ПОСТРОЕНИЕ ВЫЧИСЛИТЕЛЬНОЙ ПРОЦЕДУРЫ КОМПЛЕКСНОЙ ПРОВЕРКИ ПОДЛИННОСТИ ДАННЫХ ДЗЗ

Кузнецов А.В., Мясников В.В.
Институт систем обработки изображений РАН

Аннотация

Рассматривается задача построения вычислительной процедуры комплексной проверки подлинности данных дистанционного зондирования Земли (ДЗЗ) с использованием набора элементарных алгоритмов проверки подлинности. Указанная задача решается в рамках пассивного подхода, предполагающего определение фактов изменений (фальсификаций) данных ДЗЗ на основе их анализа.

Ключевые слова: дистанционное зондирование Земли, пассивная защита данных ДЗЗ, цифровое изображение, метаданные, элементарный алгоритм, вычислительная процедура.

Введение

Данные, получаемые с космических аппаратов (КА) при ДЗЗ, включают в себя две составляющие: цифровое изображение и соответствующие этому изображению метаданные. Изменениям после получения может подвергаться как собственно изображение, так и метаданные. Под *проверкой подлинности* данных ДЗЗ далее будем понимать комплекс мероприятий, позволяющих ответить на вопрос: «Были ли внесены изменения в данные ДЗЗ?». В настоящее время существуют два основных подхода к проверке подлинности цифровых изображений в общем случае и данных ДЗЗ в частности [1]: *активный* и *пассивный*.

Основным элементом *активного подхода* к проверке подлинности изображений являются цифровые водяные знаки (ЦВЗ) или так называемая цифровая подпись [2, 3]. Недостатком этого подхода является то, что ЦВЗ должен быть встроен в изображение во время записи/получения. Это означает, что либо механизм внесения ЦВЗ в изображение должен быть реализован при разработке бортовой аппаратуры КА, либо внесение ЦВЗ в изображение должно осуществляться на наземной станции приёма данных ДЗЗ. Независимо от места внесения ЦВЗ, итоговое изображение ДЗЗ оказывается искажённым, что является существенным недостатком, а в ряде случаев – недопустимым.

В отличие от активного, *пассивный подход* не предполагает какого-либо предварительного искажения изображения (ЦВЗ не используются). Он основан на предположении, что, даже если изменённое (фальсифицированное) изображение не содержит визуально обнаруживаемых следов изменений, их можно обнаружить путём анализа характеристик самого изображения [1, 4, 5]. В настоящее время существует большое количество алгоритмов, обеспечивающих различные способы проверки подлинности в рамках пассивного подхода [6, 7, 8]. Настоящая работа посвящена вопросам построения *вычислительной процедуры комплексной проверки подлинности данных ДЗЗ* в рамках пассивного подхода, конструируемой с использованием множества алгоритмов – *элементарных алгоритмов проверки подлинности*. Каждый из элементарных алгоритмов (ЭА) осуществляет проверку данных ДЗЗ на предмет наличия изменений – *атак* – определённого класса. Конструируемая вычислительная процедура комплексной проверки подлинности выполняет *обнаружение факта изменения* (обнаружение атаки, проверку подлинности данных ДЗЗ). Наряду с обнаружением атаки, на практике часто возникает необходимость определения типа произошедших изменений – *распознавание типа/класса изменения* (распознавание атаки). Эта задача также рассматривается в рамках настоящей работы.

Учитывая также, что различные ЭА могут отличаться как вычислительной сложностью обработки, так и качественными показателями, при построении вычислительной процедуры комплексной проверки подлинности данных ДЗЗ её показатели качества и сложности выступают в качестве ограничений и/или показателей критериев оптимальности получаемого решения.

Работа построена следующим образом. В первом разделе описана формальная постановка задачи пассивной защиты данных ДЗЗ – вводятся основные определения и элементарные алгоритмы проверки подлинности. Второй раздел посвящён описанию вычислительной процедуры комплексной проверки подлинности данных ДЗЗ и показателей критерия оптимальности. В третьем разделе приводится описание точного и приближённого алгоритмов построения вычислительной процедуры при наличии данных о статистике срабатываний ЭА. В этом разделе также описан поставленный эксперимент по сравнению разработанных алгоритмов и приведены результаты экспериментов. Четвёртый раздел посвящён описанию алгоритма построения вычислительной процедуры при отсутствии данных о статистике срабатываний ЭА. В пятом разделе приводится описание построения вычислительной процедуры распознавания атаки. Наконец, в заключение работы приведены выводы, благодарности и список использованной литературы.

1. Пассивная защита данных ДЗЗ: основные определения и элементарные алгоритмы проверки подлинности данных ДЗЗ

Задача пассивной защиты данных ДЗЗ заключается в обеспечении математических методов, алгоритмов и вычислительных средств проверки подлинности этих данных на основании анализа их содержа-

ния. Данные ДЗЗ включают растровые данные, обозначаемые далее f и метаданные, обозначаемые далее $\Im$, предоставляющие дополнительную информацию о параметрах получения космического снимка (КС). Тогда в контексте формальной постановки задачи пассивной защиты под данными ДЗЗ будем понимать $\Delta \equiv (f, \Im)$, где вторая величина (метаданные) может быть задана полностью, частично или вообще отсутствовать (обозначается $\varnothing$).

Рассмотрим элементы данных ДЗЗ по отдельности. *Цифровое изображение* определим как отображение f (или функция яркости) вида:

$$f : \mathbf{N}_M \times \mathbf{N}_N \to \mathbf{\Phi}^{\beta}, \quad (1)$$

$$f(n_1, n_2) \mapsto n,$$

где $M, N \in \mathrm{N}$ – линейные размеры изображения по вертикали и горизонтали соответственно, $\mathbf{\Phi}$ – множество значений функции яркости изображения, характеризующее способ получения изображения, β – количество спектральных каналов изображения. В качестве $\mathbf{\Phi}$ могут выступать следующие множества:

- $\mathbf{R}$;
- $\mathbf{B} = \{0,1\}$;
- $\mathbf{Z}_{2^m}$.

Для большинства данных ДЗЗ самым часто встречающимся представлением яркостей пикселей является $\mathbf{Z}_{2^m}^{\beta}, \beta \in \{1,3,4\}$. Далее обозначим множество отображений типа (1) как $F_{M,N,\beta}(\mathbf{\Phi})$.

В состав данных ДЗЗ, помимо растровых данных, обычно входят метаданные КС $\Im$, которые содержат различные характеристики процесса регистрации КС. Приведём формальное представление наиболее часто используемых параметров метаданных:

1. Дата съёмки КС $\mathbf{d} \in (\mathbf{Z}_{31} \times \mathbf{Z}_{12} \times \mathbf{N}) \cup \varnothing$ определяется как вектор из трёх компонент, соответствующих дню, месяцу и году даты регистрации КС.

2. Время съёмки КС $\mathbf{t} \in (\mathbf{Z}_{24} \times \mathbf{Z}_{60} \times \mathbf{Z}_{60}) \cup \varnothing$ определяется как вектор из трёх компонент, соответствующих времени начала регистрации КС: час, минута и секунда.

