

УДК 004.891

ИНТЕЛЛЕКТУАЛЬНЫЕ ТЕХНОЛОГИИ СЛИЯНИЯ ДАННЫХ ПРИ ДИАГНОСТИРОВАНИИ ТЕХНИЧЕСКИХ ОБЪЕКТОВ

С.М. Ковалев¹, А.Е. Колоденкова², В. Снасель³

¹ Ростовский государственный университет путей сообщения, Ростов-на-Дону, Россия
ksm@rfniias.ru

² Самарский государственный технический университет, Самара, Россия
anna82_42@mail.ru

³ Остравский технический университет, Острава, Чешская Республика
vaclav.snasel@vsb.cz

Аннотация

Слияние разнородных данных, полученных в реальном времени от различных датчиков, является важной задачей при диагностировании технических объектов. В статье рассмотрены вопросы терминологии слияния данных на основе обзора литературы, предложено новое определение термина «слияние данных». Обобщены и систематизированы научные взгляды на проблему слияния данных при диагностировании технических объектов в условиях множества разнотипных датчиков и разнородной информации. Приведена адаптированная классификация слияния данных с учётом различных критериев (отношения между датчиками, уровень абстракции данных, тип архитектуры), а также классификация структурных моделей слияния данных, разработанных для построения интеллектуальных систем слияния данных. Проведён сравнительный анализ моделей процесса слияния данных, представлены их структуры, выявлены достоинства и недостатки моделей. Отмечено, что для эффективного сбора исходных данных, поступающих от множества разнотипных датчиков, и их обработки можно использовать несколько моделей слияния данных или их комбинации. Все основные аспекты, касающиеся интеллектуальных технологий слияния данных, рассмотрены в столь полном объёме на русском языке впервые.

Ключевые слова: разнородные данные, модели слияния данных, диагностирование технических объектов, множество разнотипных датчиков.

Цитирование: Ковалев, С.М. Интеллектуальные технологии слияния данных при диагностировании технических объектов / С.М. Ковалев, А.Е. Колоденкова, В. Снасель // Онтология проектирования. – 2019. – Т.9, №1(31). – С.152-168. – DOI: 10.18287/2223-9537-2019-9-1-152-168.

Введение

С развитием сложности, масштабности, многофункциональности технических объектов (ТО) первостепенное значение приобретает проблема эффективной и безопасной эксплуатации ТО. Это связано с тем, что ухудшение технического состояния на этапе эксплуатации ТО приводит к нарушению его работоспособности (отказу), выражающемуся в изменении значений параметров его работы за пределами, регламентируемыми нормативно-техническими документами, или к полной его остановке, что может привести к значительному материальному ущербу, а также нарушению экологической обстановки.

Одним из способов решения данной проблемы является автоматизация и информатизация процессов контроля и диагностирования, в ходе которых решают такие задачи, как определение вида технического состояния, обнаружение неисправностей, прогнозирование технического состояния в условиях использования различных типов датчиков. Увеличение количества различных типов датчиков, разнородных диагностических данных привело к необ-

ходимости применения технологии слияния данных. Разработка этой технологии для диагностирования ТО даёт выигрыш по сравнению с отдельной обработкой за счёт расширения объёма получаемых диагностических данных и синергетического эффекта. Достоинствами данной технологии являются: сохранение надёжности ТО в случае отказа какого-либо датчика; рост вероятности обнаружения нештатных ситуаций; уменьшение времени принятия решений при диагностировании ТО в силу получения более полной и точной информации о ТО, поступающей от множества разнородных датчиков.

Цель статьи - дать анализ существующих технологий слияния данных для решения задач технического диагностирования ТО. Авторы попытались не только обобщить научные взгляды на проблему слияния данных [1-5], но и представить свой опыт, полученный при диагностировании ТО.

1 Проблемы слияния данных

С увеличением количества типов датчиков данных и объёма информации возникла необходимость в обработке разнородных данных для последующего анализа. Обработку таких данных понимают как соединение, комплексирование, интеграцию, объединение, сращивание (в англоязычной литературе часто называют «слияние данных» - *data fusion*) [3-7].

В настоящей работе под *слиянием данных* понимается совместная обработка разнородной информации, полученной в реальном времени от множества различных датчиков, в информацию, легко воспринимаемую человеком, на основе которой можно контролировать ТО, проводить оценку ситуации и принимать научно-обоснованные управленческие решения [8]. Несмотря на многочисленные применения технологии слияния разнородных данных при диагностировании ТО, остаётся ряд основных проблем.

- *Разнородность данных.* Данные от датчиков могут быть представлены в виде разнородных данных и характеризоваться [9]: разнообразием шкал измерения (номинальной, числовой и др.); различными типами данных (целый, вещественный, логический и др.); различной структурой представления (статистические, темпоральные, нечёткие, экспертные данные, изображения и др.); различными типами моделей баз данных (реляционные, иерархические, сетевые и др.); различной степенью достоверности, полноты и точности данных, измеряемых в различных масштабах и единицах измерения.
- *Точность датчиков.* Часто применяются многопараметрические датчики, которые способны измерять одновременно несколько физических величин, характеризующих контролируемый объект или процесс. Объединяя данные от нескольких одинаковых датчиков, можно улучшить их качество, а также повысить их точность и достоверность.
- *Выбор метода слияния данных.* Не существует универсального метода слияния данных, который был бы оптимален при всех условиях.
- *Терминология.* Спецификации данных выполняются различными экспертами, которые могут использовать одинаковые термины в различном смысле и, наоборот, для одного и того же понятия могут использовать различные названия. Для решения данной проблемы целесообразно использовать единый словарь терминов предметной области.
- *Идентификация сущностей.* Информация об одной и той же сущности (ситуации, состоянии объекта, процессе и т.п.) представляется в распределённой форме [6, 9]. Для решения проблемы целесообразно установить однозначное соответствие между описаниями, которые находятся в различных локальных датчиках, но относятся к одному экземпляру сущности предметной области.

2 Классификация слияния данных

Ввиду междисциплинарности области исследований предлагается слияние данных классифицировать по нескольким критериям (рисунок 1 адаптирован по материалам [10-12]). Слияние данных по критерию отношений между датчиками представлено на рисунке 2 (адаптирован по материалам [2, 13, 14]).



Рисунок 1 – Классификация слияния данных

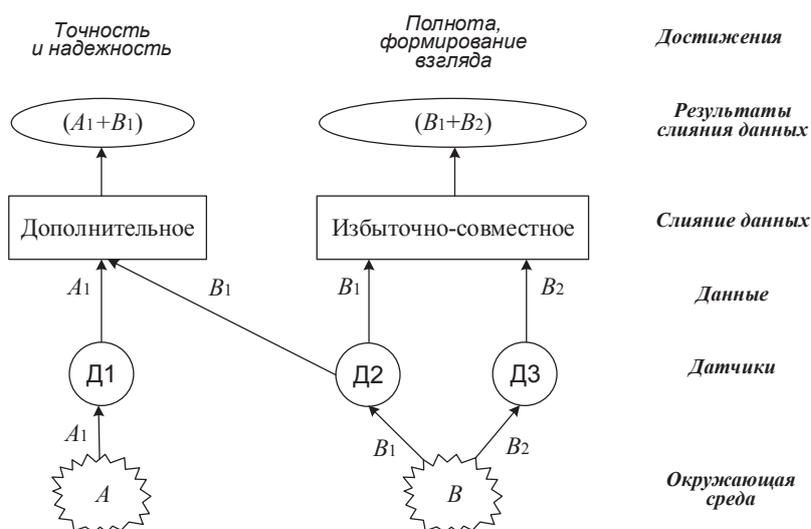


Рисунок 2 – Типы слияния данных на основе отношений между датчиками

Дополнительное (complementary) слияние – это слияние, при котором независимые датчики (D_1 и D_2 , принадлежащие различным частям пространства окружающей среды) могут быть объединены ($A_1 + B_1$) для того, чтобы дать более полное представление наблюдаемому явлению (ситуации). Данные от независимых датчиков легко объединяются.

