

АВТОМАТИЗИРОВАННОЕ УПРАВЛЕНИЕ ПОТОКАМИ ОПЕРАТИВНЫХ ДАННЫХ

© 2006 Т. И. Михеева

Самарский государственный аэрокосмический университет

Рассматривается автоматизированная информационная система сбора, обработки и анализа оперативной информации, разрабатываемая в рамках транспортной системы. Приведена модель данных, описаны инструменты проектирования базы данных оперативной информации, рассмотрены алгоритмы выявления мест концентрации объектов на карте.

Введение

Многогранность процессов, протекающих в жизни современного города, находит отражение в решениях, принятых соответствующими структурами градоуправления. Управление транспортной инфраструктурой, обеспечивающее решение не только экономических вопросов, но и вопросов, связанных с обеспечением безопасности, в настоящее время является особенно актуальным. Оперативная информация о состоянии улично-дорожной сети, обстановке (безопасности) в городе характеризуется комплексом оценок состояния множества объектов. Обеспечение сбора, накопления и обработки такой информации необходимо решать на уровне автоматизированной информационной системы для получения возможности своевременного оперативного принятия решения в условиях реального времени.

1. Автоматизированная информационная система управления потоками оперативных данных

В рамках проекта «Интеллектуальная транспортная система г. Самары» разрабатывается автоматизированная информационная система (АИС) мониторинга оперативных данных, отражающих состояние транспортной инфраструктуры города.

Основу информационного обеспечения АИС составляет совокупность распределенных данных. Сложная организация информационного обеспечения системы диктует необходимость использования *базы данных* с расширенной системой управления этими данными, выполняющей функции не только

хранения, но и поиска и манипулирования ими [1, 2].

АИС представляет собой комплекс взаимосвязанных программных модулей, каждый из которых предназначен для решения задачи обработки одного из видов оперативной информации. База данных является фундаментальным компонентом информационной системы и представляет собой поименованную совокупность данных, отображающую состояние объекта или класса объектов, их свойства и взаимоотношения [3, 4].

Модуль «Учет дорожно-транспортных происшествий (ДТП)» - подсистема, обеспечивающая сбор, хранение, фильтрацию информации, касающейся произошедших ДТП. Информация может быть отфильтрована практически по любому хранящемуся параметру: дате, времени, месту, характеру происшествия, причине его возникновения, участникам ДТП, количеству погибших или раненых. Особым видом фильтрации является формирование списка так называемых «очагов аварийности» - мест концентрации ДТП. Для этих целей разработан алгоритм, позволяющий выявлять очаги по точному или ориентировочному описанию места ДТП в пределах некоторой улицы, административного района, нескольких районов или всего города.

Модуль «Учет транспорта, находящегося в розыске» предназначен для хранения информации, касающейся транспортных средств (ТС), находящихся в угоне или совершивших какое-либо противоправное действие. Отличительной чертой данных, обрабатываемых в данном модуле, является их

возможная неполнота или даже недостоверность. Например, данные о транспортном средстве, скрывшемся с места ДТП, могут содержать неполную информацию о государственном номере ТС или искаженную информацию о марке ТС, или несколько вариантов предположительного цвета его кузова.

Модуль «Учет недостатков улично-дорожной сети» обеспечивает хранение и обработку информации о нарушениях функционирования улично-дорожной сети с позиций обслуживания инженерных сетей (вода, канализация, электричество, связь и др.) и других специальных служб (например, «Управление зеленого хозяйства»). Хранящаяся в модуле дополнительная информация о допустимых (согласно ГОСТ) временных интервалах, назначенных на устранение выявленного недостатка, и ответственных исполнителях позволяет по истечении допустимого времени и неустранении недостатка сигнализировать об этом пользователю (администратору АИС).

Модуль «Учет неисправностей технических средств организации дорожного движения» по своим функциям аналогичен модулю учета недостатков улично-дорожной сети.

Организационно-вычислительный модуль предназначен для получения новых данных из уже имеющихся в перечисленных выше модулях.

