

Школа веб-разведки. Методы обработки информации

Д.В. Ландэ, В.В. Прищепя

В мире бизнеса не для кого уже не секрет, что конкурентная разведка является важнейшим условием устойчивого развития компаний и рынков. Назовем основные цели и задачи, которые ставятся перед конкурентной разведкой для обеспечения информационной поддержки принятия решений:

- сбор и своевременное предоставление потребителям самой разноплановой информации, имеющей отношение к бизнесу;
- выявление рисков и угроз, которые могут помешать нормальному развитию бизнеса;
- нахождение информации, способствующей получению конкурентных преимуществ;
- прогнозирование влияния конкурентной среды на бизнес.

При этом можно выделить два направления конкурентной разведки – это отслеживание деятельности конкурентов, что чаще всего коррелируется с вопросами безопасности, и анализ состояния рынков. Вторая задача относится к области маркетинговых исследований. Сегодня во многих крупных компаниях идет процесс формирования служб конкурентной разведки из специалистов, ранее занимающихся информационными технологиями и маркетингом. По словам генерального директора компании "Р-Техно" Р. Ромачева, если 10 лет назад конкурентные разведчики в первую очередь проверяли наличие у бизнес-партнеров криминальных связей, то сейчас они, как и на Западе, в большей степени добывают коммерческую информацию. Т.е. можно проследить изменение центра тяжести конкурентной разведки от вопросов безопасности непосредственно, к вопросам маркетинга, что подтверждается многочисленными исследованиями. В частности, по данным компании OnConference конкурентная разведка чаще всего используется для изучения состояния рынка (74% респондентов). Включая такие функции, как поиск, сбор и анализ информации, конкурентная разведка помогает определить основные объекты и субъекты конкурентной среды, а также выявит важные взаимосвязи связи.

В зависимости от средств, которые вкладывает компания, подходы к конкурентной разведке можно условно разделить на три категории: малобюджетный (поверхностный), среднебюджетный (ответственный) и полномасштабный (профессиональный). Эти уровни определяются тремя основными факторами: штатом, занимающимся конкурентной разведкой, инструментальными средствами и источниками. Очевидно, что даже среднебюджетный подход подразумевает создание небольших подразделений, покупку некоторых программных средств и подписку на специальные источники информации.

Вместе с тем, нельзя заранее сказать, какой из подходов более адекватен реальности, не вникая в суть деятельности компании. При этом понятно, что для малых фирм в большинстве достаточно поверхностного подхода, а, например, трансконтинентальные корпорации, естественно, требуют профессионального уровня конкурентной разведки.

Казалось бы, как в случае простейшего подхода, достаточно было бы использовать в качестве программы лишь стандартный веб-браузер, а в качестве источника информации - сеть Интернет. Остальное за опытом людей и соответствующими методиками. И основания верить этому рассуждению имеются. По устоявшемуся (но, по-видимому, ошибочному) мнению, сформированному многочисленными шпионскими сериалами и детективами, вся полезная разведывательная информация добывается из строго секретных источников агентурным или оперативным путем, и обязательно сопровождается стрельбой, погонями, взломами кодов, замков и шифров, или, в крайнем случае, соблазнением прекрасной дивноокой и длинноногой носительницы секретов. На самом деле ситуация выглядит далеко не так. Известное признание адмирала Захариаса – заместителя начальника разведки ВМС США в годы Второй мировой войны, опровергает это. По его оценке 95% информации разведка военно-морских сил черпала из открытых источников, 4% - из официальных, и только 1% - из конфиденциальных источников. Справедливости ради надо сказать, что часто

именно этот один процент является тем золотым недостающим звеном, который позволяет сложить целостную картину разрозненной мозаики всех разведанных. И если, такое соотношение справедливо для военной разведки, то тем более оно будет правильным для конкурентной разведки в интересах бизнес-структур.

Разведывательная информация может быть получена из официальных источников, неофициальных открытых источников, СМИ, объявлений, рекламы, внутрифирменных, банковских, правительственных отчетов, баз данных, от экспертов, путем анализа или специальной обработки данных, текстов по прямым или косвенным признакам. Правда, при этом, количество разнородных сведений, которые необходимо переработать, чтобы получить крупницы знаний огромно, а потому в настоящее время конкурентная разведка немыслима без использования специализированных информационных технологий. Следует заметить, что интернет-аналитика сегодня имеет очень большой информационный потенциал и хорошую методическую основу, выработанную годами в сфере бизнес-аналитики

Вместе с тем, наряду с оптимистичным взглядом на открытые источники информации, следует отметить, что по некоторым оценкам более 97% критичной для бизнеса онлайновой информации невозможно найти в Интернете с помощью традиционных информационно-поисковых систем в готовом виде. К сожалению, традиционные сетевые информационно-поисковые системы не в полной мере справляются с задачами конкурентной разведки. Поэтому малобюджетный подход к конкурентной разведке объективно приводит к ряду функциональных ограничений. Т.е. при поверхностном подходе доступность большого числа источников в Интернет (зачастую избыточно большого), опыта и креативности конкурентных разведчиков, не позволяет качественно выполнить поставленные задачи без применения соответствующего инструментария и доступа к базам специальным образом подготовленным данным, для чего, в свою очередь, требуется выделение определенных средств. Кроме того, знания, навыки, контакты конкурентных разведчиков должны хорошо оплачиваться, что не всегда возможно при малобюджетном подходе. Но именно от качества персонала, занимающегося конкурентной разведкой, зависит конечный результат – информация, подготовленная для поддержки принятия решений.

