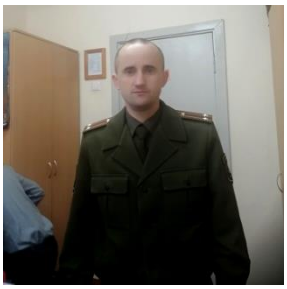
УДК 314/316:004.6

СОЦИАЛЬНАЯ ИНЖЕНЕРИЯ И ТЕХНОЛОГИИ BIG DATA



Д.М. Румянцев
Военный факультет БГУ,
преподаватель

Белорусский государственный университет, Республика Беларусь
E-mail: dimon-rum@mail.ru

Д.М. Румянцев
Военный факультет БГУ,преподаватель.

Аннотация. «Большие данные» становятся новой обсуждаемой темой в социальных науках. Начиная с 2012 г., многие определяют данное понятие через такие признаки, как объем (volume), скорость (velocity) и разнообразие (variety) [1; 2, с. 18]. Коммерческие компании, предоставляющие программы по сбору «больших данных», также подчеркивают точность (veracity) получаемой информации о поведении пользователей, по сравнению с опросами пользователей о мнениях и поведении [3]. Социальная инженерия – совокупность приёмов, методов и технологий создания такого пространства, условий и обстоятельств, которые максимально эффективно приводят к конкретному необходимому результату, с использованием социологии и психологии. Зачастую социальную инженерию рассматривают как незаконный метод получения информации, однако это не совсем так. Конечно, сегодня социальную инженерию зачастую используют для получения закрытой информации в интернете. Однако, если рассматривать современную профессиональную социальную инженерию, то область её применения вполне законна (например, она помогает достичь изначально недостижимый результат, или “программировать” для совершения полезных действий конкретного человека или группы людей. Таким образом, появляется новое направление сотрудничества между компьютерными и социальными науками, связанное с анализом и использованием результатов анализа больших данных. Эти же большие данные могут стать источником прибыли для злоумышленников. В данной работе попытаемся проанализировать влияние развития направления Big Data на безопасность пользователей в интернете.

Ключевые слова: Big Data, социальная инженерия, информационная безопасность.

Введение. В современном мире для проведения своих атак злоумышленники, применяющие техники социальной инженерии, зачастую эксплуатируют доверчивость, лень, любезность пользователей и сотрудников организаций. Защищаться от таких атак непросто, поскольку их жертвы могут не подозревать, что их обманули. Злоумышленники, использующие методы социальной инженерии, преследуют, в общем, такие же цели, что и любые другие злоумышленники: им нужны деньги, информация или ИТ-ресурсы. Big Data – социально-экономический феномен, который связан с тем, что появились новые технологические возможности для анализа огромного количества данных. Эти данные могут быть использованы против каждого отдельного пользователя интернета.

Причины роста количества атак

Популярность социальной инженерии среди злоумышленников растет потому, что нередко сами работники предприятия – люди являются самым слабым звеном в системе защиты. У данного факта много объяснений, во-первых – нередко часть работников просто недостаточно обучена, и им не хватает знаний, чтобы избежать такой атаки, а также большую роль играет и то, что большая часть предприятий думает только о защите физического

периметра от внешних угроз. При помощи сотрудника, обойдя эту внешнюю защиту, злоумышленник обходит самое большое препятствие.

Социальная инженерия является важным аспектом в контексте предприятия в целом, так как системы защиты создают для злоумышленника довольно сложно преодолеваемый барьер, и в данном случае неважно, какого именно работника удалось злоумышленнику обмануть, так как результат – доступ ко всем внутренним ресурсам, минуя барьер защиты, будет одинаковым во всех случаях. Атаки социальной инженерии нередко ориентированы на работников, у которых есть самые большие права доступа к работе с конфиденциальной информацией, однако злоумышленник нередко оценивает и потенциальные знания цели.

Одной из важных причин распространения социальной инженерии как метода атаки – это очень дешевый вид нападения, атакующий может не быть специалистом в сфере информационных технологий. Существенным фактором является также и то, что при использовании методов социальной инженерии результат нередко достигается гораздо быстрее, чем, если бы был использован иной метод для нападения, для сравнения – зачем пытаться взломать систему защиты дверь, если неподготовленный пользователь сам готов нас впустить.