Необходимые для статистического накопления и анализа оперативные данные, собранные, обработанные и отфильтрованные соответствующим модулем, собираются в вычислительном модуле для дальнейшей работы с ними. Объем перерабатываемой информации пропорционален количеству формируемых показателей. В общей сложности в АИС за один сеанс (час, сутки и т.п.) обрабатывается более 100 показателей. Помимо накопления статистической информации в АИС производится вычисление некоторых интегральных характеристик, проводится сравнительный анализ значения какого-либо показателя со значением аналогичного показателя за прошлый период (год, месяц, сутки), отмечается динамика роста и т.п.

1.1. Динамическая структура показателей

Отличительной особенностью разработанной АИС является вполне объяснимое ограничение, накладываемое областью применения (большой объем оперативно поступающих данных и постоянно изменяющаяся структура отслеживаемых параметров) на реализацию хранения данных. Динамика изменения структуры проявляется вследствие изменений характеристик реального мира, влекущих за собой введение новых исследуемых параметров, их сбор и обработку в АИС. Невозможность организации жесткой структуры хранения данных нашла свое решение в реализации гибко настраиваемой древовидной системы показателей (параметров) с возможностью размещения узлов на любом уровне и созданием произвольной глубины вложенности.

Листья-показатели могут быть двух типов: *статистические* (невывчисляемые) и *вывчисляемые*. Статистические показатели – это потоки данных, поступающие из модулей. Например, показатель «*Кража транспортного средства*: за сутки» поступает из модуля «Учет транспорта, находящегося в розыске». Вывчисляемые показатели – это потоки данных, полученные путем каких-либо вычислений или анализа аналогичных показателей за некоторый временной период. К вычислительным параметрам относится, например, показатель «*Кража транспортного средства*: за прошедший месяц», получаемый суммированием аналогичных данных за заданный период времени.

Для каждого показателя (вывчисляемого и невывчисляемого) хранится его текущее состояние по одному значению за одни сутки. Для статистических – это вводимое значение, а для вычисляемых – это полученное вследствие применения соответствующей формулы. Формула, по которой производится вычисление показателя, задается в дереве настройки показателей. Она может содержать операнды, над которыми производится операция, знаки операции +, -, *, / и скобки - «(» и «)», обеспечивающие приоритет вычислений. Этого вполне достаточно для обработ-

ки, оценки сравнительных характеристик показателей, выявления динамики роста показателей, так как позволяет вычислять такие показатели, как «Количество дорожно-транспортных происшествий (ДТП): по нарастающей с начала года за текущий месяц» или «Пострадало в ДТП: за прошлый год» = «Погибло в ДТП: за прошлый год» + «Ранено в ДТП: за прошлый год».

Вычисляемые значения образуют граф зависимостей (рис. 1), и обновление одного значения влечет за собой пересчет всех зависящих от него значений. Иногда проще отказаться от хранения вычисляемых значений, а получать их непосредственно перед отображением на экране. Такой подход оправдывает себя тогда, когда вычисляемые значения являются агрегатными функциями и могут быть получены средствами языка запросов SQL. В более сложных случаях часто оказывается выгоднее хранить вычисляемые значения в базе данных вместе с невычисляемыми и выполнять их пересчет при необходимости.

Для описываемого принципа организации данных важным является то, что все объекты (значения) сгруппированы в классы (по типам), и для каждого типа можно задать свою формулу вычисления значения. При выполнении такого условия можно выделить сущность «тип значения», характеризующую, кроме всего прочего, уникальным именем и формулой для вычисления. Формула может быть строкой, содержащей алгебраическое выражение. Например, формула вычисления количества нераскрытых преступлений на конец текущего дня ($current(quantity)$) будет записана в виде строки:

« $current(quantity_i) = previous(quantity_i) + current(credit_i) - current(debit_i)$ »,

то есть количество всех нераскрытых преступлений на конец предыдущего дня $previous(quantity)$, плюс количество нераскрытых преступлений $current(credit)$, минус количество раскрытых преступлений $current(debit)$ за текущий день. $Current$ – функция вычисления значения соответствующего параметра за текущий день (период), $previous$ – функция вычисления значения этого же параметра за предыдущий день (период). Для вычисления показателей по хранящейся формуле-строке в АИС реализован синтаксический анализатор, проверяющий структуру вводимой формулы на предмет правильности ее формирования пользователем.

