

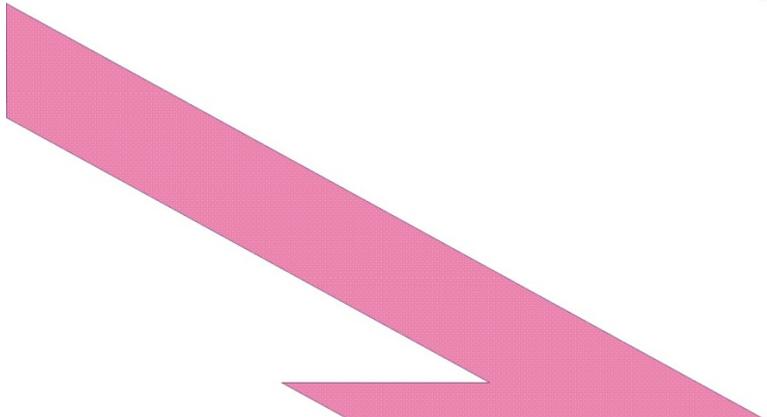


I T H E A



International Journal

INFORMATION **TECHNOLOGIES**
&
KNOWLEDGE



2017 **Volume 11** **Number 1**



**International Journal
INFORMATION TECHNOLOGIES & KNOWLEDGE**

Volume 11 / 2017, Number 1

EDITORIAL BOARD

Editor in chief: **Krassimir Markov** (Bulgaria)

Abdelmgeid Amin Ali	(Egypt)	Krassimira B. Ivanova	(Bulgaria)
Aleksey Voloshin	(Ukraine)	Larissa Zaynutdinova	(Russia)
Alexander Kuzemin	(Ukraine)	Levon Aslanyan	(Armenia)
Alexander Palagin	(Ukraine)	Luis F. de Mingo	(Spain)
Alexey Petrovskiy	(Russia)	Natalia Ivanova	(Russia)
Alfredo Milani	(Italy)	Nataliia Kussul	(Ukraine)
Arnold Sterenharz	(Germany)	Natalia Pankratova	(Ukraine)
Avram Eskenazi	(Bulgaria)	Nelly Maneva	(Bulgaria)
Darina Dicheva	(USA)	Nikolay Lyutov	(Bulgaria)
Ekaterina Solovyova	(Ukraine)	Orly Yadid-Pecht	(Israel)
Elena Chebanyuk	(Ukraine)	Rafael Yusupov	(Russia)
George Totkov	(Bulgaria)	Rumyana Kirkova	(Bulgaria)
Givi Bedianashvili	(Georgia)	Stoyan Poryazov	(Bulgaria)
Gocha Tutberidze	(Georgia)	Tatyana Gavrilova	(Russia)
Hasmik Sahakyan	(Armenia)	Vadim Vagin	(Russia)
Iliia Mitov	(Bulgaria)	Vasil Sgurev	(Bulgaria)
Irina Petrova	(Russia)	Velina Slavova	(Bulgaria)
Ivan Popchev	(Bulgaria)	Vitaliy Lozovskiy	(Ukraine)
Jeanne Schreurs	(Belgium)	Vladimir Ryazanov	(Russia)
Juan Castellanos	(Spain)	Volodimir Doncheko	(Ukraine)
Julita Vassileva	(Canada)	Martin P. Mintchev	(Canada)
Karola Witschurke	(Germany)	Yuriy Zaychenko	(Ukraine)
Koen Vanhoof	(Belgium)	Zurab Munjishvili	(Georgia)

**International Journal "INFORMATION TECHNOLOGIES & KNOWLEDGE" (IJ ITK)
is official publisher of the scientific papers of the members of
the ITHEA International Scientific Society**

IJ ITK rules for preparing the manuscripts are compulsory.
The **rules for the papers** for IJ ITK are given on www.ithea.org
Responsibility for papers published in IJ ITK belongs to authors.

International Journal "INFORMATION TECHNOLOGIES & KNOWLEDGE" Volume 11, Number 1, 2017
Edited by the **Institute of Information Theories and Applications FOI ITHEA**, Bulgaria, in collaboration with:
Institute of Mathematics and Informatics, BAS, Bulgaria; V.M.Glushkov Institute of Cybernetics of NAS, Ukraine;
Universidad Politecnica de Madrid, Spain; Hasselt University, Belgium;
St. Petersburg Institute of Informatics, RAS, Russia; Institute for Informatics and Automation Problems, NAS of the Republic of Armenia.

**Printed in Bulgaria
Publisher ITHEA®**

Sofia, 1000, P.O.B. 775, Bulgaria. www.ithea.org, e-mail: info@foibg.com
Technical editor: Ina Markova

Издател: ИТЕА®, София 1000, ПК 775, България, www.ithea.org, e-mail: info@foibg.com

Copyright © 2017 All rights reserved for the publisher and all authors.

® 2007-2017 "Information Technologies and Knowledge" is a trademark of ITHEA®

© ITHEA® is a registered trademark of FOI-Commerce Co.

ISSN 1313-0455 (printed)

ISSN 1313-048X (online)

НЕЛИНЕЙНАЯ СХЕМА КОМПРОМИССОВ В МНОГОКРИТЕРИАЛЬНЫХ ЗАДАЧАХ ПРИНЯТИЯ РЕШЕНИЙ

Альберт Воронин, Юрий Зиатдинов, Игорь Варламов

Аннотация: *Изложены теоретические основы концепции нелинейной схемы компромиссов в многокритериальных задачах оценивания и оптимизации. Предложен принцип оптимальности «подальше от ограничений» – рассматриваются как предпочтительные те решения, при которых частные критерии наиболее удалены от своих предельно допустимых значений. На этой основе проблема выбора традиционных схем компромиссов трансформируется в задачу построения единой универсальной скалярной свёртки критериев, адекватной заданной ситуации.*

Ключевые слова: *многокритериальные задачи, нелинейная схема компромиссов, Парето-оптимальность, аксиоматика.*

ACM Classification Keywords: *H.1 Models and Principles – H.1.1 – Systems and Information Theory; H.4.2 – Types of Systems.*

Введение

Решение многокритериальных задач связано с четырьмя проблемами векторной оптимизации [Voronin et al, 2013]: нормализация критериев, учёт приоритетов, определение области Парето и выбор схемы компромиссов. Только одна из этих проблем, а именно, определение области эффективных решений (области Парето) не зависит от предпочтений лица, принимающего решение (ЛПР) и имеет объективное научное обоснование. Остальные проблемы в той или иной мере требуют информации от ЛПР и, следовательно, их решение субъективно.

Это особенно относится к проблеме выбора схемы компромиссов, так как компромисс по своей природе является прерогативой человека. Суть понятия «компромисс» заключается в ответе на вопрос: *сколько единиц выигрыша по одному критерию мы готовы (по мнению ЛПР) заплатить за неизбежный проигрыш одной единицы по другому критерию (другим) в заданной ситуации?* Если такой ответ получен, то следующим шагом будет выбор схемы компромиссов для данной конкретной многокритериальной задачи, что обусловит получение искомого решения.

Выбрав схему компромиссов в форме скалярной свёртки частных критериев, можно перейти к построению конструктивного алгоритма решения многокритериальных задач. Скалярная свёртка – это математический приём сжатия информации и количественной оценки её интегральных свойств одним числом. Исторически эти вопросы рассматривал Блез Паскаль, который считается основоположником теории принятия решений. Он ещё в 1670 году ввёл два ключевых понятия теории: 1) частных критериев, каждый из которых оценивает какую-либо одну сторону эффективности решения и 2) принципа оптимальности, то есть правила, позволяющего по значениям критериев вычислять некоторую единую числовую меру эффективности решения. В качестве такого правила Паскаль предложил мультипликативную скалярную свёртку частных критериев.

Скалярная свёртка как выражение схемы компромиссов является основой для построения конструктивного аппарата решения многокритериальных задач. Существует большое разнообразие возможных схем компромиссов и, как следствие, скалярных свёрток частных критериев. Выбор конкретной схемы зависит от ситуации принятия многокритериального решения и индивидуальных предпочтений ЛПР.

Возможность решения проблемы основана на гипотезе существования некоторой *функции полезности* [Fishbern, 1978], возникающей в сознании ЛПР при решении конкретной многокритериальной задачи. Схема компромиссов в форме скалярной свёртки критериев рассматривается как математическая модель функции полезности. Практически все подходы к определению скалярной свёртки критериев сводятся к построению той или иной математической модели функции полезности ЛПР. Для этого необходимо: функцию полезности в заданных обстоятельствах 1) формализовать, 2) проанализировать, и 3) на основе этого анализа найти адекватную данной ситуации содержательную математическую модель (скалярную свёртку критериев).

Формализация функции полезности

В общем виде функция полезности ЛПР может быть представлена как $\Phi[y(x), r]$, где $x = \{x_i\}_{i=1}^n \in X$ – вектор возможных решений, определенный в допустимой области X ; $y = \{y_k\}_{k=1}^s \in M$ – вектор частных критериев, определенный в допустимой области $M = \{y \mid 0 \leq y_k \leq A_k, k \in [1, s]\}$; $A = \{A_k\}_{k=1}^s$ – вектор ограничений; $r \in R$ – вектор внешних условий, определенный на множестве возможных факторов R .

Ситуация принятия многокритериального решения зависит от факторов внешних условий r . Например, в условиях космического полёта особое внимание уделяется критерию, отражающему экономичность расхода кислорода на борту. В условиях тропиков важен критерий, характеризующий качество работы системы охлаждения двигателя, и.т.п. Обычно при решении многокритериальных задач предполагается, что вектор r является фиксированным и заданным: $r = r^\circ$. Тогда функция полезности ЛПР может быть представлена в виде

$$\Phi[y(x), r]_{r=r^\circ} = Y[y(x)]^\circ,$$

где $Y[y(x)]^\circ$ – скалярная свертка, построенная по схеме компромиссов, адекватной заданной ситуации.

Чаще всего пользуются линеаризованной моделью

$$Y[y(x)] \approx \sum_{k=1}^s a_k^\circ y_k(x),$$

где a_k° – весовые коэффициенты, которые назначает ЛПР в соответствии со своими индивидуальными предпочтениями в заданной ситуации.

Такой подход, обладая несомненным преимуществом простоты, характеризуется рядом недостатков, присущих методу линеаризации вообще. Так, линейная модель приводит к правильным результатам лишь в малых окрестностях рабочей точки, положение которой зависит от ситуации принятия многокритериального решения. Любое изменение ситуации приводит к необходимости перерасчёта весовых коэффициентов модели. Другие модели обладают своими преимуществами и недостатками. В настоящее время выбор схемы компромиссов не определяется теорией, а осуществляется эвристически, на основании индивидуальных предпочтений, профессионального опыта разработчика и сведений о ситуации, в которой принимается многокритериальное решение.

Проблема формализации выбора схемы компромиссов относится к классу слабоструктурированных задач и очень сложна в решении. При анализе возможностей

формализации выбора схемы компромиссов мы предполагаем, что существуют некоторые инварианты, правила, обычно являющиеся общими для всех ЛПР независимо от их индивидуальных склонностей, которых они одинаково придерживаются в той или иной ситуации. Согласно [Larichev, 1979], субъективность ЛПР имеет свои границы. В деловых решениях человек обязан быть рациональным, чтобы иметь возможность аргументировать мотивы своего выбора, логику своей субъективной функции полезности. Поэтому любые предпочтения ЛПР должны находиться в рамках определенной рациональной системы. Это и делает возможной формализацию.

В нашем случае предметом исследования является такая тонкая субстанция, как воображаемая функция полезности, возникающая в сознании ЛПР при решении конкретной многокритериальной задачи. При этом у каждого ЛПР функция полезности своя. Тем не менее, можно получить предпосылки для выявления вида содержательной модели функции полезности, если обнаружить и проанализировать некоторые общие закономерности, наблюдаемые в процессе принятия многокритериальных решений различными ЛПР в разных ситуациях.

Анализ функции полезности

В зависимости от вида многокритериальной задачи скалярная свертка $Y[y(x)]$ имеет различный физический смысл. В задаче векторного анализа эта свертка является *оценочной функцией*, ее величина количественно выражает меру качества многокритериального объекта при *заданных* значениях аргументов x . В задаче векторной оптимизации скалярная свертка $Y[y(x)]$ имеет смысл *целевой функции*. В результате ее экстремизации получается компромиссно-оптимальный вектор аргументов x^* . Ниже будем рассматривать задачу векторной оптимизации, считая для определенности, что все критерии $y(x)$ требуют *минимизации*. Тогда математически задача векторной оптимизации представляется в виде

$$x^* = \arg \min_{x \in X} Y[y(x)]$$

В дальнейшем удобно оперировать с относительными значениями критериев. Если вектор критериев $y(x)$ пронормирован вектором ограничений A ,

$$y_0(x) = \{y_k(x) / A_k\}_{k=1}^s = \{y_{0k}(x)\}_{k=1}^s,$$

то применяется скалярная свертка $Y[y_0(x)]$.

Теперь мы должны определиться с выбором принципа оптимальности. Нахождение экстремума взвешенной суммы частных критериев в теории принятия решений называется реализацией принципа Лапласа. Определение экстремума мультипликативной свёртки осуществляется в соответствии с принципом Паскаля. Концепция Чарнза-Купера [Charnes et al, 1961] требует минимизации расстояния от идеальной (утопической) точки до паретовской границы (принцип «*поближе к идеальной точке*»). Эти и другие принципы оптимальности имеют свои преимущества, но не содержат оценки соотношения частных критериев с их предельно допустимыми значениями (ограничениями).

Как реки имеют свои берега, так и критерии имеют свои ограничения. Навигация по рекам и морям невозможна, если существует опасность приближения к берегам (риффы, мели и пр.). При реализации известных принципов оптимальности не исключается возможность опасного приближения отдельных критериев к своим ограничениям, что весьма нежелательно. Чтобы избежать такой опасности, мы будем руководствоваться принципом «*подальше от ограничений*».

Введем понятие *напряженности ситуации* как меры близости относительных частных критериев к своему предельному значению (единице):

$$\rho_k = 1 - y_{0k}, \rho_k \in [0;1], k \in [1, s], \rho = \{\rho_k\}_{k=1}^s.$$

Проанализируем логику функции полезности, которой руководствуются ЛПР, принимая решения при различных значениях параметра напряжённости ситуации ρ .

Если мы говорим, что многокритериальное решение принимается в напряженной ситуации, то это значит, что в заданных условиях один или несколько частных критериев в результате решения могут оказаться в опасной близости к своим предельным значениям ($\rho_k \approx 0$). И если один из критериев достигнет предела (или выйдет за него), то это событие не компенсируется возможным малым уровнем остальных критериев (обычно не допускается нарушение любого из ограничений).

В этой критической ситуации необходимо всемерно препятствовать опасному возрастанию наиболее неблагоприятного (т.е. наиболее близкого к своему пределу) частного критерия, независимо от поведения в это время остальных критериев. Поэтому в достаточно напряженных ситуациях (при малых значениях ρ_k) ЛПР, если и допускает ухудшение максимального (наиболее важного в данных условиях) частного критерия на единицу, то только компенсируя это очень большим количеством единиц, которые улучшают остальные критерии. А в первом полярном случае ($\rho_k=0$) ЛПР оставляет для рассмотрения только этот единственный, наиболее неблагоприятный частный критерий, пренебрегая остальными. Следовательно, адекватным выражением схемы компромиссов в случае напряженной ситуации является минимаксная, чебышевская модель (эгалитарный принцип)

$$x^* = \arg \min_{x \in X} Y[y_0(x)]^{(1)} = \arg \min_{x \in X} \max_{k \in [1, s]} y_{0k}(x)$$

В менее напряженных ситуациях необходимо одновременное удовлетворение и других критериев, учитывая противоречивые интересы и цели системы. При этом ЛПР варьирует свою оценку выигрыша по одним критериям и проигрыша по другим в зависимости от ситуации. В промежуточных случаях выбираются схемы компромиссов, дающие различные степени частичного выравнивания частных критериев. С уменьшением напряженности ситуации предпочтения по отдельным критериям выравниваются.

И, наконец, во втором полярном случае ($\rho_k \approx 1$) ситуация настолько спокойная, что частные критерии малы и не возникает никакой угрозы нарушения ограничений. Здесь ЛПР считает, что единица ухудшения любого частного критерия вполне компенсируется равнозначной единицей улучшения любого другого критерия. Этому случаю соответствует экономичная схема компромиссов, обеспечивающая минимальные для заданных условий суммарные потери по частным нормированным критериям. Такая схема выражается моделью интегральной оптимальности (утилитарный принцип)

$$x^* = \arg \min_{x \in X} Y[y_0(x)]^{(2)} = \arg \min_{x \in X} \sum_{k=1}^s y_{0k}(x)$$

Таким образом, мы обнаружили две реперные точки (полярные значения параметра напряжённости ситуации ρ), в которых функции полезности различных ЛПП совпадают.

Проведенный анализ выявляет закономерность, в силу которой ЛПП варьирует свой выбор от модели интегральной оптимальности в спокойных ситуациях до минимаксной модели в напряженных ситуациях. В промежуточных случаях ЛПП выбирает схемы компромиссов, дающие различные степени удовлетворения отдельных критериев в соответствии со своими индивидуальными предпочтениями, но в соответствии с заданной ситуацией. Если принять выводы из приведенного анализа как логическую основу для формализации выбора схемы компромиссов, то можно предложить различные конструктивные концепции, одной из которых является концепция *нелинейной схемы компромиссов*.

Нелинейная схема компромиссов

Изложенный анализ даёт возможность заменить задачу *выбора* схемы компромиссов эквивалентной задачей синтеза некоторой *единой* скалярной свертки частных критериев, которая в различных ситуациях выражала бы разные принципы оптимальности. Сформулируем требования к синтезируемой функции $Y(y_0)$:

- она должна быть гладкой и дифференцируемой;
- в напряженных ситуациях должна выражать принцип минимакса;
- в спокойных условиях должна выразить принцип интегральной оптимальности;
- в промежуточных случаях должна приводить к парето-оптимальным решениям, дающим различные меры частичного удовлетворения критериев.

Такая универсальная свертка должна быть выражением схемы компромиссов, *адаптирующейся* к ситуации. Можно сказать, что адаптация и способность к адаптации – главная содержательная сущность исследования многокритериальных систем. Для придания способности к адаптации необходимо, чтобы синтезируемая скалярная свертка в явном виде включала характеристики напряженности ситуации ρ .

Из возможных функций, отвечающих перечисленным требованиям, выберем *простейшую*:

$$Y(\alpha, y_0) = \sum_{k=1}^s \alpha_k [1 - y_{0k}(x)]^{-1}; \alpha_k \geq 0, \sum_{k=1}^s \alpha_k = 1, \quad (1)$$

где $\alpha_k = \text{const}$ – формальные параметры, имеющие двойной физический смысл. С одной стороны – это коэффициенты, выражающие предпочтения ЛПР по отдельным критериям. С другой – это коэффициенты регрессии содержательной регрессионной модели, построенной на основе концепции нелинейной схемы компромиссов.

Таким образом, нелинейной схеме компромиссов соответствует модель векторной оптимизации, в явном виде зависящая от характеристик напряженности ситуации ρ :

$$x^* = \arg \min_{x \in X} \sum_{k=1}^s \alpha_k [1 - y_{0k}(x)]^{-1}. \quad (2)$$

В отличие от линейной модели, определенной в малой окрестности рабочей точки, нелинейная модель функции полезности ЛПР определена на всей допустимой области решений X и не требует перерасчета коэффициентов α_k при изменении ситуации.

Из выражения (2) видно, что если какой-либо относительный частный критерий, например $y_{0i}(x)$, начнет приближаться к своему пределу (единице), т.е. ситуация становится напряженной, то соответствующий член $Y_i = \alpha_i / [1 - y_{0i}(x)]$ в минимизируемой сумме возрастает настолько, что проблема минимизации всей суммы сведется к минимизации только данного наихудшего члена, т.е. критерия $y_{0i}(x)$. Это эквивалентно действию минимаксной модели. (Для исключения возможности деления на нуль в напряженных ситуациях в алгоритме оптимизации по формуле (2) можно применить следующее правило: если $y_{0i}(x) \geq 0,95$, то принимают $y_{0i}(x) = 0,95$).

Если же относительные частные критерии удалены от единицы, т.е. ситуация спокойная, то модель (2) действует эквивалентно модели интегральной оптимальности. В промежуточных ситуациях получают различные степени частичного выравнивания критериев.

Предложенная нелинейная схема компромиссов обладает свойством адаптации к ситуации принятия многокритериального решения. При этом адаптация осуществляется непрерывно, в то время как традиционный выбор схемы компромиссов происходит дискретно, в результате чего к субъективным погрешностям добавляются ошибки, связанные с квантованием схем компромиссов.

Конструкция нелинейной схемы компромиссов удобна еще и тем, что она не требует обязательной нормализации критериев. Эквивалентной формой скалярной свертки по нелинейной схеме компромиссов для минимизируемых критериев является

$$Y[y(x)] = \sum_{k=1}^s \alpha_k A_k [A_k - y_k(x)]^{-1}.$$

Более того, нелинейная схема не требует и обязательного приведения к единому способу экстремизации, например, к минимизации критериев. Для максимизируемых критериев формой скалярной свертки по нелинейной схеме компромиссов является

$$Y[y(x)] = \sum_{k=1}^s \alpha_k A_k [y_k(x) - A_k]^{-1}.$$

Если среди s критериев есть как минимизируемые, так и такие, что требуют максимизации, то скалярная свертка имеет вид

$$Y[y(x)] = \sum_{i=1}^m \alpha_i A_i [A_i - y_i(x)]^{-1} + \sum_{j=1}^n \alpha_j A_j [y_j(x) - A_j]^{-1}, \quad m + n = s,$$

где $y_i, i \in [1, m]$ – минимизируемые критерии; $y_j, j \in [1, n]$ – максимизируемые критерии; A_i – ограничения сверху для минимизируемых критериев; A_j – ограничения снизу для максимизируемых критериев.

В то же время независимо от вида скалярной свертки модель векторной оптимизации по нелинейной схеме компромиссов остается неизменной:

$$x^* = \arg \min_{x \in X} Y[y(x)].$$

Унификация

Как отмечено выше, выбор схемы компромиссов является прерогативой человека, отражением его субъективной функции полезности при решении конкретной многокритериальной задачи. Тем не менее, удалось выявить некоторые закономерности и на этой объективной основе построить скалярную свертку критериев, вид которой определяется содержательными представлениями сути изучаемого явления. Феномен же индивидуальных предпочтений ЛПР формально представлен наличием вектора α в структуре содержательных моделей (1) и (2).

Возможны различные оценки роли субъективных факторов в решении многокритериальных задач. Субъективность допустима и даже желательна, если такая задача решается в интересах конкретного человека. Действительно, если человек-оператор в эргатической системе имеет возможность подстраивать параметры рабочего места под свои индивидуальные особенности, то это повышает качество управляемого процесса; костюм, сшитый в ателье по мерке заказчика, обычно лучше, чем купленный в магазине готовой одежды, и т.п. Поэтому механизм индивидуальных предпочтений достаточно интенсивно применяется в практике решения многокритериальных задач.

Однако субъективность в их решении допустима и желательна лишь до тех пор, пока результат предназначается для конкретных ЛПР или узких коллективов людей со сходными предпочтениями. Если же он предназначен для общего использования, то обязан быть вполне объективным, унифицированным. В этих случаях механизм индивидуальных предпочтений из методики решения многокритериальных задач должен быть исключен во избежание произвола и неоднозначности результатов решения. По Гильберту, общая задача науки состоит в том, чтобы "освободить нас от случайности, предвзятости личных настроений, привычек и защитить от субъективизма". Формализация качественных понятий исключает неизбежную на интуитивном уровне рассмотрения неоднозначность толкований и, что гораздо существеннее, позволяет исследовать эти понятия математическими методами, что, как правило, уже на первых шагах приносит новые важные результаты.

Когда результат решения многокритериальной задачи предназначается для широкого использования, то он унифицируется и индивидуальные предпочтения нивелируются по статистике; становится применимым принцип недостаточного основания Бернулли-Лапласа: если априорные вероятности возможных гипотез неизвестны, то их следует положить равными, т.е. все гипотезы следует считать равновероятными. В применении к многокритериальной задаче это означает, что все весовые коэффициенты α_k , $k \in [1, s]$ в выражении для скалярной свертки нормированных минимизируемых критериев должны быть *равными*, если только нет

никаких предварительных данных о разноценности критериев. Так как в постановке задачи мы ограничились рассмотрением критериев, одинаковых по важности, то при унификации должны принять все весовые коэффициенты равными: $\alpha_k \equiv 1/s, \forall k \in [1, s]$. Тогда

$$Y(\alpha, y_0) = \frac{1}{s} \sum_{k=1}^s [1 - y_{0k}(x)]^{-1}.$$

Учитывая, что умножение на $1/s$ является монотонным преобразованием, которое, по теореме Гермейера, не изменяет результатов сравнения, переходим к унифицированному выражению для скалярной свертки критериев

$$Y(y_0) = \sum_{k=1}^s [1 - y_{0k}(x)]^{-1}. \quad (3)$$

Эту формулу рекомендуется применять во всех случаях, когда многокритериальная задача решается не в интересах какого-то одного конкретного ЛПР, а для широкого использования.

В [Voronin et al, 2013] предлагается во всех случаях начинать многокритериальное решение с использования формулы (3). Полученный результат и соответствующие ему значения частных критериев предъявляются ЛПР для оценки. Если ЛПР считает, что полученное решение его не удовлетворяет и требуется коррекция согласно его индивидуальным предпочтениям, то организуется процедура определения весовых коэффициентов $\alpha_k, k \in [1, s]$ и для оптимизации применяются формулы (1) и (2). Обращаясь к нашей аналогии, отметим, что купленный в магазине готовой одежды костюм обычно нуждается лишь в незначительной подгонке.

Учёт предпочтений ЛПР

Если многокритериальная задача решается в интересах конкретного ЛПР, то необходимо в выражении скалярной свертки определить весовые коэффициенты $\alpha_k, k \in [1, s]$, отражающие его индивидуальные предпочтения. Представим выражение для скалярной свертки по нелинейной схеме компромиссов в виде

$$Y(\alpha, y_0) = \sum_{k=1}^S Y_k(\alpha, y_0),$$

где $Y_k(\alpha, y_0) = \alpha_k [1 - y_{0k}(x)]^{-1}$. Выполнение условий $\alpha_k \geq 0, \sum_{k=1}^S \alpha_k = 1$ обеспечивает

парето-оптимальность решений.

Если проанализировать зависимости $Y_k(y_{0k})$ при различных значениях коэффициента α_k , то мы увидим, что наибольшее влияние коэффициенты α оказывают в "средних", промежуточных ситуациях. В полярных точках характеристик напряженности зависимости $Y_k(y_0)$ очень близки при любых значениях коэффициентов α . Это согласуется с выводами содержательного анализа. Действительно, в полярных точках представления различных ЛПР о схемах компромиссов совпадают (или достаточно близки), а в промежуточных ситуациях в наибольшей мере проявляются различия отдельных людей в индивидуальных предпочтениях.

Поэтому при экспериментальных оценках альтернатив нет смысла оперировать в очень напряженных или очень спокойных ситуациях, наиболее информативна средняя часть характеристики напряженности.

Для определения коэффициентов α можно применять различные подходы.

Дуальный подход

Практика решения многокритериальных задач показывает, что предположение о наличии с самого начала готовой и стабильной (хотя бы и в неявном виде) функции полезности у ЛПР справедливо далеко не всегда. Решая многокритериальную задачу, ЛПР сравнивает совокупности конкретных значений критериев при различных альтернативах, делает пробные шаги, ошибается и осмысливает соотношение между своими потребностями и возможностями их удовлетворения заданным объектом в заданной ситуации. При противоречивых критериях это соотношение по своей природе компромиссно, однако осознанной априори схемы компромиссов у ЛПР нет или она имеется пока лишь в зачаточном состоянии. Обычно представление о схеме компромиссов, необходимое для решения задачи, возникает и постепенно совершенствуется лишь в результате попыток ЛПР улучшить многокритериальное решение в серии пробных шагов. Ясно, что подразумевается наличие интерактивной компьютерной технологии, т.к. "в натуре" такая процедура обычно невозможна.

Таким образом, одновременно и взаимозависимо человек, с одной стороны, адаптируется к решаемой многокритериальной задаче, осуществляя структуризацию предпочтений и совершенствуя свое представление о функции полезности, а с другой – последовательно находит серию оптимальных в смысле текущей функции полезности решений. Взаимообусловленные процессы адаптации ЛПР к задаче и нахождения наилучшего результата носят дуальный характер и принципиально входят в методику человеко-машинного решения многокритериальных задач.

Представим схему нахождения эффективных решений в виде модели

$$x^{(\alpha)} = \operatorname{argmin}_{x \in X} \sum_{k=1}^s \alpha_k [1 - y_{0k}(x)]^{-1}, \alpha \in X_{\alpha}.$$

Выбирая различные значения параметров α из допустимой области

$$X_{\alpha} = \left\{ \alpha \mid \alpha_k \geq 0, \sum_{k=1}^s \alpha_k = 1 \right\},$$

по этой схеме получаем различные паретовские решения $x^{(\alpha)}$.

Задача заключается в такой организации интерактивной процедуры, чтобы последовательность генерируемых паретовских точек была *улучшающейся* с точки зрения ЛПР.

Как отмечалось, в начальной стадии процесса решения у ЛПР практически отсутствует не только аналитическое описание функции полезности, но и готовое априорное представление о ней. Поэтому интерактивная процедура должна быть организована как *дуальная*, а поисковый метод оптимизации должен допускать диалоговое программирование в порядковых шкалах и использовать минимальную информацию о функции полезности. Таким методом, основанным на сопоставлении предпочтений при специально рассчитываемых альтернативах, является порядковый аналог метода симплекс-планирования, обычно в модификации Нелдера-Мида.

