# ИНФОРМАЦИОННО-УПРАВЛЯЮЩИЕ КОМПЛЕКСЫ И СИСТЕМЫ

УДК 004.9

## ИСПОЛЬЗОВАНИЕ ПОЛОЖЕНИЙ ТЕОРИИ ОПАСНОСТИ В ИСКУССТВЕННЫХ ИММУННЫХ СИСТЕМАХ

Бардачев Ю.Н., Дидык А.А.

**Введение.** В последние годы компьютерные системы становятся все более и более сложными, и, следовательно, проблема защиты этих систем становится все более и более трудно разрешимой. Для противодействия злоумышленному использованию компьютерных систем были разработаны различные методики, такие как межсетевые экраны (firewalls), антивирусные программы и системы обнаружения вторжений. Сложность сетей и динамическая природа компьютерных систем оставляет большое поле деятельности для усовершенствования современных методов обеспечения компьютерной безопасности.

С недавнего времени ученые стали черпать идеи из механизмов биологических систем и, в контексте компьютерной безопасности, сосредоточили свое внимание на человеческой иммунной системе (ЧИС). Человеческая иммунная система дает пример устойчивой, распределенной системы, которая обеспечивает высокий уровень защиты от постоянных атак. Исследуя механизмы человеческой иммунной системы, можно надеяться, что данная парадигма улучшит рабочие характеристики существующих систем обнаружения вторжений.

**Целью работы** является рассмотреть возможности применения новой иммунологической парадигмы – теории опасности – для разработки искусственных иммунных систем.

**Изложение основного материала.** Начиная с 1959 года, основной постулат иммунологии утверждал, что человеческая иммунная система воздействует на объекты, которые не являются частью человеческого организма. Поэтому реакция организма на такие объекты является результатом классификации ЧИС собственных клеток как *своих* и всех остальных клеток как *чужих* [1]. ЧИС выполняет классификацию, распознавая протеины, найденные на поверхности инородных клеток (известных как антигены). Инородные клетки по структуре и форме отличаются от клеток, существующих в организме человека.

Однако, существует множество примеров, когда такой подход терпит неудачу. Например, кишечный тракт подвергнут воздействию пищи и множества различных бактерий, которые не определены как «свои» и не инициируют иммунный ответ. Кроме того, модель «свой-чужой» не может объяснить явления аутоиммунных болезней. На пример, в случае рассеянного склероза, ЧИС подвергает атакам определенные клетки, которые классифицированы как «свои». В 1994, Полли Матзингер [2] выдвинула предположение, что в этом случае ЧИС не воздействует на «своих» или «чужих», а использует защитный механизм *распознавания опасности*. Метод обнаружения опасности является основой Теории Опасности.

Теория опасности не отрицает существования разграничения на «свой-чужой», а скорее определяет, что существуют другие факторы, приводящие к инициированию иммунного ответа. Полагается, что ЧИС отвечает на определенные сигналы об опасности, произведенные в результате клеточного *некроза* (неожиданного стресса и/или смерти клетки).

Смерть клетки - естественный процесс, который происходит внутри человеческого организма в результате гомеостатической стабилизации. Этот процесс является

результатом работы предварительно запрограммированного и строго управляемого механизма, известного как *апоптоз*. Теория опасности предполагает, что в результате смерти клетки происходят различные биохимические реакции, которые в свою очередь вызывают различные сигналы опасности. Полагается, что эти сигналы могут способствовать возникновению иммунного ответа. Эта теория, вызывающая споры в рядах иммунологов, предлагает потенциальное объяснение многих явлений, где модель «свой-чужой» терпит неудачу.

Искусственные иммунные системы основаны на модели человеческой иммунной системы «свой-чужой». Алгоритмы, разработанные на основании данной модели, показали в значительной мере свою эффективность [3]. Искусственные иммунные системы были разработаны для широкого диапазона приложений от переработки данных до информационной безопасности. Во многих случаях, эти приложения дали результаты сопоставимые со стандартными методиками или даже превосходящие их.

Например, отрицательный отбор иммунных клеток в тимусе для распознавания «свой-чужой» был применен в системе Lisys и использовался как сетевой инструмент обнаружения вторжений [4]. Эта система классифицировала нормальное поведение пользователя как «свой», и любое другое поведение как «чужой».

