Раздел II. Методы защиты и технологии безопасности

УДК 004.056.5 DOI 10.18522/2311-3103-2025-3-55-62

М.А. Полтавцева, Д.В. Иванов

КЛАССИФИКАЦИЯ УЗЛОВ – ОБРАБОТЧИКОВ В СИСТЕМАХ БОЛЬШИХ ДАННЫХ В СООТВЕТСТВИИ С ПОДХОДОМ НУЛЕВОГО ДОВЕРИЯ

Кибербезопасность данных является одним из важнейших факторов успешной реализации национального проекта «Экономика данных и цифровая трансформация государства». Проблемы построения защищенных систем обработки больших данных заключаются в их гетерогенной природе, большом числе разнородных инструментов, высокой связности и высоком доверии между распределенными компонентами. Снижение внутреннего доверия и уменьшение поверхности атаки в соответствии с подходом zero-trust необходимо для повышения защищенности таких систем с наименьшим влиянием на их производительность. Целью работы является создание метода динамической классификации узлов и компонент обработки данных в гетерогенных системах больших данных на основе применения различных подходов к снижению доверия в отношении объектов, реализующих процесс обработки информации. Рассматривается подход нулевого доверия применительно к исследуемому классу систем, а также ставится задача расширенной реализации принципа минимальных привилегий уменьшения поверхности атаки. Представлена классификация узлов – обработчиков на основе выполняемых ими операций с данными, унифицированных согласно разработанной ранее концептуальной модели данных. Предлагается сопоставление узлов и применяемых в их отношении методов безопасности на основе необходимости доступа к семантике и компонентам данных для выполнения операций. На основе данной классификации разработан метод динамического определения класса узлов-обработчиков данных в процессе работы системы для ситуаций изменения компонентного состава системы обработки больших данных, типичной для многокомпонентных распределенных высоконагруженных систем. Результаты работы являются частью комплексного консистентного подхода к построению защищенных систем обработки больших данных.

Большие данные; системы обработки данных; гетерогенные системы больших данных; информационная безопасность; нулевое доверие; операции с данными; управление инфраструктурой.

M.A. Poltavtseva, D.V. Ivanov

CLASSIFICATION OF PROCESSING NODES IN BIG DATA SYSTEMS ACCORDING TO THE ZERO TRUST APPROACH

Data cybersecurity is one of the most important factors for the successful implementation of the national project 'Data Economy and Digital Transformation of the State'. The challenges of building secure big data systems lie in their heterogeneous nature, large number of heterogeneous tools, high connectivity and high trust between distributed components. Reducing the internal trust and reducing the attack surface according to the zero-trust approach is necessary to increase the security of such systems with the least impact on their performance. The aim of the paper is to create a method for dynamic classification of nodes and data processing components in heterogeneous big data systems based on the application of different approaches to trust reduction with respect to the objects realising the information processing process. The paper considers the zero trust approach as applied to the class of systems under study, as well as the task of extended implementation of the principle of minimum privilege to reduce the attack surface. The authors present a classification of nodes - handlers based on their operations with data, unified according to the previously developed conceptual data model. A comparison of nodes and security methods applied to them based on the need for access to semantics and data components to perform operations is proposed. Based on this classification, a method of dynamic node type determination during sys-

tem operation is developed for situations of changing component composition of a big data processing system, typical for multi-component distributed highly loaded systems. The results of the work are a part of the complex consistency approach to the construction of secure big data processing systems.

Big Data; data processing systems; heterogeneous big data systems; information security; zero-trust; data operations; infrastructure management.

Введение. На сегодняшний день можно говорить о синергии дух трендов: развития и повсеместной цифровизации с одной стороны и увеличения числа атак на информационные системы различной природы с другой. Появление и развитие систем больших данных во многом усугубило проблемы безопасности, так как для них свойственна концентрация большого числа разнородной конфиденциальной информации в одном месте. Также стоит отметить снижение безопасности, вызванное особенностями распределенной организации таких систем.

