УДК 004.056 doi: 10.15827/0236-235X.142.054-061 2024. Т. 37. № 1. С. 54–61

Моделирование информационных процессов систем управления большими данными для решения задач кибербезопасности

М.А. Полтавцева [1]✉, Д.П. Зегжда [1]

[1] Институт кибербезопасности и защиты информации,
СПбПУ Петра Великого,
г. Санкт-Петербург, 195251, Россия

Ссылка для цитирования
Полтавцева М.А., Зегжда Д.П. Моделирование информационных процессов систем управления большими данными для решения задач кибербезопасности // Программные продукты и системы. 2024. Т. 37. № 1. С. 54–61. doi: 10.15827/0236-235X.142.054-061
Информация о статье
Группа специальностей ВАК: 2.3.6
Поступила в редакцию: 23.08.2023 После доработки: 08.09.2023 Принята к публикации: 11.09.2023

Аннотация. Несовершенство классических моделей безопасности при их приложении к реальным системам обусловило развитие обратного подхода: моделирование систем различного класса для последующего дополнения их атрибутами безопасности. Решение задач обеспечения защищенности распределенных систем на основе таких моделей является сегодня динамически развивающейся областью научного знания. Данная статья посвящена моделированию гетерогенных систем управления большими данными для решения задач кибербезопасности. Авторы выделяют и учитывают такие ключевые особенности рассматриваемого класса систем, как использование гетерогенных структур данных и ограничение инструментов манипулирования данными прежде всего в отношении гранулярности функций безопасности при их реализации. Предложена новая графовая модель системы управления большими данными с использованием обобщенных операций над ними: объединение, разделение и преобразование. Вершины графа представляют собой структурированные фрагменты данных, а дуги – операции по их обработке вне зависимости от конкретного инструмента и типа преобразования. В отличие от аналогичных решений модель за счет обобщенных операций позволяет учесть преобразования данных внутри инструментов обработки, а также при передаче информации между ними, обеспечивая комплексное представление процесса обработки информации на уровне инженерии данных. Особенностью модели является и высокая степень возможности автоматизации ее построения на базе конкретной системы больших данных, что способствует поддержанию адекватности при эволюционных изменениях объекта моделирования. Представленная модель способствует решению широкого круга задач в области безопасности крупномасштабных гетерогенных систем управления большими данными, таких как контроль доступа, аудит, оценка защищенности. В качестве примера в работе показано использование предложенной модели для автоматизации анализа политик безопасности в данном классе систем.

Ключевые слова: большие данные, системы управления данными, СУБД, моделирование, моделирование данных, защита информации, кибербезопасность

Благодарности. Исследование выполнено за счет гранта РНФ № 23-11-20003, https://rscf.ru/project/23-11-20003/ Грант Санкт-Петербургского научного фонда (Соглашение №23-11-20003 о предоставлении регионального гранта)

Введение. Рост числа цифровых систем и обрабатываемых в них данных, расширение сферы применения обусловливают не только появление и развитие новых технологий, но и усиление влияния угроз безопасности реальным процессам со стороны злоумышленников. Утечки персональных и иных данных, искажения передаваемой информации не только могут нанести и уже наносят вред репутации или финансовой составляющей, но и несут более серьезные угрозы для общества и государства. Поэтому обеспечение безопасности систем управления большими данными от кибератак является чрезвычайно актуальной задачей.

Сложность и гетерогенность рассматриваемого класса систем не позволяют напрямую применить к ним целый ряд устоявшихся методов защиты [1]. Это усложняет решение таких задач, как контроль доступа, аудит и оценка защищенности больших данных на уровне инженерии. Поэтому прежде всего необходимо разработать метод моделирования процессов обработки информации в системах больших данных, который позволит эффективно решать задачи безопасности в условиях постоянных киберугроз. Построение адекватных комплексных моделей информационных процессов, обладающих универсальностью относительно инструментов обработки данных, дает возможность решать задачи их безопасности на основе системного подхода с использованием общепринятых положений в области защиты информации.

Моделирование информационных процессов в кибербезопасности

Моделирование информационных процессов в различных системах и средах широко применяется для решения задач обеспечения их безопасности в современных технологиях. Можно выделить несколько больших групп и направлений работ.