Избыточно-совместное (competitive-cooperative) слияние – это слияние, где каждый датчик (D_2 и D_3) поставляет информацию об одной и той же характеристике ТОО, в результате которого достигается более точное представление этой информации ($B_1 + B_2$). Различают две возможные конфигурации избыточно-совместного слияния данных: с разных датчиков или слияние измерений с одного датчика, сделанных в разные моменты времени. Избыточно-совместное слияние используется для отказоустойчивых и надёжных систем.

Согласно *уровню абстракции данных (level of abstraction)* слияние данных может быть классифицировано на три уровня (см. таблицу 1) [13].

- Низкий уровень слияния (*low-level fusion*), именуемый часто уровнем необработанных данных (*raw data level*), поступающих от датчиков. На данном уровне может применяться нечёткий фильтр Калмана.
- Средний уровень слияния (*medium-level fusion*), именуемый часто уровнем свойств (*feature level*). На данном уровне происходит слияние свойств (форма, рёбра, углы, линии, положение), в результате которого получают новые объекты, используемые для других задач, например, сегментации и распознавания. Также происходит фильтрация, нормализация, корреляция, классификация данных с использованием методов «мягких вычислений» и методов интеллектуального анализа данных.

- Высокий уровень слияния (*high-level fusion*), именуемый часто уровнем решения (*decision level*). Слияние на уровне решения происходит с использованием интеллектуальных методов: теории нечётких множеств, нейронных сетей, теории Демпстера-Шефера (*Dempster-Shafer*), Байесовского подхода, гибридных подходов, в результате которых получается совместное решение о диагностировании ТО с учётом разнородных данных, полученных от нескольких датчиков, а также мнений и решений нескольких экспертов.

Таблица 1 – Основные характеристики уровней слияния данных

Характеристики \ Уровни	Низкий	Средний	Высокий
Объём передаваемой информации	очень большой	средний	очень маленький
Информационные потери	без потерь	немного	средние
Математические вычисления	простые	средние	сложные

Как видно из таблицы 1, потеря информации увеличивается по мере её передачи с низких уровней на более высокие. При этом информации на низком уровне больше, чем на высоком, что делает методы низкого уровня дорогостоящими за счёт требований, предъявляемых к точности их вычислений.

На рисунках 3-5 показаны три типа архитектур слияния данных с несколькими датчиками (рисунки 3-5 адаптированы по материалам [14-16]).

- *Централизованная архитектура (centralized architecture)* (рисунок 3).

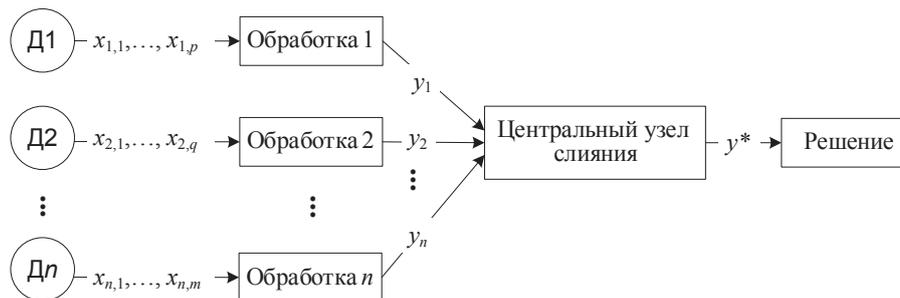


Рисунок 3 – Централизованная архитектура

Все собранные данные от датчиков (D_1, D_2, \dots, D_n) будут обработаны и отправлены в центральный узел слияния, который способен эффективно обрабатывать большие объёмы данных. Если данные правильно выровнены и связаны, а объём передаваемой информации не ограничен, то централизованная архитектура позволяет дать теоретическое оптимальное решение для оценки состояния ТО. Однако обработка всей информации на центральном узле создаёт проблемы, такие как большая вычислительная нагрузка на центральный узел, возможность сбоя работы центрального узла, негибкость изменений в архитектуре [1, 11].

Если данные представляют собой изображения, то это приводит к низкой пропускной способности отправки необработанных данных и задержки с выводом данных.

- *Децентрализованная архитектура (decentralized architecture)* (рисунок 4).

Данная архитектура состоит из нескольких локальных процессоров, которые в соответствии со своим наблюдением вычисляют параллельно локальные оценки. Центральный центр слияния отсутствует, поэтому сначала каждый локальный процессор вычисляет локальную оценку, а затем корректирует её путём интегрирования локальных оценок других процессоров для получения точной и полной оценки. Результатом применения децентрализованной архитектуры является множество решений, из которых выбирается одно наилучшее. Увеличение числа узлов приводит к увеличению стоимости связи, что является одним из недостатков децентрализованной архитектуры.

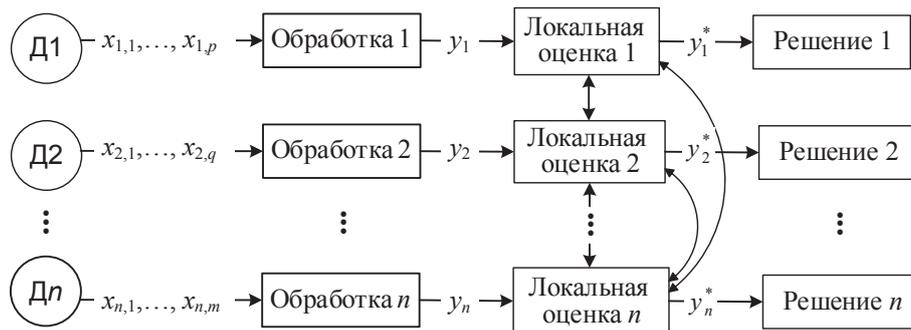


Рисунок 4 – Децентрализованная архитектура

- *Распределённая архитектура (distributed architecture) (рисунок 5).*

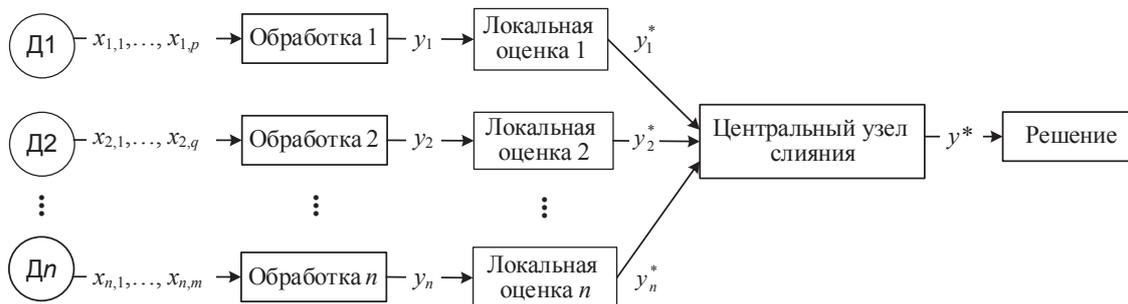


Рисунок 5 – Распределённая архитектура

Данная архитектура состоит из нескольких локальных процессоров, вычисляющих параллельно локальные оценки. Затем все собранные локальные оценки отправляются для обработки в центральный узел слияния с целью получения совместного решения. При этом локальные процессоры не сливают свои оценки с оценками, полученными от других процессоров, как в децентрализованной архитектуре, а передают их сразу в центральный узел слияния. При распределённой архитектуре за счёт распределения нагрузки достигается низкая нагрузка на каждый локальный процессор при вычислении оценки; более низкая стоимость связи, гибкость к изменениям и устойчивость к неудачам.