Однако, данный подход оказался не способным к масштабированию, как ожидалось, для использования в больших, динамических средах. Одним из объяснений такой несостоятельности данного подхода может быть то, что не все процессы, необходимые для полной функциональности иммунной системы, были включены в данную модель. Существует также ряд других факторов, объясняющих это:

- механизм отрицательного отбора несовершенен; поэтому аутореактивные реакции системы (ошибочные положительные срабатывания) неизбежны.
- граница между «своими» и «чужими» размыта, так как зачастую «свои» и «чужие» антигены совместно используют общие области.
- «свои» претерпевают изменения с течением времени. Поэтому, возможны проблемы с клетками памяти, которые позже могут оказаться неточными или даже аутореактивными.

Используя механизмы теории опасности при разработке искусственных иммунных систем, желательно учитывать следующие факторы:

- в модели, основанной на теории опасности, необходимо наличие антиген-презентирующей клетки, которая может представить соответствующий сигнал опасности.
- «опасность» – эмоциональный термин. Сигнал может не иметь никакого отношения к опасности.
- соответствующий сигнал опасности может быть позитивным (наличие сигнала) или негативным (его отсутствие).
- в биологических системах опасная зона является пространственной. В искусственных иммунных системах может использоваться другая мера близости (например, временная).
- инициализация аналога иммунного ответа в искусственной иммунной системе не должна вести к дальнейшему возникновению новых сигналов опасности. В естественных системах клетки-килеры вызывают нормальную смерть клетоки (апоптоз), а не опасность (некроз).
- для большего эффекта Матзингер предлагает примировать клетки-килеры через антиген-презентирующие клетки. В зависимости от используемой иммунной системы (это имеет смысл только для пространственно распределенных моделей) это предложение может иметь смысл.
- есть множество факторов, которые в меньшей степени относятся непосредственно к теории опасности. Например, миграция – сколько антител получает сигнал от данной антиген-презентирующей клетки.

**Использование теории опасности в задачах обнаружения вторжений.** Системы

 ISBN 7-776-8361-7

обнаружения вторжений (СОВ) разрабатываются для того, чтобы обнаруживать события, происходящие в компьютерной системе, которые могут поставить под угрозу ее целостность или конфиденциальность [5]. СОВ часто подразделяется на две категории: *обнаружение злоумышленного поведения* и *обнаружение аномального поведения*. Методики обнаружения злоумышленного поведения заключаются в описании атаки в виде шаблона или сигнатуры и поиске данного шаблона в контролируемом пространстве (например, сетевом трафике или деятельности системы). Эта методика эффективна при обнаружении уже известных атак. Однако, такой подход не годится для обнаружения новых, еще неизвестных атак.

Системы обнаружения аномального поведения основаны на модели поведения пользователя, которое рассматривается как «нормальное». Это достигается путем использования комбинации статистических методов и методов машинного обучения для исследования сетевого трафика или системных вызовов и процессов. Обнаружение новых атак с использованием подхода обнаружения аномалий является более эффективным, поскольку любое поведение, не определенное как «нормальное», классифицируется как атака или вторжение. Однако, «нормальное» поведение в больших динамических системах сложно полностью определить, и оно изменяется с течением времени. Это часто приводит к появлению значительного количества ложных тревог, известных как *false positives*. Снижение количества ложных срабатываний СОВ является ключевой проблемой, решение которой в состоянии предложить теория опасности.

Предполагается, что использование теории опасности для разработки методов обнаружения вторжений позволит создавать системы, способные эффективно реагировать как на известные угрозы информационной безопасности, так и на новые атаки, а также позволить снизить количество ложных тревог, столь обычных для систем обнаружения аномального поведения [6]. Теория опасности предполагает, что ЧИС обнаруживает сигналы опасности и продуцирует ответ, основанный на сопоставлении этих сигналов. Подобная концепция может использоваться в системах обнаружения вторжений. Это позволит реализовать систему, способную разделять поведение на *апоптическое* и *некротическое*. Апоптическое поведение может быть определено как низкоуровневые, шумовые тревоги, которые самостоятельно не формируют какого-либо значимого аномального поведения, но зачастую являются предпосылкой для атаки. Некротические тревоги могут быть результатом более серьезных атак, в результате которых имеют место существенные повреждения системы [7]. Другие виды сигналов опасности, относящиеся к самой физической системе, могут быть также использованы в этой модели. Потенциал для дальнейших разработок в этой области и успешное сопоставление таких тревог позволит значительно повысить эффективность как систем обнаружения вторжений, так и искусственных иммунных систем в целом.

**Теория опасности в задачах переработки данных.** На первый взгляд, не ясно как теория опасности может быть полезной для решения задачи переработки данных, описанной в [8], так как в этом случае не используются понятия «свой» и «чужой». В сущности, в задаче переработки данных вся система является «своей». Однако при более пристальном рассмотрении, выделение «своих» и «чужих» в этом случае не является проблемой.