Большие данные. Системы управления большими данными

Если давать краткое определение, то Большие данные – это данные, которые не помещаются в оперативную память компьютера. По сути, это определение обозначает то, что свойство “быть большим” является не самостоятельным свойством данных, а зависит от характеристики системы, применяемой для их обработки. Например, обычному человеку затруднительно запомнить какая именно температура была в нашем городе каждый день за прошедший месяц. Таким образом, три десятка значений вполне могут быть примером Больших данных. Однако вот человек уверенно сообщает “прошедший месяц был холодным”. Это сообщение несет информацию об обработанных данных: по мнению собеседника, средняя температура за прошедший месяц была ниже, чем обычно в этом месяце за несколько десятков лет. Другим примером могут быть данные об объектах, которые теоретически несут важную информацию, однако имеющие такой размер, что эти данные практически невозможно не только обработать или сохранить, но даже собрать. Рассмотрим, к примеру, набор данных, содержащий координаты и скорости молекул в воздушном столбе над территорией аэропорта. Имеются также метаданные с описанием в какой момент проводилось измерение и что это за молекула. Такой набор данных несет информацию о погодных условиях над аэропортом, включая температуру, давление, влажность, облачность, особые погодные условия – проходящий торнадо или падающий град. С другой стороны, для корректной обработки данные для всех молекул должны быть достаточно полны и репрезентативны для статистической обработки. В результате такого мысленного эксперимента мы понимаем, что для эффективной работы с большими данными нужна модель данных, позволяющая сформировать методы работы с данными. Данные могут быть различных типов.

Информацию, полученную в результате учёта или измерения каких-либо объектов или параметров, называют мастер-данными (Master Data). Например, учёт количества, замеры координат и скоростей конкретных молекул – это мастер-данные.

Транзакционные данные (в англоязычной литературе применяются термины Transactional Data, Application Specific Data, Operational Data) – это данные, отображающие результат выполнения каких-либо операций. Например, данные о взаимодействии молекул между собой, а именно о пересечении границ рассматриваемой области, о траектории конкретной молекулы, об испарении капель дождя – это транзакционные данные. Транзакционные данные описывают взаимодействие объектов друг с другом или с окружающим миром, которые можно получить при помощи обработки мастер-данных.

Ретроспективные данные (Historical data) – это данные, снабженные метками времени. Например, с одной стороны мы можем сохранять данные о координате и векторе скорости

каждой молекулы, но если у нас есть набор координат в зависимости от времени, то скорость молекулы становится лишней, она вычисляется исходя из модели, описываемой ньютоновской механикой.

Ссылочные данные (справочники, НСИ, нормативно-ссылочная информация, Reference Data, Lookup Data, Dictionaries) – это базовые неизменяемые данные, заранее известные из внешних источников, такие как нормативы, сокращения, акронимы, словари, стандарты. Например, удельные веса молекул, зависимость температуры замерзания и кипения от давления, зависимость средней скорости молекул (скорости звука) от температуры. Формат данных. Структурированные данные имеют заранее определенный формат.

Полуструктурированные или слабоструктурированные данные – это данные, зачастую собранные из различных источников. Структура данных документирована, но в зависимости от источника данных конкретный формат представления информации может быть разным. Неструктурированные данные требуют обязательной обработки и последующей валидации перед использованием. Например, данные о координатах и скоростях молекул, в которых некоторые координаты пропущены или некоторые записи повторяются, являются полуструктурированными. Нам нужно понять, почему так произошло и перед использованием либо исключить такие данные (что может привести к систематической ошибке), либо, исходя из модели данных, восстановить пропущенные значения.

Таблица 1. – Сравнительный анализ традиционной базы и больших данных.

