

## ПРЕДСКАЗАНИЕ АТРИБУТОВ ПРОФИЛЯ ПОЛЬЗОВАТЕЛЯ СОЦИАЛЬНОЙ СЕТИ ПУТЕМ АНАЛИЗА СООБЩЕСТВ ГРАФА ЕГО БЛИЖАЙШЕГО ОКРУЖЕНИЯ

В.О. Чесноков

v.o.chesnokov@yandex.ru

МГТУ им. Н.Э. Баумана, Москва, Российская Федерация

---

### Аннотация

Определение скрытых атрибутов пользователя онлайн-социальных сетей — одна из важнейших проблем анализа социальных данных. Предложен подход к определению неуказанных или скрытых атрибутов пользователя путем анализа структуры графа его ближайшего окружения и атрибутов вершин этого графа. Выполнено сравнение предлагаемого метода с другими методами на выборках графов ближайшего окружения пользователей из социальных сетей Facebook, Twitter и ВКонтакте. Алгоритм показал высокие значения  $F$ -меры, точности и полноты по предсказыванию отдельных атрибутов профиля таких, как родной город или место обучения пользователя. Использование предлагаемого алгоритма с дополнительными источниками информации позволит с высокой точностью раскрыть личность анонимного пользователя социальной сети по его связям с другими пользователями

### Ключевые слова

Социальные сети, социальный граф, выделение сообществ, предсказание профилей

Поступила в редакцию 06.06.2016  
© МГТУ им. Н.Э. Баумана, 2017

---

*Работа выполнена при поддержке гранта РФФИ № 16-29-09517 офи\_м «Методы и алгоритмы выявления сообществ и организации информационного противоборства в социальных сетях на основе байесовских и теоретико-игровых подходов с использованием графовых и фрактальных моделей»*

**Введение.** Одной из важнейших особенностей социальных сетей является тот факт, что у каждой вершины есть набор атрибутов, характеризующих ее. Это могут быть как демографические характеристики (пол, возраст, образование и др.), так и сведения об интересах личности: любимые фильмы, увлечения, политические предпочтения и т. д. Данные атрибуты имеют важное значение при анализе общественного мнения, разработке рекомендательных систем, таргетированной рекламы.

Однако в реальных сетях полная информация об атрибутах пользователя зачастую бывает недоступна из-за разных причин: настройки приватности, потери данных при передаче и др. В связи с этим возникает задача предсказания отсутствующих или неуказанных атрибутов пользователя.

В настоящее время существует несколько методов решения данной проблемы. Большинство из них можно отнести к одному из четырех подходов:

- перенос атрибутов соседних вершин путем простого голосования или по другому правилу;
- методы, основанные на машинном обучении;
- методы, основанные на выделении сообществ;
- предсказание атрибутов по предпочтениям пользователя.

В качестве примеров методов, основанных на переносе атрибутов, можно привести работы [1, 2]. В [1] продемонстрирована возможность предсказания местоположения пользователя сети Twitter по мажоритарному признаку соседей. Предложенный метод показал довольно высокие (до 86 %) значения точности, но низкую (менее 20 %) полноту. В исследовании [2] был предложен алгоритм предсказания мест работы и обучения, местоположение пользователя, который дает 60...68 % правильных ответов на выборке графов ближайшего окружения пользователей из социальной сети LinkedIn.

Результаты предсказания в работах, основанных на машинном обучении, сильно зависят от качества обучающей выборки. В исследовании [3] был проведен сравнительный анализ нескольких методов машинного обучения для предсказания пола и возраста семи миллионов пользователей сотового оператора по данным о звонках и SMS-сообщениях. При обучении на 90 % выборки лучшие методы достигают высокого значения меры  $F_1$  для пола пользователя и возраста (0,85 и 0,72). В работе [4] представлен комплексный метод, который помимо атрибутов профиля использует наиболее информативные признаки из текстов сообщений пользователя. В ней достигнута высокая доля правильных ответов по предсказанию отдельных атрибутов пользователя (пола и возраста) — 89 %.

В исследовании [5] был применен подход, основанный на выделении сообществ. При условии, что известна информация о значении некоторого атрибута для 20 % пользователей, для остальных пользователей доля правильных ответов о значении этого атрибута составила 80 %. В качестве социального графа в работе использовались подграфы социальной сети Facebook профессорско-преподавательского состава и учащихся двух университетов. Однако, как отмечено в [2], исследование общего графа социальной сети для предсказания профилей неэффективно. Более разумный подход основан на анализе графа ближайшего окружения пользователя (*ego-network*).