3 Структурные модели слияния данных

Получение нового качества данных, а именно полных, точных, своевременных данных, которые поддаются интерпретации, является наиболее сложным этапом преобразования данных интеллектуальных систем. Одним из путей решения данной проблемы явилась разработка моделей слияния данных. В настоящее время предложено несколько моде-

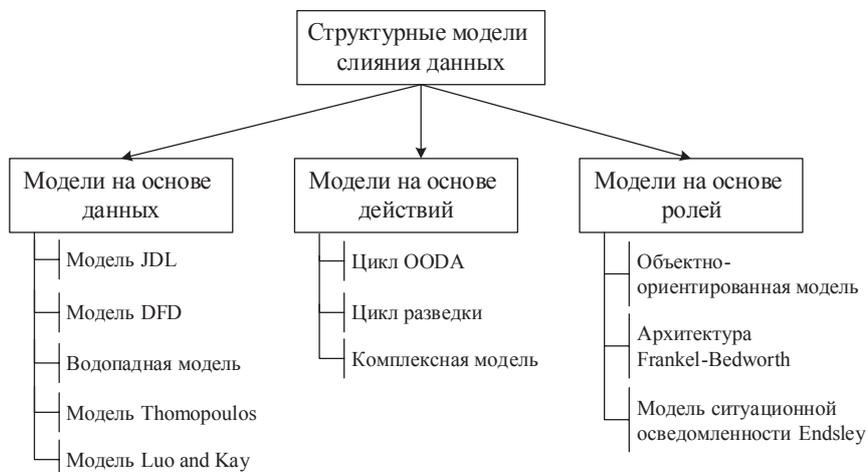


Рисунок 6 – Классификация моделей слияния данных

лей слияния данных, разработанных для выделения спецификаций, предложений и использования слияния данных в беспроводной сенсорной сети, которая представляет собой распределённую самоорганизующуюся сеть множества датчиков и исполнительных устройств, объединённых между собой посредством радиоканала. Модели слияния данных можно разделить на три основных вида: модель на основе данных, модель на основе действий и модель на основе ролей, как показано на рисунке 6 (адаптирован по материалам [17, 18]).

3.1 Модели на основе данных (data-based model)

3.1.1 Модель JDL (Joint Directors of Laboratories)

Самой известной моделью в области слияния данных является модель JDL, предложенная исследовательской группой *Joint Directors of Laboratories* совместно с Министерством обороны США [18]. Модель рассчитана на связи между исследователями слияния данных и инженерами по внедрению [19-21].

На рисунке 7 представлена структура модели слияния JDL. Модель JDL включает следующие уровни обработки данных.

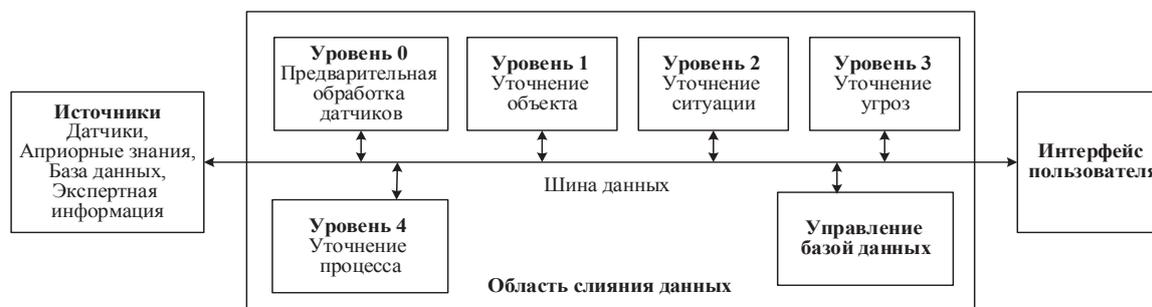


Рисунок 7 – Структура модели слияния JDL

- *Уровень 0* (предварительная обработка датчиков, *source preprocessing*) направлен на уменьшение объёма данных для последующих уровней обработки за счёт распределения данных между подходящими процессами и выбора подходящих датчиков.
- *Уровень 1* (уточнение объекта, *object refinement*) преобразует данные в согласованную структуру. На данном уровне происходит идентификация данных с применением, например, методов классификации или распознавания, а также локализация датчика; применяются алгоритмы отслеживания местонахождения объекта.
- *Уровень 2* (уточнение ситуации (*situation refinement*)) обеспечивает контекстное описание отношений между объектами и исследуемым (наблюдаемым) событием. На данном уровне используется априорная информация, знания и сведения об окружающей среде.
- *Уровень 3* (уточнение угроз (*threat refinement*)) оценивает текущую ситуацию и на её основе выявляет возможное возникновение угроз, уязвимостей и возможностей для оперативных действий.
- *Уровень 4* (уточнение процесса (*process refinement*)) отвечает за мониторинг работы системы, распределение ресурсов в соответствии с заданными целями. Могут использоваться, например, теория полезности, линейное программирование, методы, основанные на знаниях.

Модель JDL рассматривает слияние данных с точки зрения системного подхода и даёт целостное представление о структуре системы. При этом она фокусируется на данных (ввод/вывод), а не на обработке; роль человека в процессе слияния не представлена [1, 19]. Методы и алгоритмы слияния данных модели JDL, рассмотрены в работе [16].

3.1.2 Модель DFD (Data-Feature-Decision)

Предложенная В. Дасараты (*V. Dasarathy* [22]) модель DFD (рисунок 8) определяет элементы процесса слияния данных, которые задаются на основе входных (необработанных данных) и выходных данных (некоторого решения) и способствует установлению связей между данными и задачами. Подробное описание модели *DFD* дано в работах [21, 22].

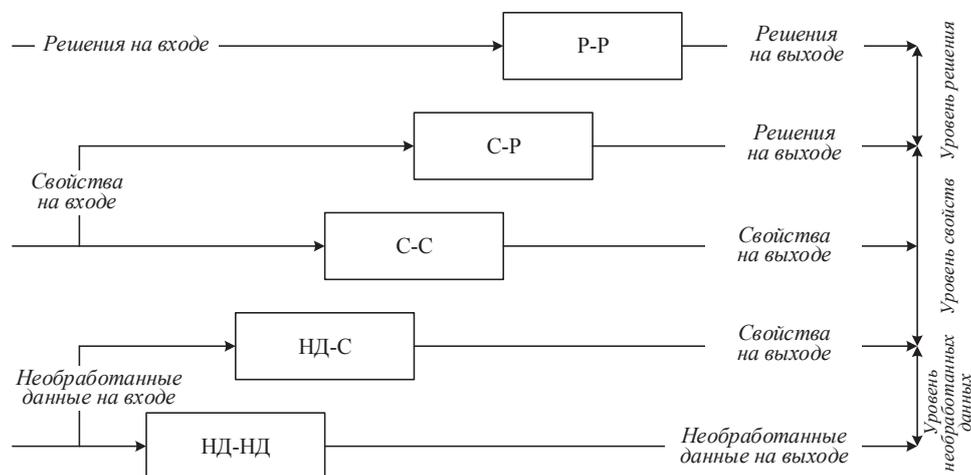


Рисунок 8 – Структура модели DFD

Модель *DFD* имеет пять уровней.

- Необработанные данные на входе - Необработанные данные на выходе (НД-НД) (*Data In-Data Out (DAI-DAO)*). Слияние осуществляется сразу после того, как данные получены с датчиков. Результатом слияния необработанных данных являются такие же данные, только более точные или надёжные.
- Необработанные данные на входе - Свойства на выходе (НД-С) (*Data In-Feature Out (DAI-FEO)*). На данном уровне для слияния используются необработанные данные, полученные от датчиков, для извлечения свойств, описывающих ТО либо ситуацию.
- Свойства на входе - Свойства на выходе (С-С) (*Feature In-Feature Out (FEI-FEO)*). На данном уровне происходит работа над набором свойств для получения нового набора свойств с целью их улучшения/уточнения или извлечения новых.
- Свойства на входе - Решения на выходе (С-Р) (*Feature In-Decision Out (FEI-DEO)*). Для получения решений используются наборы свойств ТО или ситуации.
- Решения на входе - Решения на выходе (Р-Р) (*Decision In-Decision Out (DEI-DEO)*). Данный уровень объединяет входные решения для получения новых решений.

Модель *DFD* полезна для разработки алгоритмов слияния в беспроводной сенсорной сети, однако не предоставляет системного представления, как модель *JDL*. Вместо этого модель *DFD* предоставляет детальный подход к определению задач слияния с помощью ожидаемых входных и выходных данных.

3.1.3 Водопадная модель (Waterfall model)

Модель, предложенная С. Харрисом (*C.J. Harris*) [23], представляет собой иерархическую структуру, где информация, произведённая одним модулем, используется в следующем модуле. Последний модуль (принятие решений) передаёт достаточное количество информации в модуль контроля для калибровки и настройки датчиков.

На рисунке 9 представлена структура водопадной модели, имеющая 3 уровня, каждый из которых включает два модуля и замкнутый цикл для работы в системе [24, 25].