Характеристика	Традиционная база данных	База больших данных
Объем информации	От гигабайт (10^9 байт) до терабайт (10^{12} байт)	От петабайт (10^{15} байт) до эксабайт (10^{18} байт). N="All"
Способ хранения	Централизованный	Децентрализованный
Структурированность данных	Структурирована	Полуструктурирована и неструктурирована
Модель хранения и обработки данных	Вертикальная модель	Горизонтальная модель
Взаимосвязь данных	Сильная	Слабая

Данные, в которых координаты измеряются в разных единицах измерения, числа иногда записаны словами, иногда латинскими цифрами, а иногда в виде сканированного изображения почерка лаборанта, являются неструктурированными данными. Обычно Большие данные описываются при помощи следующих характеристик. [4]

1. *Объем* (Volume) – количество сгенерированных и хранящихся данных. Размер данных определяет значимость и потенциал данных, а также то, могут ли они быть рассмотрены как Большие данные.

2. *Разнообразие* (Variety) – тип данных. Большие данные могут состоять из текста, изображений, аудио, видео. Большие данные при сопоставлении друг с другом могут дополнять отсутствующие данные.

3. *Скорость* (Velocity) – скорость. Здесь подразумевается скорость, с которой данные генерируются и обрабатываются. Очень часто Большие данные используются в режиме реального времени.

4. *Изменчивость* (Variability) – противоречивость наборов данных может препятствовать их обработке и управлению ими.

5. *Достоверность* (Veracity) – качество данных напрямую влияет на точность проведения анализа данных. Большие данные могут быть классифицированы в соответствии

с несколькими главными компонентами. Интеллект-карта, представленная на рис. 1, была составлена на основе [5].

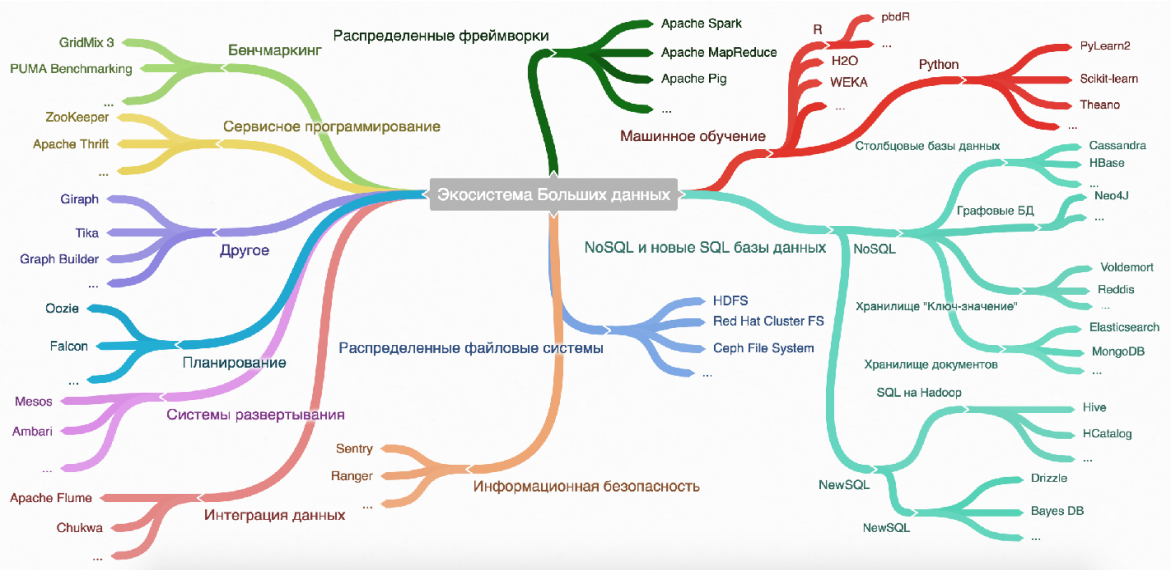


Рисунок *1.* – Интеллект-карта

Основные этапы социального взлома

Существует несколько существенных этапов, которые характерны для большинства атак с применением социальной инженерии:

–Сбор информации

Один из этапов, от которого больше всего зависит “успех операции”. На нем осуществляется сбор информации о потенциальной жертве. Широкий спектр источников позволяет получить максимум знаний и определить следующие шаги. Примеры источников и их использование для сбора инофрмации:

–Социальные сети (Вконтакте, Facebook и прочее).

Обычно, создавая страницы в социальных сетях, люди не задумываются, как много информации для злоумышленников они оставляют в открытом доступе, предоставляя тем самым под удар: ФИО, дата рождения, номер телефона, место работы, почтовый ящик, друзья. А ведь всем этим можно воспользоваться.

