamma(\varphi(y))$, а экземпляр шаблона документа должен осуществлять расстановку фрагментов документа в соответствии с матрицей описания структуры документа β .

В процессе формирования документа СПО сначала открывается экземпляр шаблона документа, который остается открытым на всем

протяжении процесса формирования документа, а экземпляр шаблона фрагментов открывается всякий раз, когда происходит формирование очередного фрагмента документа, при этом предыдущий экземпляр закрывается. Взаимодействие с СПО должно быть организовано так, что сначала СПО передает в экземпляр шаблона фрагментов данные, а потом запускает в нем соответствующий макрос, после выполнения которого, готовый фрагмент документа копируется в экземпляр шаблона документа. Для выполнения копирования СПО запускает в экземпляре шаблона документа макрос копирования фрагмента документа.

Как видно из данной обобщенной схемы, процесс формирования документа целиком возлагается на СПО, за исключение подпроцесса преобразования данных в готовые фрагменты документа и их копирования в сам документ.

Процесс формирования документа описывается алгоритмом, которым оперирует СПО. Данный алгоритм представлен на рисунке 4.

Для рассмотрения алгоритма формирования документа необходимо ввести следующие понятия:

1. Таблица параметров – данные, представляющие собой массив ограничений для запросов фрагментов, которые представлены в виде таблицы и являются результатом выполнения запроса-параметра.
2. Запрос-параметр – текст SQL-запроса, результатом которого является таблица параметров.
3. Запрос фрагмента – текст SQL-запроса, результаты выполнения которого помещаются в экземпляр шаблона фрагментов документа.

4. Таблица структуры – данные, представляющие собой вербальное описание структуры документа, представлены в виде таблицы.

Алгоритм формирования документа начинается с ввода идентификатора документа (блок 1) – происходит выбор документа, который необходимо построить.

Далее определяется таблица параметров, которая соответствует выбранному документу (блок 2), т.е. выполняется запрос-параметр, который соответствует выбранному документу.

На следующем шаге алгоритма определяется структура выбранного документа (блок 3).

Следующим шагом алгоритма организовывается цикл по строкам таблицы параметров (блок 4), тем самым осуществляется перебор всех параметров документа.

После организации цикла по строкам таблицы параметров алгоритм открывает экземпляр шаблона документа (блок 5), который будет открыт на протяжении всего периода формирования документа. Экземпляр шаблона документа предназначен для сбора фрагментов будущего документа.

Далее в алгоритме организовывается цикл по строкам таблицы структуры (блок 6), тем самым организовывается перебор всех фрагментов будущего документа, которым соответствует одна строка таблицы структуры.

Согласно рассмотренному порядку на следующем шаге алгоритма проверяется условие соответствия наименования параметра текущего запроса фрагмента, признаку рубрикатора (блок 7). Реализация проверки выполнения данного условия организована таким образом, что если параметр является рубрикатором, то его наименование начинается со следующих символов «Rubr». Особое рассмотрение рубрикаторов в алгоритме обусловлено необходимостью классификации информации в рамках формируемого документа, основным способом организации которой является вертикальная классификация.

Рассмотрим случай, когда условие, содержащееся в блоке 7 выполняется, тогда следующим шагом алгоритма является проверка факта отличия текущего рубрикатора (значение текущего параметра запроса фрагмента из таблицы параметров) от предыдущего (значение поля содержащего соответствующий руб-

рикатор в предыдущей строке таблицы параметра), что соответствует блоку 12.