Методы, основанные на анализе предпочтений пользователей требуют более разностороннюю информацию о пользователях сети, помимо их связей и атрибутов. Например, в работе [6] атрибуты профиля предсказываются по музыкальным предпочтениям, а в исследовании [7] — по положительным оценкам записей других пользователей («лайкам»). Алгоритм PGPI, представленный в [8], позволяет предсказать некоторые атрибуты профиля с долей правильных ответов более 90 %. При этом алгоритм использует ограниченное количество информации (так называемых фактов) и, кроме атрибутов пользователей, сведения об их членстве

и публикациях в виртуальных сообществах (группах), числе просмотров и положительных оценок публикаций.

Стоит отметить, что большинство алгоритмов разработаны с расчетом на то, что для обучения доступна вся информация о социальной сети и ее атрибутах [8]. Только некоторые алгоритмы предсказания профилей допускают частичное отсутствие информации об атрибутах пользователей [1, 5, 8].

**Постановка задачи.** Рассмотрим неориентированный невзвешенный граф  $G'(V', E')$  с диаметром, равным двум, и для которого есть такая вершина  $u$ , что

$$\forall v \in V', v \neq u \exists \{u, v\} \in E'.$$

Определим

$$V = V' \setminus \{u\};$$

$$E = E' \setminus \{\{u, v\} \mid v \in V\}.$$

Пусть у каждой вершины есть набор признаков (атрибутов) из множества признаков  $F$ , т. е. задана функция  $f: V' \rightarrow 2^F$ . Очевидно, что не всегда существует возможность получить полную информацию о признаках вершины (из-за ошибок при передаче данных, цензуры, при неизвестном или неуказанном состоянии и т. п.). Поэтому определим функцию  $f': V' \rightarrow 2^F$  получения признаков такую, что

$$\forall v \in V: f'(v) \subseteq f(v).$$

Пусть при этом данные о центральной вершине отсутствуют, т. е.  $f'(u) = \emptyset$ .

Задача предсказания профиля  $p$  состоит в том, чтобы максимально точно определить атрибуты центральной вершины:

$$p \subseteq F: \delta(f(u), p) \rightarrow \max,$$

где  $\delta$  — некоторая мера схожести двух множеств.

**Разработанный метод.** Предлагаемый метод основан на гипотезе о том, что сообщества образуются по модели присоединения (*affiliation model*) [9]. Гипотеза предполагает, что сообщества формируются из-за того, что вершины имеют один или несколько общих признаков, а не наоборот, т. е. атрибуты порождают сообщества.

Алгоритм выделения пересекающихся сообществ, предложенный в [10], для каждого выделенного сообщества предоставляет собой набор атрибутов, по которым оно было предположительно образовано. Он основан на переносе атрибутов. Изначально каждой вершине  $v$  из  $V$  ставится в соответствие пустое множество ключевых атрибутов  $K_v$ . На каждой итерации алгоритма проводится обход всех вершин и множества их ключевых атрибутов обновляются по следующему правилу: если квалифицированное большинство соседей вершины  $v$  имеет некий атрибут  $a$ , то он добавляется в множество ключевых атрибутов данной вершины  $K_v$ :

$$|\mathcal{N}_{v,a}| + |Q_{v,a}| > \max(2, \alpha |\mathcal{N}_v|) \Rightarrow K_v := K_v \cup \{a\},$$

где  $\alpha$  — доля, определяющая большинство для атрибута;

$$\begin{aligned} \mathcal{N}_v &= \{n \mid n \in V \wedge \{n, v\} \in E\}; \\ \mathcal{N}_{v,a} &= \{u \mid u \in \mathcal{N}_v \wedge a \in f'(u)\}; \\ Q_{v,a} &= \{u \mid u \in \mathcal{N}_v \wedge a \in K_u\}. \end{aligned}$$

Итерации прекращаются, когда на последнем шаге не было произведено ни одного изменения.

