Resource Allocation in Communication Networks

Javier A. Barria
Intelligent Systems and Networks
Electrical and Electronic Engineering Department

Content

1. Current Research
2. Monitoring of communication system behaviour
3. Support of heterogeneous traffic over the Internet
4. Path searching and routing in dynamic topology networks
5. Load control in Intelligent networks
6. Final remarks
1. Current Research 2007 (1/5)

- Intelligent monitoring of communication systems behaviour.
- Routing and scheduling mechanisms in connectionless environments.
- Resource management for future service and transport networks environments.
- Agent technology applications to network systems
- Interactive multimedia service delivery and consumption mechanisms
- Management of self-organised Networks.

Current Research (2/5)

- Intelligent monitoring of communication systems behaviour

QoS management (quality assurance)
  - Network monitoring and early anomaly detection (e.g. IP & ATM networks).
  - Self-similar traffic identification and control.
  - Network behaviour classification
Current Research (3/5)

• Routing and scheduling mechanisms in connectionless networks

• Resource management for future service and transport networks environments
  – Resource allocation and transport protocols over dynamic topology networks (e.g. LEO constellations and Ad-hoc networks).

Current Research (4/5)

• Agent technology applications to networks
  – Resource allocation in distributed Intelligent Networks.
  – Distributed algorithms to solve large-scaled systems.
  – Multi-domain roaming and dynamic QoS.

• Interactive multimedia service delivery & consumption
  – Emergent business and delivery models.
  – Content management and usage.
  – Intellectual property management and protection.
Current Research (5/5)

- Management of self-organised networks
  - Application layer distributed Proxy Servers location algorithms.
  - Application layer load balancing algorithms.
  - Network reconfiguration algorithms and dynamic routing schemes for Hybrid Networks.
  - Resource allocation with partial /uncertain information.

2. Monitoring of communication systems behaviour

[ Dr. V. Alarcon-Aquino ]

- To investigate the viability & usability of wavelet transforms in network anomaly detection.
- To develop a wavelet-based anomaly detection algorithm.
- To compare the wavelet-based approach with previous proposed anomaly detection algorithms, e.g.,
  - The Generalised Likelihood Ratio (GLR) test Using AR model.
  - Static and Adaptive Thresholding Techniques.
Monitoring (Network Anomaly Detection)

Input

IP Network/Service Performance data:
- MIB variables
- Dial IP Data
- Timing Information

Monitoring the relevant performance variables

Output

- Intelligent Alarms
- Diagnosis Messages
- Control Signals

Information for Network Manager

Analyze Data to Detect Network Anomalies

Sampling and Filtering

Trend Analysis

Anomaly Detection

Example of IP Network and Monitoring Points

Dial-End Users

Public Internet

Workstation

Router

Dial-End Users

Ethernet Switch

Network Anomaly Detector

SNMP

Monitoring Points

Alarm Time to Network Manager

HTTP - HyperText Transfer Protocol
SNMP - Simple Network Management Protocol

HTTP - Hypertext Transfer Protocol

HTTP - Hypertext Transfer Protocol
**Multiresolution Sensor**

Multiresolution Sensor:

\[ \{2^j\}_{j=0}^{J} = \text{Dyadic sequence.} \]

\[ J = \text{The number of scales in wavelet domain.} \]

\[ W_{2^j} f_i = \text{Wavelet Coefficients.} \]

\[ L = \text{Length of Sliding window.} \]

---

**Multi-resolution Data Fusion Systems**

Multi-resolution Data Fusion System:

\[ u_{ij}(t) = \text{Local decisions, } i = 1, \ldots, n \text{ and } j = 1, \ldots, J. \]

\[ b_{ij}(t) = \text{Weighted local variable.} \]

\[ v_{ij}(t) = \text{Decision of weighted variable.} \]

\[ \phi(\delta_j) = \text{Decision fusion rule.} \]
Application to IP Networks

The Monitored Dial IP Network

The Web_Latency variable and alarms obtained by the proposed wavelet-based approach

Output Multi-resolution System

The Log_Time and Data_Time variables, and the alarms obtained by the proposed wavelet-based multiresolution sensor.
Summary

- The performance of the proposed wavelet-based approach was compared with adaptive thresholding techniques and autoregressive models.

- The wavelet-based approach showed performance improvements in terms of:
  - Detection of smooth and abrupt changes.
  - Lower false alarm rate.
  - Proactive identification of anomalous conditions.

- The wavelet-based approach was able to identify changes in the AR parameters (frequency jump) and in the variance (energy jump).

References

3. The support of heterogeneous traffic over the Internet

[Dr. R. M. Salles]

- To investigate resource allocation algorithms to support different types of network applications and QoS requirements.
- To apply End-user application information to control network mechanisms: a) routing & flow control, and b) scheduling.
- To compare a novel End-user utility-based approach with previous algorithms in terms of fairness, optimality and differentiation.

