Моделирование распространения эпидемии на графах

Бишук Антон Юрьевич

Московский физико-технический институт
Факультет управления и прикладной математики
Кафедра интеллектуальных систем

Научный руководитель к.ф.-м.н. Зухба А.В.

Москва
2021 г

Цели и задачи

Цель

- Оценить влияние ограничительных мер.
  1. Для монотонных – предложить оптимальную стратегию.
  2. Для немонотонных – предложить критерий положительного влияния.
- Сформулировать задачу в терминах теории информации, и представить алгоритмы поиска вершин, проверка которых наиболее сильно уменьшит энтропию системы.

Задача

- Поиска условий при которых локдаун может пагубно влиять на эпидемию;
- Поиска оптимального тестирования и вакцинации с точки зрения уменьшения числа заболевших;
- Поиска людей, информация о которых, наибольшим образом уменьшит энтропию системы.

Граф контактов

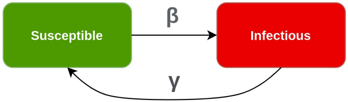
Граф контактов:

- Вершина графа – человек, а ребро – контакт между людьми.
- Предполагается, что каждая вершина графа может находиться в конечном количестве состояний.
- Состояние системы меняется во времени в соответствии с определенными правилами.

Графовые определения:

- Будем называть размером окрестности степени k вершины v – число вершин, до которых можно добраться следуя по ребрам исходного графа за k переходов.
- Число заболевших в графе на k-ом шаге можно оценить сверху как размер окрестности степени k.

SIS



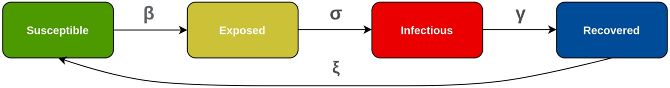
$$\begin{cases} \frac{dS}{dt} = \frac{-\beta SI}{N} + \gamma I \\ \frac{dI}{dt} = \frac{\beta SI}{N} - \gamma I \end{cases}$$

Где $S + I = N$ – вся популяция

β – вероятность заразиться

γ – вероятность выздоровить

SEIRS



$$\begin{cases} \frac{dS}{dt} = -\frac{\beta IS}{N} + \xi R, \\ \frac{dE}{dt} = \frac{\beta IS}{N} - \sigma E, \\ \frac{dI}{dt} = \sigma E - \gamma I, \\ \frac{dR}{dt} = \gamma I - \xi R. \end{cases}$$

Где $S + E + I + R = N$ – вся популяция

β – вероятность заразиться
σ – вероятность перейти из состояния зараженного в состояние больного
γ – вероятность выздоровить и приобрести иммунитет
ξ – вероятность потерять иммунитет

Описание системы

- Назовем вектор x – вектором рисков, представляющий из себя вектор длины 2^n, где каждая координата – кодировка заболевших и здоровых вершин.
- Матрицу A, в строках которой находятся состояния системы на текущей итерации, в столбцах состояния, в которые возможно перейти, в ячейках матрицы – вероятность такого перехода будем называть матрицей распространения эпидемии
- A^k - представление эпидемии на k-ом шаге. $x^T A^k$ - вектор рисков на k-ой итерации.

Теорема об «эквивалентности» вершин

Theorem (Bishuk)

Пусть дан граф $G(V, E)$, пусть в нем $n > 0$ вершин заражено, тогда число заболевших на следующей итерации монотонно от величины окрестности для всех больных вершин.

Sum.

$$\mathbb{E}(I_1) = \sum_{i=1}^{k}(i-1) \cdot \sum_{j=1}^{C_{k+\mathbf{1}}^{i}} p_{i,j}$$

Здесь $p_{i,j}$ – вероятности перехода в матрице A в группе i (группа формируется по количеству заболевших) по номеру ненулевого элемента j.
k – степень вершины i.
$\mathbb{E}(I_1)$ – математическое ожидание числа больных через одну итерацию.

Противоэпидемиологические меры для графовых структур

Для борьбы с распространением эпидемии используется ряд противоэпидемиологических мер:

- Вакцинация (Смена статуса вершины $S \to R$)
- Изоляция больных (Удаление вершины и инцидентных с ней рёбер)
- Введение lockdown (Смена графа контактов на граф, состоящий из небольших клик)

И если первые два способа сдерживания эпидемии в худшем случае могут не помочь, то в случае локдауна, ограничение может отрицательно повлиять на сдерживание эпидемии.