Затем вершины объединяются в сообщества по атрибутам из множества ключевых:

$$C_a = \{v \mid a \in K_v\},$$

т. е. каждому атрибуту ставится в соответствие множество вершин, у которых этот атрибут имеется в множестве ключевых атрибутов. Каждое из сообществ разбивается на компонентны связности, и тривиальные компоненты из одной или двух вершин отбрасываются. Наконец, сообщества, множества вершин которых совпадают, объединяются в одно, которому сопоставляется набор атрибутов всех сообществ. Таким образом, результатом работы алгоритма является список сообществ  $\mathcal{C}$ , каждому из которых поставлен в соответствие один или несколько атрибутов:

$$\mathcal{C} = \{(A, C) \mid A \subseteq F, C \subseteq V\}.$$

Поскольку центральная вершина состоит во всех сообществах графа ее ближайшего окружения, то она, скорее всего, будет иметь все атрибуты, связанные с этими сообществами. Предлагаемый метод предсказания профиля  $p$  центральной вершины заключается в объединении всех атрибутов сообществ, полученных алгоритмом:

$$p = \bigcup_{(A,C) \in \mathcal{C}} A.$$

Таким образом, данный метод объединяет два распространенных подхода к решению проблемы: выделение сообществ и перенос атрибутов. В работе [10] особое внимание уделено выделению сообществ в условиях, когда часть информации об атрибутах вершин недоступна, поэтому можно утверждать, что предложенный подход, основанный на данном алгоритме, также будет толерантен к частичному отсутствию атрибутов вершин.

Как показали эксперименты, не все атрибуты профиля могут быть предсказаны на основе информации о структуре графа и атрибутах вершин. Примером таких атрибутов может служить имя пользователя или пол. Поэтому перед применением алгоритма необходимо отбросить те атрибуты, которые невозможно предсказать.

**Метод оценки качества предсказания.** Точность предсказания предложенного метода была сравнена с несколькими известными подходами, использующими информацию о связях и атрибутах пользователей. В сравнении не рас-

смаивались методы, основанные на машинном обучении, так как оно не всегда применимо, и методы, которые используют сведения о предпочтениях пользователей, поскольку они требуют более разнородных данных, получение которых может быть затруднительным.

Первый простейший подход основан на переносе атрибутов голосованием и заключается в объединении наиболее часто встречающихся атрибутов соседей вершины. Если доля вершин, имеющих некоторый атрибут, превышает установленный порог, то считается, что и центральная вершина имеет данный атрибут.

Второй подход — анализ сообществ графа ближайшего окружения пользователя. Для каждого сообщества выбирается наиболее частый атрибут или набор атрибутов, которые присутствуют у большинства вершин в сообществе. Большинство вершин определяется порогом. В качестве алгоритмов выделения сообществ были выбраны два распространенных метода: максимизация модулярности [11] и Infomap [12]. Кроме того, были использованы три метода, которые также опираются на гипотезу об образовании сообществ по модели присоединения: AGM-fit [9], BigCLAM [13] и CESNA [14].

Третий подход основан на информации о связи между атрибутами вершин и сообществами. Алгоритм CESNA [14] предоставляет веса для каждой пары атрибут–сообщество. Чем больше вес, тем больше вероятность того, что сообщество было образовано по этому атрибуту. Значимые веса также определялись порогом.

Для тестирования подходов были использованы два набора данных от Stanford Network Analysis Project [15]: графы ближайшего окружения пользователей Facebook и Twitter. Их характеристики приведены в табл. 1. В данных наборах предоставлена информация об эталонных сообществах, поэтому во втором подходе к предсказанию профиля помимо сообществ, выделенных пятью алгоритмами, были использованы эти сообщества. Кроме того, были использованы графы ближайшего окружения 2000 случайно выбранных пользователей социальной сети ВКонтакте, полученные в [16].

Таблица 1

Характеристики наборов данных

Социальная сеть	Число графов	Среднее число			
		вершин $\bar{N}$	ребер $\bar{E}$	атрибутов $\bar{F}$	сообществ $\bar{C}$
Facebook	10	417	17 017	228	19.3
Twitter	973	138	2350	665	4,17
ВКонтакте	2000	164	1561	1665	n/a

В качестве меры качества предсказания была использована мера  $F_1$ :

$$F_1(p^*, p) = 2PR / (P + R) = 2|p^* \cap p| / (|p^*| + |p|),$$

где  $p^*$  — предсказанный профиль;  $p$  — истинный профиль;  $P = |p^* \cap p| / |p^*|$  — точность;  $R = |p^* \cap p| / |p|$  — полнота. Для всех графов были получены множе-

ства атрибутов, представляющих собой предполагаемый профиль центральной вершины. Для каждого из графов выборок и метода предсказания был выбран порог с наибольшим значением меры  $F_1$ .

