

Н.А. ЖУКОВА, И.С. СОКОЛОВ

МЕТОД ВОССТАНОВЛЕНИЯ СТРУКТУРЫ ГРУППОВОГО ТЕЛЕМЕТРИЧЕСКОГО СИГНАЛА НА ОСНОВЕ ГРАФОВОЙ МОДЕЛИ

Жукова Н.А., Соколов И.С. Метод восстановления структуры группового телеметрического сигнала на основе графовой модели.

Аннотация. Оценка функционирования сложного технического объекта во время испытаний производится на основе анализа регистрируемых датчиками параметров. Параметры передаются с объекта в центр обработки данных в форме группового телеметрического сигнала (ГТС). При передаче ГТС в ряде случаев сопровождающие их описания структуры отсутствуют или содержат неточности. Для решения задачи восстановления или проверки корректности описания авторами разработан новый подход к работе с ГТС, основанный на применении теории графов и методов корреляционного анализа. В статье рассматривается графовая модель описания ГТС, которая позволяет представить ГТС как совокупность структуры коммутации датчиков и характеристик передаваемых параметров, и метод ее построения.

Ключевые слова: групповой телеметрический сигнал, корреляционный анализ, теория графов, сравнение графов, восстановление структуры.

Zhukova N.A., Sokolov I.S. Method of reconstructing the structure of the group telemetric signals based on the graph model.

Abstract. Evaluation of a complex technical object during the test is based on analysis of recorded sensor parameters. Parameters are passed from the object to a data center in the form of group telemetric signal (GTS). When transferring GTS in some cases, their accompanying descriptions of the structure are missing or contain inaccuracies. To solve the problem of reconstructing or validating GTS structure authors describe a new approach to GTS structure analysis, based on application of graph theory and methods of correlation analysis. The article considers the graph model describing GTS, which allows submitting a GTS as a union of sensors commutation structure and characteristics of parameters, and method of GTS construction.

Keywords: group telemetric signal, correlation analysis, graph theory, graph matching, structure restore

1. Введение. Оценка функционирования сложного технического объекта во время испытаний производится на основе анализа регистрируемых датчиками параметров. Большой объем параметров не позволяет проводить обработку данных на объекте, и данные передаются в специализированные центры их обработки, которые, как правило, удалены на достаточно большое расстояние от центров испытаний. Параметры передаются с объекта в центр обработки данных в форме структурированного бинарного потока данных. Этот информационный поток принято называть групповым телеметрическим сигналом (ГТС). ГТС имеет достаточно сложную структуру, что связано с необходимостью

стью передавать параметры разных типов с разной частотой. Каждый ГТС сопровождается описанием структуры, которое позволяет на этапе обработки и анализа декомпонировать сложный информационный поток на отдельные параметры. Наличие и корректность описания структуры ГТС определяет возможность проведения анализа данных и влияет на качество получаемых результатов. Дополнительная сложность заключается в том, что в настоящее время достаточно часто происходит смена структуры ГТС между несколькими испытаниями, а также в рамках одного испытания.

При передаче ГТС в ряде случаев сопровождающие их описания структуры отсутствуют или содержат неточности. В связи с чем возникает необходимость восстановления или проверки корректности их описания. Формирование описания ГТС экспертами в «ручном» режиме с использованием специализированных средств графического представления ГТС в условиях ограниченного времени не возможно. Существующие алгоритмы автоматизированного формирования структуры ГТС, основанные на применении статистико-лингвистического подхода, не обеспечивают требуемой точности, кроме того, для их работы необходима база данных ГТС, содержащая все возможные типы ГТС. Так в работе [3] предлагается использовать в качестве модели ГТС вектор коэффициентов для аппроксимации статистико-лингвистического распределения. Подход основан на применении закона Ципфа–Мандельборта, в соответствии с которым произведение ранга некоторого слова и частоты его встречаемости в тексте является величиной постоянной, имеющей примерно одинаковое значение для любого слова из словаря. Для аппроксимации частотно-рангового распределения используется функция:

$$\Phi(r) = C \cdot r^{-\gamma \exp(d \cdot r)}, \quad (1)$$

где Φ — относительная частота слова с рангом r ; d — параметр увеличения коэффициента частотно-рангового соотношения.

Вектор-образ ГТС представляется как: $[C, \gamma, d]$.

Для восстановления структуры ГТС осуществляется поиск наиболее близкого по заданному критерию вектора-образа ГТС в базе ГТС.

К основным достоинствам модели относятся: отсутствие требований к наличию информации о внутренней организации ГТС кроме длины кадра, простота использования и низкие требования, предъявляемые к аппаратным средствам. Применение модели (1) целесообразно в случае использования ограниченного числа форматов ГТС, при

этом ГТС должны существенно различаться по структуре и передаваемым данным. В случае же наличия смены формата кадра применение статистически-лингвистического подхода имеет крайне низкую точность вследствие того, что значения вектора образа усредняются для всего ГТС. В связи с этим важной задачей является разработка метода автоматизированного восстановления описания структуры ГТС, позволяющего обеспечить точность описания структуры на уровне 95–98 % в зависимости от качества исходных данных. Для решения поставленной задачи авторами предложен новый подход к работе с ГТС, основанный на применении теории графов и методов корреляционного анализа. Применение графового представления позволяет перейти от описания потока телеметрической информации к структурному описанию поведения всей системы в целом. В статье рассматривается графовая модель описания ГТС, которая позволяет представить ГТС как совокупность структуры коммутации датчиков и характеристик передаваемых параметров, и метод ее построения.

2. Особенности формирования ГТС. Групповой телеметрический сигнал — это сигнал, получаемый объединением первичных сигналов от датчиков на основе временного разделения каналов. Рассмотрим основные принципы формирования ГТС. В сложной технической системе устанавливаются датчики, которые получают измерения в цифровом виде. В датчиках встроены аналогово-цифровые преобразователи. При этом все датчики в устройстве имеют одну и ту же разрядность.

