УДК 004.7:004.89(045)

Лукашенко В.В., к.т.н.

ИНФОРМАЦИОННАЯ МОДЕЛЬ ПРОЦЕССА ЭКСПЕРТИЗЫ МЕТОДОВ УПРАВЛЕНИЯ КОРПОРАТИВНОЙ СЕТЬЮ

Национального авиационного университета

Рассмотрены задачи системы управления корпоративной сетью. Предложена математическая модель экспертизы сети на наличие коллизий.

Введение

Одной из важных задач системы управления корпоративной сетью является задача мониторинга, которая связана с оценкой статистической информации по следующим кратериям [1]: время реакции системы; пропускная способность реального или виртуального канала связи между двумя конечными абонентами сети; интенсивность трафика в отдельных сегментах и каналах сети; вероятность искажения данных при их передаче через сеть; коэффициент готовности сети или ее определенной транспортной службы.

Без средств анализа производительности и надежности поставщик услуг сети или отдел информационных технологий предприятия не сможет ни проконтролировать, ни тем более обеспечить нужный уровень обслуживания для конечных пользователей сети.

Важно отметить, что успешное решение задач мониторинга во многом зависит от качества подбора методов экспертизы и критериев оценки статистической информации. Для решения этих задач целесообразно применить методы экспертных оценок.

Анализ существующих методов

В настоящее время не существует общепринятой научно обоснованной классификации методов экспертных оценок и тем более – однозначных рекомендаций по их применению. Однако в работе [2] при построении алгоритма обработки результатов экспертизы рекомендовано использовать порядковые шкалы. Использование порядковых шкал позволяет различать загрузку сети трафиком, когда критерий оценки не задан в явном виде.

Цель

Целью является получение устойчивых оценок эффективности при использовании математической модели экспертизы сети на наличие коллизий при заданных исходных данных относительно количественного состава и квалификации администраторов.

Оценка загрузки сети в процессе работы

В качестве примера рассмотрим алгоритмы обработки результатов экспертизы загрузки сети. Пусть m системных администраторов производят оценку n узлов по l показателям. Результаты оценки представлены в виде величин x_{ij}^{h}, где j – номер системного администратора, i - номер узла, h – номер показателя (признака) сравнения. Если оценка загрузки узлов произведена методом ранжирования, то величины x_{ij}^{h} представляют собой ранги. Если оценка загрузки узлов выполнена методом непосредственной оценки или методом последовательного сравнения, то величины x_{ij}^{h} представляют собой числа из некоторого отрезка числовой оси, или баллы. Обработка результатов оценки существенно зависит от рассмотренных методов измерения.

Рассмотрим случай, когда величины $x_{ij}^h (i = 1,...,n; j = 1,2,...,m; h = 1,2,...,l)$ получены методами непосредственной оценки или последовательного сравнения, т. е. x_{ij}^h являются числами, или баллами. Для получения от администраторов полного состояния сети можно воспользоваться средним значением оценки для каждого узла [3]

$$x_i = \sum_{h=1}^{l} \sum_{j=1}^{m} q_h x_{ij}^h k_j (i = 1,2,...,n), \quad (1)$$

где q_h - коэффициенты весов показателей сравнения узлов, k_j - коэффициенты компетентности администраторов. Коэффициенты весов показателей и компетентности администраторов являются нормированными величинами [4] $\sum_{h=1}^{l} q_h = 1; \sum_{j=1}^{m} k_j = 1.$ Коэффициенты весов показателей могут быть определены экспертным путем. Если q_{hj} - коэффициент веса h-го показателя, предоставляемый j-м администратором, то средний коэффициент веса h-го показателя по всем администратором равен [5] $q_h = \sum_{j=1}^{m} q_{hj} k_j (h = 1,2,...,l).$

Коэффициенты компетентности администраторов можно вычислить по апостериорным данным, т. е. по результатам оценки состояния загрузки узлов. Основной идеей этого вычисления является предположение о том, что компетентность администраторов сети должна оцениваться по степени согласованности их оценок с оценкой системы мониторинга сети. Алгоритм вычисления коэффициентов компетентности администраторов имеет вид рекуррентной процедуры [6]:

$$x_i^t = \sum_{j=1}^{m} x_{ij} k_j^{t-1}, \quad i = 1,2,...,n; \quad (2)$$

$$\lambda^t = \sum_{i=1}^{n} \sum_{j=1}^{m} x_{ij} x_i^t, \quad t = 1,2,...; \quad (3)$$

$$k_j^t = \frac{1}{\lambda^t} \sum_{i=1}^{n} x_{ij} x_i^t; \ \sum_{j=1}^{m} k_j^t = 1, \quad j = 1,2,...,m. \quad (4)$$

