

## TESTING THE ALGORITHM OF THE METHOD «CATERPILLAR-SSA» FOR REESTABLISHING OF TIME SERIES

*The base algorithm of the «Caterpillar-SSA» method is considered and tested.*

*Keywords: trend allocation, finding of the periodicals, noise elimination, decomposition of time series to components.*

© Вохмянин С. В., 2010

УДК 681.3

Е. А. Энгель

## ИСПОЛЬЗОВАНИЕ ИНТЕЛЛЕКТУАЛЬНЫХ АЛГОРИТМОВ ДЛЯ ОБРАБОТКИ ТЕКСТОВОЙ ИНФОРМАЦИИ\*

*С целью создания программной системы для адаптивного текстового реферирования разработаны и реализованы в виде отдельного модуля интеллектуальные алгоритмы автоматического определения жанра текста. Модуль позволяет нормализовать 45 статистических параметров: лексических, синтаксических, позиционных и дискурсивных; группировать гетерогенные параметры с помощью алгоритма K-средних; выполнять факторный анализ; ранжировать параметры, существенные для идентификации научного жанра, публицистики и беллетристики, посредством двух алгоритмов.*

*Ключевые слова: обработка текстовой информации, интеллектуальные алгоритмы, алгоритм K-средних.*

В течение прошлого десятилетия автоматическое определение жанра текста стало важной проблемой, исследованной в пределах такой научной области, как обработка естественного языка. Будучи интересной с теоретической точки зрения, задача определения жанра тесно связана с развитием информационного поиска цифровых библиотек и реферирования. Автоматическую идентификацию интернет-жанров можно считать отдельной предметной областью, которая обращается к реальной проблеме информационной перегрузки и играет существенную роль в улучшении часто неадекватных результатов работы поисковых машин.

Включение модуля автоматического определения жанра текста позволяет оптимизировать и повысить эффективность системы текстового реферирования. Стимулом для создания такого модуля стали результаты оценки эффективности следующих систем автоматического реферирования: Event Tracking Summarizer, Subject Search Summarizer, Copernic Summarizer и Open Text Summarizer. Программный продукт Event Tracking Summarizer, специально разработанный для обработки беллетристики, оказался эффективнее других систем автоматического реферирования в среднем на 15 % для беллетристики и менее эффективным для других жанров. Следовательно, возникает необходимость в создании адаптивной системы текстового реферирования на основе алгоритмов, оптимизированных для конкретного текстового жанра.

У любой NLP-системы есть модуль предварительной обработки, который в зависимости от текстовых задач обработки выполняет лексическое и синтаксическое разложение, стемминг, аннотацию и синтаксический парсинг. Результатом предварительной обработки является модель объекта, которая отражает лингвистические характеристики входного текста, например слов, фраз, предложений, параграфов. Далее лингвистические характеристики ранжируются, в результате чего получается список текстовых параметров. Параметры с самыми высокими весами затем сравниваются с эталонными моделями, хранящимися в лингвистической базе данных. Входной текстовый жанр идентифицируется в зависимости от степени соответствия между распределением параметров в этом тексте и в одной из эталонных моделей. На следующей стадии система применяет алгоритмы реферирования, оптимизированные для данного жанра (рис. 1).

Предметом данной статьи является задача определения жанра текста; алгоритмы реферирования выходят за рамки данной публикации.

**Параметры, идентифицирующие жанр.** Идентификация жанра текста основана на анализе набора параметров, являющихся лингвистическими признаками с назначенными весами, т. е. некоторыми числовыми значениями, отражающими его важность для данного текста. Следовательно, процесс идентификации жанра включает две

\* Работа выполнена в рамках Федеральной целевой программы «Научные и научно-педагогические кадры инновационной России на 2009–2013 гг.», Госконтракт 02.740.11.0663.

задачи: генерацию набора лингвистических признаков и разработку алгоритмов ранжирования признаков.

