

Обеспечение долговременной сохранности юридически значимых электронных документов

А. В. Соловьев

Федеральный исследовательский центр "Информатика и управление" РАН, Москва, Россия

Аннотация. В статье рассматривается проблема обеспечения долговременной сохранности юридически значимых электронных документов. Формализована задача обеспечения долговременной сохранности юридически значимых электронных документов как задачи оптимального управления в условиях параметрических возмущений. Результатом исследования являются математические модели и алгоритмы, позволяющие, будучи реализованными в виде программно-аппаратного обеспечения, осуществлять контроль и организацию долговременной сохранности. Приведены примеры практической реализации предложенных подходов в рамках крупных информационных систем, позволяющих комплексно решить указанную проблему. Дальнейшие перспективы исследований направлены на решение проблемы долговременной сохранности больших данных и систем, построенных по технологиям распределенных реестров.

Ключевые слова: долговременная сохранность, юридическая значимость, электронный документ, распределенные реестры, аутентичность, интерпретируемость, устойчивость, надежность

DOI 10.14357/20718632230205

Введение

Электронные документы все уверенней входят во все сферы нашей жизнедеятельности, приобретая юридическую значимость. Например, эксперимент, запущенный в ноябре 2022 года, позволил использовать электронные водительские права вместо обычных. Другой эксперимент по использованию сотрудниками ГИБДД электронного свидетельства о регистрации транспортного средства успешно проходит уже 2 года. Несколько ранее 27.12.2019 вступил в силу № 480-ФЗ «О внесении изменений в Основы законодательства Российской Федерации о нотариате и отдельные законодательные акты Российской Федерации», согласно статье 44.2 которого «удостоверенный, совершенный или выданный нотариусом документ в электронной форме подписывается

квалифицированной электронной подписью нотариуса». Электронные документы (ЭД) постепенно приобретают юридическую значимость, то есть «свойство документа выступать в качестве подтверждения деловой деятельности либо событий личного характера» (ГОСТ Р 7.0.8-2013 «Национальный стандарт Российской Федерации. Система стандартов по информации, библиотечному и издательскому делу. Дело-производство и архивное дело. Термины и определения»). Согласно Программе цифровой экономики РФ «данные в цифровой форме являются ключевым фактором производства во всех сферах социально-экономической деятельности» [1, с. 1].

Поскольку цифровая экономика прямо предполагает наличие обмена цифровыми данными между государством и отраслями цифровой экономики, государством и гражданами,

граждан между собой, то к таким цифровым данным должны быть отнесены и юридически значимые электронные документы (ЮЗЭД). Однако с использованием ЮЗЭД возникают огромные проблемы. В первую очередь, это проблема долговременной сохранности ЭД. Требования к сохранности ЮЗЭД, если их необходимо хранить годы или десятилетия, безусловно, высокие. Например, срок действия водительского удостоверения – 10 лет, срок действия свидетельства о регистрации транспортного средства может измеряться десятилетиями.

Однако ЭД, хранящиеся в программно-аппаратной среде в условиях цифровой экономики находятся, с одной стороны, в условиях динамично меняющейся цифровой среды, с другой – должны быть интерпретируемы, аутентичны, защищены в течение всего срока хранения. Данная противоречивая ситуация определяет необходимость решения важной научно-технической задачи обеспечения долговременной сохранности ЮЗЭД.

1. Обзор проблемы долговременной сохранности

Попытки решения проблемы долговременной сохранности ЮЗЭД активно предпринимаются как за рубежом, так и в РФ. Для подтверждения актуальности поставленной проблемы рассмотрим совокупность существующих методов и подходов обеспечения сохранности ЮЗЭД.

Опыт Национального управления архивов и документации США (National Archives and Records Administration (NARA)), крупнейшего в мире хранилища электронных документов по организации долговременного хранения, является значительным шагом в выработке решения проблемы [2]. Государственный департамент США и Пентагон ежегодно передают в NARA на долговременное хранение десятки миллионов электронных документов. Уже одно это говорит о том, что NARA хранит необычайно ценные ЮЗЭД [3]. Накопленный опыт NARA сложно переоценить. Так, в рамках работ по созданию архива ЮЗЭД были созданы единые стандарты хранения ЭД [4, 5], такие как:

1) библиографический формат MARC,

2) формат хранения и преобразования ЭД на оптических дисках ODISS,

3) формат архивных описаний GSA 6710 A,

4) открытый стандарт кодированного архивного описания ЭД (EAD – Encoded Archival Description) на основе XML.

В 2017 году NARA выпустил документ Universal Electronic Records Management Requirements. Положения этого документа рекомендовано использовать сотрудникам федеральных агентств США при написании технических заданий для сервисов управления ЭД [6].

С 2005 года NARA прилагает усилия для создания инструментов и методов долговременного хранения ЭД. В США удалось создать сеть из более чем 20 кластеров серверов, контролируемых NARA, для хранения ЭД. Ожидается, что эти меры защитят ЭД от потери, подмены или подделки [7]. Кроме того, в настоящее время разрабатывается функция полнотекстового поиска для сохраненных редакций [8].

Опыт NARA помог выявить основные проблемы, возникающие в связи с необходимостью длительного хранения ЭД. Первая проблема – значительное количество компьютерных форматов ЭД, что создает риск не интерпретируемости (ошибочного прочтения или невозможности прочтения) ЭД по прошествии длительного времени. Несмотря на длительное функционирование электронных архивов NARA, стало очевидным, что проблема интерпретируемости не рассматривалась на старте разработки архива [9].

Второй проблемой стало технологическое и техническое старение носителей информации. Многие ЭД передавались в NARA непосредственно вместе с электронными носителями, которые впоследствии невозможно было воспроизвести [9].

