

ФОРМАЛИЗОВАННОЕ ОПИСАНИЕ ПРОЦЕССОВ АНАЛИЗА И СИНТЕЗА ДОКУМЕНТОВ В СФЕРЕ ИНФОРМАЦИОННОГО ПРОТИВОБОРСТВА

к.т.н. В.Л. Петров, к.т.н. К.И. Хударковский, С.В. Залкин
(представил проф. А.В. Королёв)

В статье предлагается методика формализованного представления документа, основанная на определении ядра документа (системы документа) и формирования его целевой функции.

Введение. Проблема информационной безопасности является важной компонентой в общей системе национальной безопасности государства. Изменение характера угроз национальной безопасности государства, связанной с процессом информатизации и новейшими информационными технологиями, позволяет считать, что информационная борьба на современном этапе развития общества приобретает самостоятельное и исключительное значение [1, 2].

При **анализе документов** в области информационного противоборства принятие решений, как правило, осуществляется в условиях неопределенности. Неопределенность может иметь различную природу, поэтому используются различные математические методы и приемы работы с неопределенными данными и знаниями [3 – 4]: вероятностные, нечетких множеств, многозначной логики, экспертного оценивания, кластерного анализа и другие. Так, использование топологических моделей понятийных знаний [5] и применение коэффициентов интеграции [6] дают лишь частичные результаты. Анализ литературы [7, 8] показывает, что все большее число задач распознавания решается с помощью нейронных сетей.

Целью данной статьи является разработка методики формализации документа и определение его информационной направленности.

Методика формализации документа. Исходной информацией для анализа является совокупность документов D , формирование которых служит результатом решения отдельных задач системой S , имеющей вид

$$D = [d_0, d_1, \dots, d_i, \dots, d_k], \quad (1)$$

где d_0 – основной документ, характеризующий функционирование системы S ; d_k – совокупность подчиненных документов, включающих набор сегментов (информативных параметров) и участвующих в формировании всей совокупности документов D ; $i = \overline{1, k}$ – номер подчиненного документа.

Для анализа документа d_0 в работе предлагается разделить его на сегменты, в общем случае, произвольным образом. Совокупность сегментов S_0 документа d_0 имеет вид:

$$S_0 = [S_{01}, S_{02}, \dots, S_{0j}, \dots, S_{0z}], \quad (2)$$

где $j = \overline{1, z}$ – номер сегмента документа d_0 .

Методом экспертной оценки в соответствии со шкалой относительной важности Т. Саати [4] определяются коэффициенты важности сегментов S_0 . Результаты экспертной оценки могут быть представлены в виде квадратной z -мерной матрицы (табл. 1). Элементом матрицы S_{0jj} является коэффициент важности сегмента в документе.

Таблица 1

Матрица коэффициентов важности

	S_{01}	S_{02}	...	S_{0j}	...	S_{0z}
S_{01}	1	s_{012}	...	S_{01j}	...	s_{01z}
S_{02}	s_{021}	1	...	S_{02j}	...	s_{02z}
...	1
S_{0j}	s_{0j1}	s_{0j2}	...	1	...	s_{0jz}
...	1	...
S_{0z}	s_{0z1}	s_{0z2}	...	S_{0zj}	...	1

Анализ разброса оценок, даваемых экспертами коэффициентам важности сегмента, характеризуемый величиной вариации, а также степень согласованности оценок выходят за рамки работы и в статье не рассматриваются.

Следующим этапом решения задач анализа документа является определение методом экспертной оценки совокупности информативных параметров A_0 всех сегментов документа d_0 . Под информативным параметром (структурно-логическим элементом) в данной работе понимается параметр, характеристика или словосочетание, характеризующие данную предметную область. Параметр и характеристика могут иметь как количественное, так и качественное описание. Совокупность информативных параметров (структурно-логических элементов) A_0 документа d_0 имеет вид

$$A_0 = [a_{01}, a_{02}, \dots, a_{0m}, \dots, a_{0M}], \quad (3)$$

где $m = \overline{1, M}$ – номер информативного параметра (элемента) в документе.

Результаты определения информативных параметров документа d_0 могут быть представлены в табл. 2, где «1» означает, что соответствующий параметр присутствует в сегменте документа, а «0» – отсутствие параметра в сегменте. С целью последующего выделения ядра документа предлагается определить нормированные коэффициенты связности сегментов документа d_0 по количеству совпадающих информативных параметров в сегментах:

$$\bar{K}_{\hat{a}}(S_{0j}, S_{0z}) = \sum_1^M \left[1 - (\hat{a}_{S_{0j}} - \hat{a}_{S_{0z}}) \right] / \sum_1^i \hat{a}_{S_{0z}} \Big|_{\hat{a}_{S_{0j}} = a_{S_{0z}}}, \quad (4)$$

где $a_{S_{0j}}, a_{S_{0z}}$ – совпадающие элементы соответствующих сегментов.