Дуальная процедура начинается с нахождения первого ("общего") решения при $\alpha_k^0 = 1/s, k \in [1, s]$, что соответствует унифицированной модели

$$x^{(0)} = \operatorname{argmin}_{x \in X} \sum_{k=1}^s [1 - y_{0k}(x)]^{-1}.$$

Полученное решение и соответствующие ему значения частных критериев предъявляются ЛПР для оценки. Если ЛПР считает, что решение $x^{(0)}$ его не удовлетворяет и требуется коррекция согласно его индивидуальным предпочтениям, то организуется интерактивная процедура симплекс-планирования.

В каждой вершине исходного симплекса S_0 , расположенного в окрестностях решения $x^{(0)}$, рассчитываются паретовские решения $x^{(\alpha)}$ и соответствующие им значения частных критериев $u_0(x^{(\alpha)})$ предъявляются ЛПР для выбора наихудшей, с его точки зрения, вершины. По идее метода симплекс-планирования, значение функции полезности с большой вероятностью улучшится, если найти решение в новой точке, прямо противоположной худшей вершине в смысле исходного симплекса. Механизм симплекс-планирования состоит в том, что на каждой итерации текущий симплекс заменяется новым: худшая вершина отбрасывается и вместо нее в набор вводится новая, получаемая зеркальным отражением худшей точки относительно центра противоположной грани.

Так получается последовательность симплексов S_0, S_1, S_2, \dots . Согласно изложенному, последовательность генерируемых эффективных точек $x^{(\alpha)}$ является улучшающейся с точки зрения ЛПР и сходится к наилучшему, по его мнению, решению $x^* = x^{(\alpha^*)}$. Одновременно определяется вектор α^* , отражающий предпочтения данного конкретного ЛПР.

Описанный прием, по сути, сводит исходную сложную задачу синтеза многокритериального решения к последовательности более простых задач анализа специально организованной серии альтернатив. Отметим, что дуальная интерактивная процедура требует от ЛПР не многокритериальных оценок по шкалам баллов и даже не ранжирования решений, а только выбора наихудшей из s альтернатив на каждом шаге итерации. С точки зрения дуального подхода важно также, что данная процедура позволяет ЛПР возвращаться к предыдущим результатам и переосмысливать свои оценки и предпочтения. Поиск прекращается, когда ЛПР считает, что решения перестают существенно улучшаться.

Важным фактором, обуславливающим эффективность изложенного метода, представляется то, что начальная точка поиска выбирается не как произвольная точка в паретовском множестве, а как аксиоматически обоснованное решение, которое следует лишь скорректировать в соответствии с неформальными предпочтениями конкретного ЛПР. Процесс корректировки обеспечивает взаимную адаптацию: человек адаптируется к данной конкретной

многокритериальной задаче, а модель нелинейной схемы компромиссов становится отражением индивидуальных предпочтений данного человека.

Фундаментальным отличием свертки по нелинейной схеме от других известных скалярных сверток является органическая связь с ситуацией принятия многокритериального решения. По сути, предложенная свертка представляет собой нелинейную функцию регрессии (линейную по параметрам), выбранную по физическим соображениям и поэтому эффективную. Коэффициенты α в выражении для нелинейной скалярной свертки имеют смысл параметров нелинейной содержательной функции регрессии, поэтому, будучи найденными, они не изменяются от ситуации к ситуации, как в случае линейной и других известных сверток, не адаптирующихся к ситуации.

Задача определения коэффициентов α в дуальной процедуре может рассматриваться как задача синтеза *решающего правила*, которое, будучи применено формально, отражает адекватным образом логику конкретного ЛПР в любой возможной ситуации. Такая задача возникает, например, когда многокритериальная система работает в режиме советчика оператора в условиях дефицита времени. Здесь желательно, чтобы система в любой ситуации принимала такое же решение, как и данный оператор, если бы у него была возможность спокойно подумать. Аналогичные проблемы приходится решать и при разработке решающей системы интеллектуального робота, функционирующего в изменяющихся и неопределенных динамических средах, если хотят, чтобы он поступал так же, как на его месте поступил бы обучивший его человек, и т.п.

Парето-оптимальность

Покажем, что нелинейная схема компромиссов в унифицированной форме удовлетворяет условию парето-оптимальности. Дано:

1) множество допустимых решений $x \in X$, X выпуклое в E^n ;

2) решение по нелинейной схеме компромиссов

$$x^* = \left\{ x^* \mid x^* \in X; \forall x \in X : \sum_{k=1}^s [1 - y_{0k}(x^*)]^{-1} \leq \sum_{k=1}^s [1 - y_{0k}(x)]^{-1} \right\};$$

3) множество парето-оптимальных решений

$$X^K = \left\{ x' \mid x' \in X; \forall x \in X : y_{0k}(x') \leq y_{0k}(x), k \in [1, s] \right\}$$

причем хотя бы одно из указанных неравенств является строгим.

Требуется доказать, что $x^* \in X^K$.

Допустим обратное и положим, что решение x^* не принадлежит множеству X^K . Тогда найдется такое $\tilde{x} \in X$, что

$$y_{0k}(\tilde{x}) \leq y_{0k}(x^*), k \in [1, s],$$

причем некоторые (по крайней мере одно) из этих неравенств являются строгими. В этом случае при $\tilde{x} \in X$ справедливо неравенство

$$\sum_{k=1}^s [1 - y_{0k}(\tilde{x})]^{-1} \leq \sum_{k=1}^s [1 - y_{0k}(x^*)]^{-1},$$

что противоречит определению решения по нелинейной схеме компромиссов. Следовательно, $x^* \in X^K$, q.e.d.

По лемме Карлина, нелинейная схема компромиссов с весовыми коэффициентами тоже приводит к парето-оптимальному решению, если весовые коэффициенты определены на

симплексе: $\alpha \in X_\alpha = \left\{ \alpha \mid \sum_{k=1}^s \alpha_k = 1, \alpha_k \geq 0 \right\}$.

Аксиоматика

Сформулируем аксиоматику, на которой основывается концепция нелинейной схемы компромиссов. Основные системы аксиом, которым должна удовлетворять скалярная свертка в форме отношения порядка $Y(y)$, сформулированы Эрроу, Сенем, Нэшем, Милнором и Гурвицем. Классической принято считать аксиоматику Эрроу, другие системы аксиом разработаны главным образом вследствие необходимости разрешить так называемый *парадокс Эрроу*. Он состоит в том, что не существует порядка $Y(y)$, удовлетворяющего казалось бы очевидным пяти аксиомам Эрроу (приводятся на языке бинарных отношений):

- A1. Аксиома независимости от положительного линейного преобразования вектора эффективности:

$$Y\left(\{y_k(x)\}_{k=1}^s\right) \leftrightarrow Y\left(\{a_k y_k(x) + b_k\}_{k=1}^s\right), \\ \forall a_k > 0, b_k \in E^1.$$

- A2. Аксиома независимости от выбора элементов множества допустимых решений $x \in X$: если для двух допустимых векторов эффективности

$$y^1(x) = \{y_k^1(x)\}_{k=1}^{s^1}, y^2(x) = \{y_k^2(x)\}_{k=1}^{s^2}; y^1, y^2 \in F,$$

элементы $x^1, x^2 \in X$ таковы, что соблюдаются покомпонентные неравенства

$$y_k^1(x^1) \leq y_k^1(x^2), \forall k \in [1, s^1] \\ y_k^2(x^1) \leq y_k^2(x^2), \forall k \in [1, s^2]$$

и существуют компоненты, для которых эти неравенства строгие, то должны быть справедливы и выражения

$$x^1 Y(y^1) x^2, x^1 Y(y^2) x^2,$$

что означает, что решение x^1 предпочтительней решения x^2 в смысле отношений порядка как $Y(y^1)$, так и $Y(y^2)$.

- A3. Аксиома универсальности $Y(y)$ по $y \in F$: отношение порядка $Y(y)$ должно быть определено для любого допустимого вектора эффективности $y \in F$.
- A4. Аксиома отсутствия "диктатора": не существует $l \in [1, s]$ такого, что

$$y_l(x^1) \leq y_l(x^2) \rightarrow x^1 Y(y) x^2; x^1, x^2 \in X,$$

т.е. не допускается вывод $x^1 Y(y) x^2$ только на основании сравнения значений l -й компоненты вектора эффективности.

- A5. Аксиома ослабленной оптимальности по Парето:

$$y_k(x^0) \leq y_k(x) \rightarrow x^0 Y(y)x; x^0, x \in X, k \in [1, s],$$

т.е. если решение $x^0 \in X$ является недоминируемым по отношению ко всем $x \in X$ по каждому из частных критериев, то оно предпочтительно в смысле отношения порядка $Y(y)$.

Для разрешения парадокса Эрроу необходимо связать аксиоматику с напряженностью ситуации, в которой принимается многокритериальное решение. "Диктаторский" принцип согласования критериев не только допустим, но и, как следует из приведенного выше содержательного анализа, необходим в напряженных ситуациях. Поэтому вместо A4 сформулируем аксиому B4 следующим образом.

- B4. Аксиома необходимости "диктатора" в напряженных ситуациях: существуют $l \in [1, s]$ такие, что

$$y_{0l}(x) \leq y_{0l}(x^0) \rightarrow x Y(y)x^0; x^0, x \in X,$$

если решение x^0 таково, что для максимальной компоненты нормализованного вектора эффективности

$$y_{0l}(x^0) = \max_{k \in [1, s]} y_{0k}(x^0)$$

справедливо неравенство

$$1 - y_{0l}(x^0) \leq \varepsilon,$$

где ε – заданная малая величина.

Это значит, что в напряженных ситуациях вывод $x^1 Y(y)x^2$ должен быть обусловлен только l -й компонентой нормализованного вектора эффективности.

Учитывая, что парето-оптимальность нелинейной схемы компромиссов доказана, усилим аксиому A5 и сформулируем аксиому B5:

- B5. Аксиома оптимальности по Парето: если для $x^* \in X$ выполнено условие $x^* Y(y)x$, $x \in X$, то $x^* \in XK$.

Аксиому А1 ослабим, видоизменим и получим В1:

- В1. Аксиома независимости от положительного масштабного преобразования вектора эффективности:

$$Y(\{y_k(x)\}_{k=1}^s) \leftrightarrow Y(\{a_k y_k(x)\}_{k=1}^s) \forall a_k > 0.$$

Установлено, что рассмотренная скалярная свертка по нелинейной схеме компромиссов удовлетворяет модификации системы аксиом Эрроу, состоящей из аксиом В1-А2-А3-В4-В5.

Заключение

Преимуществом концепции нелинейной схемы компромиссов является возможность принятия многокритериального решения формально, без непосредственного участия человека. При этом на единой идейной основе решаются как задачи, имеющие значение для общего использования, так и те, основной содержательной сущностью которых является удовлетворение индивидуальных предпочтений ЛПР. Аппарат нелинейной схемы компромиссов, разработанный как формализованный инструмент для исследования систем управления с противоречивыми критериями, позволяет практически решать многокритериальные задачи широкого класса.

Acknowledgements

The paper is published with partial support by the project ITHEA XXI of the ITHEA ISS (www.ithea.org) and the ADUIS (www.aduis.com.ua).

Литература

- [Voronin et al, 2013] Воронин А.Н., Зиятдинов Ю.К. Теория и практика многокритериальных решений: Модели, методы, реализация. – Lambert Academic Publishing, 2013. – 305 с.
- [Fishbern, 1978] Фишберн П.С. Теория полезности для принятия решений. – М.: Наука, 1978. – 352 с.
- [Larichev, 1979] Ларичев О.И. Наука и искусство принятия решений. – М.: Наука, 1979. – 200 с.
- [Charnes et al, 1961] Charnes A., Cooper W. Management models and industrial applications of linear programming. – New York: Wiley, 1961. – 240 p.

Сведения об авторах



Воронин Альберт Николаевич – профессор, доктор технических наук, профессор кафедры Компьютерных информационных технологий Национального авиационного университета, проспект Комарова, 1, Киев-58, 03058 Украина; e-mail: alnv@voliacable.com



Зиятдинов Юрий Кашафович – профессор, доктор технических наук, зав. кафедрой компьютерных информационных технологий Национального авиационного университета, проспект Комарова, 1, Киев-58, 03058 Украина; e-mail: oberst@nau.edu.ua



Варламов Игорь Давыдович – кандидат технических наук, нач. кафедры применения космических систем и геоинформационного обеспечения Института информационных технологий Национального университета обороны Украины имени Ивана Черняховского

Non-Linear Trade-off Scheme in Multicriteria Decision-Making Problems

Albert Voronin, Yuriy Ziatdinov, Igor Varlamov

Abstract: *The theoretical basis is considered of the concept of non-linear trade-off scheme in multicriteria problems of estimation and optimization. The principle of optimality "out of limits" is proposed – those decisions are regarded as preferred in which the partial criteria are the furthest from their maximum permissible values. On this basis, the problem of selection of traditional schemes of compromise is transformed into the task of building a single universal scalar convolution of criteria appropriate to given situation.*

Keywords: *multicriteria problems, non-linear trade-off scheme, Pareto-optimality, axiomatics.*

ОБНАРУЖЕНИЕ КИБЕРУГРОЗ С ПОМОЩЬЮ АНАЛИЗА СОЦИАЛЬНЫХ СЕТЕЙ

Людмила Кириченко, Тамара Радивилова, Алексей Барановский

Аннотация: В работе проведен обзор основных методов анализа социальных сетей, которые используются для обнаружения киберугроз. Представлены основные типы угроз в социальных сетях и описаны некоторые методы защиты для их предотвращения. Рассмотрены типичные задачи анализа социальных сетей, направленные на выявление киберугроз, такие как обнаружение сообществ в сети, обнаружение лидеров и экспертов в сообществах, анализ стабильности сообществ, кластеризация текстовой информации и другие. Описаны основные классы методов теории графов и интеллектуального анализа данных, которые широко применяются в анализе социальных сетей. Показано применение методов фрактального анализа для исследования рядов показателей активности пользователей сетей.

Ключевые слова: анализ социальных сетей, киберугрозы, data mining, методы обнаружения лидеров, методы обнаружения экспертов, фрактальный анализ.

ITHEA Keywords: G.3 Probability and statistics - Time series analysis, Stochastic processes, G.1 Numerical analysis, G.1.2 Approximation - Wavelets and fractals, I.2 Artificial intelligence - I.2.0 General, K.6 Management of computing and information systems – K.6.5 Security and Protection.

Введение

Быстрое развитие социальных сетей и способность собирать информацию из них привели к заметному повышению интереса к анализу социальных сетей (АСС) и появлению новых методов, которые становятся все более популярными и используются в различных областях: для поиска экспертов, набора команды специалистов, социальных рекомендаций, маркетинга, коммуникаций, рекламы и многих других. В настоящее время АСС используется для изучения ряда экономических и организационных явлений и процессов, для борьбы с отмыванием денег, кражей личности, онлайн-мошенничеством, кибер-атаками, при расследовании незаконных операций с ценными бумагами и инвестициями, для предотвращения беспорядков и др. [Carley, 2002, Stohl, 2007, Easley, 2010, Russell, 2011].

Социальные сети все более и более широко используются в интересах информационного и психологического воздействия. Они предоставляют возможности с точки зрения влияния на формирование общественного мнения, принятие политических, экономических и военных решений, влияют на информационные ресурсы противника и распространение специально подготовленной информации (дезинформации) [Додонов, 2013, 2014]. Исходя из этого многие страны создают национальные центры кибербезопасности и документы по стратегии кибербезопасности [National Cyber Security Strategy 2016-2021 for United Kingdom, National cyber security strategy for Ukraine, The NATO Cooperative Cyber Defence Centre].

Таким образом, задача сбора информации, мониторинга и анализа социальных сетей для обеспечения информационной безопасности важна и актуальна. Целью данной работы является обзор и анализ основных задач и методов анализа социальных сетей, используемых для обнаружения, предотвращения и борьбы с угрозами в социальных сетях.

Угрозы в социальных сетях и профилактические меры

В наши дни в социальных сетях существует целый ряд различных угроз. В [Palo, 2017] описаны угрозы для социальной сети деловых кругов, в [Carley, 2002; Дзюндзюк, 2011; Матвиенко, 2011; Shantanu, 2017] описаны угрозы для социальных сетей и некоторые профилактические меры.

Спам в социальных сетях. Спам является одним из самых классических атак всех времен. Злоумышленниками генерируются фиктивные учетные записи, и тысячи запросов друзей автоматически отправляются в надежде, что кто-то их примет. После принятия атакующий может начать отправлять спам-сообщения. Даже запрос на соединение с другом позволяет отправлять короткие сообщения внутри него без какого-либо предыдущего соединения между пользователями. Для борьбы со спамом многие сообщества внедряют тесты CAPTCHA, которые решают, когда отправляется слишком много сообщений. Это должно остановить или, по крайней мере, замедлить автоматическое распространение сообщений. Кроме того, у злоумышленников часто есть возможность использовать несколько учетных записей параллельно, пока каждая из них не блокируется дневным лимитом. Большинство социальных сетей предлагают функцию для обозначения сообщений как спама и блокирования их в будущем, что помогает при условии не частого переключения записей злоумышленниками.

Угрозы социальной инженерии. Самая популярная тактика для киберпреступников. Социальные сети позволяют злоумышленникам находить конфиденциальную информацию, которая может быть использована для имущественного и морального ущерба.

Размещение приманок в социальных сетях. Идея проста: использовать ключевые слова и ссылки таким образом, чтобы спам-сообщения получали лучший список учетных записей. Некоторые нападавшие даже начали манипулировать доброжелательными сообщениями Twitter перед пересылкой. Злоумышленники ищут новые сообщения, содержащие горячие ключевые слова. Это может быть сообщение о сомнительной цели офсайда в последнем футбольном матче, с сокращенным URL-адресом, связанным с соответствующей новостной статьей. Затем мошенник принимает это сообщение, заменяет исходный сокращенный URL своей собственной ссылкой, указывающей на вредоносный сайт, и повторно читает сообщение. Это делает практически невозможным для обычных посетителей отличать хорошие и вредоносные сообщения. Следовательно, шансы, что невинный пользователь, который что-то ищет, наткнется на злонамеренную ссылку относительно высоки. Конечно, мы также видим типичные заманчивые сообщения, предлагающие ссылки на видео голых знаменитостей или взломанные программные средства, рассылаемые спамом в надежде, что кто-то найдет и нажмет на них.

Друзья. Доверие к тем, кто входит в список «друзей», всегда выше, чем случайным людям. С одной стороны, это хорошо, поскольку формируется лояльная аудитория вокруг компании, бренда или человека. С другой стороны, это возможность для злоумышленников.

Олицетворение друзей. Почти во всех социальных сетях олицетворение является реальной проблемой для всех. Поскольку сообщение, кажется, исходит от учетной записи друга, люди склонны доверять ему. Это присущее доверие и обычное любопытство приводит к высокой скорости кликов на вредоносных ссылках, что делает атаки получения пароля очень успешными. Эти сообщения об обновлениях часто содержат ссылки на другие вредоносные сайты, чтобы получить больше паролей к учетным записям.

Возможность замены человека или маскарад: наверняка, не совсем ясно, кто скрывает свои действия за именем друзей или прячется за фотографиями друзей в профиле социальной сети. Возможно по IP-адресу отправителя собрать о нем по крайней мере некоторую информацию в корреспонденции по электронной почте, которая не работает в социальной сети. Этот маскарад

возможен и на корпоративном уровне. Результатом такого вредоносного сценария может быть фишинг, организация «черного PR» или «AntiPR». Уже было множество случаев, когда было непонятно, кто создал сайт от имени какой-либо компании, и это создает проблему для оригинального бренда.

Кража паролей и фишинг. Для аутентификации в социальных сетях, использующих пароли, достаточно знать последовательность символов, и станет возможным отправлять рекламу, некоторую информацию от имени других или мотивировать получателей на любые негативные действия, в частности, на передачу ссылок на вредоносный сайт или запускать вредоносный код и выполнять другие (часто нелегальные) действия. Кроме того, некоторые компании используют социальную сеть для продвижения своих собственных продуктов, а кража паролей группы администраторов позволяет украсть саму группу. А для получения конфиденциальной информации традиционно используются фишинг, фиктивные сайты, социальная инженерия и многие другие. Защита от этих методов атаки рассматривается как DLP-система (Data Loss Prevention) и технологии репутации, которые интегрированы в различные антивирусные продукты.

Использование сервисов сокращения URL-адресов. В последние годы особенно популярны услуги сокращения URL-адресов, позволяющие маскировать нежелательный адрес сайта под короткой ссылкой. Фактически, домен перенаправляет посетителя. Сегодня идет активная борьба с этими рисками – служба сокращения URL-адресов начала использовать усовершенствованные механизмы для обнаружения спама и других угроз. Однако для пользователей социальных сетей эта угроза сохраняется: соблазнительные сообщения и предложения от уже известных контактов, которые были взломаны, часто приводят к загрузке вредоносного программного обеспечения или отображению нежелательных веб-страниц.

Использование тех же имен пользователей и паролей в корпоративной сети и внешних социальных ресурсах – эта атака также известна как «Daisy Chain». В результате взлом профилей социальной сети пользователей значительно повышает риск проникновения на корпоративные ресурсы от имени одного из сотрудников компании.

Веб-атака. Социальные сети могут использоваться хакерами для организации атак через уязвимости в браузерах, а также XSS / CSRF-атаки. Инструментами для таких атак могут быть троянские программы, поддельные антивирусы, социальные черви, вредоносные JavaScript и HTML-код, которые используются для распространения собственных списков друзей и других. Их главная цель – войти в информационную систему посетителя социальной сети, его рабочей станции или устройства и закрепиться в ней. Для защиты используются такие традиционные инструменты, как антивирусное программное обеспечение, способное работать в реальном времени и блокировать загрузку вредоносного кода.

Утечка информации и компрометация поведения сотрудников компании. Социальные сети могут использоваться для организации утечек важной информации компании, а также для подрыва ее репутации. Такая атака может вестись от внутренних сотрудников, которые недовольны руководством или специально встроенными инсайдерами. В социальных сетях люди часто ведут себя совершенно иначе, чем в корпоративной среде общения, и, возможно, шокирующая публикация и грубые реплики могут нанести определенный ущерб репутации их работодателей. DLP-системы и продукты для анализа публикаций в Интернете, предназначены для защиты от этих угроз.

Advanced Persistent Threat (APT)-атака. Исходя из вышесказанного, социальные сети могут использоваться как шлюз или источник угрозы организации, службе или другому подразделению для квалифицированных хакеров с самым современным и продвинутым вредоносным кодом, методами атак и хакерской методологией вообще. Есть много случаев, когда социальные сети стали источником информации для злоумышленников. Согласно трем хорошо известным методикам атаки: Lockheed Martin Cyber Kill Chain, Mandiant APT Attack Life Model и ISSP ThreatSCALE Model социальные сети могут использоваться на первых шагах всех этих моделей. Первый этап каждой модели - «Reconnaissance» - сбор информации о компании, ее сотрудниках, их позициях, контактах и коммуникациях, домашних адресах и т.д. С этой точки зрения социальные сети являются лучшим источником и инструментом для этого этапа. Далее, социальные сети могут быть легко использованы в качестве инструмента доставки на других этапах моделей: «Вторжение» в ThreatSCALE, «Поставка» в Cyber Kill Chain, «Проникновение» в Mandiant APT Model. Например: злоумышленник может отправить злонамеренную URL-ссылку жертве, MS Word или PDF-документ с вредоносным кодом и т.д.

Так же к основным угрозам в социальных сетях относятся следующие: содержание с признаками подстрекательства к расовой, этнической или религиозной ненависти, пропаганда тоталитарных сект; пропаганда и публичное оправдание терроризма; кибер-унижение и кибер-запугивание; популяризация и распространение наркотиков [Stohl, 2007].

Для защиты от представленных угроз службами национальной информационной безопасности, органами государственной власти, предприятиями государственного и частного сектора экономики и индивидуумами решаются следующие задачи:

- обнаружение информационных атак: определение узлов, из которых выполняется атака, оптимальное размещение сигнальных точек;
- предотвращение информационных атак: сметная стоимость нападения на объект и затраты на оборону;
- формирование и разрушение различных сетей социальной информации;
- обнаружение сообществ злоумышленников, таких как террористы, отслеживание вредоносной деятельности.

Можно выделить следующие направления противодействия информационному и психологическому воздействию виртуальных сообществ [Гриненко, 2012; Пелецишин, 2013]:

- силовые методы: закрытие серверов, формирование трафика;
- правовая и нормативная практика: уголовная ответственность организаторов и участников виртуальных сообществ;
- интернет-цензура;
- мониторинг и анализ социальных сетей.

Рассмотрим преимущества и недостатки каждого метода. Первые два метода эффективны в краткосрочной перспективе, но у них есть недостатки: отсутствие географических границ и ограничений для мгновенного распространения, сбора, обработки и использования информации – вне сферы действия законов, регулирующих правовое регулирование любого правительства; анонимность; легкодоступная изменчивость информации в электронной форме. Цензура плохо работает в демократических государствах, основанных на свободе слова.

Мониторинг и анализ социальных сетей. Методы мониторинга и анализа социальных сетей более эффективны в долгосрочной перспективе, но требуют участия специалистов в различных областях науки. Поскольку виртуальные социальные группы обладают способностью реорганизоваться, основной задачей мониторинга и анализа виртуальных сообществ, представляющих угрозу для национальной безопасности информации, является не их уничтожение, а управление и контроль их деятельности с помощью разнообразных методов.

На данный момент разработано большое количество специального программного обеспечения для мониторинга и анализа интернет-среды. Основными функциями этих систем являются: мониторинг, который обеспечивает автоматизированный поиск информации в интернет-среде и позволяет определять и изменять ключевые слова для поиска информации с использованием языков поиска информации и анализ, который включает в себя автоматическую обработку информационных потоков, выявление фактов и событий, визуализацию аналитических данных в форме дайджестов, графов, графиков и других типов отчетов.

Мониторинг относится к процессу непрерывного сбора информации из социальных сетей в целях дальнейшего анализа. Так что в научных исследованиях рассматривалось проведение поиска рассматриваемого глобальной поисковой системой для социальных сетей [Григорьев, 2007; Пелещин, 2010; 2012; 2013; Тимовчак-Максимец, 2010], и разработка коммерческих поисковых систем для специальных приложений [Горбулин, 2009; List of social networking websites, 2017; Shantanu, 2017; Смирнов, 2014], которые не учитывают особенности функционирования дискуссионных страниц.

Задачи анализа социальных сетей

В процессе анализа социальных сетей в первую очередь решаются следующие основные задачи, которые в дальнейшем декомпозируются на более подробные и точные, в зависимости от полученного результата [Fortunato, 2010, Aggarwal, 2011, Charu 2012, Батура, 2013].

Обнаружение сообществ в сети. Сообщества в сети характеризуются наличием большого числа связей между их участниками и значительно меньшим количеством связей с другими членами сети. Сообщество может соответствовать группам веб-страниц, которые имеют похожие темы [Flake, 2002], группы связанных лиц в социальных сетях [Girvan, 2002] и т. д. Простейшим случаем сообщества является такое, где каждый участник связан с каждым, а другие члены сети не общаются с членами сообщества (клика). Обнаружение сообществ (явных

и неявных) является важной задачей анализа сетей, включающей в себя классификацию членов сообщества, и, как результат, идентификацию однородных групп, групп лидеров или экспертов [Coscia, 2011, Бузун 2012, Kolomeychenko, 2014]. Обнаружение сообщества во многих случаях является кластеризацией, традиционной задачей Data Mining по отношению к различным социальным сетям. Подходы к распределению целевых групп путем выявления сообществ позволяют построить математические модели, а затем использовать модели информационного влияния и управления [Губанов, 2010]. В то же время анализ сетей исследует структуру отношений между участниками в различных областях, и обнаруживает неявные связи между ними с использованием теории графов [Ehrlich, 2005]. Более подробный обзор методов обнаружения сообщества можно найти в [Fortunato, 2010].

Анализ стабильности сообщества. Анализ явных и неявных сообществ позволяет исследовать устойчивость социальных структур. Для анализа временной стабильности групповой структуры обычно используется следующая методика. Создается первая трехмерная матрица, где строки представляют оценки взаимодействий участника со всеми другими участниками, представленные самими участниками; столбцы - это собственные оценки взаимодействия участника; временные периоды расположены на третьей оси. После этого применяются методы уменьшения размерности (например, анализ главных компонент), то есть проектирование вершин в евклидово пространство уменьшенной размерности для описания отношений между строками и столбцами матрицы. В результате возможна визуализация изменения статуса пользователя сети на фоне изменений в статусе подгрупп [Johnson, 1994]. Полученная проекция может быть сгруппирована с использованием стандартных алгоритмов кластеризации [Koren, 2003].

Обнаружение лидеров в сообществах. Поиск лидеров в сообществе является важной задачей АСС, поскольку в исследовании и моделировании информационного влияния важно иметь данные о характере взаимодействий членов сообщества, связи между ними и законами распределения информационных потоков. Согласно [Goyal, 2008], некоторый участник является лидером, если после совершения определенного действия значительное число других повторяет одно и то же действие в заданный интервал времени. Задачи обнаружения лидеров широко распространены во многих областях. Например, в [Watts, 2007] «гипотеза влиятельных членов» рассматривается в связи с маркетинговыми задачами; выбор многих лиц для предложения какого-либо продукта или инноваций [Kempe, 2003]; распространение и максимизация влияния в

конкурентных социальных сетях и привлечение последователей, вирусный маркетинг [Carnes, 2007]; распространение социального влияния [Dodds 2005, Слабченко 2013] и др.

Обнаружение экспертов в сетях. Социальная сеть может быть инструментом для поиска экспертов в конкретной области. Обнаружение экспертов связано с проблемами определения доверия и распределения влияния, а также с проблемой распространения информации в сети. С этой точки зрения распространение экспертного влияния транзитивно, т.е. влияние передается от одного узла к другому, уменьшаясь с каждым вовлеченным узлом экспертов [Bonchi 2011, Укустов, 2013]. Более подробный обзор методов экспертов по обнаружению можно найти, например, в [Charu, 2012].

