Introduction à

la Théorie de Base du Télétrafic

Mr. H. Leijon, UIT

UNION INTERNATIONALE DES TELECOMMUNICATIONS
INTERNATIONAL TELECOMMUNICATION UNION
UNION INTERNACIONAL DE TELECOMUNICACIONES

Théorie de base du télétrafic (T)

INTRODUCTION

Contenus

1. Historique

Le développement de la théorie du trafic téléphonique commença au début de ce siècle. La réalisation pionnier dans ce domaine était celle du Danois A. K. Erlang, dont ses oeuvres étaient publiées entre 1909 et 1928. Parmi ceux qui ont bâti leurs idées sur celle d'Erlang, on doit citer le suédois, Conny Palm, dont les documents, durant la période 1936 - 1946 (1957) contribuaient à donner à la théorie du trafic sa rigueur présente. Beaucoup d'autres personnes, de nationalités différentes ont aussi contribué au développement de la présente théorie.

La théorie de trafic qui peut être appliquée aux cas pratiques est basée sur la supposition d'équilibre statistique, ce qui implique qu'elle ne peut traiter que des cas avec des conditions stationnaires .

Pour les conditions non stationnaires, aucune méthode de calcul n'a encore été inventée. L'historique théorique pour traiter des cas pareils était, cependant, présenté dans la thèse de doctorat de Palm en 1943 où il faisait une étude de variations dans l'intensité d'appels . Mais maintenant, il est déjà peu possible de traiter les cas de trafic non stationnaires à l'aide de simulations informatiques. Cependant, les théories considérées ici vont être liées aux conditions stationnaires.

Les théories existantes utilisent différentes combinaisons de suppositions, et les dérivations de 1909 à ce jour se basent sur différents niveaux de savoir et utilisent partiellement différentes terminologies. Un examen direct des dérivations pour des cas différents, présentés originalement, n'offrira pas une reconnaissance de la capacité de la théorie décrivant les différents cas pratiques. Il est donc préférable de présenter la théorie du trafic dans une forme plus générale, à partir de laquelle des différents cas spéciaux peuvent être alors dérivés.

Dans la première section, donc, les caractéristiques qui sont communes aux différentes méthodes de groupage, à la fois pour le système avec perte, et pour le système avec attente, vont être tirées. Il va, donc être préférable de traiter séparément le groupe à accessibilité totale dans un système avec perte et celui dans un système avec attente. Les théories de multiplexage gradué et des systèmes de liaisons sont traitées dans les chapitres ultérieurs, où il va être facilement compris comment la théorie générale est appliquée à ces cas.

2. Envergure et nature de la théorie de trafic téléphonique

Cette théorie peut être considérée comme une théorie de file d'attente appliquée aux systèmes de télécommunications. Le concept général de la théorie des files d'attente est concernée par l'analyse mathématique des systèmes sujets aux demandes, dont les occurrences et les longueurs ne peuvent, en général, être spécifiées que d'une manière relative à la probabilité. Par exemple, vous considérez un système téléphonique, dont la fonction est d'offrir des itinéraires de communications entre des paires d'appareils téléphoniques (clients). L'offre d'un itinéraire de communication permanent entre chaque paire d'appareils téléphoniques serai astronomiquement chère et peut être impossible. En réponse à ce problème, les installations nécessaires pour établir et maintenir un itinéraire de parole entre une paire d'appareils téléphoniques, sont offerte dans un étang commun, pour être utilisé par un appel quand il est demandé, et puis retourné à l'étang après. Cela introduit la possibilité que le système sera incapable d'achever un appel demandé à cause d'un manque d'équipements disponibles en ce moment. Aussi , la question qui se pose immédiatement est : Quelle quantité d'équipements faut il fournir afin que la proportion des appels mise en attente sera d'un niveau acceptable ? Des questions similaires se posent sur beaucoup de systèmes bien différents du système téléphonique : Combien de lits doit un hôpital fournir ? Combien de terminaux de données peuvent être servis en temps-partagé ?. Ces questions partagent une caractéristique commune : dans chaque cas, le nombre de fois les services seront demandés et le temps que prennent ces demandes, les installations ne peuvent être prédites que dans un sens statistique. Malgré le fait que ces systèmes sont souvent très compliqués, il est parfois possible d'extraire de la description du système un modèle mathématique dont l'analyse produit des informations utiles

Considérez le modèle suivant. Les client demandent l'usage d'un type d'équipement particulier (serveur). Si un serveur est disponible, le client arrivant va le saisir et le maintenir pour un certain temps, après quoi le serveur va être disponible pour d'autres clients entrants ou en attente. Si le client entrant ne trouve pas de serveur disponible, il agira d'une manière spécifique telle qu'attendre ou s'en aller. Désormais, le modèle est défini selon 3 caractéristiques : Le processus d'entrée, le mécanisme du service, et la discipline de queue .