Один последовательный цикл опроса главным коммутатором всех подключенных к нему датчиков порождает последовательность измерений, называемую кадром. Под длиной кадра понимается число датчиков, непосредственно подключенных к главному коммутатору. Выделяются понятия супер- и субкоммутации. Суперкоммутация применяется для передачи параметра с частотой, кратной частоте опроса главного коммутатора. Несколько измерений от данного датчика попадает в один кадр. Субкоммутация служит для передачи параметра с частотой, субкратной частоте главного коммутатора. Измерение от данного датчика встречается раз в несколько кадров.

Смена формата кадра возникает при переключении сложной технической системы в другой режим. Новый формат характеризуется другим набором параметров, при этом внутренняя организация ГТС изменяется.

3. Постановка задачи восстановления структуры ГТС. Пусть \bar{S} — истинная структура ГТС. В общем случае \bar{S} известна и опреде-

ляется исходя из внутреннего устройства сложной технической системы. Определим через S' полученное в центре анализа ГТС описание структуры. Рассмотрим две интересующие нас ситуации:

1) если при передаче ГТС в центр анализа были допущены ошибки, или во время испытаний возникли неисправности в системе коммутаторов, S' может не соответствовать \bar{S} . В этом случае необходимо провести корректировку S' ;

2) если описание структуры было потеряно, то S' вообще не известна. В этом случае необходимо полностью восстановить структуру.

В условиях ограниченного времени получить \bar{S} не представляется возможным, поэтому в любом случае необходимо произвести восстановление структуры S ГТС. В первом случае S сравнивается с S' и на основании результатов сравнения выдвигается утверждение о наличии или отсутствии ошибок в S' . Во втором случае S принимается как \bar{S} .

Введем функцию несоответствия описания истинной структуры ГТС — $\rho(S, \bar{S})$. Тогда задачу автоматического определения структуры S можно сформулировать следующим образом: требуется найти такую структуру S , при которой $\min_{S \in \Theta} \{\rho(S, \bar{S})\}$, где Θ — множество возможных структур ГТС.

Оценить качество найденной структуры в такой формулировке затруднительно в связи с отсутствием специализированных метрик. По этой причине в [2] предлагается свести оценку близости структур к близости структурных описаний ГТС:

$$\rho(GTS, \overline{GTS}),$$

где GTS — образ ГТС, включающий в себя описание поведения параметров и описание структуры; \overline{GTS} — образ-эталон, который может быть задан экспертом или сформирован как усредненный образ ранее накопленных ГТС.

Сравнение структурных описаний ГТС позволит также уменьшить влияние ошибок восстановления структуры на расстояние между ГТС. Для структурного описания ГТС авторами предлагается использовать математический аппарат теории графов, который позволит для сравнения ГТС применить толерантные к ошибкам методы сравнения графов.

Таким образом, для решения задачи анализа структур ГТС необходимо разработать модель описания и алгоритм восстановления структуры ГТС.

4. Модели описания ГТС. В качестве моделей описания ГТС рассматриваются две основные модели:

- 1) базовая — представляет собой универсальное описание ГТС;
- 2) графовая — является описанием базовой модели в терминах графов, что позволяет применять существующие алгоритмы и инструменты работы с графами для анализа ГТС.

4.1. Базовая модель описания ГТС. Формально ГТС определяется следующим образом:

$$G = \langle P \Omega F \rangle, \quad (2)$$

где $P = \langle P_1 \dots P_N \rangle$ — множество передаваемых параметров, Ω — последовательность форматов кадров данного ГТС, F — функция назначения позиции каждому параметру.

В дальнейшем пару $\langle P F \rangle$ будем называть структурой ГТС.

Каждый параметр представляет собой временной ряд: $P_i = \rho_i(t_1), \dots, \rho_i(t_n)$, $t_1 < t_2 < \dots < t_n$, где $\rho_i(t_j)$ — значение i -го ряда в момент времени t_j . Последовательность форматов кадров — это кортеж из форматов кадров $\Omega = \langle S_1 \dots S_l \rangle$, где S_i — очередной формат кадра, а l — число различных режимов работы во время работы сложной технической системы. Каждый формат кадра состоит из трех элементов:

$$S_i = \langle L_K \text{ Sub Sup} \rangle, \quad (3)$$

где L_K — длина кадра, Sub — множество субкоммутиаций, Sup — множество суперкоммутиаций.

Множество суперкоммутиаций определяется как множество пар:

$$Sup = \left\{ (i \ f)_{0 \dots k} \right\},$$

где i — индекс слова в кадре, f — кратность суперкоммутиации (сколько раз измерение данного параметра встречается в одном кадре), k — число суперкоммутиаций в кадре.

Множество субкоммутиаций определяется как множество пар, состоящих из индекса слова в кадре и глубины субкоммутиации:

$$Sub = \left\{ (j \ d)_{0 \dots m} \right\},$$

где j — индекс слова в кадре, d — глубина субкоммутиации (сколько параметров передается в данном слове), m — число субкоммутиаций в кадре.

Каждый формат кадра задает n_i позиций для передачи параметров, что в общей сложности составляет $N = \sum_{i=1}^l n_i$. Для однозначной нумерации позиций используется следующий алгоритм:

- 1) для каждого слова в кадре выполнить пункты 2–6;
- 2) если слово является субкоммутиацией, то рекурсивно выполнить алгоритм каналов в субкоммутиаторе;
- 3) если это слово является частью суперкоммутиации, то
- 4) если это первое слово суперкоммутиации, то установить новый индекс и перейти к следующему слову;
- 5) иначе пропустить;
- 6) если это одиночный параметр, то установить новый индекс.

