



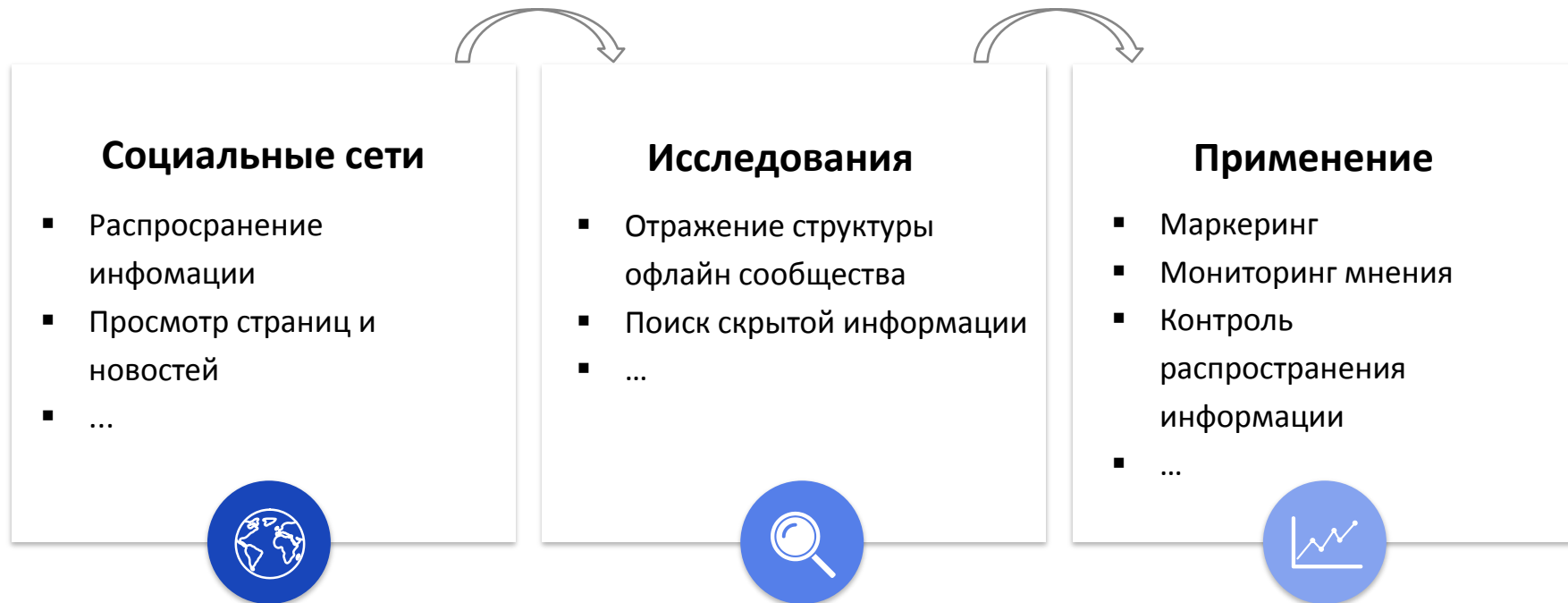
УНИВЕРСИТЕТ ИТМО

# Параллельное моделирование распространения информации внутри сообществ в социальных сетях

С. Кесарев, О. Северюхина, К. Боченина,  
oseveriukhina@gmail.com

Суперкомпьютерные дни в России  
Москва, 24-25 сентября 2018 г.

# Зачем необходимо моделирование сетей



## Цель

Создание эффективной модели распространения информации в комплексных сетях

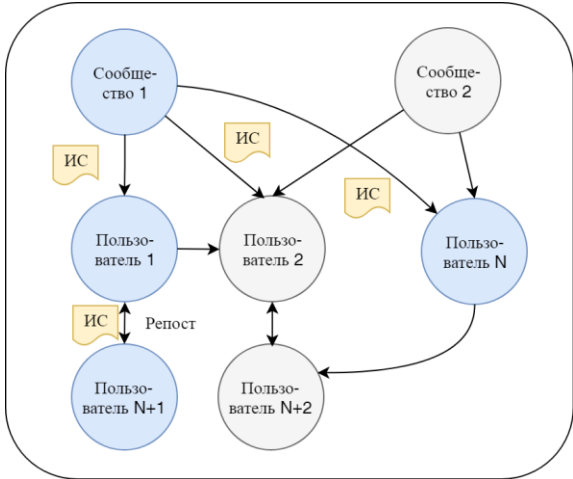
## Требования

- Мультиагентная модель
- Комплексная сеть
- Параллельная обработка
- Использование реальных данных



