



Software Engineering Conference Russia

14-15 ноября, 2019. Санкт-Петербург

Способы автоматического поиска потенциальных экспертов по социальным сетям

Евгений Липкин

Университет ИТМО

Социальная сеть — это интернет-площадка, которая позволяет зарегистрированным на нем пользователям

- размещать информацию
- устанавливать социальные связи
- коммуницировать между собой

Контент на этой площадке создается непосредственно самими пользователями

Социальные сети

1. Глобальные

ВКонтакте, Одноклассники, Facebook, Instagram

2. Специализированные

Dribbble, Github, ResearchGate, Unsplash

Термин **эксперт** используется для обозначения человека с высокой степенью квалификации или знания определенной предметной области

Вопросы



1. Можно ли использовать записи пользователей из **глобальной** социальной сети как документы, которые указывают на **потенциальный уровень экспертности** пользователя?
2. Можно ли использовать **глобальную** социальную сеть для **поиска** потенциальных экспертов?
3. Можно ли использовать **глобальную** социальную сеть для **оценивания** потенциальных экспертов?



Обзор методов обнаружения экспертов

Стратегии Valog

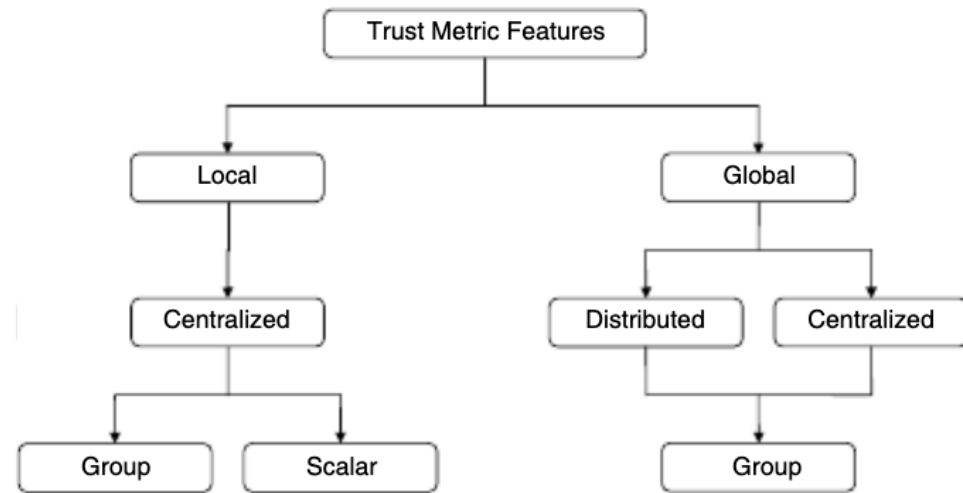
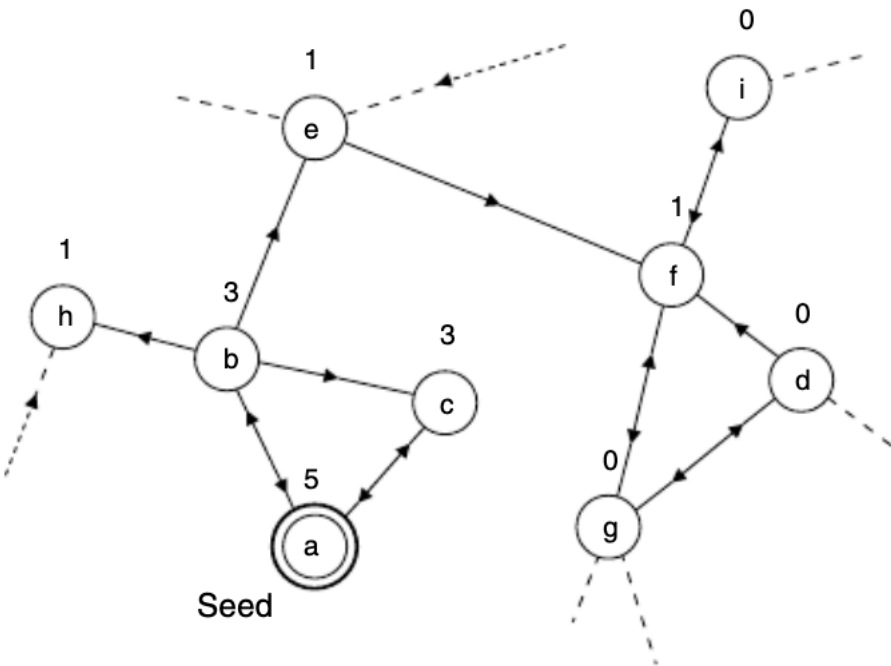
1. профиле-ориентированный подход

$$P(Q|x_i) = \prod_{t_i \in Q} P(t_i|x_i) = \prod_{t_i \in Q} \left[\sum_{d_j \in D} P(t_i|d_j) P(d_j|x_i) \right]$$