Прежде всего это моделирование процессов в сложных системах (в том числе операционных), лежащее в основе безопасности больших корпоративных систем и отчасти технологий центров управления безопасностью. Исторически оно базируется на понятии конфиденциальности, выраженном через политику безопасности [2], представляющую собой с математической точки зрения алгебру бинарных отношений между дискретными счетными сущностями (множество субъектов $\{S\}$, множество объектов $\{O\}$). Считается, что в общем случае при задании политики безопасности структура множеств $\{S\}$ и $\{O\}$ неизменна. В дальнейшем это направление развивалось в сторону усложнения понятий субъекта и объекта [3] или отношений между ними от самого факта существования через категории важности до современных сложных ассоциаций [4, 5]. В основе работ этого подхода фактически лежит логико-концептуальная модель системы обработки информации. В этом на практике и возникало серьезное противоречие. В итоге именно сама модель порождала структуру пространства S, O, а не наоборот. В условиях неоднозначности отражения модели политики безопасности на множество $S \times O$ это привело к тому, что в реальной системе невозможно доказать ни разрешимость модели безопасности, ни ее достижимость, ни однозначность.

Такая ситуация обусловила развитие обратного направления. В исследованиях [6, 7] предлагается сначала построить логико-концептуальную модель системы в виде архитектуры логических и физических связей, а затем строить по ней алгебру отношений. Рассматривая работы в этой области, можно выделить несколько принципиально различных подходов к моделированию на уровне выбора математического аппарата, а именно: функционально-семантическое моделирование (включая логико-контекстные методы, моделирование на основе информационных потоков [8], моделирование иерархическими адаптивными графами [9]), стохастические динамические модели [10], фрактальные модели [11], фазовые и информационные портреты [12], а также графические и структурные модели (как правило, на основе графов [6, 13]).

Анализ позволил сделать вывод о том, что моделирование распределенных систем является основой построения их безопасности, а модели на основе аппарата теории графов доминирующими и наиболее передовыми. Однако как они построены – на основе структурных, функциональных графов или иным образом, зависит от моделируемой системы и является уникальным отличием каждой конкретной модели. Таким образом, выбор конкретного метода моделирования в рамках общего математического аппарата определяется типом и спецификой системы, безопасность которой рассматривается.

Сегодня в области кибербезопасности отсутствуют модели для решения задач безопасности именно в отношении больших данных. Лишь в работе [14] предлагается обобщенная модель в рамках функционально-семантического подхода. Однако описан только теоретический подход, механизм построения модели такого класса на практике отсутствует. Поэтому важной задачей является на основе подхода, предложенного в [14], разработать модель и практический способ поддержания ее актуальности с учетом специфики операций информационных потоков.

Специфика моделирования процессов управления большими данными

Следует отметить, что исследователи предпринимают попытки моделирования систем больших данных. Если исключить высокоуровневые модели с высокой степенью абстракции, недостаточно детальные для решения технических задач (приведенные, в частности, в [15]), то основные работы в области моделирования процессов обработки данных сосредоточены на моделировании с использованием технологий описания бизнес-процессов [16], с применением методологии моделирования распределенных информационных систем [17], на основе больших (мета)данных [18]. Задачи безопасности исследуются в работе [19], где предлагается использовать моделирование на основе узлов и клиентских запросов, коррелирующее с исследованиями по применению к такого рода задачам теории массового обслуживания [20]. Основной проблемой указанных подходов является отсутствие учета собственно структуризации данных и ограничений ин-

струментов обработки, что не позволяет применить их на практике к решению задач, например, по гранулированному контролю доступа в системах больших данных на всех этапах жизненного цикла. Как показано в [21], прогресс привел к существенным изменениям в системах больших данных по сравнению с традиционными СУБД и требует новых подходов в обеспечении безопасности.

Выделим основные особенности систем управления большими данными, не позволяющие применить приведенные выше методы для построения функционально-семантической модели:

– различная структуризация (грануляция) данных на разных этапах жизненного цикла;

– ограничения возможностей инструментов обработки данных (технологические ограничения).

Пример различий в грануляции данных между различными СУБД в рамках системы больших данных, используемых на разных этапах жизненного цикла, приведен на рисунке 1. На нем показана грануляция в модели данных *реляционной (пост-реляционной) СУБД* (РСУБД) и системы управления данными на основе ключ-значение. При этом технологические особенности системы ключ-значение не позволяют провести детализацию доступа на достаточном уровне.