Эволюция в динамических социальных сетях. С течением времени в социальных сетях появляются новые участники, некоторые участники прекращают общение, создаются новые ссылки, некоторые ссылки устаревают, так как участники больше не общаются. Это приводит к изменениям в структуре социальных сетей в целом и некоторых сообществах в частности. Таким образом, возникают важные вопросы: в соответствии с какими правилами происходят долгосрочные изменения между основными сообществами в социальных сетях; как развиваются сообщества с течением времени; как узнать изменения, которые могут произойти, какие существуют возможности их отслеживания и представления. Для исследования динамики сети используются подходы, описанные в [Bonchi 2011, Докука 2015]. Большую роль играет моделирование эволюции сетевого графа, в котором исследуются различные стратегии построения сети. Например, в [Leskovec, 2005] было обнаружено, что со временем плотность сети возрастает по степенному закону. Среди работ, представляющих алгоритмические инструменты для анализа эволюции сети, можно выделить [Tantipathanandh, 2007], где предложены алгоритмы оценки сообщества пользователей и их изменение со временем. Основное внимание уделяется определению приблизительных кластеров пользователей и эволюционных кластеров. Более подробный обзор моделей и методов эволюции социальных сетей можно найти в [Charu 2012, Aggarwal 2014, Saoussen 2014].

Прогнозирование формирования связей. Исследования, направленные на выявление и прогнозирование возможных связей между участниками или сообществами в будущем, полезны для извлечения интересующей информации из социальной сети. Связи динамичны и могут со временем сильно изменяться. Структура сети и информация об особенностях различных

участников могут быть задействованы в процессе прогнозирования связей. Задача прогнозирования связи состоит в определении того, будут ли связаны два конкретных участника друг с другом через определенный интервал времени. Эта вычислительная задача, которая основана на анализе эволюции социальной сети во времени, и называется задачей прогнозирования формирования связей. Для ее решения используется автоматическое моделирование процесса развития социальной сети с привлечением некоторых сетевых характеристик, таких как количество общих соседей, кратчайший путь, влияние вершин, время первого входа в социальную сеть. Для решения этих задач предлагается построить множество структурных и реляционных моделей. Существуют модели предсказания ссылок на основе машинного обучения, используя личную информацию пользователей сети для повышения точности прогнозирования [Liben-Nowell, 2003]. Иногда для выявления связей между пользователями используются иерархические, стохастические (марковские) и реляционные модели. В других моделях [Kumar, 2004] в качестве основы предлагается использовать свойства пользователей, например, наличие числа многих ссылок в блогосфере, которое можно объяснить сравнением демографических групп, общих интересов или географической близостью. Обзор моделей и методов прогнозирования ссылок представлен в [Hasan, 2011, Kushwah, 2016].

Кластеризация текстовой информации на основе частотного анализа. Собирая и группируя текстовые данные из социальной сети, можно определить основные темы и события, обсуждаемые пользователями социальных сетей в разных городах и странах. В настоящее время существует множество методов, позволяющих решить проблему классификации и кластеризации текстов. На этой основе реализовано много систем с использованием семантической обработки текста. Одним из основных методов частотного анализа является подсчет количества вхождений каждого слова в документ. На основе полученной информации можно сделать так называемое «облако тегов» – визуальное представление веса слов в документе. Для корректного вычисления веса слова необходимо использовать меры, которые не только посчитают количество появлений слова в документе, но также учитывать количество вхождений слова в другие документы. Примером таких мер является TF-IDF [Ramos, 2003]. В работе исследователей, занимающихся кластерным анализом текстовой информации в различных поисковых системах, часто используется индуктивная мера Word2vec [Wang, 2014, Yu, 2014]. Принцип действия меры состоит в том, чтобы найти отношения между контекстом слова в соответствии с предположением, что слова, которые находятся в сходных контекстах,

имеют тенденцию означать сходные вещи, то есть быть семантически близкими. Word2vec анализирует использование контекстов слов и приходит к заключению, что они близки по смыслу или нет. Алгоритмы, лежащие в основе word2vec, подробно описаны в работах [Mikolov, 2013, 2013 (1)].

Основные классы методов, используемых для АСС

Среди основных классов методов, используемых в АСС, можно выделить следующие: методы анализа графов, статистические методы, интеллектуальный анализ данных, методы теории оптимизации и теории алгоритмов. Также удобно выделить отдельно методы семантического анализа и анализа текстов. В этом случае необходимо обратить внимание на поддержку системой языка, на котором общаются пользователи анализируемой социальной сети.

Методы анализа графов.

Графовые модели и методы их анализа играют важнейшую роль в анализе социальных сетей, поскольку любую социальную сеть можно математически представить в виде графа $G=(V,E)$, где V – множество вершин графа, E – множество ребер графа, N – количество вершин в графе. Графовые модели социальных сетей используются для моделирования экономических и коммуникационных связей людей, анализа процессов распространения информации, нахождения сообществ и связанных подгрупп, на которые можно разбить всю социальную сеть. В графе социальной сети вершинами являются, условно говоря, участники, а ребра означают наличие отношений между ними. Отношения могут быть как направленными, так и ненаправленными. Как правило, выделяют два типа отношений: «дружба» (люди знакомы друг с другом) и «интересы» (есть общие интересы, люди входят в одну группу по интересам).

Для анализа графовых моделей социальных сетей удобно использовать коэффициент плотности, определенный как отношение числа ребер в анализируемом графе к числу ребер в полном графе с тем же числом вершин (полный граф – это граф, в котором все вершины соединены между собой). Кроме этого, сеть могут характеризовать такие величины, как число путей заданной длины (путь – последовательность вершин, связанных между собой), минимальное число ребер, удаление которых разбивает граф на несколько частей, и другие. [Чураков, 2001]

Анализ центральности и других локальных свойств. Чтобы определить относительную важность (вес) вершин графа (т. е. насколько участник в рамках конкретной сети является влиятельным), вводят понятие центральности – меры близости к центру графа. Центральность можно определить разными способами, поэтому существуют различные меры центральности [Hanneman, 2005].

Центральность по степени (Degree centrality) определяется как количество связей, инцидентных вершине (инцидентность – связь между вершиной и ребром) Выделяют входящие и исходящие связи. Входящие связи характеризуют популярность человека, выходящие – его общительность. Полученную величину можно нормировать, разделив на общее число участников в сети. Другими словами, центральность по степени предполагает, что среди участников сети более влиятельным является тот, у кого больше друзей, либо тот, кто входит в большее количество сообществ. Тем не менее участник сети, имеющий большое количество друзей, может быть связан с остальным графом маленьким количеством ребер.

Центральность по близости (Closeness centrality) является показателем, насколько быстро распространяется информация в сети от одного участника к остальным. В качестве меры расстояния между двумя участниками используется кратчайший путь по графу. Так, непосредственные друзья участника находятся на расстоянии 1, друзья друзей – на расстоянии 2, и т. д. Далее берется сумма всех расстояний и нормируется. Полученная величина называется удаленностью вершины от других вершин. Близость определяется как величина, обратная удаленности. Центральность по близости позволяет понять, насколько близок рассматриваемый участник ко всем остальным участникам сети.

Центральность по посредничеству (Betweenness centrality). Еще одной характеристикой участника является его важность при распространении информации. Именно в этом контексте центральность по посредничеству оценивает участника. Она рассчитывается как число кратчайших путей между всеми парами участников, проходящих через рассматриваемого участника.

Центральность по собственному вектору (Eigenvector centrality) демонстрирует зависимость между центральностью участника и центральностями его друзей. Участник, который

имеет много связей с теми, у кого тоже много связей, имеет высокую центральность по собственному вектору. Таким образом, чем больше у участника друзей и чем они центральнее, тем больше его центральность. Мера центральности по собственному вектору сложна для вычисления и возможна только с помощью специализированных компьютерных программ.

Центральность можно вычислить при помощи алгоритма ссылочного ранжирования (PageRank), который используется в поисковой системе Google. В основу положен принцип «важности» веб-страницы: чем больше ссылок на страницу, тем она «важнее». Кроме того, вес самой страницы определяется весом ссылки передаваемой на нее страницы. Таким образом, PageRank – это метод вычисления веса страницы путем подсчета важности ссылок на нее [Langville, 2006].

Важными характеристиками связей сети являются *сбалансированность* и *транзитивность*. Сбалансированность – это отсутствие ситуаций типа «положительное взаимодействие (дружба, партнерство) между 1-м и 2-м участниками, а также между 1-м и 3-м, но негативное взаимодействие (вражда, соперничество) между 2-м и 3-м». Утверждается, что сбалансированные сети психологически более комфортабельны для участников и более устойчивы по сравнению с несбалансированными [Johnson, 1994]. Транзитивность – это выполнение условий вида «если есть взаимодействие между 1-м и 2-м участниками, а также между 2-м и 3-м, то имеет место взаимодействие между 1-м и 3-м. Данные характеристики описывают локальные связи участников и часто используются при анализе диад и триад.

Полезной характеристикой при анализе социальных сетей является *уровень доверия*. Алгоритм вычисления уровня доверия (TrustRank) предложен в [Gyöngyi, 2004]. Изначально он был создан для отделения информативных веб-страниц от спама. Если говорить в терминах сайтов, то сначала эксперты выбирают некоторую выборку надежных сайтов, которые принимаются за эталон. Дальнейшие действия базируются на положении, что хорошие сайты редко ссылаются на плохие, а плохие очень часто ссылаются на хорошие. TrustRank – величина, которая дает оценку того, можно ли доверять конкретному сайту, считая, что он не содержит спама. Чем больше ссылок на сайте, тем меньше доверия «передается» по каждой такой ссылке. Степень доверия сайту (TrustRank) убывает с увеличением расстояния между ним и первоначальной выборкой.

Среди наиболее важных оптимизационных задач, связанных с графами, можно выделить задачу коммивояжера [Travelling salesman problem]. Это одна из самых известных комбинаторных задач оптимизации, которая заключается в нахождении наиболее прибыльного пути, проходящего через заданные вершины хотя бы один раз, а затем возвращении к исходному. В условиях

задания указываются критерии прибыльности маршрута (кратчайшие, самые дешевые, кумулятивные критерии и т. д.) и соответствующие матрицы расстояний, стоимости и тому подобное. Одним из эффективных методов для нахождения приближенных решений задачи коммивояжера, а также решения аналогичных задач поиска маршрутов на графах является муравьиный алгоритм. Суть подхода заключается в анализе и использовании модели поведения муравьев, ищущих пути от колонии к источнику питания и представляет собой метаэвристическую оптимизацию [Ant colony optimization algorithms].

Интеллектуальный анализ данных

Многие компании стремятся проанализировать огромное количество данных социальной сети, чтобы воспользоваться этим социальным феноменом. Анализ данных в социальных сетях - одна из самых горячих тем исследования в области интеллектуального анализа данных. Применение эффективных методов интеллектуального анализа данных позволяет пользователям находить ценные, точные и полезные данные из данных социальных сетей [Cortizo, 2009, Aggarwal, 2011, Russell, 2011 Adedoyin-Olowe, 2014].

Интеллектуальный анализ данных (Data Mining) – это мультидисциплинарная область, возникшая и развивающаяся на базе таких наук как прикладная статистика, распознавание образов, искусственный интеллект, теория баз данных и др.

Основная особенность интеллектуального анализа данных - это сочетание широкого математического инструментария (от классического статистического анализа до новых кибернетических методов) и последних достижений в сфере информационных технологий. К методам и алгоритмам интеллектуального анализа относятся следующие: искусственные нейронные сети, деревья решений, символьные правила, методы ближайшего соседа и k-ближайшего соседа, метод опорных векторов, байесовские сети, линейная регрессия, корреляционно-регрессионный анализ; иерархические методы кластерного анализа, неиерархические методы кластерного анализа, в том числе алгоритмы k-средних и k-медианы; методы поиска ассоциативных правил, в том числе алгоритм Apriori; метод ограниченного перебора, эволюционное программирование и генетические алгоритмы, разнообразные методы визуализации данных и множество других [Data mining]. Большинство аналитических методов, используемые в технологиях интеллектуального анализа – это известные математические алгоритмы и методы. Новым в их применении является возможность их использования при

решении тех или иных конкретных проблем, обусловленная появившимися возможностями технических и программных средств.

Наиболее распространенными задачами Data Mining являются классификация, кластеризация, объединение, прогнозирование и визуализация. Классификация является наиболее простой и общей задачей Data Mining. В результате решения проблемы классификации обнаруживаются признаки, характеризующие группы объектов исследуемого набора данных (классов). Новый объект может быть отнесен к тому или иному классу, основанному на этих признаках. Для решения проблемы классификации используются методы ближайшего соседа и k-ближайшего соседа, байесовские сети, индукция деревьев решений и нейронных сетей [Методы классификации].

Кластеризация является логическим продолжением идеи классификации. Эта задача более сложная, функция кластеризации состоит в том, что классы объектов изначально не predetermined. Результатом кластеризации является разбиение объектов на группы. В отличие от задач классификации, кластерный анализ не требует априорных предположений о наборе данных, не накладывает ограничений на представление объектов, он позволяет анализировать различные типы данных (интервальные, частотные, двоичные). Анализ кластеров позволяет уменьшить размер данных, сделать их видимыми. Методы кластерного анализа можно разделить на две группы: иерархические и неиерархические. Каждая группа включает в себя множество подходов и алгоритмов. [Методы кластерного анализа]

Суть иерархической кластеризации состоит в последовательном объединении меньших кластеров в крупные или разделении больших кластеров на более мелкие. Преимуществом иерархических методов кластеризации является их наглядность. Иерархические алгоритмы, связаны со строительством дендрограмм, которые являются результатом иерархического кластерного анализа. Дендрограмма описывает близость отдельных точек и кластеров друг к другу, это графическая последовательность слияния (разделения) кластеров. При большом числе наблюдений методы иерархического кластерного анализа не подходят. В таких случаях используются неиерархические методы, основанные на разделении, которые являются итерационными методами фрагментации исходного набора. Во время разделения новые кластеры формируются до тех пор, пока не будет выполнено правило остановки.

Одним из самых популярных методов анализа данных является анализ основных компонентов, который исходит из прикладного статистического анализа [Principal component analysis]. Это один из основных способов уменьшить размерность пространства наблюдения, потеряв наименьшее

количество информации. Он используется во многих областях, в том числе в эконометрике, биоинформатике, обработке изображений, сжатии данных, в социальных науках.

Применение фрактального анализа для АСС

Если наблюдать во времени динамику показателей, полученных с помощью АСС, то от статических данных мы перейдем к временным рядам, к которым применимы все инструменты анализа и прогнозирования временных рядов. В настоящее время стало общепризнанным, что многие временные ряды имеют долгосрочную зависимость и фрактальные свойства. Одними из первых реальных стохастических процессов, у которых были обнаружены самоподобные свойства, были информационные потоки данных в телекоммуникационных сетях. Существует большое количество публикаций, посвященных анализу самоподобных и мультифрактальных свойств трафика и их влияния на функционирование и качество обслуживания телекоммуникационной сети (см., например, [Шелухин, 2011]). Многочисленные исследования показали, что многие биоэлектрические сигналы обладают фрактальной структурой [Stanley, 1999]. Другим примером фрактальных стохастических структур являются современные финансовые рынки. Анализируя динамику возникновения участков с различной фрактальной структурой, можно диагностировать и прогнозировать нестабильные состояния (кризисы) рынка [Peters, 1996, Соловьев, 2015]. В последние годы появились исследования динамики сообществ в социальных группах, которые показывают, что соответствующие временные ряды обладают свойствами самоподобия [Cosoι, 2009, Yang, 2009, Rybski, 2012, Qingyun, 2016].

Самоподобие случайных процессов заключается в сохранении вероятностных характеристик при изменении масштаба времени. Стохастический процесс $X(t)$ является самоподобным с параметром H , если процесс $a^{-H}X(at)$ описывается теми же законами конечномерных распределений, что и $X(t)$. Параметр H , $0 < H < 1$, называемый показателем Херста, представляет собой степень самоподобия процесса. Наряду с этим свойством, показатель $H > 0.5$ характеризует меру долгосрочной зависимости стохастического процесса. Это означает, что если временной ряд в течении какого-то времени возрастал (убывал), то с вероятностью, близкой к показателю Херста, ряд сохранит эту тенденцию в течение аналогичного промежутка времени [Feder, 1988]. Начальные моменты самоподобного случайного процесса имеют скейлинговое соотношение $E[|X(t)|^q] \propto t^{qH}$, где величина t – интервал времени.

Мультифрактальные объекты являются статистически неоднородными самоподобными объектами. Для мультифрактальных временных рядов статистическая неоднородность объекта выражается в неоднородности распределения данных ряда, т.е. наличии тяжелых хвостов плотности распределения вероятностей временного ряда. В сравнении с самоподобными, мультифрактальные процессы проявляют более сложное скейлинговое поведение: для моментов мультифрактальных процессов выполняется отношение $E[|X(t)|^q] \propto t^{qh(q)}$, где $h(q)$ – обобщенный показатель Херста, являющийся в общем случае нелинейной функцией. Значение $h(q)$ при $q=2$ совпадает со значением степени самоподобия H . Для монофрактальных процессов обобщенный показатель Херста не зависит от параметра q : $h(q) = H$. [Reidi, 2002].

В работе был проведен сравнительный фрактальный анализ для двух групп в социальной сети Facebook, найденным по ключевым словам, связанных с киберугрозами. Для проведения исследований в сети были выбраны две группы ThreatPost и Threat Signal, каждая численностью 13980 и 84375 пользователей соответственно. Для каждой из этих социальных групп были сняты данные за последние пять лет о количестве лайков, комментариев и уровне вовлеченности (показатель уровня вовлеченности аудитории в активности группы измеряется в процентном соотношении действий пользователей к охвату аудитории).

На рис.1 слева показаны временные ряды ежедневного количества лайков для обеих групп. Фрактальный анализ показал, что временные ряды показателей этих групп обладают сильными мультифрактальными свойствами. На рис. 1 справа представлены значения обобщенного показателя Херста для этих рядов. Оба ряда обладают фрактальными свойствами, персистентностью ($H_1, H_2 > 0.5$), однако ряд лайков для группы Threat Signal обладает значительно большей статистической неоднородностью, что проявляется в значительно большем диапазоне величины $\Delta h(q)$.

Исследования показали, что ряды лайков достаточно сильно коррелируют с рядами уровня вовлеченности, поэтому их соответствующие мультифрактальные характеристики очень близки. Ряды количества комментариев (рис.2, слева), в отличие от рядов лайков, также имеют фрактальные свойства, но обладают достаточно близкой между собой мультифрактальной структурой (рис.2, справа).

Можно предположить, что механизм генерации комментариев, который формирует фрактальную структуру временного ряда, достаточно сильно отличается от механизма проставления лайков. Таким образом, проведенные в работе исследования подтвердили, что многие временные ряды показателей активности пользователей социальных сетей обладают фрактальными свойствами

и применение фрактального анализа позволяет выявить различия и обнаружить характерные черты динамики разных социальных групп.

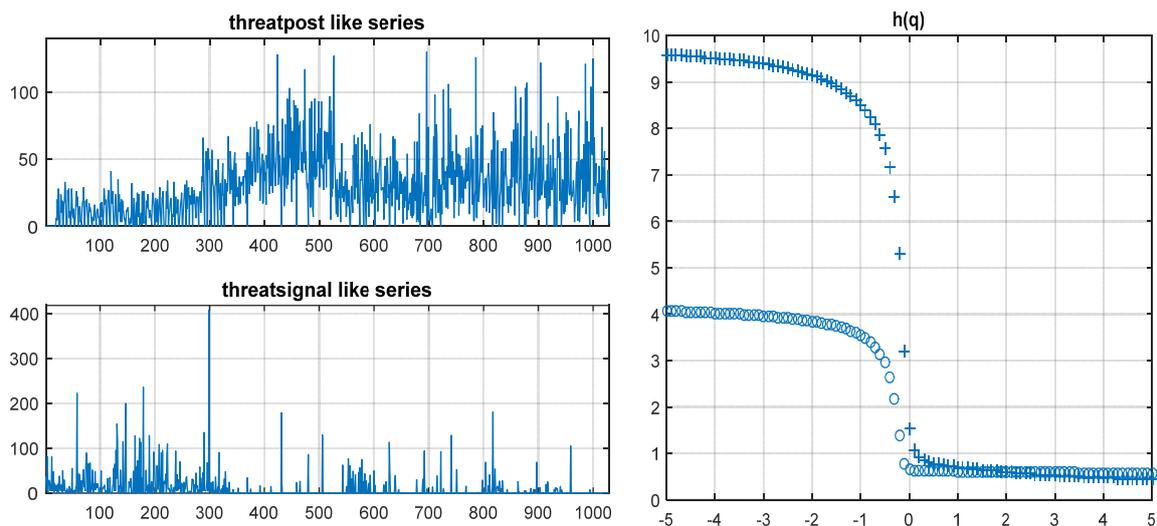


Рисунок 1. Слева: временные ряды лайков для групп; справа: соответствующие значения $h(q)$ (+ – группа Threat Post, o – Threat Signal)

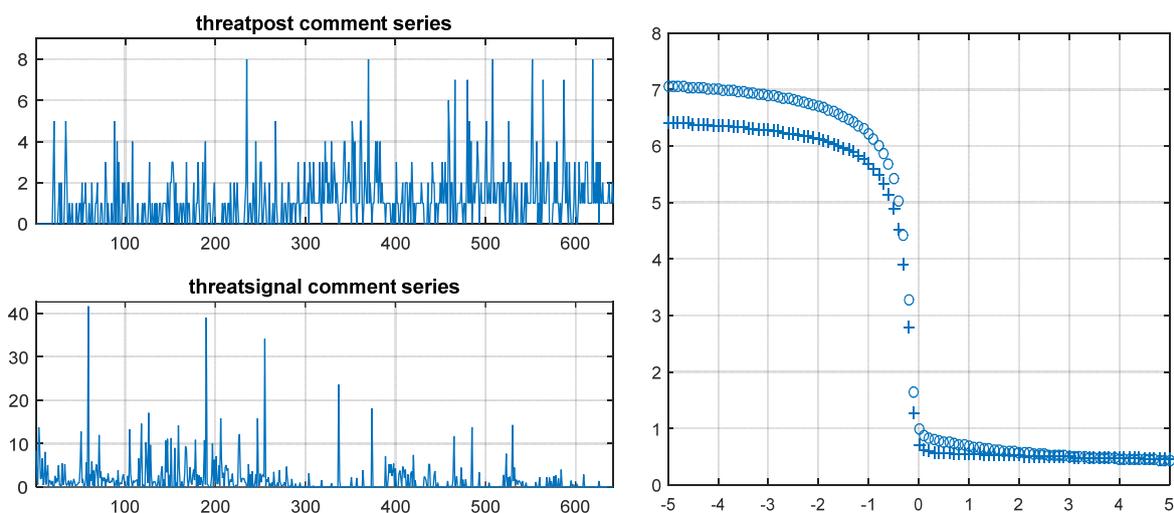


Рисунок 2. Слева: временные ряды комментариев для групп; справа: соответствующие значения $h(q)$ (+ – группа Threat Post, o – Threat Signal)

Выводы

В работе представлен обзор основных методов анализа социальных сетей, используемых для обеспечения информационной безопасности. Рассмотрены основные типы угроз и основные профилактические меры для безопасности социальных сетей. Представлены типичные задачи анализа социальных сетей, направленные на выявление киберугроз, такие как обнаружение сообществ в сети, обнаружение лидеров в сообществах, обнаружение экспертов в сетях, кластеризация текстовой информации на основе частотного анализа и другие. Кратко описаны основные методы и алгоритмы теории графов и интеллектуального анализа данных, которые используются в анализе социальных сетей. Поведенные исследования подтвердили, что многие ряды показателей активности пользователей сетей обладают фрактальными свойствами. Фрактальный анализ выявил различия в мультифрактальной структуре рядов различных показателей и социальных групп.

Bibliography

- [Adedoyin-Olowe, 2014] Mariam Adedoyin-Olowe, Mohamed Medhat Gaber, Frederic Stahl. A Survey of Data Mining Techniques for Social Network Analysis. *Journal of Data Mining & Digital Humanities*, 2014. pp.1-25.
- [Aggarwal, 2011] Aggarwal C. *Introduction to social network data analytics*. Springer US, 2011. doi: 10.1007/978-1-4419-8462-3_2
- [Aggarwal, 2014] Aggarwal C., Karthik S. *Evolutionary Network Analysis: A Survey* *ACM Computing Surveys*, Vol. 47, No. 1, Article 10, 2014.
- [Ant colony optimization algorithms]: https://en.wikipedia.org/wiki/Ant_colony_optimization_algorithms
- [Bonchi, 2011] Bonchi F., Castillo C., Gionis A., Jaimes A. *Social Network Analysis and Mining for Business Applications*, *ACM TIST*, vol. 2, iss. 3, 2011. pp.22–58.
- [Carley, 2002] Carley K., Lee J., Krackhardt D. *Destabilizing networks*. *Connections*, vol. 24, №.3, 2002. pp.79 – 92.
- [Carnes, 2007] Carnes T., Nagarajan R., Wild S.M., A. Van Zuylen. *Maximizing influence in a competitive social network: a follower's perspective* . *Proceedings of the ninth international conference on Electronic commerce*, Minneapolis, USA, 2007. 351–360.

- [Charu, 2012] Charu C. Social network data analytics. Springer Science & Business Media, 2012. 486 p.
- [Cortizo, 2009] Cortizo, J., Carrero, F., Gomez, J., Monsalve, B., Puertas, E. Introduction to Mining SM. Proceedings of the 1st International Workshop on Mining SM, 2009. pp.1 – 3.
- [Coscia, 2011] Coscia M., Giannotti F., Pedreschi D. A classification for community discovery methods in complex networks. Statistical Analysis and Data Mining, 2011. pp.512–546
- [Cosoi, 2009] Alexandru Catalin Cosoi, Carmen Maria Cosoi. A fractal approach to social network spam detection: <https://www.virusbulletin.com/conference/vb2009/abstracts/fractal-approach-social-network-spam-detection/>
- [Data mining]: https://en.wikipedia.org/wiki/Data_mining .
- [Dodds, 2005] Dodds P.S., Watts D.J. A generalized model of social and biological contagion. Journal of theoretical biology, iss. 4, 2005. pp.587–604.
- [Easley, 2010] Easley D., Kleinberg J. Networks, Crowds, and Markets: Reasoning about a Highly Connected World, Cambridge University Press, 2010. 819 p.
- [Ehrlich, 2005] Ehrlich K., Carboni I. Inside Social Network Analysis IBM Watson Research Center. New York, USA, Technical Report, 2005. pp.5–10.
- [Feder, 1988] J. Feder. Fractals. Plenum, New York, 1988.
- [Flake, 2002] Flake G.W., Lawrence S., Giles C.L., Coetzee F.M. Self-organization and identification of Web communities. Computer. , Iss. 3, 2002. pp.66–70.
- [Fortunato, 2010] Fortunato S. Community detection in graphs. Physics Reports, Vol. 486, Iss. 3–5, 2010. pp.75–174.
- [Girvan, 2002] Girvan M., Newman M.E. Community structure in social and biological networks, Proceedings of the National Academy of Sciences of the United States of America, Iss. 12, 2002. pp.7821–7826
- [Goyal, 2008] Goyal A., Bonchi F., Laks V.S. Lakshmanan. Discovering leaders from community actions. Proceedings of the 17th ACM Conference on Information and Knowledge Management, Napa Valley, California, USA, 2008. pp.499–508.
- [Gyöngyi, 2004] Gyöngyi Z., Garcia-Molina H., Pedersen J. Combating Web Spam with TrustRank. Proceedings of the International Conference on Very Large Data Bases, Vol. 30, 2004. pp.576-587.

