

ПРОГРАММНЫЙ КОМПЛЕКС ДЛЯ РАСПОЗНАВАНИЯ СПОСОБА ТРАВМИРУЮЩЕГО ВОЗДЕЙСТВИЯ В РЕЗУЛЬТАТЕ ДТП

Н. В. Огаркова, Е. Ю. Рязских

Воронежский государственный университет

Поступила в редакцию 10.10.2009

Аннотация. В статье рассмотрен программный комплекс, предназначенный для распознавания способа травмирующего воздействия в результате дорожно-транспортного происшествия. Приведена структура программного комплекса и дано описание программ, входящих в его состав.

Ключевые слова: программный комплекс, дорожно-транспортное происшествие.

Annotation. The program complex intended for recognition of a way of injuring influence as a result of road and transport incident is considered in article. The structure of a program complex is presented together with the description of the programs entering into its structure is given.

Keywords: program complex, the road and transport incident.

ВВЕДЕНИЕ

Существует категория задач, в которых необходимо определить исход какого-либо наблюдения, используя набор экспериментальных данных с известными результатами. Эти исходные данные позволяют построить модель зависимости между исследуемым результатом и выбранными признаками, определить какие из признаков оказывают влияние на исход, каково это влияние, значимо оно или нет. Такие задачи встречаются в медицине для определения диагноза больного и оправданности операции в случае тяжелых заболеваний.

Одной из актуальных задач, решаемых в судебной медицине, является определение вида травмирующего воздействия в результате дорожно-транспортного происшествия. Такими воздействиями могут быть удар, когда автомобиль сбивает пострадавшего, либо давление в случае переезда пострадавшего транспортным средством. Возможны случаи повторной травмы, когда разрушающие воздействия обусловлены вторичным переездом пострадавшего. В судебно-медицинской практике способ травмирующего воздействия может быть определен по признакам и характеру разрушения длинных трубчатых костей [1, 2].

Однако до настоящего времени не определено, какие именно признаки разрушения костей и каким именно образом наиболее це-

лесообразно использовать для определения вида травмирующего воздействия. Кроме того, в судебной медицине отсутствует единая методика последовательности анализа повреждений при автомобильной травме и не решена задача автоматизации и объективизации такого исследования [2, 3].

Таким образом, возникла необходимость в реализации комплекса программ, которые позволят определять вид травмирующего воздействия по характеристикам повреждений длинных трубчатых костей.

1. ФУНКЦИОНАЛЬНОЕ НАЗНАЧЕНИЕ СИСТЕМЫ

Программный комплекс включает трех пользователей, четыре программы и базу данных «Кости». Его структура приведена на рис. 1.

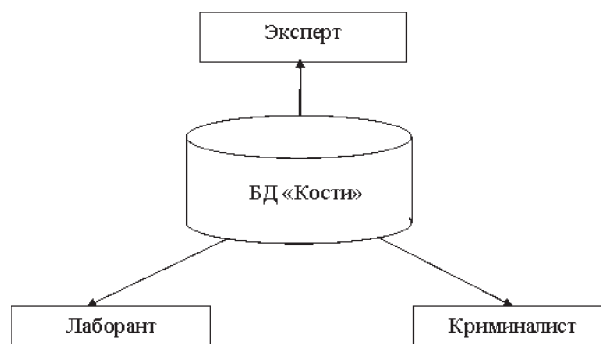


Рис. 1. Взаимодействие пользователей при определении способа травмирующего воздействия

Для того чтобы выделить основные сущности реализуемой системы необходимо построить диаграмму вариантов использования (рис. 2).

В представленной диаграмме фигурируют три сущности:

- Эксперт
 - дает описание модели;
 - формирует файл, содержащий совокупность (систему) (уравнений).
- Лаборант — заполняет файл с данными БД «Кости», то есть на основании цифровых фотографий заполняет значения атрибутов модели.
- Криминалист — определяет способ воздействия (вид воздействия) и степень доверия результатам.

Взаимодействие пользователей при использовании программного комплекса представлено на диаграмме последовательностей (рис. 3).

На диаграмме последовательностей можно выделить четыре основных вида функций системы, и, как следствие, четыре программы:

1. Эксперт: определяет структуру модели (файл «Модель.xml»).
2. Лаборант: на основании цифровых фотографий и на основании готовой структуры мо-

дели заполняет файл с данными. Результат работы программы — файл «Данные.txt» или «Данные.xls».