–Поисковики (Google, Yandex и прочее)

При должном терпении и умении искать по ключевым словам, можно найти абсолютно любую информацию. Комбинируя эту информацию с данными социальных сетей, социальный инженер сможет добиться существенных успехов.

–Коммуникация с родственниками, друзьями и даже с самой жертвой.

Пожалуй, один из самых сложных и эффективных методов сбора информации. Ведь для этого злоумышленник должен быть очень сильно подкован, чтобы не выдать себя.

–Профилирование

Этап, на котором осуществляется оценка полученной информации. Злоумышленник должен определиться с методами взаимодействия с жертвой (телефон, почта, живое общение и др.), найти уязвимые места для манипуляции и выбрать технику проведения атаки.

–Выполнение атаки

Заключительная часть атаки. Злоумышленник на основе сведений, полученных на первом этапе, и смоделированных действий на втором, реализует атаку. В случае успеха операции, цикл из трех этапов может повторяться несколько раз. Это позволяет

совершенствовать модель и реализацию с каждым разом, получая все больше и больше информации.

Техники проведения атак

Претекстинг. Данный вид атак представляет собой набор действий, проведенный по определенному, заранее готовому сценарию (претексту). Данная техника предполагает использование голосовых средств, таких как телефон, Skype и т.п. для получения нужной информации. Как правило, представляясь третьим лицом или притворяясь, что кто-то нуждается в помощи, злоумышленник просит жертву сообщить пароль или авторизоваться на подготовленной веб-странице, тем самым заставляя цель совершить необходимое действие или предоставить определенную информацию. В большинстве случаев данная техника требует каких-либо изначальных данных об объекте атаки (например, персональных данных: даты рождения, номера телефона, номеров счетов и др.)

Quid pro quo. Данный вид атаки подразумевает звонок злоумышленника в компанию по корпоративному телефону. В большинстве случаев злоумышленник представляется сотрудником технической поддержки, опрашивающим, есть ли какие-нибудь технические проблемы. В процессе "решения" технических проблем, мошенник "заставляет" цель вводить команды, которые позволяют хакеру запустить или установить вредоносное программное обеспечение на машину пользователя.

Сбор информации из открытых источников. Использование социальной инженерии требует умения собирать о человеке необходимую информацию. Основным способом получения персональной информации стал её сбор из открытых источников, главным образом из социальных сетей. К примеру, такие сайты, как «Facebook», «VK», содержат огромное количество данных, которые люди и не пытаются скрыть. Как правило, пользователи не уделяют должного внимания вопросам безопасности, оставляя в свободном доступе данные и сведения, которые могут быть использованы злоумышленником. Даже ограничив доступ к информации на своей странице в социальной сети, пользователь не может быть точно уверен, что она никогда не попадет в руки мошенников. Например, бразильский исследователь Нельсон Новаес Нето показал, что существует возможность стать другом любого пользователя «Facebook» в течение 24 часов, используя методы социальной инженерии. В ходе эксперимента исследователь выбрал «жертву» и создал фальшивый аккаунт человека из ее окружения - ее начальника. Сначала он отправлял запросы на дружбу друзьям друзей начальника жертвы, а затем и непосредственно его друзьям. Через 7,5 часов исследователь добился добавления в друзья от «жертвы». Тем самым, исследователь получил доступ к личной информации пользователя, которой тот делился только со своими друзьями.

Дорожное яблоко. Этот метод атаки представляет собой адаптацию троянского коня, и состоит в использовании физических носителей. Злоумышленник подбрасывает "инфицированный" USB-носитель, в месте, где носитель может быть легко найден (туалет, лифт, парковка). Носитель подделывается под официальный, сопровождается подписью или снабжается корпоративным логотипом и ссылкой на официальный сайт компании. Сотрудник по незнанию может подобрать носитель и вставить его в компьютер, чтобы удовлетворить своё любопытство.

