

# СИСТЕМА ПРИНЯТИЯ ОПЕРАТИВНЫХ РЕШЕНИЙ НА ОСНОВЕ РЕДУКЦИИ КОРПОРАТИВНЫХ ДАННЫХ

**А. С. Мороз**

*artsiom.maroz@gmail.com;*

*Научный руководитель — А. Н. Вальвачёв, кандидат технических наук,  
доцент*

В докладе обсуждается проблема усовершенствования методов принятия решений в крупномасштабных компаниях в условиях быстрого роста динамики и неопределенности глобальной среды. Показано, что одна из причин снижения качества управляющих решений, включая запаздывание и высокий уровень субъективизма, заключается в увеличении объема данных, поступающих от удаленных подразделений компании. Предлагается новый подход к принятию решений на основе редукции больших данных, обеспечивающий уменьшение времени синтеза решений при сохранении их качества и снижении стоимости обработки.

**Ключевые слова:** крупномасштабные компании; большие данные; редукция данных, принятие решений; интеллектуальные системы.

## **ВВЕДЕНИЕ**

В XXI веке в процессе глобализации организационно-технические системы претерпели существенные изменения. К важнейшим из них можно отнести появление крупномасштабных компьютеризированных компаний (КМК) с большим количеством подразделений в странах с выгодными для бизнеса условиями [1]. Для управления КМК используются системы поддержки принятия решений (СППР, DSS) различного типа. В СППР преобладают монолитные архитектуры и плановые задачи, использующие постоянные источники данных. Процесс проектирования, разработки и развертывания СППР требует значительных затрат времени и средств [2].

До 2020 г. СППР были достаточно эффективны. После начала пандемии COVID-19 КМК приступили к реструктуризации, пытаясь адаптироваться к быстрым глобальным изменениям и потере конкурентных преимуществ [3]. Поток данных резко увеличился, что вызвало удорожание обработки и запаздывание принятия решений.

В новых условиях стали востребованы новые типы СППР, обладающие гибкой архитектурой, способные решать оперативные задачи на основе гетерогенных источников в минимальное время с минимумом затрат. Ведущие университеты и компании начали исследования в этом направлении. В данной работе предложен вариант решения, основанный на идее редукции больших данных по принципам *small data*. Данный

подход способствует сокращению времени принятия решений без потери их качества и исключению запаздывания.

### ПОСТАНОВКА ЗАДАЧИ

Пусть имеется КМК, включающая центральный офис ( $A$ ) и множество территориально распределенных подразделений ( $B1, B2, \dots Bn$ ). Бизнес-процессы каждого подразделения описываются достаточно большим объемом данных  $vX$ , которые формируются локальными СППР ( $bDSS$ ) и в плановом режиме отправляются к СППР ( $aDSS$ ) главного офиса для синтеза управляющих решений.

Количество подразделений постоянно увеличивается ( $n \rightarrow \infty$ ). Соответственно увеличивается объем данных ( $vX$ ), необходимых для синтеза решений ( $U$ ) и время синтеза ( $tU$ ) (рис. 1).

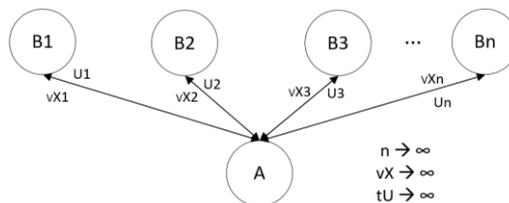


Рис. 1. Традиционная схема принятия решений в КМК

В центре за время  $t$  определяют состояния подразделений  $V = f(X)$  и синтезируют управляющие решения  $U = f(V)$ . При  $vX \rightarrow \infty$  время  $t \rightarrow \infty$ . Увеличение времени обработки приводит к запаздыванию решений и, соответственно, снижению их качества.

Требуется разработать методику и программную систему для синтеза управляющих решений, обеспечивающую быстрое и экономное развертывание, ускорение синтеза управляющих решений при увеличении количества подразделений и объема первичных данных.

### ПОДХОД К РЕШЕНИЮ

За основу решения примем следующие положения:

- используем принцип редукции, т.е. сокращение объема больших данных до уровня, обеспечивающего принятие решений без потери их качества [4];
- редукцию данных будем производить в два этапа: делегируем обработку по принципу small data от СППР главного офиса к локальным СППР [4] и минимизируем объем передаваемого в центр объема данных за счет отправки только критических значений, требующих оперативной реакции главного офиса [2, 4];

- оперативные задачи будем рассматривать как настройку над плановыми задачами [7];
- скорость развертывания увеличим за счет замены монолитного подхода к архитектуре на микросервисный [5];
- стоимость разработки снизим за счет использования адаптированного открытого программного обеспечения [6];
- для практической реализации вышеуказанных положений используем принципы архитектуры верхнего уровня (high level architecture) [7].

## АЛГОРИТМ СИНТЕЗА УПРАВЛЯЮЩИХ РЕШЕНИЙ

Схема решения поставленной задачи включает следующие шаги.

Шаг 1. Анализ структуры компании, формулирование оперативных задач для каждого подразделения.

Шаг 2. Определение плановых показателей деятельности  $\tilde{X}$  для каждого подразделения  $B$ , необходимых для решения оперативных задач.

Шаг 3. Разработка алгоритма  $f_1$  для отображения значений  $\tilde{X}$  в показатели  $X$  в рамках универсальной шкалы  $X = f_1(\tilde{X}) \in [0.00 \dots 1.00]$ .

Шаг 4. Разработка алгоритма синтеза состояния подразделения  $V = f_2(X)$ .

Шаг 5. Разработка алгоритма синтеза управляющего решения  $U = f_3(V)$ .

