

Система автоматического реферирования и определения тем научно-технических текстов

БАКИЕВА АЙГЕРИМ МУРАТОВНА

Новосибирский государственный университет (Новосибирск), Россия

e-mail: m_aigerim0707@mail.ru

БАТУРА ТАТЬЯНА ВИКТОРОВНА

Институт систем информатики имени А.П. Ершова СО РАН (Новосибирск), Россия

e-mail: tatiana.v.batura@gmail.com

Работа посвящена методам автоматического реферирования, которые используют представление текста в виде графа. Фрагменты текста, такие как слова, предложения, абзацы, представляются в виде узлов графа, а отношения между узлами, например, риторические отношения, обозначаются ребрами. Предлагаемый метод формирует текст реферата на основе наиболее значимых предложений исходного документа. Значимость предложений определяется в процессе риторического анализа, который выполняется с помощью словаря, содержащего дискурсивные маркеры. Реферат оригинального текста составляется из предложений, содержащих маркеры определенного типа и ключевые слова текста.