УДК 681.3.06

Е.Е. Слядников

Модель распределенной информационно-телекоммуникационной системы с пакетной передачей данных

Сформулирована модель распределенной информационно-телекоммуникационной системы с пакетной передачей данных для труднодоступных объектов. Показано, что такая система описывается степенным законом распределения для вероятности реализации связи между двумя узлами и проявляет свойства безмасштабной сети (малого мира), обладая ближней структурой, как у однородных систем, и дальней структурой подобно случайным системам.

Ключевые слова: распределенная информационно-телекоммуникационная система, модель малого мира.

Предметом исследования данной статьи являются распределенные информационно-телекоммуникационные системы с пакетной передачей данных. При проведении анализа таких систем необходимо учитывать следующие особенности: системы являются многоканальными, их каналы связи имеют различную физическую природу, системы имеют иерархическую структуру и распределенную топологию [1].

В работе [2] исследовалась распределенная информационно-телекоммуникационная система с пакетной передачей данных для труднодоступных объектов, исходя из предположения, что она имеет упорядоченную пространственную конфигурацию и строго определенные взаимодействия между ее составляющими. Развивая подход [2], в настоящей работе приняты во внимание эффекты неупорядоченности (случайности), которые присущи любой реальной системе связи. Хотя наличие или отсутствие связи является исключительно геометрическим свойством, не имеющим отношения к тепловым флуктуациям, многие характеристики рассматриваемой системы, среди них разрыв связей и возникновение перколяций, критические индексы и применимость методов ренорм-групп, схожи с характеристиками фазовых переходов [3, 4].

Техническая и программная конфигурация системы передачи данных определяется исходя из топологии и условий местности ее развертывания и строится на основе иерархической комбинации трех функционально-ориентированных аппаратно-программных комплексов (АПК): АПК-ЦСД – центра сбора данных (обычно размещается в управлениях или территориальных гидрометеоцентрах); АПК-КРС – кустовых центров сбора данных (обычно размещается на одной из наиболее доступных радирующих станций); АПК-МЕТЕО – для метеостанций.

Простейшую модель рассмотренной выше распределенной системы с пакетной передачей данных для труднодоступных объектов можно сформулировать как совокупность множества узлов и множества ребер, соединяющих узлы. Множество узлов системы состоит из пунктов сбора первичных данных, кустовых центров сбора данных, центров сбора данных. Ребра системы образованы пучком каналов связи, которые могут иметь различную физическую природу. Топологию сети пакетной передачи данных удобно представлять с помощью плоского графа, состоящего из множества узлов и множества ребер связи, соединяющих смежную пару узлов. Примем во внимание эффекты случайности, присущие любой реальной системе связи, тогда плоский граф становится случайным, т.е. в его узлах существует конечная вероятность возникновения связи между смежными узлами.

Рассмотрим случайный граф, между соседними узлами которого существуют связи и часть x этих связей произвольно удаляется. В такой системе случайность определяется параметром x – частью связей, которые были удалены, или частью $p = 1 - x$ связей, которые остались. Для $x = 0$ или $p = 1$ непрерывная сетка связей покрывает всю систему полностью, и ситуация не изменится, даже если x незначительно увеличится, а p уменьшится. При $p = 0$ сетки связей не существует, и картина существенно не изменится и при малых ненулевых значениях p. Поэтому разумно предположить, что имеется некоторое критическое значение p_c такое, при котором для $p < p_c$ непрерывной бесконечной траектории, идущей через узлы по связям, не существует, а реализуются лишь отдельные изолированные кластеры из связанных узлов. В случае $p > p_c$ существуют бесконечные кластеры, покрывающие всю систему.

Критическое значение p_c известно как порог протекания, это явление называется протеканием по связям (перколяцией), поскольку если представить, что связи – это поры в некотором материале, то существование бесконечного кластера приводит к протеканию жидкости через этот материал. Перколяционный переход – это своего рода геометрическое фазовое превращение, в котором критическая концентрация p_c разделяет фазу конечных кластеров при $p<p_c$ и фазу бесконечных кластеров при $p > p_c$. Это аналогично температурным фазовым превращениям, где концентрация p играет роль температуры.