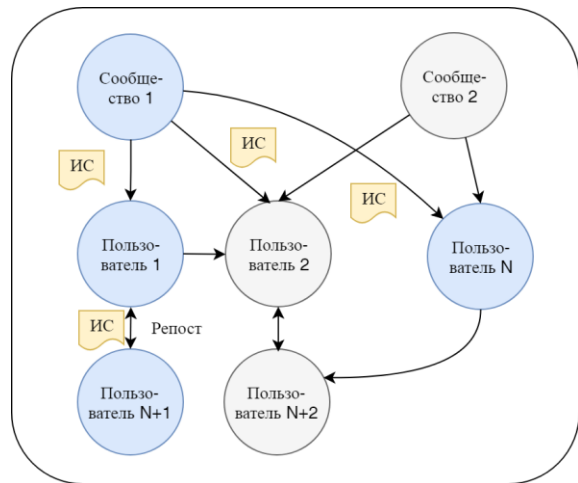
# Описание модели

## Сущности



# Описание модели

## Сущности



## Внутренние подмодели

### Модель генерации сообщений

Появление новых сообщений



### Модель активности

Состояние пользователей



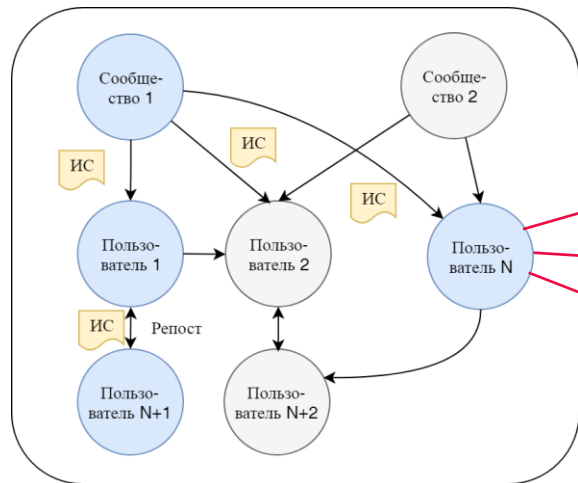
### Модель реакции

Реакция пользователей



# Описание модели

## Сущности



## Внутренние подмодели

### Модель генерации сообщений

Появление новых сообщений



### Модель активности

Состояние пользователей



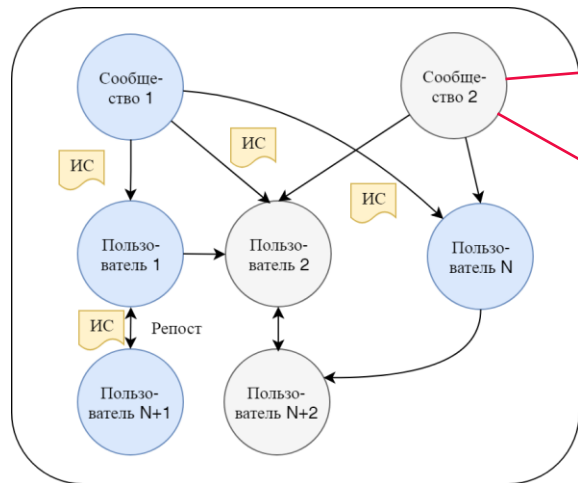
### Модель реакции

Отклик на сообщение



# Описание модели

## Сущности



## Внутренние подмодели

### Модель генерации сообщений

Появление новых сообщений



### Модель активности

Состояние пользователей

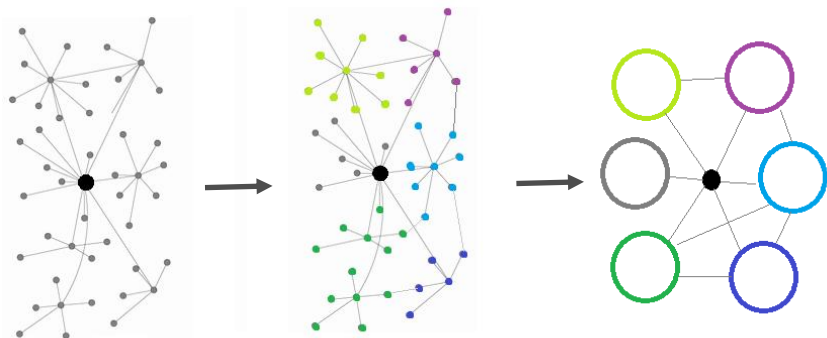


### Модель реакции

Отклик на сообщение



# Распространение информации



Мастер узел (M):

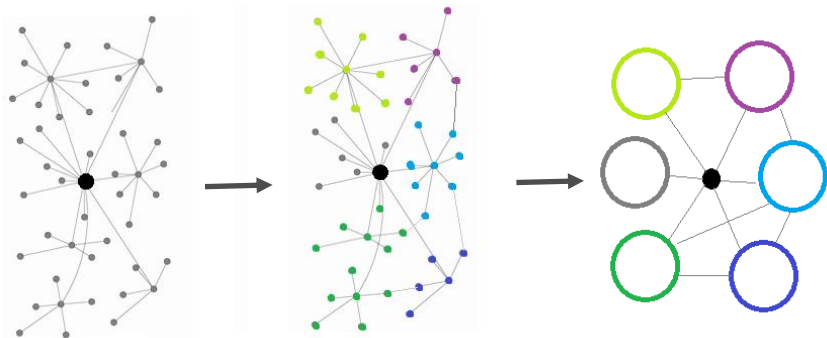
- хранение статистики
- генерация новостей

Ведомый узел –лист (L):

- хранение подсети
- распространение новостей на подсети



# Распространение информации

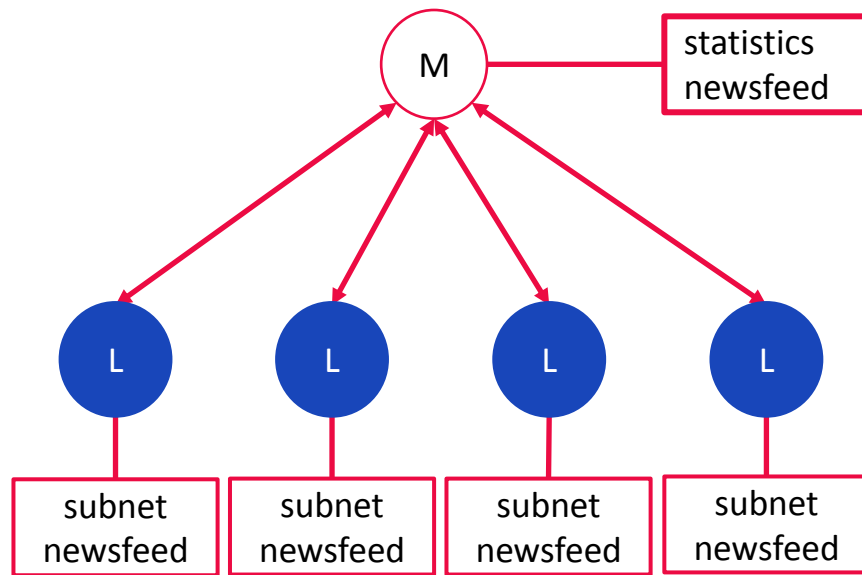


Мастер узел (M):