2. документо-ориентированный подход

$$P(Q|x_i) = \sum_{d_j \in D} P(Q|d_j) P(d_j|x_i) = \sum_{d_j \in D} \left[\prod_{t_i \in Q} P(t_i|d_j) \right] P(d_j|x_i)$$

Модель Ziegler-Lausen

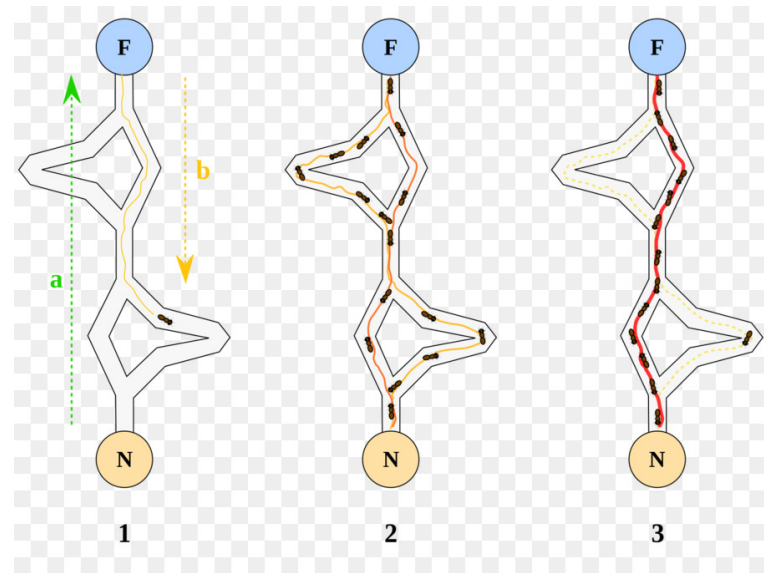


Ziegler C.N., Lausen G. Propagation models for trust and distrust in social networks // Information System Frontiers 7 (4-5), 2005. pp. 337–358.

Муравьиный алгоритм

(ant colony optimization)

$$P_i = \frac{l_i^q f_i^p}{\sum_{k=0}^N l_k^q f_k^p}$$



где P_i – вероятность перехода по пути i ;

l_i – величина, обратная весу (длине) i -го перехода;

f_i – количество феромона на i -ом переходе;

q – величина, определяющая «жадность» алгоритма;

