Использование параметрически заданных устойчивых распределений вероятностей в задаче нахождения надежного кратчайшего пути

А.А. Агафонов[1], **В.В. Мясников**[1], **А.И. Максимов**[1]

[1]Самарский национальный исследовательский университет им. академика С.П. Королева, Московское шоссе 34А, Самара, Россия, 443086

Аннотация. В работе рассматривается задача нахождения надежного кратчайшего пути в зависящей от времени стохастической транспортной сети, максимизирующего вероятность прибытия в пункт назначения в течение заданного интервала времени. Для описания времени прохождения сегментов дорожной сети предлагается использовать параметрически заданное устойчивое распределение вероятностей Леви. Использование устойчивых распределения позволит перейти от операции вычисления свертки к пересчету параметров плотности распределения, что значительно сократит время работы алгоритма. Экспериментальный анализ показал, что использование устойчивых распределений позволяет аппроксимировать истинное значение вероятности прибытия из заданной вершины в пункт назначения с помощью функции распределения Леви с малой ошибкой.

1. Введение

Навигационная задача нахождения кратчайшего пути в транспортной сети продолжает оставаться одной из наиболее актуальных задач в транспортных системах. Хотя существующие работы исследуют эту задачу в различных постановках, в т.ч. рассматривая зависящие от времени и стохастические транспортные сети, коммерческие системы работают с детерминированными сетями. Учет не только ожидаемого времени движения, но и дисперсии времени, т.е. надежности маршрута, делает задачу оптимальной маршрутизации вычислительно сложной.

В данной работе рассматривается задача нахождения надежного кратчайшего пути в зависящей от времени транспортной сети. Задача рассматривается в следующей постановке: определить оптимальную стратегию навигации, максимизирующую вероятность прибытия в пункт назначения в течение заранее определенного промежутка времени. В работе предлагается использовать устойчивое распределение вероятностей для описания времени прохождения сегментов дорожной сети, что позволит заменить операцию вычисления свертки на пересчет параметров плотности распределения.

2. Обзор литературы

В зависящих от времени стохастических транспортных сетях время прохождения дорожного сегмента представляется как случайная величина с функцией распределения, зависящей от времени [1]. В качестве решения задачи навигации рассматривается априорный кратчайший путь [2] или адаптивная стратегия навигации (дерево маршрутов)

[1, 3]. Условие оптимальности маршрута движения также может быть сформулировано в зависимости от используемой целевой функции:

- минимизация ожидаемого времени движения [1],
- максимизация вероятности прибытия в пункт назначения в течение заранее определенного промежутка времени [3, 4],
- минимизация промежутка времени, необходимого для прибытия в пункт назначения с указанной вероятностью [5].

В работе решается задача нахождения стратегии навигации, максимизирующей вероятность прибытия в пункт назначения за указанный интервал времени. Часто эта проблема обозначается как SOTA (Stochastic On-Time Arrival).

В статье [4] авторы сформулировали SOTA-проблему как задачу стохастического динамического программирования, для ее решения был применен стандартный метод последовательных аппроксимаций. Однако данный метод обладает плохой сходимостью. В качестве альтернативы, в [6] был предложен алгоритм дискретной аппроксимации для задачи SOTA, который сходится за конечное число шагов и работает за псевдополиномиальное время.

В статье [3] предложено точное решение SOTA-проблемы для сетей, в которых время прохождения сегментов является положительной величиной. Как и в работе [4], одним из этапов алгоритма является вычисление свертки, что является основной вычислительно сложной задачей. В общем виде свертка не может быть вычислена аналитически, и поэтому требуется схема дискретной аппроксимации. Решение, представленное в [3], позволяет осуществлять пакетное вычисление свертки, что является более эффективным, чем использование стандартного алгоритма, использованного в [6].

В работе [7] представлена модификация решения [3], учитывающая текущую и прогнозную информацию о параметрах транспортных потоков в сети.