- хранение статистики
- генерация новостей

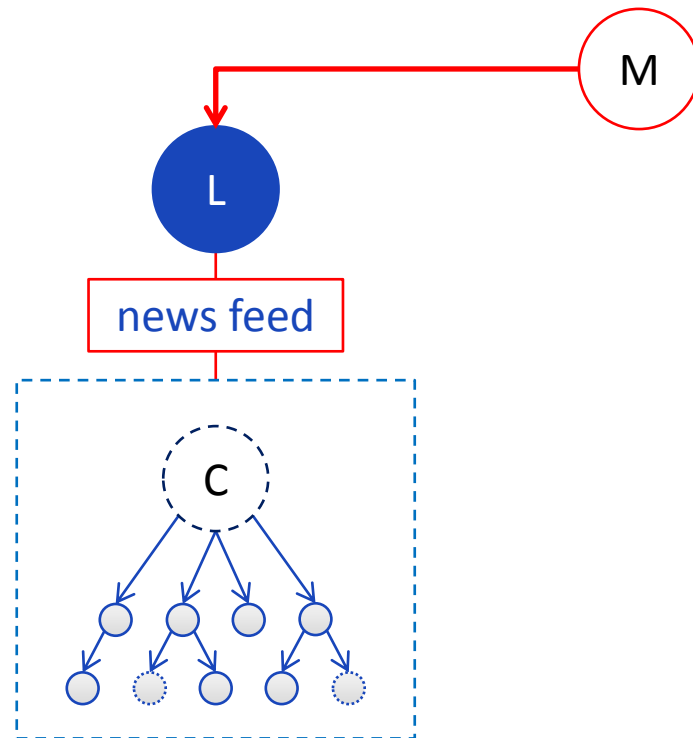
Ведомый узел –лист (L):

- хранение подсети
- распространение новостей на подсети



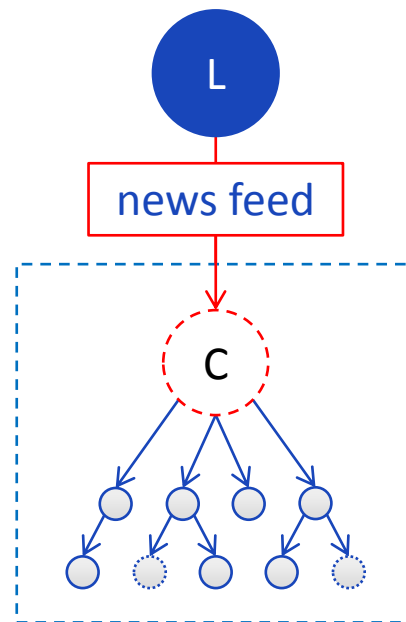
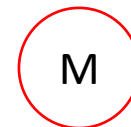
# Распространение информации на листе

- получение новых сообщений от Мастера;



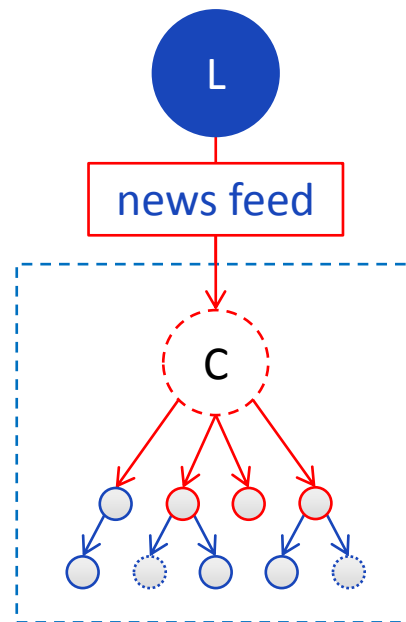
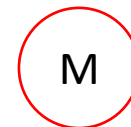
# Распространение информации на листе

- получение новых сообщений от Мастера;
- обработка каждой новости в ленте:



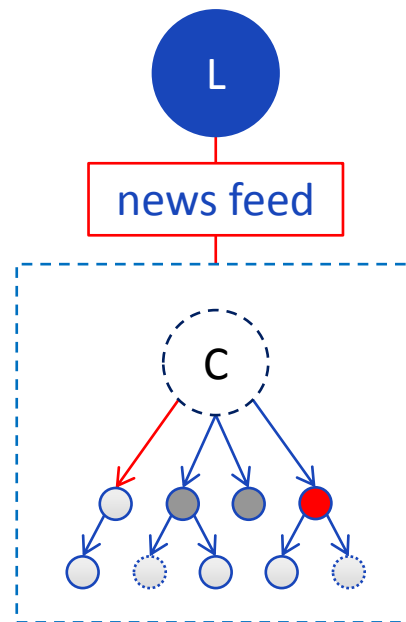
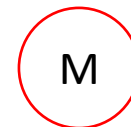
# Распространение информации на листе

- получение новых сообщений от Мастера;
- обработка каждой новости в ленте:
  - определение потенциальных зрителей для новости;
  - активные увидят новость на этой итерации;