Пусть I_P , I_S — множества индексов параметров и индексов позиций в ГТС, тогда функция назначения позиции каждому параметру будет определяться как

$$F : I_P \rightarrow I_S, \text{ где } I_S = 1 \dots N, I_P = 1 \dots k.$$

4.2. Графовая модель ГТС со сменой формата кадра. В основу графовой модели ГТС положен нагруженный граф отношений (НГО) (*Attributed Relational Graph* — ARG), который эффективно применяется для структурного описания объекта [5, 7]. Данный граф определяется как кортеж из шести элементов:

$$G = \langle V \ E \ A_V \ A_E \ \alpha_V \ \alpha_E \rangle,$$

где V и E — множества вершин и ребер НГО; A_V и A_E — множества атрибутов для вершин и ребер; α_V и α_E — функции ассоциации для каждой вершины и каждого ребра соответствующего атрибута.

Атрибуты для узла или ребра имеют следующий вид

$$t(p_1, \dots, p_{k_t}),$$

где t — тип атрибута, который задан как элемент конечного алфавита T ; (p_1, \dots, p_{k_t}) — кортеж параметров, заданный также на конечных множествах P_1, \dots, P_{k_t} (тип каждого множества P_i зависит от предметной области).

Число атрибутов k_t и их последовательность в кортеже (p_1, \dots, p_{k_t}) зависят от типа атрибута. Могут существовать особые типы атрибутов t^0 , для которых k_t может быть равен 0, т.е. атрибут не имеет никаких параметров. Тип атрибута используется для задания различных типов узлов (или ребер), а параметры несут информацию, характеризующую узлы (или ребра) данного типа.

В [1] вводится графовая модель ГТС, основанная на нагруженном графе отношений, которая имеет вид:

$$GTS = \langle V \ E \ T_V \ A_V \ \alpha_V \ T_E \ \alpha_E \rangle, \quad (4)$$

где V и E — множества вершин и ребер НГО для представления ГТС, T_V — алфавит типов вершин, A_V — множество атрибутов вершин, T_E — множество атрибутов ребер, α_V и α_E — функции ассоциации соответствующих атрибутов.

Множество атрибутов ребер определяется как:

$$T_E = \{e^0, r^0\},$$

где t^0 — атрибут ребер «непосредственно связан», r^0 — атрибут ребер «элемент заменяется».

Множество алфавит типов вершин T_V имеет вид:

$$T_V = \{d, c\},$$

где d — тип атрибута «датчик», c — тип атрибута «коммутатор».

Атрибут для коммутаторов — это кортеж имен вершин-потомков:

$$P_C = \langle nv_1 \ \dots \ nv_m \rangle,$$

где $nv_1 \ \dots \ nv_m$ — имена вершин $v_1 \ \dots \ v_m \in V$.

Для каждой вершины v_i устанавливается связь с коммутатором

$$(c, v_1) \dots (c, v_m) \in E.$$

Для задания атрибутов датчиков предлагается применить метод гистограмм, в рамках которого строится гистограмма частоты встречаемости каждого значения информационного слова для данного параметра. Таким образом, кортеж атрибута для датчиков имеет следующий вид:

$$P_D = \langle p_0 \dots p_l \rangle, \quad (5)$$

где $p_0 \dots p_l$ — вероятность появления значений $0 \dots l$, $l = 2^{L_{ИС}}$;

При этом кортеж P_D обладает следующим свойством:

$$\sum_{k=0}^l p_k = 1.$$

Для уменьшения размерности пространства P_D можно использовать ранее рассмотренную частотно-ранговую аппроксимацию (2).

5. Общий подход к построению графовой модели ГТС. При определении структуры ГТС в общем случае априорная информация о ГТС отсутствует. Для входного структурированного бинарного потока $B = \langle I_1 \dots I_k \rangle$ информационных слов I_j , который представляет собой ГТС, определенный как (1), требуется построить графовую модель. Задача построения графовой модели ГТС подразделяется на три подзадачи:

- 1) задачу сегментации ГТС, т.е. определения моментов времени, в которые наблюдается смена кадра;
- 2) задачу построения модели каждого из участков без смены кадра;
- 3) задачу объединения совокупности построенных моделей отдельных участков.

Для решения перечисленных задач авторами разработан метод построения графовой модели ГТС, основанный на двух основных алгоритмах:

- 1) алгоритме без учета смены кадра,
- 2) алгоритме с учетом смены кадра.

Алгоритм построения графовой модели ГТС для простых структур применяется при наличии априорной информации об отсутствии смены кадра в ГТС. Алгоритм для построения модели составных

структур является расширением алгоритма построения модели для простых структур и не предъявляет требований к наличию априорной информации.

К разрабатываемым алгоритмам предъявляются следующие требования:

— точность описания результирующей структуры должна составлять не менее 98 % для ГТС, число сбойных кадров в котором не превышает 20 %;

— скорость работы алгоритма должна обеспечивать возможность экспресс-обработки данных;

— процесс формирования структуры ГТС должен быть автоматизирован и не требовать участия эксперта.

Ниже детально рассмотрен каждый из предлагаемых алгоритмов.

6. Алгоритм построения графовой модели ГТС без смены формата кадра. Алгоритм построения графовой модели ГТС предполагает последовательное выполнение двух операций: 1) определение формата кадра и 2) представление ГТС с известным форматом кадра в виде графа.

Алгоритм определения формата кадра состоит из трех последовательно применяемых алгоритмов поиска соответственно для следующих параметров: 1) длины кадра, 2) субкоммутации, 3) суперкоммутации.

В результате применения алгоритма определяются следующие основные характеристики кадра: длина кадра, глубины субкоммутаций и индексы слов, слова с суперкоммутацией и их кратность. На основании этой информации осуществляется построение графа (3) без значений атрибутов для датчиков. Для определения значений атрибутов применяется статистико-лингвистический подход.

Обобщенный алгоритм построения графовой модели ГТС состоит из следующих этапов:

1) определить длину кадра с использованием алгоритма поиска длины кадра;

2) определить схему коммутаций кадра с применением алгоритмов поиска суб- и суперкоммутаций;

3) для каждого параметра и датчика вычислить их атрибуты в модели (3);

4) построить граф ГТС с использованием алгоритма представления ГТС с известной структурой кадра в виде графа.