В статье [8] авторы представили несколько методов ускорения алгоритма решения SOTA-проблемы, включая усовершенствованные алгоритмы вычисления свертки с помощью быстрого преобразования Фурье и алгоритмы вычисления свертки с нулевой задержкой [9], а также методы определения оптимального порядка вычисления стратегии навигации.

В данной работе для описания времени прохождения дорожного сегмента предлагается использовать устойчивое распределение вероятности Леви (Levy), что позволит заменить вычислительно сложную операции свертки на пересчет параметров функции распределения.

3. Основные обозначения и постановка задачи

Мы рассматриваем зависящую от времени стохастическую улично-дорожную сеть в виде ориентированного графа $G = (N, A, P)$, где N - множество вершин графа, A - множество ребер, P вероятностное описание времени прохождения сегментов сети.

В зависящих от времени стохастических сетях вес каждого сегмента $(i, j) \in A$ обычно представляется как случайная величина $T_{ij}(\tau)$ с плотностью вероятности $p_{ij}^{\tau}(t)$ зависящей от времени.

Оптимальная стратегия навигации определяется как стратегия максимизации вероятно-сти прибытия в конечную вершину $d \in N$ при наличии временного бюджета T . Пусть $u_i(t)$ - вероятность достижения конечной вершины d из вершины i за время t. Тогда оптимальная стратегия навигации формулируется следующим образом [3].

$$
\begin{aligned}
&u_i^\tau(t) = \max_{j \in N \wedge (i,j) \in A} \int_0^t p_{ij}^\tau(\theta) u_j^{\tau+\theta}(t-\theta) d\theta, \quad \forall i \in N \setminus \{d\}, t \in [0,T], \tau \geq 0 \\
&u_d^\tau(t) = 1, \quad t \in [0,T], \tau \geq 0
\end{aligned}
\tag{1}
$$

Для решения проблемы (1) используется дискретный алгоритм:

Алгоритм 1: Дискретный алгоритм решения SOTA

Шаг 0. Инициализация
$k = 0$
$u_i^k(x) = 0, \quad \forall i \in N,\ i \neq d,\ x \in \mathbb{N},\ 0 \leq x \leq \frac{T}{\Delta t}$
$u_d^k(x) = 1, \quad x \in \mathbb{N},\ 0 \leq x \leq \frac{T}{\Delta t}$
Шаг 1. Обновление
for $k = 1, 2, \cdots, L$ **do**
$\tau^k = k\delta$
$u_d^k(x) = 1, \quad x \in \mathbb{N}, 0 \leq x \leq \frac{T}{\Delta t}$
$u_i^k(x) = u_i^{k-1}(x), \qquad \forall i \in N, i \neq d, (i,j) \in A, x \in \mathbb{N}, 0 \leq x \leq \frac{\tau^k - \delta}{\Delta t}$
$u_i^k(x) = \max_j \sum_{h=0}^{x} p_{ij}(h) u_j^{k-1}(x-h)$
$\forall i \in N, i \neq d, (i,j) \in A, x \in \mathbb{N}, \frac{(\tau^k - \delta)}{\Delta t} + 1 \leq x \leq \frac{\tau^k}{\Delta t}$
end

В алгоритме Δt - интервал дискретизации, δ минимальное время прохождения дорожного сегмента в сети.
Выбор следующей вершины j с использованием бюджета времени t и вероятностей $u_i(x)$ производится следующим образом:

$$
j = \arg\max_{i \in N} u_i(t). \tag{2}
$$

Наиболее вычислительно сложным этапом работы алгоритма является вычисление свертки $\sum_{h=0}^{x} p_{ij}(h) u_j^{k-1}(x-h)$. В работе [7] в качестве плотности вероятности $p_{ij}(t)$ распределения времени движения на дорожном сегменте использовалось логнормальное распределение. В настоящей работе в качестве описания веса сегмента предлагается использовать устойчивое распределение Леви, что позволит заменить операцию свертки на пересчет параметров функции распределения.

