Learning to Estimate Traffic Matrix for M2M Management

Rossi Kamal, Mi Jung Choi and Choong Seon Hong
Department of Computer Engineering, Kyung Hee University, Korea
(rossi and cshong)@khu.ac.kr and mjchoi@kangwon.ac.kr

In this paper, first we have formulated M2M service provider's business goal as parameter estimation problem. Then, we have derived Bayesian lower bound based on Van Tree Inequality to estimate traffic matrix for M2M management. At last, we have performed numerical analysis to show how devised mechanism outperforms non-Bayesian approach in terms of Mean Square Error (MSE).

1. Introduction

Machine to Machine/Machine-Type Communication stands for intelligent interaction among wirelessly connected Smart-Devices without or with the least possible human-involvement. The recent penetration of smartphones boosts the projection how billions of smart-devices is going to play a key role in ongoing decade[1].

M2M/MTC has brought an abrupt business paradigm shift in internet community as well. The increasing demand of personalized services, which reflect the context of user or his surrounding environment, has accelerated the appearance of numerous service providers. The complex management task of inferring high-level application/network context from low-level traffic/flow is, merely shared by network operators and service providers, rather than done by only operators[1].

In this context, Big-Data or the monitoring information of M2M devices, networks and services is in center of the focus for operators or service-providers. Operators have to learn from Big-Data to control scalable flow in datacenters, which provide content or perform processing/computational task of M2M Services. Service providers on the other hand, have to infer service-usage from user's context, for example, his emotion and/or environmental information and/or service-usage frequency, etc.

This paper was supported by NIA(National Information Society Agency). Dr. CS Hong is the corresponding author)

Traffic matrix is a famous tool utilized by operators or service-providers to estimate network context from traffic flow. We have estimated traffic matrix that is capable to infer service usage from user emotion and environmental information. However, the approach is linear filtering with deterministic approach[2]-[4], which in might not be feasible in applicable field. Real world applications are of more stochastic nature, where the context of user or user-group is variable with respect to time. Therefore, we have formulated Bayesian Cramer-Rao bound to estimate how the service-usage predication can be varied in real time stochastic case.

This paper is organized as follows: Section II introduces traffic matrix for M2M Management. Section III formulates the business goal as parameter estimation problem and Bayesian lower bound for estimation, and Section V presents numerical analysis for the scalability of devised mechanism.

2. Traffic Matrices for M2M Management

In this section, we have introduced traffic matrices for M2M management. The traffic matrices that can be used by service providers to infer-service usage from network traffic.

M2M service providers want to monitor service-usage (for example: YouTube-usage) based on choices of interest, for example, user emotion: happiness, sadness, boredom, relaxation; user's location; weather of that location.

Preference matrix($\mu^{COI}_{Service}$) is used to calculate most common user-group for any specific service, given choices of interest. It might be applicable to most common usage for any user for given choices of interest.

$$\mu_{Service}^{COI} = \frac{1}{N}\sum_{N} x_{Service}^{COI}$$

Variance-Covariance matrix($\Sigma_{Service}^{COI}$) is used to calculate the variance of any user-group from common user-group of a specific service, for choices of interest.

For example, lets us consider YouTube service-usage for Happiness and Relaxation. Variance-Covariance matrix is used to estimate joint user-groups who prefers YouTube when they are both happy and relaxed. We can estimate, for this specific user-group, how isthe chance of happy or relaxed or how happiness and relaxation are correlated.

$$\Sigma_{Service}^{COI} = \frac{1}{N}\sum_{N} (x_{Service}^{COI} - \mu_{Service}^{COI})^{T} (x_{Service}^{COI} - \mu_{Service}^{COI})$$

$$N(x|\mu_{Service}^{COI}, \Sigma_{Service}^{COI}) = \frac{(x_{Service}^{COI} - \mu_{Service}^{COI})^{T} (\Sigma_{Service}^{COI})^{-1} (x_{Service}^{COI} - \mu_{Service}^{COI})}{(2\pi)^{n/2}\sqrt{\Sigma_{Service}^{COI}}}$$

Therefore, estimating objective, according to multivariate Gaussian distribution, is as follows

Given preference and variance-covariance matrices for any service, thereby, if we are given any user-group, we calculate their chance of being common or if not common how they differ from common.