Для нахождения значений атрибутов необходимо декоммутировать ГТС на отдельные параметры и подсчитать частоту встречаемости

значений каждого слова, предварительно выполнив упорядочение слов. Параметры аппроксимации полученной зависимости частоты встречаемости слова от порядкового номера с использованием функции (1) представляют собой значения атрибутов параметров и датчиков в модели. Перечисленные этапы алгоритма рассмотрены ниже.

6.1. Алгоритм поиска длины кадра. Этот алгоритм основан на утверждении, что ГТС имеет явно периодический характер. В данном случае период сигнала — это длина кадра, поэтому предлагается использовать корреляционный анализ для ее нахождения.

Корреляционная функция детерминированного сигнала с конечной энергией представляет собой интеграл (в бесконечных пределах) от произведения двух копий сигнала, сдвинутых друг относительно друга на время τ [4]:

$$B_s(\tau) = \int s(t) \cdot s(t - \tau) dt \quad (6)$$

где $s(t)$ и $s(t - \tau)$ — значения сигнала в моменты времени t и $t - \tau$ соответственно.

В случае дискретного сигнала, представляющего собой последовательность показаний в моменты времени $t_i, i = 1, 2, \dots, n$, выражение (6) принимает следующий вид:

$$B_s(\tau) = \frac{1}{n - \tau} \sum_{i=\tau}^n s_i \cdot s_{i-\tau} \quad (7)$$

где s_i и $s_{i-\tau}$ — значения сигнала в моменты времени i и $i - \tau$ соответственно.

Для поиска длины кадра необходимо последовательно для всех потенциальных длин кадра Ξ вычислить значения корреляционной функции. В качестве потенциальных длин кадра в соответствие с [8] рассматриваются длины кадра от 2 до 16384/(длина информационного слова). Максимальные пики корреляции приходятся на длины, кратные фактической длине кадра.

Для дополнительного контроля авторами предлагается использовать проверку наблюдаемости маркера. Рассмотрим более подробно, на чем основана эта проверка. Как было указано выше, одна из особенностей формирования ГТС — наличие в начале каждого кадра маркера — константного параметра, значение которого неизменно на протяжении всего сеанса испытаний. При этом маркер может быть длиной в одно-два информационных слова или более. Поэтому проверка заключается в подсчете математического ожидания изменения маркера

между соседними кадрами и сравнения данного значения с пороговым значением. Формально это можно записать как:

$$E(M_i - M_{i+1}) < \delta,$$

где M_i, M_{i+1} — значения маркера в соседних кадрах, δ — пороговое значение.

Ниже предлагаемый алгоритм представлен в виде последовательности действий:

- 1) для всех потенциальных длин кадра $\xi \in \Xi$ вычислить значения корреляционной функции (7);
- 2) для максимальных значений $\xi_{\max} \in \Xi$ последовательно провести проверку по наблюдаемости маркера. В итоге будет получено множество Ξ' кратных длин кадра;
- 3) выбрать наименьшее общее кратное $\bar{\xi} \in \Xi'$ — длину кадра.

6.2. Алгоритм поиска субкоммутации. Предлагаемый алгоритм основан на определении потенциальных слов с субкоммутацией с последующим корреляционным анализе на предмет определения глубины субкоммутации. Субкоммутация используется для передачи медленно меняющихся функциональных параметров, для которых характерна плавность изменений (отсутствие разрывов в производной).

Одной из отличительных особенностей медленно меняющихся параметров является невысокое значение математического ожидания изменения значений параметра в соседних кадрах:

$$M[K_i] = \frac{1}{N} \sum_{j=1}^{N-1} |K_{i,j} - K_{i,j-1}|, \quad (8)$$

где $K_{i,j}$ — значение i -го слова в j -м кадре, N — длина ГТС в кадрах.

При субкоммутации двух медленно меняющихся параметров в результирующем параметре появляются разрывы в производной и резко возрастает математическое ожидание изменения значений параметра в соседних кадрах. Слова, для которых значение функции (8) превышает заданное пороговое значение, относятся к словам с потенциальной субкоммутацией.

После того как выделены потенциальные слова с субкоммутацией, определяется глубина субкоммутации. Для этого строится временной ряд из значений каждого из параметров, и определяются значения корреляционной функции (4). Для построенной корреляционной функции находится период, который и является глубиной субкоммутации.

Алгоритм поиска субкоммутаций можно представить в виде следующей последовательности действий:

- 1) вычислить математическое ожидание (8);
- 2) выделить множество Θ максимальных значений функции (6.3) в соответствие с порогом σ ;
- 3) для всех $\theta \in \Theta$ вычислить корреляционную функцию (7) и найти период корреляции τ ;
- 4) если $\tau > 1$, то добавить субкоммутацию к найденным, в ином случае считать, что субкоммутация отсутствует.

6.3. Алгоритм поиска суперкоммутации. Данный алгоритм, аналогично алгоритму поиска субкоммутаций, использует свойство плавности изменения медленно меняющихся параметров.

Суперкоммутация представляет собой передачу одного и того же параметра через постоянное число слов. Пусть параметр 1 передается с кратностью суперкоммутации 2. Тогда, если рассмотреть последовательность половин кадров для всего ГТС, то математическое ожидание изменения элементов в столбце с суперкоммутацией будет малой по сравнению с изменениями в словах без суперкоммутации.

Предположим, что в i -м слове имеется j -й кратности суперкоммутация, тогда для подкадров длиной L_K/q будет наблюдаться достаточно малое среднее значение изменения i -го слова, в то время как при длине кадра, кратной $j \cdot n$, где $n = 2, 3, \dots$, будет наблюдаться резкий скачек среднего значения изменения слов.