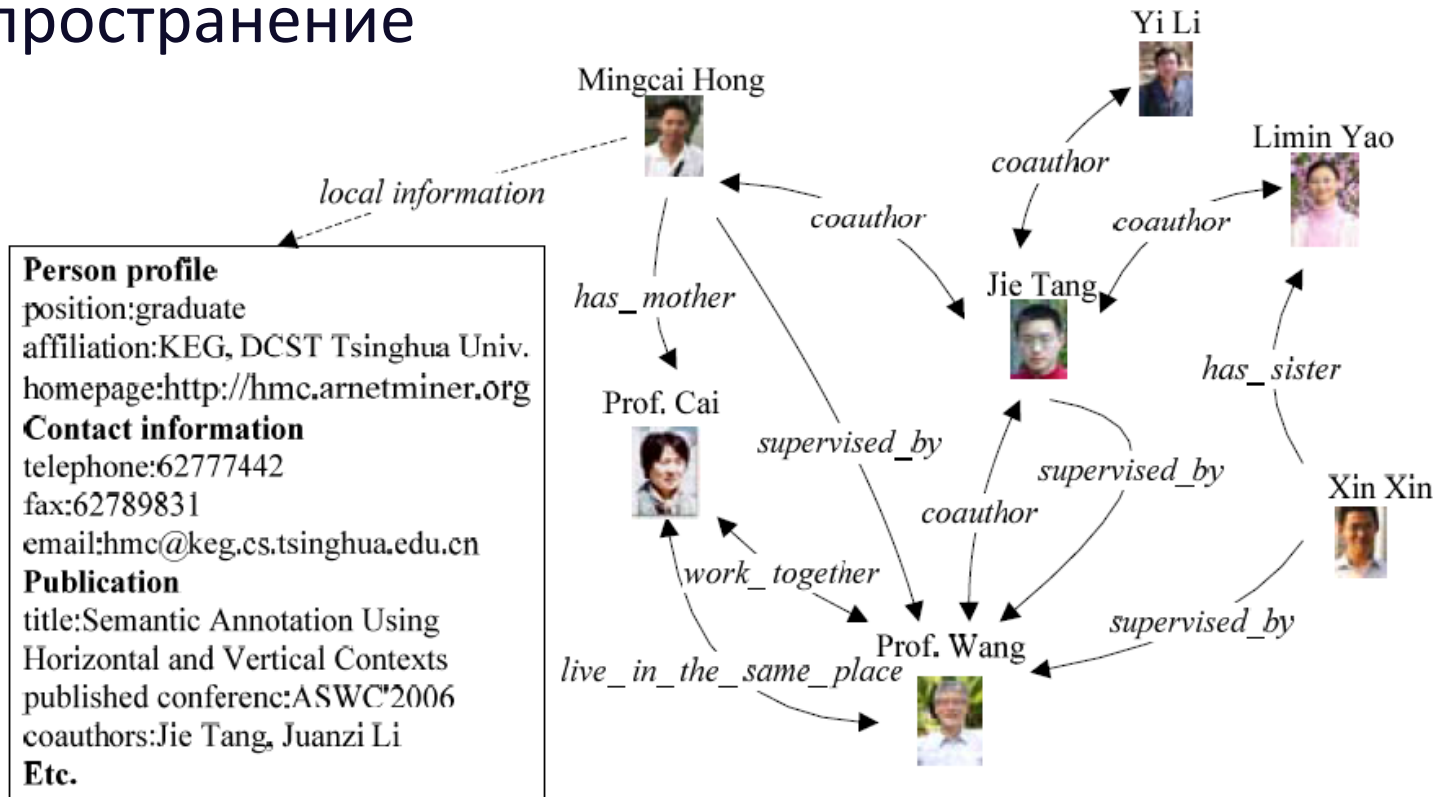
p – величина, определяющая «стадность» алгоритма;

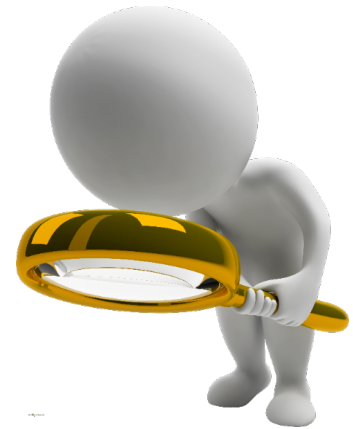
$q + p = 1$.