- [Hanneman, 2005] Hanneman R., Riddle M. Introduction to social network methods. Riverside, CA: University of California. 2005. <http://faculty.ucr.edu/~hanneman/nettext> .
- [Hasan, 2011] Mohammad Al Hasan, Mohammed J. Zaki. A Survey of Link Prediction in Social Networks. *Social Network Data Analytics*, 2011. pp.243-275.
- [Johnson, 1994] Johnson J., Ironsmith M. Assessing Children's Sociometric Status: Issues and the Application of Social Network Analysis. *Journal of Group Psychotherapy, Psychodrama & Sociometry*, vol. 47(1), 1994. pp. 36–49.
- [Kempe, 2003] Kempe D., Kleinberg J., Tardos É. Maximizing the spread of influence through a social network. *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, Washington, USA, 2003. pp.137–146.
- [Kolomeychenko, 2014] M. I. Kolomeychenko, A. A. Chepovskiy, A. M. Chepovskiy. An algorithm for detecting communities in social networks, *Fundamentalnaya i prikladnaya matematika*, Vol. 19, No. 1, 2014. pp. 21—32.
- [Koren, 2003] Koren Y. On Spectral Graph Drawing . *Proceedings of the 9th International Computing and Combinatorics Conference*, Springer, 2003. pp. 496–508
- [Kumar, 2004] Kumar R., Novak J., Raghavan P., Tomkins A. Structure and Evolution of Blogspace *Commun. ACM*, vol. 47, No. 12, 2004. pp.35–39.
- [Kushwah, 2016] Ajay Kumar Singh Kushwah, Amit Kumar Manjhvar. A Review on Link Prediction in Social Network *International Journal of Grid and Distributed Computing*, vol. 9, No. 2, 2016. pp.43-50.
- [Langville, 2006] Langville Amy N., Meyer Carl D. *Google's PageRank and Beyond: The Science of Search Engine Rankings*. Princeton University Press, 2006. 224 p.
- [Leskovec, 2005] Leskovec J., Kleinberg J., Faloutsos C. Graphs over time: densification laws, shrinking diameters and possible explanations. *Proc. 11th ACM SIGKDD Intern. Conf. on Knowledge Discovery in Data Mining*, NY, 2005. pp.177–187.
- [Liben-Nowell, 2003] Liben-Nowell D., Kleinberg J. The Link Prediction Problem for Social Networks. *Proceedings of the 12th International Conference on Information and Knowledge Management*, N. Y.: ACM Press, 2003. pp. 556–559.
- [List of social networking websites, 2017] List of social networking websites, 2017. URL: https://en.wikipedia.org/wiki/List_of_social_networking_websites

- [Mikolov, 2013 (1)] Mikolov T., Sutskever I., Chen K., Corrado G., Dean J. Distributed Representations of Words and Phrases and their Compositionality, 2013. <https://papers.nips.cc/paper/5021-distributed-representations-of-words-and-phrases-and-their-compositionality.pdf>
- [Mikolov, 2013] Mikolov T., Chen K., Corrado G., Dean J. Efficient Estimation of Word Representations in Vector Space, 2013. <https://arxiv.org/pdf/1301.3781.pdf>
- [National Cyber Security Strategy 2016-2021 for United Kingdom]: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/564268/national_cyber_security_strategy.pdf.
- [National cyber security strategy for Ukraine]: <http://www.president.gov.ua/documents/962016-19836>
- [Peters, 1996] Edgar E. Peters. Chaos and Order in the Capital Markets: A New View of Cycles, Prices, and Market Volatility. Edgar E. Peters. Wiley, 2 edition, 1996.
- [Principal component analysis]: https://en.wikipedia.org/wiki/Principal_component_analysis.
- [Qingyun, 2016] Qingyun Liu, Xiaohan Zhao, Walter Willinger, Xiao Wang, Ben Y. Zhao, Haitao Zheng Self-Similarity in Social Network Dynamics ACM Trans. Model. Perform. Eval. Comput. Syst., Vol. 2, No. 1, Article 5, 2016.
- [Ramos, 2003] Ramos J. Using tf-idf to determine word relevance in document queries. Proceedings of the first in-structional conference on machine learning, 2003.
- [Russell, 2011] Matthew A. Russell. Mining the Social Web: Analyzing Data from Facebook, Twitter, LinkedIn, and Other Social Media Sites. O'Reilly, 2011. 332 p.
- [Rybski, 2012] Diego Rybski, Sergey V. Buldyrev, Shlomo Havlin, Fredrik Liljeros, Herna A. Makse Communication activity in a social network: relation between long-term correlations and inter-event clustering Scientific reports, 2012, 2 : 560 | DOI: 10.1038/srep00560
- [Saoussen, 2014] Saoussen Aouay, Salma Jamoussi, Faiez Gargouri, Ajith Abraham. Modeling Dynamics of Social Networks: A Survey. Sixth International Conference on Computational Aspects of Social Networks (CASoN), 2014. pp.49-53.
- [Shantanu, 2017] Shantanu Ghosh. Top seven social media threats, 2017. URL: <http://www.computerweekly.com/tip/Top-seven-social-media-threats>
- [Stanley, 1999] H. E. Stanley, L. A.N. Amaral, A.L. Goldberger, S. Havlin, P.Ch. Ivanov, C.-K. Peng. Statistical Physics and Physiology: Monofractal and multifractal approaches. Physica A 270, 1999. pp. 309-324.

- [Stohl, 2007] Stohl C., Stohl M. Networks of Terror: Theoretical Assumptions and Pragmatic Consequences. *Communication Theory*, Vol. 17, 2007. pp.93 – 124.
- [Tantipathananandh,2007] Tantipathananandh C., Berger-Wolf T., Kempe D. A framework for community identification in dynamic social networks .*Proc. 13th ACM SIGKDD Intern. Conf. on Knowledge Discovery and Data Mining*, NY, 2007. pp.717–726
- [The NATO Cooperative Cyber Defence Centre]: <https://ccdcoe.org/cyber-security-strategy-documents.html>.
- [Travelling salesman problem]: https://en.wikipedia.org/wiki/Travelling_salesman_problem.
- [Wang, 2014] Wang H. Introduction to Word2vec and its application to find predominant word senses, 2014. <http://compling.hss.ntu.edu.sg/courses/hg7017/pdf/word2vec%20and%20its%20application%20to%20wsd.pdf>.
- [Watts, 2007] Watts D.J., Dodds P.S. Influentials, Networks, and Public Opinion Formation. *Journal of consumer research*, iss. 4,2007. pp.441–458.
- [Yang, 2009] Christopher C. Yang, Marc Sageman Analysis of terrorist social networks with fractal views *Journal of Information Science*, V. 35, Issue 3, 2009. pp. 299–320.
- [Yu, 2014]Yu M., Dredze M. Improving lexical embeddings with semantic knowledge. *Association for Computational Linguistics (ACL)*, 2014. pp.545-550.
- [Батура, 2013] Батура Т.В. Модели и методы анализа компьютерных социальных сетей. Программные продукты и системы. - №3 (103), 2013. С.130-137.
- [Бузун, 2012] Бузун Н., Коршунов А. Выявление пересекающихся сообществ в социальных сетях М.: Институт системного программирования РАН, 2012. 18 с. <http://www.twirpx.com/file/987306>
- [Горбулін, 2009] Горбулін В. П., Додонов О. Г., Ланде Д. В. Інформаційні операції та безпека суспільства: загрози, протидія, моделювання: монографія. Інтертехнологія, 2009. – 164 с.
- [Григорьев, 2007] Григорьев А. Н., Ландэ Д. В., Бороденков С. А., Мазуркевич Р. В., Пацьора В. Н. InfoStream. Мониторинг новостей из Интернет: технология, система, сервис. К.: ООО “Старт-98”, 2007. 40 с.
- [Гриненко, 2012] Гриненко І, Прокофьева – Янчиленко Д. Вплив віртуальних спільнот на інформаційну безпеку: сучасний стан та тенденції розвитку. Правове, нормативне та метрологічне забезпечення систем захисту інформації в Україні, № 1 (23), 2012. С. 18 – 23.

- [Губанов, 2010] Губанов Д., Новиков Д., Чартишвили А. Социальные сети: модели информационного влияния, управления и противоборства. Москва, 2010. 225 с.
- [Дзюндзюк, 2011] Дзюндзюк В. Б. Віртуальні співтовариства: потенційна загроза для національної безпеки / В. Б. Дзюндзюк // Державне будівництво [Електронне видання]. № 1, 2011. Режим доступу до журн. : [http:// www.kbuara.kharkov.ua](http://www.kbuara.kharkov.ua). – Назва з екрана
- [Додонов, 2009] Додонов О. Г., Ланде Д. В., Путятін В. Г. Інформаційні потоки в глобальних комп'ютерних мережах. Київ: Наукова думка, 2009. 295 с.
- [Додонов, 2013] Додонов А. Г., Ландэ Д. В., Прищепа В. В., Путятин В. Г. Конкурентная разведка в компьютерных сетях. Киев: ИПРИ НАН Украины, 2013. 248 с.
- [Додонов, 2014] Додонов А. Г., Ландэ Д. В., Путятин В. Г. Компьютерные сети и аналитические исследования. Киев: ИПРИ НАН Украины, 2014. 486 с.
- [Докука, 2015] Докука С. В., Валева Д. Р. Статистические модели для анализа динамики социальных сетей в исследованиях образования. Вопросы образования, № 1, 2015. С.201-213
- [Матвиенко, 2011] Матвиенко Ю. А. Деструктивные сетевые социальные структуры как средство информационной войны и угроза безопасности России. Информационно-аналитический портал «Геополитика», 2011. [Электронный ресурс]. Режим доступу до журн.: <http://old.geopolitica.ru/Articles/1218>. – Загл. с екрана.
- [Методы классификации]: <http://www.intuit.ru/studies/courses/6/6/lecture/176>.
- [Методы кластерного анализа]: <http://www.intuit.ru/studies/courses/6/6/lecture/182>.
- [Пелещишин, 2010] Пелещишин О. П. Інформаційні технології обліку та пошуку онлайн-спільнот у задачі соціального маркетингу. Вісник Національного університету "Львівська політехніка". Серія економічна, № 44, 2010. С. 50 – 59.
- [Пелещишин, 2012] Пелещишин А. М., Серов Ю. О., Березко О. Л., Пелещишин О. П., Тимовчак-Максимець О. Ю., Марковець О. В. Процеси управління інтерактивними соціальними комунікаціями в умовах розвитку інформаційного суспільства: монографія. Львів: Видавництво Львівської політехніки, 2012. 368 с.
- [Пелещишин, 2013] Пелещишин О. П. Аналіз та протидія загрозам маркетинговій позиції підприємства в онлайн-спільнотах. Захист інформації, № 3 (15), 2013. С. 217 – 224.

- [Слабченко, 2013] Слабченко О. О., Сидоренко В. Н., Пономарчук Р. А.. Методы и алгоритмы выявления сообществ потенциальных абитуриентов и их лидеров в социальных сетях. Вестник Кременчугского национального университета, №1 (78), 2013. С.53-61.
- [Смирнов, 2014] Смирнов А. И., Григорьев В. Р., Кохтюлин И. Н., Куроедов Б. В., Сандаров О. В. Глобальная безопасность в цифровую эпоху: стратегемы для России. М.: ВНИИ геосистем, 2014. 394 с.
- [Соловйов, 2015] Соловйов В.М. Мережні міри складності соціально-економічних систем. Вісник Черкаського університету, сер. «Прикладна математика. Інформатика», № 38 (371), 2015. с.67-79.
- [Тимовчак-Максимець, 2010] Тимовчак-Максимець О. Ю. Методи використання розширених можливостей глобальних пошукових систем в задачі пошуку споживацького досвіду в онлайн середовищах. Вісник Національного університету "Львівська політехніка": Інформаційні системи та мережі, № 689, 2010. с. 323 – 331.
- [Укустов, 2012] Укустов С. С., Кравец А. Г. Подход к решению задачи идентификации влиятельных разработчиков в социальной сети гитхаб. Известия Волгоградского Государственного Технического Университета, Выпуск № 15 (102), 2012. с.61-66.
- [Чураков, 2001] Чураков А. Н. Анализ социальных сетей. Социальные исследования, № 1, 2001. С. 109–121.
- [Шелухин, 2011] Шелухин О. И. Мультифракталы. Инфокоммуникационные приложения. М.: Горячая Линия, Телеком, 2011. 578 с.

Информация об авторах



Людмила Кириченко – д.т.н., профессор Харьковського національного університету радіоелектроніки; пр. Науки 14, 61166, Харків, Україна; e-mail: lyudmyla.kirichenko@nure.ua.

Основные области научных исследований: самоподобные и мультифрактальные временные ряды, фрактальный анализ, вейвлет-анализ, детерминированные хаотические системы.



Алексей Барановский - к.т.н., старший преподаватель Национального технического университета Украины "Киевский политехнический институт"; пр. Победы 37, 03056, Киев; email: o.baranovskyi@kpi.ua .

Основные области научных исследований: фрактальный анализ, анализ рекуррентных диаграмм, теория безопасности, информационная безопасность, информационные операции и воздействия.



Тамара Радивилова – к.т.н., доцент Харьковского национального университета радиоэлектроники; пр. Науки 14, 61166, Харьков, Украина; e-mail: tamara.radivilova@gmail.com.

Основные области научных исследований: самоподобные и мультифрактальные временные ряды, телекоммуникационные системы, управление трафиком, информационная безопасность

Detecting cyber threats through social network analysis

Lyudmyla Kirichenko, Tamara Radivilova, Oleksii Baranovskyi

Abstract: *The paper reviews the basic methods of social networks analysis which are used to detect cyber threats. The main types of threats in social networks are presented and some methods for their prevention are described. Typical tasks of social networks analysis aimed at identifying cyber threats, such as community detection in networks, detection of leaders and experts in communities, stability analysis of community, clustering text information, etc. are considered. The basic classes of methods of graph theory and data mining, which are widely used in the analysis of social networks are described. The application of fractal analysis methods to study the series of network users activity indicators is shown.*

Keywords: *analysis of social networks, fractal analysis, cyber threats, data mining, detection leader methods, detection experts methods.*

СОПОСТАВЛЕНИЕ СЕМАНТИЧЕСКИХ ИНФОРМАЦИОННЫХ РЕСУРСОВ WEB НА ОСНОВЕ ОНТОЛОГИЧЕСКОГО АНАЛИЗА

Юлия Рогушина

Аннотация: В работе проанализированы современные средства представления распределенных знаний. Рассмотрены проблемы, возникающие в процессе управления такими знаниями в интеллектуальных информационных системах, ориентированных на работу в Web. Обоснована необходимость использования онтологий и актуальность развития семантических технологий, направленных на их поддержку, в частности, стандарты и языки Semantic Web. Рассматривается задача сопоставления онтологий, результаты такого сопоставления и частный случай этой задачи, когда сравниваются онтологии, являющиеся развитием одной и той же начальной онтологии. Рассмотрены семантические Wiki-ресурсы, как источник информации для пополнения и усовершенствования онтологий предметных областей, отображаемых с помощью набора Wiki-страниц. Проведен анализ компонентов онтологий предметных областей и семантической разметки Wiki-ресурсов (на примере Semantic MediaWiki), предложен метод их автоматизированного сопоставления для усовершенствования баз знаний распределенных интеллектуальных систем.

Ключевые слова: онтология предметной области, автоматическая генерация онтологий, онтология, Wiki-технологии.

ITHEA Keywords: I.2.4 Knowledge Representation Formalisms and Methods.

Введение

Использование распределенных знаний является одним из определяющих факторов эффективности современных информационных систем. Это в свою очередь обуславливает актуальность исследований, направленных на приобретение и накопление знаний, решение проблем распознавание, логического вывода новых знаний для поддержки принятия решений. Интеллектуальные информационные системы (ИИС), которые работают в открытом распределенном информационном пространстве, требуют постоянного пополнения и

обновления знаний, которые поступают из внешней среды, так как в ИИС использование знаний о предметной области (ПрО) и методы их обработки играют решающую роль и определяют полезность получаемых результатов.

Из-за сложности получения знаний важное значение приобретает проблема обеспечения их интероперабельности и повторного использования.

В настоящее время для интероперабельного представления знаний в ИИС широко применяются онтологии, которые обеспечивают повторное использование полученных знаний в различных приложениях [Gruber, 1991]. Онтологии базируются на фундаментальном теоретическом базисе дескриптивных логик, для них уже существуют общепринятые стандарты описания, языки и программные средства. Особый интерес представляет применение онтологического анализа как основы для обработки распределенных знаний. В связи с этим значительное количество исследований связано с теоретическим базисом онтологий, их построением, усовершенствованием, с получением знаний из онтологических структур, а также с другими важными аспектами менеджмента онтологий, которые значительно различаются по цели и назначению такого анализа [Uschold, 1996]. Одним из важных направлений таких исследований является интеграция онтологий с другими информационными ресурсами Web. Такие свойства Web, как гетерогенность и динамичность, обуславливают ряд проблем, связанных с пополнением, использованием и оценкой онтологий.

Управление знаниями в Web

Проблема управления знаниями, которые являются составляющей различных ИИС, связана с интеллектуальным анализом данных и распознаванием свойств информационных объектов (ИО), которые используются и обрабатываются в предметной области, интересующей пользователя. В связи с тем, что современные ИИС используют и генерируют большие по объему и сложные по структуре совокупности знаний, возникает проблема разработки эффективных методов управления ими [Holtshouse, 2013].

Понятие «управление знаниями» («knowledge management») было введено американским ученым Карлом Виигом по аналогии к таким терминами, как «управление данными» и «управление информацией». Цель управления знаниями заключается в том, чтобы преобразовать процесс создания баз знаний из искусства в инженерную дисциплину. Это требует разработки соответствующих методов, языков и инструментов [Studer, 1998], в частности, связанных с построением и использованием онтологий. *Управление знаниями*

представляет собой совокупность процессов, связанных с эффективным созданием, сохранением, распространением и использованием знаний для достижения поставленных целей.

В процессе управления знаниями в Web необходимо решать такие проблемы, как:

- интеграция гетерогенных знаний, полученных из разных источников (например, интеграция онтологий или сопоставление различных видов семантической разметки, поиск по метаописаниям);
- поиск противоречий между знаниями, оценка их достоверности и надежности (например, сравнение знаний, извлеченных из контента различных ИР, сравнение двух онтологий);
- получение новых знаний из уже имеющихся (например, с помощью дедуктивного, традиционного или индуктивного логического вывода) и их представление в форме, понятной пользователю (например, визуализация онтологических структур или семантических сетей);
- поиск знаний, необходимых конкретному пользователю (например, персонализированный поиск, учитывающий знания о пользователе и его задаче, или семантический поиск сложных информационных объектов и сервисов);
- автоматизация создания метаданных, которые корректно отображают контент ИР (например, семантическая разметка ИР, – как текстовых, так и мультимедийных) на уровне содержания, и эффективный поиск в таких метаописаниях.

Любая ИИС может поддерживать определенное подмножество функций управления знаниями. Для Web-ориентированных ИИС эта задача усложняется тем, что нужно осуществлять управление распределенными и гетерогенными знаниями [Гладун, 2016]. Следует учитывать динамичность и гетерогенность среды Web: динамичность требует постоянного обновления знаний о ПрО, которые необходимо извлекать из открытых информационных ресурсов (ИР), а гетерогенность – интеграции полученной информации и ее проверки на непротиворечивость. Таким образом, возникает необходимость в средствах пополнения и усовершенствования онтологий путем анализа ИР. Не менее важная задача – обнаружить различия в онтологиях, описывающих одну и ту же ПрО. Это позволяет формализованно описать несоответствия в подходах различных пользователей, и, если надо, разрешить их.

В отличие от традиционных ИИС, те системы, которые ориентированы на работу в Web, для своего функционирования нуждаются в получении информации из внешних ИР. Доступ к некоторым из них ИИС может получать непосредственно, но в большинстве случаев в качестве

посредников используются внешние информационно-поисковые системы (ИПС). Однако внешние ИПС позволяют лишь построить множество ИП, из которых можно получить информацию, и не решают проблему извлечения нужных знаний из этих ресурсов [Гладун, 2013]. Одним из наиболее известных направлений в исследованиях управления знаниями в Web является проект Semantic Web, предложенный Тимом Бернес-Ли. Основная идея этого проекта заключается в преобразовании Web в глобальную базу знаний и разработке соответствующих стандартов, языков и программных средств для обработки информации в этих ресурсах на семантическом уровне [Berners-Lee, 2006].

Анализ современных исследований в области управления знаниями показывает, что, несмотря на значительные достижения в сфере Data Mining и Text Mining, извлечение знаний из естественно-языковых неструктурированных ИП остается крайне трудоемкой и полностью не автоматизируемой задачей. Значительно более эффективно осуществляется извлечение знаний из ИП, содержащих семантическую разметку, которую можно сопоставить с теми или иными структурами формализованного представления знаний.

Для достижения практически полезных результатов необходимо ориентироваться на те средства семантической разметки, которые уже сегодня получили достаточно широкое распространение, имеют достаточно высокую выразительную мощность, надежную реализацию и удобный пользовательский интерфейс. Этим требованиям удовлетворяют семантические Wiki, в частности, Semantic MediaWiki [Völkel, 2006]: эта платформа сегодня широко используется, сформированные на ее основе ИП часто обновляются и быстро растут. Использование таких источников обуславливается тем, что Wiki-ресурсы динамически обновляются всем сообществом пользователей, имеют четко определённую и простую для понимания структуру и обеспечивают обработку информации на семантическом уровне, и, таким образом, обеспечивают технологическую платформу для группового управления знаниями [Majchrzak, 2006].

Встроенные средства Semantic MediaWiki позволяют строить онтологии, представленные на языке OWL, но эти средства недостаточно гибки для того, чтобы построенную с их помощью онтологию можно было интероперабельно применять в других ИИС. Поэтому возникает необходимость в разработке более совершенных методов формирования и пополнения онтологий на основе Wiki-ресурсов, а также сравнения онтологий, сформированных и усовершенствованных на основе различных состояний Wiki-ресурсов.

Постановка задачи

Чтобы обеспечить интеллектуальные информационные системы возможностью импортировать, модифицировать и обновлять онтологические знания, извлекаемые из внешних семантически размеченных Wiki-ресурсов, необходимо разработать методы и средства, обеспечивающие решение следующих задач:

- построить взаимно-однозначное соответствия между основными элементами онтологии ПрО (классы и подклассы, экземпляры, свойства объектов и данных, их значение) и Semantic MediaWiki (категории, семантические свойства, их значения и связи);
- разработать средства и критерии сопоставления онтологий, являющихся усовершенствованиями одной и той же онтологии ПрО, пополненной знаниями из различных Wiki-ресурсов, семантически размеченных элементами начальной онтологии.

На первом этапе необходимо построить (или выбрать) начальную онтологическую модель ПрО, которая будет использоваться для семантической разметки. Для этого следует разработать метод преобразования элементов онтологии в конструкции Semantic MediaWiki (категории и семантические свойства). На следующем этапе необходимо обеспечить возможность уточнения и усовершенствования начальной онтологии ПрО на основе анализа семантически размеченных Wiki-ресурсов, то есть разработать метод преобразования конструкций Semantic MediaWiki в онтологию, представленную на языке OWL. Затем следует проанализировать полученную онтологию, оценить ее свойства и соответствие представлениям пользователя о ПрО. Итеративное повторение этих действий должно обеспечить формирование адекватной онтологии ПрО, которую можно использовать в различных ИИС в качестве базы знаний.

Основные семантические компоненты Semantic MediaWiki

Semantic MediaWiki (SMW) позволяет пользователям добавлять семантические аннотации к Wiki-страницам, используя дополнительные элементы разметки, превращая MediaWiki в семантический ресурс: категории, семантические свойства и запросы. SMW использует концепцию семантических свойств (для создания данных) и семантических запросов (для использования данных). Пользователи размечают статьи категориями и свойствами, чтобы информация становилась доступной для запросов. Формальная модель Wiki-ресурса состоит из следующих элементов [Rogushina, 2016]:

- множества Wiki-страниц $P = P_{user} \cup P_{categ} \cup P_{template} \cup P_{spec}$, где P_{user} – множество страниц, созданных пользователями, P_{categ} – множество страниц, которые описывают категории, $P_{template}$ – множество страниц, которые описывают шаблоны, P_{spec} – множество других специальных страниц;
- $L = \{"link"\}$ множество из одного элемента, который описывает ссылку одной Wiki-страницы этого ресурса на другую (хотя в Wiki-ресурсах предусмотрены и ссылки на другие виды страниц, в рамках данной модели они не учитываются).

Семантические свойства обеспечивают привязывание данных к Wiki-страницам. Каждое свойство имеет тип, название и значение, кроме того, ему соответствует отдельная Wiki-страница в специальном пространстве имен, которая позволяет задавать тип свойства, определять его положение в иерархии свойств, а также документировать его использование. Свойства, в отличие от категорий MediaWiki, имеют не только название, но и значение. В текст Wiki-страницы семантическое свойство вставляется в формате [[Имя свойства::Значение свойства]]. Семантические свойства используются для того, чтобы определить смысловую нагрузку ссылок на другие страницы.

Формальная модель семантически обогащенных Wiki-ресурсов является более сложной и включает ряд элементов, связанных с семантическими свойствами [Rogushina, 2016]:

$$W_s = \langle P = P_{user} \cup P_{categ} \cup P_{template} \cup P_{sem_prop} \cup P_{spec}, L = \{"link"\} \cup L_{sem_prop} = \{i\} \rangle \quad (1)$$

где P_{sem_prop} – множество страниц, которые описывают семантические свойства Wiki-страниц, некоторые из которых являются семантически определенными ссылками на другие Wiki-страницы: $P_{sem_prop_page} \subseteq P_{sem_prop}$, а другие связывают страницы со значениями различных типов данных (эти типы данных определяются на страницах семантических свойств).

Основные семантические компоненты онтологий

Онтологический анализ позволяет превращать описание определенного представления о внешнем мире в набор терминов и правил их использования, пригодных для машинной обработки. Онтология представляет собой явную спецификацию концептуализации. Онтологический подход позволяет анализировать окружающий мир независимо от формы представления знаний о нем [Guarino, 1995]. Онтологию можно рассматривать как базу знаний (БЗ) специального вида с семантической информацией об определенной ПрО. Компоненты, из которых складываются конкретные онтологии, зависят от парадигмы представления, но практически все модели онтологий содержат определенные концепты (понятие, классы), свойства концептов (атрибуты, роли), отношение между концептами (зависимости, функции) и ограничения использования, которые определяются аксиомами. В общем случае формальная модель онтологии O – это упорядоченная тройка $O = \langle T, R, F \rangle$, где T – множество понятий ПрО; R – множество отношений между ними; F – множество функций интерпретации понятий и отношений. Эта формальная модель может быть конкретизирована в зависимости от назначения и сферы применения онтологии.

Сейчас в ИИС для представления онтологий чаще всего используются различные диалекты языка OWL. Онтология OWL (Web Ontology Language) – это последовательностью аксиом и фактов, а также ссылок на другие онтологии.

Фундаментальные понятия определенной ПрО должны соответствовать *классам* онтологии. Каждый экземпляр в онтологии OWL принадлежит к классу owl:Thing, а каждый созданный пользователем класс автоматически является подклассом owl:Thing. Специфичные для ПрО корневые классы определяются простым объявлением именованного класса. OWL также позволяет задать пустой класс: owl:Nothing. Определения могут быть расширяемыми и распределенными. Фундаментальным таксономическим конструктором для классов является rdfs:subClassOf. Он связывает отдельный класс с общим. Если X – подкласс Y , то каждый представитель X – также представитель Y . Отношение rdfs:subClassOf является транзитивным. Если X – подкласс Y и Y – подкласс Z , тогда X – подкласс Z .

Определение класса состоит из двух частей: названия (или ссылки на него) и списка ограничений. Каждое выражение, которое непосредственно помещается в определении класса, уточняет свойства представителей этого класса. Представители класса принадлежат пересечению указанных ограничений. Для определения экземпляра достаточно объявить его членом какого-либо класса.

Свойства дают возможность утверждать общие факты о членах классов и об экземплярах. Они представляют собой бинарные отношения. Различают отношения между представителями классов и RDF-литералами или типами данных, определенными XML Schema, и отношения между представителями классов.

При определении свойства существует много способов ограничить это отношение. Можно задать его домен и диапазон. Свойство может быть определено как специализация (подсвойство) уже существующего свойства. Возможны и более сложные ограничения. Свойства, также как и классы, могут быть организованы иерархически. Формальная семантика OWL содержит описание того, как получить логические следствия, имея такую онтологию, то есть получить факты, которые не представлены в ней непосредственно, но тем не менее обусловлены ее семантикой.

Семантику онтологических языков обычно раскрывают через теорию моделей. В частности, она определяет функцию интерпретации, которая проецирует каждый элемент онтологии на множество D , которое называют областью интерпретации.

Формальная онтология позволяет анализировать окружающий мир независимо от формы представления знаний о нем. Формальные модели онтологий позволяют отображать различные метасвойства их элементов онтологий, такие как идентичность (*identity*), стойкость (*rigidity*), согласованность (*unity*) и зависимость (*dependence*), обеспечивая возможность онтологического концептуального анализа ПрО [Guarino, 2000]. Анализ различных средств представления онтологий и их формальных моделей, которые предлагают технологии Semantic Web, показывает, что они значительно отличаются своими выразительными возможностями и своей сложностью: RDF Schemas предлагает простейший уровень для представления онтологий, а OWL Full – наиболее сложный. Выбор средства представления онтологии зависит от специфики проблемы, для которой она разрабатывается. Например, для семантического поиска обычно достаточно простейшего уровня представления знаний, тем не менее некоторые задачи, связанные с нахождением совокупности ИО со сложной структурой и многочисленными условиями, требуют более сложных возможностей для отображения знаний соответствующих ПрО.

Будем использовать формальную модель

$$O = \langle X = X_{cl} \cup X_{ind}, R = r_0 \cup \{r_i\} \cup \{p_j\}, F, T \rangle \quad (2)$$

Эта модель состоит из таких элементов:

- $X = X_{cl} \cup X_{ind}$ – множество концептов онтологии, где X_{cl} – множество классов, X_{ind} – множество экземпляров классов, таких, что $\forall a \in X_{ind} \exists A \in X_{cl}, a \in A$;
- $R = r_{ier_cl} \cup \{r_i\} \cup \{p_j\}$ – множество отношений между элементами онтологии, где r_{ier_cl} – иерархическое отношение, которое может устанавливаться между классами онтологии и свойствами классов и характеризуется такими свойствами, как антисимметричность и транзитивность, $r_{ier_cl} : X_{cl} \rightarrow X_{cl}$; $\{r_i\}$ – множество объектных свойств, которые устанавливают отношение между экземплярами классов: $r_i(a, a \in X_{ind}) = b, b \in X_{ind}$, $r_i : X_{ind} \rightarrow X_{ind}$; $\{p_j\}$ – множество свойств данных, которые устанавливают отношения между экземплярами классов и значениями: $p_i(a, a \in X_{ind}) = t, t \in T$, $p_i : X_{ind} \rightarrow Const$, причем внутри множеств объектных свойств и свойств отношений также могут существовать иерархические отношения $r_{ier_obj} : \{r_i\} \rightarrow \{r_i\}$ и $r_{ier_data} : \{p_j\} \rightarrow \{p_j\}$;
- F – множество свойств классов онтологии, экземпляров классов и их свойств, которые могут применяться для логического вывода (например, эквивалентность, отличие);
- T – множество типов данных (например, строка, целое).

Выбор именно такой модели онтологии обуславливается следующими причинами. Во-первых, она наиболее полно соответствует интуитивному представлению об онтологиях, заложенному в пользовательском интерфейсе редактора онтологий Protégé [Tudorache, 2013]. Во-вторых, эта модель достаточно легко интегрируется с моделью тезауруса задачи, который используется для семантического поиска. В-третьих, эту модель позволяет сопоставить онтологическое представление знаний о ПрО с семантическими конструкциями Semantic MediaWiki: для некоторых элементов может быть сформировано взаимно-однозначное соответствие, а сопоставление других требует дополнительных преобразований, но также может быть хотя бы частично автоматизировано.

Рассмотрим подробнее эти соответствия. Для этого надо проанализировать основные семантические элементы Semantic MediaWiki. Следует отметить, что, хотя в Semantic MediaWiki предусмотрено автоматическое построение онтологий по семантически размеченным Wiki-страницам, но в текущих версиях это построение учитывает слишком мало семантических параметров, а его результаты мало пригодны для дальнейшего использования.