Организация базы данных для хранения такой структуры при последовательном (день за днем) внесении данных в систему проблем с пересчетом не вызывает. Однако при внесении изменений «задним числом» приходится пересчитывать все вычисляемые значения, зависящие от измененных (рис. 1). При этом возникают две задачи: определить значения, требующие пересчета, и определить порядок пересчета значений.

Введем отношение зависимости между значениями. Пусть запись $A.зависит_от(B)$ означает, что значение A зависит (не обязательно напрямую) от значения B . Пусть запись $X.вычисляется_через(Y)$ означает, что в формулу для вычисления X непосредственно входит Y . Тогда можно записать условие зависимости значения A от значения B :

$A.зависит_от(B)$, когда

либо $A.вычисляется_через(B)$,

либо $A.вычисляется_через(C)$ и $C.зависит_от(B)$.

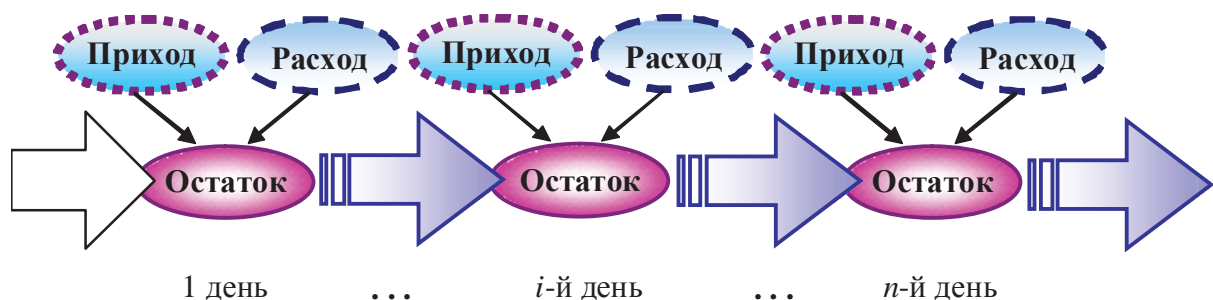


Рис. 1. Схема вычислений значений

Используя это правило, можно применить разновидность обхода графа зависимостей в ширину для определения множества значений, требующих пересчета.

Для определения порядка пересчета значений примем интуитивно понятное утверждение: *никакое значение за текущий день не зависит от значений за последующие дни*. Следовательно, чтобы не нарушить порядок вычисления значений, нужно отсортировать их по датам и затем вести вычисления в порядке возрастания даты. Чтобы определить последовательность вычислений значений внутри какого-либо дня, данные сортируются по признаку зависимости друг от друга в пределах этого дня.

1.2. Структура базы данных

Структура базы данных (некоторой ее части), обеспечивающей хранение оперативной информации, приведена на рисунке 2 и демонстрируется в концепции модели «сущность-связь» – ER-модели (от англ. *ER – Entity-Relationship model*), предложенной Ченом [5].

АИС оперативного учета предполагает регулярное (ежедневное) пополнение данных. Для каждого типа значения можно указать день, когда было внесено первое значение этого параметра; значения параметра до этого дня системе были неизвестны. Аналогич-

но можно указать и верхнюю границу (последний день), на который системе известно значение конкретного типа. Каждый день работы системы отодвигает эту границу, «открывая» новые дни.