Подход Zhang, Tang, Li

1. Инициализация
2. Распространение

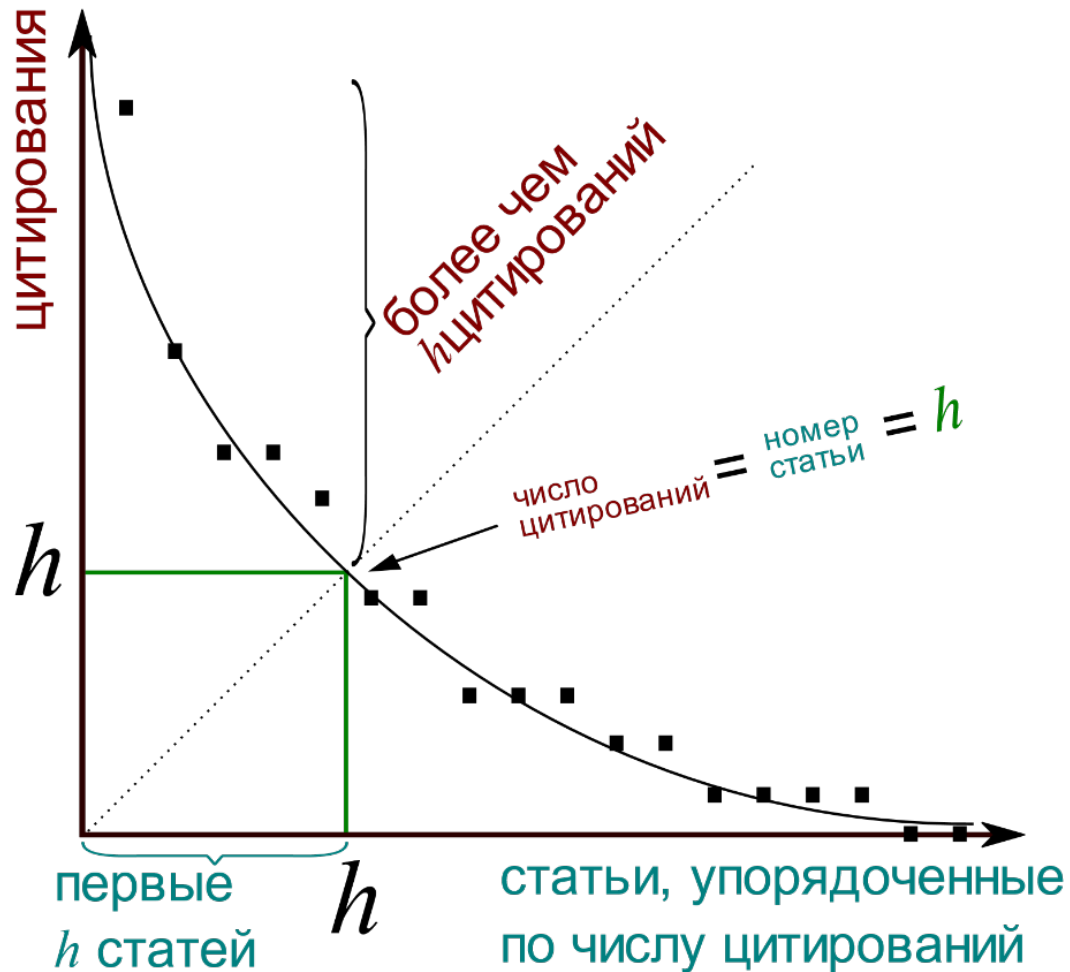
$$s(v_i)^{n+1} = s(v_i)^n + \sum_{v_j \in U} \sum_{e \in R_{ji}} w((v_j, v_i), e) s(v_j)^n$$





