Network Complexity Research Group
Internet-Draft
Intended status: Informational
Expires: April 21, 2014

R. Sircar
Ericsson
M. Behringer
Cisco

October 21, 2013

Using Entropy as a Measure for Changes in Network Complexity
draft-sircar-complexity-entropy-00.txt

Abstract

For the evaluation of network designs it is desirable to express their complexity in objective, measurable metrics. Previous work has shown that a large number of distinct, partly dependent scales play a role in overall complexity. This document proposes the use of multi-scale entropy metrics to describe the complexity of a network. We observe that the complexity of a network which undergoes no changes over a longer time period is constant. Conversely, when a network undergoes changes entropy is increasing; this is independent on whether the changes make the network more or less complex. In other words, also a simplification effort increases complexity temporarily.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at http://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on April 21, 2014.

Table of Contents

1. Introduction and Problem Statement

When designing networks, low complexity is an often cited goal. There are complexity metrics for some aspects of a network, for example graph complexity or software complexity. But there is no scientific way to determine the overall complexity of a network.

Every network has a lifecycle of its own. A network is envisaged and architected based on the requirements to provide certain services. Based on the architecture, it is then designed, implemented and finally managed. Each of these phases in the lifecycle of the network is impacted by the complexity of the network.

The overall complexity of a network can be broken down into smaller parts, such as software complexity (of the operating systems), configuration complexity, or the complexity to trouble-shoot problems. On a component level, tradeoffs can be observed between various parts of the network as described in [draft-ncrg-network-design-complexity]. For example various elements of state, such as forwarding state show a tradeoff with other network properties, such as optimal forwarding behaviour.

All networks undergo further changes based on network growth, changes in the requirements or needs as well as goal, obsolescence of equipment, addition of services, change of vendors or technology or evolution of technology. In a steady state condition of the network, complexity remains constant.

Each of these changes is impacted by existing complexity of the network. The change impacts the complexity level of the network.

Thus, change in the network becomes an important parameter to be defined and measured. This document introduces a term Network Complexity Entropy to measure this change of network's complexity. This is based on multi-scale entropy. In this document we describe network complexity using multi-scale entropy analysis. The underlying assumption is that complexity is related to change. We postulate that a network with a decreasing number of changes over time gradually becomes less complex to operate. Conversely, an increasing number of changes in the network means that the network is becoming more complex.

To support our assumption we define a set of measurable variables which influence the complexity of the network, for example configuration state or resilience of the network. The overall network entropy is a function of those variables. As those variables become stable over time, the overall entropy goes down. In our interpretation this is illustrating the decreasing operational complexity of the network over time. We believe that entropy is a good approach to capture time based aspects of network complexity.

2. Components of Network Complexity

The document "A Framework for Defining Network Complexity" [I-D.irtf-ncrg-complexity-framework] gives a number of examples of network components, such as configuration, protocol state, operating systems, network hardware such as routers and transmission equipment, etc. These components are used in this document as variables in for the multi-scale entropy analysis.

For example, consider the configuration of all network devices. While the network is evolving and growing, there will be permanent change to the global configuration of the network. However, assuming no changes in the services provided, for some period of time the overall operating system state could be unchanged, because the same OS is deployed in more locations. The entropy of the configuration is increasing in this example, whereas the entropy of the operating systems is constant.

There is no complete list of network components that should be considered for the analysis proposed in this document, as XXX offers only a categorisation with some examples. The actual analysis will therefore depend on the variables chosed, and is likely to represent only a partial view of the change in complexity of a network.

The following section explains in detail how multi-scale entropy analysis can be applied to measure changes in network complexity.

3. Multi-Scale Entropy Analysis

Entropy was first defined in thermodynamics to model real world phenomena. Later it was used by Shannon to define the expected value of information contained in a message. There has been a good amount of research where entropy has been used to measure topological structural complexity of the network. [XXX add references] In each of these definitions, change was measured over regular periodicity of univariate time series on a single scale. Certain authors have used similar definitions or mathematical constructs to assess structural complexity of the underlying traffic or signal generating mechanisms. These are very useful to evaluate repetitive patterns which are generally quite predictable (e.g., periodic). Some of these approaches have also been used for completely unpredictable (e.g., uncorrelated random) signals, but the results are not always very intuitive.