Важным аспектом предложенной модели является возможность ее интеграции с формальной моделью тезауруса задачи, который используется для семантического поиска.

Применение онтологического анализа в семантической разметке

При обработке информации на естественном языке возникает необходимость в использовании знаний о ПрО, для чего часто используются специализированные онтологии. Основное их назначение в таких задачах – обеспечить связь между фрагментами текста на естественном языке (ЕЯ) и понятиями предметной области (например, классами или экземплярами онтологии). В частности, широко используются тезаурусы и лингвистические онтологии [Loukachevitch, 2014]. Сегодня наиболее распространенными подходами к представлению лингвистических знаний являются традиционные информационно-поисковые тезаурусы, тезаурус WordNet [Pedersen, 2004] и специализированные формальные онтологии.

Тезаурус – это словарь основных понятий языка, которые выражаются отдельными словами или словосочетаниями, с определенными семантическими связями между ними. Как правило, тезаурусы используются для семантической обработки естественно-языковых ИР, например, при семантическом поиске или при семантической разметки. С тезаурусами тесно связаны лексические онтологии, которые представляют собой БЗ онтологического типа о понятийной системе и лексико-терминологическом составе ПрО.

Семантическая разметка связывает текст X с набором тегов разметки при помощи соответствующего языка разметки. Под текстом X , $X = \langle x_1, \dots, x_n \rangle$ будем понимать конечную непустую последовательность символов $x_i \in A, i = \overline{1, n}$, разделенную на отдельные слова разделителями, из конечного множества B , $B \subseteq A$.

Разметка связывает произвольный фрагмент текста $\langle x_p, \dots, x_{p+q} \rangle, 1 \leq p \leq n, 1 \leq p+q \leq n, q \geq 1$ с тэгом $t_i \in T$. Тэги разметки – элементы конечного множества T , $T = \{t_1, \dots, t_m\}$. Наборы тэгов может формироваться из разных источников в зависимости от того, какой язык разметки используется и какая ПрО должна быть охарактеризована. Примеры языков разметки: SGML, RDF Schema, тэги микроформатов, элементы Dublin Core, множество терминов произвольной онтологии.

Разметку называют *семантической*, если для элементов множества тэгов T описана их семантика. Выразительные возможности семантической разметки в значительной мере зависят от набора тэгов, которые используются для этой разметки, их объема, видов отношений между этими тэгами и возможностей логического вывода на этих структурах. Наибольшую

выразительность и универсальность как источник тэгов для семантической разметки имеют онтологии: если в качестве тэгов используются термины онтологии, то это дает возможность установить связь между фрагментами этого текста и понятиями ПрО, которая отображена в онтологии, то есть определить семантическую структуру этого текста [Vargas-Vera, 2002] .

Для автоматизированного выполнения и использования семантической разметки целесообразно использовать специализированную *лексическую онтологию* (ЛО) [Лесько, 2010]. Лексическая онтология – это легкая онтология, которая содержит два основных класса: «понятие ПрО» и «словоформа». Между экземплярами этих классов существует отношение «соответствует». Для каждого экземпляра класса «понятие ПрО» должен существовать хотя бы один экземпляр класса «словоформа», что находится с ним в отношении «соответствует»: $\forall t_i \in T \exists l_{i_1} \in \text{'словоформа'} = t_i$. Принципиальное отличие ЛО от обычных лексических онтологий заключается в том, что она строится для конкретной онтологии ПрО и поэтому как экземпляры класса «лексема» в нее включаются только те понятия, которые связаны с понятиями соответствующей онтологии, а не все понятия, которые присутствуют в естественном языке. Это значительно уменьшает объем ЛО и, соответственно, сокращает время обработки текста.

Задача сопоставления онтологий

Методы согласования онтологий помогают пользователям находить сходства и расхождения между онтологиями или интегрировать две онтологии в результирующую онтологию, которая содержит элементы обеих начальных онтологий. Для этого необходимо определить соответствия между концептами в онтологиях, обеспечить операции отображения, выравнивание и объединение. *Отображение* (сопоставление) онтологий (ontology mapping) – установление соответствия между несколькими онтологиями. *Выравнивание* онтологий (ontology alignment) – установление разных видов соответствия между онтологиями. *Интеграция* онтологий (ontology merging) – создание новой согласованной онтологии или фрагмента онтологии с двух или больше имеющихся онтологий.

При сопоставлении онтологий возникают проблемы их неоднородности. *Синтаксическая* неоднородность – онтологии построены на разных языках представления знаний или с использованием разных формализмов представление знаний, например, с помощью OWL и F-логики. *Терминологическая* неоднородность – расхождения в именах, которые ссылаются на одни и те же сущности в разных онтологиях. *Концептуальная* (семантическая) неоднородность – расхождение в моделировании одной и той же ПрО из-за использования

разных (хотя, возможно, эквивалентных) аксиом для определения концептов или же вследствие использования совсем разных понятий.

Различия между онтологиями могут быть обусловлены целью их создания и критерием классификации объектов и не являются ошибкой или противоречием, а характеризуют представления относительно ПрО разных пользователей и в разные моменты времени. Поэтому в многих случаях самое нахождение таких отличий является важным результатом .

Проблема сопоставления онтологий обусловлена тем, что:

- элементы онтологии (классы, свойства, связи, объекты) с одинаковыми именами могут иметь разную семантику;
- элементы онтологии (классы, свойства, связи, объекты) с одинаковой семантикой могут иметь разные имена.

Методы сопоставления онтологий подразделяют на лингвистические, структурные, статистические и семантические.

Лингвистический анализ определяет сходство между сущностями на основе сравнения их имен и анализа синонимии. Такой вид анализа является начальным этапом установления соответствий между сущностями онтологий.

Структурный анализ сравнивает сходства связей между сущностями (в частности и иерархическими). Оценка схожести двух сущностей двух онтологий может базироваться на позициях этих сущностей в иерархии классов: если подклассы и надкласс двух сущностей двух онтологий подобны, то и сами такие сущности тоже могут быть подобными. Анализ сходства по перекрестным связям для определения сходства между сущностями проводится так: если класс A1 связан с классом B1 связью типа R1 в одной онтологии, а класс A2 связан с B2 связью типа R2 в другой онтологии, и если известно, что B1 и B2 – похожи и R1 и R2 – похожи, то можно предположить схожесть A1 и A2. Так же можно говорить и о сходствах типов связей между R1 и R2, если известно, что A1 и A2 – похожи и B1 и B2 – похожи. Например, если известно, что классы «Человек» и «Лицо» – подобны и отношение «работать» и «быть сотрудником» тоже подобны, причем в первой онтологии класс «Человек» связан отношением «работать» с классом «Организация», а во второй – класс «Лицо» связан отношением «быть сотрудником» с классом «Компания» и «работать», тогда классы разных онтологий «Организация» и «Компания» являются подобными.

Статистический анализ базируется на использовании имеющихся экземпляров двух классов для оценки экстенционального соответствия этих классов. Для нахождения соответствия между сущностями используются такие диагностические правила:

- класс C1 эквивалентен классу C2, если невозможно найти такой экземпляр O1 класса C1, который бы не принадлежал к классу C2, и наоборот;
- класс C1 является подклассом класса C2, если невозможно найти экземпляр O1 класса C1, который бы не принадлежал к классу C2, и класс C1 не эквивалентен классу C2.

Логический анализ выявляет надклассы для тех классов, которые сопоставляются, и анализирует наложенные на них ограничения. Например, в одной онтологии может быть класс «Ребенок», который является видовым классом для класса «Человек» с ограничением на свойство «Возраст»<16, а другая онтология содержит класс «Несовершеннолетнее лицо», который является видовым классом для класса «Лицо» с ограничением на «Возраст лица»<18. Вследствие анализа соответствия между классами «Ребенок» и «Несовершеннолетнее лицо» обнаруживаются родительские классы «Человек» и «Лицо». По наличию информации о соответствии этих классов осуществляется сравнение ограничений, наложенных на эти родительские классы. Для этого сравниваются свойства классов «Возраст» и «Возраст лица»: если эти свойства похожи, то проводится сравнение наложенных ограничений «<16» и «<18». Следствием такого сравнения является вывод, что класс «Ребенок» является подклассом класса «Несовершеннолетнее лицо». После получения локальных соответствий между сущностями можно определять глобальное соответствие между сущностями.

Инструментальные средства поиска соответствия между онтологиями классифицируют по назначению:

- объединение онтологий для создания новой (PROMPT, Chimaera, OntoMerge);
- определение функции преобразование из одной онтологии на другую (OntoMorph);
- обнаружение пар подобных концептов в онтологиях (OBSERVER, FCA-Merge);
- определение правил отображения для связи релевантных частей онтологий (ONION).

Следует отметить, что все эти средства позволяют только строить предположения о наличии соответствий, которые могут оказаться как истинными, так и ложными.

Задача сопоставления (сравнения) терминов онтологий является начальным этапом для построения гипотез относительно сходства между структурными элементами онтологий и требует дальнейшей обработки и проверки пользователем. Неверное сопоставление может быть результатом того, что в близких Про используется похожая терминология, но сами

термины имеют разное значение. В таких случаях экземпляры могут иметь свойства, а классы – подклассы и сверхклассы с подобными названиями, но с различной семантикой.

Первый этап сопоставления терминов онтологий можно выполнить несколькими способами:

- если имена этих элементов тождественны;
- если имена элементов семантически тождественны, например, являются синонимами или переводом на другой язык (такое сопоставление может выполняться автоматизированно с помощью лингвистических методов и находится вне рассмотрения в данной работе);
- непосредственно пользователем на основании его представлений относительно специфики ПрО (например, синонимичные термины в рамках определенной ПрО);
- при помощи логического вывода из тождественности между другими элементами онтологий [Mishra, 2010], т.е. путем анализа свойств классов и экземпляров классов этих онтологий, с помощью которых из имеющихся соответствий логически выводятся другие соответствия, либо путем анализа отношений между теми классами и экземплярами, для которых уже установлены соответствия.

Если результаты последнего способа всегда истинны (например, если установлено соответствие между классом $x_1 \in X_{cl_1}$ онтологии O_1 и классом $x_2 \in X_{cl_2}$ онтологии O_2 и если в онтологии O_1 содержится информация относительно того, что $x_1 \in X_{cl_1}$ эквивалентен классу $y_1 \in X_{cl_1}$, тогда можно считать, что установлено соответствие между классом $y_1 \in X_{cl_1}$ онтологии O_1 и классом $x_2 \in X_{cl_2}$ онтологии O_2) и такой вывод не требует дополнительного подтверждения от пользователя, то результаты, полученные остальными способами, имеют лишь вероятностную достоверность, и потому могут считаться истинными лишь после подтверждения пользователем. Например, если установлено соответствие между классами $x_1 \in X_{cl_1}$ онтологии O_1 и $x_2 \in X_{cl_2}$ онтологии O_2 , и между классами $y_1 \in X_{cl_1}$ онтологии O_1 и $y_2 \in X_{cl_2}$ онтологии O_2 , а также в онтологии O_1 $x_1 \in X_{cl_1}$ является подклассом $y_1 \in X_{cl_1}$, тогда можно предположить, что $x_2 \in X_{cl_2}$ также является подклассом $y_2 \in X_{cl_2}$. Тем не менее, онтология O_2 может связывать эти классы иным образом, в зависимости от критерия классификации.

Это подтверждает важность сведений о свойствах элементов онтологии (например, о тождественности классов), которые не имеют аналогов в семантических Wiki-ресурсах – таким образом, выразительная мощность онтологий значительно выше.

Если не удалось установить ни одного соответствия между элементами онтологий, по их сопоставление заканчивается полной неудачей, то есть считается, что представленные в них знание не пересекаются. В таком случае интеграция онтологий может выполняться путем их непосредственного объединения, которое заключается в объединении пространств имен всех элементов и отношений между ними.

Результатом сопоставления онтологий O_1 и O_2 может быть следующее:

- онтологии O_1 и O_2 тождественны;
- онтология O_1 является подмножеством онтологии O_2 ;
- онтологии O_1 и O_2 пересекаются, то есть имеют общую часть (на практике это обычно означает, что они происходят от одной онтологии);
- схемы онтологий O_1 и O_2 тождественны, но их экземпляры не совпадают (можно выделить случаи, в которых множества экземпляров пересекаются, не пересекаются и одно является подмножеством другого);
- схема онтологии O_1 является подмножеством схемы онтологии O_2 ;
- схемы онтологий O_1 и O_2 пересекаются;
- онтологии O_1 и O_2 отличаются, но количество отличий не превышает определенную количественную меру семантической близости между онтологиями, то есть можно назвать онтологии O_1 и O_2 версиями одной онтологии.

Для того, чтобы предлагать алгоритмы распознавания этих ситуаций, нужно предоставить их формальное определение.

Будем считать, что онтологии O_1 и O_2 *тождественны*, если существует однозначное взаимное сопоставление для всех классов, экземпляров классов, их свойств и их значений.

Будем считать, что онтологии O_1 и O_2 *имеют тождественные схемы*, если существует однозначное взаимное сопоставление для всех классов, их свойств и их значений.

С практической точки зрения наибольший интерес вызывает поиск расхождений между онтологиями и оценка семантической значимости этих расхождений. Другая важная проблема, связанная с найденными расхождениями, – это оценка возможности непротиворечащего объединения таких онтологий и нахождение их пересечения.

После того, как отличия между онтологиями найдены, возникает проблема количественной оценки этих отличий. Количественная оценка может выделять разные типы отличий – например, отличия в использовании иерархических отношений могут считаться более значительными по сравнению с отличиями в использовании отношений, специфичных для ПрО.

Кроме того, часто возникает необходимость сопоставления онтологий с другими семантическими структурами или с различными ИР, неявно содержащими соответствующие знания. Такие сопоставления в чем-то подобны правилам преобразования схем для интероперабельных баз данных. В частности, выделяют несколько видов сопоставления элементов онтологий [Studer, 1998]:

- сопоставление *переименования* (Renaming) используется для преобразования специфичных терминов ПрО в термины, связанные с методами;
- сопоставление *фильтрации* (Filtering) обеспечивает средства для выбора подмножества экземпляров ПрО для соответствующего понятия метода;
- сопоставление класса (Class) предоставляет функции для вычисления экземпляров соответствующего понятия метода из определений понятий в приложении, а не из экземпляров в приложении.

Еще одна важная задача, которая находится вне рамок классического управления знаниями, – это сопоставление онтологий с естественными тестами. Ее подзадачами являются разметка ЕЯ-текста терминами онтологии, пополнение онтологии знаниями, добытыми из размеченного ЕЯ-текста, и вычисления меры семантической близости между текстом и онтологией. Решение этой задачи нуждается в учет специфики отдельных ЕЯ, и потому существующие средства и методы решения этой задачи должны разрабатываться для каждого языка в отдельности.

Построение лингвистической БЗ, которое позволяет соотносить фрагменты ЕЯ-текста с терминами онтологии, также находится вне рассмотрения данной работы (следует отметить, что на сегодня же существует определенное количество таких БЗ, в том числе и для украинского языка, и средств использования этих БЗ в семантической разметке, которые пригодны для использования в данной задаче). Кроме лингвистической БЗ, для сопоставления терминов онтологии с ЕЯ-текстом целесообразно использовать средства лингвистического анализа (например, для решения омонимии и определения частей предложения).

С точки зрения управления знаниями большой интерес вызывает обратная задача – по ЕЯ-тексту, который семантически размечен терминами онтологии, усовершенствовать эту онтологию знаниями, которые помещаются в этом тексте. Такое усовершенствование может обеспечить как установление новых связей между существующими элементами онтологии

(например, если фрагменты одного предложения связанные с экземпляром класса и , так и пополнение онтологии новыми элементами.

Сопоставление онтологий и семантических Wiki-ресурсов

Проблема сопоставления онтологий и семантических Wiki-ресурсов возникает в нескольких случаях. Во-первых, при создании семантических Wiki-ресурсов необходимо вначале сформировать набор категорий и семантических свойств. Но встроенные средства Semantic MediaWiki не позволяют ни визуализировать эту информацию, ни оценить ее целостность и непротиворечивость. Поэтому целесообразно вначале построить онтологию той ПрО, которая отображается в Semantic MediaWiki, а затем использовать эту онтологию в качестве основы для семантической разметки. Во-вторых, семантически размеченные Wiki-ресурсы намного более динамичны по сравнению с онтологиями – в их усовершенствовании и обновлении может участвовать широкий круг пользователей, и поэтому они могут быть полезны для усовершенствования соответствующей онтологии ПрО.

Рассмотрим более детально соответствия между элементами онтологии ПрО и страниц Semantic MediaWiki. Некоторые такие соответствия взаимно-однозначны и могут выявляться автоматически, некоторые – требуют дополнительных уточнений от пользователя. Для более строгого описания этих соответствий воспользуемся формальными моделями (1) и (2).

Table 1. Соответствия между основными элементами онтологий и Wiki-ресурсов

Semantic MediaWiki	Онтология	Отображение из Wiki в онтологию	Отображение из онтологии в Wiki
Категория	Класс	Взаимно-однозначное $P_{\text{categ}} \rightarrow X_{\text{cl}}$	Многозначное $X_{\text{cl}} \rightarrow P_{\text{categ}} \cup P_{\text{template}}$
Иерархия категорий	Иерархия классов	Многозначное	Взаимно-однозначное
Wiki-страница	Экземпляр класса	Многозначное $P_{\text{user}} \rightarrow X_{\text{ind}}$	Взаимно-однозначное $X_{\text{ind}} \rightarrow P_{\text{user}}$
Ссылка на другую Wiki-страницу	Объектное отношение	Взаимно-однозначное $L = \{\text{"link"}\} \rightarrow R$	Взаимно-однозначное $R \rightarrow L = \{\text{"link"}\}$

Семантическое свойство типа «страница»	Объектное свойство	Взаимно-однозначное $P_{sem_prop_page} \rightarrow \{r_i\}$	Взаимно-однозначное $\{r_i\} \rightarrow P_{sem_prop_page}$
Семантическое свойство любого другого типа	Свойство данных	Взаимно-однозначное $P_{sem_prop} \rightarrow \{p_i\}$	Взаимно-однозначное $\{p_i\} \rightarrow P_{sem_prop}$
Шаблон	Класс	Взаимно-однозначное $P_{template} \rightarrow X_{cl}$	Многозначное $X_{cl} \rightarrow P_{categ} \cup P_{template}$

Представление значений семантических свойств отображается в лексическую онтологию: для пополнения ЛО словоформами используется информация из конструкции Semantic MediaWiki: [[семантическое свойство:: термин |словоформа]].

Таким образом, если уже построена онтология на языке OWL, то ее достаточно просто использовать в Semantic MediaWiki. Но обратный процесс не может быть полностью автоматизирован. Более того, при автоматической генерации онтологии по Semantic MediaWiki будут утеряны содержащиеся в OWL-онтологии сведения о характеристиках классов и свойств, которые не имеют аналогов Wiki (в частности, об эквивалентности классов и свойств, их непересекаемости, об их области значения и определения). В то же время, часть контента Semantic MediaWiki не может быть непосредственно трансформирована в онтологии. Например, тот факт, что в страницах использован один и тот же шаблон, говорит о том, что на эти страницах описаны информационные объекты одного типа, но для отображения этого в онтологии надо создать специфичный класс и связать его с элементом страницы. Но затем по такой онтологии невозможно понять, что надо создать в Wiki – шаблон или категорию: выбор зависит от пользователя, т.к. для ИО со специфичной, но формализованной структурой целесообразно создавать шаблоны, а во всех остальных случаях – страницы, которые будут отнесены к какой-либо категории. Кроме того, нельзя связать с классом онтологии не всю страницу, а ее конкретный фрагмент (кроме тех случаев, когда у него есть подзаголовок).

Практическое использование

Предложенный подход к управлению знаниями был апробирован при разработке интеллектуальной системы информационного и когнитивного сопровождения функционирования Национальной рамки квалификаций, в которой онтологический анализ использовался для описания компетенций [Rogushina, 2012].

На основе анализа естественных языковых описаний национальных и европейских рамок [Lundqvist, 2011] была построена онтология, которая отображала эталонную модель рамки квалификаций. Эта модель описывает семантические свойства и отношения ИО, которые связаны с результатами обучения. Онтология формализует отношение между этими ИО и устанавливает их иерархию. Пользователь семантически размечает контент Wiki-страниц элементами этой онтологии: определяет категорию (или набор категорий) любого документа и идентифицирует отдельные элементы его контента (рис.1). Использование онтологической модели рамки квалификаций обеспечивает автоматизировать обработку результатов обучения и интегрируя знания, необходимые пользователям [Rogushina, 2016]. В частности, наличие онтологии позволяет пользователям представить структуру категорий и семантических свойств, которые можно использовать в запросах.

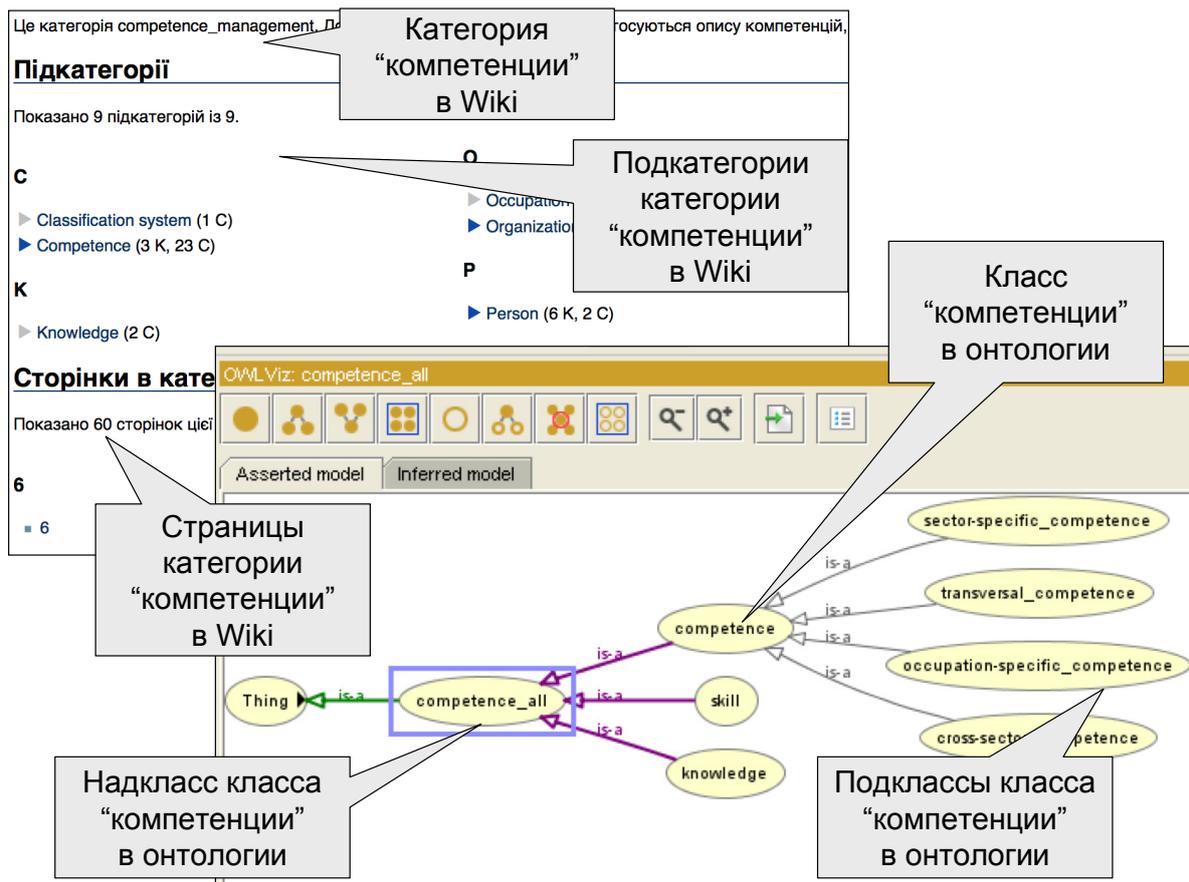


Рис.1. Соответствие между классом онтологии и категорией Wiki «Компетенция»

В общем случае для того, чтобы сопоставлять онтологии, нужно установить соответствие между их элементами. При сопоставлении версий онтологий задача значительно упрощается, так как большая часть соответствий между именами элементов устанавливается автоматически (имена не меняются), но необходимо найти возникшие отличия и оценить их влияние на Wiki-ресурс.

При этом важно, чтобы онтологическая модель соответствовала актуальному представлению Wiki-ресурса – иначе запросы не будут релевантны потребностям пользователей (например, если семантическое свойство было переименовано, а в запросе используется старое имя, то значительная часть информации не будет обнаружена). Такая ситуация может возникнуть и в том случае, если разработчики Wiki-ресурса и его пользователи применяют разные версии онтологической модели. При этом возникает необходимость сопоставления версий онтологий.

Благодарности

В данном исследовании использованы результаты, полученные автором при выполнении государственной темы III-2-12 “Разработка методов и средств поддержки построения интеллектуальных информационных систем в семантической Веб-среде” (Институт программных систем НАН Украины) и темы “Разработка интеллектуальной системы информационного и когнитивного сопровождения функционирования национальной рамки квалификаций” (Мелитопольский государственный педагогический университет).

Работа опубликована при частичной поддержке *ITHEA ISS* (www.ithea.org) и *ADUIS* (www.aduis.com.ua).

Выводы

Предложенный в работе подход к сопоставлению элементов онтологии с семантическими Wiki-ресурсами позволяет, с одной стороны, использовать преимущества онтологического представления информации для управления знаниями в интеллектуальных информационных системах (например, в различных Wiki-справочниках), а с другой, обеспечивает автоматизированный и простой для пользователей способ пополнения и усовершенствования онтологий различных предметных областей. Представляется целесообразным апробировать предложенные методы на более широком классе прикладных задач и обеспечить их соответствующими инструментами.

Литература

- [Berners-Lee, 2001] Berners-Lee T., Hendle, J., & Lassil, O. The semantic web. *Scientific american*, 284(5), 28-37, 2001. – <https://pdfs.semanticscholar.org/566c/1c6bd366b4c9e07fc37eb372771690d5ba31.pdf>
- [Davies, 2002] Davies J., Fensel D., van Harmelen F. *Towards the Semantic Web: Ontology-driven knowledge management*. – John Wiley & Sons Ltd, , England, 2002.
- [Gladun, 2013] Gladun A., Rogushina J., Valencia-García R., Martínez-Béjar R. Semantics-driven modelling of user preferences for information retrieval in the biomedical domain. *Informatics for health and social care*. V.38, N.2, 2013. – P.150-170.
- [Gruber, 1995] Gruber, T.R. Toward principles for the design of ontologies used for knowledge sharing? (.). Toward principles for the design of ontologies used for knowledge sharing?. *International journal of human-computer studies*, 43(5-6), 907-928, 1995. – <http://eolo.cps.unizar.es/docencia/doctorado/Articulos/Ontologias/Toward%20Principles%20for%20the%20Design%20of%20Ontologies%20Used%20for%20Knowledge%20Sharing-Gruber1995.pdf>.
- [Guarino, 1995] Guarino N. Formal ontology, conceptual analysis and knowledge representation. *International journal of human-computer studies*, 43(5-6), 625-640.1995. – https://www.researchgate.net/profile/Nicola_Guarino/publication/2368416_Forma_l_Ontology_Conceptual_Analysis_and_Knowledge_Representation/links/5593060d08ae16f493ee4d94.pdf.
- [Guarino, 2000] Guarino N., Welty C. A formal ontology of properties. *Proc. of International Conference on Knowledge Engineering and Knowledge Management*, 97-112, 2000. – <http://cuiwww.unige.ch/isi/cours/aftsi/articles/01-guarino00formal.pdf>
- [Holtshouse, 2013] Holtshouse, D. K. *Information technology for knowledge management*. U. M. Borghoff, & R. Pareschi (Eds.). Springer Science & Business Media, 2013. – <http://tocs.ulb-tu-darmstadt.de/55883427.pdf>
- [Loukachevitch, 2014] Loukachevitch N., Dobrov, B. RuThes linguistic ontology vs. Russian wordnets. In *Proc.of Global WordNet Conference GWC-2014*, 2014. – <http://www.aclweb.org/anthology/W/W14/W14-0121.pdf>.
- [Lundqvist, 2011] Lundqvist K. Ø., Baker K., Williams, S. (2011). Ontology supported competency system. *International Journal of Knowledge and Learning*, 7(3-4), 197-219. – <http://centaur.reading.ac.uk/23841/1/authorFinalVersion.pdf>.

- [Majchrzak, 2006] Majchrzak A., Wagner C., Yates D. Corporate wiki users: results of a survey. In Proc. of the 2006 international symposium on Wikis, 99-104, 2006. – <http://stu.hksyu.edu/~wkma/notes/jour395/wagner2004.pdf>
- [Mishra, 2010] Mishra R B, Sandeep K. Semantic Web Reasoners and Languages, Springer, 2010.
- [Pedersen, 2004] Pedersen T., Patwardhan S., Michelizzi J. WordNet: Similarity - measuring the relatedness of concepts. Proc. Of the 19th National Conference on Artificial Intelligence (AAAI-04), 1024-1025, 2004.
- [Rogushina, 2012] Rogushina J., Gladun A. Ontology-based competency analyses in new research domains. Journal of Computing and Information Technology. V.20, N. 4, 2012. – P.277-293.
- [Rogushina, 2016] Rogushina J. Semantic Wiki resources and their use for the construction of personalised ontologies. Proc. of the 10th International Conference of Programming UkrPROG'2016, 196-203, 2016. – <http://ceur-ws.org/Vol-1631/188-195.pdf>
- [Studer, 1998] Studer R., Benjamins V. R., Fensel D. Knowledge engineering: principles and methods. Data & knowledge engineering, 25(1-2), 161-197, 1998. – https://www.researchgate.net/profile/V_Richard_Benjamins/publication/222305044_Knowledge_engineering_principles_and_methods_Data_Knowl_Eng_251-2161-197/links/0fcfd50c3673c0368e000000/Knowledge-engineering-principles-and-methods-Data-Knowl-Eng-251-2161-197.pdf.
- [Tudorache, 2013] Tudorache T., Nyulas C., Noy N. F., Musen M. A. (2013). WebProtégé: A collaborative ontology editor and knowledge acquisition tool for the web. Semantic web, 4(1), 89-99. – <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3691821/>.
- [Uschold, 1996] Uschold M., Grüninger M. Ontologies: Principles, Methods and Applications. Knowledge Engineering Review, 11(2), 93–155, 1996.
- [Vargas-Vera, 2002] Vargas-Vera M., Motta E., Domingue J., Lanzoni M., Stutt A., Ciravegna F. MnM: Ontology driven tool for semantic markup. Proc. Workshop Semantic Authoring, Annotation & Knowledge Markup (SAAKM 2002), 43-47, 2002. – <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.20.3590>
- [Völkel, 2006] Völkel M., Krötzsch M., Vrandečić D., Haller H., Studer, R. Semantic wikipedia. Proc. of the 15th international conference on World Wide Web. 585-594, 2006. – <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.103.9471&rep=rep1&type=pdf>.