Проанализируем переход от однородного графа к случайному. Однородный граф – это граф, в котором каждый узел связан с k ближайшими соседями, а в случайном графе связан с k случайными соседями. Два параметра характеризуют свойства «ближнего порядка» графа и «дальнего порядка» графа, содержащего N вершин и $Nk/2$ хорд. Параметр «дальнего порядка» – это среднее минимальное число хорд $h(l)$, необходимое для того, чтобы перейти из одной вершины в другую, находящуюся на расстоянии l, и оно определяет кратчайший путь между двумя вершинами. Характер зависимости кратчайшего пути от l различен для однородных и случайных графов в то время как для однородного графа $h(l) \sim l$, а для случайного графа $h(l) \sim \ln(l)$. Можно также определить коэффициент кластеризации C, который является средней долей соседних вершин, соединенных с заданной вершиной. Этот параметр определяет ближний порядок и очевидно, что он приближается к единице для однородного графа и мал для случайного графа. Отсюда следует, что h и C велики для однородного графа и малы для случайного графа. Еще один тип графа, занимающий промежуточное положение между однородным и случайным, предложен в [5].

При формулировке модели информационно-телекоммуникационные системы с пакетной передачей данных для труднодоступных объектов возьмем за основу модель малого мира [5]. Рассмотрим граф G с N вершинами и K ребрами, который является невзвешенным (ребра эквивалентны), редким ($K \ll N(N-1)/2$) и связанным. Поэтому G можно представить, просто задавая матрицу смежности NЧN, чей матричный элемент a_{ij} равен 1, если есть ребро, соединяющее вершину i с вершиной j, и равен 0, если ребра нет. Важным свойством G является степень вершины i, т.е. число k_i ребер, выходящих из вершины i. Среднее значение k_i равно $k=2K/N$. Когда матрица a_{ij} задана, ее можно использовать для вычисления матрицы путей с самой короткой длиной d_{ij} между двумя вершинами i и j. Тот факт, что G – связанный граф, подразумевает, что матричные элементы d_{ij} положительные и конечные для любого $i \neq j$. Для количественного описания структурных свойств G вводятся две различных величины: характерная длина пути L и коэффициент кластеризации C. L есть среднее расстояние между двумя вершинами $L = [N(N-1)]^{-1}\sum d_{ij}$, а C – локальная характеристика, равная $C = N^{-1}\sum C_i$. Здесь C_i – число ребер, составляющих G_i – подграф ближайших соседей вершины i, разделенный на максимальное возможное число $k_i\,(k_i-1)/2$.

При описании реальной распределенной информационно-телекоммуникационной системы с пакетной передачей данных для труднодоступных объектов ограничения модели малого мира (невзвешенность, редкость, связанность) являются неприемлемыми, поэтому необходимо формулировать более общую модель. Покажем, что, во-первых, поведение такой системы можно описать с помощью единственной переменной с ясным физическим значением – *эффективность передачи информации* E; во-вторых, величины $1/L$ и C могут быть получены как предельные случаи переменной E соответственно в глобальном и локальном масштабе; в-третьих, можно убрать перечисленные ограничения на модель. Представим распределенную сеть связи с пакетной передачей данных как взвешенный (и, возможно, даже нередкий и несвязанный) граф G. Такой граф нуждается в двух матрицах: матрице смежности a_{ij} и матрицы l_{ij} физических расстояний, имеющей смысл некоторой метрики. Число l_{ij} может быть расстоянием между этими двумя вершинами или силой их возможного взаимодействия. Предполагаем, что l_{ij} известно, даже если в графе нет никакого ребра между i и j. Конечно, в специфическом случае невзвешенного графа $l_{ij} = 1$ для любого $i \neq j$. Элемент матрицы путей с самой короткой длиной d_{ij} между двумя вершинами i и j является наименьшей суммой физических расстояний по всем возможным путям в графе от i до j.