**Результаты** предсказания профилей с использованием всех атрибутов представлены в табл. 2. Значения меры  $F_1$  оказались очень низкими для всех методов. Стоит отметить, что несмотря на это предложенный алгоритм все равно дает приемлемые результаты для выборки Facebook.

Таблица 2

**Средние значения меры  $F_1$  предсказания профилей с использованием всех атрибутов**

Метод предсказания с использованием всех атрибутов	Социальная сеть		
	Facebook	Twitter	ВКонтакте
Простое большинство	0,46	0,16	0,06
Эталонные сообщества	0,52	0,19	Нет данных
Infomap	0,49	0,18	0,17
Modularity Max.	0,53	0,19	0,15
AGM-fit	0,52	0,19	0,19
BigCLAM	0,57	0,23	0,26
CESNA	0,54	0,22	0,22
CESNA (веса)	0,31	0,19	0,23
Предлагаемый метод	0,60	0,27	0,32

Множества атрибутов были отфильтрованы: атрибуты, которые невозможно предсказать, были отброшены. Для набора данных Facebook были оставлены такие атрибуты, как сведения об образовании пользователя (высшем и среднем) и его родном городе. Для выборки графов из сети Twitter были выбраны 100 самых популярных хэштегов во всей выборке, число которых было снижено до 86 после ручной проверки. Для набора данных из сети ВКонтакте были оставлены сведения о среднем и высшем образовании и месте работы. Результаты для выборок только с предсказываемыми атрибутами приведены в табл. 3.

Таблица 3

**Средние значения меры  $F_1$  предсказания профилей с использованием предсказываемых атрибутов**

Метод предсказания с использованием предсказываемых атрибутов	Социальная сеть		
	Facebook	Twitter	ВКонтакте
Простое большинство	0,77	0,64	0,42
Эталонные сообщества	0,76	0,76	Нет данных
Infomap	0,73	0,64	0,56
Modularity Max.	0,79	0,68	0,58
AGM-fit	0,78	0,67	0,62
BigCLAM	0,86	0,72	0,73
CESNA	0,84	0,72	0,68
CESNA (веса)	0,47	0,75	0,77
Предлагаемый метод	0,84	0,87	0,86

Для выборок, содержащих только предсказываемые атрибуты, значения меры  $F_1$  для всех подходов значительно выше. Предложенный алгоритм дает один из наилучших результатов для набора данных из сети Facebook и существенно превосходит другие подходы на выборках из сетей Twitter и ВКонтакте. При этом полученные значения меры  $F_1$  довольно близки к единице.

Если сфокусироваться на отдельных атрибутах (табл. 4), то предложенный алгоритм показывает результаты, очень близкие к единице. Таким образом, данный алгоритм может быть использован для предсказания профилей с очень высокой точностью.

Таблица 4

**Результаты предсказания некоторых атрибутов профилей с использованием только предсказываемых атрибутов**

Социальная сеть	Атрибут	Среднее значение		
		меры $F_1$	точности	полноты
Facebook	Образовательное учреждение	0,902	0,967	0,875
	Родной город	1,000	1,000	1,000
ВКонтакте	Средняя школа	0,919	0,910	0,973
	Факультет (кафедра)	0,994	0,995	0,996
	Место работы	0,994	0,997	0,994

**Заключение.** В настоящей работе был предложен метод предсказания атрибутов профиля пользователя социальной сети. Разработанный метод использует как информацию о связях пользователя, так и сведения о его атрибутах. Он превосходит другие методы по мере  $F_1$  на выборках из трех социальных сетей и позволяет предсказывать некоторые атрибуты профиля с точностью, близкой к единице.

Предложенный подход может быть использован для раскрытия личности анонимных пользователей онлайн-социальных сетей. Например, если пользователь имеет связи с другими пользователями, но при этом не указал о себе никаких данных, в том числе ни имени, ни фотографии, то он считает, что его личность не может быть раскрыта. Однако с использованием предложенного алгоритма некоторые его атрибуты могут быть предсказаны с высокой точностью по связям в социальной сети. Очевидно, что большинство людей имеют уникальную комбинацию этих атрибутов. Используя предсказанный профиль и вспомогательные источники данных такие, как списки учащихся в образовательном учреждении или телефонный справочник, аналитик может раскрыть личность пользователя с высокой точностью.