В работе М. Сантини (URL: [http://www.nltg.brighton.ac.uk/home/Marina.Santini/06\\_04\\_d\\_MSantini\\_JADT\\_2006.pdf](http://www.nltg.brighton.ac.uk/home/Marina.Santini/06_04_d_MSantini_JADT_2006.pdf)) приводятся убедительные доказательства отсутствия характеристических или, по крайней мере, общепринятых наборов текстовых параметров для идентификации жанра, однако общая тенденция состоит в том, чтобы использовать содержательные, а не служебные термины. Это объясняется тем, что число служебных слов ограничено, они встречаются в любом тексте и различия в их распределении обеспечивают надежные данные о различиях жанра. М. Сантини использует такие параметры, как признаки части речи, 50 самых общих слов, символы пунктуации, теги HTML, длина веб-страницы, которые определяют аспекты жанра (URL: [http://www.itri.brighton.ac.uk/~Marina.Santini/MSantini\\_PhD\\_Thesis.zip](http://www.itri.brighton.ac.uk/~Marina.Santini/MSantini_PhD_Thesis.zip)). Эти параметры лежат в основе различных методов ранжирования и обеспечивают достаточно эффективную идентификацию интернет-жанров.

Предлагаемый в данной статье модуль отличается от существующих систем распознавания интернет-жанров в следующих аспектах: входной файл для автоматического реферирования имеет формат (.txt) и особенности расположения документа (например, теги HTML) незначительны; главная идея состоит в том, чтобы разработать адаптивные алгоритмы реферирования, основанные на

параметрах, существенных для данного жанра, именно поэтому параметры содержания также должны быть учтены; служебные слова считают словами остановки и игнорируются во время текстового реферирования.

Несмотря на то что в лингвистике существуют различные интерпретации классификации жанра, большинство ученых соглашается с наличием ряда основных различий между беллетристикой и научной литературой (URL: <http://www.evartist.narod.ru/text14/17.htm>). Базируясь на этих различиях, нами были составлены 45 параметров, разбитых на четыре группы (табл. 1).

Лексические параметры включают средние частоты имен собственных и различные типы местоимений в тексте (A/T) и предложениях (A/S). Гипотеза нашего исследования заключалась в том, что средние частоты имен собственных в беллетристике намного выше, чем в научной литературе, потому что беллетристика сосредоточена на названиях персонажей, мест и событий, которые составляют главные элементы мира воображения. Это же относится и к местоимениям, потому что они служат заменами для имен собственных и существительных и для фраз, составленных из имен собственных, например таких, как Джон Смит. Неопределенные местоимения более типичны для научной литературы, так как они являются проявлениями безличного повествования. Нами также полагалось, что соединительные слова, которые выражают причинно-следственные отношения, т. е. при-

