Шипилов Д.В., Кремер А.А.

ФОРМАЛИЗАЦИЯ МОДЕЛИ УПРАВЛЕНИЯ РЕСУРСАМИ СЕТЕЙ ИНТЕГРИРОВАННОГО ОБСЛУЖИВАНИЯ В ВИДЕ ЗАДАЧИ ДИСКРЕТНОГО ПРОГРАММИРОВАНИЯ

denis@relex.ru

В [1] представлен подход к проблеме управления потоками в сетях интегрированного обслуживания как задаче оптимального распределения вычислительных ресурсов сетевого узла, выделенных для выполнения задачи сетевого управления.

В исходной постановке задачи выделяются два основных структурных компонента разрабатываемой программной системы – *сетевые узлы* и связанные с ними *потоки*.

Рассматриваются совокупности узлов, являющихся частью среды передачи данных с негарантированной (“best-effort”) характеристикой обслуживания, взаимодействующих посредством множества потоков и представимых в виде связных ориентированных графов. Требование связности этого графа подразумевает, что в совокупность попадают лишь узлы, между любой парой которых существует цепь (другие узлы и потоки между ними) их соединяющая. Рассматриваются только симплексные потоки.

На низком уровне рассматриваются подграфы, включающие один центральный узел и смежные с ним узлы (включаются только дуги инцидентные центральному узлу). Именно на этом уровне решается рассматриваемая оптимизационная задача: для всех узлов в системе определено множество классов распределяемых ресурсов

$$R = \{R_1, R_2, ..., R_m\},\ |R| = m,$$

и связанных с ними ограничений. Узел занимается обработкой потоков

$$S = \{S_1, S_2, ..., S_n\},\ |S| = n,$$

перераспределяя выделенные ему ресурсы $\boldsymbol{R}$ в соответствии с требованиями каждого потока.

Поток описывается двумя основными параметрами:

- *входной интенсивностью* l (количество блоков данных генерируемых приложением в единицу времени), в общем случае являющейся функцией времени;
- *величиной задержки потока* на узле назначения $\Delta \boldsymbol{t}$, определяемое в момент создания потока из требования (1), и являющееся константной величиной все время существования потока.

Основное требование, предъявляемое к системе при обработке потоков, связывающее перечисленные выше параметры, выражается равенством (1):

$$l(t) = l'(t + \Delta t) \qquad (1)$$

где $l(\boldsymbol{t})$ – входная интенсивность потока, $l'(\boldsymbol{t})$ - выходная интенсивность, $\Delta \boldsymbol{t}$ – величина задержки. Нарушение этого равенства трактуется как отказ в обслуживании по отношению к обрабатываемому потоку.

На основании (1) вычисляется вектор минимальных требований по каждому из ресурсов:

$$Q^i{}_{\min} = \{Q^i_{\min_1}, Q^i_{\min 2}, \ldots, Q^i_{\min_m}\},\ \left|Q^i{}_{\min}\right| = m,$$

где $\boldsymbol{Q}^{\boldsymbol{i}}_{\boldsymbol{min}}$ – вектор минимальных требований для $\boldsymbol{i}$-го потока. Причем:

$$\sum_{i=1}^{n} Q^i_{\min k} \le R_k, \qquad (2)$$

где $\boldsymbol{R_k}$ – количество $\boldsymbol{k}$-го ресурса, доступное узлу для распределения между потоками. Нарушение требования (2) трактуется как ситуация исчерпания ресурсов и также влечет за собой отказ в обслуживании.

В качестве критерия функционирования разрабатываемой системы принято:

$$\min P_{отк.} = \Phi(P_{пер.}, P_{рес}) \qquad (3)$$

где $\boldsymbol{P}_{отк.}$ – вероятность отказа в обслуживании, $\boldsymbol{P}_{пер.}$ – вероятность нарушения условия (1) по всем потокам, $\boldsymbol{P}_{рес.}$ – вероятность исчерпания ресурсов (2) по всем потокам.

В процессе работы система должна принять решение – сформировать общий вектор распределения ресурсов между потоками:

$$Q = \{Q^1, Q^2, ..., Q^n\},\ |Q| = n, \qquad (4)$$

где

$$Q^i = \{Q_1^i, Q_2^i, ..., Q_m^i\},\ |Q^i| = m, \qquad (5)$$

причем с одной стороны вычисляемое распределение ресурсов между потоками, также как и минимальные требования, должно подчиняться неравенству (2):

$$\sum_{i=1}^{n} Q_k^i \le R_k, \qquad (6)$$

с другой, условию (1), что можно выразить так:

$$Q_k^i \le Q_{\min k}^i,\ i = \overline{1,n} \text{ и } k = \overline{1,m}. \qquad (7)$$

Далее, вернувшись к исходной постановке задачи (об эффективном управлении в сетях) и проанализировав ее специфику, следует отметить, что множество $\boldsymbol{R}$ *неоднородно*, в том смысле, что включает как действительные, так и целочисленные компоненты. Т.е. по сути, данное множество распадается на два:

$$R = R^Z \cup R^R, \qquad (8)$$

$$R^r = \{r_i \mid r_i \in R\},\ |R^r| = u,\ i = \overline{1,u}$$

$$R^z = \{z_j \mid z_j \in Z\},\ |R^z| = v,\ j = \overline{1,v}$$

$$u + v = m,$$

где $\boldsymbol{R^Z}$ – подмножество целых компонентов $\boldsymbol{R}$, $\boldsymbol{R^R}$ – подмножество действительных компонентов $\boldsymbol{R}$, $\boldsymbol{v}$ и $\boldsymbol{u}$ – соответствующие размерности этих множеств.

Получаем, что:

1. исходная задача сводится к необходимости вычисления вектора распределения ресурсов $\boldsymbol{Q}$;
2. при заданных ограничениях (6) и (7);
3. с целевым функционалом (3);
4. и свойством множества $\boldsymbol{R}$ (8).

А это фактически представляет собой формулировку задачи дискретного программирования в её общей форме.

Если рассматривать геометрическую интерпретацию задач этого класса, то суть их сводится к отысканию точки в гиперпространстве размерности равной числу компонентов оптимизируемого вектора. У нас согласно определениям (4) и (5) размерность множества решений $\boldsymbol{D}$ составит:

$$|D| = m \cdot n, \qquad (9)$$

т.е. с появлением каждого нового потока в системе сложность решения будет возрастать на величину $\boldsymbol{m}$ количества классов ресурсов. Такая ситуация фактически сводит на нет эффективность использования стандартных процедур решения задач дискретного программирования. Поэтому целесообразно понизить ее размерность, исключив каждые **(*m-1*)** компонентов введением функции свертки для потоков:

$$j^{\,i} = \Gamma(Q_1^i, Q_2^i, \mathbf{K}, Q_m^i)\,,\ i = \overline{1,n}\,, \qquad (10)$$

Далее необходимо переформулировать (3), (6) и (7) и решать задачу относительно введенных переменных-сверток.

Список использованных источников

1. Кравец О.Я., Шипилов Д.В. Управление потоками в сетях интегрированного обслуживания как задача об эффективном распределении вычислительных ресурсов сетевых узлов // Технологии интернет - на службе обществу. Саратов: СГТУ, 2003.