## ЛИТЕРАТУРА

1. Davis C.A.J., Pappa G.L., de Oliveira D.R.R., de Zima A.F. Inferring the location of twitter messages based on user relationships // T. GIS. 2011. Vol. 15. No. 6. P. 735–751. DOI: 10.1111/j.1467-9671.2011.01297.x  
URL: <http://onlinelibrary.wiley.com/doi/10.1111/j.1467-9671.2011.01297.x/abstract>

2. *Li R., Wang C., Chang K. C.-C.* User profiling in an ego network: co-profiling attributes and relationships // Proc. 23rd Int. Conf. on World Wide Web. WWW '14. New York: ACM, 2014. P. 819–830. URL: <http://wwwconference.org/proceedings/www2014/proceedings/p819.pdf>
3. *Dong Y., Yang Y., Tang J., Yang Y., Chawla N.V.* Inferring user demographics and social strategies in mobile social networks // Proc. 20th ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining. KDD '14. New York: ACM, 2014. P. 15–24.
4. *Анализ социальных сетей: методы и приложения / А. Коршунов, И. Белобородов, Н. Бузун, В. Аванесов и др.* // Труды Института системного программирования РАН. 2014. Т. 26. № 1. С. 439–456.  
URL: <http://cyberleninka.ru/article/n/analiz-sotsialnyh-setey-metody-i-prilozheniya>
5. *Mislove A., Viswanath B., Gummadi K.P., Druschel P.* You are who you know: inferring user profiles in online social networks // Proc. 3d ACM Int. Conf. on Web Search and Data Mining. WSDM '10. New York: ACM. 2010. P. 251–260.
6. *Chaabane A., Acs G., Kaafar M.* You are what you like! Information leakage through users' interests // Proc. Annual Network and Distributed System Security Symposium. 2012.
7. *Kosinski M., Stillwell D., Graepel T.* Private traits and attributes are predictable from digital records of human behavior // Proc. of the National Academy of Sciences. 2013. Vol. 110. No. 15. P. 5802–5805. DOI: 10.1073/pnas.1218772110 URL: <http://www.pnas.org/content/110/15/5802.full>
8. *Dougnon R.Y., Fournier-Viger P., Nkambou R.* Advances in artificial intelligence // Proc. 28th Canadian Conf. on Artificial Intelligence. Canada: Springer International Publishing, 2015. P. 84–99.
9. *Yang J., Leskovec J.* Community-affiliation graph model for overlapping network community detection // 12th IEEE Int. Conf. on Data Mining, ICDM 2012. 2012. P. 1170–1175. DOI: 10.1109/ICDM.2012.139 URL: <http://ieeexplore.ieee.org/document/6413734>
10. *Чесноков В.О.* Выделение пересекающихся сообществ в социальных графах по мажоритарному признаку соседей // ЛОМОНОСОВ — 2016. XXIII Международная научная конференция студентов, аспирантов и молодых ученых. 2016. М.: МАКС Пресс, 2016. С. 49–51.
11. *Clauset A., Newman M.E.J. Moore C.* Finding community structure in very large networks // Phys. Rev. E. 2004. Vol. 70. No. 6. P. 1–6. DOI: 10.1103/PhysRevE.70.066111  
URL: <http://journals.aps.org/pre/abstract/10.1103/PhysRevE.70.066111>
12. *Rosvall M., Bergstrom C.T.* Maps of random walks on complex networks reveal community structure // Proc. of the National Academy of Sciences. 2008. Vol. 105. No. 4. P. 1118–1123. DOI: 10.1073/pnas.0706851105 URL: <http://www.pnas.org/content/105/4/1118.full>
13. *Yang J., Leskovec J.* Overlapping community detection at scale: a nonnegative matrix factorization approach // Proc. of the 6th ACM Int. Conf. on Web Search and Data Mining. WSDM '13. New York, 2013. P. 587–596. DOI: 10.1145/2433396.2433471  
URL: <http://dl.acm.org/citation.cfm?doid=2433396.2433471>
14. *Yang J., McAuley J.J., Leskovec J.* Community detection in networks with node attributes // 2013 IEEE 13th Int. Conf. on Data Mining. 2013. P. 1151–1156.
15. *Leskovec J., Krevl A.* SNAP datasets: stanford large network dataset collection // Stanford Network Analysis Project: веб-сайт. URL: <https://snap.stanford.edu/data> (дата обращения: 12.01.2017).
16. *Чесноков В.О., Ключарёв П.Г.* Выделение сообществ в социальных графах по множеству признаков с частичной информацией // Наука и образование: научное издание МГТУ им. Н.Э. Баумана. Электрон. журн. 2015. № 9. С. 188–199. DOI: 10.7463/0915.0811704