В условиях ограниченных ресурсов аналитикам в области конкурентной разведки зачастую необходимо находить необычные и альтернативные источники информации, которые не лежат на поверхности. Кроме профессиональных навыков работы с открытой частью Интернет, проиндексированной поисковыми системами, необходимо пользоваться ресурсами так называемого «скрытого» или «невидимого» веб, т.е. тех веб-ресурсов, которые по каким-то причинам не охватываются современными поисковыми системами. Именно к ресурсам скрытого веб относится большинство баз данных, размещенных в Интернет, отчеты из которых формируются динамически, по запросам пользователей. Оказывается необходимой для конкурентной разведки информации в Интернет значительно больше, чем ее охватывают универсальные поисковые машины. В отличие от "познаваемой" части Интернет, "скрытый" оказался во много раз более объемным. Практическое решение проблем невидимого Интернет позволяет, по некоторым оценкам, в три раза увеличить полноту сбора документов, касающихся конкретных объектов (предприятий, персон, технологий). Вместе с тем, традиционные поисковые системы чаще всего могут назвать адреса баз данных, но не подскажут, какие документы конкретно содержатся в них. Типичный пример - информационно-поисковые системы по украинскому (<http://www.rada.gov.ua/>) или российскому законодательству (<http://www.kodeks.ru/>). Многие поисковые системы как глобальные, так и локальные - описаны на сайтах Search Engine Watch (<http://www.searchenginewatch.com/>) и Search Engines Ru (<http://www.searchengines.ru/>). На этих сайтах приведены, среди прочих, и поисковые системы "скрытого" веба.

Порталы управления знаниями

Именно благодаря современным инструментальным средствам информационные ресурсы Интернет могут обеспечить целенаправленную аналитическую работу, получение информационно-аналитических материалов достаточно высокого качества в разумные сроки.

Профессиональный подход к конкурентной разведки требует использования соответствующих инструментальных средств. Одним самых перспективных (и недешевых) направлений таких систем являются системы «управления знаниями» (knowledge management) и "обнаружение знаний в базах данных" (knowledge discovery in databases) или иначе, Data и Text Mining, - "глубинный анализ данных или текстов", которые все шире применяются для задач конкурентной разведки. Современные порталы управления знаниями обеспечивают решение целого комплекса задач, среди которых сбор информации об объектах, определение их взаимосвязей, выявление тенденций. Функциональные возможности таких систем позволяют выполнять диагностику и прогнозирование развития ситуации. В дополнение к возможностям глубинного анализа данных и текстов, в порталах знаний используется также человеческий опыт, знания экспертов.

В настоящее время хорошо известна система Lotus Discovery Server фирмы IBM – программный продукт, предназначенный для управления знаниями в корпоративных порталах. Система находит и идентифицирует связи, а также управляет интеллектуальным капиталом корпорации. Благодаря анализу информации Lotus Discovery Server может определять области экспертных знаний, находя и организуя динамические связи между информацией, людьми и их деятельностью.

Среди самых развитых систем управления знаниями, применяемых для решения задач конкурентной разведки, можно назвать систему Hummingbird Enterprise™ канадской компании Hummingbird. Из множества компонент системы можно выделить Hummingbird Portal - платформу, позволяющую интегрировать информацию из информационного хранилища и приложения в едином веб-интерфейсе. Эта платформа, как и ранее названный портал IBM Lotus, является полнофункциональным порталом знаний.

В последнее время все основные западные брэнды, специализирующиеся на разработке хранилищ и баз данных, корпоративных системах управления, расширили свои линейки продуктов модулями business intelligence (BI) или, дословном переводе - деловой разведки. О наличии таких модулей заявляют SAP, ORACLE, SAS, IBM и другие брэнды. Еще одна флагманская платформа для конкурентной разведки - это система американской корпорации Documentum, предназначенная для управления неструктурированной информацией, хранящейся в виде файлов различных форматов. Система Documentum (EMC Platform) основана на трехуровневой архитектуре, включающей хранилище содержания - репозитарий, службу управления содержанием - контент-сервер и клиентские приложения для работы с контентом. Система Documentum позволяет хранить и управлять неструктурированным контентом - Web-контентом, XML-документами, мультимедиа-данными. Репозитарий системы должно обеспечивать как безопасность, так и открытость хранения контента, позволять объединять корпоративные данные в единую корпоративную информационную среду.

Для решения информационно-аналитических задач в настоящее время также широко используется система Cognos Business Intelligence корпорации Cognos. Решение Cognos BI базируется на идеологии OLAP. Одна из особенностей системы - это ее возможность интеграции с компонентами других информационных систем, в том числе необходимых для систем финансово-экономического планирования и управления клиентской базой. В этом случае обеспечиваются широкие возможности сбора и консолидации данных из внутренних и внешних источников.