Основной частью базы данных, отвечающей за реализацию гибкой древовидной структуры, является кортеж из пяти таблиц: *Param* («Имя параметра»), *ParamKind* («Тип параметра»), *ParamValue* («Значение параметра»), *Division* («Имя подразделения-клиента, приславшего данный параметр»), *ValueKind* («Значение типа параметра»). Рассмотрим их подробнее. *Param* отвечает за содержимое дерева (за исключением конечных листьев и значений, хранящихся в этих листьях). Сформировать дерево позволяет наличие полей *IParent* («Идентификатор родительского узла»), ссылающегося на узел-предок, и *VcName* («Наименование»), хранящего название узла. Дополнительный компонент *ValueKind* хранит названия возможных листьев дерева. Их комбинация *Param+ValueKind* образует гибкое, легко расширяемое дерево, которое реализуется в таблице конечных показателей *ParamKind*.

Они формируются на основании пары *Param.Id_Param:ValueKind.Id_Kind*. Вычисляемость параметра определяется наличием формулы в поле *VcFormula* («Формула»). Таб-

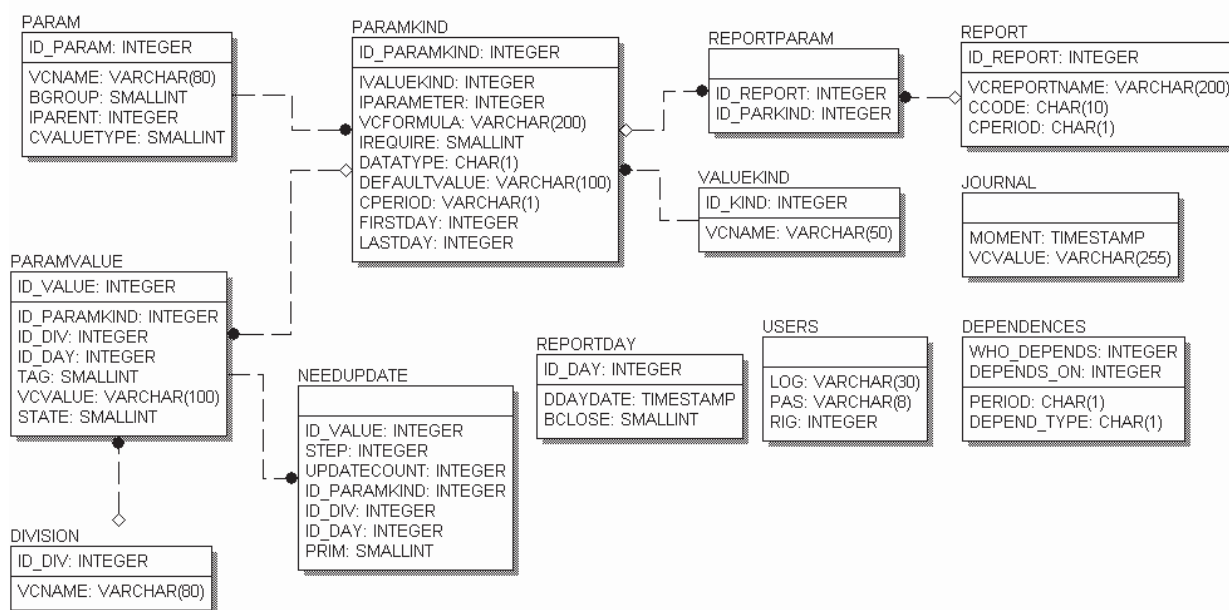


Рис. 2. ER-модель данных

лица *ParamValue* хранит данные для каждого конечного показателя на каждый день (вычисляемые или статистические). АИС поддерживает клиент-серверную технологию обработки данных, т.е. на множество хранимых значений оказывает влияние, в каком подразделении (*Division*) зафиксировано данное значение показателя. Отсюда вытекает необходимость в идентификации района города (таблица «Подразделение» *Division: ParamValue.Id_Div=Division.Id_Div*). Аналогично определяются значения подразделения, уникальное значение для какого-либо показателя за один день обеспечивает поле *Id_Day=ReportDay.Id_Day* в таблице *ReportDay* («Отчет за день»).