**Чесноков Владислав Олегович** — аспирант, ассистент кафедры «Информационная безопасность» МГТУ им. Н.Э. Баумана (Российская Федерация, 105005, Москва, 2-я Бауманская ул., д. 5).

**Просьба ссылаться на эту статью следующим образом:**

Чесноков В.О. Предсказание атрибутов профиля пользователя социальной сети путем анализа сообществ графа его ближайшего окружения // Вестник МГТУ им. Н.Э. Баумана. Сер. Приборостроение. 2017. № 2. С. 66–76. DOI: 10.18698/0236-3933-2017-2-66-76

**PREDICTING ATTRIBUTES OF USER PROFILE IN SOCIAL NETWORKS BY ANALYZING COMMUNITIES OF THEIR EGO-NETWORK**

V.O. Chesnokov

v.o.chesnokov@yandex.ru

**Bauman Moscow State Technical University, Moscow, Russian Federation**

**Abstract**

In online social networks, a user is allowed to specify a lot of personal information — attributes. Some users provide only a part of whole information, or do not provide any information about themselves at all. Due to that, inferred hidden attributes are one of the fundamental problems of social analysis. The study proposes a new approach to user's hidden or unspecified attributes prediction. The method is based on analysis of the user's ego-network structure and attributes of its social graph vertices. The developed method was compared with other methods according to three datasets of users' ego-networks from Facebook, Twitter and VKontakte social networks. It showed high values of F-measure, precision and completeness for predicting the chosen attributes of the user profile such as hometown or school. Using this method with additional data sources an analyst with high precision can reveal the identity of an anonymous social network user by their relations with other users

**Keywords**

*Social networks, social graph, community detection, profile prediction*

**REFERENCES**

- [1] Davis C.A.J., Pappa G.L., de Oliveira D.R.R., Arcanjo F. de Zima. Inferring the location of twitter messages based on user relationships. *T. GIS.*, 2011, vol. 15, no. 6, pp. 735–751. DOI: 10.1111/j.1467-9671.2011.01297.x  
Available at: <http://onlinelibrary.wiley.com/doi/10.1111/j.1467-9671.2011.01297.x/abstract>
- [2] Li R., Wang C., Chang K. C.-C. User profiling in an ego network: co-profiling attributes and relationships. *Proc. 23rd Int. Conf. on World Wide Web. WWW '14.* New York, ACM, 2014, pp. 819–830.  
Available at: <http://www.conference.org/proceedings/www2014/proceedings/p819.pdf>
- [3] Dong Y., Yang Y., Tang J., Yang Y., Chawla N.V. Inferring user demographics and social strategies in mobile social networks. *Proc. 20th ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining. KDD '14.* New York, ACM, 2014, pp. 15–24.

[4] Korshunov A., Beloborodov I., Buzun N., Avanesov V., et al. Social network analysis: methods and applications. *Trudy Instituta sistemnogo programmirovaniya RAN* [Proceedings of ISP RAS], 2014, vol. 26, no. 1, pp. 439–456 (in Russ.).

Available at: <http://cyberleninka.ru/article/n/analiz-sotsialnyh-setey-metody-i-prilozheniya>

[5] Mislove A., Viswanath B., Gummadi K.P., Druschel P. You are who you know: inferring user profiles in online social networks. *Proc. 3d ACM Int. Conf. on Web Search and Data Mining. WSDM '10*. New York, ACM, 2010, pp. 251–260.

[6] Chaabane A., Acs G., Kaafar M. You are what you like! Information leakage through users' interests. *Proc. Annual Network and Distributed System Security Symposium*, 2012.