Шаг 6. Разработка алгоритма визуализации  $U$  в форме, отвечающей требованиям пользователя  $Vis = f_4(U, user)$ .

Шаг 7. Построение предметной области:  $A, B, \tilde{X}, X, V, U, f_1, f_2, f_3, f_4$ .

Шаг 8. Реализация предметной области в виде системы.

Шаг 9. Применение системы для синтеза управляющих решений.

Для разработки алгоритмов  $f_1, f_2, f_3, f_4$  необходимо использовать математические методы, релевантные типу исходных данных.

## АРХИТЕКТУРА НА ОСНОВЕ ELASTIC STACK

Для реализации шагов 1-9 с учетом минимизации стоимости целесообразно использовать открытую экосистему Elastic Stack [7]. В настоящее время в стек входят: Elasticsearch + Logstash + Beats + Kibana.

Elasticsearch – это noSQL СУБД, ядро всей системы, которая сочетает в себе функции базы данных, поисковой и аналитической системы.

Logstash – это конвейер обработки данных на стороне сервера, который получает данные из нескольких источников одновременно, выполняет парсинг логов, а затем отправляет в базу данных Elasticsearch.

Beats – сборщик данных на стороне клиента для передачи их в Logstash или Elasticsearch.

Kibana – интерфейс для доступа к базе данных, анализа данных, построения и визуализации графиков.

Схема построения СППР на основе Elastic Stack показана на рис. 2.

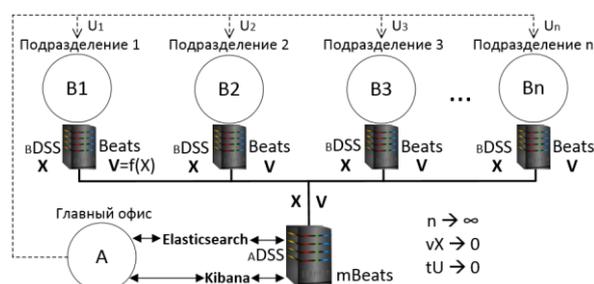


Рис.2. Схема принятия решений в КМК на основе Elastic Stack

При реализации данной схемы можно дополнительно использовать широкий спектр сервисов экосистемы Hadoop [8].

### АПРОБАЦИЯ СИСТЕМЫ

Для демонстрации работы системы и определяя результата редукции рассмотрим станцию, производящую чистую энергию на 10 ветрогенераторах, и стороннюю ремонтную компанию. Интегральный показатель состояния ветрогенератора определяет умным контроллером локального датчика и не выходит из трех диапазонов 60-100, 101-250, 251-290. После запуска станции все генераторы исправны:

0.00, 0.00, 0.00, 0.00, 0.00, 0.00, 0.00, 0.00, 0.00, 0.00

В процессе эксплуатации или после стихийных бедствий ветрогенераторы теряют работоспособность. Соответственно изменятся значения нормализованного вектора:

0.50, 0.00, 1.00, 0.50, 0.00, 0.00, 0.00, 0.00, 0.00, 0.00

Система отправляет соответствующие рекомендации ремонтной компании для исполнения и руководству станции для сведения и контроля выполнения рекомендаций. Результат представлен на рис.3.

```

*****
Традиционный метод:
Диапазон значений параметров:
[60, 100, 101, 250, 251, 290]
Входной вектор данных:
[153, 65, 270, 120, 35, 71, 84, 54, 130, 111]
Объем памяти для отправки: 76 байт
*****
Новый метод:
Нормализованный входной вектор:
[0.5, 0.0, 1.0, 0.5, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0]
Рекомендации:
- заменить ветрогенератор 2.
- выполнить ремонт ветрогенераторов: 0, 3.
Объем памяти для отправки: 16 байт
Экономия в 4.75 раз для 10 показателей.
*****

```

Рис.3. Результат оценки и синтеза рекомендаций

В аналогичных системах настройку параметров осуществляет эксперт, что обеспечивает учет местных условий и максимальную достоверность результатов. В среднем объем передаваемых данных сокращается в 4-5 раз, что существенно разгружает каналы и уменьшает расходы на хранение данных в Elasticsearch.

## **ЗАКЛЮЧЕНИЕ**

В процессе исследования возможностей применения принципа редукции больших данных для принятия оперативных решений в КМК показана возможность значительного сокращения объема передаваемых по каналам связи данных и времени принятия оперативных решений. Для снижения стоимости разработки и развертывания СППР наиболее эффективной оказалась сборка Elastic Stack. Применение СППР на основе этой сборки позволяет в энергетических и других компаниях, использующих большое количество объектов наблюдения, оперативно организовать поддержку жизненного цикла большого количества распределенных объектов (например, солнечных панелей, ветрогенераторов) при минимальной загрузке каналов.

## **Библиографические ссылки**

1. Управление развитием крупномасштабных систем (MLSD'2022): Труды XV-й междунар. конф, 26-28 сент. 2022 г., Москва – М.: ИПУ РАН, 2022. 1500 с.
2. Виссия Х., Краснопрошин В.В., Вальвачев А.Н. Принятие решений в информационном обществе. СПб: ЛАНЬ, 2019. 227 с.
3. Global economic effects of covid-19. Congressional research service. 2021. 110 p.
4. Hekler E. et al. Why we need a small data paradigm // BMC Medicine. 2019. Vol.17. P.1-9.
5. Newman S. Monolith to microservices: evolutionary patterns to transform your monolith - O'Reilly Media, 2019. 270 p.
6. Paro A. Elasticsearch 7.0 – Packt Publishing, 2019. 726 p.
7. 1516.1-2000 IEEE Standard for Modeling and Simulation (M&S) High Level Architecture (HLA). IEEE, 2000. 480 p.
8. White T. Hadoop: the Definitive Guide – O'Reilly, 2015. 600 p.