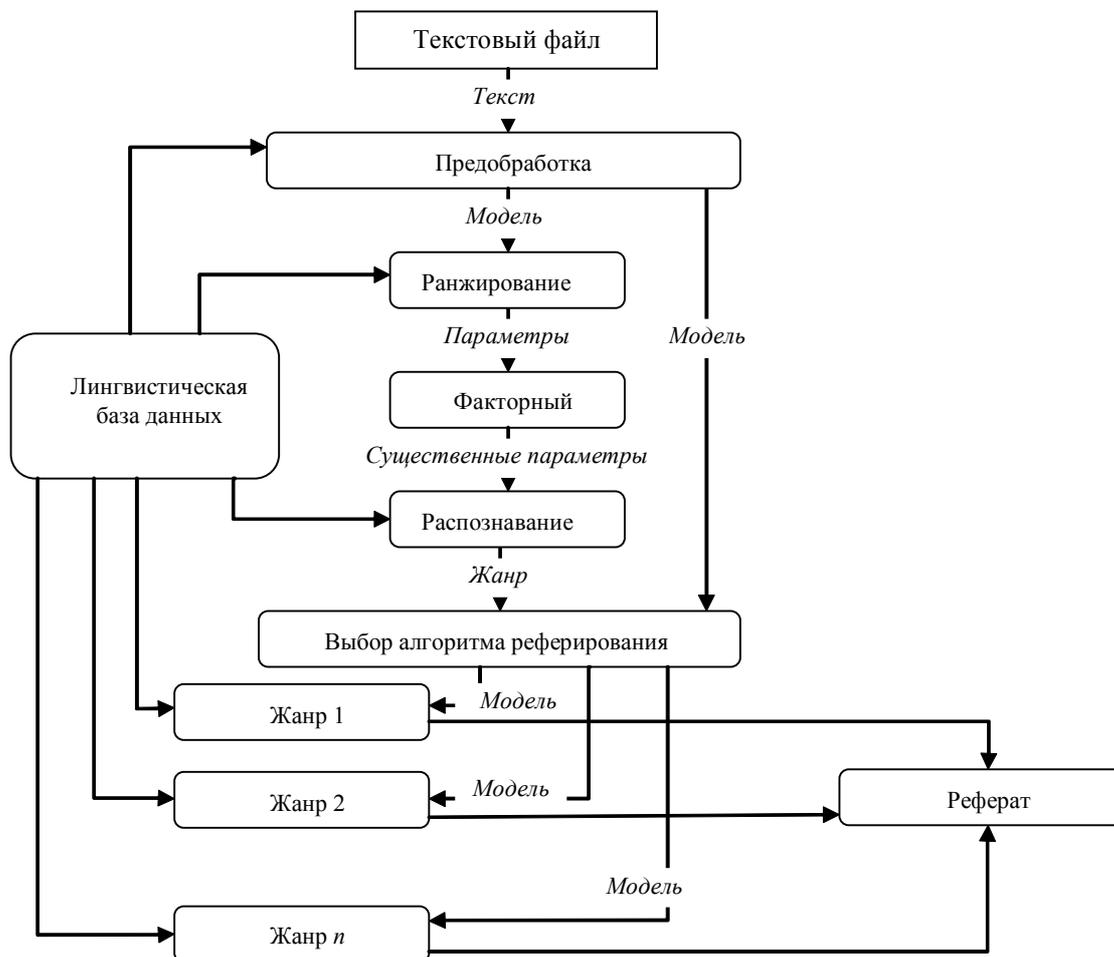


Рис. 1. Обобщенная схема адаптивной текстовой системы реферирования

чину, сроки и условия (например, *потому что, с тех пор, если* и т. д.), более характерны для научной литературы, так как они выражают ее логическую природу, в то время как соединительные слова, которые выражают контраст, время и дополнение, используются чаще в беллетристике, чтобы изобразить временную последовательность событий или обеспечить сравнительные описания. Средняя длина прямой речи в словах (AL/W) также предполагалась значительно более высокой в беллетристике.

Позиционные параметры для указанных лексических единиц были сформулированы согласно среднему значению и принципам различия, применяемым в информационной поисковой теории (URL: <http://www.criticalreading.com/fiction.htm#linkf>). Эти принципы гласят, что термины, которые встречаются в первой части текста, важнее найденных только во второй или третьей части, и термины, используемые в начале параграфа, более существенны, чем остальные. Поскольку местоимения выступают в качестве замены имен собственных и существительных, их частота должна быть выше во второй части текста, а не в начале. A/S\_1P, A/T\_1/3, A/T\_2/3 обозначают среднюю частоту имен собственных и существительных в первом предложении параграфа, 1/3 и 2/3 частях текста соответственно.

Также очевидны важные различия между следующими статистическими параметрами беллетристики и научной литературы: средней длиной предложения в словах (ASL/W), дисперсией длины предложения в словах (DSL/W), средней длиной параграфа в словах (APL/W), средней длиной параграфа в предложениях (APL/S), дисперсией длины параграфа в словах (DPL/W) и дисперсией длины параграфа в предложениях (DPL/S). Дисперсия здесь рассматривается как среднее абсолютное отклонение, представленное формулой

$$D = \frac{1}{N} \sum_{i=1}^N |x_i - M|,$$

где  $M$  – среднее значение выборки размера  $N$ .