Говоря о всех выше перечисленных западных системах следует обратить внимание на некоторую разницу в терминах. Говоря о западных продуктах, лидирующих в области Business Intelligence, следует отметить, что под этим термином, как правило, понимается набор инструментальных средств анализа статистических цифровых данных и других корпоративных отчетов и их визуализации, в отличие от competitive intelligence (конкурентной разведки), которая, является гораздо более широким направлением информационной деятельности.

Мониторинг сетевых новостей

Мониторинг сетевых новостей сегодня с полным правом рассматривается как один из видов сервиса, оказываемых в интересах служб конкурентной разведки. Задачи конкурентной разведки требуют извлечения из открытых публикаций данных, касающихся упоминания персон, фирм, брендов, географических названий, адресов, номеров телефонов и т.д. Для аналитиков интересны также параметры источников информации, тематическая направленность публикаций, их тональность, цифровая насыщенность. Отдельный интерес при поиске информации и подготовке аналитических материалов представляют инструментальные средства построения информационных портретов, подборок основных сюжетов, новых событий, дайджестов, таблиц взаимосвязей понятий и объектов, которые относятся к технологиям контент-мониторинга.

Мониторинг сетевых СМИ может поддерживать вполне определенные цели конкурентной разведки: отслеживание действий конкурентов, состояния отрасли. При этом мониторинг полезен на любом уровне, хотя для качественного проведения конкурентной разведки методами анализа текстов из Интернет необходимо построить базы данных для наблюдений и проведения исследований, сформулировать запросы. Заметим, что в любом случае не рекомендуется ограничиваться одной информационно-поисковой системой, даже для анализа такой информации, как Интернет-ресурсы. Для специальных потребностей рекомендуется также использовать законодательные, адресно-номенклатурные, ценовые базы данных, агрегаторы новостей, доступные как из Интернет, так и в локальных версиях. В дополнение к специальным, рекомендуется использовать также лучшие глобальные общедоступные информационно-поисковые системы, такие как Google, Yahoo! или Яндекс.

Запросы к поисковым системам

Покажем, как формируются запросы, относящиеся к бизнес-проблематике, на примере поисковых предписаний к доступной бесплатно системе «Яндекс-Новости» (news.yandex.ru) и системе контент-мониторинга InfoStream (www.infostream.ua).

Обычно поиск информации о компании или персоне всегда начинается с указания различных способов написания названия компании или ФИО персоны. Порой поиска в оперативных и ретроспективных данных по таким "примитивным" запросам вполне достаточно, однако задача усложняется, если необходимо исследовать состояние целой отрасли, отдельного региона, или даже страны. В таких случаях в соответствии с проблематикой строятся запросы, которые затем итеративно уточняются.

В качестве примера назовем ряд проблем, а затем поставим им в соответствие фрагменты запросов, которые могут в дальнейшем использоваться в качестве уточнений к названиям фирм, отраслей, регионов или фамилиям людей. Затем рассмотрим фрагменты текстов, публикуемые различными источниками, которые предположительно можно использовать при построении разного рода аналитических отчетов.

Ниже приведены фрагменты запросов (на информационно-поиском языке системе InfoStream), относящиеся к финансовому положению компаний:

уставной~капитал грн

уставной~фонд долл

финансовое положение

принадлежит~/2/акций

(здесь “~” – оператор близости слов, “~/2/” означает, в частности, близость на расстоянии не более двух слов).

Результаты обработки первого запроса системой «Яндекс-носости» приведены на рис. 1.