Монотонные эпидемиологические меры 1

Lemma

Тестирование и изоляция больных не ухудшает эпидемиологическую обстановку.

Proof:

- *Пусть была протестирована вершина и она оказалась здоровой, то это значит, что вершину изолировать не нужно, поскольку она не сможет никого заразить, но может заразиться, что увеличит число заболевших. Но это бы произошло с той же вероятность, если бы мы не проверяли данную вершину.*
- *Пусть мы проверили вершину (пусть v_i) и она оказалась больной. Если эта вершина не будет изолирована, то верхней оценкой прироста числа заболевших от этой вершины будет* $\sum\limits_{v_j \in V_i} \beta$*. А если эта вершина будет изолирована, то верхней оценкой будет* 0*.*

Монотонные эпидемиологические меры 2

Lemma

Вакцинирование не ухудшает эпидемиологическую обстановку

Proof:

- *Пусть мы выбрали вершину и провакцинировали её. Тогда если она в будущем будет изолированной, то вакцинация никак не повлияет на прирост числа зараженных.*
- *Пусть вершина не изолированная и имеет больные вершины, тогда вакцинация гарантированно не увеличит число заболевших на следующих итерациях по сравнению с тем, если бы мы не вакцинировали*

Критерий «плохого» локдауна

Lemma (Bishuk)

Шанс заразиться вся вершин v_i растет при локдауне тогда и только тогда, когда

$$\frac{I_k(V_i^H)B_i^H}{|V_i^H|} \geq \frac{I_k(V_i^W)B_i^W}{|V_i^W|}$$

Theorem (Bishuk)

Локдаун ухудшает эпидемиологическую обстановку (с точки зрения числа больных) тогда и только тогда, когда:

$$\mathbb{E}(\tilde{I}_{k+1} - I_{k+1}) = \sum_{i:v_i \in I_k} (\tilde{\beta}_i^k - \beta_i^k) =$$

$$= \sum_{i=1}^{|V|} [v_i \in S_k] \frac{B_i^H |V_i^W| I_k(V_i^H) - B_i^W |V_i^H| I_k(V_i^W)}{|V_i^H| \cdot |V_i^W|} > 0$$