Вычисления начинаются с $t = 1$. В формуле (2) начальные значения коэффициентов компетентности принимаются одинаковыми и равными $k_j^o = 1/m$. Тогда в соответствии с (2) оценки загрузки узлов системой мониторинга первого приближения и равны средним арифметическим значениям оценок системных администраторов:

$$x_i^1 = \frac{1}{m} \sum_{j=1}^{m} x_{ij}, \quad i = 1,2,...,n.$$

Далее вычисляется величина λ^1 по формуле (3): $\lambda^1 = \sum_{i=1}^{n} \sum_{j=1}^{m} x_{ij} x_i^1$ и значение коэффициентов компетентности первого приближения по формуле (4): $k_j^1 = \frac{1}{\lambda^1} \sum_{i=1}^{n} x_{ij} x_i^1$. Используя коэффициенты компетентности первого приближения, можно повторить весь процесс вычисления по формулам (2 – 4), и получить вторые приближения величин x_i^2, λ^2, k_j^2. Повторение рекуррентной процедуры вычислений оценок узлов и коэффициентов компетентности естественно ставит вопрос о ее сходимости. Для рассмотрения этого вопроса исключим из уравнений (2, 4) переменные k_j^{t-1} и x_i^t и представим эти уравнения в векторной форме [7]:

$$x^t = \frac{1}{\lambda^{t-1}} B x^{t-1}; k^t = \frac{1}{\lambda^t} C k^{t-1} (t = 1,2,...), \quad (5)$$

где матрицы B размерности $n \times n$ и C размерности $m \times m$ равны соответственно $B = XX', \quad C = X'X, \quad X = \|x_{ij}\|$.

Величина λ^t в уравнениях (5) определяется по формуле (3). Если матрицы B и C неотрицательны и неразложимы, то, как это следует из теоремы Перрона – Фробениуса, при $t \to \infty$ векторы x^t и k^t - сходятся к собственным векторам матриц B и C, соответствующим максимальным собственным числам этих матриц [8] $x = \lim_{t\to\infty} x^t, k = \lim_{t\to\infty} k^t$. Предельные значения векторов x и k можно вычислить из уравнений [9]:

$$Bx = \lambda_B x, \quad \sum_{i=1}^{n} x_i = 1, \quad |B - \lambda_B E| = 0,$$
$$Ck = \lambda_C k, \quad \sum_{j=1}^{m} k_j = 1, \quad |C - \lambda_C E| = 0, \qquad (6)$$

где λ_B, λ_C - максимальные собственные числа матриц B и C.

Условие неотрицательности матриц B и C легко выполняется выбором неотрицательных элементов x_{ij} матрицы X оценок объектов экспертами. Условие неразложимости матриц B и C практически выполняется, поскольку, если эти матрицы разложимы, то это означает, что администраторы и узлы распадаются на независимые группы. При этом каждая группа администраторов оценивает только узлы своей группы. Естественно, что получать групповую оценку в этом случае нет смысла. Таким образом, условия неотрицательности и неразложимости матриц B и C, а следовательно, и условия сходимости процедур (2 – 4) в практических условиях выполняются. Следует заметить, что практическое вычисление векторов групповой оценки загрузки узлов и коэффициентов компетентности проще выполнять по рекуррентным формулам (2-4). Определение предельных значений этих векторов по уравнению (6) требует применения вычислительной техники.