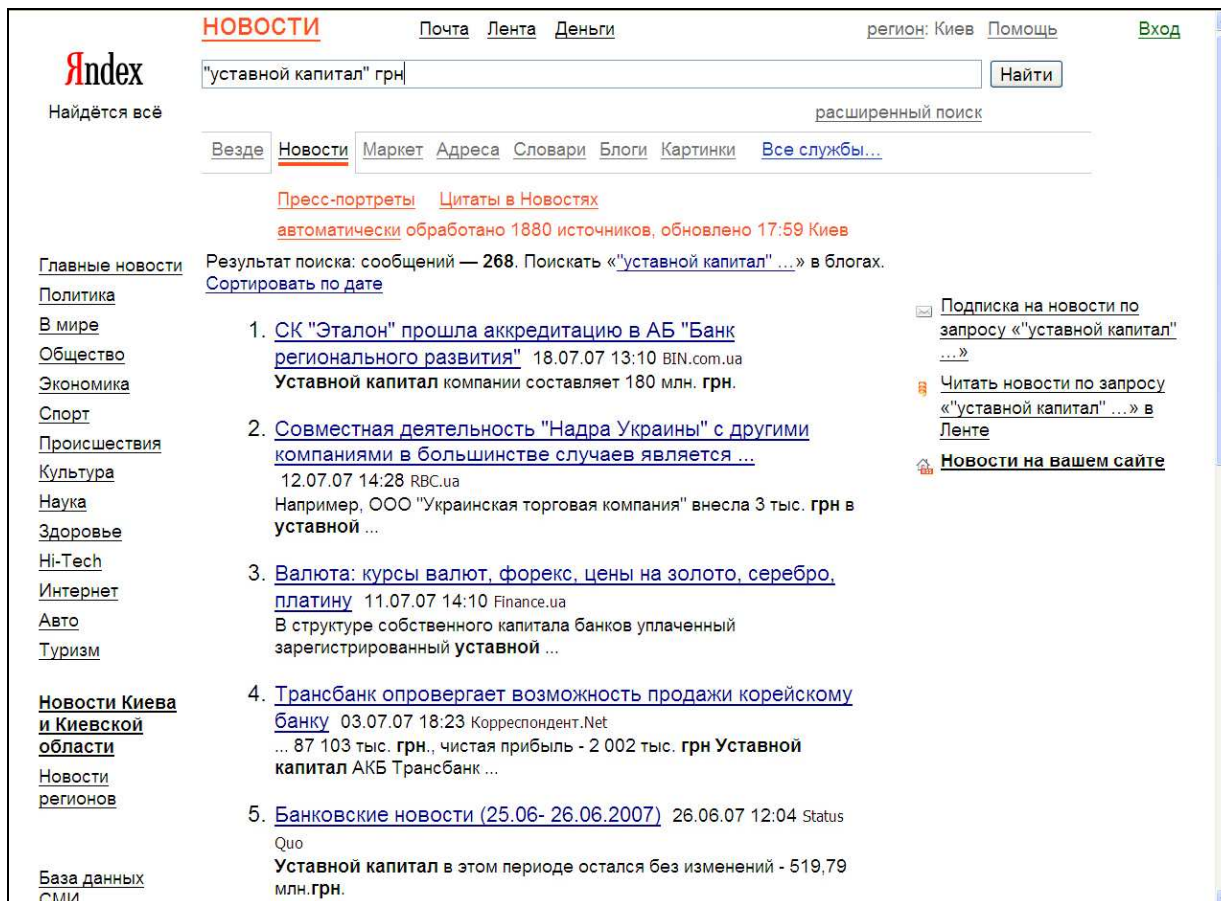


Рис. 1. Результаты обработки запроса в системе «Яндекс-новости»

После обработки приведенных выше запросов системой InfoStream были получены документы, содержащие такие фрагменты:

... Чистая прибыль банка по итогам января-сентября 2006 года составляла 13,164 млн. грн., активы банка - 3 млрд. 348 млн. 001 тыс. грн., обязательства - 2 млрд. 801 млн. 608 тыс. грн., собственный капитал – 546 млн. 394 тыс. грн., **уставной капитал** на 30 сентября с.г. составлял 482,6 млн. грн.

...они, скорее всего, продолжают свое повышение на фоне появившихся сообщений о возможном создании Газпромом и Сибирско-уральской энергетической компанией (СУЭК) совместного предприятия, в **уставной фонд** которого могут быть внесены принадлежащие газовому монополисту 13,5% акций РАО ЕЭС.

... Не хотелось бы нагонять страху, но, откровенно говоря, с учетом перспектив развития украинской экономики в частности и теории экономических циклов вообще, стоит готовиться к тому, что **финансовое положение** многих граждан - заемщиков банков серьезно ухудшится. А вместе с ним и ситуация в банковском секторе Украины.

... Как сообщалось, национальная акционерная компания (НАК) "Нафтогаз Украины" и RosUkrEnergo A. G. (Швейцария), 50% акций которого принадлежит "Газпрому", в феврале 2006 года зарегистрировали в Украине совместное предприятие (СП) ЗАО "УкрГаз-Энерго". "Нафтогазу Украины" и RosUkrEnergo **принадлежит по 50% акций СП**. RosUkrEnergo в набсовете "УкрГаз-Энерго" представляют менеджеры "Газпрома".

Информация о слияниях и приобретениях в той или иной сфере бизнеса, позволяющая следить за экспансией конкурентов в новые рыночные ниши может быть получена в результате отработки таких уточняющих запросов:

приобрел/2/акций (допустимо, например, "контрольный пакет акций")
продал~/2/пакет~акций
слияни & компан & (акци, актив)

... На украинский банковский рынок выйдет еще один российский игрок - инвестиционный банк "КИТФинанс" (Санкт-Петербург). Один из его собственников - финансовая группа "Вэб-инвест" **приобрела** более 50% **акций** днепропетровского Радабанка.

... Компания частных инвестиций Advent International - одна из ведущих мировых инвестиционных групп - **приобрела** контрольный пакет **акций** киевского завода безалкогольных напитков, одного из лидеров пищевой промышленности ОАО "Росинка".

... Фонд государственного имущества 15 февраля на Киевской международной фондовой бирже **продал 45,98% акций** отдельного комплекса "Мир" (Северодонецк, Луганская область).

... Если мы **активы** "Газпрома", стоимость которых и так превышает 300 млрд долларов, будем применять ко всем **компаниям**, используя **слияния** и поглощения, мы получим государственный монопольный капитализм, который отбросит страну далеко назад".

Для выявления публикаций об изменении финансового состояния или банкротства можно использовать такие уточняющие запросы:

выпуск~/2/акций
(увеличить~уставной)&(фонд,капитал)
продать~/2/акций
объявить~/2/банкротство

Отработка подобных запросов позволило найти документы, содержащих такие фрагменты:

... Премьер-министр Украины Виктор Янукович подписал постановление об увеличении уставного капитала ОАО "Государственный экспортно-импортный банк Украины" до 1 млрд 286 млн грн. УК увеличен за счет выпуска дополнительного **выпуска акций** в бездокументарной форме на сумму 367 млн 583 тыс. грн с сохранением в собственности государства 100% таких акций.