Третьей проблемой стала утеря семантики переданных документов, отрыв от контекста их создания и связанных документов. В 2017 году руководство NARA сообщало о создании новой модели хранения президентских ЭД, которых за год накапливается более 250 TB [10]. Для частичного решения проблем интерпретируемости и утери семантики, руководство NARA приняло решение после 31 декабря 2022 г. принимать на хранение ЭД только вместе с описательными метаданными [11].

Вслед за NARA Национальный архив Великобритании принял активное участие в решении проблемы долговременной сохранности ЮЗЭД. В 2017 году Национальный архив Великобритании представил стратегию для создания электронных архивных решений, позволяющих обеспечить долговременное хранение ЭД, переданных из государственных учреждений [12]. Однако попытки создания электронных архивов столкнулись с проблемами долгосрочного хранения, описанными выше, в результате не было создано никакого технического решения [13]. Тем не менее, усилия по созданию программно-аппаратного решения для обеспечения долговременной сохранности ЭД привели к разработке ряда нормативных документов. Один из самых известных нормативных документов [14] регламентирует:

- порядок управления доступом к ЭД, включая вопросы аудита,
- перемещение ЭД,
- протоколирование операций с ЭД,
- управление резервным копированием ЭД.

Необходимо отметить, что ни в США, ни в Великобритании при организации долговременной сохранности не используется механизм электронной подписи. Защита ЭД от искажений предполагается в многократном территориально-распределенном дублировании хранящихся ЭД, что приводит к удорожанию аппаратной составляющей системы хранения. Несколько иной подход предлагают архивы Германии для организации долговременной сохранности ЮЗЭД. Федеральное агентство по безопасности в информатике (Bundesamt für Sicherheit in der Informationstechnik (BSI, ФРГ)) подготовило документ: "Техническое руководство BSI TR-03125" [15]. В документе приводятся рекомендации по долговременному хранению ЮЗЭД, заверенных электронной подписью (ЭП) с целью обеспечения их аутентичности, целостности, конфиденциальности. Здесь следует сделать следующее важное замечание: защита ЮЗЭД с помощью ЭП содержит недостатки, связанные с меняющимися технологиями криптозащиты, а также с тем, что сроки действия сертификатов ЭП, например в Российской Федерации, максимум 5 лет (63-ФЗ «Об электронной подписи»).

Решением проблемы долговременной сохранности занимаются и другие страны, например, одним из ведущих здесь является Австралийский национальный архив, предложивший подход к долговременному хранению ЭД, основанный на «нормализации документов», т.е. приведение их к единому цифровому формату при передаче на долговременное хранение [16]. Такой подход, безусловно, интересен, но содержит и недостатки, например, риск искажения ЭД при нормализации.

Можно привести примеры и других стран (Дании, Франции, Новой Зеландии и др.), однако, формат данной статьи не позволяет сделать подробный обзор.

В РФ сравнительно недавно начали заниматься вопросом долговременной сохранности ЭД. В настоящее время Росархив совместно с Всероссийским научно-исследовательским институтом документоведения и архивного дела (ВНИИДАД) занимается разработкой документов, регламентирующих организацию долговременного хранения ЭД в рекомендательном порядке [17, 18]. В частности, данными рекомендациями предписывается нормализация ЭД при длительном хранении в формат PDF/A-1, что, правда, сразу сокращает количество форматов хранения до одного единственного, причем строго определенной версии. Учитывая то, что разработчик формата находится в юрисдикции США, несмотря на то, что формат открытый, это может вызвать риски поддержки, как самого формата, так и средств его интерпретации.

Из приведенного краткого обзора можно сделать следующие выводы:

- 1) проблема долговременной сохранности ЮЗЭД является крайне актуальной,
- 2) проблема далека от решения, т.к. отсутствует универсальное тиражируемое программно-аппаратное решение обеспечивающее сохранность ЮЗЭД.

Отсутствие программно-аппаратного решения связано с несколькими причинами:

- проблема не является чисто технической, а носит междисциплинарный характер;
- отсутствует системный подход к решению проблем долговременной сохранности ЮЗЭД;

- не все проблемы долговременной сохранности до конца изучены и систематизированы;
- отсутствует методология контроля параметров сохранности ЮЗЭД.

Тем не менее, по результатам краткого обзора можно сформулировать основные проблемы долговременной сохранности:

- проблема аутентичности (неизменности) ЮЗЭД;
- проблема интерпретируемости (читаемости, визуализации) ЮЗЭД;
- проблема изменения (старения) программно-аппаратной среды хранения ЮЗЭД;
- проблема надежности программно-аппаратной среды хранения ЮЗЭД;
- проблема информационной безопасности программно-аппаратной среды хранения ЮЗЭД.

2. Формализация задачи обеспечения долговременной сохранности юридически значимых электронных документов

Перед тем, как перейти к формальной постановке задачи обеспечения долговременной сохранности ЮЗЭД, введем ряд важных определений. *Долговременное хранение* – обеспечение сохранности ЮЗЭД в течение не менее 5 лет. *Сохранность* – свойство ЮЗЭД существовать в качестве доступного и аутентичного свидетельства (доказательства) в произвольный момент времени.

При переходе к цифровой экономике сроки хранения ЮЗЭД могут намного превосходить сроки функционирования программно-аппаратной среды их хранения. Так, например, документы по личному составу должны храниться 50 лет (до 2003 года – 75 лет). Особо ценные ЭД могут храниться еще дольше. Программно-аппаратная среда хранения ЮЗЭД будет многократно изменена за это время. Несколько раз будет обновлено как программное обеспечение поиска, индексации, хранения и отображения ЭД, так и носители информации.