[Гладун, 2016] Гладун А.Я., Рогушина Ю.В. Семантичні технології: принципи та практики. К.:ТОВ "ВД "АДЕФ-Україна", 2016. – <https://core.ac.uk/download/pdf/38468940.pdf>.

[Лесько, 2010] Лесько О.Н., Рогушина Ю.В. Использование специализированной лексической онтологии для автоматизации формирования онтологии предметной области по естественно-языковым текстам. Information Models of Knowledge. Edited by K.Markov, V.Velychko, O.Voloshin. – ITHEA, Kiev-Sofia, 93-100, 2010.

Информация об авторе

Рогушина Юлия – к.ф.-м.н., доцент, с.н.с. Института программных систем НАН Украины; ORCID <http://orcid.org/0000-0001-7958-2557>, e-mail: ladamandraka2010@gmail.com.

MATCHING OF SEMANTIC INFORMATION RESOURCES OF THE WEB ON THE BASIS OF ONTOLOGICAL ANALYSIS

Julia Rogushina

Abstract: *Modern means of representation of distributed knowledge are analyzed. The problems arising in the process of knowledge managing for intelligent information systems oriented to work on the Web are considered. The necessity of ontology use and the actuality of the development of semantic technologies aimed at their support, in particular, the standards and languages of the Semantic Web, are grounded. The problem of ontology matching, the results of such a matching and a particular case of this problem where compared ontologies are the development of the same initial ontology are considered. Semantic Wiki-resources are considered as a source of information for replenishment and improvement of ontologies of domain that is displayed by the set of Wiki pages. The analysis of components either of domain ontology or of semantic markup of Wiki-resources (on example of Semantic MediaWiki) is carried out, the method of their automated matching for the improvement of knowledge bases of distributed intellectual systems is proposed.*

Keywords: *domain ontology model, ontology, Wiki-technology, automatically generated ontologies, Subject Domain.*

СТАБИЛИЗАЦИЯ НАПРЯЖЕНИЯ РАСПРЕДЕЛИТЕЛЬНОЙ ЭЛЕКТРОСЕТИ НА ОСНОВЕ ОДНОКРИСТАЛЬНОЙ микроЭВМ

Евгений Забудский

Аннотация: Особенности развития электроэнергетики, а именно: рост единичных мощностей энергоустановок, большая протяженность электросетей, наличие слабых межсистемных связей, повышение неравномерности графиков нагрузки обуславливают необходимость совершенствования систем автоматического управления (САУ) энергетическими объектами. Опыт эксплуатации микропроцессорных САУ электроэнергетическими объектами показал их преимущества по сравнению с устройствами, выполненными на дискретных компонентах. Разработана микропроцессорная САУ, которая совместно со статическим компенсатором реактивной мощности (СКРМ) предназначена для поддержания напряжения в электроэнергетической системе в месте его установки на заданном уровне. Одновременно со стабилизацией напряжения снижаются потери мощности в системе, и улучшается режим работы электропотребителей.

Ключевые слова: микроЭВМ, система автоматического управления, управляемый реактор, распределительная электросеть.

ITHEA Keywords: J.7 Computers in Other Systems – Process control.

Введение

В автоматических устройствах управления с встроенными микроЭВМ реализуются более сложные алгоритмы управления. Запись программ в перепрограммируемое постоянное запоминающее устройство (ПЗУ) позволяет изменять алгоритмы в процессе эксплуатации при развитии энергосистем и адаптировать их к индивидуальным характеристикам конкретных электрических схем. Применение встраиваемых микроЭВМ в устройствах автоматического управления электроэнергетическими объектами дает возможность снизить ущерб от повреждения электрооборудования и повысить качество вырабатываемой электроэнергии.

Разработана, изготовлена и испытана микропроцессорная система управления углом открытия оптодиодов для управляемых реакторов (УР), электромагнитные схемы которых

рассмотрены и исследованы [Zabudsky, 2003]. Одно из назначений управляемого реактора – выполнение функций регулирующего элемента СКРМ, который предназначен для работы в распределительных сетях (Рис. 1,а,б).

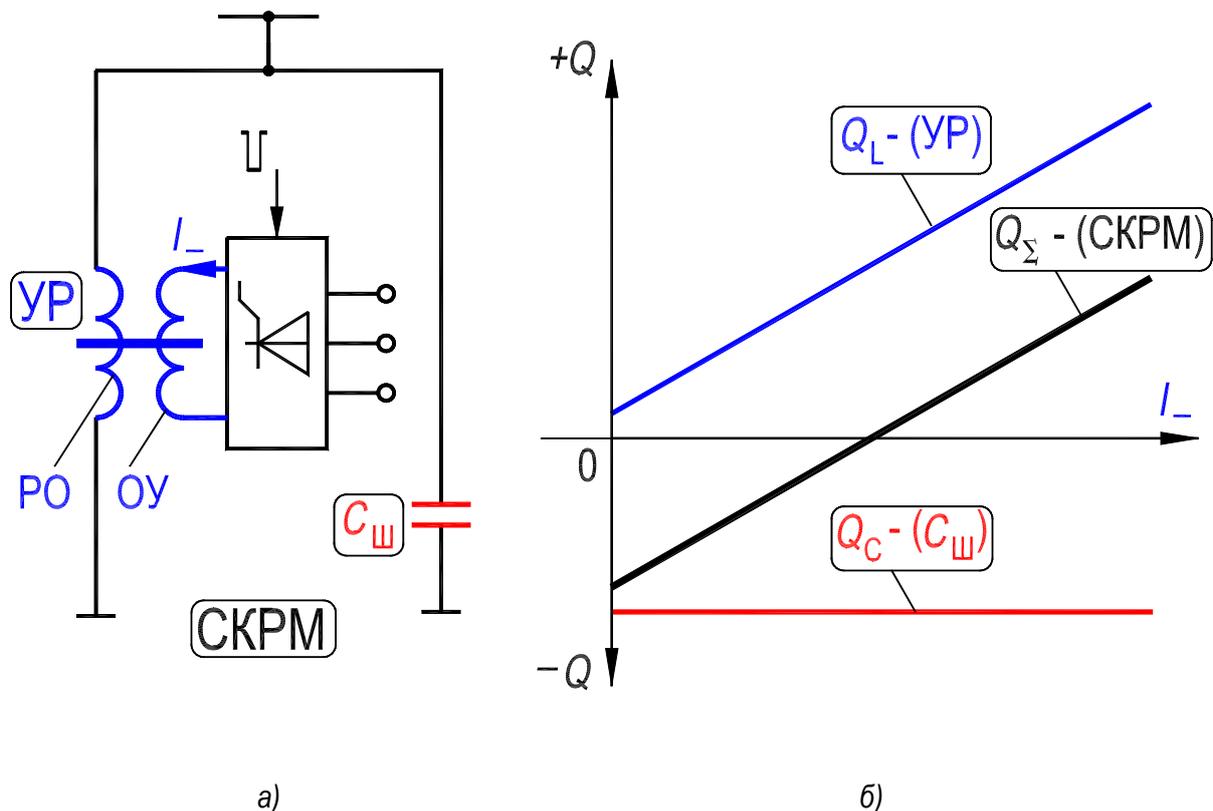


Рис. 1. Принципиальная схема СКРМ (а); характеристика реактивной мощности СКРМ (б)

Так как при компенсации реактивной мощности Q потери напряжения в сети уменьшаются, то при определенных условиях СКРМ используется не только для обеспечения баланса реактивной мощности, но и в качестве средства регулирования и стабилизации напряжения в сети в месте его установки. Это достигается регулированием потребления и выдачи реактивной мощности компенсатором посредством выработки управляющего воздействия, подаваемого на оптотиристоры в соответствии с измеряемыми переменными системы. Одновременно со стабилизацией напряжения снижаются потери мощности в электроэнергетической системе, и улучшается режим работы электропотребителей.

Микропроцессорная система управления реализована на основе однокристалльной микроЭВМ KP1816BE51 [Stashin, 1990]. Разработан алгоритм и составлена управляющая ассемблер-программа. Управление реакторами, предназначенными для работы в распределительных

сетях, осуществляется автоматически путем сравнения номинального напряжения сети с измеряемым напряжением в месте установки реактора. Управление реакторами, предназначенными для дальних высоковольтных ЛЭП, осуществляется автоматически путем сравнения натурального тока линии с протекающим по линии током в месте установки реактора [Aleksandrov, 1993]. Целью автоматического регулирования в первом случае является поддержание напряжения в сети на заданном уровне. Во втором случае целью автоматического регулирования является компенсация избыточной реактивной мощности линии.

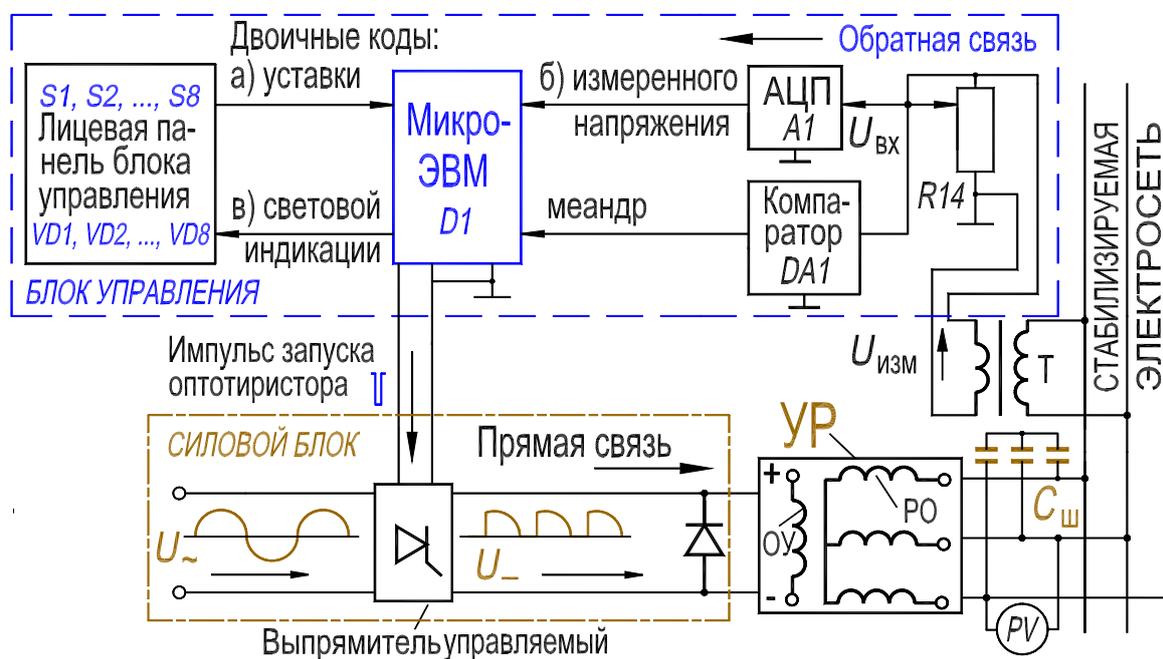


Рис. 2. Функциональная схема САУ

1. Структура и принцип действия САУ

Целью действия САУ является автоматическая стабилизация напряжения распределительной электросети в месте установки реактора на уровне соответствующем номинальному напряжению с заданной точностью приблизительно $\pm 1\%$. Это реализуется за счет изменения тока подмагничивания $I_{\text{п}}$ совмещенных управляемых реакторов и реакторов-трансформаторов. При уменьшении напряжения сети необходимо уменьшать ток подмагничивания до тех пор, пока напряжение сети не увеличится до нормируемого уровня, а при увеличении напряжения – необходимо соответственно увеличить ток подмагничивания. САУ состоит из блока управления,

силового блока и источника постоянного стабилизированного напряжения +5 В, +15 В и -15 В. На Рис. 2 показана функциональная схема САУ.

Однокристалльная микроЭВМ КР1816ВЕ51 (*D1*), на основе которой реализован блок управления, имеет следующие аппаратные средства: процессор, в состав которого входят восьмиразрядное арифметико-логическое устройство и схемы аппаратной реализации команд умножения и деления; перепрограммируемое ПЗУ емкостью 4 Кбайт, ОЗУ данных емкостью 128 байт; два 16-битных таймера/счетчика; программируемые схемы ввода/вывода (32 линии); блок двухуровневого векторного прерывания от пяти источников; асинхронный канал дуплексного последовательного ввода/вывода информации со скоростью до 375 Кбит/с; генератор; схема синхронизации и управления.

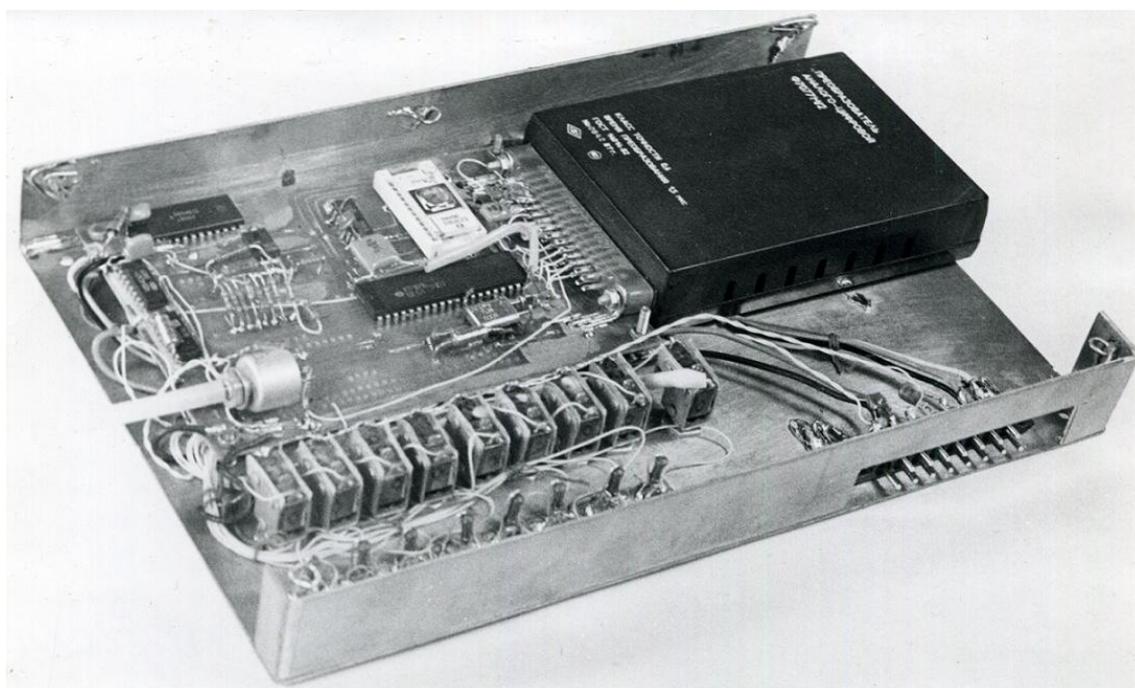
МикроЭВМ серии 1816 требуют одного источника электропитания $+5\text{ В} \pm 10\%$, рассеивают мощность $\sim 1,5$ Вт. Исходные тексты программ для микроЭВМ составляются на языке ассемблера. МикроЭВМ может работать в диапазоне частот от 1,2 до 12 МГц, при этом минимальный цикл выполнения команды равен 1 мкс, а быстродействие – 1 млн коротких операций в секунду.

Однокристалльная микроЭВМ КР1816ВЕ51 обладает значительными функционально-логическими возможностями и представляет собой эффективное средство автоматизации на основе применения средств и методов обработки данных и цифрового управления разнообразных, в том числе и электроэнергетических объектов и процессов.

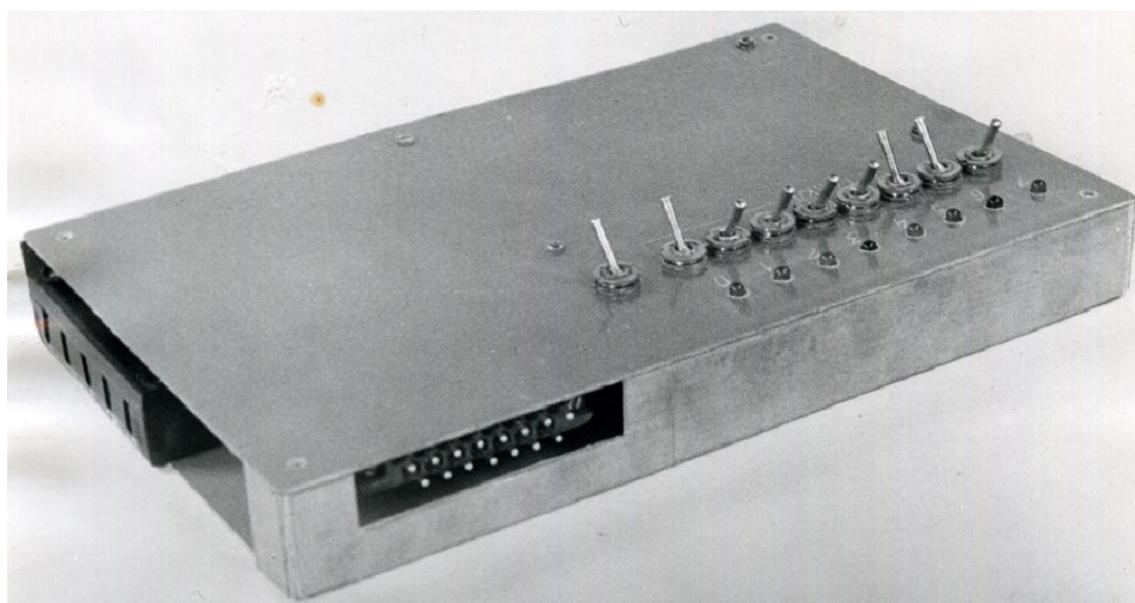
В состав блока управления входит также АЦП (преобразователь) Ф7077М/2 (*A1*), который выполняет функцию преобразования измеряемого напряжения стабилизируемой электросети в цифровой двоичный код. Преобразователь *A1* характеризуется следующими основными техническими данными: принятый диапазон преобразуемого напряжения от $-1,02$ до $+1,02$ В; класс точности – 0,6; номинальная ступень квантования – 4 мВ; время преобразования не более 1,5 мкс; время цикла кодирования не более 1 мкс; выходной код преобразователя – двоичный, параллельный девятиразрядный. Преобразователь *A1* реализован на основе полупроводниковой интегральной схемы К1108ПВ1 – десятиразрядного, прецизионного, быстродействующего АЦП.

Питание блока управления САУ осуществляется от источника питания +5 В, +15 В, -15 В, а силовой блок питается от сети переменного тока. Габаритные размеры блока управления 260x180x40 мм, его масса составляет не более 1 кг. Фотографии блока управления приведены на Рис. 3,а,б. Посредством оптодиодов, которые включены в два плеча мостового управляемого выпрямителя питающего обмотку управления (ОУ) реактора, осуществляется

воздействие блока управления на силовой блок; с другой стороны, посредством оптодиристоров осуществляется гальваническая развязка этих блоков, т.е. слаботочной и силовоточной цепей.



а)



б)

Рис. 3. Фотографии блока управления: вид на плату – (а); вид на лицевую панель – (б)

Принцип действия САУ заключается в следующем. Стабилизируемое напряжение распределительной электросети через трансформатор напряжения Т поступает на АЦП А1 (Ф7077М/2), где преобразуется в цифровой двоичный код, который сравнивается процессором с цифровым двоичным кодом (уставкой) задающимся с помощью переключателей S1, S2, ..., S8, установленных на лицевой панели блока управления (Рис. 2 и 3,б). При неравенстве этих кодов определяется знак рассогласования и соответствующим образом изменяется угол открытия оптоисториков в каждом полупериоде сети, от которой питается силовой блок, что приводит к изменению постоянной составляющей тока протекающего в ОУ реактора и в свою очередь к изменению напряжения стабилизируемой электросети [Zabudsky, 1994]. Система автоматического регулирования, совместно с управляемым реактором, образуют замкнутую систему автоматического поддержания заданного уставкой напряжения электросети.

2. Основы функционирования САУ

При разработке алгоритма функционирования САУ использован математический аппарат цифровой (дискретной) обработки аналоговых величин, который сводится к следующему.

Среднее квадратичное значение напряжения за половину периода вычисляется согласно формуле:

$$U = \sqrt{(2/T) \int_0^{T/2} u^2 dt}, \quad (1)$$

где u – мгновенное значение измеряемого напряжения; T – период изменения стабилизируемого напряжения, при частоте 50 Гц равный 0,02 с.

Для цифровой обработки аналоговой величины осуществляется дискретизация непрерывной временной функции напряжения распределительной электросети, т.е. интеграл в (1) заменяется суммой:

$$U = \sqrt{(2/T) \sum_{n=0}^k u_n^2 \Delta t}, \quad (2)$$

где u_n – значение n -ой выборки измеряемого напряжения стабилизируемой электросети, которое подается на вход АЦП А1; n – текущий номер выборки, т.е. измерения напряжения; k – количество выборок в течение половины периода.

Примем, что преобразование аналоговой величины в цифровую форму осуществляется через равные промежутки времени $\Delta t = \text{const}$, тогда (2) записывается в виде:

$$U = \sqrt{(1/k) \sum_{n=0}^k u_n^2}. \quad (3)$$

Значение количества выборок k принято из соображения укладки его в один байт ($k \sim 210$) и обеспечения достаточной точности. В этом случае промежуток между двумя выборками определяется как

$$\Delta t = T/2k \approx 48 \text{ мкс}. \quad (4)$$

Использованный АЦП преобразует мгновенное значение переменного напряжения, поступающего на его вход, в цифровой двоичный код, снимаемый с выхода за 1,5 мкс. Изменение напряжения на входе за это время много меньше веса младшего двоичного разряда АЦП (для данного уровня и частоты входного напряжения), поэтому можно утверждать, что АЦП работает в режиме преобразования постоянного напряжения.

На восходящем участке функции $u^2(t)$ преобразование напряжения в цифровой код осуществляется АЦП $A1$ с недостатком, а на нисходящем – с избытком. Однако при достаточно большом числе выборок за полпериода эти погрешности взаимно компенсируются вследствие симметрии зависимости $u^2(t)$.

Достоинством используемого аппарата преобразования аналоговой величины в дискретную является устойчивость по отношению к высокочастотным помехам.

Для оценки рассогласования между измеренным напряжением распределительной электросети и эталонным напряжением U_3 , заданным уставкой, и выработки соответствующего управляющего воздействия вычисляется на основе (3) среднее значение квадрата измеренного напряжения за полпериода по формуле:

$$U_{\text{cp}} = \left(\sum_{n=0}^k u_n^2 \right) / k. \quad (5)$$

Среднее квадратичное значение напряжения распределительной электросети за положительный полупериод составляет:

$$U_{\text{д}} = \sqrt{U_{\text{cp}}}. \quad (6)$$

Таким образом, необходимо согласно (5) осуществить программно накопление суммы квадратов двоичных кодов k выборок измеренного напряжения, а затем определить среднее значение этой суммы. Далее необходимо программно двоичный код уставки возвести в квадрат. Результат последнего действия соответствует квадрату эталонного напряжения U_3 , заданного уставкой. На

основе сопоставления вычисленных двоичных кодов величин U_{cp} и U_3^2 определяется знак рассогласования и вырабатывается соответствующее управляющее воздействие.

Вышеприведенный математический аппарат реализован в управляющей ассемблер-программе, которая состоит из основной программы и шести подпрограмм: 1) подпрограмма обработки прерываний от таймера; 2) подпрограмма вычисления квадратного корня; 3) подпрограмма сдвига вправо двухбайтового числа; 4) подпрограмма умножения двухбайтовых чисел; 5) подпрограмма деления многобайтовых чисел; 6) подпрограмма сдвига влево многобайтового числа. В Приложении представлены на Рис. П1 – структурная диаграмма основной ассемблер-программы и на Рис. П2 – структурная диаграмма подпрограммы обработки прерываний от таймера.

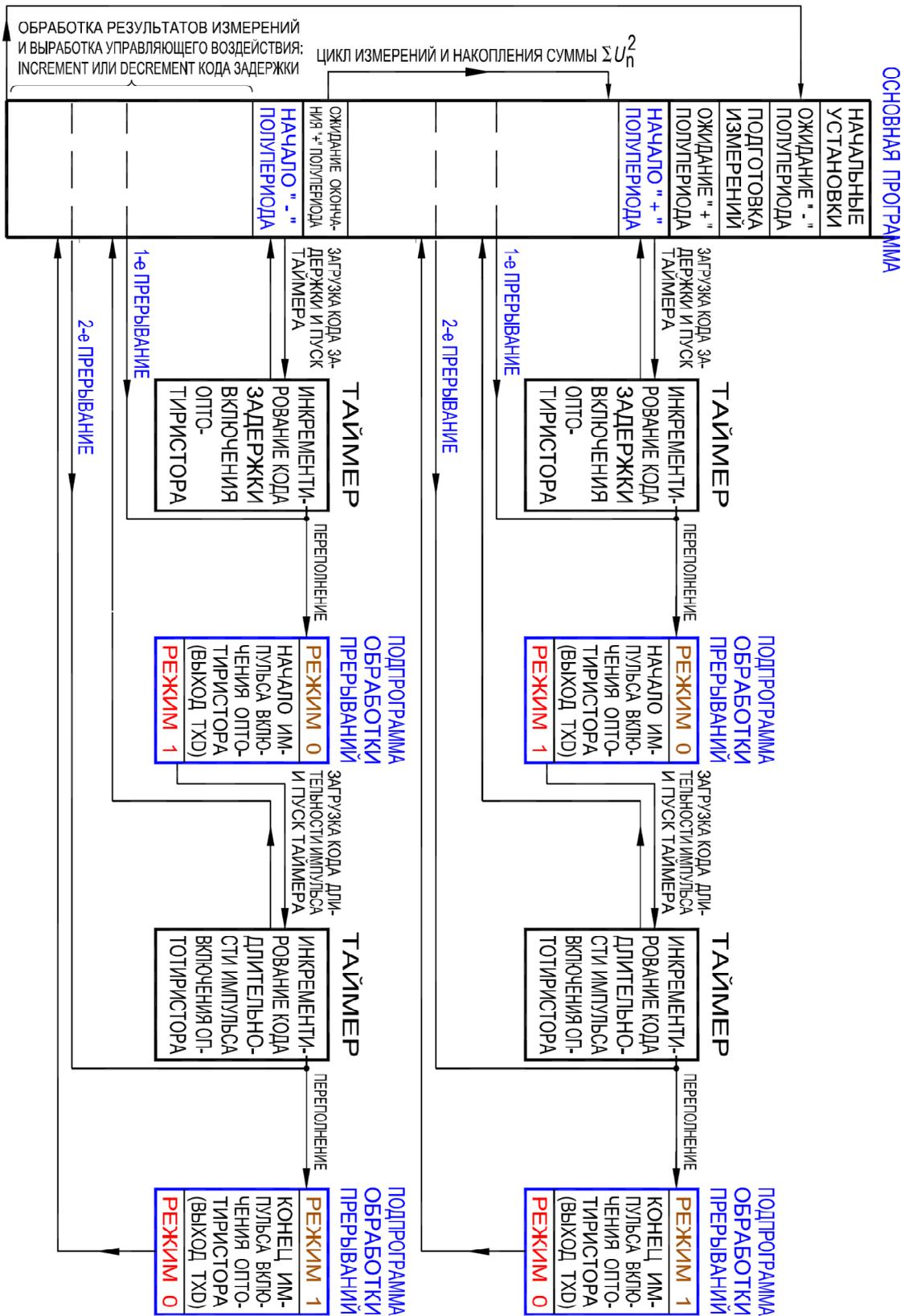
Подпрограмма обработки прерываний от таймера работает в двух режимах: режимы "0" или "1", которые определяются признаком, т.е. значением бита #08₁₆ в памяти данных. Если бит установлен в "0", то имеет место режим "0" подпрограммы, а если бит установлен в "1", то – режим "1". Режим "0" характеризуется тем, что в нем начинается импульс запуска оптодирикторов, длительность этого импульса установлена программно в 200 мкс, а в режиме "1" импульс запуска оптодирикторов прекращается. Исходный текст программ составлен в мнемокодах ассемблера микроЭВМ и посредством кросс-ассемблера оттранслирован в шестнадцатеричный код. Машинный код программ помещен в ППЗУ 573РФ2.

3. Алгоритм работы САУ

Система автоматического управления работает следующим образом: 1) переключатель S9 "Сброс-работа" устанавливается в положение "Сброс"; 2) на блок управления подается питание + 5 В, + 15 В и – 15 В;

3) силовой блок подключается к сети переменного тока; 4) на переключателях S1, S2, ..., S8 набирается двоичный код уставки; 5) микроЭВМ (D1) переходит из состояния сброса к работе по управляющей программе. В результате в электросети в месте расположения УР устанавливается в соответствии с ВАХ реактора заданное значение напряжения, отображаемое на вольтметре PV (Рис. 2).