Представленные в табл. 1 параметры позволяют обнаруживать содержательные термины во время последующего текстового реферирования. В теории информационного поиска и реферирования самые важные содержательные термины – существительные. Личные, взаимные, рефлексивные, вопросительные местоимения заменяют существительные, которые могут быть обнаружены с помощью решения анафоры; демонстративные и притяжательные местоимения предварительно изменяют существительные, служа, таким образом, их индикаторами.

**Экспериментальные результаты.** Чтобы проверить адекватность параметров, нами было проведено три ряда экспериментов, основанных на соответствии тестового лингвистического корпуса справочному корпусу. В качестве справочного лингвистического корпуса использовались 130 работ беллетристики, загруженных с сайта Gutenberg (URL: [http://www.gutenberg.org/wiki/Main\\_Page?Sess=25607a1aa5](http://www.gutenberg.org/wiki/Main_Page?Sess=25607a1aa5)). Для первого ряда экспериментов тестовый корпус состоял из работ одного автора. На основе представленных выше параметров справочный корпус был разделен на две, три, четыре и пять групп, чтобы узнать, находились ли работы, написанные одним и тем же автором, в пределах одной группы. Гипотеза была следующей: если параметры были выбраны адекватно, то такие работы должны принадлежать одной и той же группе, если же эти работы находятся в пределах различных групп, то это свидетельствует о несоответствии параметров. Во втором ряду экспериментов тестовый корпус был представлен статьями одного научного журнала. Статьи добавлялись к справочному корпусу, который был разделен на то же самое число групп. Для третьего ряда экспериментов тестовый корпус составляли статьи, взятые с одного новостного сайта.

Эксперименты показали, что когда число групп в справочном корпусе было от одной до четырех, тексты тестового корпуса находились в пределах той же самой группы. Когда число групп равнялось пяти, тестовый корпус

Таблица 1

Параметры для распознавания жанра текста

№	Тип параметра	Лингвистические единицы	Числовое значение
1–8	Лексический	Имена собственные, личные, демонстративные, взаимные, притяжательные, рефлексивные, неопределенные местоимения	A/T
9–16		Имена собственные, личные, демонстративные, взаимные, притяжательные, рефлексивные, неопределенные местоимения	A/S
17		Прямая речь	AL/W
18	Синтаксический	Фразы, составленные из имен собственных	A/T
19–26	Позиционный	Имена собственные, личные, демонстративные, взаимные, притяжательные, рефлексивные, неопределенные местоимения	A/S_1P
27		Имена собственные	A/T_1/3
28–34		Личные, указательные, взаимные, притяжательные, рефлексивные, неопределенные местоимения	A/T_2/3
35–39	Дискурсивный	Соединительные слова, которые выражают причинно-следственные и последовательно-временные отношения	A/T
40–45	Статистический	Предложение	ASL/W
		Предложение	DSL/W
		Параграф	APL/W
		Параграф	DPL/W
		Параграф	APL/S
		Параграф	DPL/S

классифицировался неправильно. Табл. 2, 3, 4 предоставляют информацию о текстах, которые классифицировались неправильно, и о параметрах, которые способствовали неправильной классификации тестового корпуса.

Необходимо отметить, что при числе групп, равно пяти, только один текст тестового корпуса классифицировался неправильно. Дальнейшее увеличение числа групп ухудшает классификацию.

Чтобы провести эксперименты, были составлены списки местоимений и дискурсивных терминов, создано оригинальное программное обеспечение и разработаны следующие алгоритмы:

- алгоритм распознавания имен собственных;
- алгоритм распознавания фраз с именами собственными;
- алгоритм распознавания первых предложений в параграфах;

Таблица 4.