In summary, all the above mentioned approaches are based on ergodic theory for dynamical systems with time as an invariant measure. As suggested by Lloyd Demetriusa et. al. XXX , the importance of entropy - and its applicability to network theory - rests on three fundamental properties:

1. Network entropy is an invariant of the dynamical system. It characterizes the structure and the ergodic behaviour of a dynamical system operating on the network.

2. Network entropy is positively correlated with robustness.

3. Evolutionarily stable states are characterized by extremal values of network entropy. Maximal values of entropy arise where evolution changes Complexity.

In this draft, we reuse the definitions and research done previously to Network Complexity Entropy.

Networks and their evolution are truly complex and long-range correlations at multiple spatial and temporal scales may be required to measure entropy. The multi-scale entropy (MSE) method proposed by Costa et. al. [XXX] explicitly quantifies the amount of structure (correlation)in real world time series. This should help in defining the underlying system os system's complexity. This approach has been used by many researchers in recent times to identify such structural complexities in Internet Traffic flows XXX .

The MSE method evaluates sample entropy of coarse grained (averaged over increasing sequential segment lengths) univariate time series;

the underlying idea is that coarse graining defines temporal scales, hence a system without structure would exhibit a rapid decrease in entropy with an increase in time scale. The existing MSE algorithm has been proven to be able to distinguish between time series with different degrees of complexity and its extensions have included more rigorous definitions of time scales.

For a discrete random variable X, taking values {x1, ..., xn} with probabilities {p1, ..., pn}, the information entropy or Shannon entropy of X is then defined as the mean information content, yielding S(X) = Sum {(over i=1...n) pi log pi}. Logarithm used here is the natural logarithm.

Since the networks generate variable amount of data as well as noise, MSE uses Approximate Entropy (ApEn) Algorithm [XXX] developed by Steve Pincus. This has been used successfully in various disciplines. In MSE or MultiScale Entropy, Costa et. al uses ApEn, but constructs multiple coarse-graining time series by averaging the data points within non-overlapping windows of increasing length. Now, let us define "Network Complexity Entropy", such that this is calculated using MSE algorithm. Thus, Network Complexity Entropy is the MSE of the selected variables measured over time.

The goal of every network designer would be to optimize this Function. Then, the goal of the Network Complexity Entropy is to:
Minimize [Network Complexity Entropy for topological structure, Network Complexity Entropy for Traffic, Network Complexity Entropy for Control Plane State,...] ---- our Objective Function.

Now, we need to now formulate 'n' Network Complexity Entropy for the network in study:

1. Network Complexity Entropy for Topology - Structural Complexity

2. Network Complexity Entropy for Systems of System - Complexity due to overall Network Structure and interfaces including VNO and multi-vendor aspects

3. Network Complexity Entropy for Control Plane State - Complexity due to Signaling, Routing and other Protocol Complexity

4. Network Complexity Entropy for Traffic and Optimal Forwarding State / Paths

5. Network Complexity Entropy for Configuration State

6. Network Complexity Entropy for Policy Architecture

7. Network Complexity Entropy for Cost and Human impact - Complexity

due to financial systems, management systems and Systems Integration etc.

4. Applying MSE Analysis to Complexity

[now we define a (small) set of the above variables, and create an entropy function with those. We model this function, and show results of the modelling.]

5. Validation of the Approach

[Ideally, we want to pick a real-life network, and check the above results against this network. Not sure this is feasible...]

6. Future Work

[undoubtedly, there will be a lot of open questions...]

7. Informative References

[I-D.irtf-ncrg-complexity-framework]
Behringer, M. and G. Huston, "A Framework for Defining Network Complexity", draft-irtf-ncrg-complexity-framework-00 (work in progress), February 2013.

Authors' Addresses

Rana P. Sircar
Ericsson

Michael H. Behringer
Cisco