3. Bayesian Lower Bound to Estimate Traffic Matrix for M2M Management

In this section, first we formulate M2M service provider's business goal as a parameter estimation problem. Then, we present the Bayesian lower bound to estimate traffic matrices with the least possible MSE.

Service providers want to recommend service based on their business goals(x). However, they can observe popularity criteria (θ), for example, service-usage frequency, real-time feedback, service-usage time to learn how their recommendation meet the business goals. Therefore, they want to maximize the likelihood when they estimate popularity criteria to achieve business goal. Thus, the problem is formulated as an inverse problem or maximum likelihood estimation problem.

Cramer-Rao bound is frequently used to estimate the lower bound of an unbiased estimator. An estimator which satisfies the Cramer-Rao lower bound, is called an efficient estimator. However, it is impractical to consider the existence of an unbiased estimator in real world applications, which are all biased in the true sense.

Therefore, we have applied Bayesian Cramer-Rao bound to resolve this by putting a prior assumption about the parameter, which gives more accurate results in posterior distribution. Now, we formulate the Bayesian Cramer-Rao bound for traffic matrix estimation, based on Van-Tree Inequality[5].

Van Tree Equality Theorem: Let θ be a random variable and x be the observation vector/business goal.

The mean square of any estimate $\hat{\theta}(X)$ satisfies the inequality

$$E\{[\hat{\theta}(X) - \theta]^2\} \geq \left(E\left\{\left[\frac{\partial \ln p_{x,\theta}(X,\theta)}{\partial\theta}\right]^2\right\}\right)^{-1}$$

$$= \{-E[\frac{\partial^2 \ln p_{x,\theta}(X,\theta)}{\partial\theta^2}]\}^{-1}..(1)$$

Derivation of Bayesian Lower Bounds: The probability is a joint density and the expectation is over both x and θ.

Let us assume

1. $\frac{\partial p_{x,\theta}(X,\theta)}{\partial\theta}$ is absolutely integrable with respect to X and θ.

2. $\frac{\partial^2 p_{x,\theta}(X,\theta)}{\partial^2\theta}$ is absolutely integrable with respect to X and θ.

3. The conditional expectation of the error, given random parameter θ is

$$B(X) = \int_{-\infty}^{\infty} [\hat{\theta}(X) - \theta]\, p_{x,\theta}(X|\theta) dX..(2)$$

Let us also assume that

$$\lim_{\theta\to\infty} B(\theta) p_{\theta}(\theta) = 0\cdots(3)$$

$$\lim_{\theta\to-\infty} B(\theta) p_{\theta}(\theta) = 0..(4)$$

Let us multiply both sides of (2) by $p_\theta(\theta)$ and then differentiate with respect to θ.

$$\frac{d}{d\theta}[p_\theta(\theta)B(\theta)] = -\int_{-\infty}^{\infty} p_{x,\theta}(X,\theta)\,dX + \int_{-\infty}^{\infty} \frac{\partial p_{x,\theta}(X,\theta)}{\partial\theta}[\hat{\theta}(X) - \theta]\,dX$$

..(5)

Now by integrating with respect to θ,

$$p_\theta(\theta)B(\theta) = -1 + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\partial p_{x,\theta}(X,\theta)}{\partial\theta}[\hat{\theta}(X) - \theta]\,d\theta\,dX \text{..(6)}$$

Assumption (3) yields the left hand side of this equation is zero.