Le processus d'entrée décrit la séquence des demandes de service. Souvent, par exemple, le processus d'entrée est spécifiée selon la distribution des longueurs de temps entre les instants consécutifs d'arrivée de clients. Le mécanisme du service est la catégorie qui inclue les caractéristiques telles que le nombre de serveurs et la longueur de temps que mettent les clients en les retenant. Par exemple, les clients pourraient être servis par un serveur unique, chaque client maintient le serveur pour une même longueur de temps. La discipline de queue spécifie la disposition de clients bloqués (ceux qui trouvent tous les serveurs occupés). Par exemple, il peut être supposé que les clients bloqués quittent le système immédiatement ou qu'ils attendent dans une queue, et ils sont servis à partir de la queue dans l'ordre de leur arrivée.

Maintenant considérez le modèle suivant. Deux villes sont interconnectées par un groupe de n circuits téléphoniques interurbains (serveurs). Supposons que tous ceux qui arrivent et trouvent les circuits occupés, n'attendent pas mais quittent le système immédiatement (techniquement, il n'y a pas de queue qui se produit). Quelle proportion des appels d'arrivée(clients) seront incapables de trouver un circuit libre (et ainsi sont rejetés)?

Nous espérons dériver la formule qui va prédire la proportion d'appels perdus, comme une fonction de la demande; c'est-à-dire nous espérons donner une formule qui permet d'estimer le nombre de circuits nécessaire pour répondre à un critère de service pré-spécifié à partir d'une estimation de la charge du trafic téléphonique généré entre les deux villes. L'énorme valeur pratique de tout modèle qui mène à une formule pareille est évidente.

Maintenant, nous allons donner une dérivation heuristique de la formule voulue, utilisant une conception de grande importance en science et en ingénierie, celle de la conservation de flux. La dérivation suivante est heuristique, alors personne n'est supposé la comprendre complètement, la "dérivation" est un argument de plausibilité et est correcte dans certaines circonstances.

Quand le nombre de clients dans le système est j, le système est dit être en état E_j ($j = 0, 1, ...n$). Soit P_j la proportion de temps durant laquelle j circuits sont occupés, P_j est la proportion de temps que le système reste en état E_j. Désignons par λ le taux d'arrivée des appels; λ est le nombre moyen de demandes de service par unité de temps. Considérez d'abord le cas $j < n$. Puisque les appels arrivent avec le taux global λ, et puisque la proportion de temps que le système reste en état E_j est P_j, le taux auquel la transition $E_j \to E_{j+1}$ se produit (le nombre moyen de telles transitions par unité de temps) est alors $\lambda \cdot P_j$. Maintenant, on considère le cas quand $j = n$. Puisque l'état E_{n+1} représente un état physiquement impossible (il y a seulement n circuits) la transition $E_n \to E_{n+1}$ se produit avec un taux de transition Zéro. Ainsi le taux pour le quel la transition ascendante $E_j \to E_{j+1}$ se produit est $\lambda \cdot P_j$ quand $j = 0,1 .... n - 1$ et Zéro quand $j = n$.

Considérons maintenant les transition descendantes.

$$E_{j+1} \to E_j \qquad (j = 0, 1, ..., n - 1)$$

Supposez que le temps moyen de prises (la longueur moyenne de temps qu'un appel maintient un circuit) est τ, alors si un circuit individuel est occupé, le nombre moyenne d'appels arrivant durant un temps écoulé τ est 1 ; le taux d'arrivée pour un appel unique est alors $1/\tau$. Semblablement, si deux appels progressent simultanément et la durée moyenne d'un appel est τ, le nombre moyen d'appel arrivant durant un temps écoulé τ est 2; le taux d'arrivée pour deux appels simultanés est alors $2/\tau$. Par ce raisonnement, donc, le taux d'arrivée pour $j + 1$ appels simultanés est $(j + 1)/\tau$. Puisque le système est à l'état E_{j+1} une proportion de temps P_{j+1}, nous concluons que la transition descendante $E_{j+1} \to E_j$ se produit au taux de $\frac{j+1}{\tau} \cdot P_{j+1}$ transitions par unité de temps ($j = 0, 1, ... n - 1$).

