

О ЦЕЛЕСООБРАЗНОСТИ ПРИМЕНЕНИЯ СИСТЕМОЛОГИЧЕСКОГО КЛАССИФИКАЦИОННОГО ПОДХОДА ДЛЯ АНАЛИЗА ЧРЕЗВЫЧАЙНЫХ СИТУАЦИЙ

ОГРАДИНА Ю. Н.

Рассматриваются вопросы применения системологического подхода для решения сложных слабоформализуемых задач на примере анализа и прогнозирования чрезвычайных ситуаций. Обосновывается применение системологического классификационного анализа для отражения в концептуальной модели глубинных знаний об объектах проблемной области, учета существенных свойств объектов. Анализируются возможности автоматизации обработки архива описаний чрезвычайных ситуаций для построения модели их анализа и прогнозирования.

В настоящее время мы являемся свидетелями быстрого развития вычислительной техники и информационных технологий: появляются новые области применения, постоянно увеличивается объем информации, сложность решаемых задач, преобладают сложные слабоструктурированные проблемные области (ПО) и плохо формализуемые задачи и, соответственно, развиваются способы обработки информации, позволяющие более эффективно использовать имеющуюся информацию для решения подобных задач.

Одной из наиболее острых проблем современного общества является усиление влияния человека на окружающую среду, увеличение числа чрезвычайных ситуаций (ЧС), в том числе техногенных. Проявляются противоречия между развитием общества и природой.

Масштабы и темпы развития цивилизации влияют на общий баланс природного потенциала. Реакция окружающей среды на вторжение техногенной сферы приобретает все более ощутимые для человечества последствия, особенно, когда происходит резонанс общего влияния природных и техногенных факторов [1].

По данным Международной федерации Красного Креста количество ЧС в 90-х годах XX века в 3 раза больше, чем в 60-х. Ущерб увеличился с 52 до 497 миллиардов долларов США [2].

Анализ природно-техногенной ситуации показывает, что часто действия по предотвращению угроз, уменьшению их уровня или ликвидации последствий неадекватны из-за недооценки угрозы или незнания сценария развития ситуации. Поэтому увеличивается актуальность задач предупреждения проявлений ЧС, предотвращения их возникновения и оперативного реагирования [1]. Эти задачи

связаны с анализом слабоструктурированной ПО и их решение требует привлечения современных методов концептуального моделирования ПО.

Для решения сложных слабоструктурированных задач был разработан новый системологический когнитивный метод концептуального классификационного моделирования произвольных ПО — системологический классификационный анализ, использующий метод концептуального моделирования слабоструктурированных ПО на основе естественной классификации (ЕК) и “учитывающий операциональные критерии ЕК и детерминантный подход к анализу ПО, который дополняет имеющиеся методы концептуального моделирования способами определения существенных свойств объектов и учета глубинных знаний о ПО” [3, с. 145].

Таким образом, применение системологического подхода в процессе разработки систем, основанных на знаниях (СОЗ), позволяет отразить в концептуальной модели и базе знаний СОЗ поверхностные ситуативные знания и, следовательно, повысить качество базы знаний, качество решения задачи.

Результаты применения системологического классификационного метода для проблемной области ЧС приведены в работе [3, с. 148-152]:

1) результат детерминантного анализа ПО — выбор оснований построения ЕК ЧС, соответствующих существенным свойствам:

- по виду вызвавшего ЧС отрицательного влияния (причина ЧС);
- по виду нарушаемого установившегося процесса (следствие ЧС);
- по виду необходимых экстренных мер (динамика процесса развития ЧС);

2) фрагмент построенной ЕК ЧС по выбранным основаниям;

3) сравнительный анализ полученной ЕК ЧС с существующими классификациями ЧС [3, с. 157-158].

На официальных сайтах Министерства Украины по вопросам чрезвычайных ситуаций и вопросам защиты населения от последствий Чернобыльской катастрофы и Министерства РФ по делам гражданской обороны и чрезвычайным ситуациям (www.mns.gov.ua, www.emercom.gov.ru) хранится информация о чрезвычайных ситуациях, произошедших с 1997г. Архивы чрезвычайных ситуаций содержат описание нескольких тысяч ЧС (порядка 8-10 тысяч для РФ и порядка 5-7 тысяч для Украины), произошедших за последние несколько лет на территории РФ и Украины, и мер по их ликвидации. Кроме того, на сайтах приведена информация, относящаяся к работе министерства (например, проекты, пресса), статистика за разные периоды.