1.3. Формирование отчетов

Функция формирования отчетов обеспечивает такую же гибкость работы, как и работа с параметрами, т. к. основана на том же принципе организации данных. Пользователь имеет возможность формировать любые шаблоны для отчетов, составлять и сохранять запросы [6]. В АИС эта возможность реализована на двух таблицах. Первая таблица *Report* («Отчет») – собственно отчет, работает с полями: название отчета, код и периодичность формирования. Вторая – набор полей, по которым идет выборка данных для формирования данного отчета (*ReportParam*). Таким образом, пользователю для формирования нового отчета достаточно создать шаблон (рис. 3) и период, за который формируется отчет. Отчет создается в форме *Excel*-файла с определенными (заданными пользователем) заголовками, содержимое которых определяется из АИС (для этого в системе предусмотрены специальные инструменты, рис. 4).

2. Система учета и анализа дорожно-транспортных происшествий

В потоке оперативных данных среди прочей информации в АИС поступает информация о дорожно-транспортных происшествиях [7, 8]. Для описания ДТП используются его основные характеристики: место и время возникновения – и вспомогательные: вид, причина, степень тяжести, ущерб, количество пострадавших и др. Помимо вычисления «обыкновенных» аддитивных показателей для показателей ДТП «вычисляются» так называемые *очаги аварийности*. Очаг аварийности – это участок улично-дорожной сети, характеризующийся высоким уровнем аварийности: более двух ДТП в год.

Рассмотрим подробнее внутреннюю структуру модуля «Учет ДТП». Модуль предназначен для хранения информации о ДТП и местах их концентрации – «очагах аварийности». Очаг аварийности хранится в базе данных как совокупность ДТП и характеризуется моментами возникновения и исчезновения (для ликвидированных очагов). Центральной утилитой в модуле «Учет ДТП» является утилита редактирования списка ДТП (рис. 5). Данные представляются в табличном виде. Область видимости информации можно ограничить населенным пунктом, возможна фильтрация и сортировка по запросу пользователя: по месту происшествия, по дате, причине ДТП и др.

Анализ ДТП с точки зрения поиска очагов аварийности сводится к группировке ДТП по дате и месту происшествия. Для чего необходимо переработать несколько тысяч записей о ДТП. При ручном поиске такая работа достаточно трудоемка. Автоматизирован-

ДАТА	#ДАТА#		
ПОДРАЗДЕЛЕНИЕ	<i>Норма выставления постов ДПС</i>	<i>Фактически выставлено постов ДПС</i>	<i>Всего задействовано постов ДПС за сутки</i>
#ПОДР#	>9<	>10<	>11<

Рис. 3. Структура шаблона отчета для показателя «Посты дорожно-патрульной службы (ДПС)»