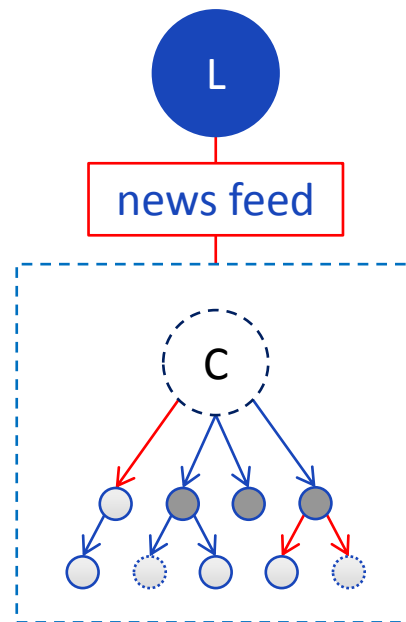
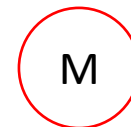
# Распространение информации на листе

- получение новых сообщений от Мастера;
- обработка каждой новости в ленте:
  - определение потенциальных зрителей для новости;
  - активные увидят новость на этой итерации;
  - модель реакции;



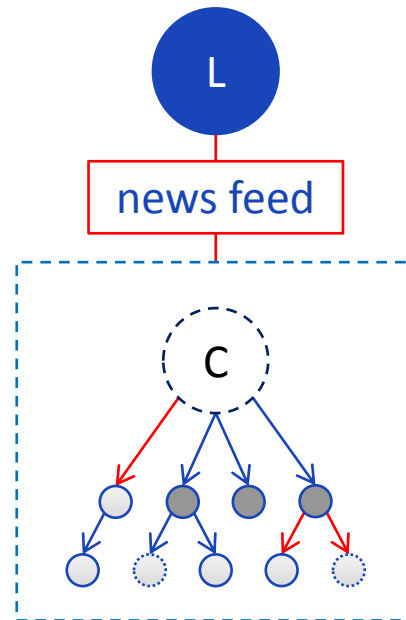
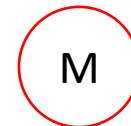
# Распространение информации на листе

- получение новых сообщений от Мастера;
- обработка каждой новости в ленте:
  - определение потенциальных зрителей для новости;
  - активные увидят новость на этой итерации;
  - модель реакции;
  - обработка подписчиков распространителей:



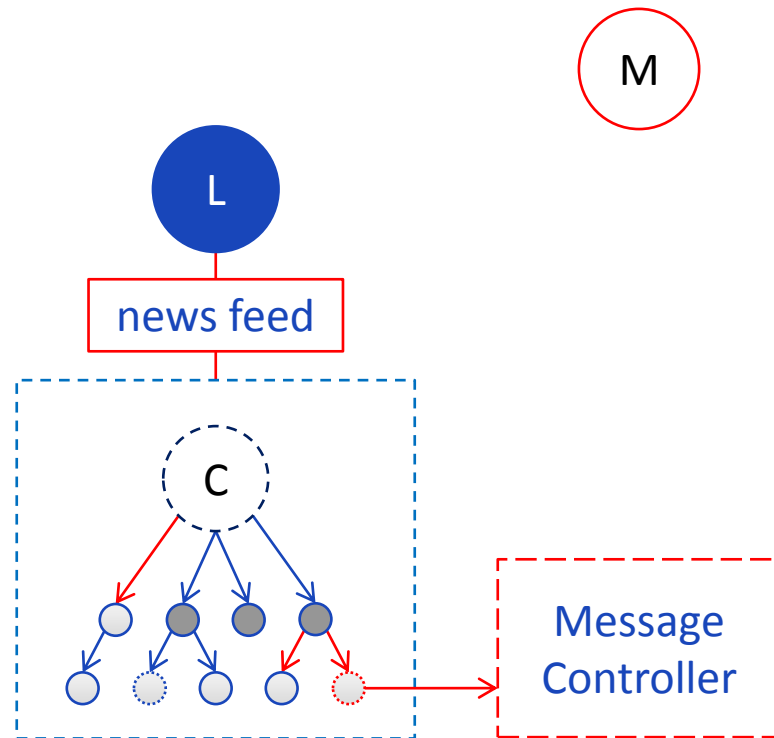
# Распространение информации на листе

- получение новых сообщений от Мастера;
- обработка каждой новости в ленте:
  - определение потенциальных зрителей для новости; активные увидят новость на этой итерации;
  - модель реакции;
  - обработка подписчиков распространителей:
    - если пользователь не видел новость, он становится потенциальным зрителем;



# Распространение информации на листе

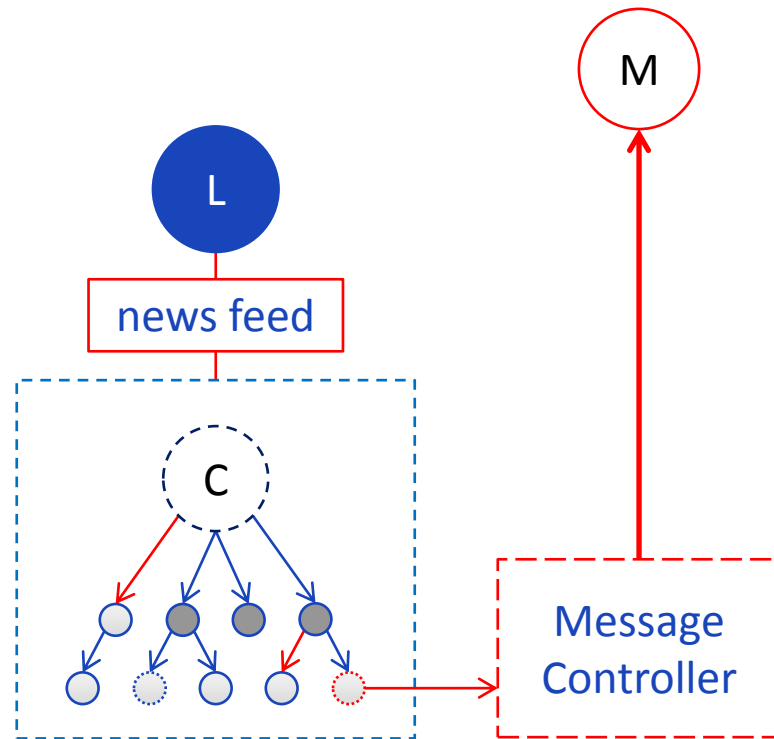
- получение новых сообщений от Мастера;
- обработка каждой новости в ленте:
  - определение потенциальных зрителей для новости; активные увидят новость на этой итерации;
  - модель реакции;
  - обработка подписчиков распространителей:
    - если пользователь не видел новость, он становится потенциальным зрителем;
    - пользователи, расположенные на других узлах добавляются в буфер;