Ключевую сложность при обеспечении защиты систем обработки и хранения больших данных представляет собой их гетерогенная природа: сочетание различных инструментов и способов обработки информации в одном жизненном цикле. В силу большого числа самостоятельных компонентов, как правило – доверенных в отношении друг друга, для этого класса решений характерен рост вероятности реализации угроз, в том числе - угроз конфиденциальности данных, со стороны внутреннего привилегированного нарушителя. Поэтому построение безопасности этого класса систем на базе подхода нулевого доверия (zero-trust) является важной задачей. Целью данной работы является формирование динамического метода классификации узлов систем обработки и хранения больших данных на основе выполняемых ими информационных операций с целью определения требований к архитектуре безопасности и методу защиты.

Применение подхода минимального доверия к узлам – обработчикам данных в системах больших данных. Понятие «доверия» в информационной безопасности до конца не определено и имеет разное значение в зависимости от контекста [1]. С одной стороны, под доверием понимают уверенность в действиях и/или корректности с некоторой стороны (участника) [2], с другой – общепризнанным с 2010-х годов подходом zero-trust или нулевого доверия [3]. В любом случае необходимо отметить, что понятие доверия тесно связано с рисками информационной безопасности и может быть соотнесено с компонентами информационных систем [4], а следовательно – и систем управления большими данными. Так доверие в системах больших данных также имеет два значения: доверие пользователя (отправителя или получателя данных) к системе обработки и хранения больших данных, как уверенность в характеристиках данных при работе с системой: полноте, целостности, конфиденциальности и доступности данных; доверие компонентов системы управления больших данных друг к другу как между уровней представления, так и внутри этих уровней [5]. С этой точки зрения под степенью доверия в системах управления больших данных в соответствии подходом нулевого доверия (zero trust) будем понимать степень безусловной уверенности в характеристиках безопасности участников обработки данных и самой информации.

Реализация принципа минимальных привилегий [6], как и уменьшение поверхности атаки, приводит к уменьшению уязвимости систем обработки и хранения больших данных [7] и, таким образом, снижению рисков информационной безопасности и реализации подхода нулевого доверия [8]. Формирование и реализация оптимальной политики безопасности в системах обработки и хранения больших данных происходит в условиях не только ограничений бизнес-логики, определяющей не снижаемые риски в рамках конкретной технологии, но и в условиях технологических ограничений [9]. Поэтому корректно говорить скорее о стремлении к нулевому доверию (в идеологическом смысле), чем о его реализации. Особенностью данных систем является наличие группы привилегированных пользователей – администраторов, имеющих высокий уровень доступа к данным (в том числе, в силу несовершенства механизмов разграничения доступа) [10]. В таких условиях необходимо стремиться к уменьшению доступа привилегированных пользователей, выведя их за границы безусловного доверия в отношении обрабатываемых данных.

Несмотря на то, что приведенная выше концепция нулевого доверия, широко принята в информационной безопасности, а снижение доверия к среде в различных областях остается де-факто стандартом безопасности [11, 12], сегодня в силу сложности практической реализации она подвергается критике и нуждается в адаптации применительно к каждому конкретному классу систем [13]. В свою очередь минимизация доверия в отношении компонентов обработки информации является сложной задачей для систем больших данных, которым характерно несколько распределенных сред и инструментов, интегрированных в рамках управления общим циклом обработки и хранения данных.

Можно выделить три подхода для снижения требуемого уровня доверия компонентам обработки данных [14]: использование защитного контура или защита периметра [15, 16], сквозное шифрование уровня приложения [17, 18], обфускация (маскирование) данных [19]. На их основе в системе больших данных формируется гибридная архитектура безопасности. Следующим шагом становится классификация узлов обработчиков для распределения по архитектурным компонентам: обрабатывающим открытые данные, маскированные данные, зашифрованные данные.

Классификация узлов-обработчиков данных. Для решения поставленной задачи необходимо разделить узлы – обработчики данных для того, чтобы определить технологическую необходимость в доступе к семантике данных при выполнении операций. Рассмотрим необходимость доступа к семантике для всех типов операций, определенных в общей концептуальной модели данных для гетерогенных систем больших данных [20].

Создание Create:({d},Key)→di={Key,Value} само по себе подразумевает доступ к значению Value создаваемого агрегата, а значит – доступ к семантике.

