*A.P. Mikhalkova, A.S. Zaytsev*

**Baysean Approach Appliance for Early Detection of Insider Threat**

*Keywords: insider, Bayesian network.*

Purpose of report is to increase the information security level of organization from insider threats. This report presents the method of detection and determinationof the type of supposed insider threatwith the help of Bayesian approach.

*А.П. Михалькова, А.С. Зайцев*

# ПРИМЕНЕНИЕ БАЙЕСОВСКОГО ПОДХОДА ДЛЯ РАННЕГО ОБНАРУЖЕНИЯ ВНУТРЕННИХ НАРУШИТЕЛЕЙ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ

В настоящее время существует множество исследований, посвященных противодействию внешним угрозам информационной безопасности (ИБ), разработаны полноценные методологии и специализированное программное обеспечение. Вопрос противодействия внутренним угрозам ИБ проработан недостаточно хорошо, но в то же время 77,6% руководителей по информационным технологиям (ИТ) и ИБ заявляют, что основная опасность для организаций связанна с внутренними угрозами, а именно с утечкой информации ограниченного доступа, нелояльным или преступным поведением сотрудников и пр. [1]

Основой успешного противодействия внутреннему нарушителю (ВН) является его раннее обнаружение. События и факты, знание которых позволяет с некоторой долей вероятности утверждать, что конкретный сотрудник реализует или может реализовать ту или иную угрозу ИБ, называются *индикаторами угрозы*. В зависимости от способа получения индикаторы делятся на *поведенческие* и *технические*: поведенческие отражают социальное поведение потенциального инсайдера, технические – действия в информационной системе (ИС) организации.

В случае мотивированного инсайдерского нарушения поведенческие индикаторы, как правило, предшествуют индикаторам техническим [10]. Для угрозы ИТ-саботажа сначала появляется недовольство сотрудника, которое впоследствии находит выход в совершении диверсии. Для угроз ИТ-мошенничества и ИТ-шпионажа совершению преступления предшествует период подготовки и рационализации, что не может не отразиться на поведении инсайдера. Поведенческие индикаторы в меньшей степени применимы к угрозе кражи интеллектуальной собственности (ИСб), т.к. в части случаев кражи ИСб инсайдер не считает свои действия нарушением, поэтому его поведение не меняется. По тем же причинам поведенческие индикаторы практически не применимы к немотивированным нарушителям ИБ. Немотивированного нарушителя также тяжело обнаружить и по техническим индикаторам, поскольку большая часть угроз для данной модели реализуется без предварительной подготовки и непреднамеренно.

Задачей выявления внутренних нарушителей ИБ является сбор поведенческих и технических индикаторов угроз и оценка вероятности того, что конкретный сотрудник реализует ту или иную угрозу ИБ (при условии, что присутствие некоторых индикаторов не удается определить однозначно). При превышении определенного порога вероятности реализации угрозы сотрудник переходит в перечень потенциальных инсайдеров, для которых проводится выбор применимых контрмер. Необходимо отметить, что

важно определить оптимальный пороговый уровень вероятности реализации угрозы, при котором будет достигнут минимум ошибок первого и второго рода.

Таким образом, опираясь на существующие исследования, все модели угроз, кроме немотивированного нарушителя ИБ, поддаются раннему обнаружению. Для составления математической модели раннего обнаружения немотивированного нарушителя ИБ требуются более глубокие исследования его поведения и возможностей проявления им технических индикаторов.

## Возможности применения сетей Байеса для выявления потенциальных внутренних нарушителей

Сеть Байеса представляет собой направленный ациклический граф, каждой вершине которого соответствует случайная переменная. Если узлы (переменные) не соединены дугами, то их считают условно независимыми [6]. Если из вершины $A$ выходит дуга в вершину $B$, то вершину $A$ называют родителем вершины $B$, а вершину$B$ – потомком вершины $A$. Множество вершин-родителей вершины $v_i$ обозначим *parents*($v_i$) [3]. Соответственно, если $V$ – множество всех вершин, а $v_i$ – значение $i$-й вершины, то полное совместное распределение вероятности можно записать следующей формулой:

$$P(v_1, \dots, v_n) = \prod_{i=1}^{n} P(v_i | \, parents(v_i)). \quad (1)$$

С математической точки зрения сеть Байеса – это модель для представления вероятностных зависимостей, а также отсутствия этих зависимостей. При этом связь $A \rightarrow B$ является причинной, когда событие $A$ является причиной возникновения $B$, т.е. влияет на значение, принятое $B$ [6].