Тестирование и изоляция

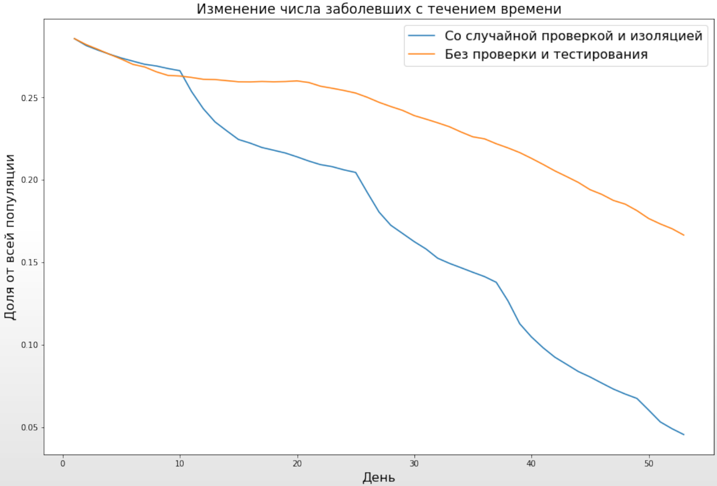


Рис.: Изменение числа заболевших с течением времени

Вакцинация

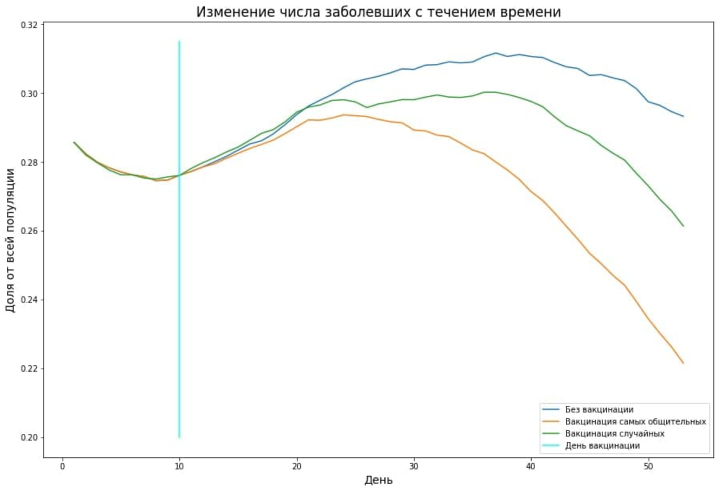


Рис.: Изменение числа заболевших с течением времени

Локдаун

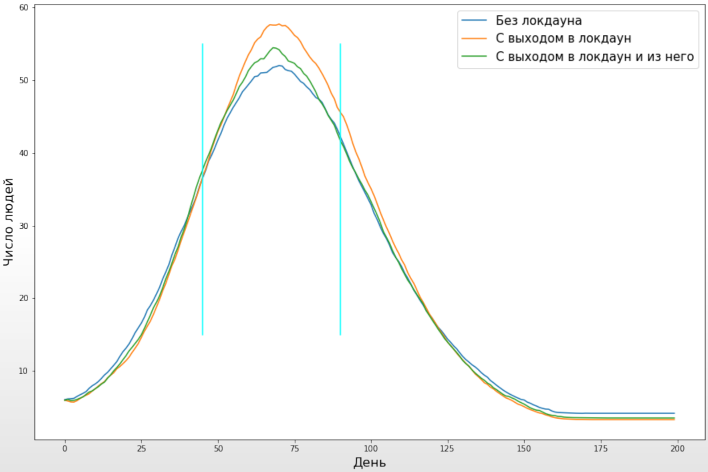


Рис.: Изменение числа заболевших с течением времени

Энтропия и пересчет вероятностей

Граф это система, в вершинах, которой стоят вероятности того, что вершина имеет метку 1 (вершина больна). Иными словами чем меньше вероятность находится в вершине, тем выше вероятность того, что в вершине находится метка 0.

Тогда под энтропией будем понимать следующую величину:

$$\mathbb{H}(\{x_i\}_{i=1}^n) = -\left(\sum_{i=1}^n p_i \log(p_i) + \sum_{i=1}^n (1-p_i)\log(1-p_i)\right)$$

Известно распределение вероятностей на прошлой итерации $\{p_i^{t-1}\}_{i=1}^n$. Пересчёт вероятностей на следующей итерации:

$$p_i^t = p_i^{t-1}(1-\gamma) + (1-p_i^{t-1})\beta e^{-c/\sum\limits_k p_{i_k}},$$

где p_{i_k} – вероятности соседей вершины v_i

Алгоритм аккумуляции информации

Максимум одного слагаемого энтропии достигается в точке $x_{opt} = \frac{1}{2}$.

Проверка человека, вероятность которого ближе всего к x_{opt} максимально уменьшит энтропию системы, получаемую от одного слагаемого. Тогда мы можем проверять людей так, чтобы вероятности их соседей наиболее вероятно приблизятся к $\frac{1}{2}$.

Пусть p_i^t – вероятность того, что в i-ой вершине находится метка 1 до того, как мы кто-то был протестирован.

$\tilde{p}_i^t$ – вероятность того, что в i-ой вершине находится метка 1 после того, как кто-то был протестирован.

Тогда необходимо найти такую вершину $x_{i'}$: $\forall i \hookrightarrow |p_i^t - \frac{1}{2}| > |\tilde{p}_{i'}^t - \frac{1}{2}|$ и для $\forall i \neq i' \hookrightarrow |\tilde{p}_i^t - \frac{1}{2}| > |\tilde{p}_{i'}^t - \frac{1}{2}|$

Алгоритм аккумуляции информации

Для простоты записи будем считать, что $c = 1$, тогда необходимо решить следующую оптимизационную задачу:

$$\left(-\frac{1}{2} + (1-\gamma)p_i^{t-1} + \beta(1-p_i^{t-1})(q_j^{t-1}\exp\left(\frac{-1}{\sum_k p_{i_k}^{t-1} + 1 - q_j^{t-1}}\right) + (1-q_j^{t-1})\exp\left(\frac{-1}{\sum_k p_{i_k}^{t-1} - q_j^{t-1}}\right))\right)^2 \to \min$$