4. Предлагаемая модель

4.1. Устойчивое распределение Леви

В теории вероятностей распределение называется устойчивым, если линейная комбинация двух независимых случайных величин с этим распределением имеет то же распределение с точностью до коэффициента сдвига и масштаба.
Функция распределения $F(x)$ называется устойчивой, если для любых действительных чисел $a_1 > 0$, $a_2 > 0$, b_1, b_2 найдутся числа $a > 0$, b такие, что имеет место равенство: $F(a_1 x + b_1) * F(a_2 x + b_2) = F(ax + b)$, где $*$ - операция свёртки.
В работе используется устойчивое распределение Леви. Плотность вероятности распределения Леви для области определения $x \geq \mu$ имеет вид:

$$f(x;\mu,c)=\sqrt{\frac{c}{2\pi}}\;\frac{e^{-c/2(x-\mu)}}{(x-\mu)^{3/2}}, \tag{3}$$

где μ - коэффициент сдвига, c - коэффициент масштаба.
Функция распределения имеет вид

$$F(x;\mu,c)=\operatorname{erfc}\left(\sqrt{c/\left(2(x-\mu)\right)}\right), \tag{4}$$

где $\operatorname{erfc}(z)$ - функция ошибок.
Если $X_1 \sim \text{Levy}(\mu_1, c_1)$, $X_2 \sim \text{Levy}(\mu_2, c_2)$, то $X_1 + X_2 \sim \text{Levy}(\mu, c)$, где

$$\begin{aligned} &\mu=\mu_1+\mu_2, \\ &|c|=\left(\sqrt{c_1}+\sqrt{c_2}\right)^2. \end{aligned} \tag{5}$$

4.2. Вычисление свертки

Рассмотрим подробнее операцию свертки в алгоритме 1. Введем обозначение

$$u_{ij}^k(x)=\sum_{h=0}^{x} p_{ij}(h)u_j^{k-1}(x-h). \tag{6}$$

Тогда операцию свертки можно записать в виде

$$u_i^k(x)=\max_j \sum_{h=0}^{x} p_{ij}(h)u_j^{k-1}(x-h)=\max_j u_{ij}^k(x). \tag{7}$$

Рассмотрим сначала выражение $u_{ij}^k(x)$. Учитывая, что $u_d^k(x)=1$, для ребер графа, входящих в конечную вершину d, мы можем получить значение:

$$u_{md}^k(x)=\sum_{h=0}^{x} p_{md}(h)u_d^{k-1}(x-h)=\sum_{h=0}^{x} p_{md}(h)=P_{md}(x), \quad \forall m\in V:\exists(m,d)\in E, \tag{8}$$

где $P_{md}(x)$ - функция распределения.
Далее, для предыдущих вершин графа $i:(i,m)\in E$, получим:

$$\begin{aligned} u_{im}^k(x)&=\sum_{h=0}^{x} p_{im}(h)u_m^{k-1}(x-h)=\sum_{h=0}^{x} p_{im}(h)\sum_{s=0}^{x-h} p_{md}(s)= \\ &=p_{im}(0)\sum_{s=0}^{x} p_{md}(s)+p_{im}(1)\sum_{s=0}^{x-1} p_{md}(s)+\dots+p_{im}(x)p_{md}(0)= \\ &=\sum_{l=0}^{x} p_{im}(l)p_{md}(x-l)+\sum_{l=0}^{x-1} p_{im}(l)p_{md}(x-1-l)+\dots+\sum_{l=0}^{x-x} p_{im}(l)p_{md}(x-x-l)= \\ &=\sum_{n=0}^{x}\sum_{l=0}^{x-n} p_{im}(l)p_{md}(x-n-l)=\sum_{n=0}^{x} p_{im+md}(x-n)=P_{im+md}(x), \end{aligned}$$

где $p_{im+md}(t)$ - плотность вероятности суммы случайных величин.