Maintenant nous appliquons le principe de conservation de flux. Si le système est en état d'équilibre statistique, c'est-à-dire si la proportion relative de temps que le système reste en chaque état doit être une quantité stable, alors la transition ascendante $E_j \to E_{j+1}$ doit se produire au même taux que la transition descendante $E_{j+1} \to E_j$. Ainsi nous avons ce qu'on appelle les équations d'équilibre statique.

$$\lambda \cdot P_j = \frac{j+1}{\tau} \cdot P_{j+1} \qquad (j = 0, 1, j......., n - 1) \qquad \text{(TIN 2.1)}$$

Ces équations peuvent être resolues par récurrence ; le résultat qui exprime chaque P_j selon la valeur P_o est: :

$$P_j = \frac{(\lambda \cdot \tau)^j}{j!} \cdot P_o \qquad (j = 0, 1, ..., n) \qquad \text{(TIN 2.2)}$$

Puisque les nombres $\{P_j\}$ sont des proportions, leur somme doit être égale à l'unité :

$$P_o + P_1 + \text{K} + P_n = 1 \qquad \text{(TIN2.3)}$$

Utilisation l'équation de normalisation (TIN 2.3) ensemble avec l'équation (TIN 2.2), nous pouvons déterminer P_o :

$$P_o = \left(\sum_{k=0}^{n} \frac{(\lambda \cdot \tau)^k}{k\,!} \right)^{-1} \qquad \text{(TIN 2.4)}$$

Ainsi nous obtenons pour la proportion P_j de temps, que j circuits sont occupés, la formule :

$$P_j = \frac{(\lambda \cdot \tau)^j \,/\, j\,!}{\sum_{k=0}^{n} (\lambda \cdot \tau)^k \,/\, k\,!} \qquad (\, j = 0,1,....n \,) \qquad \text{(TIN 2.5)}$$

Une observation importante à faire à partir de la formule (TIN 2.5) est que les proportions $\{P_j\}$ dépendent du taux d'arrivée λ et le temps moyen de maintien τ seulement à travers le produit $\lambda \cdot \tau$. Ce produit est une mesure de la demande faite sur le système; il est souvent appelé la charge offerte et donné par le symbole $A = \lambda \cdot \tau$. Les valeurs numériques de A sont exprimées en unités appelées Erlangs (erl), d'après le mathématicien danois A. K. Erlang, le premier à publier la formule (TIN 2.5) en 1917 . Quand $j = n$ dans la formule (TIN 2.5) , le côté de droite de l'équation devient la formule d'Erlang avec perte, bien conne, indiquée en Europe par $E_{1n}(A)$:

$$E_{1n}(\,A\,) = \frac{A^n \,/\, n!}{\sum_{k=0}^{n} A^k \,/\, k\,!} \qquad \text{(TIN 2.6)}$$

Nous allons dériver ces résultats plus attentivement plus tard. Ce qu'il faut signaler ici, est que quelques résultats mathématiques, potentiellement utiles ont été dérivés à l'aide seulement de raisonnement heuristique. La question à la quelle il faut répondre maintenant est : dans quelles conditions sont ces résultats valides?

Plus précisément, quelles suppositions à propos du processus d'entrée et du mécanisme du service sont demandées pour la validité des formules (TIN 2.5) et (TIN 2.6)? Peut-elle être justifiée, l'affirmation que le taux de transition descendante est proportionnel à la réciproque du temps moyen de maintien? Quelle est la relation entre la proportion P_j de temps que j appels sont en progression et la proportion Π_j, disons, celles des appels arrivants qui trouvent j autres appels en progression? A quel point est-elle applicable l'analyse de la conservation du flux? Comment manipuler les opérations pour les quelles ce type d'analyse est inapplicable?

Les questions de cette natures parfois nécessitent des arguments mathématiques très sophistiqués. Pourtant nous allons garder une attitude moyenne quand à l'usage des mathématiques avancées. La matière doit être accessible pour un élève qui comprend la théorie de probabilité appliquée et les domaines mathématiques associés.

3. Modelage mathématique

A partir de ce qu'on a vu, nous comprenons que la théorie de trafic consiste en modelage mathématique d'un système de télécommunications (ou en partie) et son comportement quand il fait des demandes ou il les subit. Une telle théorie, toute, est un exercice de modelages : nous ne pouvons pas installer un système vraiment identique, et alors, plutôt, avec des entrées bien définies et nous analysons cela. Pour cela, la validité et l'utilité de la théorie que nous développerons se base entièrement sur la réponse à cette question : comment le modèle est satisfaisant ? Si nous avons un peu de confiance en cela, alors, comment améliorer nos mathématiques de manière à avoir un peu de confiance dans les résultats finaux de la théorie.

Afin de monter note modèle, nous devons attentivement prendre en considération un tas de points:

1) Quelle est exactement la partie d'un système qui nous intéresse ? Pouvons nous séparer la section nous concernant et la regarder en isolation ? ou bien devons nous traiter tout le système à la fois ?

2) Quel est le comportement technique précis de la section en question, en ce qui concerne les temps d'opération, les limitations d’accès, les temps morts, la réponse détaillée à une demande, etc ?

