

**Принцип нечеткой классификации древнерусских текстов (на материале
«Поучения» Луки Жидяты)**

Научный руководитель – Бородкин Леонид Иосифович

Харламова Ксения Михайловна

Выпускник (магистр)

Московский государственный университет имени М.В.Ломоносова, Исторический факультет, Кафедра исторической информатики, Москва, Россия

E-mail: kseniaharlamova@yandex.ru

В последние десятилетия существуют две тенденции в развитии методов анализа текста: с одной стороны, это преобладающий традиционный, текстологический подход к вопросам истории текста [4]; с другой — формальные методы, возникающие ввиду интереса исследователей к изучению разного рода сетей [1].

Надо сказать, что древнерусские тексты в процессе рукописного копирования подвергались тем или иным изменениям со стороны переписчиков. Это могла быть лексическая, стилистическая правка, вставки или пропуски фрагментов текста, предложений и слов, а также разного рода ошибки и описки, носящие бессознательный характер. В текстологическом исследовании при использовании формальных методов происходит сличение списков произведения, в результате чего выделяются различия между текстами списков, проводится их анализ, и, наконец, строится генеалогическое древо списков (так называемая стемма) или более общая формальная классификация списков, благодаря которой можно решить те или иные проблемы, связанные с атрибуцией (определением авторства произведения), уточнением датировок (от установления дат списков зависят даты редакции, а от тех и других — даты собственно произведений), выделением значимых этапов в истории текста и, наконец, реконструкцией архетипа произведения.

Большинство формальных методов основано на предположении, что чем меньше различий содержится в текстах пары списков, тем генеалогически ближе стоят эти списки. Причем некоторые методы классификации списков, в частности метод групп, предполагают, что у каждого списка есть только один протограф, все ошибки которого список наследует, и одинаковые ошибки не содержатся в списках, имеющих в качестве своих протографов независимые списки.

Однако тот или иной список может иметь черты необязательно лишь одной группы, а двух и более, текстуально близких, и, кроме того, в разных группах стеммы (древа списков) могут быть одни и те же ошибки; поэтому С. В. Шпирко был предложен метод анализа текстов, учитывающий эту проблему и имеющий в основе теорию нечетких множеств, — в нашей работе он и используется. По сравнению с методом групп он обладает следующими особенностями. Во-первых, в модель вводится понятие нечеткости, что позволяет учитывать силу связи между списками (на заданном уровне достоверности). Во-вторых, предложенный подход учитывает возможность наличия нескольких протографов у любого из списков. В-третьих, на основе данного подхода приводится формальный алгоритм нечеткой классификации. Нечеткая классификация разбивает множество реальных списков на классы и хронологически их упорядочивает. Списки, оказавшиеся в одном классе, являются эквивалентными (схожими). Можно сказать, что предложенный алгоритм минимизирует внутриклассовую «неоднородность» [5].

На основе попарного сличения списков строится матрица нечеткого отношения, в которую вносятся подсчитанные доли ошибок. Далее выделяются текстологически близкие группы, причем степень детализации получающейся стеммы регулируется специальным

внешним параметром — порогом уверенности, с которой списки объединяются в нечеткие классы и между которыми устанавливаются генеалого-преемственные связи [6].

Объектом исследования является «Поучение» Луки Жидяты — древнерусское произведение духовно-нравственного и назидательного характера, ставящее своей целью рассказать, как следует поступать, чтобы жизнь была мудрой и благочестивой, и дать нравственные ориентиры и примеры для подражания. Создание «Поучения» Луки Жидяты относят к XI веку. Текст сохранился в составе сборников конца XIV — XVII века в двух видах: летописном (9 списков) и четьем (10 списков). По внешним признакам тексты четьей группы отличаются от летописного вида и между собой объемом, характером вставок и сокращений. Памятник ввел в оборот Р. Ф. Тимковский по Летописи начала XVI века из собрания П. П. Бекетова, ныне утраченной. Текст перепечатали И. Е. Евсеев и С. Бугославский [2]. Последний в начале XX века привел разночтения по спискам произведения и составил классификацию списков, используя формально-текстологический подход, развитый в рамках филологической школы В. Н. Перетца. После него подобную работу провел В. В. Мильков в 2010-х годах, детальнее рассмотревший списки и предложивший иную, вторую классификацию, которая уточняла результаты, полученные С. Бугославским [3].

Перед нами же встает вопрос, какая из этих двух классификаций будет подтверждена после проведения анализа списков «Поучения» с использованием методов теории нечетких множеств и если будут выявлены противоречия получившейся классификации с двумя существующими (то есть ни одна из этих двух не будет подтверждена), то в чем будут состоять отличия третьей классификации. Меняя порог уверенности, можно выделять новые группы, уточняя их состав, но в целом представляется, что сильных расхождений с результатами традиционного анализа нет.

Источники и литература

- 1) Бородкин Л.М., Милов Л.В. О некоторых аспектах автоматизации текстологического исследования (Закон Судный людем) // Математические методы в историко-экономических и историко-культурных исследованиях / Ред. И.Д. Ковальченко. М., 1977. С. 235–280.
- 2) Бугославский С. Поучение еп. Луки Жидяты по рукописям XV–XVII вв. // ИОРЯС, 1914, т. 18, кн. 2, с. 196–237.
- 3) Дергачева И. В., Мильков В. В., Милькова С. В. Лука Жидята: святитель, писатель, мыслитель. М.: Мир философии, 2016. С. 416.
- 4) Лихачев Д. С. Текстология: на материале русской литературы X–XVII вв. // при участии А. А. Алексеева и А. Г. Боброва; РАН, Ин-т рус. лит. (Пушкинский Дом). СПб.: Алетейя, 2001. С. 758.
- 5) Шпирко С.В. Применение теории нечетких множеств к задаче генеалогической классификации в текстологическом исследовании // Историческая информатика: Информационные технологии и математические методы в исторических исследованиях и образовании. 2013. №3. С. 39–51.
- 6) Шпирко С.В. Методы текстологической генеалогической классификации: матрица близости vs матрица нечеткого отношения // Историческая информатика: Информационные технологии и математические методы в исторических исследованиях. 2016. № 3–4. С. 24–32.