3. Эксперт: используя файлы «Модель.xml» и «Данные.*» (*.txt, *.xls), формирует комплекс статистических моделей (уравнений) с возможностью сохранения в файл «Комплекс.xml» или «Комплекс.xls».

4. Криминалист: получает конечный результат, т.е. после обработки файлов «Модель.xml» и «Комплекс.xls», а также данных некоторой кости, которые вводятся с клавиатуры или читаются из файла «Данные.*» (*.xls или *.txt), определяет способ (вид) воздействия и степень доверия результатам.

Как следствие, структуру программного комплекса можно представить в виде схемы, изображенной на рис. 4.

2. ПРОГРАММА «ЭКСПЕРТ: ОПРЕДЕЛЕНИЕ МОДЕЛИ»

Эксперт может задать одну или несколько моделей. При описании структуры модели должен быть определен полный перечень признаков, используемых в дальнейшем. В основном эта информация требуется только

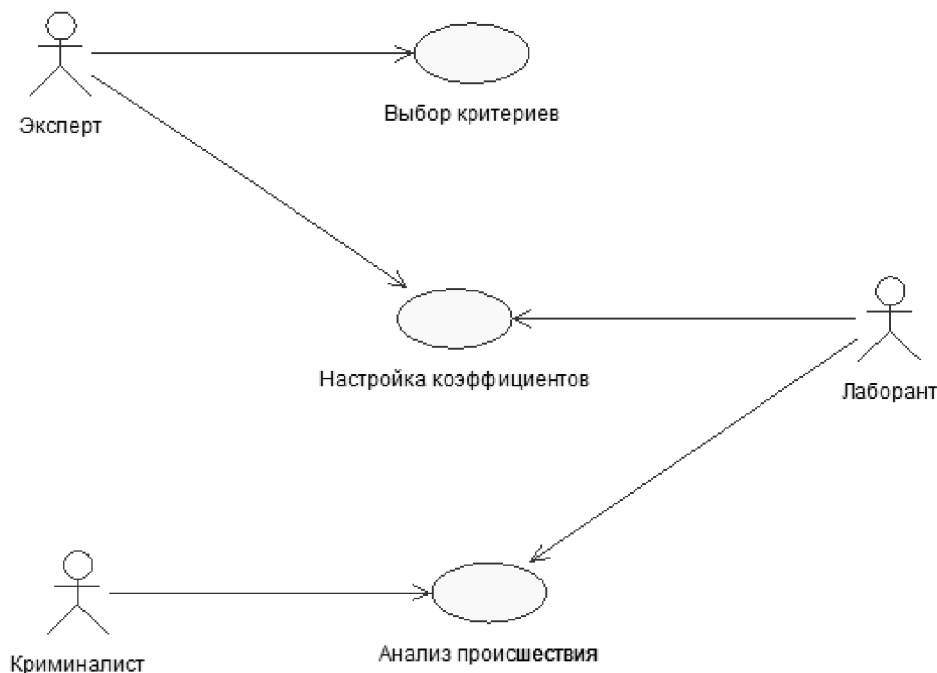


Рис. 2. Диаграмма вариантов использования

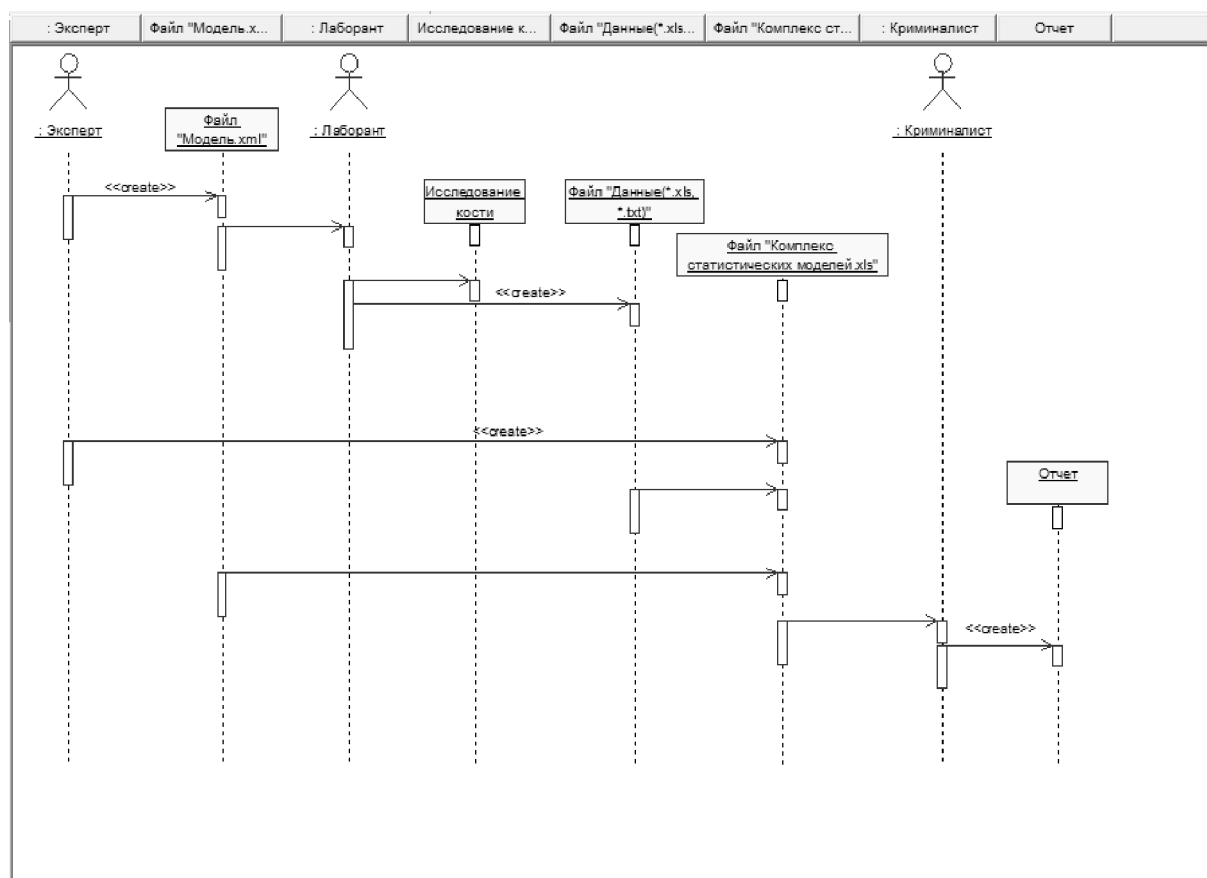


Рис. 3. Диаграмма последовательностей

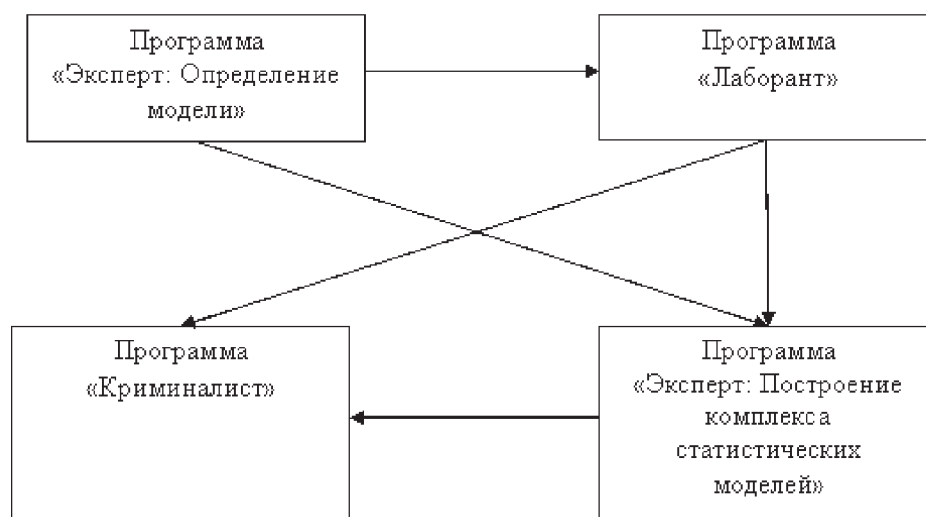


Рис. 4. Структура программного комплекса

для того, чтобы программы могли корректно отображать значения результатов или содержимое отчетов.

Для каждого признака Эксперт должен определить:

— идентификатор (номер по порядку);

— полное название;

— сокращенное название;

— тип признака;

— в зависимости от типа признака список допустимых значений (только для некоторых типов);

— единицы измерения.

Для отображения результатов будут использоваться Сокращенные названия.

