Metcalfe's Law

Pranav Kasela, Federico Moiraghi

Abstract

Sarnoff was an American businessman and a pioneer of the radio and television, he is credited with the Sarnoffs law, which states that the value of a broadcast system grows linearly with the number of users. Robert Metcalfe born in 1946 is an electrical engineer, applied mathematician, computer scientist and businessman, he is the co-founder of Ethernet together with David Boggs.
Sarnoffs law was the base of the Metcalfes law which states that in the case of Ethernet the growth is quadratic with the linear growth of users. The law was formulated by George Glider in 1993 after a discussion with Metcalfe. This law is mathematically correct but has some strong hypothesis for business applications.
Reeds law states that the utility of large network grows exponentially with the size of network 2^n, which is more than the number of users (n) and the possible pair of connection (n^2). Both these laws have been criticized by a MIT mathematician, Andrew Odlyzko, who states that both Reed and Metcalfe overstates the growth of internet since they dont take in account the human tendency to saturate the number of connections they make with others using the Dunbars number. He empirically states due to different analogies that the growth should be $n \cdot \log n$ for n very large and states that Metcalfes law is valid only for small network (small n). We will explain Odlyzkos hypothesis and generalize Metcalfes law using the mathematical series concept.
We will prove the law stated above and their economical implications and explain in which case they have a real evidence.

Math Proofs

We start with the Sarnoff's rule, which is pretty simple. It states that the value of a broadcast network is linearly proportional to it's users, so basically if we have n users we will have that the value growth is $V \cdot n$ (where V is the value of a connection). A simple example would be a radio or companies like Netflix.

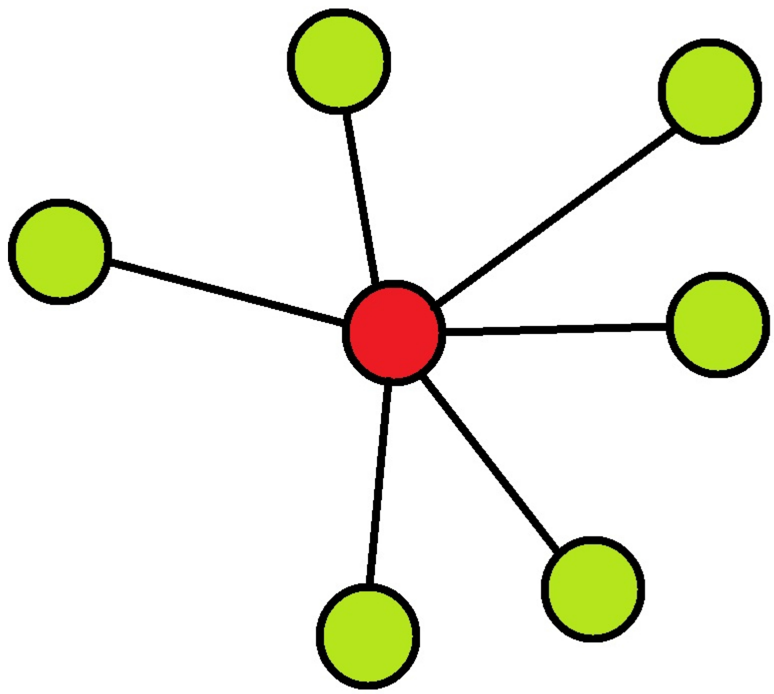


Figure 1: Red is the broadcaster, green are the users.

Proof : The proof is trivial since we have n users connected to single node, which is considered to be the streamer. Assume every connection has value V then we will just sum V n-times obtaining a value of $\sum_{i=1}^{n} V = V \cdot n$ $\square$

Next is Metcalfe's Law, which states that the growth rate is proportional to the square of number of users.
Proof : The idea is similar to the Sarnoff's law: fixing one of the n nodes,

we can see that using Sarnoff's law (or common sense) we have $n-1$ connections.