При долговременном хранении в условиях цифровой экономики ЮЗЭД становятся объектом управления, т.к. требуется их стабилизация

в условиях меняющейся среды хранения. Тогда задача долговременной сохранности будет представлять собой задачу оптимального управления в условиях параметрических возмущений [19, 20].

Для решения задачи оптимального управления необходимо научиться контролировать параметры среды хранения ЮЗЭД, разработать алгоритмы компенсации параметрических возмущений для стабилизации объекта управления [21], т.е. ЮЗЭД. По сути, такой подход используется при создании цифровых двойников, которые позволяют контролировать объекты-оригиналы, предсказывать и своевременно решать выявленные проблемы [22, 23].

В данном исследовании выявлены основные проблемы долговременной сохранности, что позволяет сформулировать математическую постановку задачи обеспечения сохранности ЮЗЭД в следующем виде.

Дано:

- 1) Множество ЮЗЭД $UZD = \{uzd_i\}$.
- 2) Множество параметрических возмущений среды хранения UZD : $E = \{\varepsilon_q\}$:
 - a) ε_1 – нарушение аутентичности UZD ,
 - b) ε_2 – нарушение интерпретируемости UZD ,
 - c) ε_3 – изменения программно-аппаратной среды UZD ,
 - d) ε_4 – нарушение надежности хранения UZD ,
 - e) ε_5 – нарушение информационной безопасности UZD .
- 3) Множество информационных систем $IS = \{IS_k\}$ подверженных E и осуществляющих хранение UZD .

Найти:

- 1) Множество математических моделей $\mu = \{\mu_r\}$, контроля параметров сохранности UZD .
- 2) Множество алгоритмов $A = \{A_r\}$, обеспечивающих сохранность UZD на заданном исходном уровне Lev_0 : $A_r(d_i) = Lev_0$.
- 3) Множество программно-аппаратных реализаций алгоритмов A : $S = \{S_r\}$, обеспечивающих сохранность UZD на исходном уровне Lev_0 : $S_r(uzd_i) = Lev_0$.

3. Математические модели контроля параметров сохранности

Если ЮЗЭД становятся при формализации постановки задачи обеспечения сохранности объектом управления, то задача контроля сохранности ЮЗЭД как объектов управления должна представлять собой задачу оптимального выбора по многим критериям.

В данном случае критериями становятся параметрические возмущения среды хранения ЮЗЭД или же связанные с ними соотношением $\mu_r = 1 - \varepsilon_q$ параметры сохранности μ_r .

Действительно, пусть сохранность ЮЗЭД uzd_i характеризуется функцией $\mu(t)$, значение которой представляет собой вероятность сохранности ЮЗЭД uzd_i на уровне Lev_0 в произвольный момент времени t . Тогда можно утверждать, что задача обеспечения сохранности ЮЗЭД uzd_i формулируется, как задача достижения максимума функции $\mu(t)$ на произвольном временном интервале t .

Или максимум функции $\mu(t)$: $M = \max_{t \in [t_0, \infty]} \mu(t)$, где t_0 – момент времени создания uzd_i в некоторой исходной информационной системе IS_0 .

Тогда $\mu(t) = \alpha(t)^{\omega_1} \zeta(t)^{\omega_2} \varphi(t)^{\omega_3} \rho(t)^{\omega_4} \sigma(t)^{\omega_5}$, где $\sum \omega_i = 1$, $\omega_i > 0$, $i = [1, 5]$, где $\alpha(t)$, $\zeta(t)$, $\varphi(t)$, $\rho(t)$, $\sigma(t)$ – соответственно оценки вероятностей сохранения аутентичности, интерпретируемости, устойчивости к изменениям (старению) программно-аппаратной среды хранения, надежности, устойчивости к угрозам нарушения информационной безопасности UZD .

Математические модели контроля параметров сохранности ЮЗЭД, приведенные ниже.

3.1. Математическая модель оценки аутентичности

Сохранение аутентичности (неизменности) ЮЗЭД в течение всего срока хранения является важнейшей задачей обеспечения долговременной сохранности.

Предположим, что мы нашли способ подтвердить аутентичность ЮЗЭД с помощью ЭП. Ниже покажем, что такой способ возможен. Тогда нужно оценить вероятность не взлома ЭП ($\alpha_1(t)$), а также вероятность того, что ЭП юридически значима, т.е. «свойство

электронной подписи, обеспечивающее условия признания электронного документа, подписанного электронной подписью, равнозначного документу на бумажном носителе, подписанном собственноручной подписью» (ГОСТ Р 7.0.8-2013) не утрачено ($\alpha_2(t)$). Тогда оценку сохранения аутентичности ЮЗЭД можно представить следующими соотношениями:

$$\alpha_1(t) = 1/N_{DS} \sum_{i=1, N_{DS}} (1/2 - 2^{n-1} \prod_{ii=1, n} (p_{ii}(t) - 1/2)) \quad (1)$$

$$\alpha_2(t) = 1/N_{DS} \sum_{i=1, N_{DS}} (Vf_i(t) (N_{sert_dbi}(t)/N_{serti}) \sum_{j=1, N_{serti}} S_{true_ij}(t) ((T_{dij} - T_{dcij}(t))/T_{dij})) \quad (2)$$

$$\alpha(t) = \min(\alpha_1(t), \alpha_2(t)), \quad (3)$$

где: $p_{ii}(t)$ – вероятность подбора правильной комбинации одного произвольного бита пары ключ-шифртекст длины n ,

$Vf_i(t)$ – актуальность i -й ЭП на момент времени t ,

T_{dij} – срок действия j -го сертификата i -й ЭП,

$T_{dcij}(t)$ – время, прошедшее с начала действия сертификата,

$S_{true_ij}(t)$ – актуальность j -го сертификата i -й ЭП на момент времени t ,

N_{DS} – общее количество сертификатов ЭП,

$N_{sert_dbi}(t)$ – количество доступных сертификатов i -й ЭП на момент времени t ,

N_{serti} – требуемое количество сертификатов i -й ЭП.