Описание каждой ЧС включает время и место ЧС, ее описание, причины возникновения (если уста-

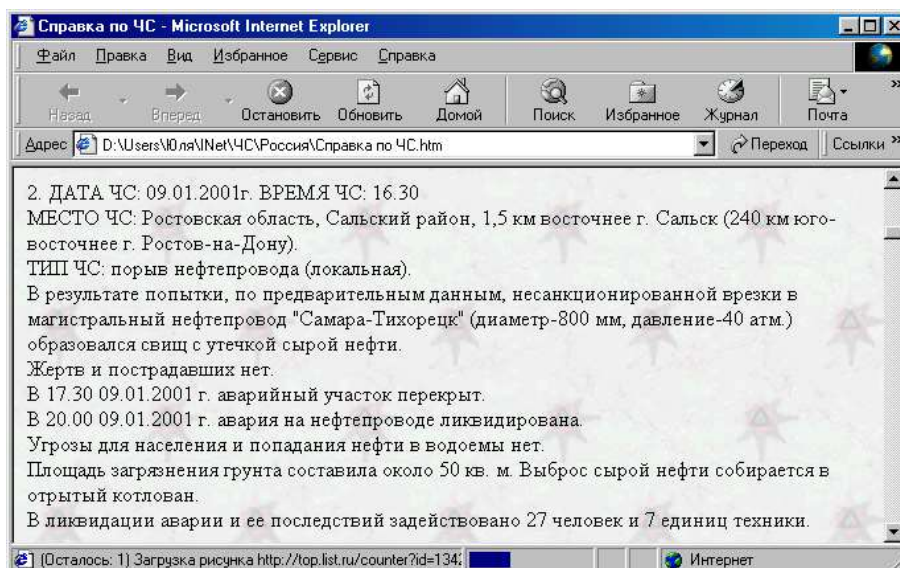


Рис. 1. Пример описания ЧС на WEB-странице МЧС РФ

новлены) и меры, предпринятые для ее ликвидации. Пример информации, представленной на сайте [4], приведен на рис. 1.

Как видно из примера, текстовые описания ЧС полуструктурированы. В каждом описании выделены как отдельные поля дата, время, место и тип ЧС. Затем идет текст описания ЧС. Вид описания зависит от типа ситуации. Для техногенных ЧС приведены причины (если они установлены), последствия и меры, которые были приняты для ликвидации. Описание природных ситуаций также содержит поля даты, времени, места и типа ЧС, а также информацию, относящуюся к произошедшей ЧС. Например, для землетрясений приведены координаты эпицентра, магнитуда, сила землетрясения в эпицентре, сведения о жертвах и разрушениях. Объем описаний различный: от 2-3 предложений и более.

Описание каждой ЧС встречается только один раз, но в более поздней информации могут быть приведены меры по ее ликвидации (если все ее последствия не удалось устранить сразу).

На сайте МЧС Украины [3] описание каждой ЧС представляет собой сплошной текст, в котором есть указание времени и места ЧС, ее причины, следствия и меры, предпринятые для ее ликвидации. Пример приведен на рис. 2.

Следует отметить, что автоматическое выделение отдельных частей описания может производиться на основе ключевых слов или шаблонов, выделенных на этапе предварительного анализа описаний. В каждом описании можно выделить причину произошедшей ЧС, непосредственное ее описание, последствия (материальный ущерб и людские жертвы), меры по ликвидации в соответствии с основаниями деления в построенной концептуальной классификационной модели (ККМ) ЧС.