Уничтожение Delete:(di)→∅ не предполагает доступа к семантике уничтожаемого агрегата.

Включение Incl:(di,{d})→di’ в общем случае не требует доступа к семантике включаемых фрагментов, однако подразумевает доступ к значению Value получаемого фрагмента, или, по крайней мере, его составляющим.

Исключение Extr:(di’,{Keyj})→(di, dj). Операция исключения в ряде случаев может быть выполнена без доступа к семантике, однако в большинстве случаев такой доступ потребуется. Проблема заключается в том, что при исключении нужно выделить новый агрегат из существующего, а значит получить доступ к содержимому (путь и составному) его значения (Value). Поэтому к этой операции относится, например, поиск. На более низком уровне детализации операция исключения может быть реализована с использованием сравнения (как например поиск) что также следует учитывать. В целом отнесем операцию исключения к требующим доступа к семантике.

Преобразование Transform:(di)→dj операция явно предполагающая доступ к семантике данных (например, со стороны пользователя или приложения, расчет показателей на основе данных).

Итоговая систематизация узлов – обработчиков данных в виде нечеткой классификации на основе технологического уровня и доступа к семантике для реализации принципа минимизации доверия представлена на рис. 1.

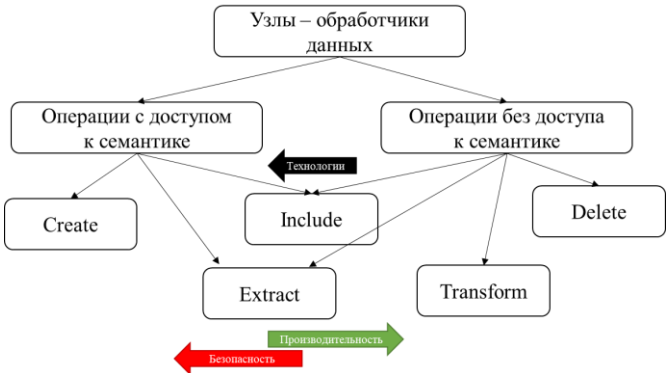

Рис. 1. Классификация узлов-обработчиков данных

Возможность выполнения операций включения и исключения без доступа к семантике зависит от двух факторов: особенностей выполнения операции на нижележащем уровне, уровне инструментов обработки и хранения данных и технологических возможностей, или наличия эффективного выполнения данной операции над зашифрованными данными без использования специфических алгоритмов (либо же интеграцией таких алгоритмов в узлы – обработчики данных).

Несмотря на то, что при использовании гомофонного шифрования достаточно сложные операции с данными могут производиться без доступа к семантике, их производительность остается низкой и их эффективное применение на сегодняшний день не может рассматриваться в широком круге систем обработки и хранения больших данных.

Классификация узлов – обработчиков данных на основе принципа минимизации доверия основывается в первую очередь на доступе к семантике данных, однако есть еще такой аспект, как динамичность систем обработки и хранения больших данных. Можно выделить по крайней мере два аспекта такой динамики:

1. Изменение маршрутов данных, и, как следствие, изменение операций, выполняемых на узле – обработчике (а значит, и его класса в соответствии с приведенной классификацией).

2. Изменение состава узлов – обработчиков из-за отказов и замены двух типов: отказа/замены/добавления узлов кластера в рамках используемого инструмента обработки и/или хранения данных, отказа/замены/добавления инструментов обработки и/или хранения данных.

Новые узлы или инструменты при подключении к системе должны быть автоматически отнесены к одной из приведенных выше категорий (классов). Следовательно, возникает требование их динамической классификации «на лету» на основе текущих данных, состояния системы и процессов обработки информации. Таким образом, требованиями к методу динамической классификации узлов-обработчиков данных на основе принципа минимизации доверия являются:

1. Учет конкретных операций, выполняемых согласно процессу, на классифицируемом узле обработки данных;

2. Учет технологических возможностей по защите данных на узле – обработчике на основе рассмотренных подходов: переносом в защищенный контур, маскированием или шифрованием;

3. Приоритезация рассмотренных подходов с точки зрения производительности и безопасности.