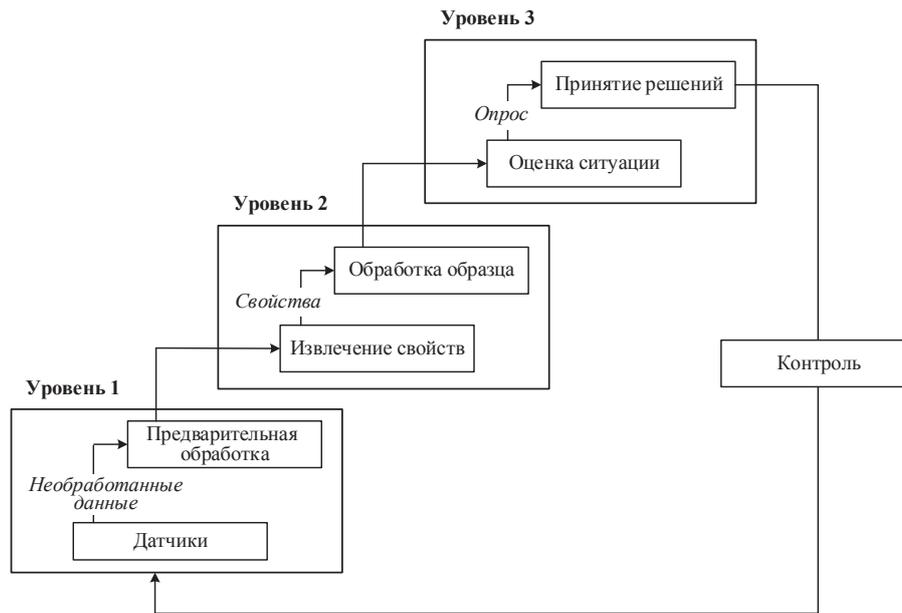


Рисунок 9 – Структура водопадной модели

- *Уровень 1.* Данные из окружающей среды собираются и обрабатываются. На следующий уровень поступают не только обработанные данные, но и информация о датчиках.
- *Уровень 2.* Извлекаются основные свойства ТО, а затем сливаются, тем самым снижая количество передаваемых данных и увеличивая их информационную насыщенность.
- *Уровень 3.* Разрабатываются различные сценарии ситуаций, а также действия для конкретных ситуаций, действий в понятной форме, осуществляются оценки ситуаций. Принимается решение на основе информации, собранной с предыдущих уровней, составляются возможные пути действий для определённой ситуации.

Водопадная модель является более точной в анализе процесса слияния данных, чем другие модели. Однако она имеет ряд недостатков: отсутствие описания обратной связи потока данных; указаний на то, должны ли датчики быть параллельными или последовательными, хотя обработка информации – последовательная [24, 26].

В работах [4, 27] предложен механизм обратной связи с тем, чтобы информация, поступающая к датчикам из модуля принятия решений, постоянно обновлялась.

3.1.4 Модель Томопулоса (Thomopoulos Model)

Данная модель была предложена С. Томопулосом (*S.C. Thomopoulos*, 1989) [28] и представляет трёхуровневую модель, сформированную на основе уровня сигнала (*signal level*), уровня доказательств (*evidence level*) и уровня динамики (*dynamics level*) (рисунок 10).

На каждом уровне собранные данные сталкиваются с ранее обработанными и сохранёнными данными, сохраняя при этом заданный порядок, что означает необходимость решения проблем с задержкой или ошибками в передаче данных. В зависимости от применения эти уровни слияния могут быть реализованы последовательно или поочередно [4, 25].

После того, как датчики проследили за ТО, они выводят свои измерения. Затем

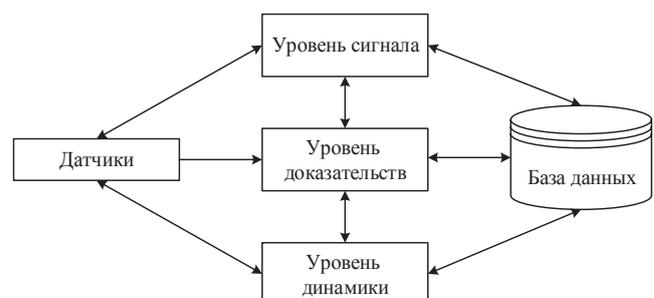


Рисунок 10 – Структура модели Thomopoulos

уровень сигнала обрабатывает эти данные, выполняя корреляции из-за отсутствия математической модели. Поэтому в процессе обучения собранные данные соотносятся с информацией, ранее сохранённой в базе данных. На уровне доказательств данные объединяются на различных уровнях вывода на основе статистической модели и оценки, требуемой пользователем (например, принятие решений или тестирование гипотез). Недостатком данной модели является отсутствие математической модели, описывающей процесс сбора данных.

3.1.5 Модель Ло и Кей (Luo and Kay)

Данная модель представлена Ло и Кей (Luo & Kay, 1998) [29] как общая структура слияния данных, основанная на иерархической модели, но отличающаяся от водопадной модели (рисунок 11). В данной модели данные от датчиков попадают на различные центры слияния, тем самым повышая уровень представления от необработанных данных к решениям.

Из рисунка 11 видно, что в системе присутствуют три функции, которые обычно используются в процессе слияния: выбор датчика, модель ситуации, преобразование данных. Линии от системы к каждому узлу слияния представляют собой любой из возможных сигналов [4, 24, 27]. Данная модель предполагает параллельный ввод и обработку данных, которые могут входить в систему на разных уровнях слияния. Модель основана на децентрализованной архитектуре и не предполагает контроля обратной связи [24].

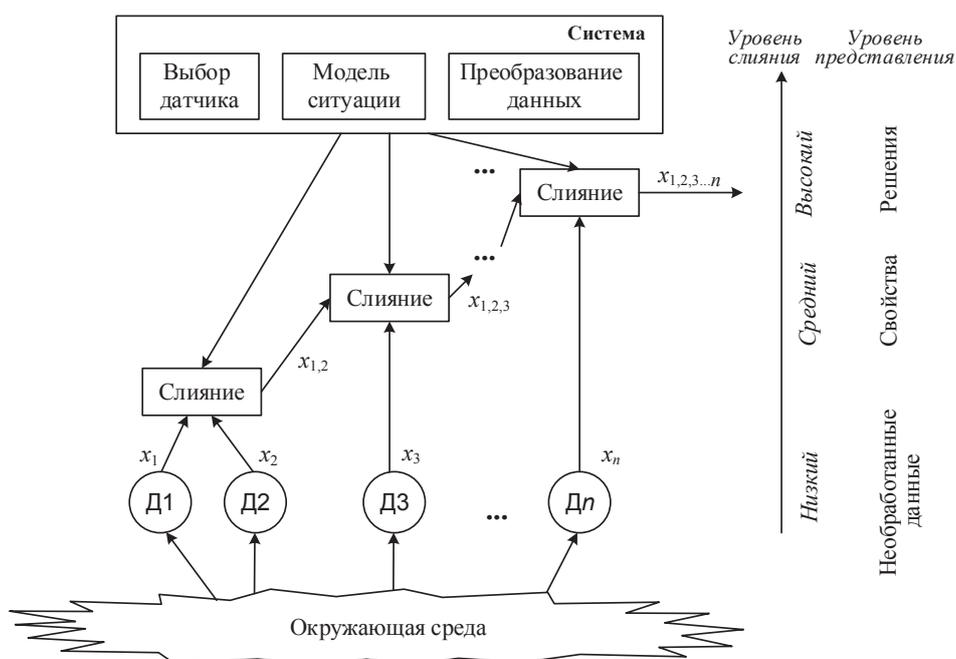


Рисунок 11 – Структура модели Luo и Kay

3.2 Модели на основе действий (activity-based model)

Данные модели определяются на основе определённой последовательности действий, которые указываются и должны выполняться интеллектуальной системой слияния данных.

3.2.1 Цикл OODA

Цикл OODA (O – observe, O – orient, D – decide, A – act), предложенный Дж. Бойдом (J.R. Boyd) [30], рассматривается в качестве единой типовой модели цикла принятия решений для систем командования и управления.