Рассмотрим теперь случай, когда системные администраторы производят оценку множества объектов методом ранжирования. Пусть величины x_{ij} есть ранги. Обработка результатов ранжирования заключается в построении обобщенной ранжировки. Для построения такой ранжировки введем конечномерное дискретное пространство ранжировок и метрику в этом пространстве. Каждая ранжировка множества узлов j- м администратором есть точка R_j в пространстве ранжировок. Очевидно, что $a_{kk} = 0$, поскольку каждый узел эквивалентен самому себе. Элементы матрицы $\|a_{kl}\|$ антисимметричны: $a_{kl} = -a_{lk}$. Достаточно простым способом построения обобщенной ранжировки является способ суммирования рангов. Он заключается в ранжировании объектов по величинам сумм рангов, полученных каждым объектом от всех администраторов. Для матрицы рангов $\|r_{ij}\|$ составляются суммы [9] $r_i = \sum_{j=1}^{m} r_{ij}$, $i = 1, 2, \ldots, n$. Далее объекты упорядочиваются по цепочке неравенств $r_1 < r_2 < \ldots < r_n$. Для учета компетентности администраторов достаточно умножить каждую i-ю ранжировку на коэффициент компетентности j-го администратора $0 \le k_j \le 1$. В этом случае вычисление суммы рангов для i-го объекта производится по следующей формуле [9]: $r_i = \sum_{j=1}^{m} r_{ij} k_j$, $i = 1, 2, \ldots, n$. Обобщенная ранжировка с учетом компетентности администраторова строится на основе упорядочения сумм рангов для всех узлов.

При ранжировании объектов администраторы обычно расходятся во мнениях по решаемой проблеме. В связи с этим возникает необходимость количественной оценки степени согласия администраторов. Получение количественной меры согласованности мнений администраторов позволяет более обоснованно интерпретировать причины в расхождении мнений. В настоящее время известны две меры согласованности мнений группы администраторов: дисперсионный и энтропийный коэффициенты согласия.

Дисперсионный коэффициент согласия. Рассмотрим матрицу результатов ранжировки n узлов группой из m администраторов $\left\| r_{ij} \right\|$ (j=1,…,m; i=1,…,n), где r_{ij} - ранг, присваиваемый j-м администратором i-му узлу. Составим суммы рангов по каждому столбцу. В результате получим вектор с компонентами [9]:

$$r_i = \sum_{j=1}^{m} r_{ij} \qquad i = 1,2,\ldots,n. \tag{7}$$

Величины r_i рассмотрим как реализации случайной величины и найдем оценку дисперсии. Как известно, оптимальная по критерию минимума среднего квадрата ошибки оценка дисперсии определяется формулой [9]:

$$D = \frac{1}{n-1}\sum_{i=1}^{n}\left(r_i - \bar{r}\right)^2, \tag{8}$$

где $\bar{r}$ - оценка математического ожидания, равная:

$$\bar{r} = \frac{1}{n}\sum_{i=1}^{n} r_i. \tag{9}$$

Дисперсионный коэффициент согласия W определяется как отношение оценки дисперсии (8) к максимальному значению этой оценки

$$W = \frac{D}{D_{max}}. \tag{10}$$

Коэффициент согласия изменяется от нуля до единицы, поскольку $0 \le D \le D_{max}$.

Вычислим максимальное значение оценки дисперсии для случая отсутствия связанных рангов (все узлы различны). Предварительно покажем, что оценка математического ожидания зависит только от числа узлов и количества администраторов. Подставляя в (9) значение r_i из (7), получаем:

$$\bar{r} = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{m} r_{ij}. \tag{11}$$

Рассмотрим вначале суммирование по i при фиксированном j. Это есть сумма рангов для j-го администратора. Поскольку администратор использует для ранжировки натуральные числа от 1 до n, то, как известно, сумма натуральных чисел от 1 до n равна:

$$\sum_{i=1}^{n} r_{ij} = \frac{n(n+1)}{2}. \tag{12}$$

Подставляя (12) в (11), получаем $\bar{r} = \frac{1}{n}\cdot\sum_{j=1}^{m}\frac{n(n+1)}{2} = \frac{(n+1)m}{2}$. Таким образом, среднее значение зависит только от числа администраторов m и количества узлов n. Для вычисления максимального значения оценки дисперсии подставим в (8) значение r_i из (7) и возведем в квадрат двучлен в круглой скобке. В результате получаем

$$D = \frac{1}{n-1}\left[\sum_{i=1}^{n}\left(\sum_{j=1}^{m} r_{ij}\right)^2 - 2\bar{r}\sum_{i=1}^{n}\sum_{j=1}^{m} r_{ij} + n\bar{r}^2\right].$$

Учитывая, что из (11) следует $\sum_{i=1}^{n}\sum_{j=1}^{m} r_{ij} = n\bar{r}$, получаем [9]:

$$D = \frac{1}{n-1}\left[\sum_{i=1}^{n}\left(\sum_{j=1}^{m} r_{ij}\right)^2 - n\bar{r}^2\right]. \tag{13}$$