При этом маскирование (обфускация данных) и шифрование – это операции, зависимые, в свою очередь, от операций, которые выполняются над данными каждым отдельным узлом – обработчиком.

Для проведения такой классификации следует определить статус узлов – обработчиков $Node = \{n_1, \dots, n_N\}$ относительно архитектурных компонентов: множества маскированных узлов ($Node_m$), множества узлов применяющих сквозное шифрование ($Node_{en}$) и множества узлов функционирующих в защищенном контуре и проводящих обработку открытых данных ($Node_{op}$).

Сама по себе задача выбора оптимального метода маскирования является зависимой не только от операций над данными, но и от характеристик самих данных [21] и не может быть универсализирована. Однако практика применения маскирования позволяет выделить круг задач, в которых использование этого подхода можно унифицировать. Примером таких ситуаций служит обработка типовых корпоративных данных [22]. В этом случае конфиденциальные данные заменяются на идентификатор, сохраняющий их основные свойства: уникальность, статистическое распределение и др. В значимом числе случаев для решения задач и выполнения операций над данными достаточно знать характеристики данных, а не их значения. То есть, эти операции априори выполняются без доступа к семантике. При этом для получения маскированных значений $md_{i,j}$ на основе открытых исходных данных $d_{i,j}$на практике используются различные подходы:

Преобразование по определенному обратимому алгоритму:

$$\left(md_{i,j} = Fm_k\left(d_{i,j}\right) | (Fm_k \in Ms)\right) \vee \left(\exists (Fm_k^{-1}) | \left(d_{i,j} = Fm_k^{-1}\left(d_{i,j}\right)\right)\right). \quad (1)$$

Обратимое преобразование с использованием ключа:

$$\left(md_{i,j} = Fm_k\left(d_{i,j}, Key\right) | (Fm_k \in Ms)\right) \vee \left(\exists (Fm_k^{rev}) | \left(d_{i,j} = Fm_k^{rev}\left(d_{i,j}, Key\right)\right)\right). \quad (2)$$

Преобразование на основе таблицы подстановки:

$$\left(md_{i,j} = Fm_k\left(d_{i,j}, Tb\right) | (Fm_k \in Ms)\right) \vee \left(\exists (Fm_k^{rev}) | \left(d_{i,j} = Fm_k^{rev}\left(d_{i,j}, Tb\right)\right)\right). \quad (3)$$

Причем второй случай (2) описывает, фактически, маскирование с использованием методов шифрования, а в третьем случае (3) может использоваться простая функция сопоставления по порядку маскируемого значения с маскирующим. Отметим, что первый случай, представленный (1), является наименее приемлемым, т.к. требует сокрытия алгоритма маскирования, а не ключа. Выбор между вторым и третьим случаем должен производиться в условиях текущей стоимости и доступности ресурсов: объемов памяти и вычислительной мощности.

Если $d_{i,j} \in D$ – определенный тип фрагментов данных, обрабатываемый узлом $n_j \in Node$, то возможность его обработки в том или ином виде определяется производимыми с ним операциями. Определим подмножество операций с данными, не требующих доступа к семантике, в первую очередь с учетом маскирования, Op^{S-} как включение ($Incl$), исключение ($Extr$) и удаление ($Delete$). Остальные операции: создание ($Create$), трансформация ($Transform$) отнесем к подмножеству Op^{S+}, или требующих доступа к семантике.