Поскольку система с пакетной передачей данных параллельна (каждая вершина одновременно посылает информацию по сети, через ее ребра), то эффективность связи ε_{ij} между вершинами i и j тогда может быть определена как величина, обратно пропорциональная самому короткому расстоянию: $\varepsilon_{ij} = 1/\ d_{ij}$ для любого i, j. Когда нет никакого пути в графе между i и j, $d_{ij} = +\infty$ и, следовательно, $\varepsilon_{ij} = 0$. Средняя эффективность передачи информации в графе G может быть определена как

$$E(G) = [N(N-1)]^{-1} \sum_{i \neq j \in G} \varepsilon_{ij} = [N(N-1)]^{-1} \sum_{i \neq j \in G} d^{-1}{}_{ij}. \qquad (1)$$

Чтобы нормировать E, рассмотрим идеальный случай G_{id}, в котором граф G имеет все $N(N-1)/2$ возможных ребра. В таком случае информация распространяется самым эффективным способом для $d_{ij} \geq l_{ij}$ для любого i, j, и E принимает максимальное значение при $E(G_{id})=[N(N-1)]^{-1}\sum l^{-1}{}_{ij}$. Эффективность $E(G)$, используемая ниже, всегда делится на $E(G_{id})$ и поэтому $0 \leq E(G) \leq 1$. Хотя равенство $E=1$ выполняется, когда есть ребра между всеми вершинами, реальные сети тоже могут достигнуть высокого значения E.

Используем единственную величину E, чтобы одновременно анализировать и локальное, и глобальное поведение системы, а не две различных переменные L и C. Величина E в (1) есть глобальная эффективность G, и поэтому определим ее как E_{glob}. Определим локальную эффективность как среднюю эффективность локальных подграфов, $E_{\text{lok}}=N^{-1} \sum E(G_i)$. Так как i не принадлежит G_i, локальная эффективность E_{lok} показывает насколько система толерантна к ошибкам (сбоям). Она показывает, насколько эффективна связь между первыми ближайшими соседями i, когда i удален.

Таким образом, определение распределенной информационно-телекоммуникационной системы с пакетной передачей данных для труднодоступных объектов может быть качественно сформулировано в терминах информационного потока: система имеет высокие E_{glob} и E_{lok}, т.е., очень эффективна в глобальной и локальной связи. Это определение правильно, как и для невзвешенного, так и для взвешенного графа и может быть применено к несвязанным и (или) нередким графам.

Введем вероятность $P(l) \sim l^{-\delta}$ существования связи между двумя вершинами в графе на расстоянии l. Если $\delta = 0$, $P(l)$ постоянная и связи всех размеров одинаково вероятны, как в случайном графе. С другой стороны, если величина δ очень велика, то дальние связи малы, поэтому получается однородный граф, содержащий только ближние связи. При δ, принимающей промежуточное значение, получаем модель малого мира. Получается, что при $\delta < \delta_1$ система ведет себя как случайный граф, при $\delta < \delta_1 < \delta_2$ – как граф малого мира, а при $\delta > \delta_2$ – как однородный граф. Предположим, что степенной закон распределения $P(l) \sim l^{-\delta}$ вероятности реализации связи между вершинами выполняется для распределенной системы с пакетной передачей данных. Центральная предельная теорема утверждает, что для независимых случайных событий с конечным средним a и дисперсией σ случайная переменная x подчиняется нормальному закону распределения:

$$P(x) = (2\pi\sigma)^{-1/2} \exp[-(x-a)^2 / 2\sigma].$$

Колоколообразная кривая нормального закона распределения, вероятно, является наиболее часто встречаемой в литературе. В нашем случае вместо закона нормального распределения имеет место степенной. То есть наша система – безмасштабная система, т. к. степенной закон является масштабно-инвариантным и не включает в себя характерные длины. Эти случаи аналогичны равновесной системе вблизи критической точки, в которой малые длины не имеют значения, а важна только характерная длина, являющаяся длиной корреляции. Последняя расходится в критической точке, поэтому в этом состоянии характерная длина отсутствует.

Наш граф, состоящий из N вершин, имеет k связей с другими вершинами. Разные вершины могут иметь разное количество связей, поэтому необходимо ввести распределение количества связей k, функцию $P(k)$. Для безмасштабных систем $P(k)$ является степенной функцией вида $P(k) \sim k^{-\gamma}$. Чтобы показать, как проявляется степенной закон, сопоставим каждой вершине i случайное число x_i [x_i принадлежит $(0,\infty)$], которое назовем рангом вершины i и будем использовать для обозначения приоритетности этой вершины. Пусть функция $f(x)$ определяет распределение рангов и представляет собой степенную функцию. Кроме того, предположим, что вероятность существования связи между двумя вершинами i и j равна $\Psi(x_i, x_j)$. В случайных графах эта вероятность будет одинакова и постоянна для всех соединенных вершин: $\Psi(x_i, x_j) = b$.