На рисунке 12 представлен цикл *OODA*, предполагающий многократное повторение четырёх последовательных взаимосвязанных действий [17, 20, 21].

- Наблюдать (*observe*). Осуществляется сбор данных из доступных различных датчиков.
- Ориентироваться (*orientate*). Осуществляется слияние собранных данных с тем, чтобы получить интерпретацию текущей ситуации.
- Решать (*decide*). Определяется план действий в ответ на понимание ситуации.
- Действовать (*act*). Выполняется план.

Данный цикл имеет обратную связь, позволяет проводить полный обзор и разделять системные задачи. Однако из-за своей структуры не обеспечивает надлежащее представление слияния данных [20], а также не показывает результат этапа «действовать» на других этапах.

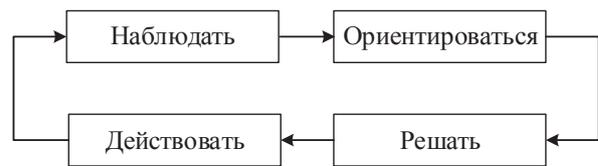


Рисунок 12 – Цикл OODA

3.2.2 Цикл разведки (*intelligence cycle*)

Цикл разведки описывает процесс от сбора необработанных данных до выдачи готовых обработанных данных, которые могут использоваться при принятии решений [31]. На рисунке 13 представлены четыре этапа цикла разведки.

- Сбор (*collection*). Собираются необработанные данные от датчиков.
- Сопоставление (*collation*). Собранные данные анализируются, сравниваются между собой, проводится корреляция. Неважные и ненадёжные данные отбрасываются.
- Оценка (*evaluation*). Собранные данные обрабатываются, а затем анализируются.
- Распространение (*dissemination*). Результаты слияния предоставляются пользователям для выработки решений и действий в ответ на выявленные ситуации.

В отличие от цикла *OODA*, цикл разведки не делает явными этапы решения и исполнения, которые включены в этапы оценки и распространения. Цикл разведки является общим и может быть использован в любой области применения. Однако он не выполняет конкретные задачи, касающиеся слияния данных.

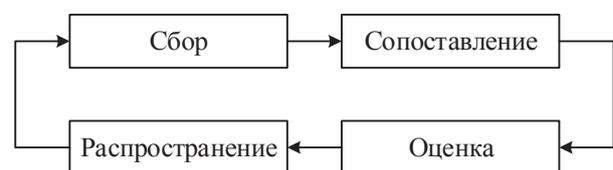


Рисунок 13 – Этапы цикла разведки

3.2.3 Комплексная модель (*Omnibus Model*)

Комплексная модель, предложенная М. Бедворсом (*M.D. Bedworth*) и Дж. Брайном (*J.C. Brien*) [32], организует этапы слияния данных точно так же как цикл разведки и *OODA*. Подробное описание модели представлено в работах [17, 21, 25, 33]. На рисунке 14 представлена структура комплексной модели, состоящая из четырёх этапов.

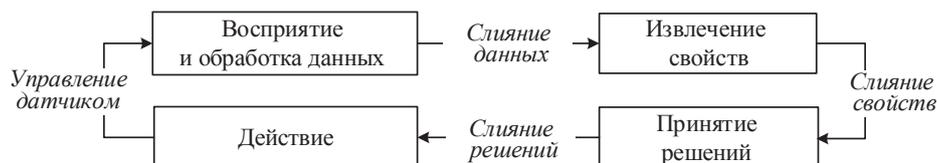


Рисунок 14 – Структура комплексной модели

- Восприятие и обработка данных, поступающих от различных датчиков.
- Извлечение свойств из собранных данных, которые затем сливаются.

- Принятие решений о наиболее подходящем плане действий.
- Действие по выполнению выбранного плана.

Комплексная модель является циклической по структуре, при этом может быть использована несколько раз для одного и того же приложения. Однако она не поддерживает декомпозицию задач на модули, поэтому её необходимо реализовывать и использовать отдельно для различных применений.

3.3 Модели на основе ролей (role-based model)

Данные модели определяются на основе ролей слияния и взаимосвязей между ними, тем самым обеспечивая более детальную модель для интеллектуальных систем слияния данных и системный подход к слиянию данных, однако при этом не указываются задачи слияния. Вместо этого для слияния данных они предоставляют набор функций и определяют отношения между ними [21].

3.3.1 Объектно-ориентированная модель (object-oriented model)

Для слияния данных М. Кокар (M. Kokar) предложил воспользоваться объектно-ориентированной моделью, состоящей из четырёх ролей [21, 34] (рисунок 15).

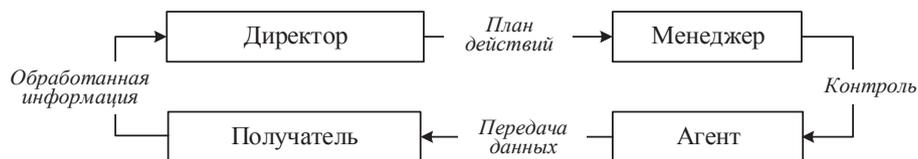


Рисунок 15 – Структура объектно-ориентированной модели

- Агент (*actor*) взаимодействует с окружающей средой, собирает данные.
- Получатель (*perceiver*) проводит анализ, оценивает полученные данные после сбора данных, а затем передаёт директору.
- Директор (*director*) строит план действий с указанием целей системы на основе анализа, предоставленного получателем.
- Менеджер (*manager*) контролирует агентов для выполнения планов, разработанных директором.

Данная модель описывает различные роли системы, однако при этом не рассматривает системные задачи.

3.3.2 Архитектура Франкеля-Бедворта (Frankel-Bedworth)

Б. Франкель предложил архитектуру, состоящую из двух саморегулирующихся процессов, у которых разные цели и роли [17, 35] (рисунок 16):

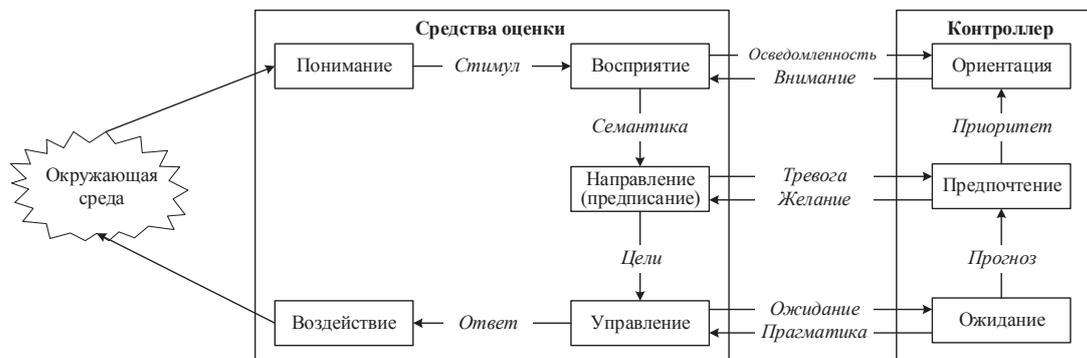


Рисунок 16 – Архитектура Frankel-Bedworth

- локальный процесс пытается достичь целей, предусмотренных глобальным процессом, и обладает ролью средства оценки, которая подобна предыдущим моделям слияния данных;
- глобальный процесс обновляет цели в соответствии с обратной связью, предоставленной локальным процессом.

3.3.3 Модель ситуационной осведомлённости (Situation Awareness Endsley)

Модель Эндсли (*M.R. Endsley*) [36] является наиболее известной моделью оценки обстановки. Она разрабатывалась для анализа деятельности операторов сложных динамических систем (самолетов, АЭС, заводов и т.п.), включая сравнительную оценку эффективности их работы, поиск причин ошибок и выработку рекомендаций по обучению для автоматизации процесса управления [20, 37].

Ситуационная осведомлённость означает возможность получения достаточно полной и точной информации о ситуациях в реальном режиме времени, необходимой для принятия решения и прогнозирования состояний. Согласно модели Эндсли состояние ситуационной осведомлённости является результатом процесса анализа и оценки ситуации и включает три уровня [38] (рисунок 17).