При этом множество Op^{S+} также может быть разбито и часть операций из него перенесена в Op^{S-}, точнее подмножество Op^{S-E}, если это операции, которые могут быть эффективно выполнены над шифр текстом в условиях используемого метода шифрования. Методы шифрования, доступные для использования в конкретной системе $En = \{en_1, \dots, en_E\}$ характеризуются двумя множествами $en_e = \langle Op_e, To_e \rangle$ где Op_e – операции, которые могут быть выполнены над шифр текстом, а To_e – временные показатели (сложность) этих операций. Каждому фрагменту данных $d_i \in D$ может быть сопоставлена пара значений $\langle op_e, to_e \rangle$ или показано отсутствие поддержки операции над указанным типом данных в конкретном метод. То есть $\left((d_i \leftrightarrow \langle op_e, to_e \rangle) \vee \left(\nexists (d_i \leftrightarrow op_e)\right)\right) \forall (en_e \in En)$. Тогда при некотором заданном граничном значении временных затрат T_{lim} должно соблюдаться ограничение $T < T_{lim}$ где $T = F\left(D_j, Op_j, En_j\right)$, а D_j, Op_j, En_j – соответственно фрагменты данных, выполняемые операции и методы шифрования поддерживаемые узлом $n_j \in Node$, а D_j, Op_j фактически представляют собой отображения $d_{i,j} \leftrightarrow Op_{i,j}$ соотносящие типы фрагментов данных и операции над ними. Однако такая детализация является достаточно сложно реализуемой на практике и на данном этапе ограничимся условно универсальным разделением Op^{S-} и Op^{S+} на основе базовых операций модели данных, как приведено выше.

В итоге разделение узлов – обработчиков данных на три типа: обрабатывающих маскированные данные, зашифрованные данные или открытые данные в защищенном контуре может быть описано следующим образом:

1. Сформировать множества Op^{S-} и Op^{S+}.
2. Определить перечень операций Op, совершаемых с каждым типом фрагмента данных на узле – обработчике ($D \rightarrow \{Op_{d_i} = \{op_1 \dots op_n\}\}$).

- Если $\left(\forall \left(op_j \in Op_{d_i}\right)\right)\left(op_j \in Op^{S-}\right)$ или $Op_{d_i} \subseteq Op^{S-}$ узел вносится в список подлежащих сокрытию путем маскирования L^{Ms}.

- Если $\left(\exists\left(op_j \in Op_{d_i}\right)\right)\left(op_j \in Op^{S+}\right)$ узел вносится в список подлежащих обработке в открытом виде L^{Op}.

3. Для каждого узла $node \in L^{Ms}$
   - Если узел не соответствует критериям маскирования: ограниченный диапазон чувствительных значений, перенести в список подлежащих шифрованию L^{En}.

4. При подключении нового узла в системе или изменении графа обработки данных выполнить п.2 и 3.

Таким образом реализуется динамическая классификация узлов – обработчиков данных в процессе работы системы на три класса с точки зрения места в архитектуре безопасной обработки информации с минимизацией доверия. Более детально метод можно представить в виде основного подпроцесса классификации и подпроцесса – динамического планировщика, реализующего динамический аспект классификации на основании событий в системе.

Заключение. Сегодня с практической точки зрения при имплементации предложенного метода можно говорить скорее не об отдельных независимых узлах обработчиков данных, а об отдельных инструментах хранения и обработки информации, интегрированных в общую среду. Каждый из этих инструментов, в свою очередь, реализует замкнутый комплекс операций с данными в, как правило, распределенной среде и не имеет внутреннего механизма разделения узлов обработчиков конфиденциальных и открытых данных, либо же данных с разным уровнем конфиденциальности. Поэтому предложенная классификация в первую очередь должна применяться не к отдельным узлам, а к инструментам обработки в гетерогенной среде больших данных, сохраняя в этих условиях свою актуальность.

В заключении стоит отметить, что разработка принципов построения и архитектур безопасности в отношении гетерогенных систем больших данных сегодня высоко востребована на практике для новых цифровых систем. Интеграция подходов безопасности в отношении архитектур обработки данных, создание гибридных архитектур, в совокупности с автоматизацией отнесения узлов (и более крупных компонент) к определенному архитектурному блоку, автоматической авторизации на этой основе и применением методов защиты – ключевой аспект построения защищенных систем обработки и управления большими данными.