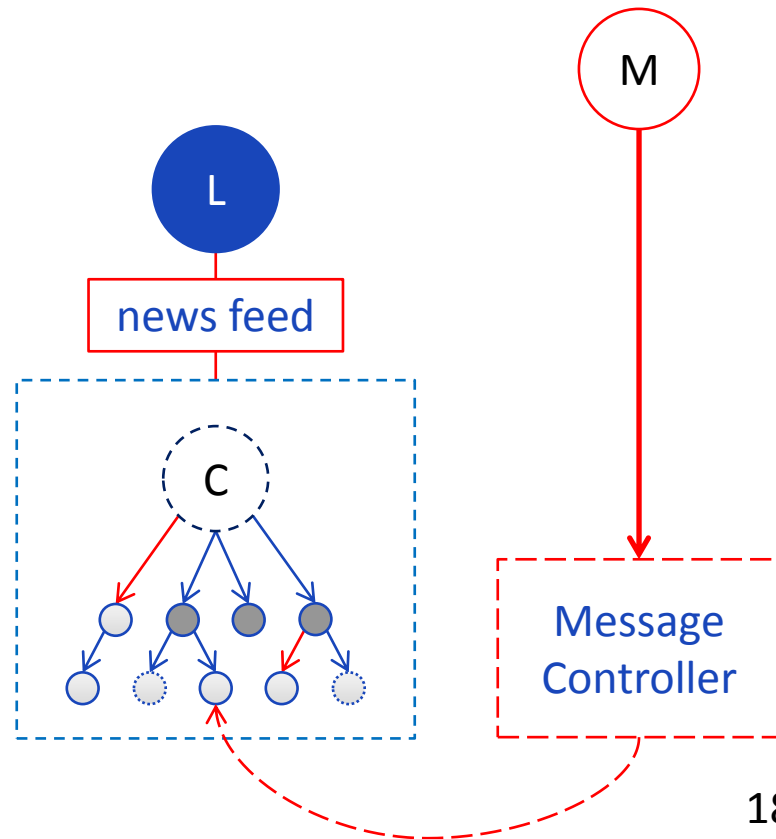
## Распространение информации на листе

- получение новых сообщений от Мастера;
- обработка каждой новости в ленте:
  - определение потенциальных зрителей для новости; активные увидят новость на этой итерации;
  - модель реакции;
  - обработка подписчиков распространителей:
    - если пользователь не видел новость, он становится потенциальным зрителем;
    - пользователи, расположенные на других узлах добавляются в буфер;
- отправка информационного сообщения и буфера на Мастер;



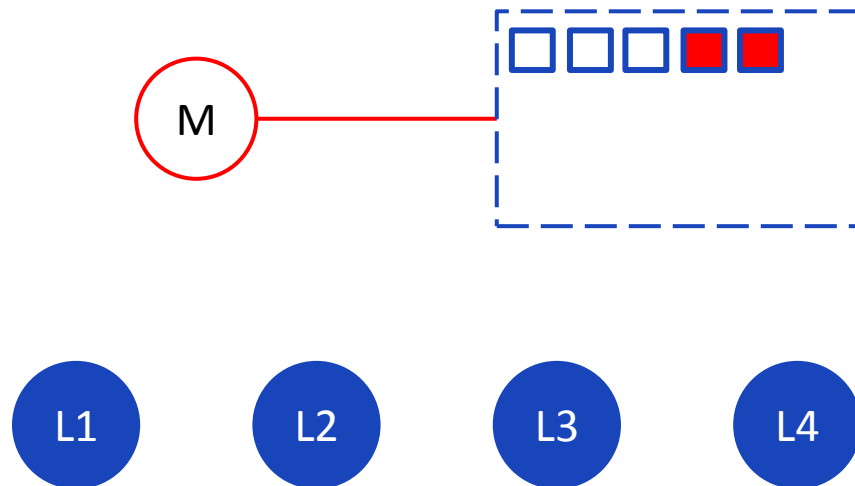
## Распространение информации на листе

- получение новых сообщений от Мастера;
- обработка каждой новости в ленте:
  - определение потенциальных зрителей для новости; активные увидят новость на этой итерации;
  - модель реакции;
  - обработка подписчиков распространителей:
    - если пользователь не видел новость, он становится потенциальным зрителем;
    - пользователи, расположенные на других узлах добавляются в буфер;
- отправка информационного сообщения и буфера на Мастер;
- получение сообщений от Мастера: список потенциальных зрителей, полученных на других листах, обновленная статистика по сообщениям.