3.2. Математическая модель оценки интерпретируемости

Как было показано выше, при обеспечении долговременной сохранности ЮЗЭД возникает проблема интерпретируемости и отображения данных спустя десятилетия. Иными словами, должна быть обеспечена высокая вероятность раскодирования ЮЗЭД через десятилетия для возможности отобразить ЮЗЭД на мониторе, распечатать и т.д. Для решения проблемы потери семантики вместе с ЮЗЭД необходимо хранить описательные метаданные, которые тоже должны быть интерпретированы (прочитаны) спустя годы.

Тем самым, необходимо, во-первых, контролировать сохранение интерпретируемости ЮЗЭД, во-вторых, обеспечить интерпретируемость всеми доступными средствами, напри-

мер, с помощью нормализации ЮЗЭД в подходящий формат долговременного хранения.

Для того, чтобы понять какую информацию о ЮЗЭД нужно хранить, необходимо определить модель ЮЗЭД долговременного хранения с точки зрения состава информации. Для начала, необходимо понимание, что такое документ при долговременном хранении. Введем необходимые определения.

Документ – структурированная информация, представляющая собой совокупность взаимосвязанных семантических блоков.

ЭД – документ, семантические блоки которого и взаимосвязи между ними представлены в цифровой форме.

Семантические блоки — некоторые фрагменты ЭД, выделенные по смысловому содержанию.

Для контроля интерпретируемости ЮЗЭД была разработана модель ЮЗЭД долговременной сохранности (4) и модель контроля (5) вероятности интерпретации ЮЗЭД ($\zeta(t)$):

$$uzdi = MD \cup OrD \cup OdfD \cup FTI \cup CLI \cup LDI \cup OpD \quad (4)$$

$$\zeta(t) = \zeta OdfD(t)^{k_1} \zeta OrD(t)^{k_2} \zeta MD(t)^{k_3} \zeta CLI(t)^{k_4} \zeta OpD(t)^{k_5} \zeta FTI(t)^{k_6} \zeta LDI(t)^{k_7}, \quad (5)$$

где k_i – весовые коэффициенты важности (назначаются экспертами) $\sum k_i = 1$, $k_i > 0$, $i=[1,7]$. На основании практического применения рекомендуется назначать следующие значения $k_1 = k_2 = 0,1$, $\sum_{i=[3,7]} k_i = 0,8$, поскольку интерпретируемость оригинала (*OrD*) ЮЗЭД (*uzdi*) и нормализованной копии (*OdfD*) являются наиболее критичными.

Семантические блоки документа при долговременном хранении:

MD – метаданные ЮЗЭД,

OrD – оригинал ЮЗЭД,

OdfD – нормализованная копия ЮЗЭД,

FTI – нормализованный текст ЮЗЭД для полнотекстового индекса,

CLI – классификаторы ЮЗЭД,

LDI – описание связей с другими ЮЗЭД,

OpD – журнал операций с ЮЗЭД.

Обозначение вероятностей интерпретации математической модели:

$\zeta OdfD(t)$ – для *OdfD*,

$\zeta OrD(t)$ – для *OrD*,

$\zeta MD(t)$ – для *MD*,

$\zeta CLI(t)$ – для *CLI*,

$\zeta OpD(t)$ – для *OpD*,

$\zeta FTI(t)$ – для *FTI*,

$\zeta LDI(t)$ – для *LDI*.

Важность разработанной модели подчеркивается также необходимостью, в том числе, и автоматического анализа структуры и семантики ЭД (например [24]).

3.3. Математическая модель оценки устойчивости к изменениям программно-аппаратной среды хранения

Как при старении носителей, так и при обновлении программно-аппаратной среды возникает необходимость корректной миграции ЮЗЭД. Существенным риском при этом является то, что высока вероятность потери ЮЗЭД по халатности или злему умыслу. При миграции важно обратить внимание на вероятность переноса метаданных ЮЗЭД, классификаторов, журнала операций, связей с другими документами, чтобы избежать потери ценной описательной информации и не оторвать ЮЗЭД от семантического контекста, в котором он был создан.

На основании модели ЮЗЭД (4) были разработаны алгоритм контроля среды хранения, а также математическая модель оценки устойчивости (6).

$$\varphi(t) = \varphi OdfD(t)^{k_1} \varphi OrD(t)^{k_2} \varphi MD(t)^{k_3} \varphi CLI(t)^{k_4} \varphi OpD(t)^{k_5} \varphi FTI(t)^{k_6} \varphi LDI(t)^{k_7}, \quad (6)$$

где k_i – назначаемые экспертами весовые коэффициенты важности $\sum k_i = 1$, $k_i > 0$, $i=[1,7]$.

Обозначения вероятностей для модели оценки устойчивости:

$\varphi OdfD(t)$ – для *OdfD*,

$\varphi OrD(t)$ – для *OrD*,

$\varphi MD(t)$ – для *MD*,

$\varphi CLI(t)$ – для *CLI*,

$\varphi OpD(t)$ – для *OpD*,

$\varphi FTI(t)$ – для *FTI*,

$\varphi LDI(t)$ – для *LDI*.

3.4. Математическая модель оценки надежности хранения

ЮЗЭД (*uzdi*) неотделим от программно-аппаратной среды хранения. Поэтому важным при организации долговременного хранения является контроль надежности программно-

аппаратной среды хранения. Модель надежности достаточно подробно описана в [25]. Здесь же кратко можно сказать, что автором доказано, что схема надежности информационной системы всегда может быть представлена в виде дерева или леса.