Максимальное значение дисперсии достигается при наибольшем значении первого члена в квадратных скобках. Величина этого члена зависит от рас положения рангов - натуральных чисел в каждой строке i. Пусть, например, все m администраторов дали одинаковую ранжировку для всех n узлов, что соответствует максимальному значению коэффициента ранговой корреляции Спирмэна ρ_s [7]:

$$\rho_s = \frac{\sum_{i=1}^{n}\left(r_{ij}-\bar{r}\right)\left(r_{ik}-\bar{r}\right)}{\sqrt{\sum_{i=1}^{n}\left(r_{ij}-\bar{r}\right)^2+\sum_{i=1}^{n}\left(r_{ik}-\bar{r}\right)^2}} = \frac{12}{n^3-n}\sum_{i=1}^{n}\left(r_{ij}-\frac{n+1}{2}\right)\left(r_{ik}-\frac{n+1}{2}\right), \quad (14)$$

где r_{ij} и r_{ik} - ранги, присваиваемые i-му узлу j-м и k-м администраторами соответственно. При $j=k$, очевидно, из выражения (14) получаем оценку дисперсии D.

В каждой строке матрицы $\left\|r_{ij}\right\|$ будут расположены одинаковые числа. Следовательно, суммирование рангов в каждой i-й строке дает m-кратное повторение i-го числа: $\sum_{j=1}^{m} r_{ij} = im.$ Теперь предположим, что администраторы дают несовпадающие ранжировки, например, для случая $m=n$ все администраторы присваивают разные ранги одному объекту. Тогда

$$\sum_{i=1}^{n}\left(\sum_{j=1}^{m} r_{ij}\right)^2 = \sum_{i=1}^{n}\left(\frac{m(m+1)}{2}\right)^2 = \frac{m^2(m+1)n}{4}.$$

Сравнивая это выражение с m^2 при $m=n$, убеждаемся, что первый член в квадратных скобках формулы (13) равен второму члену и, следовательно, оценка дисперсии равна нулю. Таким образом, случай полного совпадения ранжировок администраторов соответствует максимальному значению оценки дисперсии. Приравняв в (14) $j=k$ и подставляя (14) в (13), получаем

$$D_{max} = \frac{m^2(n^3-n)}{12(n-1)}. \quad (15)$$

Введем обозначение

$$S = \sum_{i=1}^{n}\left(\sum_{j=1}^{m} r_{ij} - \bar{r}\right)^2. \quad (16)$$

Используя (16), запишем оценку дисперсии (8) в виде

$$D = \frac{1}{n-1}S. \quad (17)$$

Подставляя (15 – 17) в (10) и сокращая на множитель $(n-1)$, запишем окончательное выражение для коэффициента согласия:

$$W = \frac{12S}{m^2(n^3-n)}. \quad (18)$$

Данная формула определяет коэффициент согласия для случая отсутствия связанных рангов. Если в ранжировках имеются связанные ранги, то максимальное значение дисперсии в знаменателе формулы (10) становится меньше, чем при отсутствии связанных рангов. Можно показать [9], что при наличии связанных рангов коэффициент согласия вычисляется по формуле:

$$W = \frac{12S}{m^2(n^3-n) - m\sum_{j=1}^{m} T_j}, \quad (19)$$

где $T_j = \sum_{k=1}^{H_j}\left(h_k^3 - h_k\right)$. В формуле (19) T_j - показатель связанных рангов в j-й ранжировке, H_j - число групп равных рангов в j-й ранжировке, h_k - число равных рангов в k-й группе связанных рангов при ранжировке j-м администраторов. Если совпадающих рангов нет, то $H_j = 0$, $h_k = 0$ и, следовательно, $T_j = 0$. В этом случае формула (19) совпадает с формулой (18). Коэффициент согласия равен 1, если все ранжировки администраторов одинаковы. Коэффициент согласия равен нулю, если все ранжировки различны, т. е. Совпадения отсутствуют. Коэффициент согласия, вычисляемый по формуле (18) или (19), является оценкой истинного значения коэффициента и, следовательно, представляет собой случайную величину. Для определения значимости оценки коэффициента согласия необходимо знать распределение частот совпадений для различных значений числа администраторов m и количества узлов n. Распределение частот для W при $3 \le m \le 20$ и $3 \le n \le 7$ приведено в [9]. Для больших значений m и n можно использовать известные статистики. При числе узлов $n > 7$ оценка значимости коэффициента согласия может быть произведена по критерию χ^2. Величина $Wm(n-1)$ имеет распределение χ^2 с $\nu = n-1$ степенями свободы.

