

# АЛГОРИТМ ПОИСКА РОДСТВЕННЫХ СВЯЗЕЙ НА ОСНОВАНИИ ДАННЫХ ИЗ СОЦИАЛЬНЫХ СЕТЕЙ

И. К. Сечко

*Белорусский государственный университет, Минск;*

*igorsechko.fish@gmail.com;*

*науч. рук. – Д. В. Щегрикович, канд. физ.-мат. наук*

В данной статье рассматривается идея построения алгоритма поиска пересечения родоводов различных интернет-пользователей, призванного решить задачу поиска родственников третьего и более кругов родства. Алгоритм предлагает пользователям ввести информацию о своих прародителях, после чего введенные данные сверяются и на основании количественной оценки сходства алгоритм определяет, есть ли между пользователями родственная связь или нет. Также введена матрица родственных соответствий, при помощи которой устанавливаются родственные соотношения. В существующих генеалогических порталах не реализуются подобные функции, поэтому в данной статье предлагается создание специализированной социальной сети или дополнение существующих социальных сетей предлагаемым алгоритмом.

**Ключевые слова:** родственный круг (круг родства); матрица родственных соответствий; информационный вес; алгоритм пересечения родоводов.

В настоящее время существует множество сайтов, предлагающих пользователю составить свое генеалогическое древо: familyspace.ru, myheritage.com, analizfamilii.ru, vgd.ru, rosgenea.ru. Однако на данных сервисах не реализуется возможность поиска пересечений в родоводах различных пользователей и установления родственных соответствий между ними. Предлагается разработать приложение и создать специализированный сервис для реализации данных возможностей.

Для реализации данной задачи был проанализирован ряд схем родства. Наиболее удобна для установления родственных соответствий схема В. Бусыгина [1]. В данной схеме были выделены родственные круги (рис. 1). Для работы предлагаемого алгоритма необходима база первичных данных (рис. 2). Приложение на основании этой информации устанавливает родственные соответствия между пользователями. Места жительства прародителей пользователи выделяют на карте, т.к. наименования административно-территориальных единиц не редко изменяются. Для последующей обработки используются первые 7 признаков, поскольку они несут наибольшую информацию о человеке (рис. 3). Пусть два пользователя создали описания своих предков:  $list_i = \{x_{i1}, x_{i2}, x_{i3}, x_{i4}, x_{i5}, x_{i6}, x_{i7}\}$ ,  $list_j = \{x_{j1}, x_{j2}, x_{j3}, x_{j4}, x_{j5}, x_{j6}, x_{j7}\}$ , где  $x_1$  – имя,  $x_2$  – фамилия,  $x_3$  – отчество,  $x_4$  – годы жизни,  $x_5, x_6$  и  $x_7$  – места жительства 1, 2 и 3 соответственно и периоды проживания там.

