

РАСПРОСТРАНЕНИЕ НОВОСТЕЙ В СОЦИАЛЬНОЙ СЕТИ

Вихрева Мария Викторовна

Студент

Факультет ВМК МГУ имени М. В. Ломоносова, Москва, Россия

E-mail: mary.vikhreva@gmail.com

Социальная сеть часто представляется в виде динамического графа, в котором вершины — пользователи сети, а рёбра — социальные связи (например, отношения дружбы или подписка).

Определим в данном графе для момента времени t :

- $I(t)$ — множество пользователей сети, запостивших новость хоть раз за t моментов времени (*активные пользователи*);
- $S(t)$ — множество пользователей сети, не постивших новость ни в один из t моментов времени (*неактивные пользователи*);
- $S(t) \cup I(t)$ — множество человек в сети (фиксировано).

Считается, что человек не может перейти из множества уже запостивших новость в множество не запостивших, и выполнено условие $I(t) \subseteq I(t+1)$ для любого t . Также предполагается, что ребра графа статичны и не изменяются со временем.

Тогда исследование распространения новости в социальной сети — это исследование динамики изменения подмножества $I(t)$ вершин графа.

А именно, ставится задача предсказать пользователей, активных к моментам времени $t_0 + 1, t_0 + 2, \dots, t_0 + n$, на основе истории постов за t_0 предшествующих моментов времени. Рассматривается как задача предсказания количества $|I(t+i+1) \setminus I(t+i)|$ пользователей, запостивших новость в момент времени t , так и задача определения в социальном графе вершин из этого множества $I(t+i+1) \setminus I(t+i)$.

Эксперименты проведены для истории постов с тегом «Higgs Boson» в сети Twitter. Считалось, что новость содержится в посте пользователя социальной сети Twitter, если пост содержит тег этой новости.

В данной работе предложен метод решения задачи в рамках математических методов обучения по прецедентам. В качестве альтернативного подхода рассматривается SI-модель распространения эпидемий [1].

Показано, что для датасета Twitter предложенный метод показывает большую точность по сравнению с некоторым классом SI-методов. А также выявлены основные факторы, влияющие на скорость распространения новости «Higgs Boson».

Литература

1. Anderson R. M. and May R. M. Infectious Diseases in Humans // Oxford University Press, Oxford, 1992
2. Montenari A. and Saberi A. The spread of innovations in social networks // Proceedings of the National Academy of Sciences, California, USA, 2010, vol. 107 no. 47.