

**С. И. Коровкин**

*(УО «ГГУ им. Ф. Скорины», Гомель)*

**МАТЕМАТИЧЕСКАЯ МОДЕЛЬ И МЕТОДЫ АНАЛИЗА  
WEB-СТРАНИЦ В СЕТИ ИНТЕРНЕТ**

В настоящее время все больше людей предпочитают получать и делиться информацией в сети интернет, они оставляют всевозможные

отзывы о товарах, различных продуктах, местах, где побывали. Что бы повысить качество предлагаемых товаров и услуг компаниям важно знать, как реагируют на их деятельность клиенты. Это очень большое количество информации, которое нужно обработать, структурировать и анализировать. В результате оценки данной информации компании могут делать выводы о дальнейших путях развития и способах привлечения новых клиентов.

Web Mining – это использование методов интеллектуального анализа данных для автоматического обнаружения веб-документов и услуг, извлечения информации из веб-ресурсов и выявления общих закономерностей в Интернете.

Целью информационного извлечения является превращение коллекции документов в легко усвояемую и проанализированную информацию. Благодаря динамике и разнообразию веб-содержимого, создание ручного режима систем информационного извлечения не представляется возможным. В связи с этим, большинство систем по извлечению данных сосредотачивают внимание на конкретные веб-сайты. Другие используют обучающие машины или методы интеллектуального анализа данных и способны извлекать веб-документы в автоматическом или полуполупри автоматическом режиме. С этой точки зрения, Web Mining является частью процесса извлечения информации из Интернета.

Сбор данных – непростая задача, так как используемые алгоритмы могут быть очень сложными, и данные не всегда доступны в одном месте. При разработке приложения следует учесть следующее: методология получения данных и взаимодействие с пользователем; проблемы с производительностью; проблемы с разнородными типами данных; хранение данных; анализ данных; визуализация данных.