Будем называть кратности, соответствующие простым числам, простыми кратностями. Интуитивно понятно, что все возможные кратности раскладываются на простые кратности. Поэтому нет необходимости проверять кратность $j \cdot n$, если в данном слове при j -й кратности не наблюдается суперкоммутация. В качестве проверки на наличие суперкоммутаций в словах авторами предлагается использовать корреляционный анализ. Отсутствие периода корреляции говорит о наличии в данном слове суперкоммутации.

Алгоритм поиска суперкоммутации можно описать следующим образом:

- 1) пусть дана очередь потенциальных кратностей суперкоммутаций Q и задано множество найденных суперкоммутаций R ;
- 2) добавить в очередь Q все возможные простые кратности в порядке возрастания;
- 3) R — пусто;

- 4) пока очередь Q не пуста, выполнять шаги 5–11;
- 5) взять очередным элемент $q \in Q$;
- 6) составить последовательность подкадров длиной L_K/q ;
- 7) для данной последовательности вычислить Δ — математическое ожидание изменения слова (5);
- 8) если $\Delta < \rho$, где ρ — некоторая пороговая величина, то вычислить τ период корреляции параметра. иначе перейти к 5;
- 9) если период $\tau \neq 1$, то перейти к шагу 5;
- 10) в ином случае добавить номер слова и кратность в R ;
- 11) добавить в Q все кратности, которые отсутствуют и можно получить из чисел в R .

6.4. Алгоритм представления ГТС с известной структурой кадра в виде графа. Если обратиться к базовой модели, то видно, что вся необходимая информация для построения графа присутствует: в данном случае это формат кадра (3). Алгоритм построения следующий:

- 1) добавить в граф вершину C — коммутатор;
- 2) добавить в граф $L_K - |Sub| - \sum_{(i, f) \in Sup} f + |Sup|$ вершин датчиков;
- 3) для каждого добавленного датчика вычислить значения атрибутов (7);
- 4) соединить все добавленные вершины с коммутатором C ребрами с атрибутом c ;
- 5) для каждой субкоммутации из Sub выполнить данный алгоритм рекурсивно;
- 6) соединить добавленные субкоммутаторы с коммутатором C ребрами с атрибутом c ;
- 7) собрать атрибут для коммутатора, выполнив для каждого слова в кадре выполнить шаги 8–10;
- 8) если слово является субкоммутацией, то добавить в атрибут имя субкоммутации;
- 9) если это слово является первым словом суперкоммутации, то добавить в атрибут имя данной суперкоммутации;
- 10) если это одиночный параметр, то добавить имя параметра.

7. Алгоритм построения графовой модели ГТС со сменой формата кадра. Этот алгоритм предполагает последовательное выполнение следующих операций:

- 1) сегментация ГТС;

2) нахождение формата кадра для участков, на которых сохраняется формат кадра;

3) объединение графовых моделей сегментов ГТС.

Задача поиска формата кадра для участков, на которых сохраняется формат кадра, решается с использованием алгоритма построения графовой модели ГТС без учета смены кадра.

Для сегментации ГТС стандартные методы сегментации неприменимы, так как ГТС носит не функциональный (природный) характер, а является сложным объединением различных процессов, каждый из которых обладает своими квазистационарными и переходными участками. В качестве критерия разбиения ГТС на сегменты используется степень различия в формате кадра ГТС на различных участках (сегментах), представленных в виде графовой модели. Для определения степени отличия графовых моделей авторами предлагается использовать метрики расстояния для графов, в частности, расстояние редактирования [5–7].

Для решения задачи сегментации ГТС осуществляется разбиение ГТС на участки одинаковой длины, равной длине минимально необходимой для нахождения субкоммутаций. Длина такого участка определяется как $L_K \cdot \varphi_{\max}$, где L_K — длина кадра, φ_{\max} — максимальная глубина субкоммутации. Для каждого участка осуществляется построение графовой модели ГТС (3). Для определения моментов смены структуры ГТС выполняется сравнение графовых моделей. Сходство графовых моделей определяется с использованием расстояния редактирования графов и говорит о сходстве их структур, участки с одинаковыми структурами относятся к одному сегменту.

В алгоритме осуществляется поиск троек последовательных участков (Y_1, Y_2, Y_3) , для которых выполняется условие:

$$d(Y_1, Y_2) + d(Y_2, Y_3) \ll d(Y_1, Y_3).$$

Такое неравенство означает, что на участке Y_2 произошла смена формата кадра. Если неравенство не выполняется, то три участка относятся к одному и тому же сегменту.

Ниже представлен алгоритм сегментации ГТС:

1) представить ГТС в виде последовательности участков размером $L_K \cdot \varphi_{\max}$;

2) для каждого участка построить графовую модель (3);

3) для каждой тройки последовательно расположенных участков (Y_1, Y_2, Y_3) определить расстояния редактирования $d(Y_1, Y_2), d(Y_2, Y_3), d(Y_1, Y_3)$;

4) выполнить проверку условия $d(Y_1, Y_2) + d(Y_2, Y_3) \ll d(Y_1, Y_3)$.

В случае, если условие выполнено, то пометить Y_2 как участок, на котором происходит смена формата кадра, и исключить его из дальнейшего рассмотрения. В противном случае объединить (Y_1, Y_2, Y_3) в один сегмент, обновив графовую модель для данного сегмента;

5) произвести объединение графовых моделей с помощью добавления ребер «элемент заменяется на» для каждого элемента, который изменился при переключении.

8. Экспериментальные исследования. Рассмотрим на примере процесс построения графовой модели ГТС. Для экспериментальных исследований был взят ГТС длиной слова 8 бит со сменой формата кадра. С ГТС были последовательно выполнены следующие операции: ГТС был разбит на участки, для каждого участка определена длина слова, определены субкоммутации, определены суперкоммутации и построена графовая модель. Ниже подробно рассмотрен анализ первого участка.

Определение длины кадра ГТС. Для ГТС были рассчитаны значения корреляционной функции (4), полученная функция представлена на рис. 1.