Таким образом, методы моделирования систем больших данных в кибербезопасности должны предусматривать отражение различий в структуризации информации и позволять преодолевать проблему грануляции данных в рамках разных СУБД и других инструментов – компонентов систем управления большими данными.

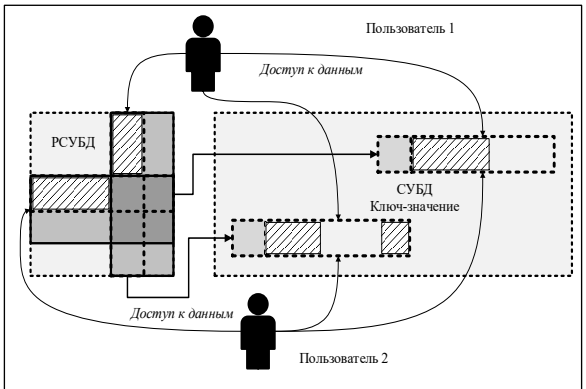

Рис. 1. Пример влияния различий в грануляции данных на разграничение доступа

Fig. 1. The effect of differences in data granulation on access control

Модель информационных процессов управления большими данными

Предлагаемая модель основана на математической теории графов, хорошо зарекомендовавшей себя при моделировании современных информационных систем и общем подходе концептуального моделирования на основе графов, предложенном в [14]. Система в целом является орграфом $G = \{V, E\}$, где V – множество вершин графа, представляющих собой структурированные фрагменты данных на различных этапах жизненного цикла, ребра графа E – операции над данными.

Для учета структурных преобразований и грануляции данных предлагается представить возможные преобразования данных в виде обобщенных операций в контексте модели. С технологической точки зрения и в рамках концепции архитектуры ANSI/SPARK информационные процессы в системах управления большими данными можно представить (без учета инфраструктуры) на логическом уровне обработки данных и уровне распределенных инструментов их обработки (рис. 2).

На верхнем уровне изображены различные структуры, содержащие данные от момента поступления из источника до получения их системой – пользователем (1–9). На нижнем – инструменты системы (1–4), реализующие преобразования данных между структурами. В пределах одного инструмента данные имеют общую структуризацию, тогда как между инструментами структуризация и грануляция данных могут отличаться в рамках общепринятых и распространенных сегодня моделей: ключ-значение, документно-ориентированное представление, семейство столбцов, колоночное, реляционное или графовое представление.

Это представление упрощает многоуровневый подход исходного исследования и позволяет связать конкретные операции модели непосредственно с процессами, происходящими в системе управления большими данными.

Таким образом, основная задача моделирования процессов в системе управления большими данными заключается в моделировании преобразований данных как внутри одной модели (и, соответственно, одного инструмента обработки), так и между структурами при взаимодействии различных инструментов. Основными технологиями оперирования данными являются варианты реализации SQL-операций, физически представленные планами запросов,