- Уровень 1 – восприятие данных и обстановки. Быстро и эффективно осуществляется сбор необходимых данных с целью получения знаний того, что происходит вокруг;
- Уровень 2 – понимание значения и важности ситуации. Осуществляется процесс слияния данных от разнородных источников с целью понимания значения собственных действий и действий других участников ситуации.
- Уровень 3 – прогнозирование будущих состояний и событий. Представление собранных и проанализированных данных в необходимой форме с целью представления сценария развития ситуации.

Модель Эндсли может применяться в различных областях, которые характеризуются высокой динамикой, большими объёмами и структурной сложностью обрабатываемых данных, необходимостью одновременного решения нескольких задач в условиях ограниченных субъективных возможностей (например, рабочей памяти), высокой степенью риска, т.е. где неудачное решение может привести к тяжелым последствиям.



Рисунок 17 – Структура модели ситуационной осведомлённости

Главные требования, которые предъявляются к системам ситуационного осведомления о состоянии сети, это: масштабируемость и возможность интеграции с существующей инфраструктурой безопасности; возможность централизованного сбора, корреляции, анализа и отображения разнородных данных от всех компонентов инфраструктуры; возможность взаимодействия специалистов различных служб; предоставление всесторонних комплексных отчётов в режиме реального времени, уведомления об угрозах, высокий уровень визуализации; возможность получения данных из внешних источников [38].

Отметим, что достаточно сложно определить универсальную модель слияния данных для конкретного использования датчика. Проектирование и разработка интеллектуальных и сенсорных систем и сбор данных с нескольких разнотипных датчиков является сложной зада-

чей. По этой причине для эффективного сбора данных и их обработки можно использовать несколько моделей слияния или их комбинации. Типичным примером является Комплексная модель, которая представляет собой сочетание трёх моделей слияния данных: цикл *OODA*, водопадная модель и *DFD*.

Проведённый анализ показал достаточно большое количество моделей слияния данных, которые могут найти широкое применение при диагностировании ТО любых отраслей промышленности и транспорта. Однако при выборе той или иной модели необходимо учитывать ряд аспектов, например: какие собраны будут данные и на каком этапе; какая преследуется цель использования той или иной модели с точки зрения требований и будущего использования; на каком этапе будут приниматься решения.

В связи с этим предлагаются следующие рекомендации: например, *JDL* модель и водопадную модель использовать для работы на высоком уровне, однако не исключено их использование и на низком уровне; модель *OODA* целесообразно применять для деятельности отдельных лиц и организаций в условиях конкурентной среды (соперничества); модель ситуационной осведомлённости использовать в случае, когда специалист, обеспечивающий эффективное управление безопасностью организации, глубоко владеет текущей обстановкой (ситуационной осведомленностью).

Заключение

Слияние данных играет важную роль при диагностировании ТО, поскольку это увеличивает вероятность обнаружения неисправностей ТО в условиях множества разнотипных датчиков, а также уменьшает время реакции экспертов на различные ситуации.

Исследованы различные структурные модели слияния данных и их принципы работы. Проведено сравнение рассмотренных моделей между собой, выявлены их преимущества и недостатки. Отмечено, что для эффективного сбора исходных данных, поступающих от множества разнотипных датчиков, и их обработки можно использовать несколько моделей слияния данных либо их комбинации.

Благодарности

Работа выполнена при поддержке РФФИ, проекты № 19-07-00263, 19-07-00195, 19-08-00152.

Список источников

- [1] **Bahador Khaleghi** Multisensor data fusion: A review of the state-of-the-art / Bahador Khaleghi, Alaa Khamis, Fakhreddine O. Karray, Saiedeh N.R. // Information Fusion. 2013. - Vol.41. - No.1. - P. 28-44.
- [2] **Fouad, M.** Data mining and fusion techniques for WSNs as a source of the big data / M.M. Fouad, N.E. Oweis, T. Gaber, M. Ahmed, V. Snasel // Procedia Computer Science. - Elsevier, 2015. - Vol.65. - P. 778-786.
- [3] **Valet, L.A.** A statistical overview of Recent Literature in Information Fusion / L.A. Valet, G. Mauris, P. Bolon // 3rd International Conference on Information Fusion. France. 2001. - P.532-536.
- [4] **Esteban, J.** A review of data fusion models and architectures: Towards engineering guidelines / J. Esteban, A. Willetts, R. Hannah, P. Bryanston-Cross // Neural Computing & Applications. 2005. - Vol.14. - No.4. - P.273-281.
- [5] **Bleiholder, J.** Data Fusion and Conflict Resolution in Integrated Information Systems / J. Bleiholder. – Potsdam: Hasso-Plattner-Institut, 2010. - 184 p.
- [6] **Goodman, I.R.** Mathematics of Data Fusion / I.R. Goodman, R.P. Mahler, H.T. Nguyen // Kluwer Academic Publishers, 1997. - 508 p.
- [7] **Вовченко, А.Е.** Методы разрешения сущностей и слияния данных в ETL-процессе и их реализация в среде Nadoop / А.Е. Вовченко, Л.А. Калиниченко, Д.Ю. Ковалев // Информатика и ее применение. 2014. - Т.8. - Вып. 4. - С.94-109.

- [8] **Долгий, А.И.** Проблемы и методы слияния разнородных данных в гибридных интеллектуальных системах / А.И. Долгий, А.Е. Колоденкова, С.М. Ковалев // Гибридные и синергетические интеллектуальные системы: материалы IV Всероссийской Поспеловской конференции с международным участием. – Калининград: БФУ им. И. Канта, 2018. - С.181-187.
- [9] **Городецкий, В.И.** Многоагентная технология принятия решений в задачах объединения данных / В.И. Городецкий, О.В. Карсаев, В.В. Самойлов // Труды СПИИРАН. – СПб.: СПИИРАН. 2002. - Вып.1. - Т.2. - С.12-37.
- [10] **Kushwaha, M.** Feature-Level Information Fusion Methods for Urban Surveillance Using Heterogeneous Sensor Networks / M. Kushwaha // Vanderbilt University, Nashville, TN, USA, 2010. - 185 p.
- [11] **Castanedo, F.** A review of data fusion techniques / F. Castanedo // The Scientific World Journal, 2013. - 19 p.
- [12] **Afnan, Alofi** A Review of Data Fusion Techniques / Afnan Alofi, Anwaar Alghamdi, Razan Alahmadi, Najla Aljuaid, M. Hemalatha // International Journal of Computer Applications. 2017. - Vol.167. - No.7. - P.37-41.
- [13] **Afef Ben, Brahim** Solving data fusion problems using ensemble approaches. - Tunisia: University of Tunisia, 2010. - 104 p.
- [14] **Elmenreich, A.W.** Sensor fusion in time-triggered systems / A.W. Elmenreich // Technische Universität Wien, Wien 2002. - 173 p.
- [15] **Xinhan, H.** Multi-sensor data fusion structures in autonomous systems: A Review / H. Xinhan, M. Wang // International symposium on intelligent control, Houston, Texas, 2003. - P.817-821.
- [16] **Hall, D.L.** An Introduction to Multisensor Data Fusion / D.L. Hall, J. Llinas // Proceedings of the IEEE. 1998. - Vol.85. - No.1. - P.6-23.
- [17] **Nakamura, E.F.** Information fusion for wireless sensor networks: methods, models, and classifications / E.F. Nakamura, A.A. Loureiro, A.C. Frery // ACM Comput. Surv. 2007. - Vol.39. - No.3. - 55 p.
- [18] **White, F.E.** A Model for Data Fusion / F.E. White // 1st National Symposium on Sensor Fusion: process. 1988. - Vol.2. - P.20-26.
- [19] **Llinas, J.** Handbook of Multisensor Data Fusion: Theory and Practice / J. Llinas, D.L. Hall, M.E. Liggins. – London: CRC Press, 2009. - 872 p.
- [20] **Foo, P.H.** High-level Information Fusion: An Overview / P.H. Foo, G.W. Ng // J. Adv. Inf. Fusion. 2013. - Vol.8. - No.1. - P.33-72.
- [21] **Abdelgawad, A.** Resource-Aware Data Fusion Algorithms for Wireless Sensor Networks / A. Abdelgawad, M. Bayoumi // Lecture Notes in Electrical Engineering. 2012. - Vol.118. - 119 p.
- [22] **Dasarathy, B.V.** Sensor fusion potential exploitation-innovative architectures and illustrative applications / B.V. Dasarathy // Proceedings of the IEEE. 1997. - Vol.85. - No.1. - P.24-38.
- [23] **Harris, C.J.** Multi-sensor Data Fusion in Defence and Aerospace / C.J. Harris, A. Bailey, T.J. Dodd // Aeronaut J. 1998. - Vol.102. - No.1015. - P.229-244.
- [24] **Veloso, M.** Multi-Sensor Data Fusion on Intelligent Transport Systems / M. Veloso, C. Bentes, F. Camara Pereira // MIT Portugal Transportation Systems Working Paper Series, 2009. - 18 p.
- [25] **Rawa, Adla.** Multi Sensor Data Fusion, Methods and Problems / Rawa Adla, Youssef Bazzi, Nizar Al-Holou // Proceedings of the 2013 International conference on parallel and distributed processing techniques and applications. - Las Vegas Nevada, USA. 2013. - Vol.1. - P.106-111.
- [26] **Bedworth, M.** The Omnibus Model: A New Model of Data Fusion / M. Bedworth, J. O'Brien // IEEE Aerospace and Electronic Systems Magazine. - 2000. - Vol. 15. - P. 30-36.
- [27] **Zegras, C.** Data Fusion for Travel Demand Management: State of the Practice & Prospects / C. Zegras, F. Pereira, A. Amey, M. Veloso, L. Liu, C. Bento, A. Biderman // 4th International Symposium on Travel Demand Management, 2008. - 98 p.
- [28] **Thomopoulos, S.C.** Sensor Integration and Data Fusion / S.C. Thomopoulos // Proceedings of SPIE 1198, Sensor fusion II: Human and machine strategies. - 1989. - P.178-191.
- [29] **Luo, R.C.** Multisensor Integration and Fusion: Issues and Approaches / R.C. Luo, M.G. Kay // Proceedings of 1988 Orlando Technical Symposium. 1989. - P.42-49.
- [30] **Boyd, J.R.** A discourse on winning and losing. Unpublished set of briefing slides available / J.R. Boyd // Air University Library: Maxwell AFB, Alabama. 1987. – <http://dnipogo.org/john-r-boyd/>.
- [31] **Shulsky, A.N.** Silent warfare: understanding the world of intelligence / A.N. Shulsky, G.J. Schmitt // Inc., New York, NY, 2002. - 248 p.
- [32] **Bedworth, M.D.** The omnibus model: A new model for data fusion / M.D. Bedworth, J.C. O'Brien // Proceedings of the 2nd International Conference on Information Fusion. ISIF, Sunnyvale. 1999. - P.437-444.
- [33] **Elmenreich, W.** An Introduction to Sensor Fusion / W. Elmenreich // Vienna University of Technology, Austria, 2002. - Vol.502. - 29 p.
- [34] **Kokar, M.M.** Reference Model for data fusion Systems / M.M. Kokar, M.D. Bedworth, C.B. Frankel // International Society for Optics and Photonics: In AeroSense. 2000. - P.191-202.