Приведем в качестве примера модель оценки надежности трехуровневой иерархической информационной системы (7):

$$\rho(t) = K_{rIS} = K_{r tl} \sum_{i=1}^{N_{ml}} (b_i K_{r cc ml i} K_{r ml i} (\sum_{j=1}^{N_{ll i}} a_{ij} K_{r cc ll ij} K_{r ll ij})), \quad (7)$$

где: $K_{r tl}$ – показатель надежности программно-аппаратного комплекса центрального уровня (дата-центр, центральный серверный комплекс и др.),

$b_i = Ob_i / \sum_{i=1}^{N_{ml}} Ob_i$ – доля объектов регионального уровня i -го региона,

Ob_i – количество объектов регионального уровня i -го региона ($\sum_{i=1}^{N_{ml}} b_i = 1$),

N_{ml} – количество объектов регионального уровня,

$K_{r cc ml i}$ – показатель надежности канала связи i -го региона с комплексом центрального уровня,

$K_{r ml i}$ – показатель надежности программно-аппаратного комплекса i -го региона,

$a_{ij} = Ob_{ij} / \sum_{j=1}^{N_{ll i}} Ob_{ij}$ – доля объектов районного уровня j -района i -го региона ($\sum_{j=1}^{N_{ll i}} a_{ij} = 1$),

Ob_{ij} – количество объектов районного уровня j -района i -го региона,

$N_{ll i}$ – количество объектов районного уровня i -го региона,

$K_{r cc ll ij}$ – показатель надежности канала связи j -района и i -го региона,

$K_{r ll ij}$ – показатель надежности программно-аппаратного комплекса j -района i -го региона.

3.5. Математическая модель оценки устойчивости к угрозам нарушения информационной безопасности

Можно определить устойчивость ЮЗЭД к угрозам нарушения информационной безопасности как вероятность сохранения аутентичности, надежности, интерпретируемости ЮЗЭД в случае реализации внешнего воздействия (реализация угроз информационной безопасности).

Тогда математическая модель оценки вероятности сохранения устойчивости ЮЗЭД к

угрозам нарушения информационной безопасности ($\sigma(t)$) может быть представлена в следующем виде (8):

$$\sigma(t) = p_{isc}(t)^{v_1} (1 - p_{cr}(t))^{v_2} p_{hra}(t)^{v_3}, \quad (8)$$

где: $\sum v_i = 1$, $v_i > 0$, $i=[1,3]$,

$p_{isc}(t)$ – вероятность реализации мер противодействия угрозам информационной безопасности,

$p_{cr}(t)$ – вероятность потери критической информации,

$p_{hra}(t)$ – оценка влияния человеческого фактора.

4. Алгоритмы обеспечения долговременной сохранности

В рамках данного исследований был разработан алгоритмический аппарат, позволяющий, в случае его реализации в виде программно-аппаратных решений $S = \{ S_r \}$, решить важнейшие проблемы обеспечения долговременной сохранности ЮЗЭД, такие как аутентичность, интерпретируемость, а также частично (в силу зависимости от организационных мер) устойчивость и надежность. К сожалению, формат статьи не позволяет привести полностью схемы этих алгоритмов, поэтому ограничимся их кратким описанием.

4.1. Алгоритм инвентаризации электронной подписи

Из приведенного выше краткого обзора проблемы, можно сделать вывод, что существуют два подхода к решению проблемы сохранения аутентичности ЮЗЭД.

Первый подход связан с организацией многократного дублирования оригинала, защищенного организационными мерами, ЭП не применяется (пример, NARA).

Второй подход связан с использованием криптозащиты. Это использование ЭП как таковой, либо же защищенных криптозащитой распределенных реестров (например, технологий блокчейн). Однако, использование технологий блокчейн будет связано во многом с доверием средствам криптозащиты иностранного производства, а также с тем, что за время длительного хранения технологии блокчейн также

изменяться и, следовательно, подтвердить аутентичность старых документов будет крайне сложно. Тем не менее, второй подход представляется предпочтительным, т.к., во-первых, отечественные средства криптозащиты достаточно развиты, во-вторых, с их помощью можно организовать заверение не только ЮЗЭД, но и журналов операций над ними.

Исходя из этих соображений, был предложен алгоритм инвентаризации ЭП для ЮЗЭД. Алгоритм предполагает использование механизма ЭП как независимое средство обеспечения аутентичности (неизменности) ЮЗЭД. Алгоритм предполагает периодическое перезаверение ЮЗЭД с помощью новой ЭП. При этом должны сохраняться сведения об авторстве сертификата старой ЭП. Кроме того, подлинность ЭП должна быть подтверждена электронным нотариатом.

Приведем последовательность шагов алгоритма.

1) Выполнить оценку вероятности нарушения аутентичности ЮЗЭД с помощью математической модели (3).

2) Зафиксировать факт нарушения в журнале аудита подлинности, если значение вероятности $\alpha(t)$ меньше некоторого заданного порогового значения αT .

3) Верифицировать существующую ЭП действующими средствами криптографической защиты.

4) Зафиксировать факт положительной проверки ЭП в журнале проверки подлинности в случае подтверждения подлинности ЭП.

5) Заверить запись в журнале аудита ЭП с указанием авторства из сертификата ЭП.

6) Извлечь данные об авторстве сертификата ЮЗЭД действующей ЭП.

7) Заверить ЮЗЭД новой ЭП.

8) Вписать в новую ЭП данные об авторе сертификата старой ЭП ЮЗЭД.

9) Сохранить старую ЭП ЮЗЭД.

10) Зафиксировать связь с новой ЭП ЮЗЭД по авторству из сертификата ЭП.

11) Зафиксировать факт заверения ЮЗЭД новой ЭП в журнале проверки подлинности.