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

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Dumitru I.A.* Zero trust security // Proceedings of the International Conference on Cybersecurity and Cybercrime-2022. – Asociatia Romana pentru Asigurarea Securitatii Informatiei, 2022. – P. 99-104.
2. *Правиков Д.И., Щербаков А.Ю.* К вопросу об изменении парадигмы информационной безопасности // Системы высокой доступности. – 2018. – Т. 14, № 2. – С. 35-39.
3. *Малинский С.В.* Концепция безопасности Zero Trust: принципы и практика внедрения // Интеллектуальные транспортные системы. – 2022. – С. 430-437.
4. *Грызунов В.В. и др.* Обеспечение информационной безопасности интегрируемых информационных систем на базе доверия // Тр. учебных заведений связи. – 2024. – Т. 10, № 4. – С. 110-125.
5. *Полтавцева М. А., Зегжда Д.П., Калинин М.О.* Многоуровневая концепция безопасности систем управления большими данными // Вопросы кибербезопасности. – 2023. – № 5. – С. 25-36.
6. *Mishra K.N. et al.* Cloud and big data security system's review principles: A decisive investigation // Wireless Personal Communications. – 2022. – Vol. 126, No. 2. – P. 1013-1050.
7. *Awaysheh F.M. et al.* Security by design for big data frameworks over cloud computing // IEEE Transactions on Engineering Management. – 2021. – Vol. 69, No. 6. – P. 3676-3693.
8. *Stafford V.* Zero trust architecture // NIST special publication. – 2020. – Vol. 800. – 207 p.
9. *Attaallah A. et al.* Analyzing the Big Data Security Through a Unified Decision-Making Approach // Intelligent Automation & Soft Computing. – 2022. – Vol. 32, No. 2.

10. *Alani M.M.* Big data in cybersecurity: a survey of applications and future trends // Journal of Reliable Intelligent Environments. – 2021. – Vol. 7, No. 2. – P. 85-114.
11. *Wang Z., Yu X., Xue P., Qu Y., Ju L.* Research on Medical Security System Based on Zero Trust // Sensors. – 2023. – Vol. 23. – 3774. – 16 c. – DOI: 10.3390/s23073774.
12. *Daah C., Qureshi A., Awan I., Konur S.* Enhancing Zero Trust Models in the Financial Industry through Blockchain Integration: A Proposed Framework // Electronics. – 2024. – Vol. 13. – 865. – 49 p. – DOI: 10.3390/electronics13050865.
13. *Fernandez E.B., Brazhuk A.* A critical analysis of Zero Trust Architecture (ZTA) // Computer Standards & Interfaces. – 2024. – Vol. 89. – 103832. – 12 p. – DOI: 10.1016/j.csi.2024.103832.
14. *Poltavtseva M.A., Platonov V.V., Semyanov P.V.* Secure data processing architectures in big data systems. December 16-17, 2024. – 2024. – P. 104-108.
15. *Zhao Y. et al.* A zone-based data lake architecture for IoT, small and big data // Proceedings of the 25th International Database Engineering & Applications Symposium. – 2021. – P. 94-102.
16. *Awaysheh F.M. et al.* Security by design for big data frameworks over cloud computing // IEEE Transactions on Engineering Management. – 2021. – Vol. 69, No. 6. – P. 3676-3693.
17. *Roy P., Kumar R.* Multilevel Security Framework based on An Onion Encryption in Public Cloud Network // 2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N). – IEEE, 2021. – P. 1442-1446.
18. *Kuhn C. et al.* Onion routing with replies // Advances in Cryp-tology–ASIACRYPT 2021: 27th International Conference on the Theory and Application of Cryptology and Information Security, Singapore, December 6–10, 2021, Proceedings, Part II 27. – Springer International Publishing, 2021. – P. 573-604.
19. *Thirumalaisamy M. et al.* Interaction of secure cloud network and crowd computing for smart city data obfuscation // Sensors. – 2022. – Vol. 22, No. 19. – Art. 7169.
20. *Полтавцева, М.А., Калинин М.О., Зегжда Д.П.* Моделирование данных в задачах информационной безопасности поли-хранилищ // Проблемы информационной безопасности. Компьютерные системы. – 2023. – № 4 (57). – С. 122-132. – DOI: 10.48612/jisp/x468-hp82-adav.
21. *Duncan G. and Stokes L.* Data masking for disclosure limitation. // WIREs Comp Stat. – 2009. – Vol. 1. – P. 83-92. – DOI: 10.1002/wics.3.
22. *Jain R.B., Puri M.* An approach towards the development of scalable data masking for preserving privacy of sensitive business data // Artificial Intelligence and Evolutionary Computations in Engineering Systems. – Springer Singapore, 2020. – Vol. 1056. – P. 733-743. – DOI: 10.1007/978-981-15-0199-9.