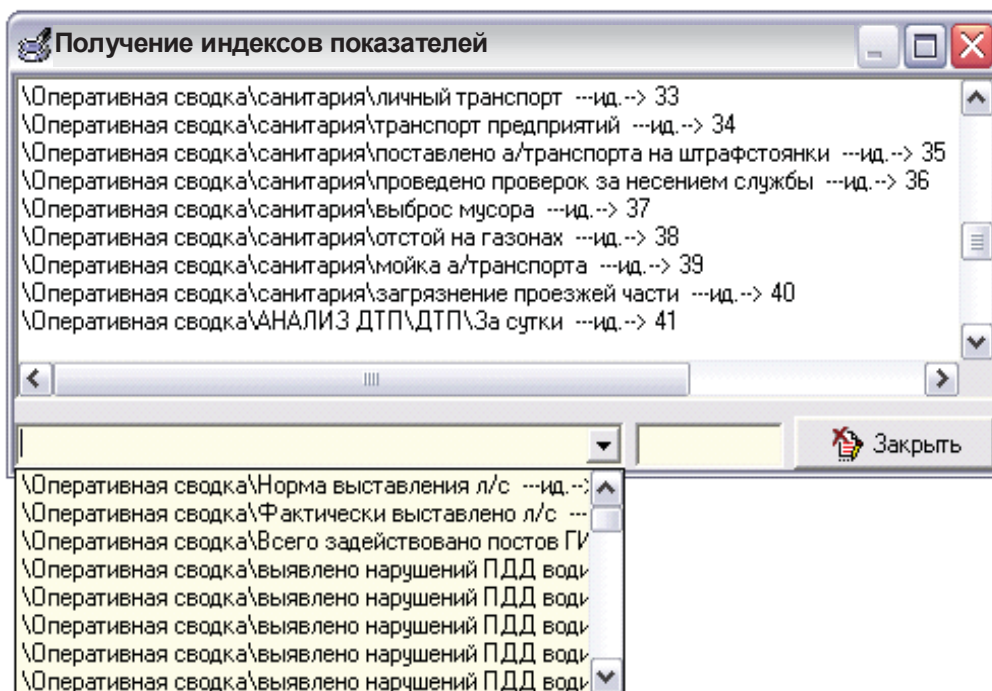


Рис. 4. Экранная копия инструмента формирования отчетов

ное выявление очагов аварийности значительно упрощает поиск, но имеет свои трудности. Каждое ДТП фиксируется по территориальной принадлежности. ДТП, произошедшее на границе районов, записывается в обоих районах. При анализе эти ДТП могут не образовывать очагов аварийности, хотя в действительности таковыми являются, т. к. топологический анализ ДТП подразумевает отсутствие границ. Еще одной проблемой является «словесное» (неформализованное) описание места ДТП. Например, перекресток, образованный пересечением улицы «А» и улицы «В», семантически однозначен перекрестку улиц «В» и «А» с точки зрения топологического

анализа, но не является одной и той же точкой города «с точки зрения компьютера». По словесному описанию места происшествия бывает довольно трудно оценить близость двух ДТП. Например, ДТП, произошедшие на разных сторонах дороги, могут оказаться «привязанными» к разным домам или даже оказаться в разных районах города.

Еще одним видом «семантического разногласия» является ситуация с распознаванием перекрестка, образованного слиянием нескольких улиц. Место возникновения ДТП можно записать, используя любое сочетание названий улиц, формирующих перекресток. Так, например, не имея дополнительных дан-

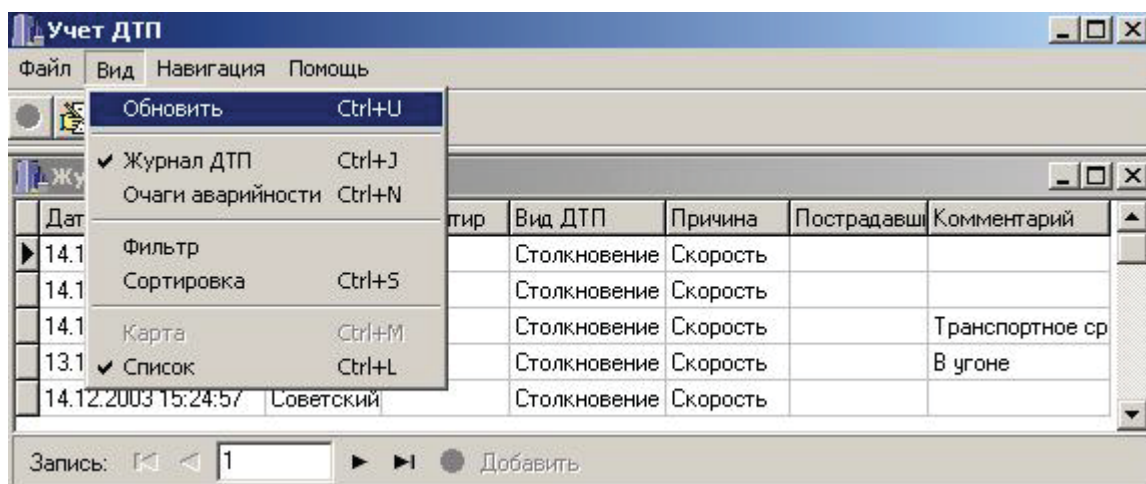


Рис. 5. Экранная копия модуля «Учет ДТП»

ных, невозможно определить, что 155-й километр трассы М-5 и 29-й километр трассы Самара-Курумоч – это одно и то же место: точка слияния двух дорог. Эти и ряд других проблем решены в АИС и не вызывают неоднозначности толкования.