Методы защиты от атак

К сожалению, невозможно предсказать какую атаку выберет атакующий, в какой период времени, кто будет жертвой, но тем не менее возможно уменьшить успешность атаки используя нижеприведенные методы защиты:

Тестирование системы защиты - это метод выявления недостатков безопасности с точки зрения постороннего человека (злоумышленника). Используя этот метод, можно обнаружить даже те недостатки защиты, которые не были учтены в самом начале, при разработке политики безопасности. При тестировании могут быть затронуты деликатные вопросы частной жизни сотрудников и безопасности организации, поэтому желательно

получить предварительное разрешение на проведение такого мероприятия.

Профессионалам в области безопасности при проведении теста необходимо иметь такое же положение, как и у потенциального злоумышленника: в их распоряжении должны быть время, терпение и максимальное количество технических средств, которые могут быть использованы злоумышленником.

Осведомленность. Осведомленность является ключевым моментом и вследствие того, что это предварительная, предупреждающая мера, нацеленная на усвоение самими служащими основных принципов и необходимых правил защиты. Разумеется, этот аспект требует обучения и тестирования сотрудников. В рамках данной меры акцентируется внимание на следующих пунктах:

1. Привлечение внимания людей к вопросам информационной безопасности;
2. Осознание сотрудниками всей серьезности проблемы и принятие политики безопасности организации;
3. Изучение и внедрение необходимых методов и действий для повышения защиты информационного обеспечения.

Мировой опыт

Атака 2016 года, от которой пострадала столица Украины, была не первой, но несколько более заметной. В конце 2015 года для нескольких украинских областей в буквальном смысле настал конец света – из-за атаки вируса на системы энергоснабжения, без электричества на 6 часов остались 300 тысяч украинцев в Прикарпатье, Киевщине и Черновцах. Виновником стал троян BlackEnergy, который смог отключить оборудование в диспетчерских облэнерго, а также уничтожил всю информацию в их компьютерных сетях. Источником заражения стали письма с вредоносными excel-документами, полученные сотрудниками компаний. Детально изучив этот инцидент, Служба безопасности Украины пришла к выводу, что за атакой стояли российские хакеры, но позже эта информация не подтвердилась. Тем не менее "черная энергия" в 2015 году стала первым в мире подтвержденным случаем хакерского вмешательства в работу энергосистем. В декабре 2016 года история повторилась, затронув в этот раз жителей столицы. злоумышленников стала подстанция "Киевская", управляемая "Укрэнерго", Киевская гидроаккумулирующая электростанция и ряд подстанций "Киевэнерго" и "Киевоблэнерго". Длительных перебоев со светом не возникло – отключение длилось чуть больше часа, однако власти забеспокоились и поручили расследование международным специалистам по кибербезопасности. Спустя полгода, 8 июня словацкая ESET и американская Dragos опубликовали отчеты, в которых загадка прошлогоднего обесточивания немного прояснилась. Оказалось, что проблемы "Укрэнерго" вызвал вирус типа Industroyer (или же Crashoverride). Его особенность в том, что "заточен" именно под системы электроэнергетики. Программа обращается напрямую к оборудованию, вызывая его отключение либо же перегруз, и выводит энергоснабжающую систему из строя. В то же время специалисты обнаружили, что вредонос угрожает не только энергетикам – код написан таким образом, что может быть модифицирован для любой другой индустрии, будь то металлургия, машиностроение, управление инфраструктурой. Когда вирус попал в энергосистему – неизвестно, но исследователи предполагают, что проникновение произошло по сходному со случаем атаки на облэнерго годом ранее принципу: посредством рассылки зараженного письма. Впрочем, необязательно это случилось именно в день отключения – по статистике кибербезопасников, во многих компаниях между атакой и ее обнаружением проходит 10 месяцев. В то же время исследователи подчеркивают, что это может быть не последняя атака Industroyer, а всего лишь его тест-драйв.

В 2014 году в Германии жертвой атаки стал один из металлургических заводов. Используя социальную инженерию, посредством фишингового письма хакеры сумели получить доступ к компьютеру одного сотрудника, с которого они смогли проникнуть во

внутреннюю сеть системы управления. В результате этого стало невозможным выключить одну из доменных печей для плавки металла, что нанесло огромный ущерб предприятию.