REFERENCES

1. *Dumitru I.A.* Zero trust security, *Proceedings of the International Conference on Cybersecurity and Cybercrime-2022*. Asociatia Romana pentru Asigurarea Securitatii Informatiei, 2022, pp. 99-104.
2. *Pravikov D.I., Shcherbakov A.Yu.* K voprosu ob izmenenii paradigmy informatsionnoy bezopasnosti [On the issue of changing the paradigm of information security], *Sistemy vysokoy dostupnosti* [High Availability Systems], 2018, Vol. 14, No. 2, pp. 35-39.
3. *Malinskiy S.V.* Kontseptsiya bezopasnosti Zero Trust: printsipy i praktika vnedreniya [Zero Trust security concept: principles and implementation practices], *Intellektual'nye transportnye sistemy* [Intelligent Transport Systems], 2022, pp. 430-437.
4. *Gryzunov V.V. i dr.* Obespechenie informatsionnoy bezopasnosti integriruemykh informatsi-onnykh sistem na baze doveriya [Ensuring *information security of integrated information systems based on trust], Tr*. uchebnykh zavedeniy svyazi [Proceedings of educational institutions of communication], 2024, Vol. 10, No. 4, pp. 110-125.
5. *Poltavtseva M. A., Zegzhda D.P., Kalinin M.O.* Mnogourovnevaya kontseptsiya bezopasnosti sistem upravleniya bol'shimi dannymi [Multi-level concept of security of big data management systems], *Voprosy kiberbezopasnosti* [Issues of Cybersecurity], 2023, No. 5, pp. 25-36.
6. *Mishra K.N. et al.* Cloud and big data security system's review principles: A decisive investigation, *Wireless Personal Communications,* 2022, Vol. 126, No. 2, pp. 1013-1050.
7. *Awaysheh F.M. et al.* Security by design for big data frameworks over cloud computing, *IEEE Transactions on Engineering Management,* 2021, Vol. 69, No. 6, pp. 3676-3693.
8. *Stafford V.* Zero trust architecture, *NIST special publication,* 2020, Vol. 800, 207 p.
9. *Attaallah A. et al.* Analyzing the Big Data Security Through a Unified Decision-Making Approach, *Intelligent Automation & Soft Computing,* 2022, Vol. 32, No. 2.
10. *Alani M.M.* Big data in cybersecurity: a survey of applications and future trends, *Journal of Reliable Intelligent Environments,* 2021, Vol. 7, No. 2, pp. 85-114.
11. *Wang Z., Yu X., Xue P., Qu Y., Ju L.* Research on Medical Security System Based on Zero Trust, *Sensors,* 2023, Vol. 23, 3774, 16 p. DOI: 10.3390/s23073774.