- [35] **Frankel, B.** Control, estimation and abstraction in fusion architectures: Lessons from human data processing / B. Frankel, M.D. Bedworth // Proceeding of the 3rd International Conference on Data Fusion, Paris, France. 2000. - Vol.1. - P.3-10.
- [36] **Endsley, M.R.** Toward a theory of situation awareness in dynamic systems / M.R. Endsley // Human Factors: The J. of the Human Factors and Ergonomics Society. 1995. - Vol.37. - No.1. - P.32-64.
- [37] **Лебедев, С.В.** Концептуальная модель подсистемы оценки обстановки интеллектуального агента реального времени / С.В. Лебедев, М.Г. Пантелеев // Известия СПбГЭТУ «ЛЭТИ». 2015. - №7. - С.41-46.
- [38] **Федулов, М.** Ситуационная осведомленность. Новый подход к проблемам комплексной безопасности / М. Федулов // Безопасность: Информационное обозрение. - 09-08-2011. - <http://www.securityinfowatch.ru/view.php?section=articles&item=4>.
-

INTELLECTUAL TECHNOLOGIES OF DATA FUSION FOR DIAGNOSTICS OF TECHNICAL OBJECTS

S.M. Kovalev¹, A.E. Kolodenkova², V. Snasel³

¹Rostov State Transport University, Rostov-on-Don, Russia
ksm@rfniias.ru

²Samara State Technical University, Samara, Russia
anna82_42@mail.ru

³VSB-Technical University of Ostrava, Ostrava, Czech Republic
vaclav.snasel@vsb.cz

Abstract

Fusion of heterogeneous data in real time is an important task in the diagnosis of technical objects. This is due to the need of taking into account not only the data coming from the sensors, but also external factors affecting the technical object. The article addresses problems of data fusion terminology on the basis of a review of the national and foreign literature. New definition of the term "data fusion" is proposed. Different scientific views of domestic and foreign experts on a problem of data fusion for diagnosing of technical objects in presence of different types of sensors and heterogeneous information are generalized and systematized. The adapted classification of data fusion, taking into account various criteria (the relations between sensors, the level of abstraction of data, architecture type), is presented. Classification of structural models of data fusion, developed for creation of intellectual systems of data fusion is given. Models of process of data fusion are investigated, their structures are presented, model merits and demerits are revealed. It is noted that for effective collecting of the basic data, arriving from of different types of sensors, and its processing it is possible to use several models of data fusion or their combination. It will allow making scientifically based management decisions when diagnosing difficult technical objects. While a lot of works of foreign researchers are devoted to separate sections of data fusion technologies, the paper presents the first full research in Russian where all the main aspects that belong to the intellectual technologies of data fusion are considered.

Key words: *heterogeneous data, models of data fusion, diagnostics of technical objects, many different types of sensors.*

Citation: *Kovalev SM, Kolodenkova AE, Snasel V. Intellectual technologies of data fusion for diagnostics of technical objects [In Russian]. *Ontology of designing*. 2019; 9(1): 152-168. – DOI: 10.18287/2223-9537-2019-9-1-152-168.*

Acknowledgment

The work was supported by RFBR grants No. 19-07-00263, 19-07-00195, 19-08-00152.

References

- [1] **Bahador Khaleghi, Alaa Khamis, Fakhreddine OK, Saiedeh NR.** Multisensor data fusion: A review of the state-of-the-art. *Information Fusion*. 2013; 41(1): 28-44.