Топологический анализ данных о ДТП проводится в АИС в двух видах: семантический анализ данных, хранящихся в базе данных, и географический анализ данных, визуализированных на электронной карте. В геоинформационной системе (ГИС) данные о ДТП и об очагах аварийности можно формировать по слоям (каждый слой содержит информацию за определенный год). Очаг аварийности отображается в ГИС как совокупность ДТП [9].

Для поиска мест концентрации ДТП на электронной карте используется несколько алгоритмов. Поскольку указанные пользователем координаты на карте (или координаты от GPS) считаются наиболее достоверной информацией о месте возникновения ДТП, то предпочтение отдается их группировке по координатам X:Y. Предварительно осуществляется триангуляция: разбиение плоскости карты на треугольники, в вершинах которых находятся ДТП. Затем для каждого ДТП анализируются соседние с ним ДТП на предмет возможной принадлежности к очагу аварийности (это необходимо для отсекаания ДТП, произошедших на параллельных улицах или во дворах домов).

Однако точные координаты ДТП известны не всегда, т. к. иногда место происшествия задается «текущим километром дороги» (ДТП вне города) или ориентиром (например, номером световой опоры). Для каждого типа ориентира существует оценка расстояния между ДТП, которая используется при группировке ДТП в очаги аварийности. Так, например, для ориентиров типа «№ дома» или «№ световой опоры» дополнительно можно указать расстояние до ДТП. При использовании одного ориентира расстояния до него недостаточно, необходимо еще знать направление. Однако, имея несколько ориентиров и расстояний до них, можно определить место ДТП. Если не знать точные координаты самих ориентиров, то нельзя определить коор-

динаты X:Y ДТП, но можно оценить близость двух ДТП, привязанных к этим ориентирам.

Поиск очагов осуществляется «с накоплением», т. е. при каждом поиске корректируется имеющаяся картина очагов аварийности. Поскольку на корректировку влияют лишь вновь добавленные или измененные ДТП, то можно вручную изменять привязку ДТП к очагам – при пересчете внесенные вручную изменения не теряются.

Инструмент генерации отчетов для модуля «Учет ДТП» предназначен для получения статистического анализа состояния аварийности и вывода результатов на печать или в файл-документ Microsoft Word либо Excel. В АИС предусмотрено получение сравнительных характеристик аварийности и тяжести ДТП в зависимости от времени суток, дня недели, времени года на основе указанной пользователем выборки. Кроме того, распечатку списка ДТП или очагов аварийности можно делать непосредственно из редактора списка ДТП.

3. Клиент-серверная организация АИС

Технически АИС реализована как клиент-серверная система, серверная часть которой (находясь резидентной в памяти) отвечает за пересчет и занесение в базу данных значений вычисляемых показателей за указанный период. Это увеличивает скорость работы клиентских приложений, устраняя необходимость в дополнительных запросах к серверу. Клиентское приложение работает уже с конечным набором данных, находящимся в базе данных.

Для облегчения передачи потока данных из одного подразделения в другой предусмотрена возможность экспорта/импорта и резервного копирования базы данных. Инструменты экспорта и импорта представлены отдельными исполняемыми модулями. Выделение их из основного модуля обусловлено прежде всего тем, что существуют разные способы экспорта и импорта: ручной и автоматический. При ручном способе данные экспортируются в файл, который переносится на дискете или передается по сети. При автоматическом способе клиент подключается к удаленному серверу и передает ему данные, которые импортируются в удаленную базу без

участия пользователя с серверной стороны. Автоматический режим удобен при передаче данных из удаленных (районных) подразделений в центральный (городской) отдел по DialUp-сети. Клиенту возвращается результат импорта. Поскольку объем экспортируемых данных может оказаться довольно большим, то перед передачей данные сжимаются с помощью алгоритма VZIP2.