Аналогично могут быть посчитаны значения $u_{ij}^k(x)\ \forall i, j \in V$, что позволяет заменить операцию свертки в алгоритме 1 на вычисление значения функции распределения. Значение коэффициентов масштаба и сдвига рассчитываются по формуле (5).
Далее необходимо получить оценку функции $u_i^k(x) = \max_j u_{ij}^k(x)$. Будем аппроксимировать значение $u_i^k(x)$ функцией распределение Леви. Обозначим аппроксимируемую функцию как $F^*(x)$, искомую функцию как $F(x; \hat{\mu}, \hat{c})$. Тогда задача аппроксимации, в результате решения которой оцениваются искомые параметры $\alpha, \hat{\mu}, \hat{c}$, заключается в минимизации ошибки вида:

$$J = \epsilon^2 = \frac{1}{2} \sum_j \left(F^*(x_j) - \alpha F\left(x; \hat{\mu}, \hat{c}\right)\right)^2 \to \min_{\alpha, \hat{\mu}, \hat{c}} \quad (9)$$

5. Экспериментальные исследования

Целью экспериментальных исследований является сравнение оценок вероятности прибытия в пункт назначения, полученных с помощью вычисления свертки по формуле (7) и с помощью аппроксимационного приближения функцией распределения Леви $F(x; \hat{\mu}, \hat{c})$.
Примеры аппроксимации вероятности прибытия $u_i^k(x)$ представлены на рисунке 1. На рисунке 1 слева показан пример случая, когда вершина i связана с двумя вершинами в графе (т.е. существует два возможных маршрута движения из вершины i), справа - с тремя вершинами в графе.

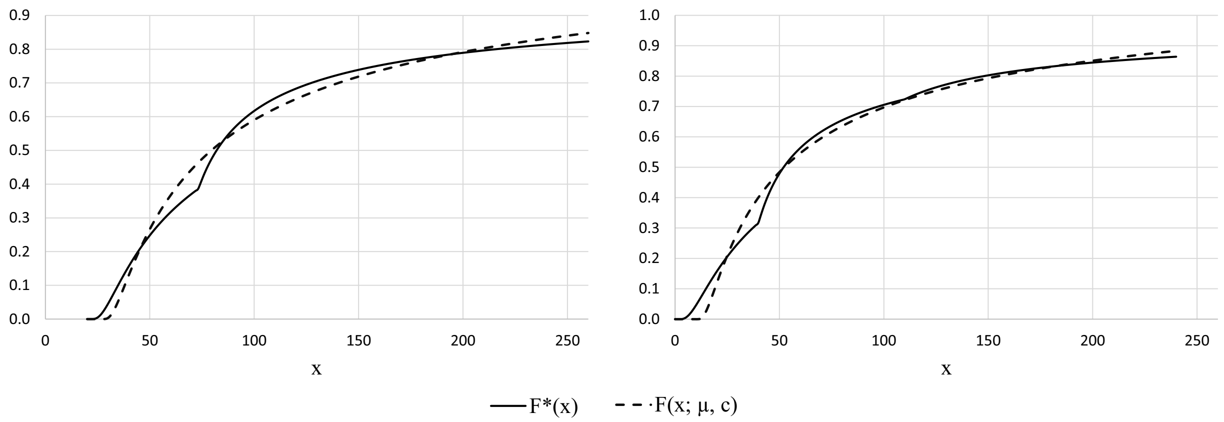

Рисунок 1. Аппроксимация вероятности прибытия для случая двух и трех связанных вершин.

Далее были проведены эксперименты по оценке ошибки аппроксимации. Для 100 случайно выбранных параметров распределения Леви μ, c была запущена процедура аппроксимации и посчитано среднеквадратичное отклонение по формуле

$$\mathrm{RMSE} = \frac{1}{n} \sqrt{\sum_{t=1}^{n} \left(x_t - \hat{x}_t\right)^2} \quad (10)$$

где x_t - точное значение, $\hat{x}_t$ - аппроксимация.
В результате экспериментов было получено, что средняя ошибка аппроксимации составляет

$$RMSE = 0.0174.$$

6. Заключение

В работе предложена модификация дискретного алгоритма решения задачи нахождения маршрута движения в зависящей от времени стохастической транспортной сети. Для уменьшения времени работы алгоритма предложено использовать устойчивое распределение Леви для описания времени прохождения сегментов дорожной сети. Такой подход позволяет перейти от вычислительно сложной операции свертки к вычислению значения функции распределения для оценки вероятности прибытия в вершину назначения за указанное время.

