

МОДЕЛЬ РЕПУТАЦИИ ИНТЕРНЕТ-ИСТОЧНИКА**А.П. Кузнецова, А.А.Гавриченко (Владимир)**

В условиях постоянно растущего уровня информатизации общества лицо, принимающее решение (ЛПР) сталкивается с проблемой обработки все большего количества информации из разнообразных источников с различным уровнем надежности содержащейся в них информации. Для того чтобы предсказать, насколько верно выработанное в ходе информационной деятельности решение, необходимо оценить уровень качества использованной при этом информации. Эту проблему решает грамотно построенная аналитическая работа, которая предназначена для оценки информации и подготовки принятия решений.

Информационно-аналитическая работа составляет основное содержание повседневной работы администратора безопасности (АБ), осуществляющего основные функции управления ИБ в ИТКС. Эффективная аналитическая деятельность АБ возможна только при наличии достаточно полной, своевременной и достоверной информации не только об ИТКС, но и об ее окружении. При анализе конкретного информационного сообщения необходимо, в том числе, учесть влияние источника информации, поставившего данное сообщение, на уровень его надежности. Заметим, что надежность – т.н. уровень доверия к источнику информации – зависит от типа задачи и модели предметной области, поэтому далее мы будем рассматривать информационно-аналитическую работу с точки зрения решаемых АБ задач.

Одним из самых сложных и малоизученных типов источников информации являются интернет-источники (ИнТИ) как современный тип технических средств. С одной стороны – легкость в получении информации от них сопровождается, как правило, низким ее качеством. Однако, использование ИнТИ в аналитической работе АБ – задача актуальная и имеет явную практическую востребованность. Тем не менее, явно не хватает методик и средств ранжирования ИнТИ по качеству получаемой от них информации.

Интернет-источники отличаются большим разнообразием, что позволяет в зависимости от решаемой задачи выбрать тот или иной источник информации. Анализ типовых функций АБ в аналитической деятельности и их информационное содержание, информационное наполнение известных ИнТИ позволяет выделить конечное множество категорий интернет-источников, необходимых и достаточных для аналитической деятельности АБ:

- К1. Официальные сайты государственных органов (предприятий/учреждений).
- К2. Официальные сайты научно-исследовательских и образовательных организаций.
- К3. Официальные сайты негосударственных организаций.
- К4. Блоги.
- К5. Тематические сайты (информационные порталы).
- К6. Форумы.
- К7. Новостные сайты.

На уровень доверия к ИнТИ вне зависимости от его типа оказывает сильнейшее влияние его репутация – комплексный показатель, складывающийся из нескольких факторов. Репутация говорит о том, насколько авторитетен и объективен данный источник. Особенность показателя репутации состоит в том, что он не зависит от предметной области. Задача состоит в следующем: на основе множества имеющихся в наличии интернет-источников, множества факторов, влияющих на показатель репутации ранжировать ИнТИ для конкретно решаемой задачи. Для этого необходимо сформировать интегральный показатель *репутации интернет-источника*, который будет учитывать степень важности каждого из факторов.

Для выявления наиболее значимых факторов, влияющих на показатель репутации, была выполнена коллективная экспертиза с согласованием экспертных оценок. Были выделены 4 наиболее существенных фактора по экспертному мнению. В результате наиболее важными факторами были названы следующие:

