

## ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ ОБРАБОТКИ АНКЕТНЫХ ДАННЫХ – ПРОБЛЕМЫ И ПЕРСПЕКТИВЫ

С.Н. Мартышенко, Н.С. Мартышенко

Владивостокский государственный университет экономики и сервиса

Анкетные данные являются основным источником информации при проведении социологических и маркетинговых исследований. В последние годы постоянно растет количество исследователей использующих в своей работе анкетные опросы. Однако методы сбора и обработки данных чаще всего не соответствуют современным достижениям научного знания в этой области. Такую ситуацию, в первую очередь, можно объяснить отсутствием доступного программного обеспечения пригодного для анализа анкетных данных на современном уровне – **первая проблема**.

Между тем, данные анкетных опросов обладают рядом специфических особенностей, отличающих их от данных полученных другими способами. В работе [Мартышенко С.Н., Мартышенко Н.С. Информационные технологии повышения качества данных при исследовании социально-экономических систем // Надежность и качество. Труды Международного симпозиума — Пенза, 2007. — Т.1. — С. 40-43.] определяется десять таких особенностей и этот список можно было бы еще расширить. Главной особенностью анкетных данных является необходимость обработки данных различной природы. Разработка методов обработки данных различной природы является бурно развивающимся научным направлением [Орлов А.И. Нечисловая статистика. – М.: МЗ-Пресс, 2004. – 513 с., Ю.Н.Толстова Анализ социологических данных. Методология, дескриптивная статистика, изучение связей между номинальными признаками. – М.: Научный мир, 2000. – 352с.]. Сегодня есть ряд методов и алгоритмов, позволяющих решать такие задачи. Многие из них можно найти в книге [Загоруйко Н.Г. Прикладные методы анализа данных и знаний. – М.: – Н-ск: изд-во ин-та математики СО РАН, 1999. – 270с.]. Однако, если эти методы и получили программную реализацию, то как отмечает автор книги найти такие программы практически невозможно. Даже если такие программы были бы доступны, вряд ли их можно было бы использовать для решения практических задач. Такие программы, как правило, пишутся для решения одной отдельной задачи в чисто научных целях. Решение практических задач требует комплексного использования методов обработки. И вряд ли из-за одной задачи найдется много желающих осваивать сложную программу. Возникает **вторая проблема обработки анкетных данных – отсутствие технологий**. Так в работе известного отечественного ученого А.И. Орлова – автора большого количества работ по прикладным вопросам статистики отмечается, что в научной литературе вопросам рассмотрения технологий обработки статистических данных уделяется явно недостаточное внимание [Орлов А.И. Эконометрика: учеб. пособ. для вузов – М.: «Экзамен», 2002. – 576 с.]. Обычно все внимание сосредотачивается на том или

ином элементе технологической цепочки, а переход от одного элемента к другому остается в тени. Тот же А.И. Орлов утверждает, что говорить о полной автоматизации всего процесса анализа статистических данных говорить преждевременно, потому что слишком много нерешенных проблем, вызывающих дискуссии среди статистиков.

Поэтому можно сделать вывод, что сегодня в нашей стране создание вполне автономного программного средства обработки анкетных данных, включающего современные методы обработки многомерных данных не реально. С одной стороны, на такую разработку не найти достаточно средств, с другой стороны, большинство потенциальных пользователей и не готово сразу перейти к использованию сложной наукоемкой технологии в полном объеме.

Решение проблемы разработки технологий обработки анкетных данных мы видим в поэтапном освоении пользователями современных методов анализа данных. Поэтому мы пошли по пути разработки программных приложений к наиболее распространенному среди пользователей программному пакету EXCEL [Мартышенко С.Н. Совершенствование математического и программного обеспечения обработки первичных данных в экономических и социологических исследованиях / С.Н. Мартышенко, Н.С. Мартышенко, Д.А. Кустов // Вестник ТГЭУ. – 2006. – № 2 – С. 91–103.]. Разработанный комплекс программ объединяет единая идеология, основой которой являются понятия проекта и модели данных. Такой подход позволяет пользователям постепенно осваивать новые методы и технологии обработки данных в рамках знакомой среды EXCEL. Комплекс только расширяет возможности EXCEL. Обширный арсенал функций по обработке и представлению данных EXCEL остается доступен пользователю. А сможет ли какой-нибудь автономный пакет представить такие возможности, и нужно ли конкурировать с разработчиками EXCEL. Особое место в разработанном нами комплексе отведено многомерным методам повышения качества данных и многомерной классификации.

С целью популяризации разработанного комплекса программных средств предусмотрена возможность работы с наиболее интересными программами без соблюдения дополнительных условий, накладываемых пакетом, то есть автономная работа программ вне пакета. На самом деле, соблюдение условий облегчает работу со всеми программами комплекса, обеспечивает целостность данных и исключает многие ошибки, которые может допускать пользователь в процессе работы с программами.

Дальнейшие перспективы развития технологии обработки анкетных данных мы видим в популяризации современных методов обработки данных. Необходимо убедить разработчиков в необходимости размещать свои программы для свободного доступа в Интернет. Разработчики должны накапливать на своих сайтах, как свои программы, так и программы других разработчиков или ссылки на них. Стремление закрыть свои программы ставит под большое сомнение декларируемое качество и возможности программ. Целесообразно организовать обмен пакетами реальных данных, чтобы и другие исследователи могли проводить апробацию своих программ на более обширном материале.

Информационные технологии обработки данных в нашей стране будут развиваться гораздо быстрее, если нам удастся наладить свободный обмен информацией не только на уровне публикаций, но и на уровне программ и данных.