Чтобы определить вероятность принадлежности сотрудника к тому или иному классу внутренних нарушителей, для каждого класса реализована сеть Байеса, применение которой имеет следующие преимущества [5]:

- простота построения и интерпретации;
- работа с заведомо неточными и неполными данными;
- обучение в процессе работы с низкими вычислительными затратами.

Входами каждой сети являются индикаторы, проявляемые потенциальным внутренним нарушителем ИБ. Выходом каждой сети является вероятность принадлежности сотрудника к конкретному классу внутренних нарушителей ИБ.

Иначе говоря, каждая вершина представляет собой некоторую случайную величину, которая может принимать два значения: единица, если индикатор наблюдался, ноль – в противном случае. Дуги сети Байеса, в свою очередь, представляют собой вероятностные зависимости величин, задаваемые с помощью таблицы условной вероятности. Значения таблицы условной вероятности для каждой вершины определяются по формуле (1).

## Обучение сетей Байеса в условиях неполноты и неточности статистических данных

Под обучением сети понимается её регулярный пересмотр и переоценка с учетом получаемых выборочных данных об исследуемом явлении или процессе.

Существует множество алгоритмов, на основании которых производится обучение сети Байеса. Среди них обычно выделяют такие методы, как метод релевантных

векторов, EM-алгоритм, методы Монте-Карло и т.п. Входами каждой сети являются проявляемые нарушителями индикаторы. Но в большинстве случаев наблюдателю затруднительно судить о проявлении части индикаторов. Таким образом, в модели появляются скрытые данные. Поэтому при моделировании ВН ИБ самым подходящим для обучения является EM-алгоритм, так как он имеет следующие преимущества:

1) линейное увеличение сложности при росте объёма данных [9];
2) устойчивость к шумам [9];
3) возможность работы со скрытыми данными;
4) множество реализаций на различных языках программирования.

EM-алгоритм – алгоритм, используемый для нахождения оценки максимального правдоподобия параметров вероятностных моделей, в случае, когда некоторые переменные не наблюдались [7].Алгоритм состоит из двух шагов:

1) E-шаг (expectation step);
2) M-шаг (maximization step).

Пусть мы имеем набор данных $Y$. Из этих данных $X$ наблюдались, а $Z$ – нет, т.е. $Z$ – скрытые данные. Следовательно, $Y = X \cup Zh$ – это значение скрытой переменной. Перед началом работы алгоритма ей задаётся некоторое предполагаемое начальное значение.

На E-шаге вычисляется условное математическое ожидание $Q(h)$ логарифма правдоподобия набора переменных от параметра $h$:

$$Q(h) = E[\ln p(Y\,|h)|X). \quad (2)$$

Фактически, на этом шагеопределяется ожидаемое значение всех переменных по текущему значению параметра $h$, а на M-шаге вычисляется максимизирующее условное матожидание $Q(h)$:

$$h_1 = argmax_h Q(h) \quad (3)$$

Т.е. находится следующее приближение параметра $h$ при значении $Q(h)$, полученном на E – шаге. Эти шаги выполняются до тех пор, пока последовательность $h_k$ не будет сходиться.

Таким образом, с помощью EM-алгоритма находится новая оценка максимального правдоподобия параметров модели на основе выборочных данных. С помощью полученной оценки можно провести пересмотр сети и её априорных вероятностей. Следовательно, чем больше инцидентов используется для обучения сети, тем точнее она сможет определить ВНИБ.

## Статистическая база и выделение основных индикаторов

Для обучения сети разработана статистическая база, состоящая из более 100 различных инсайдерских инцидентов. Все инциденты взяты из открытых источников, база постоянно пополняется.

Как было отмечено ранее, не все типы ВН ИБ поддаются раннему обнаружению, поэтому на данный момент целесообразно рассматривать следующие модели ВН ИБ:

1) ИТ-саботаж;
2) мошенничество на руководящей должности;
3) мошенничество на неруководящей должности;
4) шпионаж;
5) кража ИСб.

Для каждого типа ВН ИБ на основе экспертного мнения, различных исследований ВН [2,4,8] и выборки инцидентов были выделены технические и поведенческие индикаторы. Ниже приведены индикаторы для одного из типов ВН ИБ:

***Мошенничество на руководящей должности***

- *Поведенческие индикаторы:*
  - финансовые проблемы;
  - неожиданный доход;
  - состояние стресса без видимых на то причин.
- *Технические:*
  - подозрительные транзакции;
  - мошеннические операции;
  - финансовое несоответствие;
  - фальсификация документов.