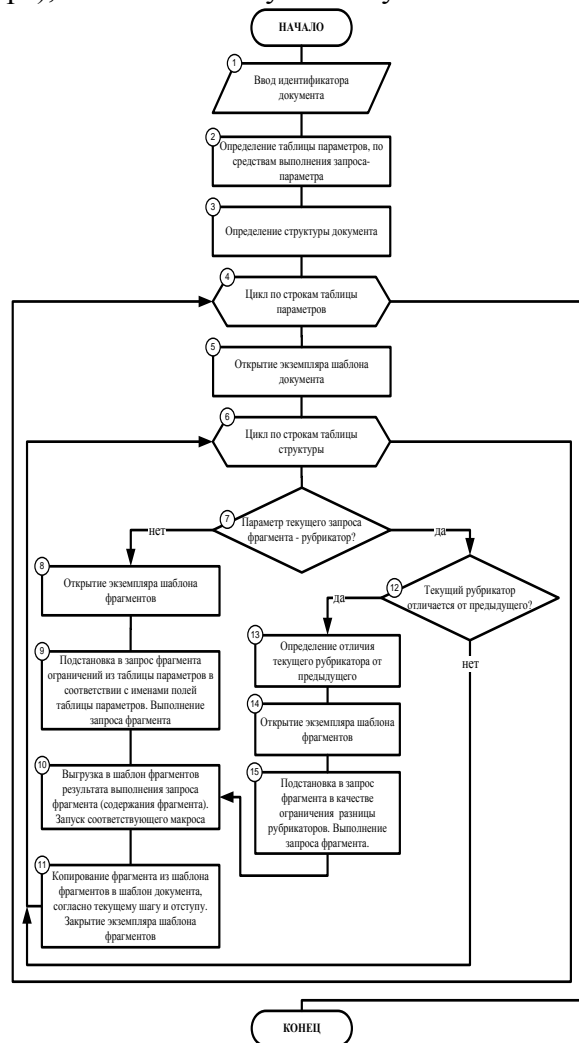


Рисунок 4 – Алгоритм формирования документа

Проверка данного условия необходима для того, чтобы алгоритм пропустил построение текущего фрагмента документа (узла классификатора) в случае, если значение рубрикатора в текущей строке таблицы параметров не отличается от его значения в предыдущей строке.

То есть, если условие в блоке 12 не выполняется, то алгоритм переходит к блоку 6, тем самым осуществляется переход к следующей итерации цикла по строкам таблицы структуры.

Если же, условие блока 12 выполняется, то следующим шагом алгоритма является определение отличия текущего рубрикатора от предыдущего (блок 13). Пример выполнения данной операции представлен на рисунке 5

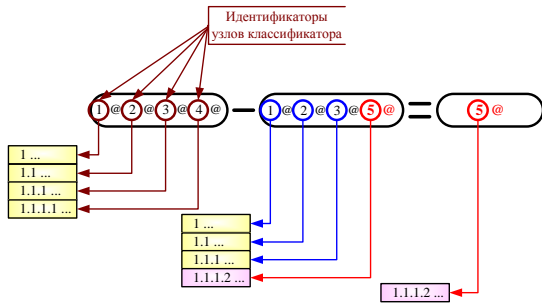


Рисунок 5 – Пример операции определения отличия одного рубрикатора от другого.

Таким образом, результатом операции сравнения текущего и предыдущего рубрикатора является подстрока текущего рубрикатора, которая содержит идентификаторы не встречающихся ранее узлов классификатора (разница рубрикаторов). Выполнение данного шага алгоритма исключает повторное отображение узлов классификатора в документе.

На следующем этапе алгоритма открывается экземпляр шаблона фрагментов (блок 14), который предназначен для построения текущего фрагмента документа.

Далее алгоритм переходит к блоку 15, где осуществляется подстановка в запрос фрагмента в качестве ограничения разницы рубрикаторов определенной в блоке 13, после чего готовый запрос фрагмента выполняется.

Далее алгоритм переходит к блоку 10, где происходит выгрузка в шаблон фрагментов данных, необходимых для построения фрагмента — содержания фрагмента (результат выполнения текущего запроса фрагмента) и запуск соответствующего макроса, идентификатор которого содержится в текущей строке таблицы структуры. Результатом выполнения данного шага алгоритма является готовый фрагмент документа.

Следующим шагом алгоритма является копирование готового фрагмента в шаблон документа, согласно текущему шагу и отступу, а также закрытие экземпляра шаблона фрагментов (блок 11). При копировании фрагмента документа используются следующие типы шагов: новая строка; вправо; вниз; вверх вправо.

Иллюстрация данных типов шагов представлена на рисунке 6.

Таким образом, основной задачей данного шага алгоритма является последовательное формирование документа, в соответствии с

его структурой, после чего алгоритм переходит к следующей итерации цикла по строкам таблицы структуры.