... По словам представителя "КИТ Финанс" Вадима Бараусова, через Радабанк компания намерена продавать ипотеку и другие розничные продукты, планируется **увеличение уставного капитала** и развитие региональной сети.

... НАСТА планирует продать часть компании, ее президент Гарри Делба заявил еще летом 2006 года. Однако, по его словам, компания готовилась **продать лишь 25% акций**, притом что стоимость СК оценивалась в \$200 млн.

... Вспомним декабрь 2001 года, когда американская топливно-газовая корпорация Enron Corporation неожиданно для всех **объявила о своем банкротстве**. Это самое крупное в мире банкротство исчислялось суммой в 50 миллиардов долларов!

Контент-мониторинг

Процесс конкурентной разведки можно рассматривать как построение и исследование сети из исследуемых объектов и связей между ними. Результаты такого анализа должны представлять собой информационно-аналитические отчеты, которые могут быть использованы для принятия решений. Аналитическая информация может быть представлена в виде отдельных статей, дайджестов, сюжетных линий, семантических сетей, таблиц взаимосвязей ключевых понятий, компаний, лиц, технологий и т.п. Для эффективного решения этих задач, естественно, необходимы специальные информационные технологии, обеспечивающие возможность извлечения фактов из текстовых массивов и обработки необходимой информации, что в свою очередь вызвало поток предложений систем со стороны разработчиков программного обеспечения.

Наиболее прогрессирующими методами конкурентной разведки сегодня являются мониторинг, прогнозирование, оперативный анализ ситуаций на рынках. Охват, обобщение больших динамических информационных массивов для решения задач конкурентной разведки требует качественно новых подходов. Возникла необходимость создания методов мониторинга информационных ресурсов, тесно связанных с традиционной методологией контент-анализа. В связи с этим стали актуальными технологии контент-мониторинга - содержательного анализа информационных потоков с целью получения необходимых качественных и количественных срезов, который, в отличие от контент-анализа, производится непрерывно во времени. Т.е. методы контент-мониторинга - это адаптация классических методов контент-анализа к условиям динамических информационных массивов, например потоков информации из Интернет.

В отличие от систем интеграции информации, которые реализуют идею сбора и накопления всей доступной информации, как из внутренних, так и из внешних источников, системы контент-мониторинга позволяют выявлять неочевидные закономерности в документальных массивах данных или текстов - так называемые латентные (скрытые) знания. В целом эти технологии еще определяют как процесс обнаружения в "сырых" данных ранее неизвестных, но полезных знаний, необходимых для принятия решений. Системы этого класса позволяют осуществлять анализ больших массивов документов и формировать предметные указатели понятий и тем, освещенных в этих документах. Характерная задача конкурентной разведки, обычно включаемая в системы контент-мониторинга - это нахождение исключений, то есть поиск объектов, которые своими характеристиками сильно выделяются из общей массы. Еще один класс важных задач, решаемых в рамках этой технологии - это моделирование данных, ситуационный и сценарный анализ, а также прогноз. И наконец, для обработки и интерпретации результатов большое значение имеет визуализация, которая обычно используется как средство представления контента всего массива документов, а также для реализации механизма навигации по семантическим сетям, который может применяться при исследовании как отдельных документов, так и их классов.

Типичная задача контент-мониторинга - построение диаграмм динамики появления понятий во времени. Рассмотрим, как в системе InfoStream отслеживались кризисные явления на рынке сахара в Украине в конце мая 2005 года. Для этого был составлен запрос *"кризис & сахар & Украина"*, который был введен через веб-интерфейс системы и получена соответствующая гистограмма (рис. 2).

Англогичные графики появления понятий во времени дают некоторые российские аналитические системы, например «Артефакт» или «Галактика Zoom». В частности, «Галактика ZOOM» - это одна из ведущих российских систем, обеспечивающих решение задач анализа больших массивов текстовой информации (что с успехом используется в практике конкурентной разведки). Эта система позволяет выполнять процедуры мониторинга информационного пространства, распределения по тематике информационного потока, исследование, анализ и выявление основных тенденций. «Галактика ZOOM». На

рис. 3 приведена диаграмма появления понятий, соответствующих тематике «мобильная связи».