Тогда сам алгоритм определения лучшей вершины для проверки:

1. Упорядочим $\{p_j^{t-1}\}_{j=1}^n$.
2. Для каждой вершины на итерации t получаем точку в которой достигается минимум $(\tilde{p}_i^t - \frac{1}{2})^2 \to \min$
3. Бинарным поиском ищем какая из вероятностей на прошлой итерации ближе всего к оптимуму, для неё считаем значение функции и запоминаем вершину, если функция уменьшилась
4. В конце алгоритма получим вершину, которую необходимо проверить, чтобы больше всего информации саккумулировать в вершине.

Алгоритм поиска вершины с большим вкладом в энтропию

Теперь будем рассматривать другой подход. Мы будем тестировать рассматриваемую вершину и смотреть как от этого поменяется энтропия всей системы за счет изменения энтропии соседних вершин.

- Пройдемся по всем рёбрам, записывая в вершины вероятности соседей.
- Пересчитываем вероятности для каждой вершины на основе вероятностей с прошлой итерации.
- На основе новых вероятностей считаем H_i – вклад i-ой вершины в энтропию системы.
- ройдемся по всем рёбрам и для каждой вершины считаем величину $H_i^t + \sum\limits_{i_k} H_{i_k}^t$.
- Теперь для каждой вершины считаем, что мы проверили её, тогда посчитаем матожидание вклада в энтропию системы окрестности проверенной вершины. Она равна $p_i \sum\limits_{i_k} \tilde{H}_{i_k} + (1 - p_i) \sum\limits_{i_k} \tilde{\tilde{H}}_{i_k}$

Алгоритм поиска вершины с большим вкладом в энтропию

- И тогда необходимо найти ту вершину, для которой величина

$$\left| H_i^t + \sum_k H_{i_k}^t - (p_i \sum_{i_k} \tilde{H}_{i_k} + (1 - p_i) \sum_k \tilde{\tilde{H}}_{i_k}) \right|$$

  будет максимальной, поскольку тогда мы найдем ту вершину, проверка, которой даст наибольшее уменьшение энтропии системы.

Данный алгоритм работает за $O(E + V)$

Алгоритм аккумуляции

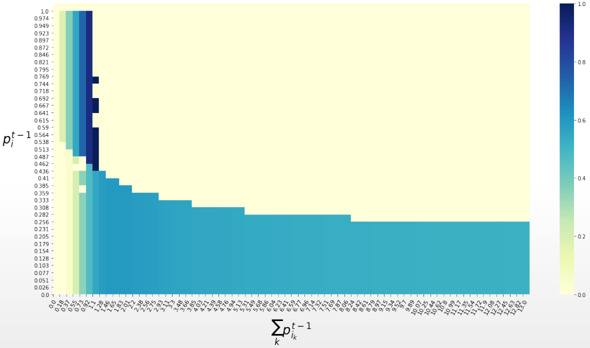


Рис.: Распределения оптимальных вероятностей для проверки от параметров вершины

Анализ ошибки

С точки зрения мат ожидания алгоритм оптимален, однако мы работаем с вероятностными распределениями, поэтому в качестве ошибки будет рассматриваться дисперсия.

Выводы и результаты

- Был изучен эффект отрицательного влияния локдауна. Была сформулирован и доказан критерий пагубного влияния локдауна.
- Была сформулирована и доказана теорема об эквивалентности людей для симуляции эпидемии.
- Сформулирована задача в терминах теории информации.
- Был предложен и реализован и экспериментально проверен алгоритм для аккумуляции неизвестной информации о системе в определенных вершинах.
- Был предложен алгоритм выбора вершины, проверка которой даёт наибольшее уменьшение энтропии системы.

Литература

1. *Moreno Y., Pastor-Satorras R., Vespignani A.* Epidemic outbreaks in complex heterogeneous networks //The European Physical Journal B-Condensed Matter and Complex Systems. – 2002. – Т. 26. – №. 4. – С. 521-529.
2. *Pastor-Satorras R.* et al. Epidemic processes in complex networks //Reviews of modern physics. – 2015. – Т. 87. – №. 3. – С. 925.