**Экспериментальные результаты, полученные по тестовому корпусу из 10 статей каждого новостного сайта (число групп – 5, размер справочного корпуса – 140)**

Неправильно классифицированный текст (заголовок)	Автор	Параметры, способствующие неправильной классификации	Число текстов автора
The Forsyte Saga, Volume II. Indian Summer of a Forsyte and In Chancery	John Galsworthy	Средняя длина прямой речи (больше справочной для данного жанра)	4
Sister Carrie	Theodore Dreiser	Средняя длина параграфа в словах (меньше справочной для данного жанра)	4
Ersatz Eternal	A. E. Van Vogt	Средняя длина параграфа и дисперсия предложений (меньше справочной для данного жанра)	3
Liza of Lambeth	W. Somerset Maugham	Средняя длина параграфа в словах (меньше справочной для данного жанра). Дисперсия длины параграфа в предложениях и словах (меньше справочной для данного жанра)	3
The Happy Prince	Oscar Wilde	Средняя длина параграфа в предложениях (меньше справочной для данного жанра). Дисперсия длины параграфа в предложениях и словах (меньше справочной для данного жанра)	3
The Prince and The Pauper, Complete	Mark Twain (Samuel Clemens)	Среднее число личных местоимений в тесте (меньше справочного для данного жанра)	4
The Pickwick Papers	Charles Dickens	Среднее число личных местоимений в предложениях (меньше справочного для данного жанра)	11
The Jungle Tales of Tarzan	Edgar Rice Burroughs	Среднее число личных и указательных местоимений в тесте (меньше справочного для данного жанра). Средняя длина прямой речи (меньше справочной для данного жанра). Средняя длина параграфа в предложениях (больше справочной для данного жанра)	8

Таблица 2

**Экспериментальные результаты, полученные по тестовому корпусу, составленному из работ одного и того же автора (число групп – 5, размер справочного корпуса – 130)**

Текст	Автор	Параметры, способствующие неправильной классификации
Language Learning & Technology. 2003. Vol. 7, № 1. P. 46–70 (URL: <a href="http://llt.msu.edu/vol7num1/greenfield/default.html">http://llt.msu.edu/vol7num1/greenfield/default.html</a> ) Collaborative E-mail Exchange for Teaching Secondary Esl: a case Study in Hong Kong	Roseanne Greenfield	Среднее число имен собственных в предложениях и в первых предложениях параграфов (больше справочного для данного жанра)

Таблица 3

**Экспериментальные результаты, полученные по тестовому корпусу, включающему по 6 статей каждого научного журнала (число групп – 5, размер справочного корпуса – 136)**

Analysis: Taliban talks strategy attracts backers (URL: <a href="http://edition.cnn.com/2008/WORLD/asiapcf/10/28/analysis.taliban/index.html">http://edition.cnn.com/2008/WORLD/asiapcf/10/28/analysis.taliban/index.html</a> )	CNN	Среднее число указательных местоимений в предложениях и в первых предложениях параграфов (больше справочного для данного жанра)
---	-----	---

- алгоритм распознавания прямой речи;
- алгоритм разбиения текста на части в следующих отношениях: 1/3, 2/3.

Для текстового разложения использовались программное обеспечение tokenizer и splitter (URL: <http://www.criticalreading.com/fiction.htm#linkf>).

Распознавание имен собственных выполнялось согласно следующему алгоритму:

1) выбираются все символы с начальной заглавной буквой. Если символ не находится в начальной позиции в предложении, то он маркируется как имя собственное и попадает в список имен собственных;

2) проверяются все символы, которые открывают предложения. Если символ находится в списке имен собственных, то его маркируют как имя собственное.