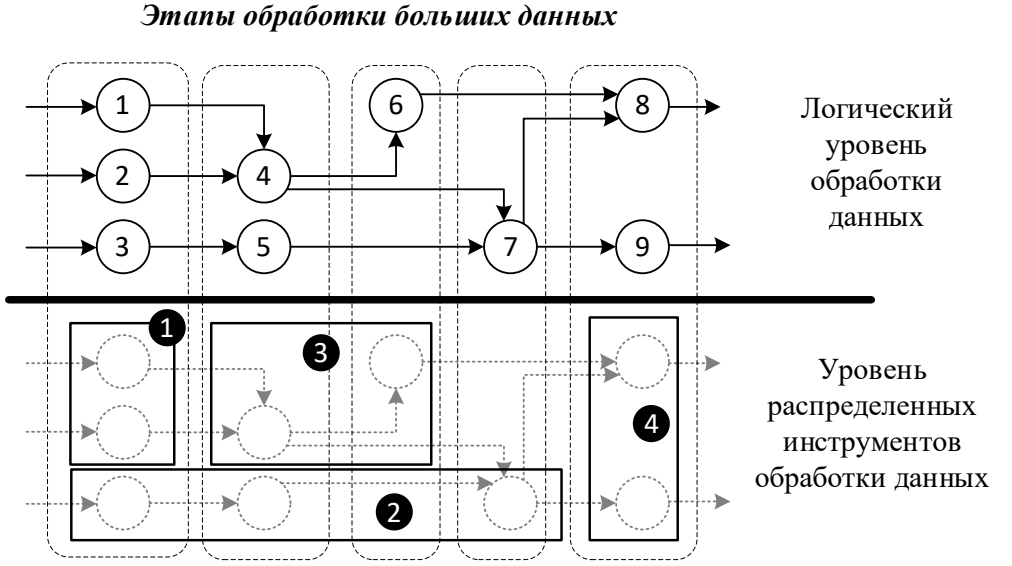

Рис. 2. Уровни представления процессов обработки информации в системах управления большими данными

Fig. 2. Representation levels of information processing in big data management systems

содержащими такие основные этапы, как выборка, проекция и соединение, и различные реализации технологии Map-Reduce. Анализ языков запроса и физической реализации операций с данными позволяет свести все действия с данными в системе больших данных к трем типам операций (исключая удаление и создание): объединение (∪), разделение или извлечение (\), преобразование данных (F()).

При объединении или извлечении (разделении) фрагмента данных сохраняется семантика как исходного, так и извлекаемого фрагмента (речь идет о преобразовании множеств фрагментов данных). Преобразование собственно данных фактически означает создание нового фрагмента на основе предшествующего. С точки зрения информационной безопасности важны некоторые аспекты.

1. При объединении и разделении данных их семантика сохраняется без изменений, и эта операция может быть произведена без доступа к семантике (например, над гомоморфно зашифрованными данными).

2. При преобразовании данных для получения нового значения доступ к семантике исходных данных обязателен. Однако полученное значение также (в общем случае) сохраняет некоторую связь с породившими его. Например, среднее значение за период связано с отдельными значениями в этот период. Данный факт важно учитывать при оценке возможного логического вывода над данными.

Таким образом, множество вершин и ребер графа G можно описать следующим образом:

$$E = \{op_i \mid (op_i = \cup) \vee (op_i = \setminus) \vee (op_i = F(d_{i-1}))\},$$

$$V = \{d_i \mid i = [0, N]\},$$

где d_i – i-й фрагмент данных; N – общее число различных фрагментов в системе.

Важно указать, что в данном случае под фрагментом понимается некоторый тип элементов данных с общей семантикой, синтаксисом (форматом) и маршрутом обработки. В рамках модели два элемента данных (например, два показателя с одного датчика температуры) будут считаться разными, если имеют разный порядок или маршрут обработки.

В данном случае получение каждого выходного фрагмента данных из некоторого множества входных можно представить в виде подграфа исходного графа G. Основным преимуществом такой модели для решения задач кибербезопасности больших данных является возможность отслеживания жизненного цикла фрагментов данных с учетом их преобразования в процессе обработки.

Применение модели информационных процессов управления большими данными для анализа политик безопасности

В качестве примера использования представленной модели служит решение задачи обеспечения разграничения доступа в системе больших данных. Для учета таких атак, как

атака логического вывода, и для минимизации утечек информации необходимо обеспечить порядок разграничения доступа, при котором для получателей данных, с одной стороны, закрыты источники, доступ к которым им запрещен, а с другой, соблюдена бизнес-логика работы. Если получателям будут назначены права в отношении результирующих фрагментов, как часто происходит на практике, может быть нарушено первое условие, если в отношении исходных – второе. Обеспечение рационального разграничения доступа в этом случае крайне сложно и трудоемко, а поддержание его тем более.

Для автоматизации процесса анализа политик безопасности в данном примере определяются также S – множество субъектов системы; $D_{вх} \subseteq V$, $D_{вых} \subseteq V$ – множества входных и выходных фрагментов данных соответственно; $P = \{P_1, P_2, \ldots, P_{|S|}\}$ – множество прав доступа субъектов S к объектам в вершинах V, где $P_i = (p_{i,1}, p_{i,2}, \ldots, p_{i,j}, \ldots, p_{i,|V|})$ – вектор доступа субъекта s_i ко всем фрагментам данных ($p_{i,j} = 0$, если s_i не имеет доступа к объекту j, иначе $p_{i,j} = 1$). Также определяются правила переноса прав доступа внутри модели. Отметим, что эти сведения могут быть автоматически получены через запросы к системным файлам и каталогам инструментов обработки данных.