Средняя инцидентность вершины ранга x равна

$$<k(x)>=\int_0^\infty \Psi(x.y)f(y)dy = NF(x)\,, \qquad (2)$$

поэтому $x = F^{-1}(<k>/N)$.

С другой стороны, по определению

$$<k>=\int_0^\infty kP(k)dk\,. \qquad (3)$$

Поскольку

$$P(k) = f(x)dx/dk\,,$$

то, объединив уравнения (2) и (3), получим

$$P(k) = f[F^{-1}(<k>/N)]d[F^{-1}(<k>/N)]/dk\,.$$

Пусть $\Psi(x_i,\ x_j)= (x_i\ x_j)/\ x^2{}_{\max}$, где $x_{\max}$ – максимальный ранг в сети. Тогда

$$<k>=(Nx/x^2_{\max})\int_0^\infty yf(y)dy = Nx<x>/x^2_{\max} \text{ и } P(k) = (x^2_{\max}/N<x>)f(k[x^2_{\max}/N<x>])\,. \qquad (4)$$

Из соотношения (4) следует, что распределение инцидентности $P(k)$ полностью определяется распределением рангов $f(x)$, поэтому если $f(x)$ подчиняется степенному закону распределения, то аналогичное распределение будет у инцидентности $P(k)$.

Для дальнейшего анализа модели системы используем теорию среднего поля Ландау для фазовых превращений [4]. Для этого сопоставим системе термодинамический потенциал G, который является функцией параметра порядка Δ и внешнего поля h: $G=G(\Delta,h)$. Как обычно, используем условие, чтобы параметр порядка обращался в ноль в неупорядоченной фазе.

Вершина с инцидентностью k испытывает на себе воздействие в среднем k соседних вершин, т.е. кроме внешнего поля h на нее действует среднее поле kh. Разложение Ландау для $G(\Delta,h)$ содержит оба поля h и kh, и это разложение можно записать как

$$G(\Delta,h) = -\Delta h + \sum_k P(k)\Phi(\Delta,k\Delta)\,. \qquad (5)$$

То есть оно зависит от функции распределения $P(k)$. Дальнейший анализ (5) основывается на разложении в ряд Тейлора функции $\Phi(x,\ y)$ для малых x и y

$$\Phi(x,y) = \sum_m \sum_l \varphi_{ml} x^m y^l$$

и исследовании первых членов этого ряда. Коэффициенты данного разложения имеют вид

$$(d^nG(\Delta,h)/d\Delta^n)_{\Delta=0} = n!\sum_{l=0}^{n} \varphi_{n-l,l} <k^l>\,. \qquad (6)$$

Средние значения $<k^l>$, входящие в уравнение (6), полностью определяются функцией распределения $P(k)$. Для степенного распределения $P(k)\sim k^{-\gamma}$ получаем, что

$$<k^l>=\int_1^\infty k^l P(k)dk \approx k^{l-\gamma+1}\,, \text{ при } k\to\infty\,.$$

Это означает, что все моменты, для которых $l \geq \gamma$, расходятся. Функция (5) содержит в качестве коэффициентов эти моменты, поэтому теория среднего поля Ландау не работает. Равновесное состояние разрушается сильными флуктуациями. В графе наиболее сильно связанные вершины, в которых имеют место сильные флуктуации локальной инцидентности, являются причиной того, что теория среднего поля неприменима. С увеличением γ в $P(k)\sim k^{-\gamma}$ случайное число вершин с большим числом связей уменьшается, и фазовое превращение происходит по типу среднего поля. Расчеты показывают, что это происходит при $\gamma > 5$, а при $\gamma = 5$ появляются логарифмические поправки. Таким образом, G будет содержать следующие члены:

$$\Delta^3 \ln\Delta \text{ при } \gamma = 5,$$

$$\Delta^{\gamma-1} \text{ при } 2<\gamma < 3 \text{ и } 2<\gamma < 3,$$

$$\Delta^2 \ln\Delta \text{ при } \gamma = 3.$$

Полученные результаты позволяют сделать следующие выводы:

1. Распределенная информационно-телекоммуникационная система с пакетной передачей данных для труднодоступных объектов проявляет свойства безмасштабной сети (малого мира), обладая ближней структурой, как у однородных систем, и дальней структурой подобно случайным системам.