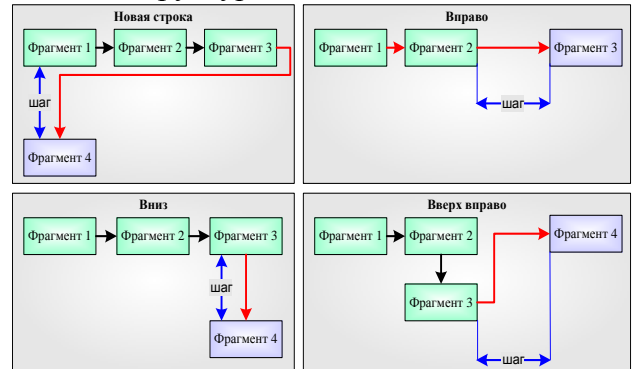


Рисунок 6 – Типы шагов при копировании фрагментов документа

Далее рассмотрим случай, когда условие блока 7 не выполняется. В этом случае алгоритм переходит к блоку 8, который идентичен блоку 14. Далее алгоритм переходит к блоку 9, который аналогичен блоку 15, за исключением того, что в качестве параметра запроса фрагмента используется не отличие рубрикаторов, а конкретное значение из таблицы параметров. Следует отметить, что в рассматриваемом алгоритме используется особый порядок подстановки параметров из таблицы параметров в запрос фрагмента. Иллюстрация данного процесса представлена на рисунке 7



Рисунок 7 – Порядок подстановки параметров в запрос фрагмента

Согласно рисунку 8 порядок подстановки параметров в запрос фрагмента заключается в том, что наименование параметра в тексте запроса фрагмента подменяется значением поля текущей строки таблицы параметров, имя которого идентично наименованию параметра запроса фрагмента, за исключением случая с рубрикатором, который был описан ранее.

После алгоритм проходит рассмотренные ранее блоки 10 и 11, после которых так же переходит к следующей итерации цикла по строкам таблицы структуры.

По завершению цикла по строкам таблицы структуры (блок 6), алгоритм переходит к следующей итерации цикла по строкам таблицы параметров, по окончанию которого алгоритм завершается.

Выводы. Таким образом, разработанная технология обладает следующими характеристиками:

- за счет алгоритмизации всех процессов обеспечивает автоматизированное формирование форм представления данных;
- обеспечивает автоматизированные разработку и корректировку структуры документов;
- обеспечивает сохранение структур документов, и алгоритмов подготовки исходных данных для них;
- основывается на использовании библиотек базовых простейших правил (алгоритмах) построения и визуализации;
- обеспечивает независимость структур документов от изменения данных;
- за счет внедрения соответствующих шаблонов фрагментов обеспечивает возможность использования элементов дизайнерского искусства и основных возможностей графических редакторов;
- требует изменения только элементов информационного обеспечения;

- обеспечивает возможность построения сложных структур документов, путем создания их на основе взаимосвязей различных по содержанию и сложности фрагментов.

Из этого следует, что разработанная технология удовлетворяет описанным выше требованиям. В связи с этим, можно сделать вывод, что реализация разработанной технологии позволит оперативно и с минимальными затратами ресурсов разрабатывать новые формы представления данных путем построения структур документов из отдельных относительно простых фрагментов, а также осуществлять автоматизированное формирование этих документов.

На основе результатов анализа существующих и разрабатываемых автоматизированных систем управления, реализующих функции поддержки принятия решений, а также с учетом положительного опыта использования разработанной технологии автору видится целесообразным внедрение данной технологии в данные системы. Внедрение разработанной технологии необходимо проводить как одно из базовых средств автоматизации.

Данная технология реализована в программно-техническом комплексе (ПТК) «Арбат-НВ-Центр» и успешно эксплуатируется при решении функциональных задач, связанных с использованием ПТК. В настоящее время данная технология используется при формировании документов планирования перспектив развития системы вооружения Вооруженных Сил Российской Федерации.

Список использованных источников

- 1 Соловьев А.В. «Обзор средств генерации отчетов». // Научный сервис в сети Интернет
- 2 Большаков С.А. «Технологии конечного пользователя для формирования выходных документов информационных систем». // Научный сервис в сети Интернет
- 3 Большаков С.А. «Как лучше и быстрее генерировать отчеты информационных систем». // Научный сервис в сети Интернет
- 4 В. О. Ким, Д. А. Шалашов. «Автоматизированное формирование полнотекстовых баз данных на основе документов Word». М: МИФИ, 2007.
- 5 Осипов С.А. «Автоматизация написания сложных отчетов». // Научный сервис в сети Интернет, 2003.
- 6 Зеленская Е.В. «Автоматическая генерация выходных форм FoxPro приложений на основе абстрактных спецификаций». // Научный сервис в сети Интернет
- 7 Хрусталева Е.Ю., Чумичкин А.А., «Адаптивная технология подготовки форм представления данных в автоматизированных системах управления развитием продукции военного назначения». Прикладная информатика. 2009. №1.