3. Положение Солнца характеризуется вектором $\boldsymbol{\alpha} = (\alpha_{az}, \alpha_{el})^T$, $\boldsymbol{\alpha} \in \mathrm{R}^2 \cup \varnothing$, который состоит из значений двух углов в системе координат «широта/долгота»: первый описывает положение в горизонтальной α_{az} плоскости, а второй – в вертикальной α_{el}. Высота солнца над Землёй считается бесконечно большой по сравнению с остальными метрическими характеристиками в системе «спутник-Земля».

4. Положение космического аппарата описывается кортежем $\mathbf{p} = (\phi_{az}, \phi_{el}, h_{alt}), \mathbf{p} \in \mathbf{R}^3 \cup \varnothing$, который включает в себя значения двух углов в системе координат «широта/долгота»: первый описывает положение в горизонтальной ϕ_{az} плоскости, а второй – в вертикальной ϕ_{el} – и высоту спутника над поверхностью Земли h_{alt}.

5. Координаты территории съёмки описываются вектором $\mathbf{s} = (s_1, s_2, ..., s_k)^T, \mathbf{s} \in \mathbf{R}^k \cup \varnothing, k \in \{4,6\}$. Значение k зависит от типа космического аппарата. В зависимости от значения параметра k территория съёмки описывается 2 способами:

- параллелограммом – 4 вершины;
- полигоном с 4 вершинами, у которого одна пара противоположных граней параллельна, а две другие описываются кривыми 2-го порядка – каждая по 3 точкам, итого, 6 вершин.

6. Тип КА t, с которого был получен КС, будет принадлежать перечислимому типу K.

7. Тип конкретного оптического сенсора l, расположенного на КА, будет принадлежать перечислимому типу L. Так как на КА часто бывает установлено более одного оптического сенсора, то будем описывать сенсоры в виде вектора, размерность которого равна количеству установленных сенсоров.

Таким образом, $\Im = (\mathbf{d}, \mathbf{t}, \boldsymbol{\alpha}, \mathbf{p}, \mathbf{s}, k, l)$ – кортеж, характеризующий метаданные КС. Каждый элемент кортежа может принадлежать пустому множеству $\varnothing$, если этот элемент отсутствует в наборе метаданных снимка. Будем обозначать элемент кортежа $\Im_j$, где j – порядковый номер этого элемента в кортеже $\Im$.

j	Элемент кортежа $\Im$
0	**d**
1	**t**
2	**α**
3	**p**
4	**s**
5	k
6	l

В общем случае данные ДЗЗ полностью описываются парой $\Delta(f, \Im)$, что соответствует растровому изображению, полученному с космического аппарата, и его метаданным. На практике полнота метаданных $\Im$ зависит от поставщика КС или от того, как КС был получен (не напрямую от поставщика, а от дилера или представителя). Множество данных ДЗЗ обозначим $\mathbf{\Lambda}, \Delta(f, \Im) \in \mathbf{\Lambda}$.

Введём индикаторную функцию $I(\Im_j) \in \mathbf{B}$, показывающую наличие конкретного параметра метаданных:

$$I(\Im_j) = \begin{cases} 1, \Im_j \neq \varnothing, \\ 0, \text{иначе}. \end{cases}$$

В самом общем случае под *элементарным алгоритмом (ЭА) a* проверки подлинности данных ДЗЗ D понимается вычислительная процедура, которая

на основании конкретных данных ДЗЗ Δ указывает факт их подлинности: 1 – подлинные, 0 – не подлинные. Иными словами, a осуществляет однозначное отображение следующего вида:

$$a : \mathbf{\Delta} \to \mathbf{B} . \quad (2)$$

Обозначим $\mathbf{A}$ множество отображений типа (2), то есть множество алгоритмов проверки подлинности ($a \in \mathbf{A}$). В зависимости от конкретных реализаций алгоритмов реализация отображения (2) может быть конкретизирована, то есть определена через другие формализованные отображения. Для этого введём ряд понятий.

Параметром ЭА будем называть функцию, реализующую отображение вида:

$$p : \mathbf{A} \to \mathbf{R} .$$

Перечислим параметры ЭА, которые будут использоваться в дальнейшем:

1) *вычислительная сложность ЭА* $u : \mathbf{A} \to \mathbf{R}$ – функция, вычисляющая оценку общего числа арифметических операций алгоритма a;
2) *корректность выполнения ЭА* характеризуется двумя параметрами:
   - первый описывает ошибку первого рода, $p_0 : \mathbf{A} \to \mathbf{R}$,
   - второй описывает ошибку второго рода, $p_1 : \mathbf{A} \to \mathbf{R}$;
3) *число срабатываний ЭА* $c : \mathbf{A} \to \mathbf{N}$ на заданной выборке данных – этот параметр характеризует, сколько раз ЭА a обнаруживает подделку среди анализируемых данных ДЗЗ D (без учёта факта правильности этого обнаружения);
4) *число запусков ЭА* $c_{all} : \mathbf{A} \to \mathbf{N}$ на заданной выборке данных – этот параметр характеризует, сколько раз ЭА a запускался в целях проверки подлинности анализируемых данных ДЗЗ Δ.

В дальнейшем будет также использоваться следующий параметр ЭА, являющийся комбинацией параметров c и c_{all}. Под *частотой срабатывания ЭА* $fr : \mathbf{A} \to \mathbf{R}, fr(a) = c(a) / c_{all}(a)$ на заданной выборке данных будем понимать, как часто ЭА a обнаруживает подделку среди общего количества запусков ЭА.

Дополнительно введём следующие отображения:

1) отображение $Q : \mathbf{A} \times \mathbf{Z}_7 \to \mathbf{B}$, которое для конкретного ЭА $a \in \mathbf{A}$ и номера $j \in \mathbf{Z}_7$ элемента кортежа метаданных $\Im$ указывает на необходимость использования при выполнении a j-го элемента метаданных:

$$Q(a, j) = \begin{cases} 1, \text{ необходим } \Im_j ; \\ 0, \text{ не используется}; \end{cases}$$

2) отображение $R : \mathbf{A} \times \mathbf{Z}_7 \to \mathbf{B}$, которое для ЭА $a \in \mathbf{A}$ и номера $j \in \mathbf{Z}_7$ элемента кортежа метаданных $\Im$ указывает, является ли результатом выполнения a j-й элемент метаданных:

$$R(a, j) = \begin{cases} 1, \Im_j \text{ является результатом}, \\ 0, \text{ не является результатом}. \end{cases}$$

Разрабатываемые ЭА могут состоять из последовательности отдельных отображений, каждое из которых формирует на выходе промежуточные данные. В качестве промежуточных результатов работы ЭА могут выступать:

- список координат объектов;
- изображение – результат аналитической обработки f;
- оценки параметров обнаруженных преобразований изображения $\mathbf{c} \in \mathbf{C}$;
- оценки параметров метаданных $\tilde{\Im}$.

Для того чтобы приводить набор промежуточных операций алгоритмов проверки подлинности к единому типу, введём понятие проекции.

Проекцией будем называть отображение вида:

$$\mathrm{Pr}_{b_j} : (b_1, ..., b_K)^T \to b_j ,$$

где b_j – произвольные величины.