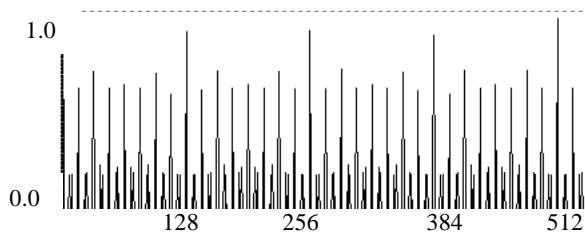


Рис. 1. Значение корреляционной функции.

Пики корреляционной функции приходятся на смещения (упорядочены по убыванию): 512, 128, 256, 768, 896, 384, 640, 288. Последовательно все смещения проверены на наблюдаемость маркера. Данное условие выполнено для смещений равных 512, 128, 256, 768, 896, 384, 640. На рис. 2 представлен график частоты изменения первых бит при длине кадра 128. Видно, что для первых 16 бит частота изменения практически равна нулю, т.е. они являются словами маркера.

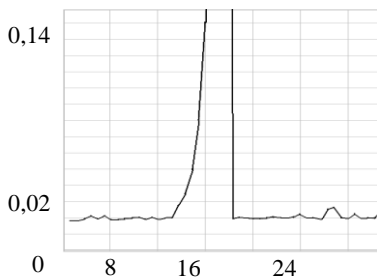


Рис. 2. Частота изменения бит в кадре.

Наименьшее общее кратное для найденной последовательности смещений равно 64, следовательно, длина кадра принимается равной 64.

Поиск субкоммутаций. Для поиска субкоммутаций осуществляется построение функции математического изменения слов в кадре (рис. 3).

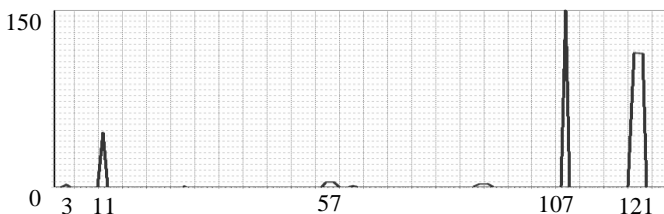


Рис. 3. Математическое изменение слов в кадре.

На рис. 2 необходимо выделить пики в словах 3, 11, 57, 58, 59, 107, 121, 122, 124 — они являются потенциальными словами с субкоммутацией.

Графики параметра и корреляции для параметра в 11-м слове даны на рис. 4. Пики в словах наблюдаются в словах, кратных 8, что свидетельствует о наличии субкоммутации глубиной 8 в 11-м слове.

Поиск суперкоммутаций. Во время поиска суперкоммутаций последовательно рассматриваются подкадры, длины которых субкратны длине кадра. На рис. 5 представлено математическое ожидание для подкадров длиной 64 (слева) и 32 (справа).

Четвертое слово, как при полукадре длиной 64, так и при 32 имеет одинаково малое среднее изменение, т.е. данное слово является частью суперкоммутации кратностью как минимум 4.

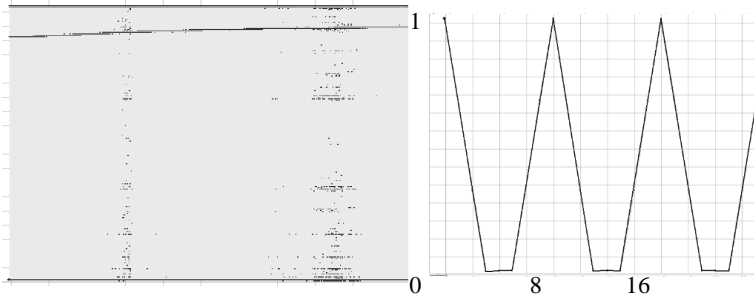


Рис. 4. Графики параметра (слева) и корреляции (справа) для параметра в 11-м слове.

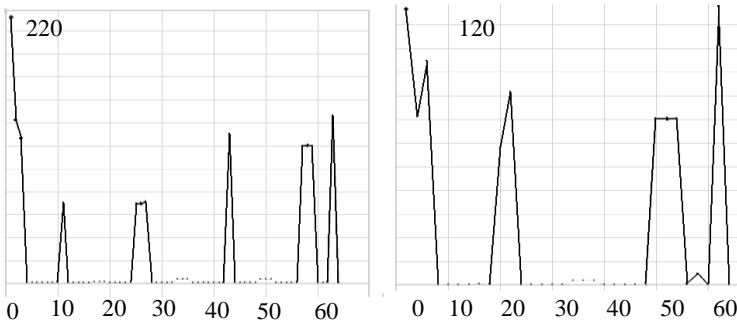


Рис. 5. Математическое ожидание для подкадров длиной 64 (слева) и 32 (справа).

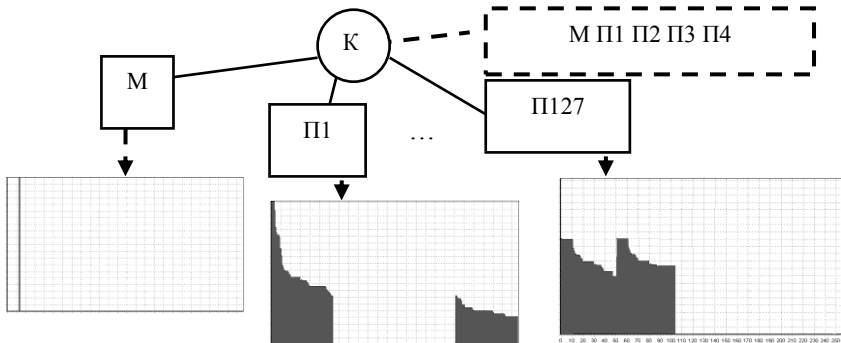


Рис. 6. Фрагмент дерева коммутации с добавленными атрибутами.

Далее для каждого параметра подсчитывается частота встречаемости слов, и преобразуются полученные ранее значения субкоммутации суперкоммутации в дерево коммутации. На рис. 6 дан фрагмент дерева коммутации с добавленными атрибутами.