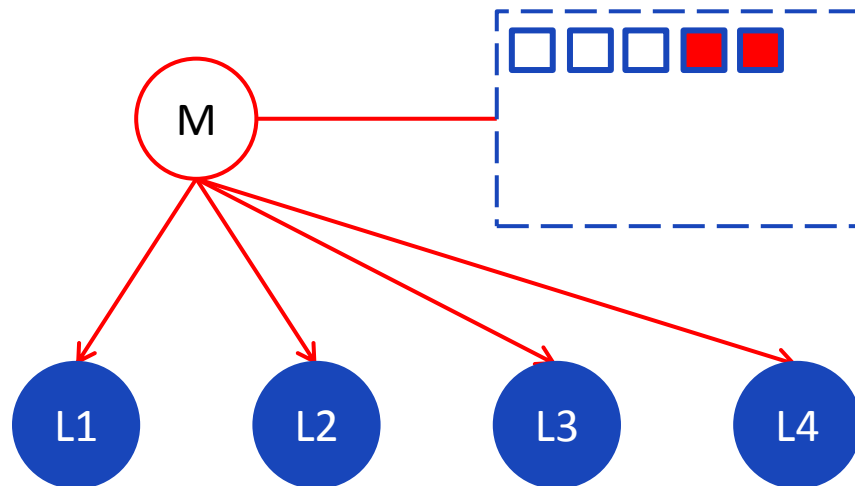
# Распространение информации на Мастере

- создание новостей;



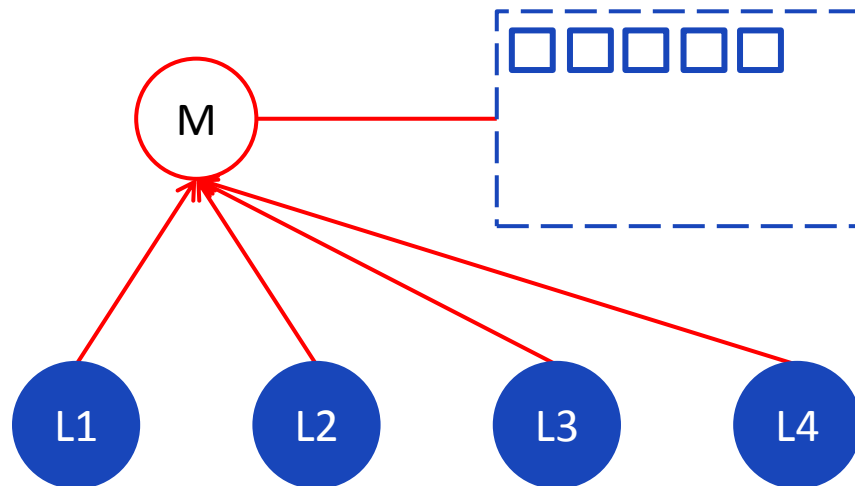
## Распространение информации на Мастере

- создание новостей;
- отправка новостей на листья;



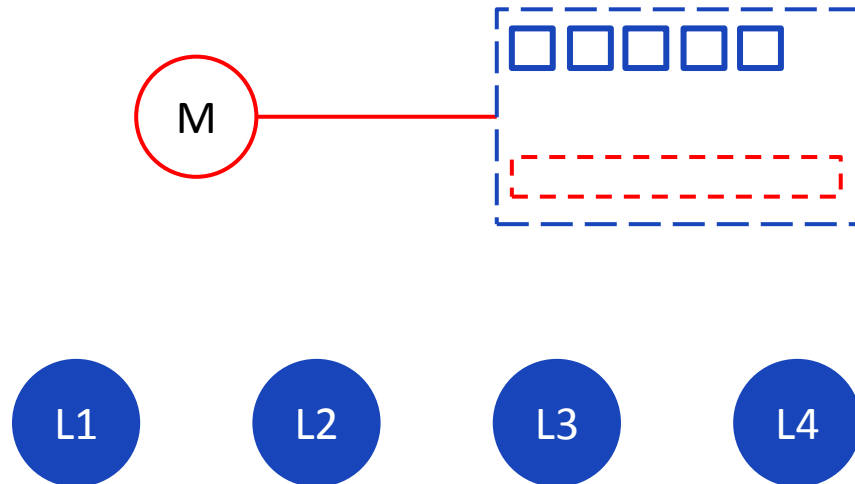
# Распространение информации на Мастере

- создание новостей;
- отправка новостей на листья;
- получение информационных сообщений от листьев;



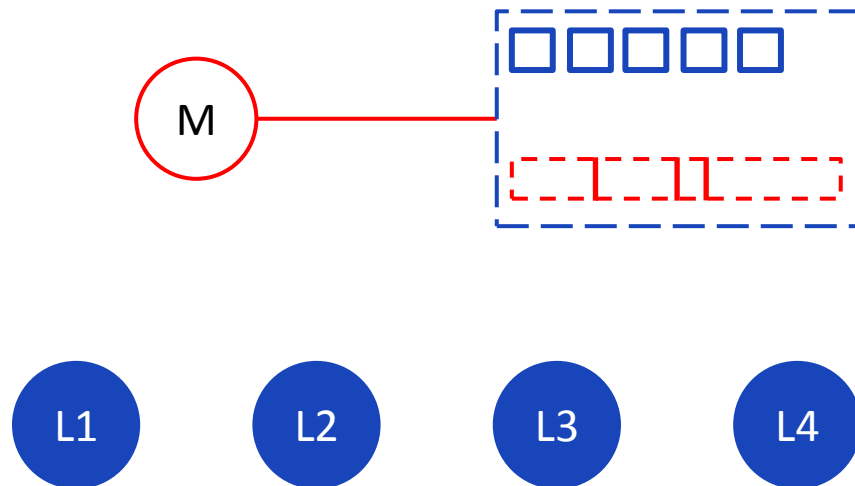
## Распространение информации на Мастере

- создание новостей;
- отправка новостей на листья;
- получение информационных сообщений от листьев;
- выделение памяти для получения новостей согласно информационным сообщениям;