Однако самой известной атакой на критическую инфраструктуру стоит считать вирус Stuxnet, который был создан для срыва ядерной программы Ирана. Это была скоординированная атака израильских и американских спецслужб. Созданный ими червь заставлял работать центрифуги на заводе по обогащению урана на полной скорости, при этом сотрудники видели у себя на экранах нормальный режим работы. Это привело к многочисленным поломкам и миллиардным убыткам.

Big-Data в кибербезопасности

–Изучение логов

Чем же дата-анализ может пригодиться на передовой кибербезопасности? В первую очередь – исследованием логов систем, где уже произошли кибератаки. Понятно, что чем больше массив данных, тем точнее вывод, позволяющий вычислить портрет хакера, возможный характер атаки и предотвратить ее. В компании-производителе сетевого оборудования и решений безопасности Cisco говорят, что атаки зловредов на системы становятся все более изощренными – раньше целью было просто "найти и вырубить", теперь преступники используют все больше комбинированных средств – трояны, шифровальщики, программы- вымогатели. Самыми уязвимыми местами в компании по-прежнему считают не столько операционные системы, сколько несвоевременно обновляемое ПО. "Часто "дыры" обнаруживаются в критичных программах типа Apache или OpenSSH, где были обнаружены 16 уязвимостей, не исправленных в среднем в течении 5 лет", – отмечают в Cisco.

–Real-time анализ

Второй – не менее эффективный способ противостояния – реагирование на угрозы в реальном времени. Руководитель (или Министр обороны, как он себя называет) компании Splunk, поставляющей ПО для дата-анализа, Монци Мерза говорит, что все чаще к ним обращаются за системами наблюдения и анализа поведения пользователей. Цель состоит в том, чтобы искать необычные шаблоны среди цифрового "шума" организации. Система постоянно обрабатывает этот шум, и если что-то выбивается из него – сигнализирует аналитику. Почему это связано с большими данными? Система должна суметь свести воедино события в больших временных масштабах и из нескольких источников. Ведь даже когда злоумышленникам удается успешно обойти традиционные средства защиты, они неизбежно оставляют небольшие следы в сети – будь то неудачная попытка входа в систему, или рост трафика с определенного адреса. Если такое поведение наблюдается несколько дней подряд – есть повод задуматься. "Изначально такой подход использовали военные, теперь мы пытаемся внедрить это в промышленных масштабах", – рассказывает Монци. Наверное, самый известный пример такой системы – Palantir, созданный ЦРУ для борьбы с терроризмом. Только изначально он искал ключевые слова и поведение, типичные для террористов, а сейчас делает то же самое, но и в части киберугроз. Сама система оценивается в $20 млрд, и это не предел. По словам Монци, такие услуги часто считаются хорошими инвестициями, поскольку предприятия, которые допускают утечки данных из-за нарушения протоколов безопасности, теряют доверие клиентов. Или, что еще хуже – может всплыть чувствительная коммерческая информация, что чревато штрафами и судами.

–Банки и перспективы

Среди передовиков дата-анализа в противостоянии кибертеррору Монци особенно выделяет даже не промышленников – банкиров. Например, до того, как дата стала мейнстримом, в августе 2011 года Visa представила новый аналитический движок для борьбы с мошенничеством. Вместо того, чтобы анализировать лишь 2% трансакций, компания загрузила в систему все свои данные, и выявила потенциальные возможности ежегодного мошенничества с фишингом карт и взломом банкинга в размере $2 млрд. Да, это всего 0,06%

от общего оборота, но даже на таком уровне это недопустимо. И компания отмечает, что к 2017 году в 98% случаев системе удается определить мошенническую трансакцию и заблокировать ее еще до момента потери денег. Возможность дата-аналитики сопоставлять данные из широкого диапазона источников за длительный период времени называют возможным выходом и в пострадавшей от Industroyer "Укрэнерго". Понятно, что обработки в режиме реального времени достичь будет крайне сложно, но временное окно хотя бы в несколько часов сможет заранее предупредить о возможных атаках наинфраструктуру и обнаружить зловред вовремя, а не спустя 10месяцев, о которых говорят аналитики. Показательно, что направление становится актуальным для крупных компаний. Только в работу с данными для прогнозирования кибератак в ближайшие годы вложат $1 трлн во всем мире. Это недешевое решение, доступное большим игрокам рынка. Но результаты этих исследований доступны и малому/среднему бизнесу, а также обычному населению. Многие компании безвозмездно делятся результатами таких исследований с антивирусными компаниями. Так, доступ к прогрессивной вирусной дата-аналитике имеют в Avira, Avast, ESET, Microsoft и многие другие разработчики антивирусного ПО. Именно поэтому специалисты по кибербезопасности рекомендуют в первую очередь устанавливать проверенный антивирус с доступом к вирусной онлайн-базе. В ней есть частичка той "бигдаты", которая сможет защитить инфраструктуру.