2. Безмасштабная сеть очень неоднородная и очень немного вершин (концентратов) имеют большое количество связей, а большинство вершин имеют лишь несколько связей. Такая структура довольно типична для распределенной информационно-телекоммуникационной системы с пакетной передачей данных для труднодоступных объектов, которая соединяет множества узлов связи, используя концентраторы (например, АПК-КРС) в качестве промежуточных станций для соединения маршрутов. Динамика этих сетей такова, что отвечает принципу «богатый богатеет еще больше» или «концентратор получает все».

3. Такая сеть описывается степенным законом распределения для вероятности реализации связи между двумя вершинами и имеет топологическую структуру, заметно отличающуюся от обычной, которая характерна для сетей, подчиняющихся нормальному закону распределения. Главное условие, которое приводит к степенному закону распределения, – это избирательное присоединение, обозначающее, что новые связи стремятся соединить уже существующие вершины, имеющее множество других связей.

4. Распределение инцидентности полностью определяется распределением рангов, поэтому если распределение рангов подчиняется степенному закону распределения, то аналогичное распределение будет у инцидентности.

5. Для явлений малых миров небольшая доля дальних связей приводит систему в класс универсальности, характерной для приближения среднего поля. Однако в безмасштабных системах, несмотря на наличие дальних связей, ситуация иная. Поведение по типу среднего поля проявляется только при достаточно больших значениях γ в функции распределения $k, \gamma > 5$, в то время как при малых значениях γ, критические свойства не зависят от модели, но не универсальны, так что критические индексы зависят от γ.

6. Важное отличие между экспоненциальной и безмашстабной сетью – это различная реакция на повреждение. Под повреждением понимается устранение некоторых вершин и всех связей, идущих от вершины. Безмасштабные сети весьма устойчивы по отношению к случайным повреждениям, однако очень чувствительны к намеренным повреждениям, направленным на концентраторы, которые приводят к разрушению большого количества связей и нарушению взаимодействия между различными частями сети. Если агрессивное воздействие будет направлено на уничтожение концентраторов, то даже если устранить менее 10 % таких узлов, сеть связи распадется на несвязанные между собой кластеры.

Литература

1. Сонькин М.А. Архитектура и общая технология функционирования территориально распределенных аппаратно-программных комплексов с пакетной передачей данных / М.А. Сонькин, Е.Е. Слядников // Известия Том. политех. ун-та. – 2006. – Т. 309, № 5. – С. 131–139.

2. Сонькин М.А. Об одном подходе к оптимизации функционирования многоканальной системы связи для труднодоступных объектов / М.А. Сонькин, Е.Е. Слядников // Вычислительные технологии. – 2007. – Т. 12. – Спец. вып. 1. – С. 17–22.

3. Хакен Г. Синергетика. – М.: Мир, 1980. – 404 с.

4. Ландау Л.Д. Статистическая физика. Часть 1. / Л.Д. Ландау, Е.М. Лифшиц. – М.: Наука, 1989. – 521 с.

5. Watts D.J. Collective dynamics Small-World Networks/D.J. Watts, S.H. Strogatz // Nature. – 1998. – Vol. 393, – № 4. – P. 440–442.

Слядников Евгений Евгеньевич
Д.ф.-м.н., с.н.с. отдела проблем информатизации
ТНЦ СО РАН, профессор каф. электронных приборов ТУСУРа
Тел.: (382-2) 49-13-12
Эл. адрес: opi@hq.tsc.ru

E.E. Slyadnikov

Model of the distributed information-telecommunication system with batch data transmission

The model of the distributed information-telecommunication system with batch data transmission for remote objects is formulated. It is shown, that such system is described by the sedate law of distribution for probability of realization of connection between two units and shows properties without scale networks (the small world), possessing near structure, as at homogeneous systems, and distant structure, similarly to casual systems.

Keywords: distributed information-telecommunication system, model of small world.