12) Заверить запись в журнале проверки с помощью ЭП с указанием авторства лица, проводившего проверку.

13) Зафиксировать факт нарушения аутентичности ЮЗЭД в журнале проверки подлинности, если хотя бы одна проверка 1) - 12) не дала положительного результата.

В случае установления факта нарушения сохранности ЮЗЭД требуется детальное расследование причин нарушения аутентичности ЭД.

Предложенный подход, по сути, является промежуточным между чистым использованием ЭП и технологиями распределенных реестров, т.к. позволяет, с одной стороны, учитывать изменения средств криптозащиты, с другой, в предложенном подходе предполагается формирование взаимосвязанных цепочек из ЭП, каждая из которых действительна на момент проверки (подтверждения подлинности). Если предположить, что изменения средств криптозащиты происходят не скачкообразно, и в каждый конкретный момент времени в течение календарного года доступны средства проверки старых ЭП, то алгоритм инвентаризации можно проводить не реже 1 раз в год или чаще по мере необходимости.

4.2. Алгоритм инвентаризации носителей информации

В настоящее время не существует носителей для хранения ЮЗЭД, способных без проблем функционировать в течение десятилетий. Кроме того, существует еще проблема технологического старения, когда абсолютно сохраненный носитель невозможно будет со временем прочитать по причине отсутствия читающих устройств. Причина этого в том, что производителям ИТ-технологий со временем становится экономически невыгодно поддерживать старые, слабо востребованные технологии и устройства хранения.

Автором предложен алгоритм инвентаризации носителей информации для решения проблемы изменения (старения) программно-аппаратной среды. Важным допущением алгоритма является то, что ЮЗЭД хранится в нескольких копиях на нескольких независимых носителях, как это принято в регламенте резервного хранения ЭД [17, 18].

Приведем последовательность шагов алгоритма.

1) Проверить работоспособность устройства хранения.

2) Перейти к шагу 3), если устройство неисправно; перейти к шагу 5), если устройство в порядке.

3) Найти резервное устройство.

4) Зафиксировать факт потери ЮЗЭД в журнале аудита, если устройство не найдено - окончание алгоритма; перейти к шагу 1), если устройство найдено.

5) Проверить возможность миграции (устойчивости к изменениям) ЮЗЭД по модели (6).

6) Перейти к шагу 8), если миграция возможна; перейти к шагу 3), если это невозможно.

7) Зафиксировать результат проверки в журнале аудита. Подтвердить результат с помощью ЭП верификатора.

8) Проверить интерпретируемость ЮЗЭД по модели (5).

9) Перейти к шагу 11), если ЮЗЭД интерпретируется; перейдите к шагу 3), если нет.

10) Зафиксировать результат проверки в журнале аудита. Подтвердить результат с помощью ЭП верификатора.

11) Проверить подлинность ЮЗЭД по алгоритму инвентаризации ЭП.

12) Перейти к шагу 14), если ЮЗЭД является аутентичным (подлинным); перейти к шагу 3), если он не является подлинным.

13) Зафиксировать результат проверки в журнале аудита. Подтвердить результат с помощью электронной подписи верификатора.

14) Создать необходимое количество копий ЮЗЭД на новых носителях, если часть носителей повреждена (см. неуспешные проверки 2), 6), 9) 12)).

15) Проверить интерпретируемость, аутентичность и стабильность после создания копий.

16) Зафиксировать результат проверки в журнале аудита. Подтвердить результат с помощью ЭП верификатора.

5. Реализации предложенных математических моделей и алгоритмов

Наиболее полно предложенные математические модели и алгоритмы обеспечения долговременной сохранности ЮЗЭД реализованы в электронном архиве Пенсионного Фонда РФ, который функционирует более 15 лет в 80 регионах РФ и позволяет организовывать защи-

щенное долговременное хранение ЮЗЭД. Сроки хранения документов в этом архиве составляют от 5 до 75 лет. Внедрение электронного архива позволило в 10 раз повысить оперативность поиска ЮЗЭД и организовать надежное хранение свыше 500 млн. ЮЗЭД (свыше 50% документов фонда), заверенных ЭП, поступающих от работодателей и ФНС. До сих пор не отмечено нарушение аутентичности и интерпретируемости ЮЗЭД. В несколько раз сократились арендуемые площади для хранения документов, т.к. более половины из них – ЮЗЭД.

На основе предложенных математических моделей и алгоритмов обеспечения долговременной сохранности ЮЗЭД созданы другие программные реализации и методическое обеспечение для АО «Газпромбанк», ООО «Юридическая фирма Городисский и Партнеры», ООО «Когнитивные технологии».

Заключение

В данной статье рассмотрены проблемы организации длительного хранения юридически значимых электронных документов (ЮЗЭД). Задача обеспечения долговременной сохранности формулируется как задача оптимального управления в условиях параметрических возмущений. Предложены математические модели и алгоритмы, позволяющие контролировать параметры среды хранения (аутентичность, интерпретируемость, устойчивость, надежность) ЮЗЭД и организовывать их сохранность (аутентичность и устойчивость к изменениям среды хранения).

Предложенные математические модели и алгоритмы доведены до уровня программно-методических решений. Эти решения реализованы и используются в промышленной эксплуатации в рамках крупных территориально-распределенных информационных систем.

Преимуществом представленных подходов и алгоритмов является то, что они позволяют комплексно решать научно-техническую проблему организации долговременной сохранности ЮЗЭД. На основании предложенных подходов и алгоритмов можно строить программные решения, что подтверждено практикой, способные проводить автоматизированную количественную оценку параметров

сохранности ЮЗЭД, комплексно решать такие проблемы сохранности ЮЗЭД, как обеспечение аутентичности, интерпретируемости, надежности и устойчивости.