IP Network: DiffServ Architecture, OSPF/MPLS

Problems:

a) How to allocate transmission capacity among different applications?
b) How to schedule packets from different applications?
Utility-based framework

Definition

\[ u = f(Q) : \mathbb{R}^N \rightarrow [0, 1] \]

Utility as a function of QoS vector

\[ Q = \{ q_0, q_1, \ldots, q_N \} \]

QoS vector with \( N \) different QoS parameters

\[ q(A) : \mathbb{R}^M \rightarrow \mathbb{R} \]

QoS parameter \( q \) as a function of \( A \)

\[ A = \{ a_0, a_1, \ldots, a_M \} \]

Resource vector with \( M \) different resources

Examples of utility functions

Utility \( \Leftrightarrow \) Application performance

Optimality and Fairness

• Fairness (why?)
  – Pre-requisite for some applications: real-time tele-stock trading, tele-voting, etc…
  – Deliberately unfair systems are also unstable;
  – Protection against misbehaving users;
  – Avoids intrinsic system discrimination: TCP vs UDP, large flows, long routes, etc…

• Optimality and Fairness
  – Undesirable optimal points: e.g. the solution of (*)
    \[ \max_{x \in \mathcal{X}} \{ W = \sum_{i=1}^{N} u_i(x) \} \]
    – \( W \) is likely to exclude some users (very low or even zero utility).
    – Need to balance the two goals: utility fairness criteria
3.a. Single-path utility routing (SPUR) problem

Objective: find $Y_0$ and $\vec{r}$

Subject to: $\sum_{j=1}^{K} r_j a_{ij} + r_0 y_i \leq C(i), 1 \leq i \leq L$

Where: $Y_0$: path selected to the incoming flow
$\vec{r}$: vector of flow allocation ($r_0, r_1, \ldots, r_K$)
$C(i)$: capacity of link $i$
$\{a_{ij}\}$: routing matrix
$u_j$: flow’s $j$ utility function ($j = 0, 1, \ldots, K$)

Assumptions: - single-path routing;
- network state: $K$ flows already in the system;
- utility functions of general concavity.

Simulation Results

- SPUR provided the highest values for the minimum utility in the system;
- SPUR was able to differentiate flows according to the utility functions.
3.b. Packet switching: Utility based scheduling (UBS)

**Goal:** Derive a service discipline for adaptive applications which is fair and can be controlled by application utility functions

**Application:** DiffServ networks (e.g. substituting the PQ discipline)

**Characteristics:** utility fair, robust, configurable and sensitive to loads variations

\[ D_j = \frac{1}{2} (u_j(t) - s_j(t) + D_j^{\text{last}}(t)) \]

**Utility Functions**

\[ u_j = \delta_j (D_j^{\text{max}} - D_j) \]

**Measurement** \[ \rightarrow \]

**Decision** \( \max \min_j \{u_j(t)\} \)

**Input traffic**

---

**UBS example:** Implementation of Proportional DiffServ

- Relative performance guarantee (the problem)
  - Class performance “spacing” defined by weights \( \delta_j \)
  - Weighted delays \( W_j = \delta_j D_j \)
  - Current proposal: waiting time priorities (WTP)
  - Objective: equalise weighted delays

- The proposal
  - Apply utility-based schedule with linear functions

\[ u_j = \left( \frac{\delta_j}{\delta_j^{\text{max}}} \right) (D_j^{\text{max}} - D_j) \]

- Simulation results
  - Three classes of services with corresponding weights \( \delta_1 = 1; \ \delta_2 = 2; \ \delta_3 = 4 \)
Packet Scheduling Results

Balanced Scenario:
same loads to all classes

Unbalanced Scenario:
different class loads

Goal set for the service (3 Classes)
\[ \delta_1 = 1, \delta_2 = 2 \text{ and } \delta_3 = 4 \]
relative ratios: \[ K_{12} = \delta_2 / \delta_1 \text{ and } K_{23} = \delta_3 / \delta_2 \]

Summary

- The concept of utility function has been used as a mean of allocating resources.
- Algorithms to solve the combined routing and flow control problem have been proposed.
- The framework was extended to the packet level where scheduling disciplines were proposed and used to emulate the *Proportional DiffServ* providing better results than current available solutions.
- There is evidence that utility based algorithms lead to a more desirable network state in terms of fairness, optimality and differentiation.
References


4. Path Searching and Routing in Dynamic Topology Networks

• To provide QoS constrained path search and routing in the presence of imprecise network information.
• To provide call admission control and routing in the presence of imprecise network information.
• To provide efficient trade-off between the path discovery success ratio and search message overhead.

LEO Satellite System example

A rendering of the originally-announced Teledesic design with 840 active satellites in 21 planes of 40. The Boeing redesign uses 288, in 12 planes of 24.

Reference: http://www.ee.surrey.ac.uk/Personal/L.Wood/constellations/teledesic.html
Problem Background (LEO)

• In connection-oriented networks with dynamic topology, paths with the most stable links are selected to minimise the cost of readapting to topology changes.
• A call is admitted to the network provided that there are sufficient resources in the links that constitute the selected path.
• Admitting new calls to the network generates revenue for the network operator but increases the forced termination probability of handover calls.