Обзор методов оценивания экспертов

Индекс Хирша

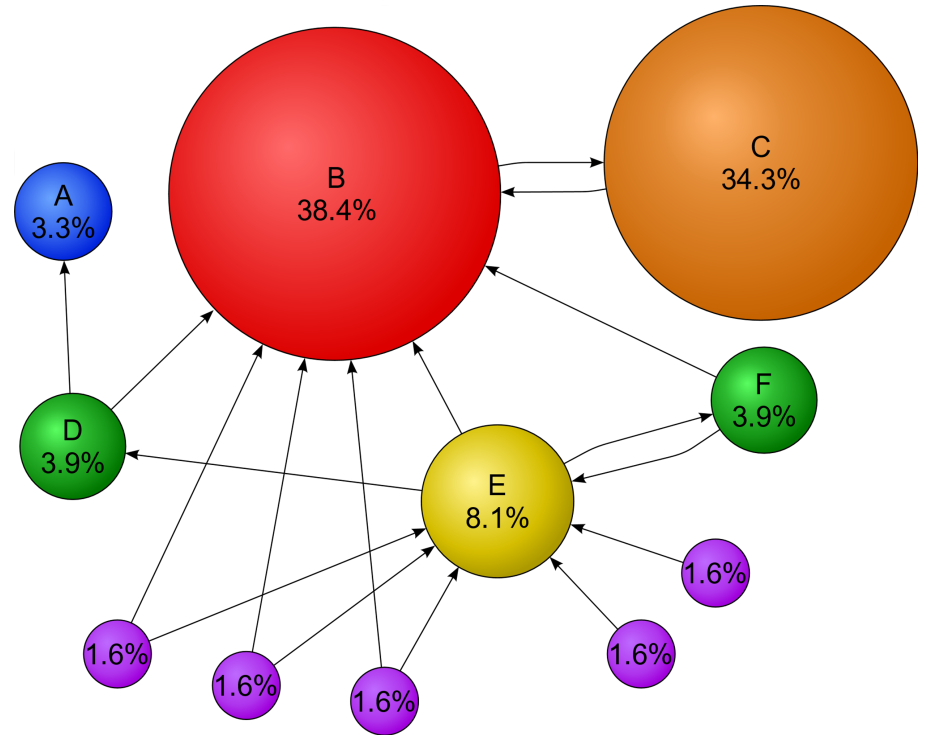


Hirsch J.E. An index to quantify an individual's scientific research output // Proceedings of the National Academy of Sciences of the United States of America, 2005. V. 19253, pp. 16569-16572.

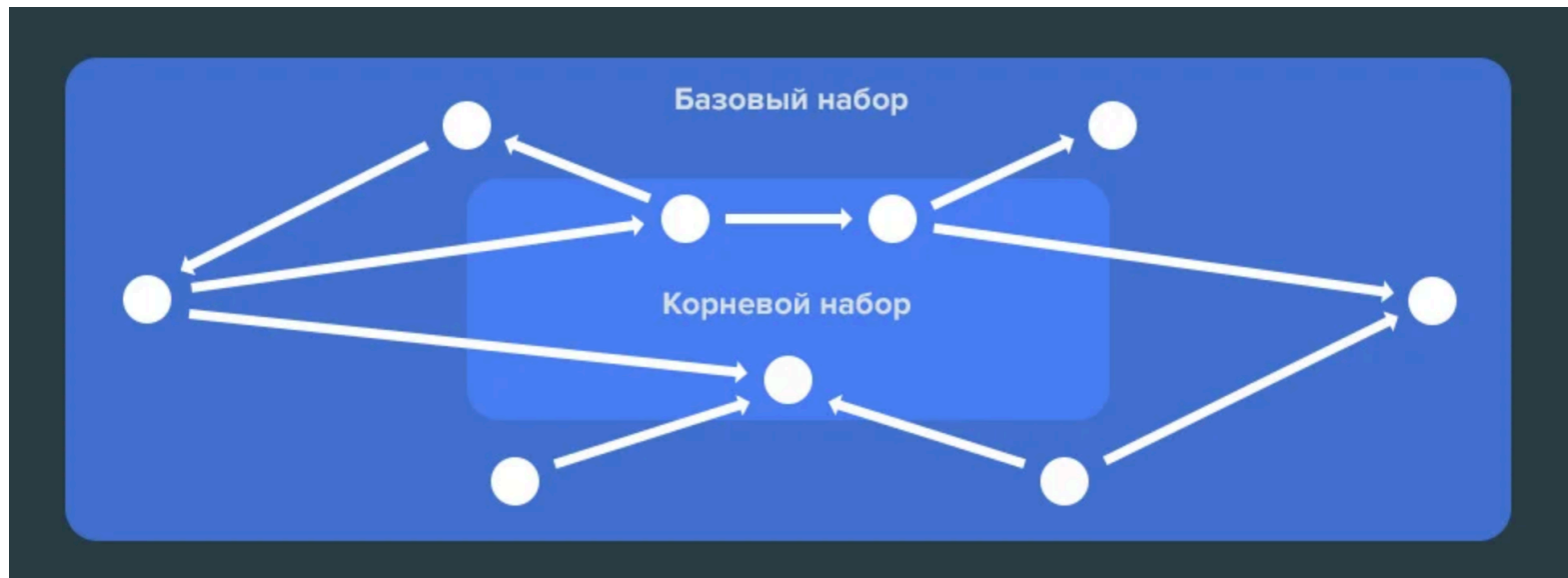
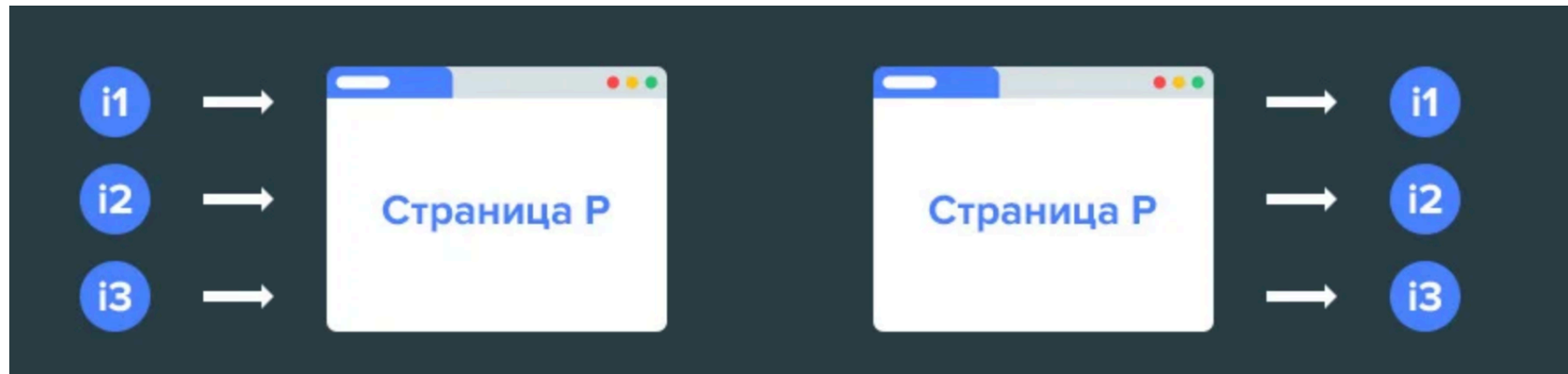
Алгоритм PageRank