Все признаки можно разделить на следующие типы (категории):

- Результирующие — признаки, значения которых могут быть либо заданы пользователем (например, для обучающей выборки), либо определены в результате работы ПК. К таким признакам могут относиться, например, «ВидВ» (Вид воздействия), «ППТ» (Признак повторной травматизации), «Ротация». Поскольку для решения задачи используется логистическая регрессия, т.е. результирующий признак может принимать только два значения, то допустимыми в ПК являются следующие значения: 0 — первое из допустимых, 1 — второе из допустимых, -1 — значение не определено (его надо определить в результате работы ПК).

- Фильтрационные (описательные или характеризующие) — признаки, которые могут быть использованы для фильтрации набора «Данные.*» по горизонтали. Фактически, каждый из этих признаков относится к перечисленному типу данных, т.е. значения признаков перечисляет пользователь (например, «Вид кости» — {0 — «Бедро», 1 — «Голень»}; «Пол» — {0 — «Мужской», 1 — «Женский»}). Такой признак, кроме значений 0, 1, 2... может принимать еще одно значение: -1 — «Значение не определено», то есть признак может принимать любое из перечислимых значений.

- Факторные — признаки, которые могут быть использованы (или используются) для получения результата. Для таких признаков могут быть выделены следующие категории:

- Измеряемые (их значения могут быть определены при анализе характера разрушений, т.е. должны быть введены пользователем). Они в свою очередь делятся на:

- количественные (например, «Количество трещин первого порядка» или «Количество ХУ-образных трещин»);

- линейные (например, «Длина зоны долома»);

- квадратичные («Площадь сколов»);

- угловые (например, «Угол отхождения микротрещины от расклинивания»).

- Вычисляемые (значения вычисляются с помощью выражений и элементарных арифметических операций и функций).

- Групповые — выделяются группы однотипных элементов, а на их основании могут определяться такие характеристики как:

- количество элементов в группе,

- сумма значений элементов группы;

- среднее значение элементов группы;

- СКО значений в группе.

Таким образом, эксперт дает описание модели, т.е. перечисляет все признаки, которые будут учитываться при построении модели. Модель будет сохранена в файле формата xml.

3. ПРОГРАММА «ЛАБОРАНТ»

Лаборант на основании цифровых фотографий и на основании готовой модели, созданной Экспертом, заполняет файл с данными БД «Кости», т.е. записывает значения признаков в файл формата xml или txt, полученных опытным путем. Эти значения представляют собой измерения ряда параметров разрушения длинных трубчатых костей, которые могут рассматриваться как признаки, используемые для распознавания вида травмирующего воздействия.

4. ПРОГРАММА «ЭКСПЕРТ: ФОРМИРОВАНИЕ КОМПЛЕКСА СТАТИСТИЧЕСКИХ МОДЕЛЕЙ»

Эксперт используя файлы «Модель.xml» и «Данные.*» (*.txt, *.xls), формирует комплекс статистических моделей.

В случаях, когда зависимая переменная принимает конечное число значений, применяется логистическая регрессия.

Логистическая регрессия — это разновидность множественной регрессии, общее назначение которой состоит в анализе связи между несколькими независимыми переменными (факторными признаками) и бинарной зависимой переменной, принимающей только два значения. Целью логистической регрессии является построение модели прогноза вероятности события в зависимости от независимых переменных x_1, \dots, x_m . Если полученная вероятность близка к нулю, то принимается решение о большей вероятности альтернативного события.

Функция логистического распределения равна

$$F(z) = \frac{e^z}{1 + e^z} = \frac{1}{1 + e^{-z}}, \quad (1)$$
$$-\infty < z < \infty,$$

Вероятность того, что произойдет случайное событие W_0

$$P(W_0) = F(\mathbf{x}\mathbf{a}) = \frac{e^{\mathbf{x}\mathbf{a}}}{1 + e^{\mathbf{x}\mathbf{a}}} = \frac{1}{1 + e^{-\mathbf{x}\mathbf{a}}}, \quad (2)$$

где $\mathbf{x} = (x_0, x_1, x_2, \dots, x_m)$ — вектор значений независимых переменных ($x_0 = 1$), т.е. множество анализируемых признаков. Множество коэффициентов регрессии — вектор

$$\mathbf{a} = (a_0, a_1, a_2, \dots, a_m). \quad (3)$$

Для расчета коэффициентов логистической регрессии будет использоваться метод наименьших квадратов и метод максимального правдоподобия.

Согласно методу максимального правдоподобия в качестве оценки неизвестного параметра \mathbf{a} принимается такое значение $\mathbf{a}^* = \mathbf{a}^*(y_1, \dots, y_n)$, которое максимизирует функцию L .