Для ранжирования параметров и кластеризации корпуса необходимы более сложные алгоритмы. Главной проблемой является гетерогенный характер параметров: у некоторых из них были вероятностные оценки, в то время как другие представляли средние значения. Чтобы разрешить эту проблему, равно как и проблему кластеризации корпуса, использовался метод  $K$ -средних, группирующий  $N$  объектов признакового пространства в постоянное число  $k$  кластеров,  $k < N$  (URL: [http://www.cs.duke.edu/courses/spring\\_07/cps296.2/papers/kMeans-socg.pdf](http://www.cs.duke.edu/courses/spring_07/cps296.2/papers/kMeans-socg.pdf)).

Предположим, что корпус включает  $N$  текстов и для каждого из них вычисляются  $M$  средних значений параметров текста. Для их нормализации используется линейная нормализация относительно минимального и максимального значения, которая может быть представлена формулой

$$X_{\text{norm}} = (X - X_{\text{min}}) / (X_{\text{max}} - X_{\text{min}}),$$

где  $X$  – исходное среднее значение параметра;  $X_{\text{max}}$  и  $X_{\text{min}}$  – максимальное и минимальное значение параметра соответственно.

После нормализации тексты характеризуются набором параметров, представляющих точки в многомерном единичном кубе. Далее осуществляется кластеризация полученного множества точек с помощью авторской программы, реализующей метод  $K$ -средних.

Также был разработан метод выделения существенных параметров, идентифицирующих конкретный жанр. Метод состоит из трех этапов.

*Первый этап* – составление обучающего и тестового лингвистических корпусов текстов, пропорционально представляющих каждый жанр. Тексты были отобраны из американского (ANC) и британского лингвистических корпусов (BNC) (табл. 5).

*Второй этап* – осуществление факторного анализа параметров за два шага.

Первый шаг формирует вес  $w_{1i}$ . Для распознавания жанра текста используются две нейросети бинарной классифика-

ции: выходной сигнал первой сети равен 1, если жанр текста беллетристика; в противном случае классификацию продолжает вторая нейросеть, выходной сигнал которой равен 1, если жанр текста научный, и 0, если жанр текста публицистика. Далее осуществляется анализ значимости нейросетевых входов, которые представляют собой числа в диапазоне от 0 до 1. Затем параметры ранжируются следующим образом: каждому параметру  $i$  присваивается вес  $w_{1i}$ , соответствующий значимости его как нейросетевого входа.

Второй шаг формирует вес  $w_{2i}$ . Для этого вычисляются средние значения параметров для каждого жанра. Далее определяется вес  $i$ -го параметра  $w_{2i}$ :

$$|X(G1) - X(G2)| / \max(X(G1), X(G2)), \quad (1)$$

где  $X(G1), X(G2)$  – среднее значение параметра  $i$  для первого и второго жанра соответственно. Полученный список параметров  $w_{2i}$  нормализуется.

*Третий этап* – отбор из каждого жанра текста  $n$  наиболее значимых параметров для идентификации жанра текста, осуществляемый по формуле

$$\frac{\sum_{i=1}^M w_{ji}}{M} = K, \quad w_{ji} \in S, \quad \text{если } w_{ji} \geq K, \quad (2)$$

где  $w_{ji}$  – вес  $i$ -го параметра, полученного на шаге;  $S$  – множество существенных параметров;  $M$  – общее число параметров текста;  $K$  – порог, т. е. параметры, веса которых превышают или равны  $K$ , являются наиболее значимыми при идентификации жанра текста. Наборы параметров, наиболее значимых для идентификации жанров текста (беллетристики, публицистики и научного жанра), представлены в табл. 6.

Анализ представленных выше данных подтверждает следующие гипотезы:

- соединительные слова, которые выражают временные отношения, важны для беллетристики;
- соединительные слова, которые выражают условия, важны для научных текстов;
- неопределенные местоимения важны для научной литературы.

Гипотеза о важности имен собственных для беллетристики была частично подтверждена. Оказалось, что распределение имен собственных еще более важно для газетных текстов.

Гипотеза о важности причинно-следственных связей для научных текстов не подтвердилась.

