В данной работе мы применяем методы, основанные на анализе вероятностей катастроф [1], чтобы оценивать риски особо значимых событий в интернет сети.

Постановка задачи: есть некий канал связи интернет-сети, на котором установлено оборудование, считывающее характеристики этого канала. Характеристики в такой постановке можно моделировать как некий временной ряд значений замеров оборудования этих характеристик. Событием, риск которого нужно оценивать, будем считать превышение значений характеристикой некоторого порога.

В качестве данных мы использовали запись характеристик канала на территории ВМК МГУ, предоставленную кафедрой АСВК. Данные представляют из себя многомерный временной ряд с измерениями нескольких характеристик канала связи, проводимые на протяжении около 6.5 часов с частотой в 1 секунду, или 23900 измерений суммарно.

Для оценки рисков катастрофических событий мы используем следующие теоремы:

Теорема 1. *Пусть S_n - сумма n случайных величин, не обязательно одинаково распределенных. Предположим, что случайная величина N_n имеет геометрическое распределение с параметром p_n, причем $p_n \to 0$ при $n \to \infty$. Предположим, что существуют конечные $\gamma > 0$ и $b > 0$ такие, что*

$$\frac{S_n}{bn^\gamma} \Longrightarrow 1 \quad (n \to \infty),$$

где $\Longrightarrow$ означает сходимость по распределению. Тогда

$$\lim_{n\to\infty} \sup_{x\geq 0} \left| P(p_n^\gamma S_{N_n} \geq bx) - \exp\{-x^{1/\gamma}\} \right| = 0.$$

Эта теорема является версией закона больших чисел для случайных сумм неодинаково распределенных случайных величин, в частности описывая закон распределения сумм такого рода и подвязывая его к распределению Вейбулла-Гнеденко. Также, параметры b и γ имеют смысл частоты катастрофы и степени изменения частоты катастроф со временем.

Второй теоремой является теорема Балкема-Пикандса-Де Хаана, обосновывающая использование обобщенного распределения Парето для моделирования распределения критических значений характеристики.

Идейно, первая теорема дает модель вероятности катастрофического события в момент времени, но требует функцию распределения этих катастроф, что дает нам вторая теорема. Таким образом, можно в явном виде вычислительно дешево расчитать эти вероятности.

Практически, метод работает для оценки рисков в совокупности на всем временном ряду, но также предоставляется модификация метода, которая пересчитывает вероятности на скользящем окне. Благодаря этому, метод способен моделировать изменение параметров b и γ со временем, а значит более точно описывать риски с учетом изменения состояния сети во времени.

Данный метод оказался крайне эффективен в моделировании ряда характеристик, таких как средняя задержка и доля потерь пакетов за единицу времени при различных критических порогах, тем самым позволяя адаптировать метод для различных уровней устойчивости сети. Дальнейшим направлением является исследование других характеристик и применение алгоритма для встраивания в симулированный или реальный сетевой маршрутизатор и оценка работы такого подхода.