Модель системы управления большими данными задается как направленный граф GraphRepresent() через конфигурационный файл или модуль считывания системы распределенного аудита. Для каждого субъекта вычисляются права доступа ко всем фрагментам данных в зависимости от направления анализа от источников данных к получателям или наоборот. Они сохраняются в специальный список AddSubject(). В результате аналитик безопасности получает сформированную таблицу прав доступа и может оценить выполнение политики безопасности, значительно сэкономив время на анализе системы. Модель в этом случае позволяет подбирать оптимальные настройки прав доступа, правила их переноса и конфигурации системы управления данными, оценивая все изменения путем моделирования для дальнейшего переноса итогового (наилучшего) результата на реальную систему. Пример промежуточного вывода рассчитанных прав доступа для оценки аналитиком безопасности представлен на рисунке 3.

Приведенный пример использования предложенной модели не единственный. Модель и способ моделирования процессов обработки

```
Permissions out:
Subject                |16|17|18|
-----------------------+--+--+--+
Manager                |+ |+ |+ |
-----------------------+--+--+--+
System Administrator   |+ |- |- |
-----------------------+--+--+--+
Analyst team           |+ |- |- |
-----------------------+--+--+--+
Accounting             |- |- |- |
-----------------------+--+--+--+
Laboratory             |- |- |- |
-----------------------+--+--+--+
```

Рис. 3. Пример результата определения доступа получателей данных к их исходным фрагментам

Fig. 3. The result of determining data recipient access to original data fragments

информации в гетерогенных системах управления большими данными позволяют также проводить анализ защищенности, аудит безопасности системы и при использовании соответствующих криптографических средств осуществлять обработку больших данных в рамках концепции нулевого доверия.

Заключение

Моделирование современных сложных цифровых систем является важным и неотъемлемым этапом при проектировании новых методов и средств обеспечения их безопасности в условиях растущего числа кибератак. Представленная графовая модель соответствует общим принципам и подходам, применяемым сегодня при моделировании объектов информатизации и цифровых решений в области информационной безопасности и систем управления большими данными.

Ее новизна заключается в детализации многоуровневого представления до двух базовых уровней, из которых модель фактически охватывает логический уровень операций с данными, а также в конкретизации операций с данными до трех типов основных преобразований, соответствующих преобразованиям в нижележащих СУБД. Предложенные авторами операции с данными – обобщение, разделение и преобразование – позволяют описывать все преобразования, проводимые с данными в современных системах управления и комплексах различных СУБД, в достаточной степени детализации для решения таких задач, как аудит, контроль доступа и др.

Представление в форме орграфа преобразований данных соответствует общим тенденциям графового моделирования распределенных (в том числе киберфизических) систем. Использование единого математического аппарата позволяет интегрировать модель в представленные ранее модели для более сложных комплексов и информационных систем как независимый компонент.

Модель отличается универсальностью с точки зрения решения задач безопасности, соответствующей использованному подходу к моделированию, и при этом обладает высокой степенью автоматизации, что на практике является ее значительным преимуществом. Данные для построения могут быть как заданы декларативно на этапе проектирования, так и получены на реальной системе в автоматическом режиме через использование систем распределенного аудита.

Определенным недостатком модели все еще является упрощенное представление структур данных. Хотя для целого ряда рассматриваемых задач большей степени математической формализации структур данных не требуется, при решении более общих, фундаментальных задач в данной области и обеспечении комплексного контроля и безопасности может потребоваться разработка не только модели процессов управления большими данными, представленной в данной работе, но и полноценной модели данных концептуального уровня аналогично современным СУБД.

Список литературы

1. Naeem M., Jamal T., Diaz-Martinez J., Butt Sh.A., Montesano N. et al. Trends and future perspective challenges in big data. In: Advances in Intelligent Data Analysis and Applications. SIST, 2022, vol. 253, pp. 309–325. doi: 10.1007/978-981-16-5036-9_30.

2. Чернов С.Б., Новикова О.С. Обеспечение безопасности данных в условиях цифровой экономики // Экономические науки. 2020. № 8. С. 104–109. doi: 10.14451/1.189.104.