При наличии связанных рангов χ^2 распределение с $\nu = n-1$ степенями свободы описывается выражением [9]:

$$\chi^2 = \frac{12S}{mn(n+1) - \frac{1}{n-1}\sum_{j-1}^{m} T_j}.$$

Энтропийный коэффициент согласия определяется по формуле [9]: $W = 1 - \frac{H}{H_{max}}$, где Н – энтропия, вычисляемая по формуле:

$$H = -\sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij} \log p_{ij}, \qquad (20)$$

а H_{max} - максимальное значение энтропии. В формуле для энтропии p_{ij} - оценки вероятностей j-го ранга, присваиваемого i-му узлу. Эти оценки вероятностей вычисляются в виде отношения количества администраторов m_{ij}, приписавших узлу O_i ранг j, к общему числу администраторов [8]: $p_{ij} = \frac{m_{ij}}{m}$. Максимальное значение энтропии достигается при равновероятном распределении рангов, т. е. когда $m_{ij} = m/n$. Тогда:

$$p_{ij} = \frac{m}{mn} = \frac{1}{n}. \qquad (21)$$

Подставляя это соотношение в формулу (20), получаем

$$H_{max} = -\frac{1}{n} \log \frac{1}{n} \sum_{i,j=1}^{n} = n \log n. \qquad (22)$$

Коэффициент согласия изменяется от нуля до единицы. При $W_Э = 0$ расположение узлов по рангам равновероятно, поскольку в этом случае $H = H_{max}$. Данный случай может быть обусловлен либо невозможностью ранжировки узлов по сформулированной совокупности показателей, либо полной несогласованностью мнений администраторов. При $W_Э = 1$, что достигается при нулевой энтропии $(H = 0)$, все администраторы дают одинаковую ранжировку. Действительно, в этом случае для каждого фиксированного узла O_i все администраторы присваивают ему один и тот же ранг j, следовательно, $p_{ij} = 1$, а $p_{kj} = 0$ $(k \neq j, k = 1,2,...,n)$. Поэтому и $H = 0$.

Сравнительная оценка дисперсионного и энтропийного коэффициентов согласия показывает, что эти коэффициенты дают примерно одинаковую оценку согласованности системных администраторов при близких ранжировках. Однако если, например, вся группа администраторов разделилась в мнениях на две подгруппы, причем ранжировки в этих подгруппах противоположные (прямая и обратная), то дисперсионный коэффициент согласия будет равен нулю, а энтропийный коэффициент согласия будет равен 0,7. Таким образом, энтропийный коэффициент согласия позволяет зафиксировать факт разделения мнений на две противоположные группы. Объем вычислений для энтропийного коэффициента согласия несколько больше, чем для дисперсионного коэффициента согласия.

Выводы

Используя предлагаемую марематическую модель экспертизы сети на наличие коллизий при заданных исходных данных относительно количественного состава и квалификации администраторов, можно получить достаточно устойчивые оценки эффективности. Достоинством предложенной методики является возможность текущего оценивания загрузки всей сети в процессе ее работы.

Список литературы

1. Кульгин М. В. Компьютерные сети. Практика построения / Кульгин М. В. – [2-е изд.]. – СПб : Питер, 2003. – 464 с.
2. Трайнев В.А. Параметрические модели в экспертных методах оценки при принятии решений. – М.: Прометей, 2003. – 232 с.
3. Игнатьев А. Ю. Применение метода экспертных оценок при разработке технологии принятия управленческих решений // Актуальные вопросы современного управления и статистики. 2001 . Вып. 4. – С. 12-22.
4. Панкова Л.А., Петровский А.М., Шнейдерман М.В. Организация экспертиз и анализ экспертной информации. – М.: Наука, 1984. – 120 с.
5. Моисеев Н.Н. Математические задачи системного анализа. – М.: Наука, 1981. – 487 с.
6. Литвак Б.Г. Экспертная информация. Методы получения и анализа. – М.: Радио и связь, 1982. –184 с.
7. Орлов А.И. Теория принятия решений. – М.: «Март», 2004. – 656 с.
8. Бешелев С.Д., Гурвич Ф.Г. Математико-статистические методы экспертных оценок. – М.: Статистика, 1980. – 263 с.
9. Евланов Л.Г., Кутузов В.А. Экспертные оценки в управлении. – М.: Экономика, 1978. – 133 с.