$$I(P_i) = \sum_{P_j \in B_i} \frac{I(P_j)}{L_j}$$

$$H_{ij} = \begin{cases} \frac{1}{L_j}, & \text{if } P_j \in B_i \\ 0, & \text{otherwise} \end{cases}$$



Алгоритм HITS



Kleinberg, J. M. (1999). Authoritative sources in a hyperlinked environment. *Journal of the ACM (JACM)*, 46(5), 604-632.



Обзор эксперимента

Описание эксперимента

Социальная сеть: vk.com

Данные: лайки, шэры, посты, репосты участников пабликов

Методы обнаружения: стратегии Balog, подход Zhang, Tang, Li

Методы оценивания: h-index, PageRank

Предметные области: Technical and Computing, Business and Industry

Вопросы



1. Можно ли использовать записи пользователей из **глобальной** социальной сети как документы, которые указывают на **потенциальный уровень экспертности** пользователя?
2. Можно ли использовать **глобальную** социальную сеть для **поиска** потенциальных экспертов?
3. Можно ли использовать **глобальную** социальную сеть для **оценивания** потенциальных экспертов?

Результаты эксперимента

	h-index				PageRank			
	Likes		Shares		Likes		Shares	
	Init	Prop	Init	Prop	Init	Prop	Init	Prop
FIT. Education	114	67	37	20	594	494	56	51
EF. Education	106	189	18	47	663	1367	38	108
FIT. T&C	73	40	17	13	314	225	30	27
EF. T&C	47	94	10	25	195	499	15	89
FIT. B&I	73	51	12	15	318	350	18	39
EF. B&I	90	148	17	37	496	945	34	52



ИТОГИ

Ответы на вопросы



Можно ли использовать записи пользователей из **глобальной** социальной сети как документы, которые указывают на **потенциальный уровень экспертности** пользователя?

Можно ли использовать **глобальную** социальную сеть для **поиска** потенциальных экспертов?

Можно ли использовать **глобальную** социальную сеть для **оценивания** потенциальных экспертов?

Ответы на вопросы



Можно ли использовать записи пользователей из **глобальной** социальной сети как документы, которые указывают на **потенциальный уровень экспертности** пользователя?

Да

Можно ли использовать **глобальную** социальную сеть для **поиска** потенциальных экспертов?

Да

Можно ли использовать **глобальную** социальную сеть для **оценивания** потенциальных экспертов?

Непонятно

Результаты

Исследование показывает, что исследователей могут интересовать не только **лайки** и **посты**, которые изучает большинство людей, но также и **шэры** и **репосты**

Исследователи также могут быть заинтересованы в использовании **алгоритмов ранжирования веб-страниц** для определения веса агентов в социальном графе



Контакты

1. Евгений Липкин
2. Университет ИТМО
3. E-mail: elipkin@itmo.ru