Reduced LEO Satellite System Model

\[ \lambda_{j,N} = \text{New call arrival rate of class } - j \]

\[ \lambda_{j,HO} = \text{Handover call arrival rate of class } - j \]
**MDP routing in ISLs**

\[ g_s = \text{reward gain of route } k_i \]
\[ = \sum_{x \in k_i} q(x, \pi) \]

**ISLs = Intersatellite links**

**Problem decomposition**

- Periodic update of network topology is decomposed into sequential snapshots.
- In each snapshot, a stable topology is considered. The topology contains links that satisfy:

\[
\text{Prob}\{\text{path holding time} \leq \text{time until link is not part of the topology}\} > \text{Threshold}
\]

- Network model is decomposed into a set of independent link admission control tasks.
- Each call admission control task is formulated as a semi-Markov decision process (SMDP).
Results: Blocking probabilities

- MHA: Finds a path with a minimum number of hops.
- MA: 1) Finds set of all min-hop paths 2) Selects path with minimum congested link.
- MCA: Finds minimum cost path. Cost of link = 1/vacancy, where vacancy = number of free channels in the link.
- SMDP: Admits and allocate calls to routes according to SMDP policy.

Results: Revenue

- MHA
- MA
- MCA
- SMDP
Summary (LEO Satellites)

• A Markov decision process (MDP) framework has been applied to call admission decisions and route allocation in communication networks with sequential topologies.
• SMDP approach shows higher net revenue and lower blocking probability when compared to other end-to-end routing schemes in dynamic networks with symmetric traffic demands.
• Asymmetric traffic demand case has been studied and results are similar.

Reinforcement Learning

\[ x_t = \text{actual state} \]
\[ y_t = \text{observed state} \]
\[ a_t = \text{action taken (number of tickets transmitted at source node)} \]
\[ r(x_t, a_t) = \text{reward returned from environment} \]
Summary (MANETs)

- Imprecise network information affects the ability to perform path resource reservation.
- The problem of overhead control in a route discovery process for MANETs with delay-constrained services is formulated as a control problem.
- Reinforcement learning (RL) is used to approximate an optimal memoryless deterministic policy.
- RL approach shows higher average reward when compared to the route discovery schemes that use e.g. ticket based proving and flooding, under different degrees of imprecision and mobility.

References

5. Load control in Intelligent Networks

• To investigate the applicability of multi-agent systems (MAS) to solve distributed (e.g. multi-domain) Intelligent network (IN) resource allocation (RA) problems.
• To implement novel MAS distributed resource allocation schemes in connection oriented and connectionless environments.
• To compare developed MAS approaches with previously proposed algorithms.

Problem Background

• The system:
  – To control distributed SSP contending for distributed SCP process utilization.
• Aim:
  – To control SCP process utilization (capacity) below a pre-defined level.
  – To maximize the expected network revenue.
• Assumptions:
  – Input rate selection throttling mechanism is used where the maximum of calls of type $j$ that SSP $k$ is allowed to send to SCP $i$.
  – Quasi-stationary input traffic.

SSP = Service switching point
SCP = Service switching point
IN Load Control using a MAS

SSP: Service Switching Point
SCP: Service Control Point
IP: Intelligent Peripheral
SDP: Service Data Point
SS7: Switching System No.7

Distributed Resource Allocation

Agents
- Individuals that will be competing for resources

Auctioneers
- These control individual markets for goods (e.g., bandwidth or service logic)
Auction Process (WALRAS)

Agents
Excess Demand Curves (as a function of price)

Auctioneer
Aggregate Demand Curve

Good Price

Record P*
P*=Market Clearing Price

Agent types used in the Model

Consumer agents (maximize their individual utility functions)
- Utilities must be convex w.r.t. allocations.
- Consumer constrained by a budget.

Producer agents (maximize their profits)
- Production function determines the transfer function between input and output commodities.
- Production function must be convex w.r.t. allocation.
Computational Economy Model (IN load control)

SCP Process Utilisation

Average SCP Processor Utilisation

Agent Controlled
ACG Control
Load Profile

Time/Seconds
Average Call set-up delay

Cumulative Network Revenue
Summary

• The IN load control problem was presented as a revenue maximising problem.
• A distributed solution using a “market-based” mechanism was proposed.
• The solution is compared to a standard Automatic Call gapping (ACG) mechanism.
• The market-solution was found to be more stable, lower call set-up delay and greater revenue.

References

7. Final remarks

2. Monitoring of communication system behaviour.
3. Support of heterogeneous traffic over the Internet.
5. Load control in Intelligent networks.
6. Final remarks.

8. More Information

Dr. J. A. Barria
Intelligent Systems and Networks
Department of Electrical and Electronic Engineering
Imperial College London

Email: j.barria@imperial.ac.uk
Phone: +44 (0)20 7594 - 6275
Fax: +44 (0)20 7594 - 6302