На рис. 7 представлен параметр с субкоммутацией, на котором хорошо видны моменты переключения режимов системы: вертикальными линиями выделены сегменты, в которых были определены переключения режимов работы системы. Видно, что вследствие наличия в начале сеанса и в самом конце большого число помех были выделены некорректные сегменты (первый и последние).

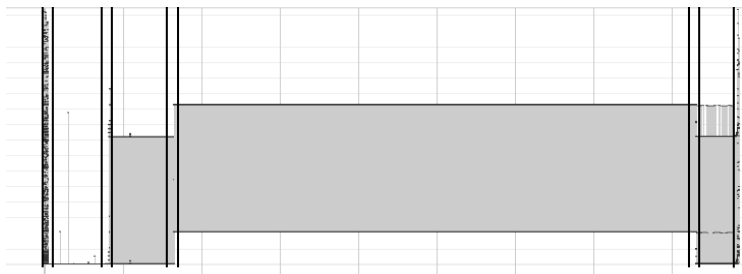


Рис. 7. Результат сегментации ГТС.

Для проверки эффективности предложенных алгоритмов проведены экспериментальные исследования на 40 модельных ГТС, сгенерированных на основе реальных параметров. При генерации ГТС на 20 ГТС накладывался 20%-й шум, остальные ГТС содержали порядка 2 % шума. Результаты экспериментальных исследований приведены в табл. 1.

Таблица 1. Эффективность алгоритмов, %

Характеристика ГТС	Алгоритм поиска		
	длины кадра	субкоммутаций	суперкоммутаций
Стандартные ГТС (уровень шума порядка 2 %)	100	96	95
Зашумленные ГТС (уровень шума от 2 до 20 %)	85	85	84
Сильно зашумленные ГТС (уровень шума выше 25 %)	68	65	62
Большое число суперкоммутаций (около 50 %)	100	96	92
Большое число субкоммутаций (около 50 %)	100	94	93

9. Заключение. В статье определена задача восстановления структуры ГТС. Для ее решения предложена графовая модель описания ГТС, которая позволяет представить ГТС как совокупность струк-

туры коммутации датчиков и характеристик передаваемых параметров, и метод ее построения. Такое представление позволяет перейти от описания потока телеметрической информации к структурному описанию поведения всей системы в целом.

В статье представлен общий подход к построению графовой модели ГТС. Проведенные экспериментальные исследования подтвердили эффективность предлагаемого подхода при условии содержания шума менее 20 %. При уровне шума 20–25 % полученная структура требует уточнения и дополнения для дальнейшего использования.

Дальнейшее направление развития рассматриваемого подхода связано, в первую очередь, с повышением устойчивости алгоритмов к уровню шуму в ГТС.

Литература

1. *Балтрашевич В. Э., Соколов И. С.* Применение графовой модели для представления структуры группового телеметрического сигнала // Сб. докл. науч.-техн. конф. профессорско-преподавательского состава университета: студентов, аспирантов и молодых ученых. Санкт-Петербург, 26 января—6 февраля 2010. СПб.: Изд-во СПбГЭТУ «ЛЭТИ», 2010. 98–103 с.
2. *Балтрашевич В. Э., Витол А. Д., Жукова Н. А., Соколов И. С.* Интеллектуальный подход к анализу структуры и семантики ГТС // Тр. 12-й национальн. конф. по искусственному интеллекту с международным участием КИИ-2010, 20–24 сентября 2010 г., г. Тверь, Россия. Т. 2. М., 2010. С. 47–56.
3. *Гептнер В. В., Горбачева Н. В., Жукова Н. А., Соколов И. С.* Идентификация телеметрических параметров с использованием нейронных сетей // Нейрокомпьютеры: разработка, применение. 2009, № 11. С. 39–44.
4. *Айфичер Э., Джервис Б.* Цифровая обработка сигналов: практический подход: Пер. с англ. М.: Вильямс, 2004
5. *Bunke H., Allermann G.* Inexact graph matching for structural pattern recognition // Pattern Recognition Let. 1983. 1. P. 245–253 с.
6. *Sanfeliu A., Fu K.* A distance measure between attributed relational graphs for pattern recognition // IEEE Transactions on Systems, Man, and Cybernetics (Pt. B). 1983. 13 (3) P. 353–363.
7. *Riesen K, Bunke H.* Approximate graph edit distance computation by means of bipartite graph matching // Image and Vision Computing. 2009.27. P. 950–959.
8. Telemetry Group. Document 106-05 Part I: Telemetry standards, Range Commanders Council, White Sands Missile Range, New Mexico, 88002, 2005.

Жукова Наталия Александровна — канд. техн. наук, сотрудник научно-исследовательской лаборатории объектно-ориентированных геоинформационных систем Учреждения Российской академии наук Санкт-Петербургский институт информатики и автоматизации РАН (СПИИРАН), ассистент кафедры математического обеспечения ЭВМ факультета компьютерных технологий и информатики Санкт-Петербургского государственного электротехнического университета (СПбЭТУ). Область научных интересов: методы статистической и интеллектуальной обработки сложных нестационарных процессов, проектирование и разработка распределенных систем обработки измерительных данных. Число научных публикаций — 40. gna@oogis.ru.; СПИИРАН,

14-я линия В.О., д. 39, Санкт-Петербург, 199178, РФ; р.т. +7(812)328-3337, факс +7(812)328-4450.

Zhukova Nataly Alexandrovna — PhD in CS, researcher, Research laboratory of object-oriented geo-information, St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences (SPIIRAS), lecturer, Software Department, Faculty of Computing Technologies and Informatics, St. Petersburg Electrotechnical University (SPETU). Research interests: methods of statistical and intellectual processing of non-stationary processes, development of distributed systems for measuring data processing. The number of publications — 40. gna@oogis.ru; SPIIRAS, 39, 14th Line V.O., St. Petersburg, 199178, Russia; office phone +7(812) 328-01-79, fax +7 (812) 329-08-63.