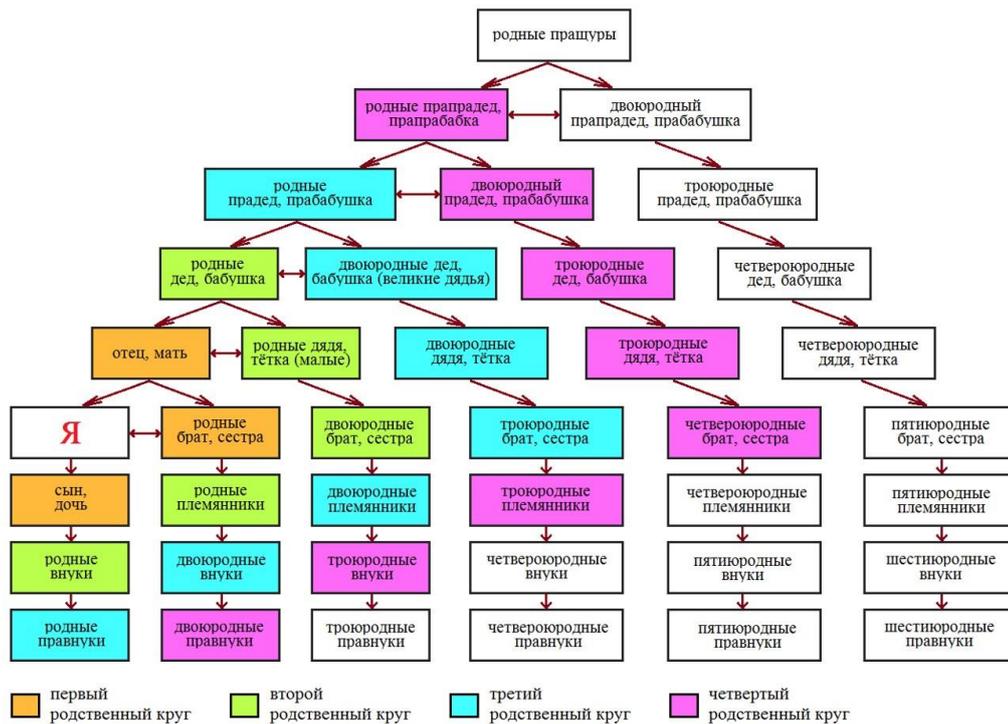


Рис. 1. Схема родства по В. Бусыгину с выделенными родственными кругами

Имя <input style="width: 100%;" type="text"/> Фамилия <input style="width: 100%;" type="text"/> Отчество <input style="width: 100%;" type="text"/> Годы жизни <input style="width: 50%;" type="text"/> - <input style="width: 50%;" type="text"/> Место жительства 1 <input style="width: 100%;" type="text"/> <small>указать на карте</small> Даты проживания <input style="width: 50%;" type="text"/> - <input style="width: 50%;" type="text"/> Место жительства 2 <input style="width: 100%;" type="text"/> <small>указать на карте</small> Даты проживания <input style="width: 50%;" type="text"/> - <input style="width: 50%;" type="text"/> Место жительства 3 <input style="width: 100%;" type="text"/> <small>указать на карте</small> Даты проживания <input style="width: 50%;" type="text"/> - <input style="width: 50%;" type="text"/>		Профессия <input style="width: 100%;" type="text"/> <small>Выберите</small> ▾ Периоды <input style="width: 50%;" type="text"/> - <input style="width: 50%;" type="text"/> Интересы <input style="width: 100%;" type="text"/> <small>Выберите</small> ▾ Школа <input style="width: 100%;" type="text"/> <small>Выберите</small> ▾ Учебное заведение <input style="width: 100%;" type="text"/> <small>Выберите</small> ▾ Служба в армии, участие в войнах: место <input style="width: 100%;" type="text"/> <small>Выберите</small> ▾ время <input style="width: 50%;" type="text"/> - <input style="width: 50%;" type="text"/> Участие в иных исторических событиях <input style="width: 100%;" type="text"/> <small>Выберите</small> ▾ Награды <input style="width: 100%;" type="text"/> <small>Выберите</small> ▾ Родство по схеме В. Бусыгина <input style="width: 100%;" type="text"/> <small>Выберите</small> ▾
---	--	--

Рис. 2. Структура базы первичных данных

Для определения сходства описываемых людей будем сравнивать признаки из каждого списка. При этом первые четыре сравниваются соответственно, а последние три – перебором. Перебором сравниваются места проживания людей. Для учета сходства введем информационные веса признаков, поскольку разные признаки имеют разную значимость информации.