Последовательность работы управляющей программы, а, следовательно, и всей системы управления, рассмотрим на основе блок-схемы взаимодействия основной программы с таймером и подпрограммой обработки прерываний, инициируемых по переполнению таймера (Рис. 4).



ис. 4. Блок-схема взаимодействия основной программы с таймером и подпрограммой обработки прерываний, происходящих при переполнении таймера

Управляющая программа составлена так, что в течение каждого положительного полупериода напряжения $U_{вх}$, подаваемого на вход АЦП $A1$, осуществляется цикл измерений, т.е. накопление суммы квадратов кодов мгновенных значений измеряемого напряжения (5) и подсчет количества циклов измерений, а в течение каждого отрицательного – осуществляется обработка результатов этих измерений (5 и 6), их анализ и коррекция текущего значения кода задержки включения оптодиодов.

В начале работы программы выполняются начальные установки адреса стека, номера банка памяти данных, подготовка прерывания от таймера T/CO , который входит в состав микроЭВМ $D1$. Затем программа анализирует состояние выхода компаратора $DA1$ (554СА3). Если на выходе компаратора логическая "1", то на входе АЦП $A1$ действует отрицательная полуволна напряжения $U_{вх}$ (Рис. 5,а,б). После подготовки в этом полупериоде цикла измерений, ожидается начало положительного полупериода.

4. Работа САУ в положительном полупериоде

1) В начале положительного полупериода производится загрузка таймера T/CO кодом задержки включения оптодиодов и пуск таймера. Код задержки включения оптодиодов устанавливается программно в заданном диапазоне от $38_{16} = 56_{10}$ до $EC_{16} = 236_{10}$, причем с увеличением кода сама задержка, т.е. угол открытия оптодиодов, уменьшается, т.к. сигнал (импульс) включения оптодиодов передается по переполнению таймера. Диапазон задержек включения составляет от 1 мс до 9 мс от начала полупериода изменения напряжения $U_{вх}$ на входе АЦП $A1$. Таймер, после своего включения, работает в режиме "0", согласно которому через каждые 48 мкс (при частоте кварцевого резонатора 8 МГц) к его содержимому добавляется 1. Таким образом, всему диапазону задержек соответствует примерно 180 единиц;

2) после запуска таймера согласно управляющей программе начинается накопление суммы квадратов двоичных кодов мгновенных значений измеряемого напряжения (5);

3) по истечении заданного в таймере времени задержки, происходит первое прерывание работы управляющей программы в положительном полупериоде и переход к подпрограмме обработки прерывания от таймера, которая вначале анализирует признак режима своей работы;

4) в режиме "0" работы подпрограммы обработки прерываний устанавливается в "0" выход "TXD – выход передатчика" микроЭВМ D1 (Рис. 5,з), что соответствует началу импульса тока через светодиод оптодиристора, т.е. включению последнего. В таймер загружается код длительности импульса запуска (~ 200 мкс), устанавливается в "1" признак режима работы подпрограммы и пускается таймер;

5) происходит возврат в основную программу (первое прерывание ее работы закончено), где продолжается накопление суммы квадратов двоичных кодов мгновенных значений измеряемого напряжения (5), (Рис. 5,в);

6) по истечении ~ 200 мкс, т.е. после переполнения таймера, происходит второе прерывание работы основной программы в течение положительного полупериода и передача управления подпрограмме обработки прерывания от таймера;

7) подпрограмма обработки прерываний останавливает таймер и анализирует бит режима, который установлен в "1". Находясь в режиме "1" подпрограмма устанавливает в "1" выход "TXD – выход передатчика" микроЭВМ D1 (Рис. 5,з), что приводит к окончанию импульса тока через светодиод оптодиристора. После установки бита режима работы подпрограммы в "0" осуществляется возврат в основную программу (завершено второе прерывание ее работы);

8) согласно работе основной программы продолжается накопление суммы квадратов двоичных кодов мгновенных значений измеряемого напряжения (5), и подсчет количества циклов, т.е. количества выборок мгновенных значений измеряемого напряжения в течение его положительного полупериода.

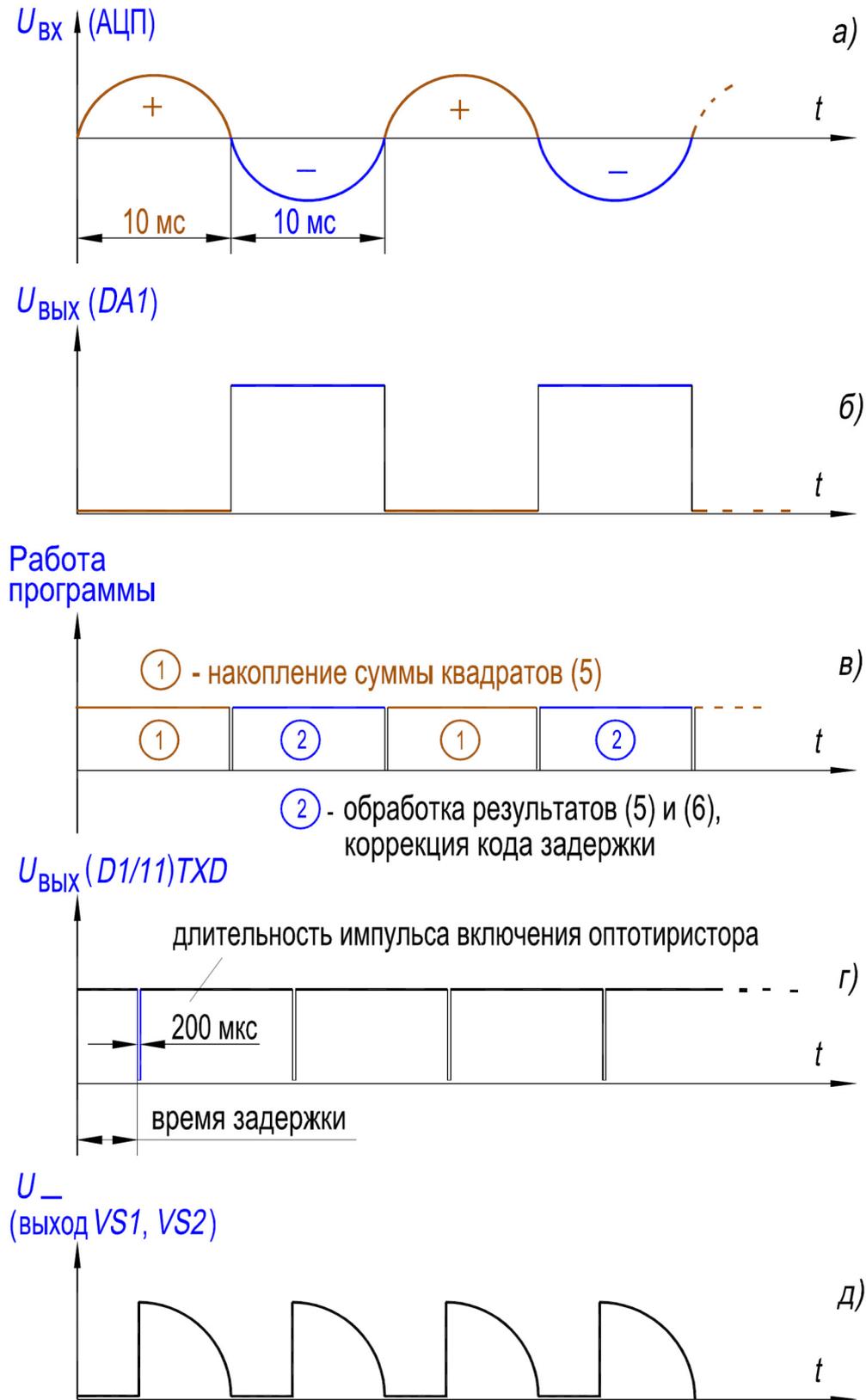


Рис. 5. Временная диаграмма работы САУ

5. Работа САУ в отрицательном полупериоде

1) В начале отрицательного полупериода также производится загрузка таймера T/CO кодом задержки включения оптотиристора и пуск таймера. Работа таймера и подпрограммы обработки прерываний от таймера происходит также, как и в положительный полупериод;

2) при работе основной программы в отрицательный полупериод изменения напряжения $U_{вх}$, подаваемого на вход АЦП $A1$, осуществляется, кроме обработки прерываний от таймера, следующее:

а) вычисляется двоичный код, соответствующий напряжению $U_{ср}$ (5);

б) вычисляется двоичный код, соответствующий напряжению U_d (6), который отображается на светодиодах $VD1, VD2, \dots, VD8$, расположенных на лицевой панели блока управления (Рис. 2 и Рис. 3,б)

в) вычисляется квадрат двоичного кода уставки, т.е. эталонного напряжения $U_э$;

г) на основе сравнения кодов соответствующих $U_{ср}$ и $U_э^2$ формируется управляющее воздействие, т.е. корректируется (декрементируется или инкрементируется) код задержки включения оптотиристоров.

Если $U_{ср} > U_э^2$, то для восстановления равенства между ними, необходимо увеличить ток подмагничивания, поэтому программно увеличивается значение текущего кода задержки включения оптотиристора на 1, а сама задержка, т.е. угол открытия оптотиристоров, при этом уменьшается.

Если $U_{ср} < U_э^2$, то для восстановления равенства между ними, необходимо уменьшить ток подмагничивания, поэтому программно уменьшается значение текущего кода задержки включения оптотиристора на 1, а сама задержка, т.е. угол открытия оптотиристоров, при этом увеличивается.

Если $U_{ср} \sim U_э^2$, то значение кода задержки и самой задержки колеблются возле некоторого среднего значения, обеспечивая выполнение равенства с точностью до единицы младшего разряда. При выполнении равенства $U_{ср} \sim U_э^2$ на светодиодах $VD1, VD2, \dots, VD8$ высвечивается двоичный код среднего квадратичного напряжения распределительной электросети U (6), который соответствует коду $U_э$, зафиксированному переключателями $S1, S2, \dots, S8$ (Рис. 2 и 3,б);

д) после каждой коррекции кода задержки его значение сравнивается с максимальным EC_{16} и минимальным 38_{16} значениями и при выходе кода задержки за эти пределы, ему присваивается

соответственно максимальное или минимальное значение, а коду, высвечиваемому на светодиодах, – значение FF_{16} или 01_{16} . Коды FF_{16} или 01_{16} индицируются на светодиодах $VD1$, $VD2$, ..., $VD8$ для указания на то, что система управления вышла из режима (зоны) стабилизации. На Рис. 5 приведена временная диаграмма работы САУ, а на Рис. 6 – характеристика стабилизации.

В диапазоне $U_{вх.мин} \leq U_{вх} \leq U_{вх.мак}$ имеет место стабилизация, т.е. САУ поддерживает равенство $U_{р.с} = U_{ст}$, т.к. в этой зоне, в соответствии с ВАХ реактора, изменение напряжения $U_{вх}$ на входе АЦП А1 полностью компенсируется соответствующим током подмагничивания $I_{\underline{}}$, протекающим по обмотке управления реактора. Этот ток регулируется системой за счет коррекции кода задержки включения оптодиодов.

Процесс стабилизации возможен в допустимом "окне" задержек включения оптодиодов, составляющем 1...9 мс.

Для того чтобы войти в зону стабилизации необходимо, например, изменить с помощью делителя $R14$ напряжение на входе АЦП А1 или же изменить код уставки с помощью переключателей $S1$, $S2$, ..., $S8$. При этом импульс запуска оптодиодов перемещается в разрешенный диапазон задержек. Примерно через 1...2 секунды система входит в режим стабилизации, о чем свидетельствует высвечивание на индикаторах кода, равного коду уставки.

Максимально возможный код уставки определяется из выражения $[255/(2)^{1/2}] \cdot 0,9 \sim 160_{10} = 1010000_2$, где 0,9 – коэффициент запаса. При большем коде уставки амплитуда стабилизируемого напряжения на входе АЦП А1 становится больше напряжения насыщения АЦП, что приводит к большой погрешности.

6. Результаты экспериментальных исследований

Блок управления и силовой блок САУ изготовлены в соответствии с принципиальными электрическими схемами на Рис. 7, а, б.

Опытный образец совмещенного управляемого реактора изготовлен согласно электромагнитной схеме Рис. 8 [Zabudsky, 1994].

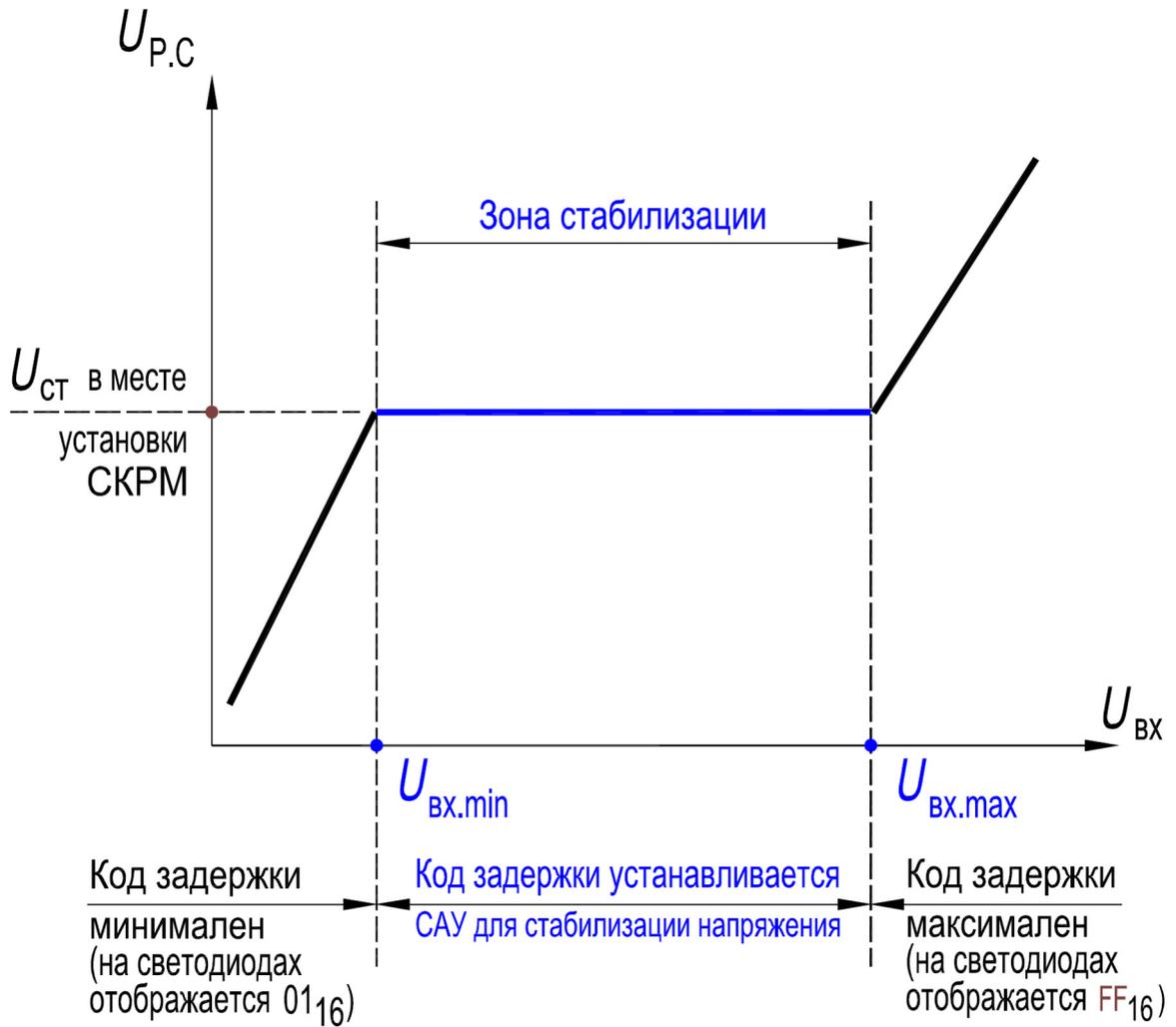


Рис. 6. Характеристика стабилизации САУ

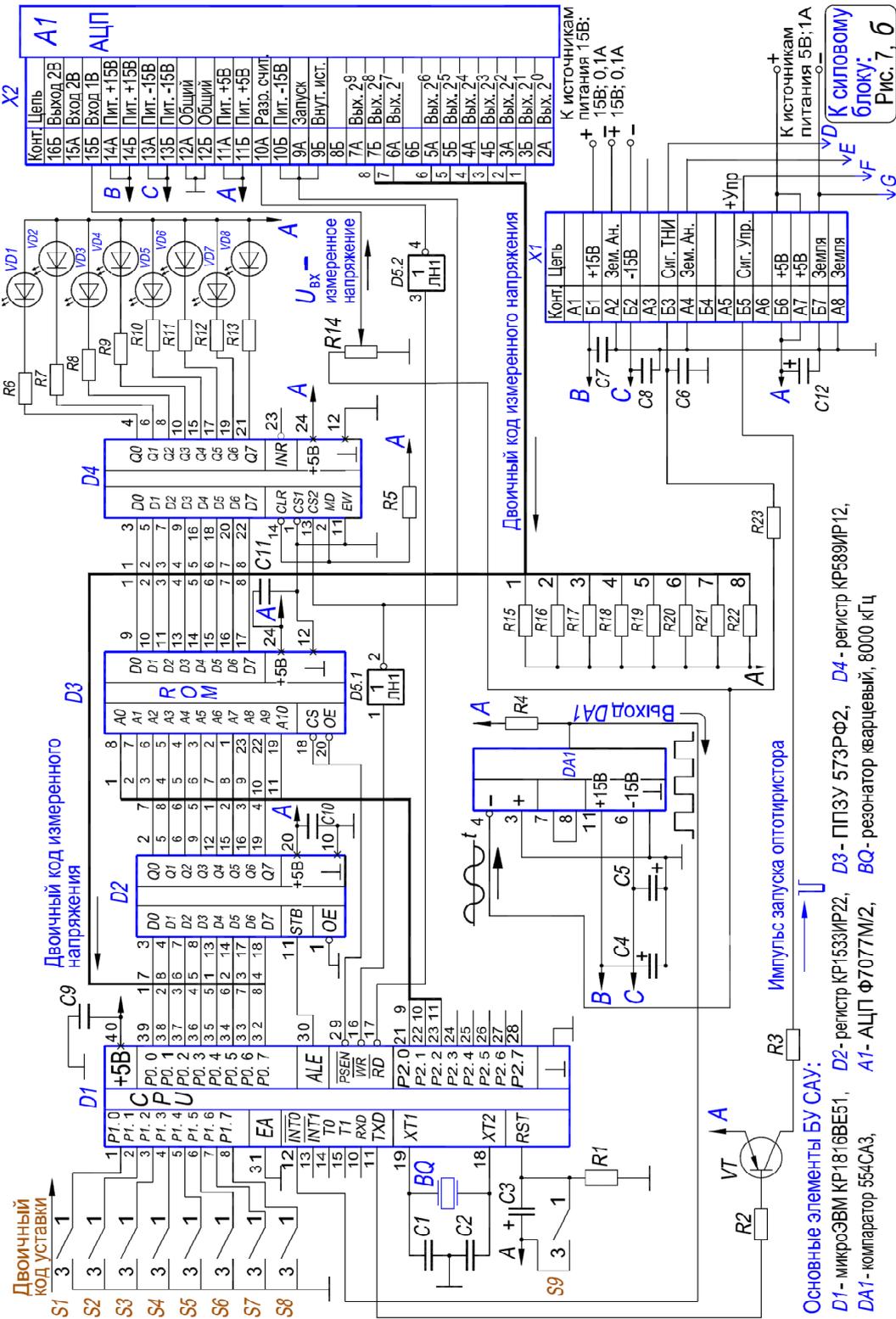


Рис. 7. Принципиальная электрическая схема САУ: блок управления – (а); силовой блок – (б)

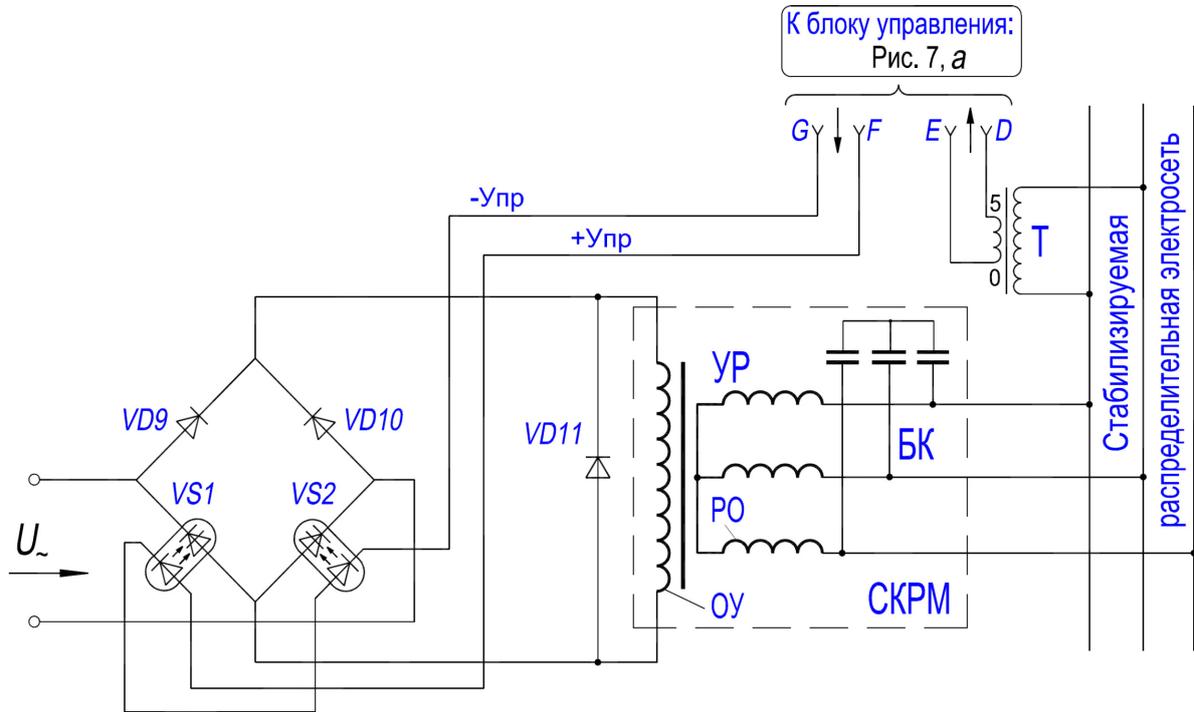


Рис. 7. Принципиальная электрическая схема САУ: блок управления – (а); силовой блок – (б)

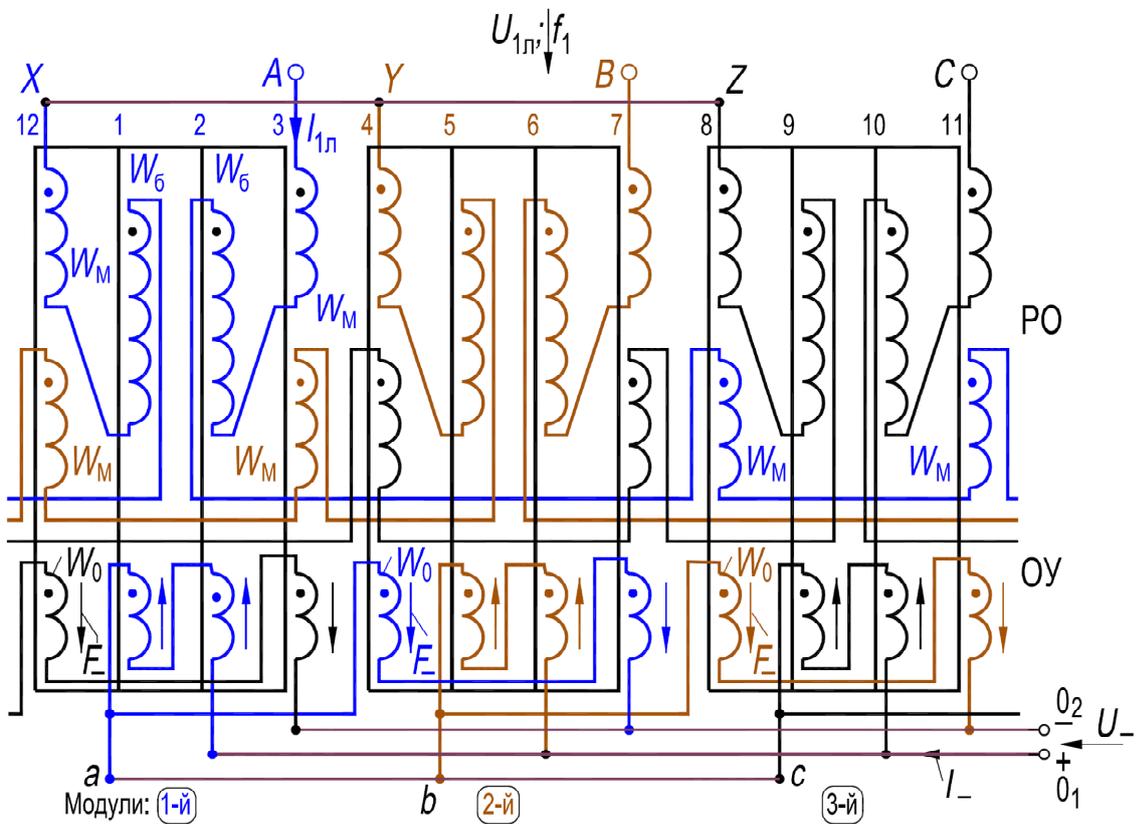


Рис. 8. Электромагнитная схема совмещенного УР

Активная часть реактора состоит из трех одинаковых модулей и характеризуется следующими основными данными: количество планарных магнитнонесвязанных сердечников – 3; количество стержней в одном сердечнике – 4; длина стержня – $10 \cdot 10^{-2}$ м; поперечное сечение стержня – $10 \cdot 10^{-4}$ м²; материал магнитопровода – сталь 3413, $0,35 \cdot 10^{-3}$ м; сочленение стержней и ярм – шихтованное внахлест; число витков в катушках трехфазной рабочей обмотки – $W_6 = 142$, $W_m = 82$; число витков в катушке обмотки управления – $W_0 = 197$; число эффективных витков в катушках обмоток – 2; диаметр провода – $1,32 \cdot 10^{-6}$ м².

Испытания проводились при отсутствии и наличии перемычки, соединяющей зажимы a, b, c обмотки управления ОУ (Рис. 8). При разомкнутых зажимах a, b, c обмотка содержит три параллельные ветви, а при замкнутых зажимах – шесть ветвей. Кроме того, в процессе исследования катушки обмотки соединялись встречно-последовательно, при этом в ОУ отсутствуют параллельные ветви. При разомкнутых зажимах a, b, c реализуется режим свободного намагничивания по 4 и 8-й гармоникам магнитного поля (СН_4_8), а при замкнутых – режим намагничивания по 2, 4, 8 и 10-й гармоникам магнитного поля (СН_2_4_8_10). При встречно-последовательном соединении катушек ОУ реализуется режим вынужденного намагничивания (ВН) по всем высшим гармоникам поля [Zabudsky, 2016].

Испытания проводились при отсутствии и наличии перемычки, соединяющей зажимы a, b, c обмотки управления ОУ (Рис. 8). При разомкнутых зажимах a, b, c обмотка содержит три параллельные ветви, а при замкнутых зажимах – шесть ветвей. Кроме того, в процессе исследования катушки обмотки соединялись встречно-последовательно, при этом в ОУ отсутствуют параллельные ветви. При разомкнутых зажимах a, b, c реализуется режим свободного намагничивания по 4 и 8-й гармоникам магнитного поля (СН_4_8), а при замкнутых – режим намагничивания по 2, 4, 8 и 10-й гармоникам магнитного поля (СН_2_4_8_10). При встречно-последовательном соединении катушек ОУ реализуется режим вынужденного намагничивания (ВН) по всем высшим гармоникам поля [Zabudsky, 2016].

На Рис. 9 представлены ВАХ реактора $U_{1л} = f(I_{1л}, I_-)$ для режимов ВН, СН_4_8 и СН_2_4_8_10. Как следует из экспериментальных зависимостей при $U_{1л} = 380$ В и токе подмагничивания $I_- = 6$ А диапазон регулирования реактора в режиме СН_2_4_8_10 ~ на 35 % больше, чем в режиме ВН и ~ на 21 % больше, чем в режиме СН_4_8. Режим СН_2_4_8_10 выгодно отличается не только увеличенным диапазоном регулирования тока $I_{1л}$, но и улучшенной формой кривой его, а также сниженными магнитными потерями.