Рис. 2. Динамика появления документов, соответствующих заданному понятию в системе InfoStream

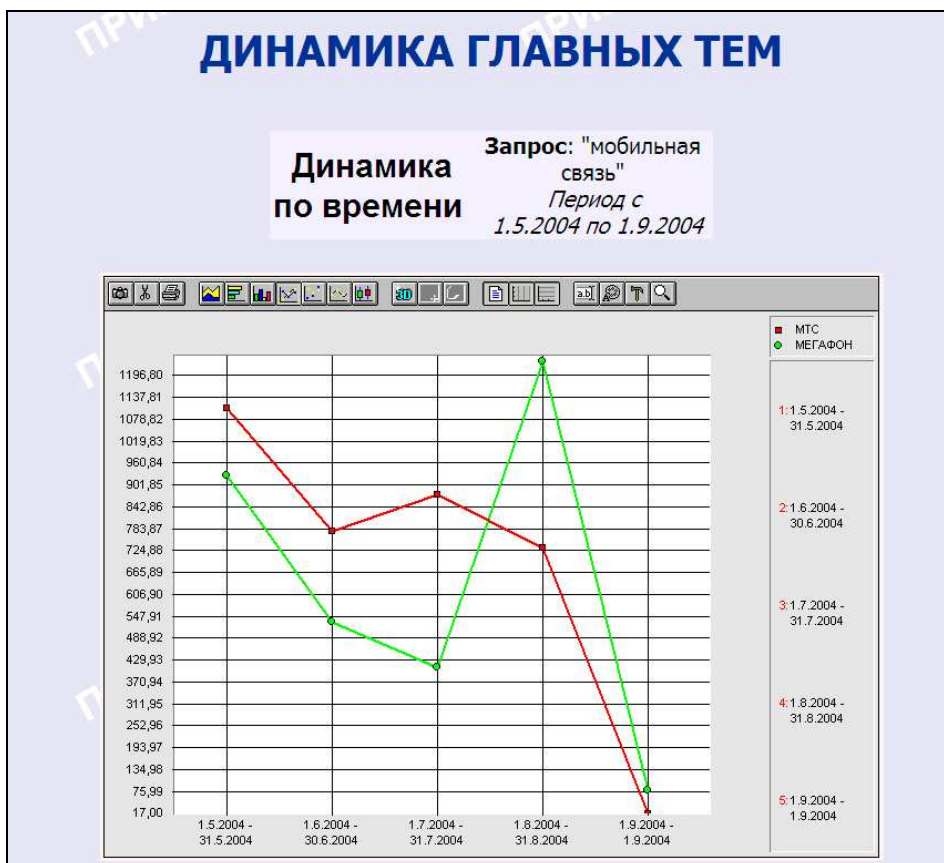


Рис. 3. Диаграмма появления понятий во времени в системе «Галактика Zoot»

Одна из возможностей систем контент-мониторинга заключается в автоматическом выявлении сюжетных цепочек из потоков документов, например, содержащих максимальное количество ценовой информации по данному рынку.

В частности, в системе InfoStream для получения основных сюжетов, относящихся к рынку нефтепродуктов можно ввести запрос *"(нефтепродукты | бензин) & цены"*, уточнив его специальными признаками *"numb.medium | numb.large"*, означающее в InfoStream средний или высокий уровень заполненности документов цифровой информацией (рис. 4).

Обзор основных сюжетов (16 февраля 2007)

"(нефтепродукты | бензин) & цены" (большая цифровая насыщенность, цифровая насыщенность) Распечатать

документов - 208, сюжетов - 70

- РАБОТА УДОВЛЕТВОРИТЕЛЬНАЯ**

Политико-консультативный совет (ПКС) при одесском городском голове признал удовлетворительной работу управления внутренней политики Одесского городского совета, передает корреспондент ИА "Контекст-Причерноморье". Такое решение принято 7 февраля на заседании Политико-консультативного совета при одесском городском голове. Начальник управления внутренней политики Одесского городского совета // Независимая газета "Юг" 2007.02.10 16:10

(15) [Похожие документы >>](#)

 - ОСТРОВ РАЗДОРА // Независимая газета "Юг" 2007.02.14 22:10
 - ТЕХОСМОТР НЕ ПОДОРОЖАЕТ // Независимая газета "Юг" 2007.02.14 22:10
 - КОРЗИНА С ПРОРЕХОЙ // Независимая газета "Юг" 2007.02.14 22:10
 - "ГЛАЗА Б МОИ ТЕБЯ НЕ ВИДЕЛИ..." // Независимая газета "Юг" 2007.02.14 22:10
 - УТОНУЛ ШКИПЕР // Независимая газета "Юг" 2007.02.14 22:10
- Росстат: с 29 января по 4 февраля средние потребительские цены на автомобильный бензин снизились на 0,1 проц**

Как сообщает Федеральная служба госстатистики, с 29 января по 4 февраля потребительские цены на автомобильный бензин снизились на 0,1 проц. Средняя потребительская цена на бензин автомобильный в целом по России на 5 февраля снизилась на 0,02 руб. // ИА "Финмаркет" 2007.02.13 13:20

(14) [Похожие документы >>](#)

 - Потребительские цены на бензин в России за неделю снизились на 0,1% // ИА "АКС - Реальный сектор" 2007.02.14 12:03
 - Цены на бензин в РФ снизились на 0,1% // "Век" 2007.02.13 16:14
 - Цены на бензин в РФ снизились на 0,1% // "RAMBLER" 2007.02.13 16:01
 - Цены на бензин в России за неделю снизились на 0,1% // "КМ-Новости" 2007.02.13 15:29
 - Цены на бензин в России: последние данные // "Газета.Ru" 2007.02.13 15:22
- Саудовская Аравия спровоцировала резкое падение цен на нефть**