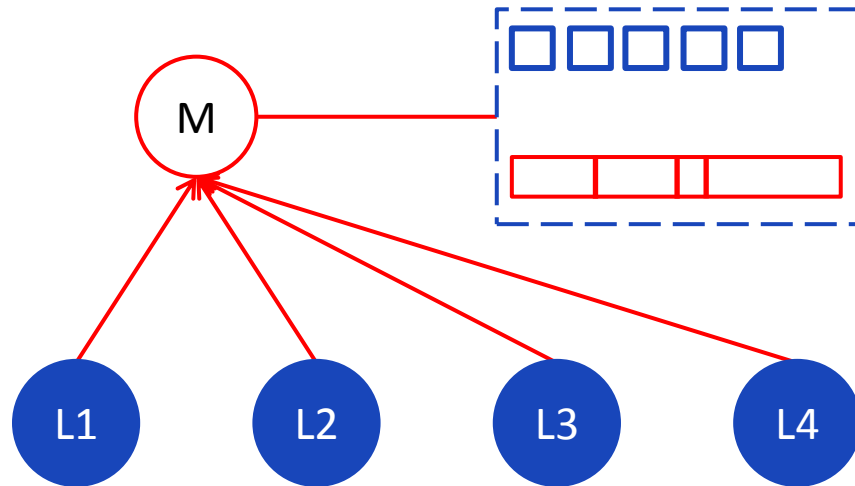
## Распространение информации на Мастере

- создание новостей;
- отправка новостей на листья;
- получение информационных сообщений от листьев;
- выделение памяти для получения новостей согласно информационным сообщениям;
- создание `MPI_Datatype` для каждого Листа согласно информационным сообщениям;



## Распространение информации на Мастере

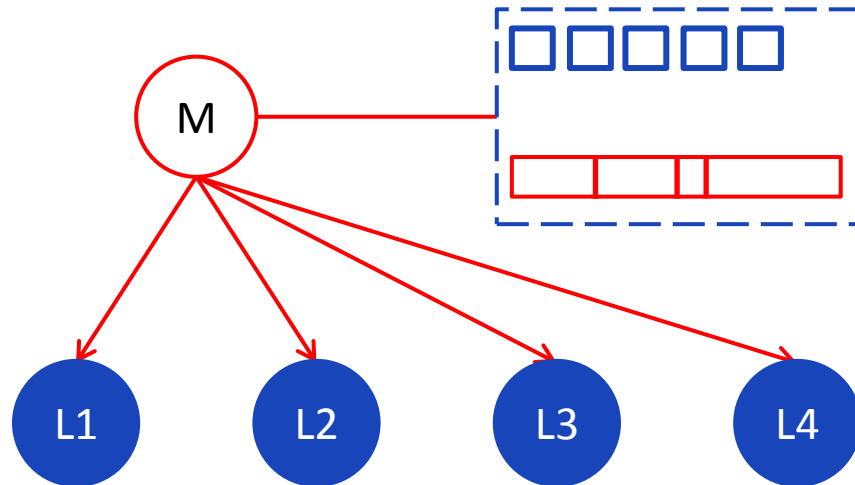
- создание новостей;
- отправка новостей на листья;
- получение информационных сообщений от листьев;
- выделение памяти для получения новостей согласно информационным сообщениям;
- создание `MPI_Datatype` для каждого Листа согласно информационным сообщениям;
- получение (`post`, `potential reader`) буфера от Листов в консолидированный пул;





## Распространение информации на Мастере

- создание новостей;
- отправка новостей на листья;
- получение информационных сообщений от листьев;
- выделение памяти для получения новостей согласно информационным сообщениям;
- создание `MPI_Datatype` для каждого Листа согласно информационным сообщениям;
- получение (`post`, `potential reader`) буфера от Листов в консолидированный пул;
- отправка на каждый Лист данных, соответствующих Листу.



## Описание данных

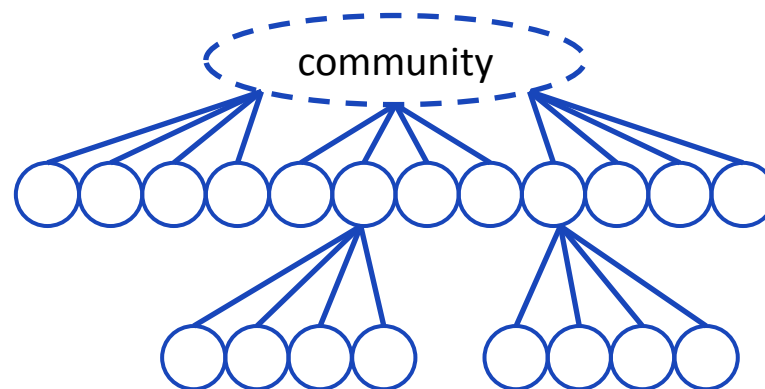
Эксперимент 1: данные ВК

- Большое сообщество ВК
- Все подписчики  
(294 345 users)
- Все друзья подписчиков

Всего: 33 768 036 пользователей

Эксперимент 2: Искусственный набор

- Множество сообществ
- Каждое сообщество имеет фиксированное число подписчиков
- Каждый подписчик имеет фиксированной число друзей