- [2] **Fouad M, Oweis NE, Gaber T, Ahmed M, Snasel V.** Data mining and fusion techniques for WSNs as a source of the big data. *Procedia Computer Science*: Elsevier. 2015; 65: 778-786.
- [3] **Valet LA, Mauris G, Bolon P.** A statistical overview of Recent Literature in Information Fusion. 3rd International Conference on Information Fusion; 2001: P. 532-536.
- [4] **Esteban J, Willetts A, Hannah R, Bryanston-Cross P.** A review of data fusion models and architectures: Towards engineering guidelines. *Neural Computing & Applications*. 2005; 14(4): 273-281.
- [5] **Bleiholder J.** Data Fusion and Conflict Resolution in Integrated Information Systems. – Potsdam: Hasso-Plattner-Institut; 2010.
- [6] **Goodman IR, Mahler RP, Nguyen HT.** Mathematics of Data Fusion. Kluwer Academic Publishers; 1997.
- [7] **Vovchenko AE, Kalinichenko LA, Kovalev DU.** Methods of entity resolution and data fusion in the ETL-process and their implementation in the Hadoop environment [In Russian]. *Informatics and her application*. 2014; 4: 94-109.
- [8] **Dolgy AI, Kolodenkova AE, Kovalev SM.** Problems and methods of heterogeneous data fusion in hybrid intelligent systems [In Russian]. Hybrid and synergetic intelligent systems: Proceedings of the IV Russian Pospelov conference with international participation. – Kaliningrad: IK BFU; 2018: 181-187.
- [9] **Gorodetski VI, Karsayev OV, Samoilo VV.** Multi-agent data fusion technology [In Russian]. SPIIRAS Proceedings. – SPb: SPIIRAS 2002; 1: 12-37.
- [10] **Kushwaha M.** Feature-Level Information Fusion Methods for Urban Surveillance Using Heterogeneous Sensor Networks. Vanderbilt University, Nashville, TN, USA; 2010.
- [11] **Castanedo F.** A review of data fusion techniques. *The Scientific World Journal*; 2013.
- [12] **Afnan Alofi, Anwaar Alghamdi, Razan Alahmadi, Najla Aljuaid, Hemalatha M.** A Review of Data Fusion Techniques. *International Journal of Computer Applications*; 2017: 167 (7): 37-41.
- [13] **Afef Ben B.** Solving data fusion problems using ensemble approaches. Tunisia: University of Tunisia; 2010: 104.
- [14] **Elmenreich AW.** Sensor fusion in time-triggered systems. Technische Universität Wien, Wien; 2002: 173.
- [15] **Xinhan H, Wang M.** Multi-sensor data fusion structures in autonomous systems. International symposium on intelligent control. Houston, Texas; 2003: 817-821.
- [16] **Hall DL, Llinas J.** An Introduction to Multisensor Data Fusion. Proceedings of the IEEE. 1998; 85(1): 6-23.
- [17] **Nakamura EF, Loureiro AA, Freyre AC.** Information fusion for wireless sensor networks: methods, models, and classifications. *ACM Comput. Surv.* 2007; 39(3): 1-55.
- [18] **White FE.** A Model for Data Fusion. 1st National Symposium on Sensor Fusion: process 1988; 2: 20-26.
- [19] **Llinas J, Hall DL, Liggins ME.** Handbook of Multisensor Data Fusion: Theory and Practice. – London: CRC Press; 2009.
- [20] **Foo PH, Ng GW.** High-level Information Fusion: An Overview. *Inf. Fusion* 2013; 8 (1): 33-72.
- [21] **Abdelgawad A, Bayoumi M.** Resource-Aware Data Fusion Algorithms for Wireless Sensor Networks. *Lecture Notes in Electrical Engineering* 2012; 118: 1-119.
- [22] **Dasarathy BV.** Sensor fusion potential exploitation-innovative architectures and illustrative applications. Proceedings of the IEEE 1997; 85 (1): 24-38.
- [23] **Harris CJ, Bailey A, Dodd TJ.** Multi-sensor Data Fusion in Defence and Aerospace. *Aeronaut J.* 1998; 102(1015): 229-244.
- [24] **Veloso M, Bento C, Camara Pereira F.** Multi-Sensor Data Fusion on Intelligent Transport Systems. MIT Portugal Transportation Systems Working Paper Series; 2009.
- [25] **Rawa Adla, Youssef Bazzi, Nizar Al-Holou.** Multi Sensor Data Fusion, Methods and Problems. Proceedings of the 2013 International conference on parallel and distributed processing techniques and applications. Las Vegas Nevada, USA 2013; 1.1:106-111.
- [26] **Bedworth M, O'Brien J.** The Omnibus Model: A New Model of Data Fusion. *IEEE Aerospace and Electronic Systems Magazine* 2000; 15: 30-36.
- [27] **Zegras C, Pereira F, Amey A, Veloso M, Liu L, Bento C, Biderman A.** Data Fusion for Travel Demand Management: State of the Practice & Prospects. 4th International Symposium on Travel Demand Management; 2008.
- [28] **Thomopoulos SC.** Sensor Integration and Data Fusion. Proceedings of SPIE 1198, Sensor fusion II: Human and machine strategies; 1989: 178-191.
- [29] **Luo RC, Kay MG.** Multisensor Integration and Fusion: Issues and Approaches. Proceedings of 1988 Orlando Technical Symposium; 1989: 42-49.
- [30] **Boyd JR.** A discourse on winning and losing. Unpublished set of briefing slides available. Air University Library: Maxwell AFB, Alabama 1987. - <http://dnipogo.org/john-r-boyd/>.
- [31] **Shulsky AN, Schmitt GJ.** Silent warfare: understanding the world of intelligence. Inc., New York, NY; 2002.
- [32] **Bedworth MD, O'Brien JC.** The omnibus model: A new model for data fusion. Proceedings of the 2nd International Conference on Information Fusion. ISIF, Sunnyvale; 1999: 437-444.
- [33] **Wilfried E.** An Introduction to Sensor Fusion. Vienna University of Technology, Austria 2002; 502: 29.

- [34] **Kokar MM, Bedworth MD, Frankel CB.** Reference Model for data fusion systems. International Society for Optics and Photonics: In AeroSense; 2000: 191-202.
- [35] **Frankel B, Bedworth MD.** Control, estimation and abstraction in fusion architectures: Lessons from human data processing. Proceeding of the 3rd International Conference on Data Fusion, Paris, France 2000; 1: 3-10.
- [36] **Endsley MR.** Toward a theory of situation awareness in dynamic systems. Human Factors: The J. of the Human Factors and Ergonomics Society. 1995; 37(1): 32-64.
- [37] **Lebedev SV, Pantelev MG.** A conceptual model of the subsystem of assessment of the situation an intelligent agent real-time [In Russian]. Proceedings of Saint Petersburg Electrotechnical University «LETI». 2015; 7: 41-46.
- [38] **Fedulov M.** Situational awareness. A new approach to integrated security [In Russian]. Security: informational review. 09-08-2011. - <http://www.securityinfowatch.ru/view.php?section=articles&item=4>.

Сведения об авторах



Ковалев Сергей Михайлович, 1954 г. рождения. Окончил Таганрогский радиотехнический институт в 1976 г., д.т.н. (2002). Профессор кафедры «Автоматика и телемеханика на железнодорожном транспорте» Ростовского государственного университета путей сообщения. Член Российской ассоциации искусственного интеллекта. В списке научных трудов более 200 работ в области интеллектуальных и нечётких систем, нечётко-темпоральных баз знаний, мягких вычислений и экспертной поддержки принятия решений.

Sergey Mikhailovich Kovalev (b. 1954) graduated from the Taganrog radio engineering Institute (Taganrog-city) in 1976, D. Sc. Eng. (2002). He is Professor at Rostov State Transport University (Department of Automation and Remote Control at the Railway on Transport). He is member of Russian Association of Artificial Intelligence. He is co-author of about 200 scientific articles

and abstracts in the field of intelligent and fuzzy systems, fuzzy temporal knowledge bases, soft computing, expert and decision support.



Колоденкова Анна Евгеньевна, 1982 г. рождения. Окончила Уфимский государственный авиационный технический университет в 2004 г., д.т.н. (2017). Заведующая кафедрой «Информационные технологии» Самарского государственного технического университета. Член Российской ассоциации искусственного интеллекта, заместитель председателя научного совета Российской ассоциации нечётких систем и мягких вычислений. В списке научных трудов более 150 работ в области атомной энергетики, программной инженерии, системного анализа, интеллектуальных систем, мягких вычислений и экспертной поддержки принятия решений.

Anna Evgenievna Kolodenkova (b. 1982) graduated from the Ufa State Aviation Technical University (Ufa-city) in 2004, D. Sc. Eng. (2017). She is Associate Professor, Head of «Information

technologies» Department at Samara State Technical University. She is member of Russian Association of Artificial Intelligence and Deputy Chairman of the scientific Council of the Russian Association of fuzzy systems and soft computing. She is co-author of about 150 scientific articles and abstracts in the field of nuclear energy, software engineering, system analysis, intelligent systems, soft computing and expert decision support.



Вацлав Снасель, 1957 г. рождения. Профессор (2006), получил докторскую степень в области алгебры и геометрии в Масариковом университете в Брно Чешской Республики. Ректор Оставского технического университета. Является членом IEEE, ACM, AMS и SIAM. В списке научных трудов более 400 работ, посвященных искусственному интеллекту, управлению знаниями, поиску информации, интеллектуальному анализу данных, которые применяются к различным проблемам реального мира.

Vaclav Snasel (b. 1957). Professor (2006), he received Ph.D. degree in Algebra and Geometry from Masaryk University, Brno, Czech Republic. Rector of VSB-Technical University of Ostrava. He is a member of IEEE, ACM, AMS and SIAM. He is co-author of about 400 scientific articles. He works in a multi-disciplinary environment involving artificial intelligence, knowledge

management, information retrieval, data mining and applied to various real-world problems.