В перспективе планируется развить математический и алгоритмический аппарат для решения проблемы долговременной сохранности больших данных и систем, построенных по технологиям распределенных реестров.

Литература

1. Программа «Цифровая экономика Российской Федерации». Утверждена распоряжением Правительства Российской Федерации от 28 июля 2017 г. № 1632-р. М.: 2017 – 88 с.
2. Рысков, О.И. Об основных направлениях деятельности зарубежных архивных органов в области исследования и нормативного регулирования работы с электронной документацией // Секретарское дело. – 2005. – № 3. – С. 76.
3. Афанасьева, Л.П. Автоматизированные архивные технологии // Федеральное агентство по образованию. Государственное Образовательное учреждение высшего профессионального образования Российский Государственный Гуманитарный университет – М.: 2005. – С. 114.
4. Рысков О.И. Основные направления деятельности национальных архивов США и Соединенного Королевства Великобритании и Северной Ирландии в области управления электронными документами правительственных учреждений // Отечественные архивы. – 2004. – № 3.
5. US National Archives Blog. 2019, URL: <http://blogs.archives.gov/records-express/2013/11/01/opportunity-for-comment-transfer-guidance-bulletin/> (дата обращения 20.01.2023).
6. U.S. National Archives and Records Administration. 2017. Universal Electronic Records Management (ERM) Requirements, URL: <https://www.archives.gov/records-mgmt/policy/universalermmrequirements> (дата обращения 20.01.2023).
7. Miller, J. NARA to suspend development of ERA starting in 2012 [Электронный ресурс] / J. Miller, FederalNewsRadio.com, URL: <http://www.federalnewsradio.com/?sid=2204570&nid=35> (дата обращения 23.01.2023).
8. Lipowicz, A. NARA officials defend searchability of electronic archive [Электронный ресурс] / A. Lipowicz // Federal Computer Week. URL: <http://fcw.com/articles/2011/11/01/nara-officials-defending-searchability-of-electronic-archive.aspx> (дата обращения 23.01.2023).
9. Carlstrom, G. Is DoD's new pay system fair? [Электронный ресурс] G. Carlstrom // FederalTimes.com, URL: <http://federaltimes.com/index.php?S=3502888> (дата обращения 24.01.2023).
10. National Archives Announces a New Model for the Preservation and Accessibility of Presidential Records. U.S. National Archives and Records Administration [Электронный ресурс] – 2017 – URL: <https://www.archives.gov/press/press-releases/2017/nr17-54> (дата обращения 25.01.2023).
11. Draft National Archives Strategic Plan. U.S. National Archives and Records Administration [Электронный ресурс] – 2017 – URL: <https://www.archives.gov/about/plans-reports/strategic-plan/draft-strategic-plan> (дата обращения 25.01.2023).
12. Суровцева Н.Г. Хранение электронных документов: зарубежный опыт // Вестник культуры и искусства. – 2017. – №4(52). – с. 17-23.
13. Open Government Partnership UK National Action Plan 2013 to 2015. London SW1A 2AS – 2013 – 58 p.
14. Typical requirements for automated electronic document management systems. Specification MoReq // Office for Official Publications of the European Communities as INSAR Supplement VI, ISBN 92-894-1290-9.
15. Preservation of Evidence of Cryptographically Signed Documents // BSI Technical Guideline TR-03125 – Version 1.2 – Federal Office for Information Security – 2015 – 183 p.
16. Netwerk Digitaal Erfgoed n.d. [Электронный ресурс] – URL: <https://netwerkdigitaalerfgoed.nl/> (дата обращения 26.01.2023).
17. Рекомендации по комплектованию, учету и организации хранения электронных архивных документов в архивах организаций. – М.: ВНИИДАД, 2013 – 58 с.
18. Рекомендации по комплектованию, учету и организации хранения электронных архивных документов в государственных и муниципальных архивах. – М.: ВНИИДАД, 2013 – 33 с.
19. Емельянов С.В. Системы автоматического управления с переменной структурой. – М.: Наука, 1967 г. – 336 с.
20. Емельянов С.В., Костылева, Н.Е., Матич, Б.Л., Миловилов, Н.Н. Системное проектирование средств автоматизации. – М.: Машиностроение, 1978.
21. Емельянов С.В. Новые типы обратной связи. – М.: Наука, Физматлит, 1997. – 352 с.
22. Кириллов Д.С., Барчукова, Т.А. Цифровые двойники как основа цифровой трансформации промышленных предприятий // В сборнике: Актуальные вопросы экономики и управления. 2021. с. 161-164.
23. Рахманов М.Л., Шишкин, А.В. Современные цифровые технологии и цифровой двойник // Качество и жизнь. 2021. № 2 (30). с. 57-59.
24. Grigoriev, O.G. et al. TITANIS: A Tool for Intelligent Text Analysis in Social Media // In: Kovalev S.M., Kuznetsov S.O., Panov A.I. (eds) Artificial Intelligence. RCAI 2021. Lecture Notes in Computer Science, vol 12948. pp. 232-247. Springer, Cham. DOI: https://doi.org/10.1007/978-3-030-86855-0_16.
25. Solovyev, A.V. Reliability Assessment Method for Geographically Distributed Information Systems / G.P. Akimova, A.V.Solovyev, I.A. Tarkhanov // The IEEE 12th International Conference on Application of Information and Communication Technologies / AICT 2018 (17-19 Oct. 2018, Almaty, Kazakhstan), IEEE, 2018, P.188-191.