Результаты экспериментального анализа позволяют сделать вывод, что использование устойчивого распределения позволяет аппроксимировать точное значение вероятности прибытия из заданной вершины в пункт назначения с помощью функции распределения Леви с высокой точностью.

7. Литература

[1] Gao, S. Optimal ime-dependent networks / S. Gao, I. Chabini // Transportation Research Part B: Methodological. – 2006. – Vol. 40(2). – P. 93-122. DOI: 10.1016/j.trb.2005.02.001.

[2] Fu, L. Expected shortest paths in dynamic and stochastic traffic networks / L. Fu, L. Rilett // Transportation Research Part B: Methodological. – 1998. –Vol. 32(7). – P. 499-516. DOI: 10.1016/ S0191- 2615(98)00016-2.

[3] Samaranayake, S. A tractable class of algorithms for reliable routing in stochastic networks / S. Samaranayake, S. Blandin, A. Bayen // Transportation Research Part C: Emerging Technologies. – 2012. – Vol. 20(1). – P. 199-217. DOI: 10.1016/j.trc.2011.05.009.

[4] Fan, Y. Optimal routing for maximizing the travel time reliability / Y. Fan, Y. Nie // Networks and Spatial Economics. – 2006. – Vol. 6(3-4). – P. 333-344. DOI: 10.1007/s11067-006-9287-6.

[5] Nie, Y. Shortest path problem considering on-time arrival probability / Y. Nie, X. Wu // Transportation Research Part B: Methodological. – 2009. – Vol. 43(6). – P. 597-613. DOI: 10.1016/j.trb.2009.01.008.

[6] Nie, Y. Arriving-on-time problem: Discrete algorithm that ensures convergence / Y. Nie, Y. Fan // Transportation Research Record. – 2006. – Vol. 1964. – P. 193-200. DOI: 10.3141/1964-21.

[7] Агафонов, А.А. Метод определения надёжного кратчайшего пути в зависящей от времени стохастической сети и его применение в геоинформационных задачах управления транспортом / А.А. Агафонов, В.В. Мясников // Компьютерная оптика. – 2016. – Т. 40, № 2. – С. 275-283. DOI: 10.18287/2412-6179-2016-40-2-275-283.

[8] Samaranayake, S. Speedup techniques for the stochastic on-time arrival problem / S. Samaranayake, S. Blandin, A. Bayen // DRPOS. – 2012. – Vol. 25. – P. 83-96. DOI: 10.4230/OASIcs.ATMOS.2012.83.

[9] Gardner, W.G. Efficient convolution without input-output delay / W.G. Gardner // AES: Journal of the Audio Engineering Society. – 1995. – Vol. 43(3). – P. 127-136.

Благодарности

Работа выполнена при частичной финансовой поддержке грантов РФФИ № 18-29-03135-мк, № 18-07-00605 А.

The use of stable probability distributions in the reliable routing problem

A.A. Agafonov[1], **V.V. Myasnikov**[1], **A.I. Maksimov**[1]

[1]Samara National Research University, Moskovskoe Shosse 34A, Samara, Russia, 443086

Abstract. In this paper, we consider the reliable shortest path problem in a time-dependent stochastic transportation network. The problem is to find a routing policy that maximizies the probability of arriving at the destination point on time. It is proposed to use parametrically defined stable probability distribution Levy to describe the travel time of road segments. The use of stable distributions allow us to replace the convolution operation with the distribution value, and significantly reduce the execution time of the algorithm. Experimental analysis have shown that the use of stable distributions allows to approximate the exact value of the arrival probability to a destination with a low approximation error.