Соколов Игорь Сергеевич — аспирант кафедры математического обеспечения ЭВМ факультета компьютерных технологий и информатики Санкт-Петербургского государственного электротехнического университета (СПбЭТУ), заместитель начальника научно-исследовательского отдела Научно-инженерный центр Санкт-Петербургского электротехнического университета — ОАО «НИЦ СПб ЭТУ». Область научных интересов: интеллектуальная обработка данных, распознавание образов с применением теории графов, программная инженерия, обработка телеметрической информации, онтологический инжиниринг. Число научных публикаций — 6. igor.s.sokolov@gmail.com, СПбГЭТУ ул. Профессора Попова, д. 5, Санкт-Петербург, 197376, РФ, р.т. +7(904)619-4830.

Sokolov Igor Sergeevich — P.G. of Software Department, Faculty of Computing Technologies and Informatics, St. Petersburg Electrotechnical University (SPETU), head deputy of the research department of Open Joint Stock Company «Research & Engineering Center of Saint-Petersburg Electrotechnical University» (R&EC ETU). Research interests: Data Mining, pattern recognition using graph theory, software engineering, processing of telemetry data, ontological engineering. The number of publications — 6. igor.s.sokolov @ gmail.com, ETU Professor Popov street, Building 5, St. Petersburg, 197376, Russia, office phone +7 (904) 619-4830.

Рекомендовано СПИИРАН, НИЛ ООГИС, д -р техн. наук, проф. В.Ю. Осипов.
Статья поступила в редакцию 10.12.2010.

РЕФЕРАТ

Жукова Н.А., Соколов И.С. Метод восстановления группового телеметрического сигнала на основе графовой модели.

Оценка функционирования сложного технического объекта во время испытаний производится на основе анализа регистрируемых датчиками параметров. Параметры передаются с объекта в центр обработки данных в форме структурированного бинарного потока данных. Этот информационный поток принято называть групповым телеметрическим сигналом (ГТС). ГТС имеет достаточно сложную структуру, что связано с необходимостью передавать параметры разных типов с разной частотой. Каждый ГТС сопровождается описанием структуры, которое позволяет на этапе обработки и анализа декомпозировать сложный информационный поток на отдельные параметры. При передаче ГТС в ряде случаев сопровождающие их описания структуры отсутствуют или содержат неточности. В связи с чем возникает необходимость восстановления или проверки корректности их описания.

Существующие алгоритмы автоматизированного формирования описания ГТС, основанные на применении статистико-лингвистического подхода, не обеспечивают требуемой точности, кроме того, для их работы необходима база данных ГТС, содержащая все возможные типы ГТС.

В качестве моделей описания ГТС в статье предлагаются две основные модели — базовая и графовая. Базовая модель описания ГТС представляет собой универсальное описание ГТС. Графовая модель описания ГТС является описанием базовой модели в терминах графов, что позволяет применять существующие алгоритмы и инструменты работы с графами для анализа ГТС. В основу графовой модели ГТС положен нагруженный граф отношений, который эффективно применяется для структурного описания объекта.

Предлагаемый подход к восстановлению структуры ГТС для входного структурированного бинарного потока информационных слов основан на построении графовой модели. В рамках данного подхода последовательно решаются три подзадачи:

- 1) задача сегментации ГТС, т.е. определение моментов смены кадра;
- 2) задача построения модели каждого из участков без смены кадра;
- 3) задача объединения совокупности построенных моделей отдельных участков.

Для проверки эффективности предложенных алгоритмов проведены экспериментальные исследования на 40 модельных ГТС, сгенерированных на основе реальных параметров. При генерации ГТС на 20 ГТС накладывался 20%-й шум, остальные ГТС содержали порядка 2 % шума. Проведенные экспериментальные исследования подтвердили эффективность предлагаемого подхода при условии содержания шума менее 20 %. При уровне шума 20–25 % полученная структура требует уточнения и дополнения для дальнейшего использования.

SUMMARY

Zhukova N.A., Sokolov I.S. Method of reconstructing the structure of the group telemetric signals based on the graph model.

The assessment of complex technical object functioning during tests is produced on the basis of parameters analysis registered by sensors. Parameters are transferred from object to the data processing center in the form of the structured binary data stream. It is assumed to name this information stream a group telemetric signal (GTS). GTS has a rather complicated structure since there is necessity to transfer parameters of different types with different frequency. Every GTS is accompanied by structure declaration which at the stage of processing and analyzing provides decomposition information for obtaining separate parameters from the complex information stream. In some cases when transferring GTS its accompanying structure declaration may be missing or containing inaccuracies. Because of this, there is a necessity of sanity check or reconstruction of this declaration.

Existing algorithms of the automated GTS declaration construction which are based on the usage of the statistical and linguistic approach don't provide required accuracy. Besides, a GTS database, containing all possible GTS types, is necessary for their operation.

This paper suggests two main models as models of GTS — the basic model and the graph model. The basic model of GTS represents the general-purpose GTS description. The graph model is the description of basic model in terms of graphs that allows applying existing algorithms and graph processing tools to the GTS analysis. GTS graph model is based on attributed relational graph which is efficiently utilized for the structural description of an object.

The offered approach to the GTS structure reconstruction for the input structured binary stream builds the graph model. Within the offered approach three sub-tasks sequentially are being solved:

- GTS segmentation task, i.e. determination of the moments of frame change;
- construction of the independent models for each segment;
- merging of constructed models set.

The efficiency of the offered algorithms was confirmed by experimental researches on 40 GTS, generated on the basis of real parameters. At generation of research data on 20 GTS 20 % noise was superimposed. On other GTS 2 % noise was superimposed. The experimental research confirmed the efficiency the offered approach provided that the amount of noise is less than 20 %. At the amount of 20–25 % the received structure demands adjustment for the further usage.