12. *Daah C., Qureshi A., Awan I., Konur S.* Enhancing Zero Trust Models in the Financial Industry through Blockchain Integration: A Proposed Framework, *Electronics,* 2024, Vol. 13, 865, 49 p. DOI: 10.3390/electronics13050865.
13. *Fernandez E.B., Brazhuk A.* A critical analysis of Zero Trust Architecture (ZTA), *Computer Standards & Interfaces,* 2024, Vol. 89, 103832, 12 p. DOI: 10.1016/j.csi.2024.103832.
14. *Poltavtseva M.A., Platonov V.V., Semyanov P.V.* Secure data processing architectures in big data systems. December 16-17, 2024, 2024, pp. 104-108.
15. *Zhao Y. et al.* A zone-based data lake architecture for IoT, small and big data, *Proceedings of the 25th International Database Engineering & Applications Symposium,* 2021, pp. 94-102.
16. *Awaysheh F.M. et al.* Security by design for big data frameworks over cloud computing, *IEEE Transactions on Engineering Management,* 2021, Vol. 69, No. 6, pp. 3676-3693.
17. *Roy P., Kumar R.* Multilevel Security Framework based on An Onion Encryption in Public Cloud Network, *2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N).* IEEE, 2021, pp. 1442-1446.
18. *Kuhn C. et al.* Onion routing with replies, *Advances in Cryp-tology–ASIACRYPT 2021: 27th International Conference on the Theory and Application of Cryptology and Information Security, Singapore, December 6–10, 2021, Proceedings, Part II 27*. Springer International Pub-lishing, 2021, pp. 573-604.
19. *Thirumalaisamy M. et al.* Interaction of secure cloud network and crowd computing for smart city data obfuscation, *Sensors,* 2022, Vol. 22, No. 19, Art. 7169.
20. *Poltavtseva, M.A., Kalinin M.O., Zegzhda D.P.* Modelirovanie dannykh v zadachakh informatsionnoy bezopasnosti polikhranilishch [Data modeling in problems of information security of polystorage facilities], *Problemy informatsionnoy bezopasnosti. Komp'yuternye sistemy* [Problems of Information Security. Computer Systems], 2023, No. 4 (57), pp. 122-132. DOI: 10.48612/jisp/x468-hp82-adav.
21. *Duncan G. and Stokes L.* Data masking for disclosure limitation, *WIREs Comp Stat*., 2009, Vol. 1, pp. 83-92. DOI: 10.1002/wics.3.
22. *Jain R.B., Puri M.* An approach towards the development of scalable data masking for preserving privacy of sensitive business data, *Artificial Intelligence and Evolutionary Computations in Engineering Systems*. Springer Singapore, 2020, Vol. 1056, pp. 733-743. DOI: 10.1007/978-981-15-0199-9.

Полтавцева Мария Анатольевна – Санкт-Петербургский политехнический университет Петра Великого; e-mail: poltavtseva@ibks.spbstu.ru; г. Санкт-Петербург, Россия; тел.: +78125527632; д.т.н.; доцент; профессор института кибербезопасности и защиты информации; ORCID 0000-0001-9659-1244.

Иванов Денис Вадимович – Санкт-Петербургский политехнический университет Петра Великого; e-mail: vanov@ibks.spbstu.ru; г. Санкт-Петербург, Россия; тел.: +78125527632; к.т.н.; доцент института кибербезопасности и защиты информации; ORCID 0009-0008-7331-9721.

Poltavtseva Maria Anatolyevna – Peter the Great St. Petersburg Polytechnic University; e-mail: poltavtseva@ibks.spbstu.ru; Saint Petersburg; Russia; phone: +78125527632; dr. of eng. sc.; associate professor; professor at the Institute of Cyber Security and Information Protection; ORCID 0000-0001-9659-1244.

Ivanov Denis Vadimovich – Peter the Great St. Petersburg Polytechnic University; e-mail: vanov@ibks.spbstu.ru; Saint Petersburg; Russia; phone: +78125527632; cand. of eng. sc.; associate professor at the Institute of Cyber Security and Information Protection; ORCID 0009-0008-7331-9721.

УДК 004.7 DOI 10.18522/2311-3103-2025-3-62-81

А.М. Маевский, В.А. Рыжов, Т.А. Федорова, И.В. Кожемякин, Н.М. Буров

СТОХАСТИЧЕСКАЯ ДИНАМИЧЕСКАЯ МОДЕЛЬ ПОДВОДНОЙ БЕСПРОВОДНОЙ СЕНСОРНОЙ СЕТИ, ОСНОВАННАЯ НА ЛУВЕНСКОМ АЛГОРИТМЕ КЛАСТЕРИЗАЦИИ

Подводные беспроводные сенсорные сети (ПБСС) играют важную роль в мониторинге океанических процессов, подводной навигации, экологическом контроле и обеспечении безопасности. Однако особенности подводной среды, такие как высокая затухаемость сигналов, ограниченные ресурсы энергии и изменяющаяся топология сети, создают значительные сложности в организации эффективной передачи данных. Для оптимизации работы сети и продления ее срока службы используется метод кластеризации, позволяющий группировать узлы, снижать нагрузку