Начало новой торговой недели на нефтяном рынке было отмечено резким падением цен на нефть. Так, нефтяные фьючерсы 12 февраля 2007 г. подешевели за один день более чем на 2 доллара за баррель, а ценовые показатели упали в ходе торгов ниже 57 долл./барр "Нефть России", 13 февраля Начало новой торговой недели на нефтяном рынке было отмечено резким падением цен на нефть // "Номад-кочевник" 2007.02.14 06:41

(13) [Похожие документы >>](#)

 - Мировые цены на нефть снова повысились // "Экономические новости Украины" 2007.02.14 18:17
 - Мировые цены на нефть снова повысились // "Финансовые новости" 2007.02.14 18:17
 - Саудовская Аравия спровоцировала резкое падение цен на нефть // "What's New?" 2007.02.13 18:26
 - На мировом рынке резкое падение цен на нефть // "Деньги" 2007.02.13 17:15
 - Саудовская Аравия спровоцировала резкое падение цен на нефть // "Энергетика и промышленность России" 2007.02.13 12:18

Рис. 4. Цепочка основных сюжетов в системе InfoStream

В системе «Яндекс-новости» также предусмотрено построение сюжетных цепочек из тематических новостных сообщений (рис. 5).

Перспективным направлением развития технологий конкурентной разведки является также автоматическое извлечение понятий из неструктурированных текстов, а также построение таблиц взаимосвязей и гистограмм распределения понятий. На рис. 6 показан пример визуализации связей между отдельными организациями, построенной с помощью полномасштабной системы конкурентной разведки X-Scif.

Как правило, таблицы взаимосвязей понятий строятся как статистические отчеты, отражающие близость (совместную встречаемость в новостных сообщениях или близость по сопутствующему контексту, другим понятиям) отдельных понятий. Эти таблицы можно рассматривать как неориентированные графы и применять к ним соответствующие математические методы. Как правило, узлы этих графов – коэффициенты, которые пропорциональны количеству документов входного информационного потока, одновременно соответствующие обоим понятиям, или количеству других понятий, употребляемых совместно с данными понятиями.

Яндекс
Найдётся всё

НОВОСТИ Почта Лента Деньги регион: Киев Помощь Вход

бензин цена

расширенный поиск

Везде **Новости** Маркет Адреса Словари Блоги Картинки Все службы...

[Пресс-портреты](#) [Цитаты в Новостях](#) [автоматически обработано 1880 источников, обновлено 17:29 Киев](#)

Результат поиска: сообщений — **76757**. Поискать «бензин цена» в блогах.
Сортировать по дате

1. **Потребительская корзина подорожала на 19%** 17:52 СаратовИнформ
... наибольшее влияние оказал рост **цен** на продукты переработки зерна и изделий из ...
... на 14,2%) и доски обрезные (на 17,2%). **Цены на бензин** выросли в среднем на одну ...

- **Индекс потребительских цен** в Республике Коми в ... 16:50 ИА РБК Санкт-Петербург
- **В Коми в июле хлеб подорожал на 9 процентов** 15:44 КомиОнлайн
- **Сюжет полностью (108) >>**

2. **На "сговоровшихся" владельцев бензоколонок подадут в суд** 17:29 ТРК Белый город
В конце июля в регионе существенно выросли **цены на бензин**, дизельное топливо и ...
... нет ли сговора между продавцами, когда **цены** на товары поднимают, предварительно ...

- **Кто взводит ценовой курок бензинового пистолета?** 06:08 07 21:19 Промышленные ведомости
- **Сюжет полностью (2) >>**

3. **Беларусь: Розничные цены на бензин повышены** 16:48 Ukraine Daily
Концерн "Белнефтехим" с 6 августа повысил розничные **цены** на нефтепродукты в ...
Об этом сообщает "БЕЛТА" со ссылкой на "Белнефтехим". С 6 августа розничные **цены** ...

- **Потребительские цены на бензин в России за неделю** ... 16:43 РесБалт
- **Цены на бензин** развили черепашую скорость 14:57 Neftgaz.RU
- **Сюжет полностью (54) >>**

4. **Цены на бензин в США снизились** 15:56 Ореанда
Розничные **цены на бензин** в США упали третью неделю подряд, опустившись еще на 3 ...
Средняя стоимость обычного неэтилированного **бензина** по стране составила \$2,84.

- **Fibo Group: Снижение котировок нефти в понедельник** ... 14:15 Quote.Ru
- **Цены на бензин в США** снизились третью неделю ... 12:31 Press-uz.info
- **Сюжет полностью (18) >>**

5. **Штрафы - действенный инструмент ФАС** 15:22 Бел.Ру
У журналистов возникли вопросы по поводу **цен**, сложившихся в области на ...
... года существенно выросли **цены** на муку, **бензин** и дизельное топливо, сжиженный газ ...

- **Сюжет полностью (8) >>**

6. **"Лукойл" вызвали в суд конкуренты** 14:42 KM-News
... информации обвинения, стало общее повышение **цен на бензин** и крупные потери среди ...
... вернуть незаконно полученный доход, уплатить штраф и снизить **цены на топливо**.