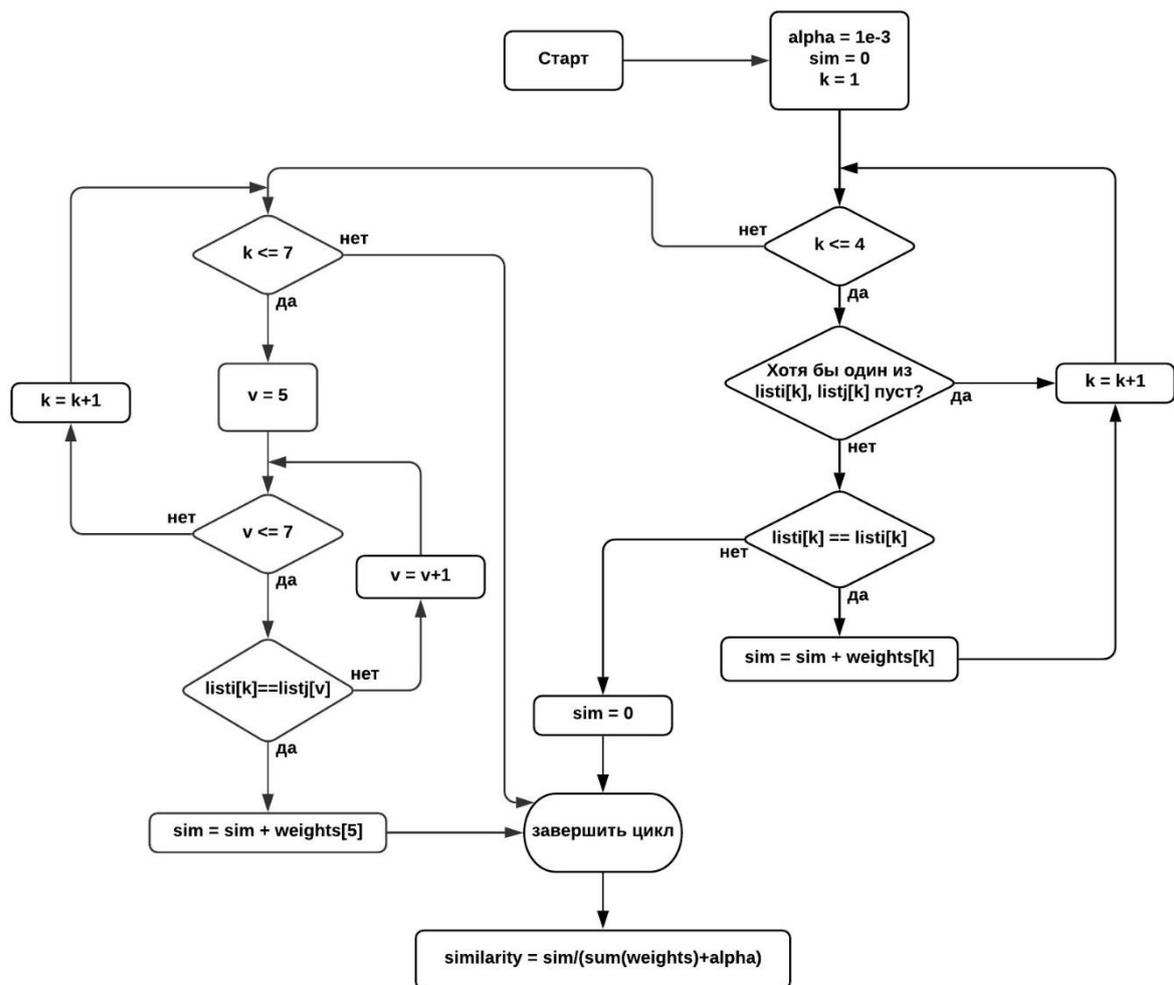


Рис. 3. Блок-схема основной части алгоритма поиска пересечения родоводов

Например, встретить разных людей с одним и тем же именем гораздо больше шансов, чем с одинаковой фамилией. Поэтому совпадение фамилии несет больший информационный вес.

Алгоритмом подсчитывается относительный информационный вес совпавших признаков. Если относительный вес превышает определенный порог, то это означает, что различные пользователи, вероятно, описали одного и того же прародителя. Далее алгоритм определяет, в каком родственном соответствии находятся данные пользователи, и приложение высылает им об этом сообщение.

Для установления родственных соответствий предлагается использовать матрицу (рис. 4). Если один и тот же прародитель первому пользователю приходится дедом (второй столбец), а второму – прадедом (третья строка), то первый пользователь по отношению ко второму является двоюродным дядей.

Для реализации задачи поиска родственников третьего и более кругов родства предлагается создание социальной сети «Родственники», которая кроме стандартных функций социальных сетей будет иметь специа-

лизированные функции: создание генеалогического древа, поиск пересечений генеалогических древ различных пользователей, рассылка пользователям сообщений о том, что они, вероятно, в  $n$ -ом круге родства друг с другом и в определенных родственных соотношениях.

	отец, мать	дед, бабушка	прадед, пра- бабушка	пра2дед, пра2бабка	пра3дед, пра3бабка
отец, мать	родные брат, сестра	родные племян- ники	двоюр. внуки	двоюр. пра- внуки	двоюр. пра2- внуки
дед, бабушка	родные дядя, тётка	двоюр. брат, сестра	двоюр. племян- ники	троюр. внуки	троюр. пра- внуки
прадед, пра- бабушка	двоюр. дядя, бабушка	двоюр. дядя, тётка	троюр. брат, сестра	троюр. племян- ники	4-юр. внуки
пра2дед, пра2бабка	двоюр. прадед, пра- бабушка	троюр. дед, бабушка	троюр. дядя, тётка	4-юр. брат, сестра	4-юр. племян- ники
пра3дед, пра3бабка	двоюр. пра2дед, пра2- бабушка	троюр. прадед, пра- бабушка	4-юр. дед, бабушка	4-юр. дядя, тётка	5-юр. брат, сестра

 первый  
родственный круг

 второй  
родственный круг

 третий  
родственный круг

 четвертый  
родственный круг

Рис. 4. Матрица родственных соответствий

Таким образом, в ходе работы над темой предложен способ оценки родства на основе матричного представления схемы Бусыгина. Также на схеме показано, какие именно ячейки входят в один и тот же круг родства. Предложена структура базы первичных данных для установления родственных соответствий между пользователями. Инновационным является указание места проживания прародителя не словесно-числовым адресом, а вводом области на карте. Предложен алгоритм для сравнения списков различных пользователей о своих прародителях и установления родственных соответствий между пользователями. На основании граф-схемы родственных связей В. Бусыгина нами составлена матрица родственных соответствий, по которой в алгоритме устанавливаются данные соответствия между пользователями. Предложена метрика, включающая информационные веса признаков, на основе которой алгоритм выдает сообщение пользователям о пересечении их родоводов.

#### Библиографические ссылки

1. Родословие: классификатор родственных связей [Электронный ресурс]. URL: <http://selvolostj.ru/ROD/rodstvo.html> (дата обращения: 24.04.2020).