[7] Kosinski M., Stillwell D., Graepel T. Private traits and attributes are predictable from digital records of human behavior. *Proc. of the National Academy of Sciences*, 2013, vol. 110, no. 15, pp. 5802–5805. DOI: 10.1073/pnas.1218772110

Available at: <http://www.pnas.org/content/110/15/5802.full>

[8] Dougnon R.Y., Fournier-Viger P., Nkambou R. Advances in artificial intelligence. *Proc. 28th Canadian Conf. on Artificial Intelligence*. Canada, Springer International Publishing, 2015, pp. 84–99.

[9] Yang J., Leskovec J. Community-affiliation graph model for overlapping network community detection. *12th IEEE Int. Conf. on Data Mining, ICDM 2012*. 2012, pp. 1170–1175.

DOI: 10.1109/ICDM.2012.139 Available at: <http://ieeexplore.ieee.org/document/6413734>

[10] Chesnokov V.O. Intersecting communities allocation in social graphs based on majoritarian neighbours characteristic. *LOMONOSOV–2016. XXIII Mezhd. nauch. konf. studentov, aspirantov i molodykh uchenykh* [LOMONOSOV–2016. XXIII Int. conf. of students, postgraduates and young scientists]. Moscow, MAKS Press Publ., 2016, pp. 49–51.

[11] Clauset A., Newman M.E.J., Moore C. Finding community structure in very large networks. *Phys. Rev. E*, 2004, vol. 70, no. 6, pp. 1–6. DOI: 10.1103/PhysRevE.70.066111

Available at: <http://journals.aps.org/pre/abstract/10.1103/PhysRevE.70.066111>

[12] Rosvall M., Bergstrom C.T. Maps of random walks on complex networks reveal community structure. *Proc. of the National Academy of Sciences*, 2008, vol. 105, no. 4, pp. 1118–1123.

DOI: 10.1073/pnas.0706851105 Available at: <http://www.pnas.org/content/105/4/1118.full>

[13] Yang J., Leskovec J. Overlapping community detection at scale: a nonnegative matrix factorization approach. *Proc. of the 6th ACM Int. Conf. on Web Search and Data Mining. WSDM '13*. New York, 2013, pp. 587–596. DOI: 10.1145/2433396.2433471

Available at: <http://dl.acm.org/citation.cfm?doid=2433396.2433471>

[14] Yang J., McAuley J.J., Leskovec J. Community detection in networks with node attributes. *2013 IEEE 13th Int. Conf. on Data Mining*. 2013, pp. 1151–1156.

[15] Leskovec J., Krevl A. SNAP datasets: stanford large network dataset collection. Stanford Network Analysis Project: website. Available at: <https://snap.stanford.edu/data> (accessed 12.01.2017).

[16] Chesnokov V.O., Klyucharev P.G. Social graph community differentiated by node features with partly missing information. *Nauka i obrazovanie: nauchnoe izdanie MGTU im. N.E. Baumana* [Science and Education: Scientific Publication of BMSTU], 2015, no. 9, pp. 188–199 (in Russ.).

DOI: 10.7463/0915.0811704

**Chesnokov V.O.** — assistant, post-graduate student of Information Security Department, Bauman Moscow State Technical University (2-ya Baumanskaya ul. 5, Moscow, 105005 Russian Federation).

**Please cite this article in English as:**

Chesnokov V.O. Predicting Attributes of User Profile in Social Networks by Analyzing Communities of their Ego-Network. *Vestn. Mosk. Gos. Tekh. Univ. im. N.E. Baumana, Priborostr.* [Herald of the Bauman Moscow State Tech. Univ., Instrum. Eng.], 2017, no. 2, pp. 66–76. DOI: 10.18698/0236-3933-2017-2-66-76



В Издательстве МГТУ им. Н.Э. Баумана  
вышло в свет учебное пособие автора  
**Л.Н. Лысенко**

**«Наведение баллистических ракет»**

Изложены научные и методологические основы наведения баллистических ракет. Рассмотрены вопросы программирования движения (задачи наведения) и информационно-навигационного обеспечения управления (задачи навигации), а также проблемы определения точности стрельбы (задачи оценки точности возмущенного движения). Показаны направления решений соответствующих задач при создании ракетных комплексов тактического, оперативно-тактического и стратегического назначения, возможные пути совершенствования баллистико-навигационного обеспечения полета ракет указанных классов.

**По вопросам приобретения обращайтесь:**

105005, Москва, 2-я Бауманская ул., д. 5, стр. 1  
+7 (499) 263-60-45  
press@bmstu.ru  
www.baumanpress.ru