Таблица 2

Информативные признаки документа d_0

Инф. парам. Сегменты	a_{01}	a_{02}	...	a_{0m}	...	A_{0m}
S_{01}	1	1	...	0	...	0
S_{02}	0	0	...	1	...	1
...
S_{0j}	1	0	0
...
S_{0z}	1	0	1

Нормированные коэффициенты связности сегментов характеризуют внутригрупповое расстояние элементов от центра группы, нормированное по числу элементов в группе и используемое в кластер-анализе. Таким образом, каждый сегмент (кластер) характеризуется соответствующим коэффициентом связности. По результатам анализа сегментов составляется квадратная z -мерная матрица коэффициентов связности сегментов (табл. 3).

Таблица 3

Матрица коэффициентов связности сегментов

	S_{01}	S_{02}	...	S_{0j}	...	S_{0z}
S_{01}	1	K_{012}	...	K_{01j}	...	K_{01z}
S_{02}	$K_{св021}$	1	...	$K_{св02j}$
...	1
S_{0j}	1
...	1	...
S_{0z}	$K_{св0z1}$	1

Анализируя матрицу коэффициентов важности сегментов документа d_0 (табл. 1) и матрицу коэффициентов связности сегментов (табл. 3), определяем ядро документа. Под ядром документа в данной работе понимается сегмент (сегменты) с совпадающими максимальными значениями коэффициентов важности и связности, отличными от единицы. Тогда ядро документа d_0 будет иметь вид:

$$X_0 = \{a_{01}, \dots, a_{0N}\}, \quad (5)$$

где N – количество информативных параметров в ядре документа d_0 .

Для выделенного ядра целесообразно методом экспертного оцени-

вания определить уровень каждого информативного параметра. Полученные сглаженные оценки предлагается использовать в качестве весовых коэффициентов при информативных параметрах в целевой функции. Тогда целевую функцию документа можно представить в виде:

$$L = l_1 a_{01} + \dots + l_n a_{0n} + \dots + l_N a_{0N} \quad (6)$$

при $l_N \geq 0$; $n = \overline{1, N}$.

Целевая функция (6) характеризует информационную направленность документа d_0 . Предложенный методологический подход к определению ядра и целевой функции документа может быть распространен на всю совокупность документов D системы S .

Выводы. Предложенная методика на основе системного подхода с использованием методов экспертного оценивания, элементов кластерного анализа, теорий множеств и матриц позволяет в формализованном виде представлять результаты анализа документа (системы документов) и определять его информационную направленность.

ЛИТЕРАТУРА

1. Руснак І.С., Телелим В.М. Розвиток форм і способів ведення інформаційної боротьби на сучасному етапі // *Наука і оборона*. – 2000. – № 2. – С. 18 – 23.
2. Попов М.О., Лук'янець А.Г. До забезпечення воєнної безпеки в умовах загрози інформаційної війни // *Наука і оборона*. – 1999. – № 2. – С. 37 – 43.
3. Ларичев О.И., Мошкович Е.М. Качественные методы принятия решений. *Вербальный анализ решений*. – М.: Наука, 1996. – 208 с.
4. Саати Т. *Принятие решений: Метод анализа иерархий / Пер. с англ. В.Г. Вогнадзе*. – М.: Радио и связь, 1993. – 386 с.
5. Вороной М.Ф. Структурирование запросов к интеллектуальным системам при организации вывода на понятийных знаниях // *АСУиИТ*. – Х.: ХГПУ. – 1999. – Вып. 58. – С. 20 – 23.
6. Левыкин В.М., Евланов М.В. Формализованное описание процессов интеграции в организационных АСУ // *АСУиИТ*. – Х.: ХГПУ. – 1999. – Вып. 58. – С. 30 – 32.
7. Каллан Р. *Основные концепции нейронных сетей*. – М.: Вильямс, 2001. – 287 с.
8. Бондарев В.Н., Аде Ф.Г. *Искусственный интеллект*. – Севастополь: Изд-во Сев. НТУ, 2002. – 615 с.

Поступила 4.11.2003

ПЕТРОВ Вадим Лукьянович, к.т.н., доцент, проф. кафедры ХВУ. В 1978 году окончил ВИРТА ПВО. Область научных интересов – информационная борьба.

ХУДАРКОВСКИЙ Константин Игоревич, к.т.н., доцент, с.н.с. НДВ (НПВО). В 1989 году окончил Харьковское ВВКИУ РВ. Область научных интересов – электромагнитная совместимость радиоэлектронной аппаратуры и техническая защита информации.

ЗАЛКИН Сергей Владимирович, начальник НИО (НПВО) ХВУ. В 1984 году окончил ВПА им. В.И. Ленина. Область научных интересов – информационно-психологические операции.