

Построение меры близости документов для тематической классификации авторефератов диссертаций

ЛЕОНОВА ЮЛИЯ ВИКТОРОВНА

Институт вычислительных технологий СО РАН (Новосибирск), Россия
e-mail: julia_leo@ngs.ru

ФЕДОТОВ АНАТОЛИЙ МИХАЙЛОВИЧ

Институт вычислительных технологий СО РАН (Новосибирск), Россия
e-mail: fedotov@sbras.ru

ФЕДОТОВА ОЛЬГА АНАТОЛЬЕВНА

Государственная публичная научно-техническая библиотека СО РАН (Новосибирск), Россия

В работе рассматривается метод тематической классификации авторефератов диссертаций. Для этого используется специально построенная мера близости документов, учитывающая специфику предметной области. В качестве шкал для определения меры предлагается брать характеристики структурных атрибутов описания авторефератов (научная новизна; положения, выносимые на защиту и т.п.). Значения весовых коэффициентов в формуле для вычисления меры близости определяются предполагаемой апостериорной достоверностью данных соответствующей шкалы.