3. Богаченко Н.Ф. Анализ проблем управления разграничением доступа в крупномасштабных информационных системах // Математические структуры и моделирование. 2018. № 2. С. 135–152.

4. Статьев В.Ю., Докучаев В.А., Маклачкова В.В. Информационная безопасность на пространстве «больших данных» // T-Comm: Телекоммуникации и Транспорт. 2022. Т. 16. № 4. С. 21–28. doi: 10.36724/2072-8735-2022-16-4-21-28.

5. Haourani L.E., Elkalam A.A., Ouahman A.A. Knowledge based access control a model for security and privacy in the Big Data. Proc. Int. Conf. SCA, 2018, art. 16, pp. 1–8. doi: 10.1145/3286606.3286793.

6. Zegzhda D., Zegzhda P., Pechenkin A., Poltavtseva M. Modeling of information systems to their security evaluation. Proc. Int. Conf. SIN, 2017, pp. 295–298. doi: 10.1145/3136825.3136857.

7. Shukla A., Katt B., Nweke L.O., Yeng P.K., Weldehawaryat G.K. System security assurance: A systematic literature review. Comput. Sci. Review, 2022, vol. 45, art. 100496. doi: 10.1016/j.cosrev.2022.100496.

8. Ashibani Y., Mahmoud Q.H. Cyber physical systems security: Analysis, challenges and solutions. Computers & Security, 2017, vol. 68, pp. 81–97. doi: 10.1016/j.cose.2017.04.005.

9. Zegzhda D., Lavrova D., Pavlenko E., Shtyrkina A. Cyber attack prevention based on evolutionary cybernetics approach. Symmetry, 2020, vol. 12, no. 11, art. 1931. doi: 10.3390/sym12111931.

10. Zegzhda P.D., Anisimov V.G., Suprun A.F., Anisimov E.G., Saurenko T.N., Los' V.P. A model of optimal complexification of measures providing information security. Automatic Control and Comput. Sci., 2020, vol. 54, pp. 930–936. doi: 10.3103/S0146411620080374.

11. Lavrova D.S., Popova E.A., Shtyrkina A.A. Prevention of DoS attacks by predicting the values of correlation network traffic parameters. Automatic Control and Comput. Sci., 2019, vol. 53, pp. 1065–1071. doi: 10.3103/S0146411619080157.

12. Lavrova D.S. Maintaining cyber sustainability in industrial systems based on the concept of molecular-genetic control systems. Automatic Control and Comput. Sci., 2019, vol. 53, pp. 1026–1028. doi: 10.3103/S0146411619080145.

13. DeLong R.J., Rudina E. MILS architectural approach supporting trustworthiness of the IIoT solutions. IIC Whitepaper, 2018, 96 p.

14. Poltavtseva M.A., Kalinin M.O. Modeling big data management systems in information security. Automatic Control and Comput. Sci., 2019, vol. 53, pp. 895–902. doi: 10.3103/S014641161908025X.

15. Константинов А., Кузнецов С., Скворцов Н. Ценность ваших данных. М.: Альпина ПРО, 2022. 410 с.

16. Пудеян Л.О., Запорожцева Е.Н., Медведская Т.К. Применение инструментов моделирования и анализа больших данных в управлении бизнес-процессами // ВАЗ. 2022. № 6. С. 229–233.

17. Мельникова Т.В., Питолин М.В., Преображенский Ю.П. Моделирование обработки больших массивов данных в распределенных информационно-телекоммуникационных системах // Моделирование, оптимизация и информационные технологии. 2022. Т. 10. № 1. С. 136–141.

18. Бурый А.С., Шевкунов М.А. Суррогатное моделирование распределенных информационных систем по большим данным // ИЭА СТР. 2019. № 5. С. 43–50.

19. Котенко И.В., Проничев А.П. Моделирование процессов обработки запросов в распределенных системах хранения больших данных // АПИНО: сб. науч. ст. 2020. Т. 1. С. 620–624.

20. Shelest M. Lower bound for average delay in transactions processing systems. Proc. XV Int. Symposium REDUNDANCY, 2016, pp. 142–145. doi: 10.1109/RED.2016.7779349.