Разные типы атак будем обозначать $w_t \in \mathbf{N}_0$, $t \in [0, T-1]$, где T – количество различных типов атак. Множество различных типов атак будем обозначать $\mathbf{W}_T$.

Введём понятие *матрицы инцидентности* $\mathbf{M}^{\mathbf{B}}_{T \times K}$, $m_{ij} \in \mathbf{B}$, $i \in [0, T-1]$, $j \in [0, K-1]$, где T – количество различных видов атак, K – количество ЭА проверки подлинности данных ДЗЗ. Данная матрица показывает, какие виды атак способны обнаруживать ЭА.

Если значение матрицы инцидентности $m_{ij} = 1$, то оно показывает, что ЭА a_j может выявить подделку на данных с применённой атакой типа i. Эти данные будут использоваться в дальнейшем при распознавании типа атаки по результатам выполнения последовательности ЭА.

Приведём формальное описание различных групп ЭА.

1. ЭА, которые осуществляют проверку данных ДЗЗ на предмет наличия атак первого класса (*ЭА первого типа*) – *дублирования фрагментов КС* [8, 9, 10] – используют в своей работе исключительно цифровое изображение для анализа, то есть $Q(a, j) = 0, \forall j \in \mathbf{Z}_7$. При обнаружении фальсификации, наряду с ответом, алгоритм формирует список, каждый из элементов которого есть множество идентичных (в смысле выбранного критерия алгоритма) фрагментов анализируемого изображения.

Под *фрагментом изображения* далее будем понимать 4- или 8-связную подобласть в $D, D \equiv \mathbf{N}_M \times \mathbf{N}_N$.

Для определённости обозначим список фрагментов из D как $L[D]$. Тогда результатом работы алгоритма является:

$$L\left[L[D]\right] = \begin{Bmatrix} \{D_0^0, D_1^0, ..., D_{R_0-1}^0\}, ..., \\ \{D_0^{s-1}, D_1^{s-1}, ..., D_{R_{s-1}-1}^{s-1}\} \end{Bmatrix}, \quad (3)$$

где s – число элементов различных «типов» фрагментов (например, дубликатов). Обозначим множество фрагментов вида (3) как $\mathbf{L}^2[D]$.

Следует также отметить, что любой ЭА в качестве входных данных может использовать не всё изображение, входящее в состав данных ДЗЗ, а один или несколько его фрагментов, которые могли быть получены по результатам запуска других алгоритмов, в том числе алгоритмов других типов. Это используется, как правило, для сокращения вычислительной сложности или выполнения требовательных к ресурсам алгоритмов.

Под ЭА первого типа будем понимать последовательное выполнение отображений a_1, a_2, a_3:

$$a = a_1 \cdot a_2 \cdot a_3,$$
$$a_1 \equiv \mathrm{Pr}_f : \mathbf{\Delta} \to \mathbf{F}^\beta,$$
$$a_2 : \mathbf{F}^\beta \to \mathbf{L}^2\left[D\right],$$
$$a_3 : \mathbf{L}^2\left[D\right] \to \mathbf{B}.$$

При этом $a_3\left(L\left[L\left[D\right]\right]\right) = \begin{cases} 1, \left|L\left[L\left[D\right]\right]\right| \neq 0 \\ 0, \left|L\left[L\left[D\right]\right]\right| = 0 \end{cases}$.

2. ЭА, осуществляющие проверку данных ДЗЗ на предмет наличия атак второго класса (*ЭА второго типа*) – *ресэмплирование, вставка фрагментов другого КС, склеивание КС, генерирование текстур на КС, компрессия КС, нарушение межканальных зависимостей, добавление шумов, размытие или повышение резкости на КС* [11,12,13,14,15,16,17] – также используют для работы только растровые данные. В отличие от ЭА первого типа в ходе анализа в качестве промежуточных данных выступают изменённые области и оценки параметров преобразований фрагментов изображения, если производились какие-либо изменения (например, матрица аффинного преобразования фрагмента, коэффициент сжатия алгоритмом JPEG).

В контексте ЭА второго типа под $L\left[(D, \mathbf{R}^s)\right] \in \mathbf{L}\left[(D, \mathbf{R}^s)\right]$ будем понимать список обнаруженных фрагментов и вектор оценок параметров обнаруженных преобразований, где s – число различных параметров преобразований. Под параметрами преобразований будем понимать матрицу аффинного преобразования, коэффициент качества JPEG, математическое ожидание и дисперсию шума.

Под ЭА второго типа будем понимать последовательное выполнение отображений a_1, a_2, a_3:

$$a = a_1 \cdot a_2 \cdot a_3,$$
$$a_1 \equiv \mathrm{Pr}_f : \mathbf{\Delta} \to \mathbf{F}^\beta,$$
$$a_2 : \mathbf{F}^\beta \to \mathbf{L}^2\left[(D, \mathbf{R}^s)\right],$$
$$a_3 : \mathbf{L}^2\left[(D, \mathbf{R}^s)\right] \to \mathbf{B}.$$

Здесь $a_3\left(L\left[(D, \mathbf{R}^s)\right]\right) = \begin{cases} 1, \left|L\left[(D, \mathbf{R}^s)\right]\right| \neq 0, \\ 0, \left|L\left[(D, \mathbf{R}^s)\right]\right| = 0. \end{cases}$

3. Группа ЭА, которые осуществляют проверку данных ДЗЗ на предмет наличия атак третьего класса (*ЭА третьего типа*) – *несоответствия освещённости объектов на КС, изменение семантических данных на КС* [18,19,20] – производит оценку параметров метаданных $\Im$, анализируя растровые данные f. Результаты работы алгоритма представляются в виде списка областей с отличающимися параметрами метаданных и, собственно, оценок этих параметров $L\left[(D, \tilde{\Im})\right] \in \mathbf{L}\left[(D, \tilde{\Im})\right]$. Это позволит определить отклонение от параметров входных метаданных и на этом основании сделать вывод о наличии изменений. Отклонение от параметров метаданных позволит определить истинные значения типа КА, типов его сенсоров, координат положения Солнца и КА:

$$a = a_1 \cdot a_2,$$
$$a_1 : \mathbf{\Delta} \to \mathbf{L}\left[(D, \tilde{\Im})\right],$$
$$a_2 : \mathbf{L}\left[(D, \tilde{\Im})\right] \to \mathbf{B}.$$

Здесь $a_2\left(L\left[(D, \tilde{\Im})\right]\right) = \begin{cases} 1, \left|L\left[(D, \tilde{\Im})\right]\right| \neq 0, \\ 0, \left|L\left[(D, \tilde{\Im})\right]\right| = 0. \end{cases}$

4. ЭА, осуществляющие поиск изображений в базе данных, схожих с искомым по характерным признакам (учитываются метаданные изображения, а также содержимое сцены) [21], формируют в качестве промежуточных данных список изображений и вероятности их сходства с искомым фрагментом, которые в дальнейшем будут обозначаться

$$L\left[(\mathbf{F}^\beta, \mathbf{R}_{[0,1]})\right] \in \mathbf{L}\left[(\mathbf{F}^\beta, \mathbf{R}_{[0,1]})\right].$$