Экспериментальные результаты показали, что программная система распознала правильно 100 % научных текстов, 96 % газетных текстов (4 % из них были идентифицированы как научные), и 86,5 % текстов беллетристики (3,5 % были распознаны как научные). Средняя ошибка классификации составила 94,75 %. Система распознавания была оценена на текстах соответствующих жанров, взятых из BNC и ANC (табл. 7 и 8).

Таблица 5

Структура обучающего и тестового лингвистических корпусов текстов

Корпус	Беллетристика		Научный жанр		Публицистика	
	Кол-во текстов	Кол-во слов	Кол-во текстов	Кол-во слов	Кол-во текстов	Кол-во слов
BNC	100	3 187 732	50	1 573 825	50	1 428 367
ANC	70	2 865 623	50	1 205 912	50	1 346 843

В отличие от существующих исследований, нацеленных на адаптивные алгоритмы распознавания интернет-жанров, эта работа была вдохновлена идеей адаптивного текстового реферирования. Одна из лучших систем такого типа – IntellexerSummarizer (URL: [http://summarizer.intellexer.com/order\\_summarizer\\_standard.php](http://summarizer.intellexer.com/order_summarizer_standard.php)) – осуществляет автоматическое распознавание таких жанров, как об-

щее, доступное, научное, экономика и политика. Система представляет коммерческий продукт и не дает информации о причинах такого выбора жанров и алгоритмах для их распознавания.

Данное исследование было проведено на заранее определенных жанрах, т. е. в нем проанализированы параметры текстов, жанр которых уже определен в BNC и ANC.

Таблица 6

**Параметры, наиболее значимые для идентификации жанров текста**

Параметры	Формула	Вес
<i>Параметры, наиболее значимые для идентификации беллетристики</i>		
Наиболее значимые параметры (нейросетевой факторный анализ):		
притяжательные местоимения	A/T	0,832 398
соединительные слова, которые выражают временные отношения	A/T	0,385 037
имена собственные	A/T_1/3	0,356 929 3
рефлексивные местоимения	A/T	0,378 697
неопределенные местоимения	A/T	0,350 852
взаимные местоимения	A/T_2/3	0,349 781
Наиболее значимые параметры (метод средних):		
взаимные местоимения	A/S	0,525 304
<i>Параметры, наиболее значимые для идентификации научных текстов</i>		
Наиболее значимые параметры (нейросетевой факторный анализ):		
притяжательные местоимения	A/S_1P	0,811 473
длина предложений	ASL/W	0,584 083
соединительные слова, выражающие условия	A/T	0,421 059
Наиболее значимые параметры (метод средних):		
притяжательные местоимения	A/T_2/3	0,498 94
неопределенные местоимения	A/S_1P	0,657 554
имена собственные	A/S	0,719 157
относительные местоимения	A/S	0,432 064
<i>Параметры, наиболее значимые для идентификации публицистики</i>		
Наиболее значимые параметры (нейросетевой факторный анализ):		
взаимные местоимения	A/S	0,643 268
притяжательные местоимения	A/S	0,585 424
имена собственные	A/S_1P	0,582 093
фразы с именами собственными	A/T	0,543 224
указательные местоимения	A/S	0,398 675
Наиболее значимые параметры (метод средних):		
имена собственные	A/S_1P	0,422 16
имена собственные	A/S	0,345 877

Таблица 7

**Качество распознавания первой нейросети**

Показатель	Беллетристика	Nonfiction
Количество текстов	170	100
Количество текстов, распознанных правильно	147	100
Количество текстов, распознанных неправильно	23	0
Процент распознавания	86,5	100

Средняя ошибка классификации – 91,5 %.

Таблица 8

**Качество распознавания второй нейросети**

Показатель	Научный жанр	Публицистика
Количество текстов	100	100
Количество текстов, распознанных правильно	100	96
Количество текстов, распознанных неправильно	0	4
Процент распознавания	100	96

Средняя ошибка классификации – 98 %.