Therefore, we get

$$E\{[\hat{\theta}(X) - \theta]^2\} \geq \left(E\left\{\left[\frac{\partial \ln p_{x,\theta}(X,\theta)}{\partial\theta}\right]^2\right\}\right)^{-1} \cdots(7)$$

We can say, equivalently,

$$E\{[\hat{\theta}(X) - \theta]^2\} \geq \{-E\left[\frac{\partial \ln p_{x|\theta}(X|\theta)}{\partial\theta^2}\right] - E\left[\frac{\partial \ln p_\theta(\theta)}{\partial\theta^2}\right]\}^{-1} \text{..(8)}$$

With equality if and only if

$$\frac{\partial \ln p_{x,\theta}(X,\theta)}{\partial\theta} = k[\hat{\theta}(X) - \theta], for\ all\ X,\theta \cdots(9)$$

Differentiating again yields an equivalent condition

$$\frac{\partial^2 \ln p_{x,\theta}(X,\theta)}{\partial\theta^2} = -k \cdots(10)$$

Rewriting in terms of posterior density yields

$$\frac{\partial^2 \ln p_{\theta|x}(\theta|X)}{\partial\theta^2} = -k \cdots(11)$$

Integrating twice and putting the result in the exponent yields

$$p_{\theta|x}(\theta|X) = \exp(-k\theta^2 + C_1\theta + C_2),\ for\ all\ X,\theta \text{..(12)}$$

This yields that a posterior probability density of θ must be Gaussian for all X for the existence of an efficient estimator.

MAP estimate will be efficient, if (10) is satisfied.

Therefore, $\hat{\theta}_{mse}(X) = \hat{\theta}_{map}(X)$ when an efficient estimator exists.

When an efficient estimator does not exist, it is not possible to know how closely the mean square error approaches the lower bound, by either using $\hat{\theta}_{mse}(X)$ or $\hat{\theta}_{map}(X)$.

4. Numerical Analysis

We have performed analysis on Matlab to check the scalability of devised linear filtering with stochastic environment. We have varied the number of popularity criteria and observed that minimum average mean square error is achieved with Bayesian Cramer-Rao bound, compared to Original Bayesian Creamer-Rao Bound and mean-square-error.

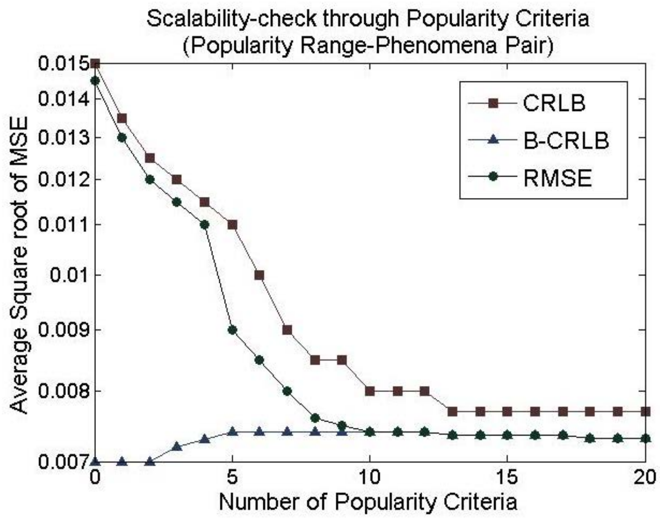

Fig.1: Numerical Analysis Result

5. Conclusion

In this paper, we have used linear filtering with stochastic environment, so that M2M service provider can infer service usage from traffic matrix and can learn to recommend personalized services, which are varied with respect to time.

6. References

[1]R.Kamal,J.H.Lee,C.K.Hwang, S.I.Moon, C.S.Hong and M.J.Choi, Psychic: An Autonomic Inference Engine for M2M Management in Future Internet APNOMS 2013, 25-27 September, Hiroshima, Tokyo, 2013.

[2] R. Kamal, C. Hong, and M. Choi, “Psychic inference engine,”http://code.google.com/p/psychic-inference-engine/, 2013, [Online; accessed 24-May-2013].

[3] “Questionnaires of e-mail survey on usage-pattern of smart-device services,” http://goo.gl/fjRTb, 2013, [Online; accessed 24-May-2013].

[4] “Dataset of traffic-info of lab-members on the 25th February,”http://goo.gl/hhTDL, 2013, [Online; accessed 24-May-2013].

[5] Van Trees, Bayesian Bounds for Parameter Estimation and Nonlinear Filtering/Tracking, IEEE Press, 2007.