Применение Big-Data в других областях

Опрос Tech Pro Research показал, что самое широкое применение технологии больших данных нашли в телекоммуникационной сфере, а также в инжиниринге, в страховании и финансах. Более подробно результаты опроса представлены на графике.

Рисунок *2*. – Результаты опроса Tech Pro Research [6]

Отечественный бизнес также начинает свою работу с большими данными, но в темпах роста значительно отстает от зарубежных коллег. По данным CNews, лишь 10 % компаний внедрили большие данные в свою деятельность, в то же время за рубежом уже 30 % организаций сделали это.

Заключение. Главное изменение за последние несколько лет в том, что данных о пользователе стало критически много, и на основе них можно делать некоторое довольно точное прогнозирование поведения людей в зависимости от некоторых внешних стимулов. Каждый пользователь оставляет о себе все цифровые следы, которые только возможно (как часто он пользуется телефоном, история его браузера, видео, которые он просматривает и др.).

Все эти данные могут быть использованы злоумышленниками в их целях, что говорит о необходимости каждого пользователя изучать и внедрять методы для повышения защиты информационного обеспечения.

Список литературы

[1.] Dumbill, E. What is big data? [Electronic resource]. – 2012. – Mode of access: http://radar.oreilly.com/2012/01/what-is-big-data.html.

[2.] Отчет AAPOR о больших данных: 12 февраля 2015 / Л. Джапек [и др.]; Американская ассоциация исследователей общественного мнения; пер. с англ. Д. Рогозина, А. Ипатовой, Е. Вьюговской. – Москва, 2015. – Режим доступа:

[3.] http://www.ibmbigdatahub.com/infographic/four-vs-big-data

[4.] The FOUR V's of Big Data [Electronic resource]. – 2015. – Mode of access: http://www.ibmbigdatahub.com/infographic/four-vs-big-data.– Date of access: 30.02015.

[5.] Hilbert, M. (2016). Big Data for Development: A Review of Promises and Challenges. Development Policy Review, 34(1), 135–17

[6.] Davy Cielen, Arno D. B. Meysman, and Mohamed Ali. Introducing Data Science. Big data, machine learning, and more, using Python tools (https://www.manning.com/books/introducing-data-science)

[7.] Tech Pro Research [Электронный ресурс]. – Режим доступа: http://techproresearch.com/topic/big-data.

SOCIAL ENGINEERING AND TECHNOLOGY BIG DATA

D.M.Rumiantsau

The military faculty of Belarusian state University, lecturer

Belarusian state University, Republic of Belarus
E-mail: dimon-rum@mail.ru

Abstract. "Big data" is becoming a new topic of discussion in the social Sciences. Since 2012, many people define this concept through such features as volume, velocity, and variety [1; 2, p.18]. Commercial companies that provide "big data" collection programs also emphasize the accuracy (veracity) of the information they receive about user behavior, compared to user surveys about opinions and behavior [3]. Social engineering is a set of techniques, methods and technologies for creating such a space, conditions and circumstances that most effectively lead to a specific desired result, using sociology and psychology. Social engineering is often seen as an illegal method of obtaining information, but this is not entirely true. Of course, today social engineering is often used to get private information on the Internet. However, if we consider modern professional social engineering, the scope of its application is quite legitimate (for example, it helps to achieve an initially unattainable result, or " program" for performing useful actions of a particular person or group of people. Thus, a new direction of cooperation between computer and social Sciences is emerging, related to the analysis and use of the results of big data analysis. This same big data can be a source of profit for hackers. In this paper, we will try to analyze the impact of the development of Big Data on the security of users on the Internet.

Keywords: Big Data, social engineering, information security.