В таком случае алгоритм будет представлять из себя отображение следующего вида:

$$a = a_1 \cdot a_2,$$
$$a_1 : \mathbf{\Delta} \to \mathbf{L}\left[(\mathbf{F}^\beta, \mathbf{R}_{[0,1]})\right],$$
$$a_2 : \mathbf{L}\left[(\mathbf{F}^\beta, \mathbf{R}_{[0,1]})\right] \to \mathbf{B},$$

при этом $a_2\left(L\left[\left(\mathbf{F}^\beta, \mathbf{R}_{[0,1]}\right)\right]\right) = \begin{cases} 1, \left|L\left[\left(\mathbf{F}^\beta, \mathbf{R}_{[0,1]}\right)\right]\right| \neq 0, \\ 0, \left|L\left[\left(\mathbf{F}^\beta, \mathbf{R}_{[0,1]}\right)\right]\right| = 0. \end{cases}$

Как было показано выше, любой ЭА проверки подлинности может быть представим в виде произведения отображений:

$$a = b \cdot c,$$

где $b : \mathbf{\Delta} \to \mathbf{L}[*], c : \mathbf{L}[*] \to \mathbf{B}$,

при этом

$$c\left(L[*]\right) = \begin{cases} 1, \left|L[*]\right| \neq 0, \\ 0, \left|L[*]\right| = 0. \end{cases}$$

Заметим также, что каждый ЭА характеризуется не только набором входных данных, но и промежуточными данными, формируемыми при анализе данных ДЗЗ.

2. Вычислительная процедура комплексной проверки подлинности данных ДЗЗ

Вычислительной процедурой комплексной проверки подлинности данных ДЗЗ назовём алгоритм проверки подлинности, осуществляющий путём последовательного выполнения ЭА проверки подлинности:

$$\begin{gathered} a_{i_0}, a_{i_1}, ..., a_{i_{K-1}}; \\ a_{i_k} \in \mathbf{A}_K, \quad i_k \neq i_j (k \neq j), \quad i_k \in \overline{0, K-1}; \end{gathered} \tag{4}$$

здесь $\mathbf{A}_K \subseteq \mathbf{A}$ – множество ЭА, используемых при построении вычислительной процедуры.

Данные ДЗЗ считаются изменёнными / фальсифицированными (проверка подлинности считается непройденной, обнаружен факт атаки), если хотя бы один из ЭА определил факт наличия изменений в данных ДЗЗ (обнаружил атаку).

Задача построения вычислительной процедуры комплексной проверки подлинности данных ДЗЗ заключается в определении оптимальной в смысле некоторого критерия последовательности ЭА (4), обеспечивающей обнаружение факта изменения данных ДЗЗ. В качестве показателей или ограничений критерия оптимальности могут выступать следующие параметры вычислительной процедуры:

- *вычислительная сложность* процедуры комплексной проверки подлинности данных ДЗЗ, вычисляемая на основании значений вычислительной сложности входящих в её состав ЭА следующим образом:

$$\begin{gathered} C_1 = u\left(a_{i_0}\right) + \left(1 - fr\left(a_{i_0}\right)\right) \cdot \\ \cdot \left(u\left(a_{i_1}\right) + ... + \left(1 - fr\left(a_{i_{K-2}}\right)\right) u\left(a_{i_{K-1}}\right)\right); \end{gathered} \tag{5}$$

- ошибки первого и/или второго рода, вычисляемые на основании соответствующих параметров ЭА, входящих в её состав:

$$\begin{gathered} C_2 = p_1\left(a_{i_0}\right) + \left(1 - fr\left(a_{i_0}\right)\right) \cdot \\ \cdot \left(p_1\left(a_{i_1}\right) + ... + \left(1 - fr\left(a_{i_{K-2}}\right)\right) p_1\left(a_{i_{K-1}}\right)\right). \end{gathered} \tag{6}$$

В приведённых выше выражениях последовательность $a_{i_0}, ..., a_{i_{K-1}}$ обозначает конкретную последовательность выполнения ЭА.

3. Построение вычислительной процедуры при известной статистике

Как было замечено в предшествующем разделе, задача построения вычислительной процедуры комплексной проверки подлинности данных ДЗЗ заключается в определении последовательности (4) ЭА, которая в смысле выбранного критерия оказывается оптимальной. Существенным моментом является то, что состав *множества используемых при построении алгоритмов* $\mathbf{A}_K \subseteq \mathbf{A}$ определяется составом и типом входных данных ДЗЗ, проверку подлинности которых требуется осуществить. При построении оптимальной последовательности ЭА следует также учитывать тот факт, что они могут быть зависимы по входным и выходным данным. В случае отсутствия такой зависимости, после выполнения алгоритма a_{i_0} может быть запущен любой из оставшихся $K-1$ алгоритмов множества $\mathbf{A}_K$. В случае появления зависимостей по данным, последовательность ЭА длины K может сократиться до длины $K' < K$, где K' – количество ЭА множества $\mathbf{A}_K$, зависящих друг от друга по входным и выходным данным.

Очевидный способ построения оптимальной в смысле выбранного критерия последовательности ЭА заключается в переборе всех возможных перестановок алгоритмов, что требует рассмотрения $K!$ вариантов. Если для $K < 10$ указанная переборная задача не представляет серьёзной вычислительной сложности, то при бо́льших значениях K поиск оптимального решения может оказаться затруднительным или вовсе невыполнимым ввиду значительных затрат времени и ресурсов. Поэтому ниже предлагается два способа решения задачи построения искомой вычислительной процедуры комплексной проверки подлинности данных ДЗЗ (поиска оптимальной последовательности ЭА):

- точный – обеспечивает построение оптимальной последовательности ЭА при сравнительно небольших значениях K;
- приближённый – обеспечивает нахождение квазиоптимального решения задачи для больших значений K.

Далее в настоящем разделе рассмотрены оба предложенных способа/алгоритма построения вычислительной процедуры комплексной проверки подлинности данных ДЗЗ, проводится анализ получаемых решений в смысле затрачиваемого времени на построение и получаемого показателя качества процедуры комплексной проверки подлинности.

Заметим также, что ниже, для того чтобы избавиться от лишних вычислений, выражения (5) и (6) приведены к следующему виду:

$$
\begin{aligned}
&u\left(a_{i_0}\right)+\left(1-fr\left(a_{i_0}\right)\right)\left(\left(u\left(a_{i_1}\right)+\ldots+\left(1-fr\left(a_{i_{K-2}}\right)\right)u\left(a_{i_{K-1}}\right)\right)\right)=\\
&=u\left(a_{i_0}\right)+\left(1-fr\left(a_{i_0}\right)\right)u\left(a_{i_1}\right)+\\
&+\left(1-fr\left(a_{i_0}\right)\right)\left(1-fr\left(a_{i_1}\right)\right)u\left(a_{i_2}\right)+\ldots+\\
&+\left(1-fr\left(a_{i_0}\right)\right)\cdot\ldots\cdot\left(1-fr\left(a_{i_{K-2}}\right)\right)u\left(a_{i_{K-1}}\right)=\\
&=\left[\upsilon_{i_k}=1-fr\left(a_{i_k}\right)\right]=\\
&=u\left(a_{i_0}\right)+\upsilon_{i_0}u\left(a_{i_1}\right)+\ldots+\upsilon_{i_0}\upsilon_{i_1}\cdot\ldots\cdot\upsilon_{i_{K-2}}u\left(a_{i_{K-1}}\right).
\end{aligned}
\qquad (7)
$$

В таком виде вычисление показателя критерия сводится к вычислению прямой суммы (7), что существенно удобнее и не требует при переборах перестановок повторных вычислений.