Например, метки «свой» и «чужой» могут быть заменены на «интересные» и «неинтересные» данные. В этом случае, искусственная иммунная система применяется как классификатор. Далее, если далее предположить, что интересные данные расположены «близко» или "рядом" с другими интересными данным, то возможно использовать механизм теории опасности. Для этого необходимо определить «близко»/"рядом". Можно использовать:

- физическую близость, например расстояние в базе данных, полученное с помощью соответствующей метрики.
- корреляцию данных, измеренную статистическими методами.

- подобие времени ввода данных в базу данных.
- размер файла.

Сигнал опасности, таким образом, может быть интерпретирован как полученная порция ценная (интересной) информации. Следовательно, стимуляции подвергаются те антитела, которые соответствуют данным, которым «близка» эта ценная информация.

Принимая эту идею, можно определить сигнал опасности как идентификатор интереса пользователя. Учитывая это определение, можно представлять различные сценарии, в которых может быть полезен сигнал опасности.

Возьмем для примера пользователя, просматривающего ряд документов. Каждый документ обладает некоторыми свойствами (например, ключевые слова, заголовок, автор, дата и т.д.). Предположим, что есть искусственная иммунная система, функционирующая как «наблюдатель», антитела которой соответствуют свойствам документа. Таким образом, «интересными» документами являются те, свойства которых соответствуют антителам иммунной системы.

Когда пользователь либо явно, либо неявно проявляет интерес к текущему документу, возникает сигнал «опасности».

Стимулируемые антитела становятся исполнительными эффекторами, и таким образом иммунная система обучается и выполняет фильтрацию, т.е. ищет другие интересные документы. Интересные документы могут быть представлены вниманию пользователя. Важно то, что понимание пользователем «интересности» документа может со временем измениться и иммунная система может своевременно адаптироваться к таким изменениям.

Этот пример является иллюстрацией того, как механизмы теории опасности могут быть использованы для разработки искусственных иммунных систем, используемых для решения задач, в которых значение понятия «опасность» не является очевидным.

**Выводы.** Теория опасности не предоставляет новых механизмов представления данных для искусственных иммунных систем. Данная теория определяет, какие данные в искусственных иммунных системах должны быть представлены и обрабатываться: внимание должно быть сфокусировано на опасных, т.е. интересных данных. Проблема заключается в выборе соответствующего сигнала опасности. Кроме того, мера физического расстояния в биологической системе должна быть интерпретирована как адекватная мера подобия или причинно-следственной связи в искусственной иммунной системе. В данной статье был приведен пример, как можно решить эту проблему в задаче переработки данных. Однако, не смотря на перечисленные трудности, применение механизмов теории опасности для построения искусственных иммунных систем в различных областях знаний позволит значительно повысить их эффективность.

 ISBN 7-776-8361-7

Over the last decade, a new idea challenging the classical self- non-self viewpoint has become popular among immunologists. It is called the Danger Theory. Possibilities of application of new approach for development of artificial immune systems used in the field of information security and data mining are considered in this paper. Substantive states of the danger theory are short given.

1. Aickelin U., Bentley P., Cayzer S., Kim J. and McLeod J., 2003, 'Danger Theory: The Link between AIS and IDS?', in Proceedings ICARIS-2003, 2nd International Conference on Artificial Immune Systems, 147-155.

2. De Castro, L.N. and Timmis, J., 2002, 'Artificial Immune Systems: A New Computational Approach, Springer-Verlag, London. UK.

3. Hofmeyr S. and Forrest S., 2000, 'Architecture for an Artificial Immune System', Evolutionary Computation, 8,(4), 443-473.

4. Matzinger P., 2002, 'The Danger Model: A Renewed Sense of Self', Science, 296, 301-305.

5. Medzhitov R. and Janeway C, 2000, 'How does the immune system distinguish self from nonself?', Seminars in Immunology, 12, 185-188.

6. Twycross J., 2004, 'Immune Systems, Danger Theory and Intrusion Detection', AISB 2004 Symposium on Immune System and Cognition (ImmCog-04) , Leeds, U.K.

7. Venter H. and Eloff J., 2003, 'A Taxonomy for Information Security Technologies', Computers & Security, 22, (4), 299-307.

8. Cayzer S., Aickelin U. A Recommender System based on the Immune Network, Proceedings of the 2002 Congress on Evolutionary Computation, 2002.