

Секция «Вычислительная математика и кибернетика»

О теоретических свойствах новой схемы кодирования неструктурированной информации

Сабурова Мария Ивановна

Студент

*Московский государственный университет имени М.В. Ломоносова, Факультет
вычислительной математики и кибернетики, Москва, Россия*

E-mail: masha-saburova@yandex.ru

Исследуются теоретические свойства новой схемы кодирования неструктурированной информации. Примерами такой информации являются различные тексты (статьи, газетные заметки, интервью), которые часто приходится анализировать в социологических, политологических и других исследованиях. Следовательно, нужна методика кодирования, которая включает в себя схему и процедуры кодирования, а также под-держана программными средствами [1].

Предлагается неструктурированную информацию кодировать в виде объединения гетерогенных графов, т.е. вершины в графах соответствуют разным единицам анализа. При этом кодирование проводится в два этапа. Результатом первого этапа является множество утверждений специального вида. Затем происходит переход к графовому представлению.

В работе доказано, что предложенная формализация утверждений гарантирует получение всех семантически обоснованных структур графа. В рамках алгебраического подхода к интеллектуальному анализу данных можно сказать, что мы ввели некоторое пространство оценок и доказали комплекс теорем, показывающих полноту нашей модели [2].

Рассматриваемая технология позволяет от гетерогенных сетей переходить далее к гомогенным, т.е. таким, в которых вершины соответствуют одной единице анализа. В социологических исследованиях обычно рассматриваются именно такие виды графов, например, графы акторов и графы понятий. Примерами, распространенными в гуманитарных сферах, служат граф согласия и граф конфликтов. Поэтому было введено и исследовано отношение совпадения (согласованности) утверждений. Соответственно, рассмотрено отношение несовпадения утверждений, когда они противоречат друг другу, находятся в конфликте.

Литература

1. Киселев С.Л., Ермаков А.Е., Плешко В.В. Поиск фактов в тексте естественного языка на основе сетевых описаний // Труды Международной конференции Диалог'2004. - Москва, Наука, 2004;
2. Рудаков К.В. Полнота и универсальные ограничения в проблеме коррекции эвристических алгоритмов классификации // Кибернетика. 1987. № 3. С. 106-109.