## Пример построения сети Байеса для выявления ВН ИБ

Общие принципы построения и функционирования сети Байеса являются одинаковыми для всех типов ВН. Поэтому в данной статье рассматривается пример построения байесовской сети доверия только для угрозы инсайдерского мошенничества на руководящей должности.

В первую очередь необходимо выделить основные поведенческие и технические индикаторы, что проведено в разделе 3 данной статьи. Обозначим поведенческие индикаторы $BIk$, а технические – $TIj$, где $k, j$ –порядковые номера индикаторов. Далее необходимо построить графическую модель, т.к. её проще интерпретировать (рис. 1).

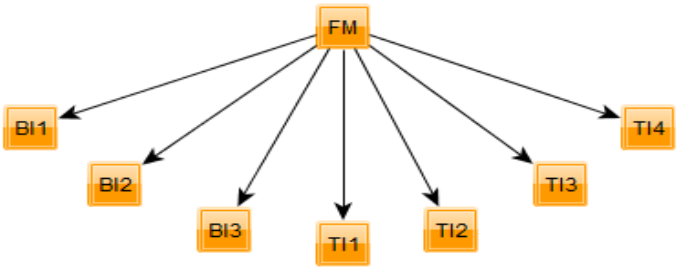


*Рис. 1. Сеть Байеса для угрозы мошенничества на руководящей должности*

На рис. 1 вершине $FM$ соответствует априорное значение вероятности принадлежности сотрудника к данному типу ВН ИБ без учета индикаторов $P(FM)$, вершинам $BI1 - BI3$ – значение вероятностей $P(BIk \mid FM)$, $TI1 - TI4$ – значения вероятностей $P(TIj \mid FM)$, где $k, j$ – порядковые номера индикаторов в списке пункта 3 данной статьи.

Рис. 1 наглядно показывает, что индикаторы и выбор модели условно зависят друг от друга, но между собой индикаторы независимы. Такое построение сети называется простым дивергентным соединением. Предполагается, что все индикаторы при таком способе построения условно независимы между собой, но в реальности это не так. Это не сильно влияет на точность выявление потенциального ВН ИБ, так как аналогичным построением и допущением обладает другая модель – наивный байесовский классификатор, который, несмотря на корреляцию классов, позволяет достаточно точно решать поставленную задачу [11].

Следующим этапом построения сети Байеса для выявления мошенничества на руководящей должности является заполнение таблицы априорных вероятностей параметров сети на основе экспертного мнения. После этого производится обучение сети с корректировкой значений вероятностей.

Для того чтобы начать работу с сетью Байеса, необходимо опросить наблюдателя о проявленных предполагаемым нарушителем индикаторах. Необходимо отметить, что для получения информации о технических индикаторах можно использовать программные комплексы мониторинга ИБ и противодействия мошенничеству. Поведенческие индикаторы можно определить посредством наблюдения. Предположим, что наблюдались только финансовые проблемы и подозрительные транзакции. Значит, значения переменных :

$$BI1 = TI1 = True$$
$$BI2 = BI3 = TI2 = TI3 = TI4 = False$$

Тогда вероятность принадлежности сотрудника к ВН ИБ данного типа вычисляется по следующей формуле:

$$P\left(FM \left| BI1, \ \dots, BI3, \ TI1, \dots, \ TI4\right.\right) =$$
$$= \frac{P\left(BI1 \left| FM\right.\right) * \dots * P(BI3|FM) * P\left(TI1 \left| FM\right.\right) * \ \dots * P(TI4 \left| \ FM\right.) * P\ (FM)}{\prod_{i,j=1}^{i=3,j=4} P(BIi|FM) * P(TIj|FM) * P(FM) + \prod_{i,j=1}^{i=3,j=4} P(BIi|\overline{FM}) * P(TIj|\overline{FM}) * P(\overline{FM})}.$$

**Заключение**

В данной статье представлен метод раннего обнаружения ВН ИБ с применением сетей Байеса. На основании экспертного мнения и других исследований разработана классификация ВН ИБ, выделены основные поведенческие и технические индикаторы для каждого типа ВН. Разработана обучаемая математическая модель на основе байесовской сети доверия для обнаружения ВН. В рамках дальнейших исследований необходимо реализовать метод с применением современных инструментов разработки и моделирования, а также провести анализ других математических методов, подходящих для решения поставленной задачи, и оценку результатов их применения.