3) Comment se comporte le flux de demandes d'entrées dans le système.

4) Quelle information voulons nous de notre modèle, et à quel point doit-elle être exacte ?

Les points sont tous mutuellement dépendants. Ils est peut être plus facile de commencer par le flux de demandes d'arrivée et faire une recherche sur sa structure : ce qui va impliquer quelques suppositions que les sources de demande, c.à.d. : abonnés - réagissent au comportement habituel du système, de leur manière habituelle. Typiquement, ce flux de demande est très largement gouverné par le hasard, de telle façon à ce que les méthodes des opérations Stochastiques seront appropriées. Il nous faut donc savoir :

a) quelle est (en langage probabilistique) l'opération d'arrivée de demandes? Et

b) quelle est la distribution de travail qu'elles apportent?

La description du processus d'arrivée peut demander plus ou moins de détails. Dans un système à faible congestion, offrant du trafic "frais" par une multitude d'abonnés indépendants, nous pouvons supposer avec beaucoup d'exactitude que (à n'importe quel taux au delà des périodes de temps nom trop longues) les arrivées sont du *hasard pur* c.à.d. : qu'elles forment un processus Poissonnien. Si nous examinons un système avec trafic débordant offert, une description plus compliquée sera nécessaire, et plus compliquée encore si nous nous attendons à une proportion significative de tentatives répétées.

Maintenant considérez le deuxième point - la distribution du travail apporté par une demande unique. En fait, cela peut avoir une variation qualitative considérable : si nous sommes concernés par l'occupation du circuit, le "travail" consiste en un maintien de temps individuel continue ; alors que si nous somme en train de modeler le contrôle commun d'un système de processus compliqué, il peut être une séquence de tâches disjointes de types et de durées largement différentes. Deux cas particulièrement communs et importants sont, cependant, où le temps de maintien a une distribution exponentielle négative, et où elle est efficacement deterministique - c.à.d. une constante .

Nous pouvons maintenant virer notre attention vers le point (4) ci-dessus : quelle information voulons nous de notre modèle? Cela exige, naturellement dans tous les cas une certaine connaissance du comportement du système (est-ce que les appels bloqués sont perdus, ou se mettent-ils en queue? par example) , et une compréhension de nature du processus d'entrée. Puisque cette entrée est stochastique, la sorties de notre modèle sera probabilistique, et peut consister en probabilités de pertes, délai moyen, occupation de processeur ou de quantités similaires. Il peut même être que c'est seulement le premier stade de la construction d'un modèle plus compliqué , le cas où il nous sera nécessaire de savoir les détails complets de la distribution des retards, des appels débordants ou de quelques autres quantités qui affectent le reste du système. A ce stade, le point (1) est considéré.

Nous sommes maintenant prêts à nous occuper des détails du système d’ingénierie, point (2) ci-dessus, et c'est à ce stade que notre modèle mathématique prend sa forme, et il devient clair si nous avons de l'espoir d'un traitement analytique. En fin, le processus total est répété jusqu' à ce que nous ayons confiance sur sa consistance et que le comportement du système est en effet compatible avec les suppositions concernant le flux de demandes entrantes et vice-versa.

Nous supposons donc qu'un modèle a été ou peut être monté. Il va nécessairement être approximatif d'une manière ou d'une autre, et nous devons estimer combien l'effet de cette approximation qui va avoir sur les résultats du modèle. Si la réponse est qu'il y a trop d'effet, nous devons reformer le modèle. La demande ultime est toujours d'un nombre ou un ensemble de nombres, c.à.d. d'un calcul numérique : tant de compromis peuvent souvent être faits entre les modèles approximatifs et les calcul approximatifs.

En effet, il ne peut pas être possible de déterminer les données d'entrée du comportement d'un abonné en tant de détails qu'il est nécessaire, car les quantités utiles sont non connues ou même non mesurables. Dans de telles circonstances, deux approches sont raisonnables : décider ce que doit faire l'homme Raisonnable et décider que c'est avec cela que le système doit être dimensionné (avec, bien sur, des garantis appropriées au système lui même !) ; ou analyser plusieurs modèles, qui diffèrent seulement dans le comportement de l'abonné, et qui présente une rangée de résultats pour prendre des décisions finales, dans d'autres côtés.

Heureusement, la plupart des quantités d’intérêt comme entrée à partir des modèles mathématiques sont robustes d'une manière remarquable quant aux variations dans le processus d'entrée (à condition bien -entendu, que certains paramètres critiques, comme le trafic total sont gardés constants) , et aussi, même des modèles assez simples donnent des résultat remarquablement exacts et utiles. Cependant, ne jamais oublier que quelque soit l'élaboration de la solution mathématique,

UN MAUVAIS MODELE SIGNIFIE DES RESULTATS NON REALISABLES