3.1. Точный алгоритм построения вычислительной процедуры комплексной проверки подлинности

Точный алгоритм основывается на методе полного перебора с отсевом подмножеств допустимых решений, заведомо не содержащих оптимальное (метод динамического программирования ветвей и границ) [22]. На первом шаге алгоритм формирует первое решение (первую перестановку), считая его потенциальным оптимумом. Далее во время перебора всех возможных комбинаций ($K!$) происходит исключение тех перестановок, для которых значение показателя критерия превышает потенциальный оптимум, полученный ранее. В случае, если находится новый оптимум, потенциальный оптимум меняется. Таким образом удаётся получать точное значение показателя критерия при меньших временных и ресурсных затратах. Результаты зависимости времени поиска минимального значения показателя критерия от числа алгоритмов представлены на рис. 1. Как можно видеть из данного рисунка, несмотря на ускорение по сравнению с полным перебором, данный метод не позволяет вычислять точное значение показателя критерия за приемлемое время для $K > 25$.

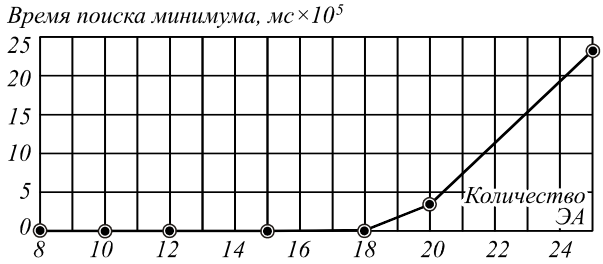

Рис. 1. Зависимость времени поиска минимального значения показателя критерия от количества ЭА в последовательности

3.2. Приближённый алгоритм построения вычислительной процедуры комплексной проверки подлинности

Для того чтобы иметь возможность строить вычислительную процедуру при $K > 25$, был реализован приближённый алгоритм, основанный на методе последовательного присоединения [22].

Суть алгоритма состоит в том, что на k-ом шаге алгоритма, когда уже определена и зафиксирована последовательность ЭА

$$a_{i_0}, a_{i_1}, \ldots, a_{i_{k-2}}$$

определяется такой ЭА $a_{i_{k-1}}$, для которого показатель критерия (5) или (6) принимает минимальное значение. Каждое слагаемое на следующем шаге алгоритма фактически зависит от значения выбранного параметра ЭА $u\left(a_{i_{k-1}}\right)$, так как коэффициенты $\upsilon_{i_0}, \upsilon_{i_1}, \ldots, \upsilon_{i_{k-2}}$ определены выбором ЭА-ов $a_{i_0}, a_{i_1}, \ldots, a_{i_{k-2}}$ на предыдущих шагах. Таким образом, в результате последовательного присоединения мы получаем упорядоченную последовательность алгоритмов (4).

Для улучшения решения, получаемого методом последовательного присоединения, была разработана его модификация. Она заключается в итерационном выполнении следующей операции в текущей последовательности: производится парная перестановка двух ЭА a_{i_k} и $a_{i_{k+n}}$ в том случае, если эта перестановка приводит к снижению значения показателя критерия. Итерации продолжаются до тех пор, пока существуют требуемые для перестановки пары алгоритмов.

Результаты сравнения приближённых алгоритмов с модификацией и без неё представлены на рис. 2 (описание постановки эксперимента приведено ниже). На представленном графике приведена величина относительного отклонения минимального значения показателя критерия, полученного при помощи приближённых алгоритмов, от минимального значения показателя критерия, полученного при помощи точного алгоритма. Величина относительного отклонения вычисляется по следующей формуле:

$$\varepsilon = \frac{\tilde{C} - C_{\min}}{C_{\max} - C_{\min}},$$

где $\tilde{C}$ – минимальное значение показателя критерия, полученное при помощи приближённого алгоритма, $C_{\max}, C_{\min}$ – точные минимальное и максимальное значения показателя критерия, вычисленные при помощи точного алгоритма.

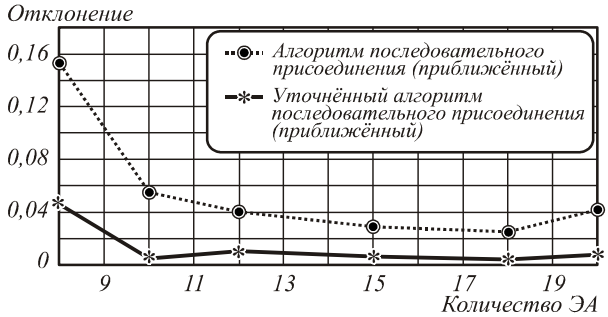

Рис. 2. Зависимость отклонения приближённого минимального значения показателя критерия от количества ЭА в последовательности

Из графиков видно, что модифицированный алгоритм позволяет в несколько раз улучшить получаемое решение по сравнению с первоначальным алгоритмом.

3.3. Сравнение точного и приближённого алгоритмов построения вычислительной процедуры

Для модифицированного приближённого алгоритма график зависимости времени построения вычислительной процедуры комплексной проверки подлинности от числа ЭА выглядит следующим образом (рис. 3).

Если сравнивать результаты, показанные на рис. 3 с результатами на рис. 1, то можно заметить, что скорость работы модифицированного приближённого алгоритма во много раз превышает скорость работы точного метода, делая приближённый алгоритм применимым на практике практически для любого реального числа ЭА.

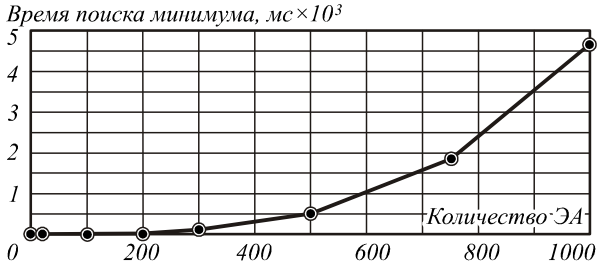

Рис. 3. Зависимость времени поиска приближённого минимального значения показателя критерия от количества ЭА в последовательности

Для проведения экспериментальных исследований разработанных методов была создана .NET библиотека. Генерирование количества ЭА и их параметров (вычислительная сложность, частота срабатывания) было произведено при помощи класса *Math.Random* встроенной библиотеки *Math*. В основе алгоритма генерации случайных чисел, реализованного в *Math.Random*, лежит субтрактивный алгоритм генератора случайных чисел Д. Кнута [23]. Сгенерированные случайные величины распределены по равномерному закону распределения.

В данной работе было реализовано 10 экспериментов, в каждом из которых случайным образом выбирались количество ЭА (на промежутке [8, 20]) и их параметры (вычислительная сложность выбиралась на промежутке [100, 10000], а частота срабатывания – на промежутке [0, 1]).