Подписка на новости по запросу «бензин цена»
Читать новости по запросу «бензин цена» в Ленте
Новости на вашем сайте

Главные новости
Политика
В мире
Общество
Экономика
Спорт
Происшествия
Культура
Наука
Здоровье
Hi-Tech
Интернет
Авто
Туризм
Новости Киева и Киевской области
Новости регионов
База данных СМИ

Рис. 5. Основные сюжеты по тематическому запросу «цена на бензин» в системе «Яндекс-новости»

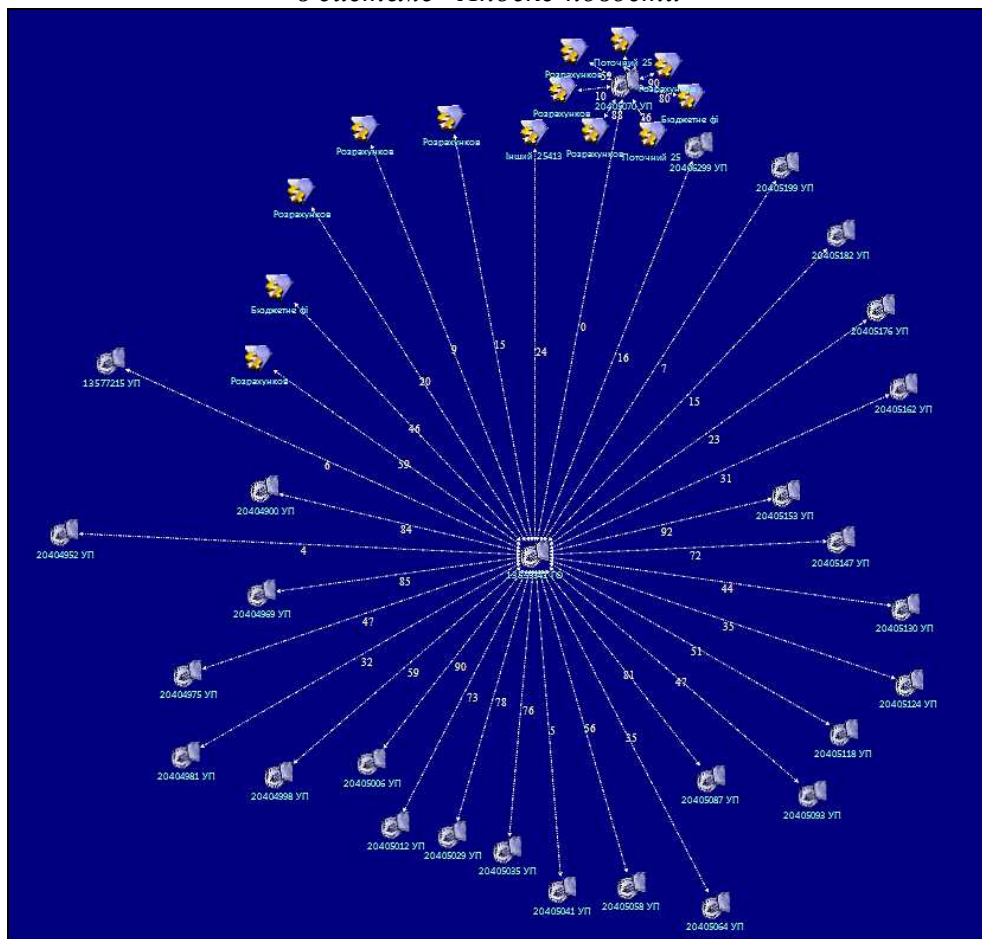


Рис. 6. Визуализация взаимосвязи компаний в системе X-Scif

Заключение

Можно констатировать, что значение конкурентной разведки постоянно растет. С одной стороны, это связано с рыночными процессами и ростом конкуренции во всех областях бизнеса, а, с другой, с бурным развитием информационных технологий. Сегодня задачи конкурентной разведки стимулируют развитие систем управления знаниями, глубинного анализа данных и текстов, с другой стороны наиболее развитые из этих систем в явном виде содержат аналитические блоки, специально ориентированные на задачи конкурентной разведки. Кроме того, законодательные акты многих стран мира все более адаптируются к проблемам свободы доступа и использования любой информации, в том числе и коммерческой. Как правило, ограничения в законодательствах развитых стран касаются лишь материалов, имеющих отношение к обороне, документов правоохранительных органов, личных и финансовых документов.

Уровни функциональности компьютерных систем конкурентной разведки, может быть очень разнообразным - от простых информационно-поисковых систем, необходимых на этапе становления, до ресурсоемких систем управления знаниями и глубинного анализа данных и текстов. Состав и объемы информации, необходимой для решения задач конкурентной разведки, бывают настолько огромными, что даже специализированные поисковые системы не всегда способны оперативно сформировать необходимый отчет. Именно для решения данной проблемы предлагаются системы управления и порталы знаний, представляющие среду для эффективного поиска и обмена знаниями.

В настоящее время даже малобюджетные варианты конкурентной разведки в сети Интернет обеспечивает доступность, огромный охват информации и высокую оперативность. Но такой подход не может заменить другие виды и инструментальные средства бизнес-разведки. Для принятия серьезных решений необходимо использование комплексных систем, которые разрешают компоновать и обобщать информацию об объекте исследований, полученную из разных источников с применением разных технологий.