Функция правдоподобия, выражающая плотность вероятности совместного появления результатов выборки y_1, y_2, \dots, y_n равна

$$L(y_1, y_2, \dots, y_n, \mathbf{a}) = P(y_1, \mathbf{a}) \cdot P(y_2, \mathbf{a}) \cdot \dots \cdot P(y_n, \mathbf{a}), \quad (4)$$

Для оценки коэффициентов логистической регрессии методом наименьших квадратов используется линеаризация в точке 0:

$$F(\mathbf{x}_i \mathbf{a}) \approx F(0) + f(0) \cdot \mathbf{x}_i \mathbf{a}, \quad (5)$$

где $f(\cdot)$ — плотность распределения, производная функции распределения $F(\cdot)$, \mathbf{x}_i — вектор значений признаков для i -го наблюдения, \mathbf{a} — вектор коэффициентов регрессии [4, 5].

При определении вида травмирующего воздействия анализ может проводиться как по всем данным, так и отдельно по каждому виду кости. В программе реализована возможность фильтрации данных при анализе. В фильтре могут использоваться только установочные признаки.

По результатам фильтрации строится комплекс статистических моделей по специальному алгоритму.

5. ПРОГРАММА «КРИМИНАЛИСТ»

Для решения задачи определения способа травмирующего воздействия необходима построенная на основе измерений фрагмента кости модель классификации, которая может быть представлена в виде таблицы, содержащей значения следующих величин:

- Коэффициент детерминации:

$$R^2 = 1 - \frac{Q_{\text{ост}}}{Q_{\text{общ}}} = \frac{Q_{\text{объясн}}}{Q_{\text{общ}}}, \quad (6)$$

где $Q_{\text{ост}}$ — остаточная сумма квадратов (сумма квадратов отклонений фактических значений от расчетных)

$$Q_{\text{ост}} = \sum_{i=1}^n (y_i - \hat{y}_i)^2. \quad (7)$$

$Q_{\text{объясн}}$ — объясненная сумма квадратов (сумма квадратов отклонений расчетных значений от среднего)

$$Q_{\text{объясн}} = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2, \quad (8)$$

а $Q_{\text{общ}}$ — общая сумма квадратов

$$Q_{\text{общ}} = Q_{\text{ост}} + Q_{\text{объясн}}; \quad (9)$$

• значение среднеквадратической ошибки, которое определяется формулой

$$E = \sqrt{\sum_{i=1}^n (y_i - \hat{y}_i)^2}; \quad (10)$$

- количество правильно классифицированных случаев;
- количество неправильно классифицированных случаев;
- коэффициенты соответствующих признаков модели.

СПИСОК ЛИТЕРАТУРЫ

1. Диагностикум механизмов и морфологии переломов при тупой травме скелета. Т. 1. Механизмы и морфология переломов длинных трубчатых костей / В. И. Бахметьев, В. Н. Крюков, В. П. Новоселов и др. — 2-е изд. — Новосибирск: Наука. Сибирская издательская фирма РАН, 2002. — 166 с.
2. Семенников В. С. Диагностика механизмов травмы по характеру и особенностям переломов длинных трубчатых костей // Актуальные вопр. суд. медицины и экспертной практики: Тез. докл. — Ростов-на-Дону, 1985. — С. 73—74.
3. Огаркова Н. В. Программа построения и анализа регрессионной модели для распознавания способа травмирующего воздействия в результате ДТП / Н. В. Огаркова, В. А. Голуб, А. С. Попов // Вестн. Воронеж. Ун-та. Сер. Системный анализ и информационные технологии. — Воронеж, 2008. — № 2. — С. 16—20.
4. Лимаренко Е. Ю. Программа построения и анализа системы математических моделей для распознавания способа травмирующего воздействия в результате ДТП: курсовая работа / Е. Ю. Лимаренко. — Воронеж, 2008. — 41 с.

Огаркова Наталья Владимировна — преподаватель кафедры программного обеспечения и администрирования информационных систем. E-mail: Nataly.Ogarкова@gmail.com

Рязских Екатерина Юрьевна — студентка 5 курса специальности «Математическое обеспечение и администрирование информационных систем» факультета ПММ ВГУ

Ogarкова N.V. — teacher of the dept. of the Software and Information system Administration, Voronezh State University. E-mail: Nataly.Ogarкова@gmail.com

Ryazhskikh E.Yu. — student of the dept. of the Software and Information system Administration, Voronezh State University