В ходе каждого эксперимента вычислялись точные и приближённые минимальные и максимальные значения показателя критерия, а также приближённое среднее значение показателя критерия. На рис. 4 представлены средние значения критерия, вычисленные по 10 проведённым экспериментам.

Как видно из графика, кривая приближённых значений, вычисленных при помощи модифицированного приближённого алгоритма, практически совпадает с кривой точных минимальных значений показателя критерия. Таким образом, приближённый алгоритм следует использовать для вычисления значения критерия при больших значениях K.

Тенденция, показанная на рис. 1, 2, 3, наблюдается как для случая, когда алгоритмы не зависят друг от друга по входным и выходным данным, так и для случая, когда такая зависимость присутствует.

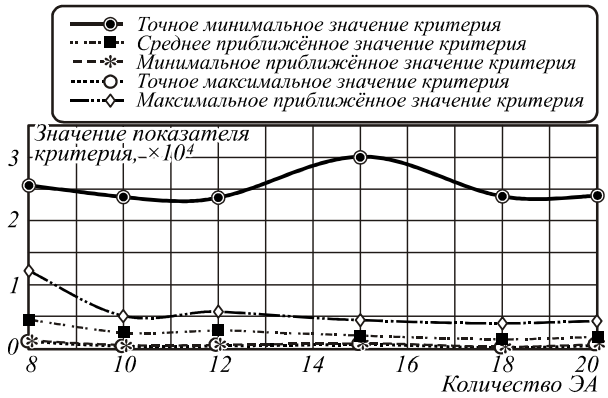

Рис. 4. Зависимость значений показателя критерия от количества ЭА в последовательности (средние значения для 10 реализаций)

4. Построение вычислительной процедуры комплексной проверки подлинности данных ДЗЗ при отсутствии статистики

Ранее была рассмотрена ситуация, когда вся информация об алгоритмах считалась известной и задача заключалась в поиске оптимальной последовательности ЭА. Теперь рассмотрим ситуацию, когда перед построением вычислительной процедуры комплексной проверки подлинности отсутствует информация о частоте срабатываний ЭА или, другими словами, не было выполнено достаточное количество запусков ЭА из $\mathbf{A}_K$. Тогда множество $\mathbf{A}_K$ можно разделить на два подмножества:

- ЭА с недостаточной информацией;
- ЭА с достаточной информацией.

При такой постановке дополнительной задачей является уточнение информации об ЭА из первого подмножества, а именно о частоте их срабатываний. В таком случае последовательность ЭА конструируемой вычислительной процедуры должна начинаться с тех ЭА, информация о количестве срабатываний которых (число запусков алгоритмов) является минимальной. При такой стратегии все ЭА множества $\mathbf{A}_K$ перейдут во второе подмножество. Дальнейшее построение вычислительной процедуры в смысле оптимальности критерия производится аналогично описанию выше.

После запуска каждого ЭА информация о количестве его запусков увеличивается на 1, а число срабатываний изменяется в зависимости от результата выполнения ЭА. Как только число запусков алгоритма удовлетворяет некоторому заранее введённому ограничению, этот алгоритм переходит в группу с достаточной информацией. Как только информация обо всех алгоритмах будет накоплена, построение последовательности будет производиться так, как описано в предыдущем разделе.

5. Построение вычислительной процедуры распознавания атаки

Наряду с задачей определения подлинности предъявляемых для проверки данных ДЗЗ, в ряде случаев требуется распознать собственно атаку, то есть определить, какие именно изменения данных ДЗЗ

были произведены. Для решения задачи распознавания атаки в данной работе будет использоваться матрица инцидентности $\mathbf{M}^{\mathbf{B}}_{T\times K}$, введённая ранее.

Распознавание типа атаки может производиться при отсутствии информации о запуске последовательности ЭА, либо по результатам выполнения вычислительной процедуры комплексной проверки подлинности. В обоих случаях задача сводится к построению решающего правила для классификации атаки, применённой к данным ДЗЗ, используя информацию из матрицы инцидентности $\mathbf{M}^{\mathbf{B}}_{T\times K}$.

Для построения решающего правила воспользуемся обратной процедурой байесовского конечного последовательного распознавания с упорядочиванием признаков [24]. Представим задачу определения типа атаки в терминах распознавания образов. Под признаками будем понимать результаты работы ЭА множества $\mathbf{A}_K$ (то есть числа {0,1}), а в качестве классов будут выступать типы атак $w_t, t \quad [0,T–1]$.

Стоимость измерения каждого признака будем трактовать как вычислительную сложность ЭА – $u(a_i)$.

Риск принятия решения о классификации замеров $a_0,...,a_n$ в класс w_t обозначим $R(a_0,...,a_n;w_t)$. В качестве минимального среднего риска последовательного решающего процесса на шаге n будем использовать следующее выражение, характеризующее вычислительную сложность процесса принятия решения:

$$\rho_n(a_0,...,a_n) = \min\begin{cases} прод.: u(a_{n+1}) + \\ \quad \sum\limits_{a_{n+1}\in\mathbf{A}_K} \rho_{n+1}(a_0,...,a_n,a_{n+1})\cdot P(a_{n+1}/a_0...a_n) \\ ост.: \min\limits_t R(a_0,...,a_n;w_t). \end{cases}$$

Используя данные матрицы инцидентности, составим таблицу для построения решающего правила в стандартном для задачи последовательного распознавания виде. Матрица будет иметь следующий вид:

	a_0	…	a_{K-1}	w_0	…	w_{T-1}
0	m_{00}	…	$m_{0,K-1}$	v_{00}	…	$v_{0,K-1}$
…	…	…	…	…	…	…
$T-1$	$m_{T-1,0}$	…	$m_{T-1,K-1}$	$v_{T-1,0}$	…	$v_{T-1,K-1}$
…	…	…	…	–	–	–
2^K	1	…	1	–	–	–

В каждой строке расширенной матрицы содержатся сначала значения матрицы инцидентности m_{ij}, а затем частоты срабатывания алгоритмов v_{ij}, полученные на стадии обучения. Первые T строк содержат информацию из матрицы инцидентности, остальные $2^K - T$ строк соответствуют неопределённым типам атак.

Далее происходит построение решающего правила по расширенной таблице в соответствии с алгоритмом, описанным в [4]. Расчёт риска производится с последнего шага. Несущественное отличие от классической постановки заключается в том, что количество атак T меньше числа различных комбинаций значений признаков 2^K. Поэтому в ходе построения решающей процедуры могут быть получены последовательности ЭА, не имеющие соответствующих им типов атак. Такие результаты будем классифицировать в неопределённый класс атак.

Выводы и рекомендации

В данной работе предложен алгоритм построения вычислительной процедуры комплексной проверки подлинности данных ДЗЗ, который позволяет строить последовательность ЭА проверки подлинности в смысле оптимальности одного из предложенных критериев. На примере проведённых экспериментов показана высокая скорость вычисления минимальных значений показателей критериев при помощи точного и приближённого алгоритмов. В работе было показано применение обратной процедурой байесовского конечного последовательного распознавания для решения задачи построения вычислительной процедуры распознавания атаки.