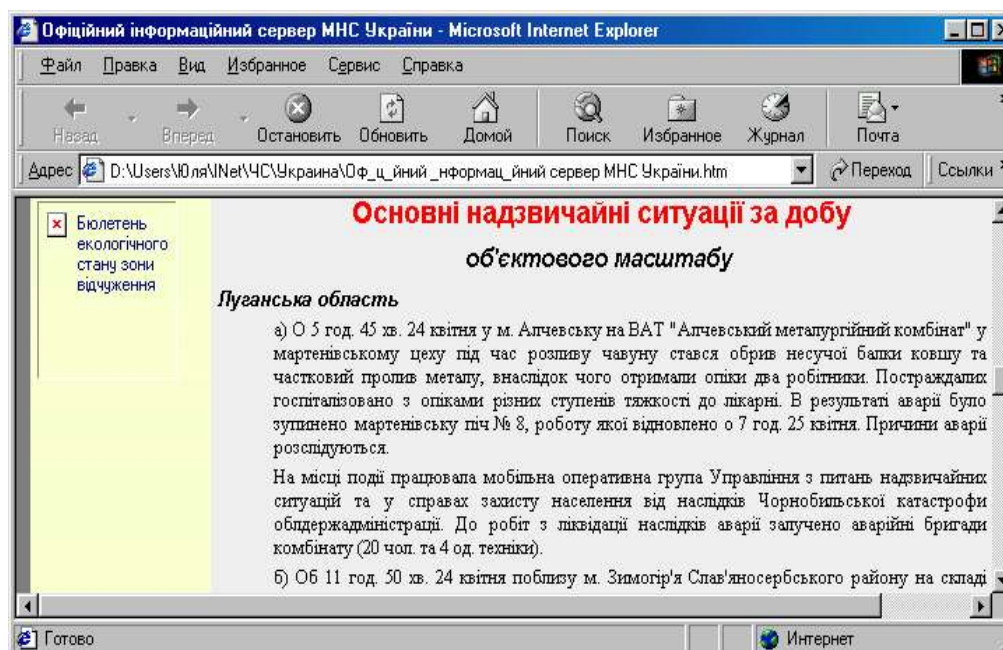


Рис. 2. Пример описания ЧС на WEB-странице МЧС Украины

Сравнительный анализ имеющихся данных о ЧС показывает, что предпочтительнее использовать для построения ККМ данные МЧС РФ, так как они лучше структурированы, содержат больше видов ЧС и больший объем архива. Затем полученные результаты могут быть использованы для структурирования данных МЧС Украины. Другой подход состоит в одновременном использовании для построения ККМ данных обоих Министерств ЧС. При этом на подготовительном этапе необходимо провести анализ архива МЧС Украины и привести его данные к виду, аналогичному данным МЧС РФ. В обоих случаях возникает задача использования данных на двух языках – русском и украинском – и задача перехода с одного языка на другой.

Информация из архива ЧС может быть использована на этапе системологического терминологического анализа ПО для выявления начальной совокупности терминов понятий, описывающих данную ПО (единичных). Это начальное множество терминов может быть использовано для определения полноты (минимального объема) совокупности терминов, полноты построенной модели.

На этапе системологического классификационного анализа информация из архива ЧС может быть использована для обеспечения четвертого критерия оптимальности классификационной модели. “Наиболее оптимальной классификация получится при учете единичных понятий, так как соответствующие системы-явления могут быть изучены лучше, чем системы-классы. Следовательно, оптимальность естественной классификационной схемы ПО с точки зрения актуальности и практичности полученных знаний может быть охарактеризована близостью нижнего уровня классификации к единичным понятиям.” [3, с.142].

В процессе обработки архива ЧС и добавления описаний конкретных ЧС в ККМ в каждом описании выделяется в соответствии с выбранными функциональными свойствами причина, следствие и меры по ее ликвидации [3, с.157-158]. Таким образом, определяется множество единичных терминов, необходимое для построения полной модели ПО, и таким образом обеспечивается полнота.

В настоящее время построение ККМ ПО производится вручную экспертом в данной ПО или инженером по знаниям. Однако для извлечения знаний о ПО и построения модели необходимо обработать большой объем – информацию из архива МЧС, литературу о ПО. Поэтому представляется целесообразным разработать программного инструментария, позволяющего уменьшить трудоемкость и время построения модели, снизить вероятность появления ошибок, автоматизировать рутинные процедуры.

В архиве собраны описания реагирования в различных типах ситуаций за несколько лет, т.е. архив содержит опыт различных людей по реагированию и ликвидации ЧС. Так как “эксперт строит базу знаний опираясь частично на прошлый опыт и примеры из литературы, есть основания полагать,

что некая программа индуктивного вывода могла бы создавать базы знаний аналогичным образом” [6, с. 145]. Для организации более эффективного использования и обработки архива ЧС, содержащего тысячи описаний зарегистрированных ЧС, представляется перспективным применить современные средства и алгоритмы анализа данных большого объема и машинного обучения.

Технология интеллектуального анализа данных (ИАД) активно развивается последнее десятилетие в связи с необходимостью обработки сверхбольших объемов информации. ИАД предоставляет средства для анализа и использования имеющейся информации при принятии решений. Современные системы ИАД используют статистические и математические алгоритмы и моделирование для нахождения закономерностей, скрытых в базе данных, и включают инструментарий и различные подходы к анализу текста и цифровых данных [7,8].

Использование ИАД позволит решить следующие задачи:

- поиск похожих текстовых документов;
- поиск информации в тексте;
- навигация в базе текстов;
- структуризация хранилищ документов;
- нахождение взаимосвязей между ключевыми понятиями текстов.

Это упростит работу с архивом описаний ЧС и позволит использовать имеющуюся информацию для эффективного принятия решений.

В литературе по ИАД выделяют шесть основных методов анализа данных: классификация, регрессия, прогнозирование временных последовательностей (рядов), кластеризация, ассоциация, последовательность. Первые три используются главным образом для предсказания, остальные удобны для описания существующих закономерностей в данных [7].

На выбор методов ИАД накладывает определенные ограничения тип данных, на которых этот метод будет применяться. Поэтому к текстовым данным можно применять ограниченное множество методов ИАД. Большинство разработанных алгоритмов направлены на обработку числовых данных и не могут быть применены к текстовым описаниям. Однако в настоящее время уже существуют модификации алгоритмов ИАД применительно к текстовым данным [8].

