

## СИСТЕМА АВТОМАТИЧЕСКОГО РЕФЕРИРОВАНИЯ ТЕКСТОВОЙ ИНФОРМАЦИИ

Коваленко А. С., Кашеев Л. Б.

Научный руководитель: канд. техн. наук, доцент Кашеев Л. Б.

Национальный технический университет «Харьковский политехнический институт» (61002, Харьков, ул. Кирпичева, 2, каф. Системного анализа и информационно-аналитических технологий, тел. (057) 707-66-54)

E-mail: [adbc@ukr.net](mailto:adbc@ukr.net)

Использование компьютеров дает возможность быстро получать и публиковать любую информацию. На сегодняшний день объемы передаваемой по Интернету информации выросли и продолжают расти дальше – по прогнозам Nokia Bell Labs в течение предстоящих четырех лет объем IP-трафика к 2022 году достигнет уровня в 330 эксабайт в месяц. Ежегодные темпы роста IP-трафика составят 25%, превышая человеческие возможности воспринимать и обрабатывать такое количество информации.

Задача реферирования документов является актуальной для любых хранилищ информации: от библиотек до интернет-порталов. Потребности в средствах автоматического реферирования испытывают корпоративные системы документооборота, поисковые машины и каталоги ресурсов Интернет, автоматизированные информационно-библиотечные системы, каналы вещания, службы рассылки новостей и многие другие [1]. Реферирование требуется и конкретному человеку, например, для быстрого ознакомления с интересующей его публикацией или с подборкой статей по определенной тематике.

Процесс автоматического реферирования текстовой информации позволяет заменить трудоемкий процесс выделения важной информации вручную. Формирование краткого смысла оригинальных текстов в виде рефератов в несколько раз повышает скорость анализа текстовых документов.

Автоматическое реферирование – извлечения наиболее важных сведений из одного или нескольких документов и генерация на их основе лаконичных и информационно-насыщенных отчетов. Существует два направления автоматического реферирования – квазиреферирование и краткое изложение содержания. Краткое изложение исходного материала основывается на выделении из текстов с помощью методов искусственного интеллекта и специальных информационных языков наиболее важной информации и порождении новых текстов, содержательно обобщающих первичные документы [2]. Квазиреферирование основано на экстракции из первичных документов с помощью определенных формальных признаков «наиболее информативных» фраз (фрагментов), совокупность которых образует некоторый квазиреферат.

1. Luhn H. The automatic creation of literature abstracts. In IBM Journal of Research and Development, Vol. 2(2), 1958. – P. 159–165.

2. Гинкул А.С. Сравнительный анализ существующих систем автоматического реферирования текста / А.С. Гинкул // Політ. сучасні проблеми науки – Київ, 2012. – С. 255.