21. Gahar R.M., Arfaoui O., Hidri M.S. Towards Big Data modeling and management systems: From DBMS to BDMS. Proc. IC_ASET, 2023, pp. 1–6. doi: 10.1109/IC_ASET58101.2023.10151190.

Software & Systems doi: 10.15827/0236-235X.142.054-061 2024, 37(1), pp. 54–61

Modeling information processes of big data management systems to solve cybersecurity problems

Mariya A. Poltavtseva [1]✉, Dmitry P. Zegzhda [1]

[1] Institute of Cyber Security and Information Protection, Peter the Great SPbPU, St. Petersburg, 195251, Russian Federation

For citation
Poltavtseva, M.A., Zegzhda, D.P. (2024) 'Modeling information processes of big data management systems to solve cybersecurity problems', *Software & Systems*, 37(1), pp. 54–61 (in Russ.). doi: 10.15827/0236-235X.142.054-061
Article info
Received: 23.08.2023 After revision: 08.09.2023 Accepted: 11.09.2023

Abstract. The imperfection of classical security models when applied to real systems has led to developing a reverse approach: modeling systems of different classes to subsequently supplement their models with security attributes. Nowadays solving distributed system security problems based on such models is a dynamically developing area of scientific knowledge. The paper considers modeling of heterogeneous big data management systems for solving modern cybersecurity problems. The authors identify and take into account such key features of the system class under consideration as using heterogeneous data structures and limitations of data manipulation tools, primarily with respect to the granularity of security functions during implementation. The paper proposes a graph model of a big data management system using generalized operations on data: merge, split and transform. Graph vertices represent structured data fragments, the arcs represent their processing operations regardless of a specific tool and a transformation type. Due to generalized operations, the model allows taking into account data transformations both within processing tools and when transferring information between them; it provides a comprehensive representation of information processing at the data engineering level. A special feature of the model is its construction automation based on a specific big data system, which helps maintaining adequacy during evolutionary changes in the modeled object. The presented model allows solving a wide range of problems in the field of security of large-scale heterogeneous systems, such as access control, auditing, security assessment. As an example, the paper shows how use the proposed model to automate the analysis of security policies in this class of systems.
Keywords: big data, data management systems, DBMS, modeling, data modeling, information protection, cybersecurity
Acknowledgements. The study was funded by the Russian Science Foundation Grant No. 23-11-20003, https://rscf.ru/project/23-11-20003/ St. Petersburg Science Foundation Grant (Regional Grant Agreement No. 23-11-20003)

References

1. Naeem, M., Jamal, T., Diaz-Martinez, J., Butt, Sh.A., Montesano, N. et al. (2022) 'Trends and future perspective challenges in big data', in Advances in Intelligent Data Analysis and Applications. *SIST*, 253, pp. 309–325. doi: 10.1007/978-981-16-5036-9_30.

2. Chernov, S.B., Novikova, O.S. (2020) 'Ensuring data security in digital economy', *Economic Sci*., (8), pp. 104–109 (in Russ.). doi: 10.14451/1.189.104.

3. Bogachenko, N.F. (2018) 'The analysis of problems of access control administration in large-scale information systems', *Math. Structures and Modeling*, (2), pp. 135–152 (in Russ.).

4. Statev, V.Yu., Dokuchaev, V.A., Maklachkova, V.V. (2022) 'Information security in the big data space', *T-Comm*, 16(4), pp. 21–28 (in Russ.). doi: 10.36724/2072-8735-2022-16-4-21-28.

5. Haourani, L.E., Elkalam, A.A., Ouahman, A.A. (2016) 'Knowledge based access control a model for security and privacy in the Big Data', *Proc. Int. Conf. SCA*, art. 16, pp. 1–8. doi: 10.1145/3286606.3286793.

6. Zegzhda, D., Zegzhda, P., Pechenkin, A., Poltavtseva, M. (2017) 'Modeling of information systems to their security evaluation', *Proc. Int. Conf. SIN*, pp. 295–298. doi: 10.1145/3136825.3136857.

7. Shukla, A., Katt, B., Nweke, L.O., Yeng, P.K., Weldehawaryat, G.K. (2022) 'System security assurance: A systematic literature review', *Comput. Sci. Review*, 45, art. 100496. doi: 10.1016/j.cosrev.2022.100496.