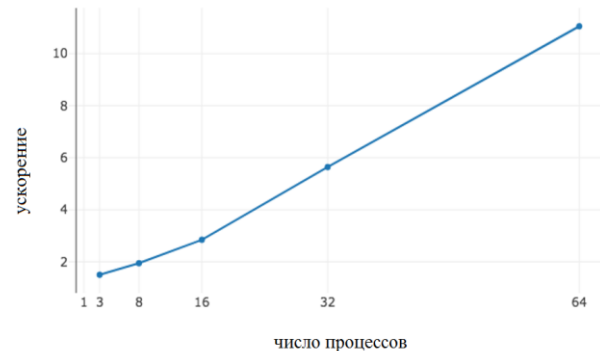


# Результаты

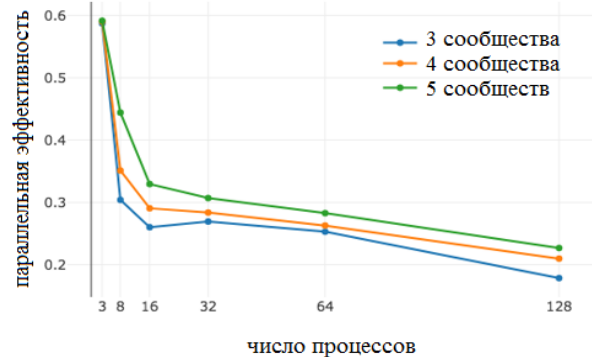
Суперкомпьютер Ломоносов (МГУ, Россия)



Эксперимент 1

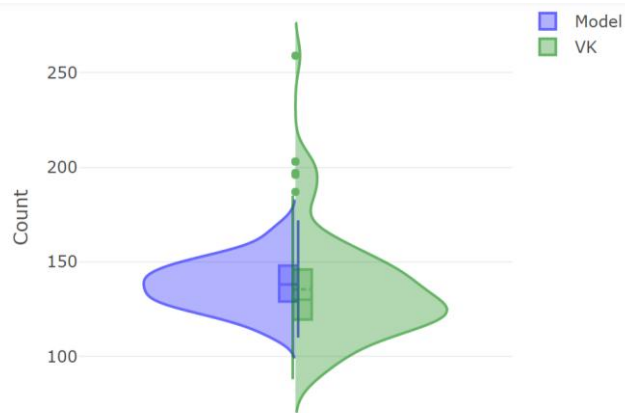


Эксперимент 2

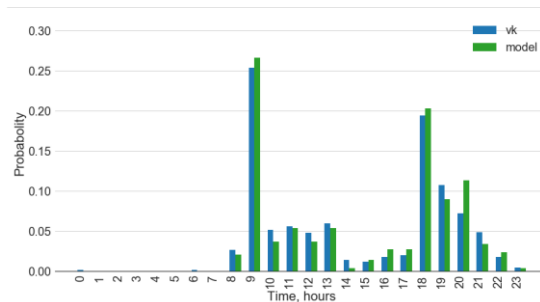


# Результаты: агрегированная динамика

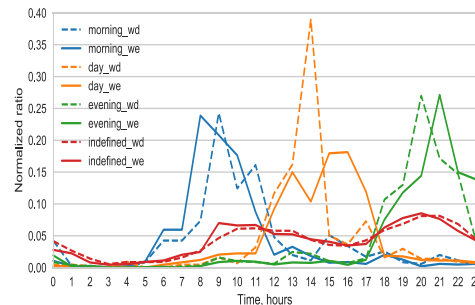
### Модель реакции



### Генеративная модель

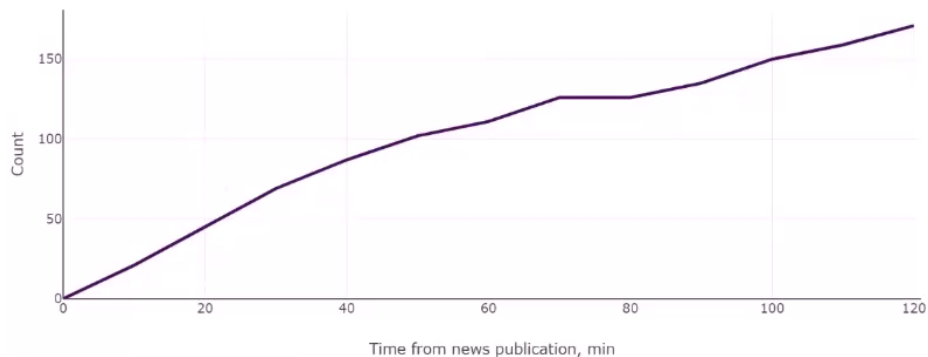


### Модель активности

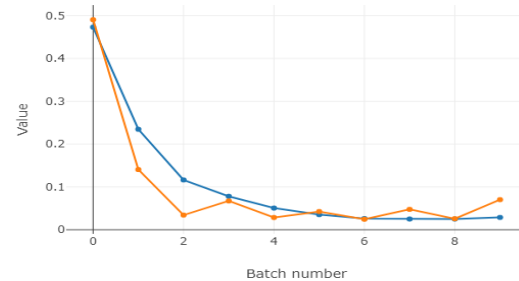


# Результаты: прогнозирование

## Прогнозирование с уточнением



## Ошибка прогноза



# Выводы

- предложен **алгоритм параллельного моделирования распространения информации** в сообществах, который позволяет изучать поведенческие модели пользователей и сообществ;
- **эффективность параллельного алгоритма** изучена для синтетических и реальных социальных сетей с использованием суперкомпьютера Ломоносов.

# Направления будущих исследований

- прогнозирование распространения информации в сообществе с **ассимиляцией данных** социальной сети;
- моделирование **эволюции пользовательских состояний** с точки зрения их уровня участия во взаимодействии с сообществом;
- улучшить **масштабируемость системы**, что сократит время вычислений для большого числа объектов.



# Спасибо за внимание!

Оксана Сеерюхина  
[oseveriukhina@gmail.com](mailto:oseveriukhina@gmail.com)

[www.ifmo.ru](http://www.ifmo.ru)

IT'sMO *re than a*  
UNIVERSITY