Как было отмечено выше, предложенная программная система делает ошибки в определении жанра и такие случаи особенно интересны. Например, если система признает текст как научный, хотя в BNC или ANC он определен как газетный, то было бы полезно проанализировать параметры, из-за которых система распознала текстовый жанр неправильно. Эти параметры могут отличать некоторые поджанры в пределах

данного жанра – так называемые d-жанры, что может наложиться на поджанры, которые различают в литературе, такие как научная фантастика, исторический роман, документ для обсуждения и т. д. Алгоритмы реферирования, оптимизированные для каждого из d-жанров, обеспечат необходимую гибкость – одно из важных современных требований для систем рассматриваемого типа.

E. A. Engel

## USING INTELLIGENT ALGORITHMS FOR THE TEXT INFORMATION PROCESSING

*For the purpose of adaptive text summarization the software for automatic genre recognition has been developed base on intelligent algorithms. The software allows creation of normalizing 45 statistic, lexical, syntactic, positional, and discursive parameters and clustering of heterogeneous parameters with the help of k-means algorithm, verification of the parameters, selection of the parameters highly salient for academic, newspaper, and fiction texts by means of two factor analysis algorithms.*

*Keywords: text information processing, intelligent algorithms, of k-means algorithm.*

© Энгель Е. А., 2010

УДК 621.771: 621.777

Н. Н. Загиров, А. А. Ковалева, Е. В. Иванов

## ТЕХНОЛОГИЯ ИЗГОТОВЛЕНИЯ ПРОВОЛОКИ С ВОЛОКНИСТОЙ СТРУКТУРОЙ ИЗ СТРУЖКИ СПЛАВА СИСТЕМЫ АЛЮМИНИЙ–МАГНИЙ–КРЕМНИЙ

*Предложена технологическая схема переработки отходов в виде сыпучей стружки сплава системы алюминий–магний–кремний в прутки и проволоку, основу которой составляют приемы порошковой металлургии. Выявлены характерные структурные особенности и дана оценка уровня механических свойств полученной проволоки.*

*Ключевые слова: сыпучая стружка, брикетирование, совмещенные прокатка-прессование, волочение, волокнистый материал, структура, механические свойства.*

В большинстве случаев, когда речь идет об использовании в различных отраслях промышленности проволоки, изготовленной из некомпактных металлических материалов, в качестве ее подразумевают сложный композиционный материал, состоящий из металлической оболочки – твердопластичного тела и порошкового сердечника – сыпучего не связного материала, представляющего смесь разнородных по составу и крупности отдельных частиц [1]. При обработке давлением металлическая оболочка находится в сложном взаимодействии с порошковым сердечником, вызывая некоторую совокупность сложного перемещения частиц порошка и их упруго-пластического взаимодействия друг с другом в массе порошковой шихты под действием внешней нагрузки.

По предлагаемой в работе технологии изготовления проволоки из сыпучей мелкой стружки (опилок) алюминиевого сплава АД31 помещению ее в металлическую оболочку не предусматривается, а сам процесс получе-

ния конечной продукции условно может быть разбит на две составляющие:

1) технологическую цепочку получения промежуточной заготовки для волочения, которая включает стадии подготовки стружки к компактированию, брикетирования ее, нагрева полученных брикетов под экструзию и горячей экструзии на пруток заданного диаметра;

2) технологический процесс изготовления проволоки, состоящий из многократно повторяющихся операций протягивания заготовки через волокнистые фильеры, а также ряда сопутствующих и вспомогательных операций.

Такой подход к переработке стружки, общая доля которой во всем вторичном сырье составляет достаточно внушительную цифру, обеспечивает, по сравнению с плавильным переделом, более высокий выход годного металла стружки. И кроме того, снижаются энергозатраты и вредные воздействия на окружающую среду, что является актуальной задачей для любого производства, независимо от его масштабов.