8. Ashibani, Y., Mahmoud, Q.H. (2017) 'Cyber physical systems security: Analysis, challenges and solutions', *Computers & Security*, 68, pp. 81–97. doi: 10.1016/j.cose.2017.04.005.

9. Zegzhda, D., Lavrova, D., Pavlenko, E., Shtyrkina, A. (2020) 'Cyber attack prevention based on evolutionary cybernetics approach', *Symmetry*, 12(11), art. 1931. doi: 10.3390/sym12111931.

10. Zegzhda, P.D., Anisimov, V.G., Suprun, A.F., Anisimov, E.G., Saurenko, T.N., Los', V.P. (2020) 'A model of optimal complexification of measures providing information security', *Automatic Control and Comput. Sci.*, 54, pp. 930–936. doi: 10.3103/S0146411620080374.

11. Lavrova, D.S., Popova, E.A., Shtyrkina, A.A. (2019) 'Prevention of DoS attacks by predicting the values of correlation network traffic parameters', *Automatic Control and Comput. Sci*., 53, pp. 1065–1071. doi: 10.3103/S0146411619080157.

12. Lavrova, D.S. (2019) 'Maintaining cyber sustainability in industrial systems based on the concept of molecular-genetic control systems', *Automatic Control and Comput. Sci*., 53, pp. 1026–1028. doi: 10.3103/S0146411619080145.

13. DeLong, R.J., Rudina, E. (2018) 'MILS architectural approach supporting trustworthiness of the IIoT solutions', *IIC Whitepaper*, 96 p.

14. Poltavtseva, M.A., Kalinin, M.O. (2019) 'Modeling big data management systems in information security', *Automatic Control and Comput. Sci*., 53, pp. 895–902. doi: 10.3103/S014641161908025X.

15. Konstantinov, A., Kuznetsov, S., Skvortsov, N. (2022) *The Value of your Data*. Moscow, 410 p. (in Russ.).

16. Pudeyan, L.O., Zaporozhtseva, E.N., Medvedskaya, T.K. (2022) 'Applying big data modeling and analysis tools in business process management', *Bull. of the Academy of Knowledge*, (6), pp. 229–233 (in Russ.).

17. Melnikova, T.V., Pitolin, M.V., Preobrazhenskiy, Yu.P. (2022) 'Modeling of large data array processing in distributed information and telecommunication systems', *Modeling, Optimization and Inform. Tech.*, 10(1), pp. 136–141 (in Russ.).

18. Bury, A.S., Shevkunov, M.A. (2019) 'Surrogate modeling of distributed information systems on Big Data'', *Information and Economic Aspects of Standardization and Tech. Regulation*, (5), pp. 43–50 (in Russ.).

19. Kotenko, I., Pronichev, A. (2020) 'Modeling query processing processes in distributed big data storage systems', *Proc. ICAIT*, 1, pp. 620–624 (in Russ.).

20. Shelest, M. (2016) 'Lower bound for average delay in transactions processing systems', *Proc. XV Int. Symposium REDUNDANCY*, pp. 142–145. doi: 10.1109/RED.2016.7779349.

21. Gahar, R.M., Arfaoui, O., Hidri, M.S. (2023) 'Towards Big Data modeling and management systems: From DBMS to BDMS', *Proc. IC_ASET*, pp. 1–6. doi: 10.1109/IC_ASET58101.2023.10151190.

Авторы

Полтавцева Мария Анатольевна [1],
д.т.н., доцент, профессор,
poltavtseva@ibks.spbstu.ru
Зегжда Дмитрий Петрович [1], д.т.н.,
профессор, чл.-корр. РАН, директор,
dmitry@ibks.spbstu.ru

[1] Институт кибербезопасности
и защиты информации, СПбПУ Петра Великого,
г. Санкт-Петербург, 195251, Россия

Authors

Mariya A. Poltavtseva [1], Dr.Sc. (Engineering),
Associate Professor, Professor,
poltavtseva@ibks.spbstu.ru
Dmitry P. Zegzhda [1], Dr.Sc. (Engineering), Professor,
Corresponding Member of RAS, Director,
dmitry@ibks.spbstu.ru

[1] Institute of Cyber Security
and Information Protection, Peter the Great SPbPU,
St. Petersburg, 195251, Russian Federation