Соловьев Александр Владимирович. Федеральное государственное учреждение "Федеральный исследовательский центр "Информатика и управление" Российской академии наук" Москва, Россия. Главный научный сотрудник, доктор технических наук. Область научных интересов: системный анализ, системы управления базами данных, теория надежности, математическое моделирование, долговременное хранение электронных документов. E-mail: soloviev@isa.ru

Ensuring the Long-Term Preservation of a Legally Significant Electronic Document

A. V. Solovyev

Federal Research Center "Computer Science and Control" of Russian Academy of Sciences, Moscow, Russia

Abstract. The article deals with the problem of ensuring the long-term preservation of legally significant electronic documents. A review of the problem and existing options for its solution is made, the problem of ensuring the long-term preservation of legally significant electronic documents is formalized as an optimal control problem under parametric disturbances. Document safety indicators and mathematical models for evaluating safety indicators are proposed. The result of the study are algorithms that allow, being implemented in the form of software and hardware, to control and organize long-term preservation. Examples of the practical implementation of the proposed approaches within the framework of large information systems are given. The advantage of the proposed algorithms and approaches is that their implementation will make it possible to comprehensively solve the important scientific and technical problem of organizing the long-term preservation of legally significant electronic documents. At the end of the article, further prospects for research aimed at solving the problem of long-term preservation of big data and systems built using distributed ledger technologies are identified.

Keywords: long-term preservation, legal significance, electronic document, distributed registries, authenticity, interpretability, stability, reliability.

DOI 10.14357/20718632230205

References

1. The Government of the Russian Federation. 2017. PROGRAM "Digital Economy of the Russian Federation" approved by the order of at July 28, 2017. № 1632-р: 88.
2. Ryskov O.I. 2005. "On the main activities of foreign archival bodies in the field of research and regulatory work with electronic documentation". Sekretarskoye delo. № 3: 76.
3. Afanasyeva L.P. 2005. "Automated Archive Technologies". Federal Agency for Education. State Educational Institution of Higher Professional Education Russian State University for the Humanities: 114.
4. Ryskov O.I. 2004. "The main activities of the national archives of the United States and the United Kingdom of Great Britain and Northern Ireland in the field of management of electronic documents of government agencies". Otechestvenniye Arhivy № 3.
5. US National Archives Blog. 2019. <http://blogs.archives.gov/records-express/2013/11/01/opportunity-for-comment-transfer-guidance-bulletin/>
6. U.S. National Archives and Records Administration. 2017 a. Universal Electronic Records Management (ERM) Requirements, <https://www.archives.gov/records-mgmt/policy/universalemrequirements>
7. Miller J. n.d. "NARA to suspend development of ERA starting in 2012" FederalNewsRadio, <http://www.federalnewsradio.com/?sid=2204570&nid=35>
8. Lipowicz A. n. d. "NARA officials defend searchability of electronic archive" Federal Computer Week, <http://fcw.com/articles/2011/11/01/nara-officials-defending-searchability-of-electronic-archive.aspx>
9. Carlstrom G. 2017. "Is DoD's new pay system fair?" Federal Times, <http://federaltimes.com/index.php?S=3502888>
10. U.S. National Archives and Records Administration. 2017 b. National Archives Announces a New Model for the Preservation and Accessibility of Presidential Records, <https://www.archives.gov/press/press-releases/2017/nr17-54>
11. U.S. National Archives and Records Administration. 2017 c. Draft National Archives Strategic Plan, <https://www.archives.gov/about/plans-reports/strategic-plan/draft-strategic-plan>

12. Suvorovtseva N.G. 2017. Storage of electronic documents: foreign experience *Vestnik kultury I iskusstva*. №4(52): 17 – 23.
13. UK Open Government Partnership. 2013. National Action Plan 2013 to 2015: 58.
14. Office for Official Publications of the European Communities as INSAR Supplement VI. Typical requirements for automated electronic document management systems. Specification MoReq. ISBN 92 894 1290 9.
15. Federal Office for Information Security. 2015. Preservation of Evidence of Cryptographically Signed Documents. BSI Technical Guideline TR-03125. Version 1.2: 183.
16. Netwerk Digitaal Erfgoed n.d. <https://netwerkdigitaalerfgoed.nl/>
17. VNIIDAD. 2013. Recommendations for the acquisition, accounting and organization of storage of electronic archival documents in the archives of organizations.
18. VNIIDAD. 2013. Recommendations for the acquisition, accounting and organization of storage of electronic archival documents in state and municipal archives.
19. Emelyanov S. 1967. Automatic control systems with variable structure. *Science*.
20. Emelyanov S., Kostileva N., Matich B. and Milovidov N. 1978. System design automation. *Mashinostroyeniye*.
21. Emelyanov S. 1997. New types of feedback. *Science, Fizmatlit*.
22. Kirillov D.S., Barchukova T.A. 2021. Digital twins as the basis for the digital transformation of industrial enterprises. *Actual issues of economics and management*: 161-164.
23. Rakhmanov M.L., Shishkin A.V. 2021. Modern digital technologies and digital twin. *Kachestvo I zhizn*. 2(30): 57-59.
24. Grigoriev O.G. et al. 2021. TITANIS: A Tool for Intelligent Text Analysis in Social Media. *Artificial Intelligence. RCAI 2021. Lecture Notes in Computer Science*. Springer, Cham. 12948: 232-247. DOI: https://doi.org/10.1007/978-3-030-86855-0_16.
25. Akimova G., Solovyev A. and Tarkhanov, I. 2018. Reliability Assessment Method for Geographically Distributed Information Systems. The IEEE 12th International Conference on Application of Information and Communication Technologies: AICT, 188-191, 2018, Almaty, Kazakhstan. <https://doi.org/10.1109/ICAICT.2018.8747055>.

Solovyev A. V. Chief Researcher, Doctor of Technical Sciences. Federal Research Center “Computer Science and Control” of Russian Academy of Sciences, 44/2 Vavilova str., Moscow, 119333, Russia. E-mail: soloviev@isa.ru