## СПИСОК ЛИТЕРАТУРЫ:


1. Безопасность информации в корпоративных информационных системах. Внутреннииеугрозы // Infowatch. 2013. с. 8.
2. Greitzer F.L., Kangas L.J., Noonan C.F. Identifying At-RiskEmployees: A Behavioral Model for Predicting Insider Threats // Richland, WA: Pacific Northwest National Laboratory. 2010. P. 8-39.
3. Maxwell Chickering D., Heckerman D.,Meek C. Bayesian Approach to Learning Bayesian Networks with Local Structure // Microsoft Research Redmond WA. 1998.p. 2.
4. Зайцев А.С., Малюк А.А. Исследование проблемы внутреннего нарушителя // Вестник РГГУ. №14. 2012. с. 117-133.
5. Атаманов А. Н. Вопросы оценки рисков информационной безопасности в автоматизированных системах // Современная наука: Актуальные проблемы теории и практики.2012. URL: http://www.nauteh-journal.ru/index.php/ru/----etn12-02/371-a (дата обращения 11.03.2015).
6. Бидюк П.И., Терентьев А.Н. Построение и методы обучения Байесовских сетей // Таврический вестник информатики и математики. №2. 2004. с. 1-3.
7. Neal R.M., Hilton G.E. A view of the EM algorithm that justifies incremental, sparse, and other variants // University of Toronto, 1999. P. 1-2.
8. Greitzer F.L., Paulson P.R., Kangas L.J., Franklin L.R., Edgar T.W., Frincke D.A. Predictive Modeling for Insider Threat Mitigation // Richland, WA: Pacific Northwest National Laboratory. 2009, p. 6-17.
9. EM – масштабируемый алгоритм кластеризации // BaseGroupLabs. URL: http://www.basegroup.ru/library/analysis/clusterization/em/ (дата обращения 12.03.2015).

10. Moore A.P., Mundie D.A., Collins M.L. A System Dynamics Model for Investigating Early Detection of Insider Threat Risk // Software Engineering Institute. CERT Program.2013. p. 4.
11. Субботин С. В., Большаков Д. Ю. Применение байесовского классификатора для распознавания классов целей // Журнал Радиоэлектроники. 2006.URL: http://jre.cplire.ru/iso/oct06/2/text.html (дата обращения 12.03.2015).

## REFERENCES:

1. Bezopasnostinformatsyivkorporatvnyhinformatsionnyhsistemah. Vnutrennieugrozy // Infowatch. 2013. P. 8.
2. Greitzer F.L., Kangas L.J., Noonan C.F. Identifying At-RiskEmployees: A Behavioral Model for Predicting Insider Threats // Richland, WA: Pacific Northwest National Laboratory. 2010. P. 8-39.
3. Maxwell Chickering D., Heckerman D., Meek C. Bayesian Approach to Learning Bayesian Networks with Local Structure // Microsoft Research Redmond WA. 1998. P. 2.
4. Zaytsev A.S., Malyuk A.A. Issledovanie problemy vnutrennego narushitelya/ Vestnik RGGU.№14.2012. p. 117-133.
5. AtamanovA. N. Voprosy otsenki riskov informatsionnoy bezopasnosti v avtomatizirovannyh sistemah // Sovremennayanauka: Aktualnyeproblemteoriiipraktiki. 2012. URL: http://www.nauteh-journal.ru/index.php/ru/----etn12-02/371-a (data obrascheniya 11.03.2015).
6. Bidyuk P.I., Terentyev A.N. Postroenie i metody obucheniya Bayesovskih setey // Tavricheskiy vestnik informatiki i matematiki, №2, 2004. P. 1-3.
7. Neal R.M., Hilton G.E. A view of the EM algorithm that justifies incremental, sparse, and other variants // University of Toronto. 1999. P. 1-2.
8. Greitzer F.L., Paulson P.R., Kangas L.J., Franklin L.R., Edgar T.W., Frincke D.A. Predictive Modeling for Insider Threat Mitigation. / Richland, WA: Pacific Northwest National Laboratory. 2009., p. 6-17.
9. EM – masshtabiruemyi algoritm klasterizatsyi // BaseGroupLabs. URL: http://www.basegroup.ru/library/analysis/clusterization/em/ (dataobrascheniya 12.03.2015).
10. Moore A.P., Mundie D.A., Collins M.L. A System Dynamics Model for Investigating Early Detection of Insider Threat Risk // Software Engineering Institute. CERT Program.2013. P. 4.
11. SubbotinS. V., BolshakovD. Y. Primenenie baesovskogo klassifikatora dlya raspoznovaniya klassovtseley // Jurnal Radioelektroniki. 2006. URL: http://jre.cplire.ru/iso/oct06/2/text.html (data obrascheniya 12.03.2015).