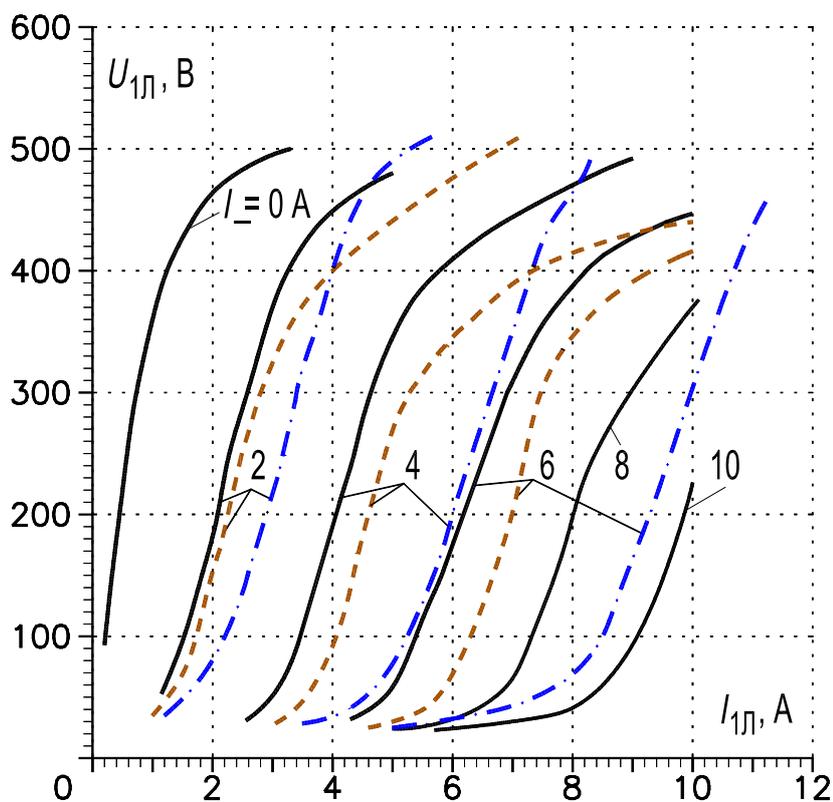
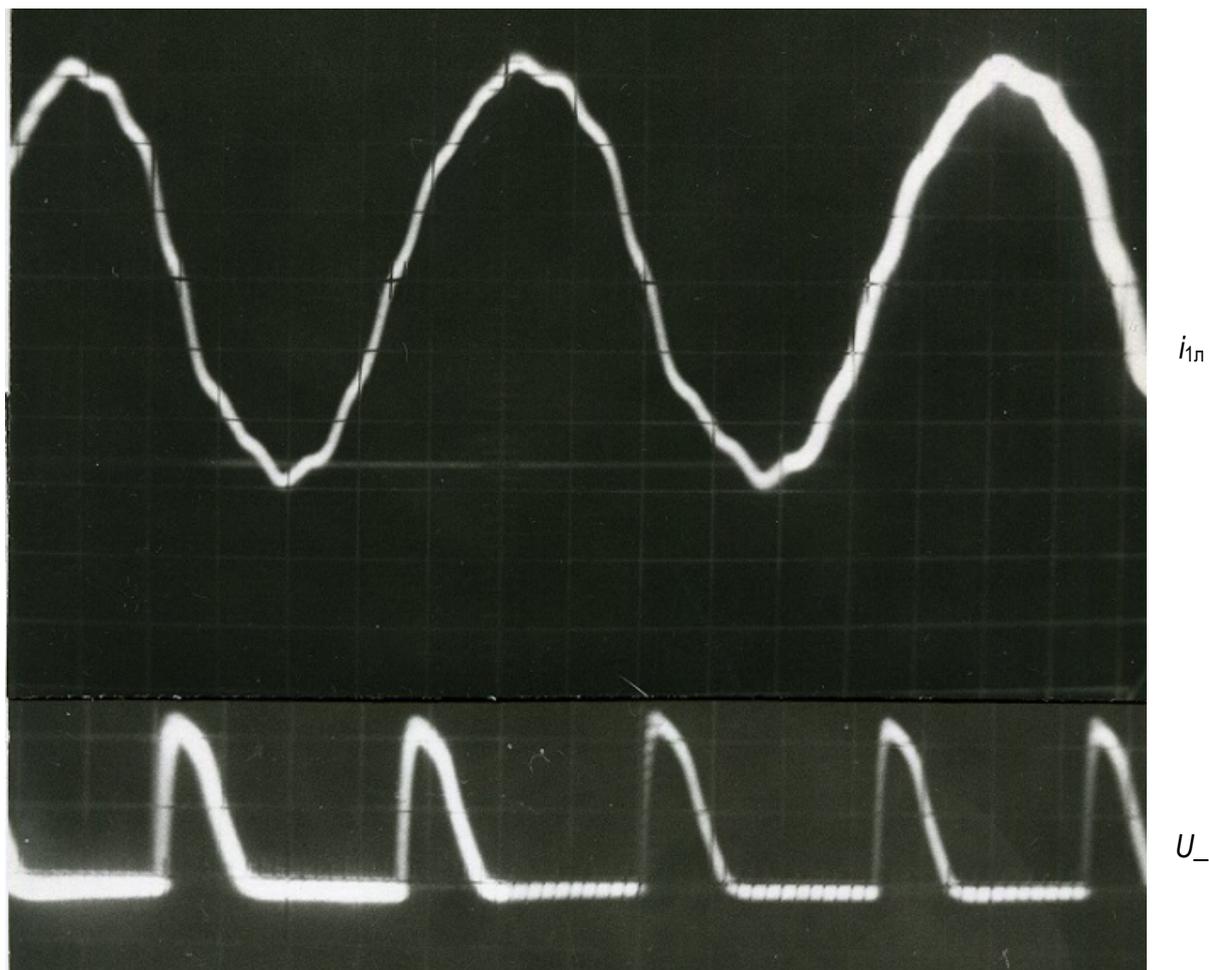


Рис. 9. Вольт-амперные характеристики реактора в режимах намагничивания:
 вынужденное — — —; свободное намагничивание по 4 и 8-й гармоникам поля — — — —;
 свободное намагничивание по 2, 4, 8 и 10-й гармоникам поля — · · · · ·

Рассмотренный реактор испытан совместно с разработанной и изготовленной микропроцессорной САУ стабилизацией напряжения электросети (Рис. 2). В опыте на обмотку управления подавалось выпрямленное напряжение $U_{\text{в}}$, которое снималось с управляемого выпрямительного моста. Величина выпрямленного напряжения регулировалась автоматически за счет изменения угла открытия оптотиристоров ТО325-12,5, включенных в мост.

Рассмотренный реактор испытан совместно с разработанной и изготовленной микропроцессорной САУ стабилизацией напряжения электросети (Рис. 2). В опыте на обмотку управления подавалось выпрямленное напряжение $U_{\text{в}}$, которое снималось с управляемого выпрямительного моста. Величина выпрямленного напряжения регулировалась автоматически за счет изменения угла открытия оптотиристоров ТО325-12,5, включенных в мост.

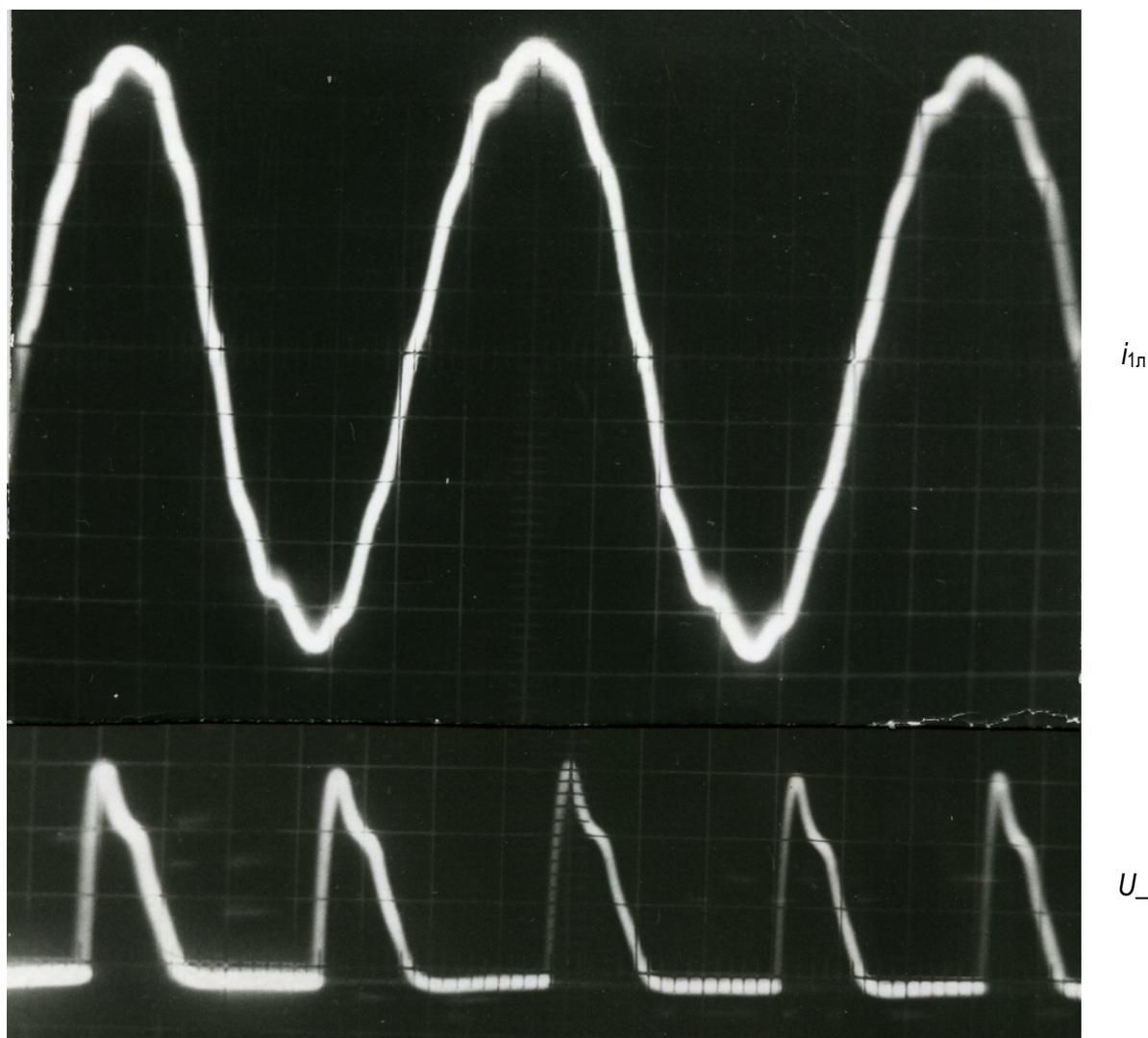
На Рис. 10,а показаны осциллограммы тока $I_{1л}$, протекающего в рабочей обмотке РО реактора, и, выпрямленного напряжения U_{-} , подаваемого на обмотку управления ОУ при значении стабилизируемого напряжения $U_{1л} = 220$ В. При этом ток в РО составлял 4 А, ток подмагничивания в ОУ – 8,3 А, а угол открытия оптотиристоров – $\sim 120^{\circ}$.



а)

Рис. 10 (а). Осциллограммы тока $i_{1л}$, протекающего в РО, и выпрямленного напряжения U_{-} , питающего ОУ реактора, соответствующие углу открытия оптотиристоров 120°

На Рис. 10,б показаны осциллограммы этих же величин, но при токе в РО 5,9 А, токе подмагничивания 12,2 А и при неизменном значении (220 В) стабилизируемого напряжения, но при меньшем значении угла открытия оптотиристоров – $\sim 111^{\circ}$. Испытания показали устойчивость стабилизации напряжения.



б)

Рис. 10. Осциллограммы тока $i_{л}$, протекающего в РО, и выпрямленного напряжения U_{-} , питающего ОУ реактора, соответствующие углу открытия оптотиристоров: 111°

Выводы

Разработана микропроцессорная САУ, которая совместно со статическим компенсатором реактивной мощности предназначена для поддержания напряжения в электроэнергетической системе в месте его установки на заданном уровне. Одновременно со стабилизацией напряжения снижаются потери мощности в электроэнергетической системе, и улучшается режим работы электропотребителей.

Эффективность программно-аппаратного комплекса определяется устойчивостью стабилизации напряжения в точке установки СКРМ, быстротой отзывчивости его, универсальностью САУ и др.

САУ может быть использована при соответствующей коррекции для управления режимами дальних высоковольтных ЛЭП с реализацией регулирования по отклонению от передаваемой мощности, а также для автоматического управления другими объектами электроэнергетики и электромеханики.

Библиография

- [Zabudsky, 2003] Забудский Е.И. Совмещенные регулируемые электромагнитные реакторы. – Москва: Энергоатомиздат, 2003. – 436 с.
- [Stashin, 1990] Сташин В.В., Урусов А.В., Мологонцева О.Ф. Проектирование цифровых устройств на однокристальных микроконтроллерах. М: Энергоатомиздат. 1990. – 236 с.
- [Aleksandrov, 1993] Александров Г.Н. Система передач из бассейна Амазонки в промышленно развитые районы Бразилии // Изв. РАН. Энергетика и транспорт. 1993. №5 С.63–72.
- [Zabudsky, 1994] Забудский Е.И., Амброзевич А.С. Микропроцессорная система управления совмещенным реактором-трансформатором. Тезисы докладов научно-техн. конференции "Создание комплексов электротехнического оборудования, высоковольтной, преобразовательной, сильноточной и полупроводниковой техники". Москва, ВЭИ, 1994. С.284--287.
- [Zabudsky, 1994] Three-phase mathed reactors transformers for electric power engineering // 9th International Power System Conference // Proceedings, volum 2, St. Petersburg, 4–7 July 1994. p.618–629.
- [Zabudsky, 2016] Забудский Е.И. Режимы намагничивания управляемых электромагнитных реакторов // Proceedings of the XVIth International Conference «Electromechanics, Electrotechnology, Electromaterials and Components». ICEEE-2016 Abstracts. – Moscow: «Знак». 19-24 September 2016. С.171–172.

Приложение

Структурные диаграммы ассемблер-программ.

Рис. П1, лист 1. Структурная диаграмма основной программы

Рис. П1, лист 2. Структурная диаграмма основной программы

Рис. П.1, лист 3. Структурная диаграмма основной программы

Рис. П2. Структурная диаграмма подпрограммы обработки прерываний, инициируемых при переполнении таймера

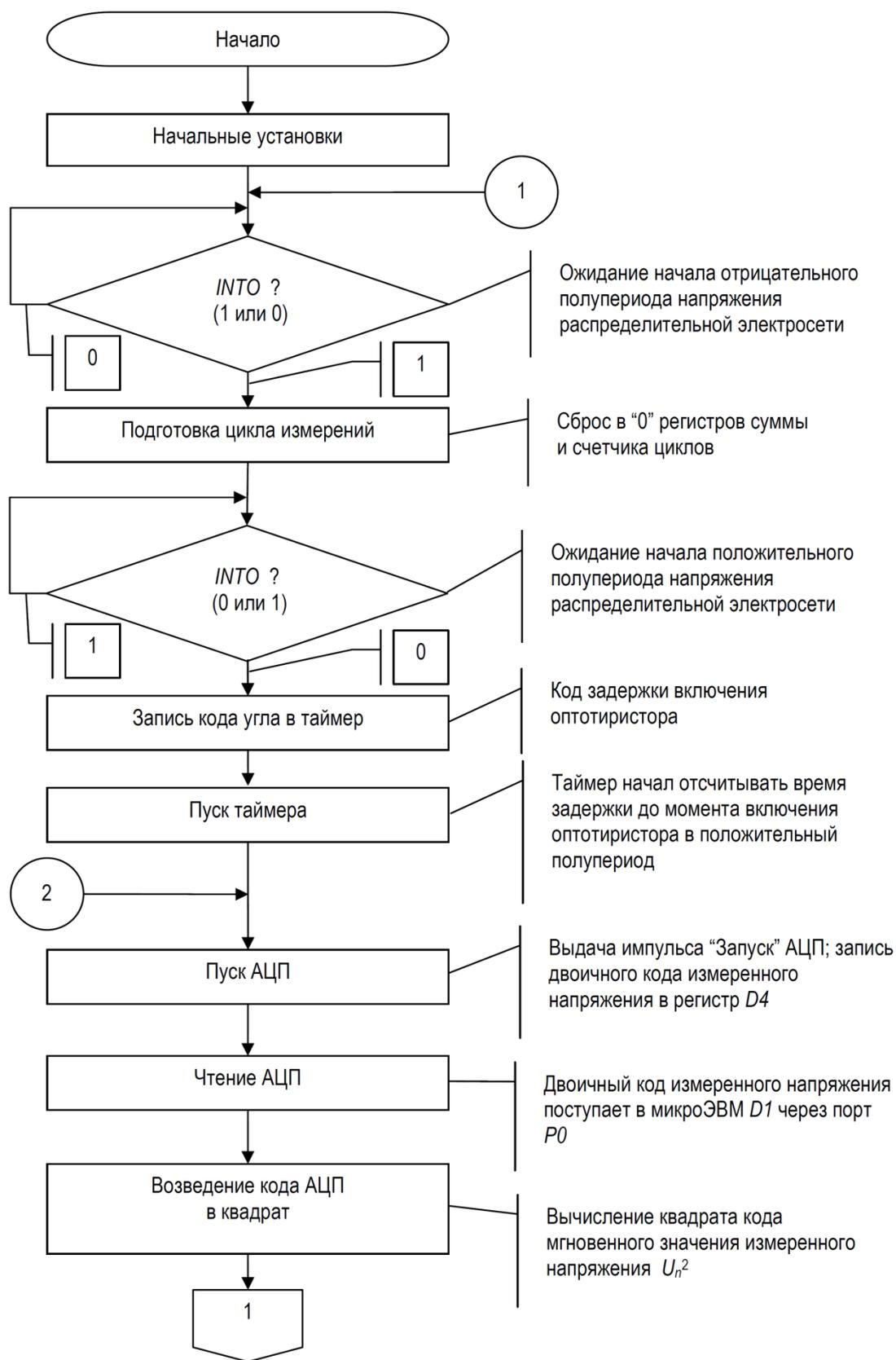


Рис. П1, лист 1. Структурная диаграмма основной программы

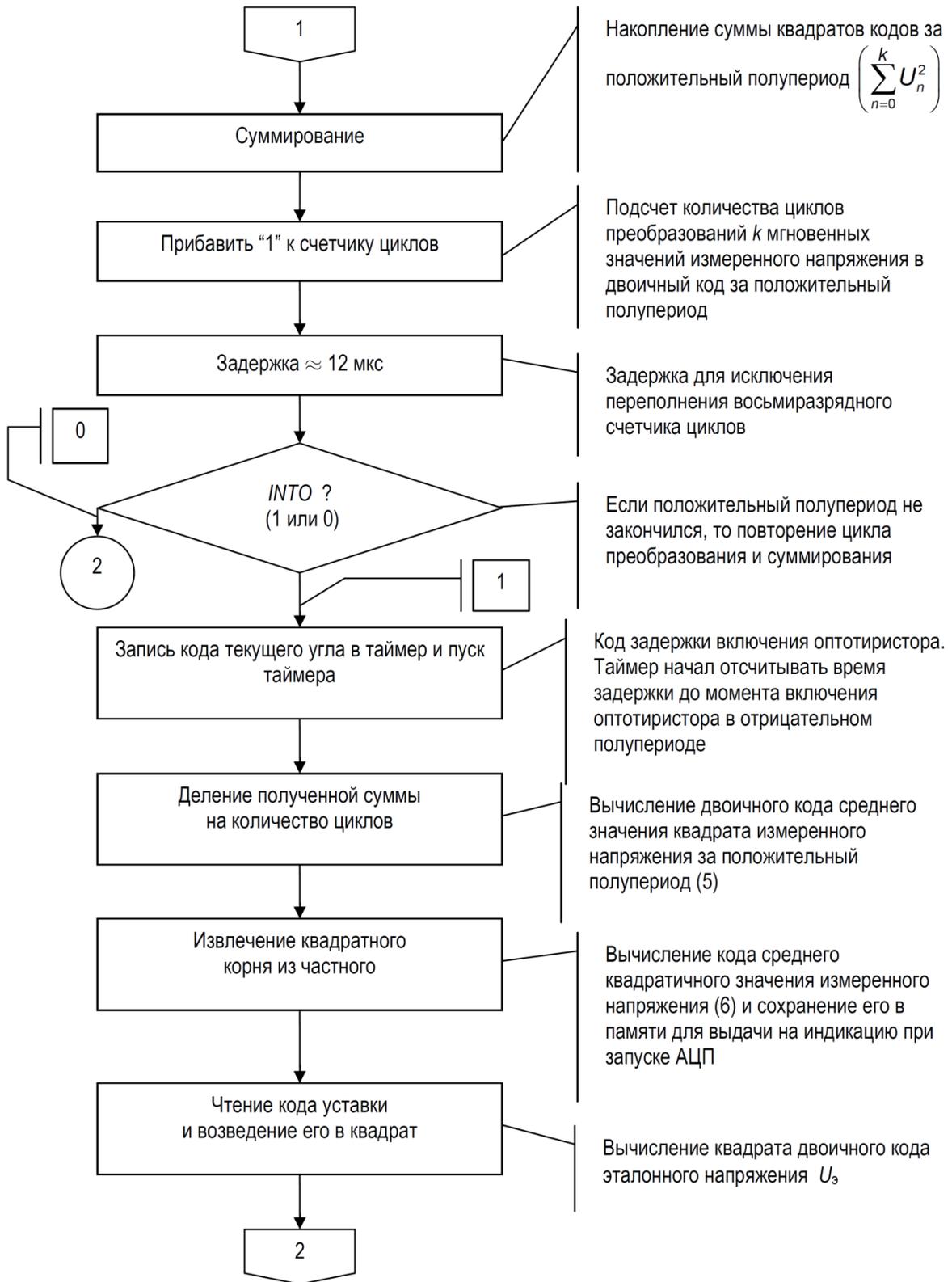


Рис. П1, лист 2. Структурная диаграмма основной программы

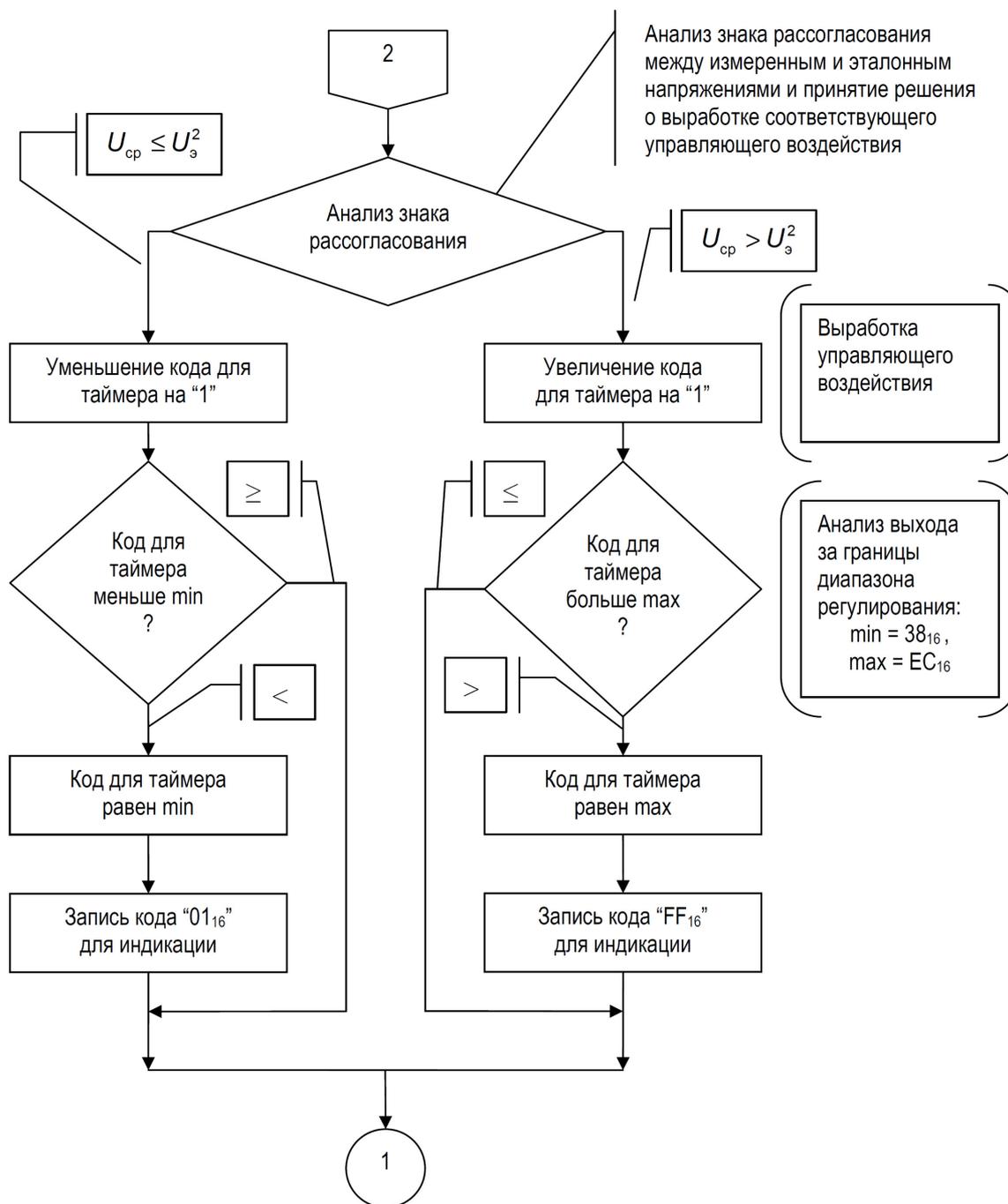


Рис. П.1, лист 3. Структурная диаграмма основной программы

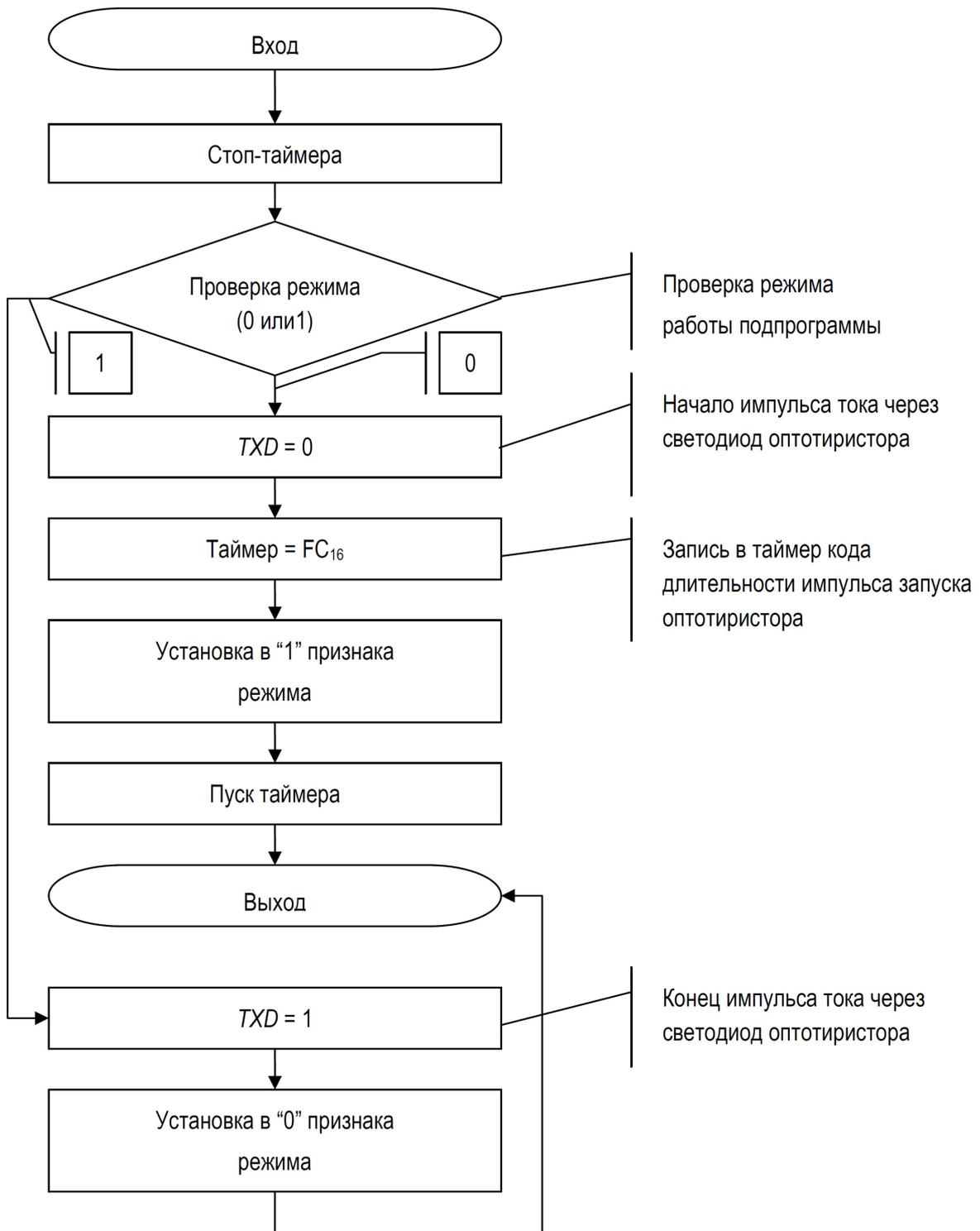


Рис. П2. Структурная диаграмма подпрограммы обработки прерываний, инициируемых при переполнении таймера

Информация об авторе



Евгений Забудский – профессор, доктор техн. наук; Российский университет дружбы народов, департамент Механики и мехатроники Инженерной академии; 117198, г. Москва, ул. Миклухо-Маклая, д. 6; e-mail: zabudskiy_ei@pfur.ru

Направление научной работы: Регулируемые электромагнитные устройства для управления режимами электроэнергетических систем; компьютерные и информационные технологии в образовании; web-сайт <http://zabudsky.ru>

Voltage Regulation of Distributed System Based on Single-Chip Microcomputer

Evgeny Zabudsky

Abstract: Significant growth of power plants unit capacity, large extent of power grids, presence of weak interconnections and uneven load schedule increase are the development features of power industry, that necessitate an improvement of energy facilities automatic control systems (ACS). Operating experience of electric power facilities microprocessor automatic control systems showed their advantages in comparison with the devices made from discrete components, for both service functions and reliability. A microprocessor control system in a combination with a static VAR compensator has been designed to regulate and stabilize voltage in an electric power system at a predetermined level. The ACS usage also leads to a reduce of power losses in the stabilized electric power system and an improvement in electrical load operations.

Keywords: single-chip microcomputer, automatic control system, controlled reactor, distribution electric network

TABLE OF CONTENT

<i>Нелинейная схема компромиссов в многокритериальных задачах принятия решений</i>	
Альберт Воронин, Юрий Зиатдинов, Игорь Варламов.....	3
<i>Обнаружение киберугроз с помощью анализа социальных сетей</i>	
Людмила Кириченко, Тамара Радвилова, Алексей Барановский.....	23
<i>Сопоставление семантических информационных ресурсов web на основе онтологического анализа</i>	
Юлия Рогушина.....	49
<i>Стабилизация напряжения распределительной электросети на основе однокристалльной микроЭВМ</i>	
Евгений Забудский.....	72
Table of content.....	100