Для работы с текстовыми данными разрабатываются системы рассуждений на основе аналогичных случаев (case based reasoning – CBR). В системе хранится полное символьное описание примеров. Для нового примера алгоритм находит в базе данных идентичный случай и приписывает новому примеру такое же решение, как решение и у найденного в БД. Если идентичный случай не найден, то алгоритм ищет случаи, имеющие описание, подобное новому примеру, и пытается полу-

чить решение для нового примера как комбинацию решений найденных аналогичных ситуаций. Если найденные решения несовместимы, то алгоритм возвращается к другим примерам. Для получения общего решения в некоторых системах используются знания о ПО.

Системы рассуждения на основе аналогичных случаев показывают хорошие результаты, однако имеют недостатки. Главный из них заключается в том, что они не создают каких-либо моделей или правил, обобщающих предыдущий опыт. В выборе решения они основываются на всем массиве доступных исторических данных, поэтому невозможно сказать, на основе каких конкретно факторов эти системы строят свои ответы. Другой недостаток - большие затраты памяти [7,8].

Система рассуждений на основе аналогичных случаев может быть реализована для анализа архива чрезвычайных ситуаций. Такая система позволит использовать опыт работы по ликвидации ЧС, учесть трудности, которые могут возникнуть в процессе ликвидации последствий ЧС некоторого типа, применить положительный опыт и учесть отрицательный, чтобы не повторять тех же ошибок. В такой системе необходимо реализовать поиск в архиве ЧС того же типа и формирование общих рекомендаций по ликвидации с учетом положительных и отрицательных примеров. При этом концептуальная модель области может использоваться для выделения существенных свойств, формирования рекомендаций для ЧС, для которых нет примеров, выработки общих рекомендаций по некоторому типу ЧС.

Применение методов ИАД к данным, описывающим ПО, дает новые возможности решения существующих проблем. Полученные в результате модели и закономерности могут быть использованы для решения задач, формулирования новых задач, лучшего понимания ПО.

В разработках по ИАД в настоящее время ведутся активные исследования в области построения быстрых, масштабируемых алгоритмов, способных обрабатывать большие объемы информации. Для поиска закономерностей используются статистические и математические алгоритмы, однако выяв-

ление в БД “значимых причинно-следственных связей требует значительных основополагающих знаний, наложенных на статистические подпрограммы [6, с.145]”.

Поэтому при работе с архивом ЧС необходимо совместное использование методов ИАД и системологического подхода. Методы ИАД позволят выявить на основе описаний фактов ЧС поверхностные знания о ПО. Применение в процессе обработки информации о ЧС концептуальной модели, отражающей знания о существенных свойствах объектов ПО, позволит учесть глубинные знания и повысить качество получаемых результатов.

Таким образом, использование методов ИАД и системологического классификационного подхода позволит повысить качество решения задач анализа, прогнозирования и оперативного реагирования в области ЧС.

Литература: 1. *Виступ міністра Василя Дурдинця на засіданні розширеної Колегії МНС у м. Львові // Надзвичайна ситуація. 2001. № 3. С. 24-27.* 2. *Сапон В. Зарубіжний досвід захисту населення і територій від надзвичайних ситуацій // Надзвичайна ситуація, 2001. № 9. С. 34-36.* 3. *Соловьева Е.А. Естественная классификация: системологические основания. Харьков: ХТУРЕ, 1999. 222с.* 4. *Урядова інформаційно-аналітична система з питань надзвичайних ситуацій. Міністерство України з питань надзвичайних ситуацій та у справах захисту населення від наслідків Чорнобильської катастрофи. Офіційний інформаційний сервер. www.mns.gov.ua.* 5. *Официальный информационный сервер МЧС Российской Федерации. www.emercom.gov.ru* 6. *Построение экспертных систем: Пер. с англ. / Под ред. Ф. Хейес-Рота, Д. Уотермана, Д. Лената. М.: Мир, 1987. 441с* 7. *Буров К. О. Обнаружение знаний в хранилищах данных // Открытые системы. 1999. №5-6. С.67-77.* 8. *Han J., Kamber M. Data Mining: Concepts and Techniques, Morgan Kaufman Publishers, 2000.*

Поступила в редколлегия 25.04.2002

Рецензент: д-р техн. наук, проф. Руденко О.Г.

Оградина Юлия Николаевна, аспирантка кафедры ПО ЭВМ ХНУРЭ. Научные интересы: знаниеориентированные технологии, концептуальное моделирование, приобретение знаний, интеллектуальный анализ данных. Адрес: Украина, 61166, Харьков, пр.Ленина, 14, к. 525, тел.: (0572) 40-95-91. E-mail: nulpz@kture.kharkov.ua