Существует несколько вариантов экспорта данных, определяющих, какие записи подлежат экспорту: данные за указанный период (фильтрация по дате ДТП или по дате изменения базы данных), ни разу не экспортированные данные, повтор предыдущего экспорта. Аналогично существует и несколько вариантов импорта, определяющих поведение модуля импорта при конфликте версий. Например, можно оставлять более новую запись или перезаписывать все записи принудительно.

Инструмент поиска очагов аварийности выполняется на одном компьютере вместе с сервером баз данных. Он не имеет пользовательского интерфейса и представляет собой WindowsNT-сервис. Поиск очагов аварийности происходит либо по запросу пользователя, либо с заданной периодичностью.

АИС учета оперативных данных по обеспечению безопасности дорожного движения реализована с использованием Borland Delphi 5.0 и Borland C++ Builder 5.0. Данные технические средства выбраны в соответствии с требованиями скорости доступа к перечисленным базам данных, так как в них уже реализованы оптимизированные компоненты для доступа к базам данных InterBase. Для хранения используется база данных Interbase 6.5, однако все методы доступа к данным локализованы в нескольких модулях, и переход от одной СУБД к другой не вызовет осложнений.

Заключение

В настоящий момент АИС используется в отделе ГИБДД УВД г. Самары и позволяет автоматизировать процесс получения оперативной информации руководством городского отдела для принятия своевременных и обоснованных оперативной обстановкой в городе решений по обеспечению безопасности.

Список литературы

1. Автоматизация проектирования вычислительных систем. Языки, моделирование и базы данных // Под ред. М. Брейера: Пер. с англ. - М.: Мир, 1979.
2. Буч Г. Объектно-ориентированный анализ и проектирование с примерами приложений на С++ / Пер. с англ. - СПб.: Невский диалект, 1999.
3. Михеева Т. И., Денисенко В. С., Чугунов И. А. Автоматизированная система обработки и хранения оперативной информации. / Математика. Компьютер. Образование // Тезисы докладов XI международной конф. - М.-Ижевск: НИЦ «Регулярная и хаотическая динамика», - 2004. - С. 53.
4. Хренов А. В., Ярцев В. С., Рудаков И. А. Учет и анализ оперативных данных по обеспечению безопасности дорожного движения // Сборник материалов Всероссийской научно-практической конференции «Безопасность – многоуровневый аспект: превентивные меры и методы». - Пенза, 2003.
5. Чен П. П. Модель «сущность-связь» – шаг к единому представлению данных // СУБД. - №3, 1995.
6. Михеева Т. И., Рудаков И. А. Автоматизация мониторинга транспортной и дорожной инфраструктуры // Труды 6 международной научно-практической конференции «Организация и безопасность дорожного движения в крупных городах». - С-Пб.: С-ПБАДИ. - 2004. - С. 93-96.
7. Методика оценки и расчета нормативов социально-экономического ущерба от дорожно-транспортных происшествий / Под ред. И. А. Венгерова // Р-031121990505-00. - М.: Трансконсалтинг, 2001.
8. Михеева Т. И., Чугунов И. А. Автоматизация учета и анализ дорожно-транспортных происшествий // Труды конференции «Организация и безопасность дорожного движения в крупных городах». - С-Пб.: С-ПБАДИ. - 2004. - С. 267-271.
9. Михеева Т. И. Геоинформационная система наблюдения за улично-дорожной сетью города / Труды 5 международной междисциплинарной научно-практической конф. «Современные проблемы в науке и образовании». - Алушта-Харьков: Харьковский нац. ун-т им. В. Н. Каразина. - 2004. - С. 202.

COMPUTER – AIDED CONTROL OF INSTANT DATE

© 2006 T. I. Mikheeva

Samara State Aerospace University

The paper analysis a computerized information system for collecting, processing and analyzing instant data developed within the framework of a transport system. A data model is proposed, instruments for designed operative information database are described, algorithms of finding object concentration sports on a map are discussed.