Благодарности

Работа выполнена при частичной финансовой поддержке:
– гранта РФФИ (проект 12-07-00021-а);
– программы фундаментальных исследований Президиума РАН «Фундаментальные проблемы информатики и информационных технологий», (проект 2.12);
– Министерства образования и науки Российской Федерации.

Литература

1. **Sridevi, M.** Comparative study of image forgery and copy-move techniques / M. Sridevi, C. Mala and S. Sanyam. – New Delhi, India: Proceedings of the Second International Conference on Computer Science, Engineering and Applications (ICCSEA 2012), 2012. – P. 715-723.
2. **Cox, I.J.** Watermarking is not cryptography / I.J. Cox, G. Doerr, T. Furon. – Proceedings of the 5th International Workshop on Digital Watermarking, 2006. – P. 1-15.
3. **Lin, E.T.** A review of fragile image watermarks / E.T. Lin, E.J. Delp // Proceedings of ACM Multimedia and Security Workshop. – 1999. – Vol. 1. – P. 25–29.
4. **Mahdian, B.** A bibliography on blind methods for identifying image forgery / B. Mahdian, S. Saic // Signal Processing: Image Communication. – 2010. – Vol. 25. – P. 389-399.

5. **Глумов, Н.И.** Обнаружение на изображениях искусственных изменений локального происхождения / Н.И. Глумов, А.В. Кузнецов // Автометрия. – 2011. – Т. 47, № 3. – С. 3-11.
6. **Popescu, A.C.** Statistical Tools for Digital Image Forensics: PhD thesis / A.C. Popescu. – Hanover, USA: Dartmouth College, Department of Computer Science, 2005. – 102 p.
7. **Fridrich, J.** Estimation of primary quantization matrix in double compressed JPEG images / J. Fridrich, J. Lukas. – Digital Forensic Research Workshop. – 2003. – P. 2-5.
8. **Bayram, S.** A Survey of Copy-Move Forgery Detection Techniques / S. Bayram, H.T. Senca, N. Memon. – NY: IEEE Western New York Image Processing Workshop, 2008. – P. 1-4.
9. **Fridrich, J.** Detection of copy–move forgery in digital images / J. Fridrich, D. Soukal, J. Lukas. – Cleveland, OH, USA: Proceedings of Digital Forensic Research Workshop, IEEE Computer Society, 2003. – P. 55-61.
10. **Huang, H.** Detection of copy–move forgery in digital images using sift algorithm / H. Huang, W. Guo, Y. Zhang. – Washington, DC, USA: Proceedings of the 2008 IEEE Pacific-Asia Workshop on Computational Intelligence and Industrial Application, IEEE Computer Society, 2008. – P. 272-276.
11. **Kirchner, M.** Fast and reliable resampling detection by spectral analysis of fixed linear predictor residue / M. Kirchner. – New York, NY, USA: Proceedings of the 10th ACM workshop on Multimedia and security, ACM, 2008. – P. 11-20.
12. **Dong, J.** Run-length and edge statistics based approach for image splicing detection / J. Dong, W. Wang, T. Tan, Y. Shi. – Busan, Korea: Digital Watermarking, 7th International Workshop, IWDW 2008, 2008. – P. 76-87.
13. **Sankar, G.** Feature based classification of computer graphics and real images / G. Sankar, V. Zhao, Y.-H. Yang. – Washington, DC, USA: Proceedings of the 2009 IEEE International Conference on Acoustics, Speech and Signal Processing, IEEE Computer Society, 2009. – P. 1513-1516.
14. **Li, C.-T.** Detection of block artifacts for digital forensic analysis / C.-T. Li // e-Forensics. – 2009. – P. 173-178.
15. **Fan, N.** A pixel-based digital photo authentication framework via demosaicking inter-pixel correlation / N. Fan, C. Jin, Y. Huang // New York, NY, USA: Proceedings of the 11th ACM Workshop on Multimedia and Security, ACM, 2009. – P. 125-130.
16. **Gou, H.** Noise features for image tampering detection and steganalysis / H. Gou, A. Swaminathan, M. Wu. – San Antonio, USA: ICIP (6), IEEE, 2007. – P. 97-100.
17. **Li, Z.** Blind detection of digital forgery image based on the local entropy of the gradient / Z. Li, J. Bin Zheng // IWDW. – 2008. – P. 161-169.
18. **Johnson, M.** Exposing digital forgeries in complex lighting environments / M. Johnson, H. Farid. // IEEE Transactions on Information Forensics and Security. – 2007. – N 3(2). – P. 450-461.
19. **Farid, H.** Image forensic analyses that elude the human visual system / H. Farid, M. Bravo. – San Jose, CA, USA: SPIE Symposium on Electronic Imaging, 2010. – 10 p.
20. **Lee, S.** Detecting false captioning using common-sense reasoning / S. Lee, D.A. Shamma, B. Gooch // Digital Investigation 3, Suppl. 1. – 2006. – P. 65-70.
21. **Taileb, M.** NOHIS-Tree: High-Dimensional Index Structure for Similarity Search / M. Taileb, S. Touati // World Academy of Science, Engineering and Technology. – 2011. – N 59.– P. 518-525.
22. **Таха, Х.А.** Введение в исследование операций / Х.А. Таха. – 6-е изд. – М.: Вильямс, 2001. – 912 с.
23. **Кнут, Д.Э.** Искусство программирования. Том 2. Получисленные алгоритмы / Д.Э. Кнут. – М.: Вильямс, 2007. – 500 с.
24. **Фу, К.** Последовательный методы в распознавании образов и обучении машин / К. Фу. – М.: Наука, 1971. – 256 с.