- Цитируемость ИнТИ (Ц) – один из параметров, по которым поисковые системы оценивают его качество. Она определяется как количество обратных ссылок на сайт, которые поставлены на других веб-ресурсах. Для измерения данного параметра существует множество онлайн seo-анализаторов.
- Посещаемость (П). Данный фактор отражает число уникальных посетителей, зашедших на сайт за определенный период. Для удобства расчетов в данной работе за меру посещаемости сайта берется количество уникальных посетителей за период, равный одному месяцу. Этот фактор также можно получить, используя инструменты для seo-анализа.
- Длительность существования (В). Для получения информации о возрасте сайта применимы различные способы. Они зависят от типа ИнТИ, особенностей архитектуры сайта и уровня домена, на котором сайт размещен. Для сайтов, размещенных на втором доменном уровне, она приравнивается к возрасту домена. Для сайтов на доменах последующих уровней длительность существования предлагается рассчитывать с даты публикации первого поста, даты регистрации пользователя или другим возможным способом получения информации о начале работы ИнТИ.
- Доверие Web (Д). Данный фактор основан на инструменте Web of Trust (WOT) – системе оценки репутации веб-сайтов, основанной на отзывах пользователей. WOT – это бесплатная надстройка к интернет-браузеру, которая при помощи цветных иконок предупреждает посетителя сайта о пользовательской оценке его репутации.

Поставим в соответствие каждому фактору свой параметр, назовем данные параметры так же, как и факторы: параметр цитируемость обозначается P^C , параметр посещаемость обозначается P^P , параметр длительность существования обозначается P^B , параметр доверие Web обозначается P^D . Поскольку каждый из факторов является величиной относительной, чтобы решить проблему измерения факторов, нормируем шкалы так, чтобы каждый из параметров изменялся в диапазоне $[0;1]$.

Для оценки уровня доверия к ИстИ по показателю репутации введем следующие формулы:

- Показатель репутация P_k^I k -го источника по параметру цитируемости p_k^I для m -й категории ИнтИ.

$$P_k^I = (p_k^I)^{\sigma_m^I / (\sigma_m^I + \sigma_m^{II} + \sigma_m^B + \sigma_m^D)} \quad (1)$$

- Показатель репутация P_k^{II} k -го источника по параметру посещаемость p_k^{II} для m -й категории ИнтИ.

$$P_k^{II} = (p_k^{II})^{\sigma_m^{II} / (\sigma_m^I + \sigma_m^{II} + \sigma_m^B + \sigma_m^D)} \quad (2)$$

- Показатель репутация P_k^B k -го источника по параметру длительность существования p_k^B для m -й категории ИнтИ.

$$P_k^B = (p_k^B)^{\sigma_m^B / (\sigma_m^I + \sigma_m^{II} + \sigma_m^B + \sigma_m^D)} \quad (3)$$

- Показатель репутация P_k^D k -го источника по параметру доверие Web p_k^D для m -й категории ИнтИ.

$$P_k^D = (p_k^D)^{\sigma_m^D / (\sigma_m^I + \sigma_m^{II} + \sigma_m^B + \sigma_m^D)}, \quad (4)$$

здесь σ_m^I , σ_m^{II} , σ_m^B , σ_m^D – коэффициенты важности соответственно факторов цитируемости, посещаемости, длительности существования и доверия Web. Они были получены методом экспертного анализа. Экспертам было предложено оценить факторы в зависимости от степени их влияния на показатель репутации для каждой категории ИнтИ.

- Интегральный показатель репутации P_k k -го интернет-источника:

$$P_k = P_k^I \times P_k^{II} \times P_k^B \times P_k^D. \quad (5)$$

В ходе апробации данной модели оценки репутации ИнтИ были получены значения на более 100 реальных ИнтИ, необходимых в деятельности АБ. Следующие ИнтИ получили наивысшие значения уровня доверия по показателю репутации: официальный сайт ЦБ РФ (www.cbr.ru) со значением $P_k = 0,91$, официальный сайт Президента России (www.kremlin.ru) $P_k = 0,89$, официальный сайт МВД России (mvd.ru) $P_k = 0,88$, интернет-портал СNews (www.cnews.ru) $P_k = 0,87$, сайт ЗАО «Лаборатория Касперского» (www.kaspersky.ru) $P_k = 0,85$.

Данная модель, представленная в виде методики как последовательность действий позволяет:

- подобрать основные интернет-источники для анализа с точки зрения эффективности применения в задачах АБ;
- выявить существенные факторы, влияющие на показатель репутации;
- определить параметры, характеризующие репутацию;
- ранжировать интернет-источники по комплексному показателю.