References

1. **Sridevi, M.** Comparative study of image forgery and copy-move techniques / M. Sridevi, C. Mala and S. Sanyam. – New Delhi, India: Proceedings of the Second International Conference on Computer Science, Engineering and Applications (ICCSEA 2012), 2012. – P. 715-723.
2. **Cox, I.J.** Watermarking is not cryptography / I.J. Cox, G. Doerr, T. Furon. – Proceedings of the 5th International Workshop on Digital Watermarking, 2006. – P. 1-15.
3. **Lin, E.T.** A review of fragile image watermarks / E.T. Lin, E.J. Delp // Proceedings of ACM Multimedia and Security Workshop. – 1999. – Vol. 1. – P. 25–29.
4. **Mahdian, B.** A bibliography on blind methods for identifying image forgery / B. Mahdian, S. Saic // Signal Processing: Image Communication. – 2010. – Vol. 25. – P. 389-399.
5. **Glumov, N.I.** Detection of Local Artificial Changes in Images / N.I. Glumov, A.V. Kuznetsov // Optoelectronics, Instrumentation and Data Processing. - 2011. - Vol. 47(3). - P. 4-12..
6. **Popescu, A.C.** Statistical Tools for Digital Image Forensics: PhD thesis / A.C. Popescu. – Hanover, USA: Dartmouth College, Department of Computer Science, 2005. – 102 p.
7. **Fridrich, J.** Estimation of primary quantization matrix in double compressed JPEG images / J. Fridrich, J. Lukas. – Digital Forensic Research Workshop. – 2003. – P. 2-5.
8. **Bayram, S.** A Survey of Copy-Move Forgery Detection Techniques / S. Bayram, H.T. Senca, N. Memon. – NY: IEEE Western New York Image Processing Workshop, 2008. – P. 1-4.
9. **Fridrich, J.** Detection of copy–move forgery in digital images / J. Fridrich, D. Soukal, J. Lukas. – Cleveland, OH, USA: Proceedings of Digital Forensic Research Workshop, IEEE Computer Society, 2003. – P. 55-61.
10. **Huang, H.** Detection of copy–move forgery in digital images using sift algorithm / H. Huang, W. Guo, Y. Zhang. – Washington, DC, USA: Proceedings of the 2008 IEEE Pacific-Asia Workshop on Computational Intelligence and Industrial Application, IEEE Computer Society, 2008. – P. 272-276.
11. **Kirchner, M.** Fast and reliable resampling detection by spectral analysis of fixed linear predictor residue / M. Kirchner. – New York, NY, USA: Proceedings of the 10th ACM workshop on Multimedia and security, ACM, 2008. – P. 11-20.
12. **Dong, J.** Run-length and edge statistics based approach for image splicing detection / J. Dong, W. Wang, T. Tan, Y. Shi. – Busan, Korea: Digital Watermarking, 7th International Workshop, IWDW 2008, 2008. – P. 76-87.
13. **Sankar, G.** Feature based classification of computer graphics and real images / G. Sankar, V. Zhao, Y.-H. Yang. – Washington, DC, USA: Proceedings of the 2009 IEEE International Conference on Acoustics, Speech and Signal Processing, IEEE Computer Society, 2009. – P. 1513-1516.
14. **Li, C.-T.** Detection of block artifacts for digital forensic analysis / C.-T. Li // e-Forensics. – 2009. – P. 173-178.
15. **Fan, N.** A pixel-based digital photo authentication framework via demosaicking inter-pixel correlation / N. Fan, C. Jin, Y. Huang // New York, NY, USA: Proceedings of the 11th ACM Workshop on Multimedia and Security, ACM, 2009. – P. 125-130.
16. **Gou, H.** Noise features for image tampering detection and steganalysis / H. Gou, A. Swaminathan, M. Wu. – San Antonio, USA: ICIP (6), IEEE, 2007. – P. 97-100.
17. **Li, Z.** Blind detection of digital forgery image based on the local entropy of the gradient / Z. Li, J. Bin Zheng // IWDW. – 2008. – P. 161-169.
18. **Johnson, M.** Exposing digital forgeries in complex lighting environments / M. Johnson, H. Farid. // IEEE Transactions on Information Forensics and Security. – 2007. – N 3(2). – P. 450-461.

19. **Farid, H.** Image forensic analyses that elude the human visual system / H. Farid, M. Bravo. – San Jose, CA, USA: SPIE Symposium on Electronic Imaging, 2010. – 10 p.
20. **Lee, S.** Detecting false captioning using common-sense reasoning / S. Lee, D.A. Shamma, B. Gooch // Digital Investigation 3, Suppl. 1. – 2006. – P. 65-70.
21. **Taileb, M.** NOHIS-Tree: High-Dimensional Index Structure for Similarity Search / M. Taileb, S. Touati // World Academy of Science, Engineering and Technology. – 2011. – N 59.– P. 518-525.
22. **Taha, H.A.** Operations Research: An Introduction (6th Edition) / H.A. Taha. – Moscow: "Williams" Publisher, 2001. – 912 p. – (In Russian).
23. **Knuth, D.E.** The art of computer programming. Volume 2 / D.E. Knuth. – Moscow: "Williams" Publisher, 2007. – 500 p. – (In Russian).
24. **Fu, K.** Sequential methods in pattern recognition and machine learning / K. Fu. – Moscow: "Nauka" Publisher, 1971. – 256 p. – (In Russian).

COPY-MOVE IMAGE FORENSICS DETECTION

A.V. Kuznetsov, V.V. Myasnikov
Image Processing Systems Institute of the RAS

Abstract

The problem of constructing a complex calculation procedure of remote sensing data authentication using a set of basic algorithms for authentication is considered in this paper. This problem is solved according to the passive approach of data authentication, which assumes that artificial changes detection (forgeries detection) based on remote sensing data analysis.

Key words: passive remote sensing data protection, digital images, metadata, optimality criterion, elementary algorithm.

Сведения об авторах

Кузнецов Андрей Владимирович, родился в 1987 году. В 2010 году окончил Самарский государственный аэрокосмический университет (СГАУ) с отличием по специальности «Прикладная математика и информатика». В настоящее время работает стажёром-исследователем в Институте систем обработки изображений РАН, является аспирантом СГАУ. Круг научных интересов включает обработку и анализ изображений, обнаружение локальных изменений на изображениях, распознавание образов, геоинформатику. Имеет 17 публикаций, в том числе 4 научных статьи.

E-mail: *kuznetsoff.andrey@gmail.com* .

Andrey Vladimirovich Kuznetsov (b. 1987) graduated with honours (2010) from the S. P. Korolyov Samara State Aerospace University (SSAU), majoring in Applied Mathematics and Informatics. He works as a researcher in Samara Image Processing Systems Institute of the Russian Academy of Sciences (IPSI RAS), also studies as a postgraduate student in SSAU. His research interests are currently focused on image processing and analysis, local images changes detection, pattern recognition, geoinformatics. He has 17 publications, including 4 scientific papers.

Мясников Владислав Валерьевич, 1971 года рождения. В 1994 году окончил Самарский государственный аэрокосмический университет (СГАУ). В 1995 году поступил в аспирантуру СГАУ, в 1998 году защитил диссертацию на соискание степени кандидата технических наук, а в 2008 – диссертацию на соискание степени доктора физико-математических наук. В настоящее время работает ведущим научным сотрудником в Федеральном государственном бюджетном учреждении науки Институт систем обработки изображений РАН и одновременно профессором кафедры геоинформатики и информационной безопасности СГАУ. Круг научных интересов включает цифровую обработку сигналов и изображений, компьютерное зрение, распознавание образов, искусственный интеллект и геоинформатику. Имеет более 100 публикаций, в том числе 40 статей и две монографии (в соавторстве). Член Российской ассоциации распознавания образов и анализа изображений.

E-mail: *vmyas@smr.ru* . Страница в интернете: http://www.ipsi.smr.ru/staff/MyasVV.htm

Vladislav Valerievich Myasnikov (1971 b.), graduated (1994) from the S.P. Korolyov Samara State Aerospace University (SSAU). He received his PhD in Technical sciences (2002) and DrSc degree in Physics & Maths (2008). At present he is a leading researcher at the Image Processing Systems Institute of the Russian Academy of Sciences and holds a part-time position of Associate Professor at the Department of Geoinformatics and Information Security at SSAU. The area of interests includes digital signals and image processing, geoinformatics, neural networks, computer vision, pattern recognition and artificial intelligence. He's list of publications contains about 100 scientific papers, including 40 articles and 2 monographs. He is a member of Russian Association of Pattern Recognition and Image Analysis.

Поступила в редакцию 17 апреля 2013г.