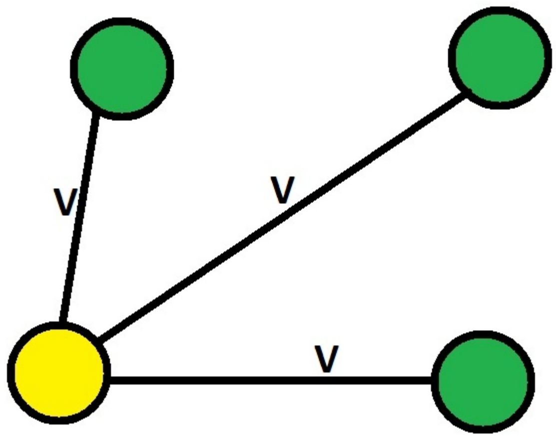


Figure 2: we are connecting the yellow node to the greens.

Assuming that every connection has the same value V, we have that from the first node we are gaining $V \cdot (n-1)$ value; now, fixing a second node, we obtain again $n-1$ connections; but remember that we have already considered the connection with the first node, so we have $n-2$ connections, thus we generate $V \cdot (n-2)$ value.

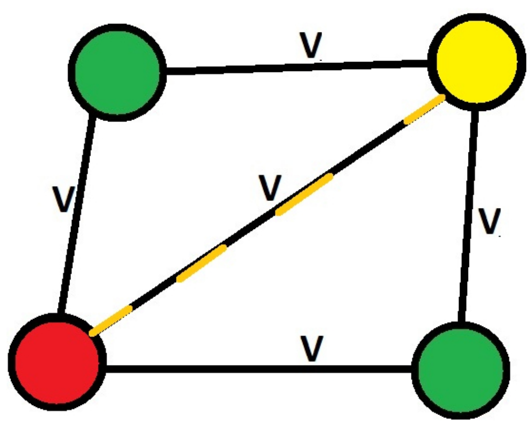


Figure 3: we can see that connecting the yellow we are doubling the dashed yellow connection to the red node, so we need to remove it in our calculations.

So, proceeding this way and summing all the values, we have that the total value is $\sum_{i=1}^{n-1} V \cdot i = V \cdot \sum_{i=1}^{n-1} i = V \frac{n(n-1)}{2} \propto \frac{V}{2} n^2$ $\square$

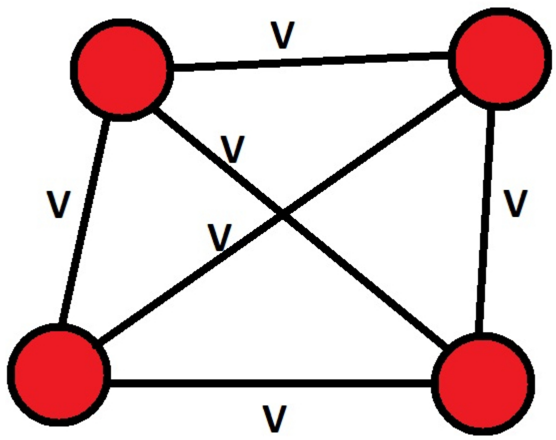


Figure 4: A totally connected network with n=4

Now we try to prove a general formula, without assuming that every connection has the same value and assuming that n is large enough; otherwise the optimal solution is Metcalfe's Law, which has a strong hypothesis and cannot be satisfied if the number of users is very high simply because we cannot connect one of them to all others. For example if we had only 10 users on a social media then most probably all of them will connect between themselves without discrimination; but if we were to have millions of users, without using some algorithm to extrapolate value out of them, we cannot have that these millions (it would be $\propto$ million^2 using Metcalfe's law) of connection will have the same value: it's not only a time restriction but a space restriction too. Since the human behavior on a large scale is similar, just see the Big Data mindset, we might use the following generalization: fix a user, sort the others on the value of connection and call the values V_1 for the first connection, V_2 for second, ... , V_{n-1} for the last one.
Since we assumed this to be true for every user, we have that the total value of network is $\propto n \cdot \sum_{i=1}^{n-1} V_n$. We can see if we have $V_{n-1} = ... = V_2 = V_1$ the Metcalfe hypothesis we have the $\sum_{i=1}^{n-1} V_n = n \cdot V$ thus the network value is $\propto V \cdot n^2$ which was the Metcalfe's Law. The network value is based on the series $\sum V_i$ if number of users n is big enough, so if the series converges

we will have that the value connection is linear, if it diverges we will need to study the asymptotic behavior of the series. Surprisingly, Sarnoff's law is not only valid for broadcast networks but can be true also for fully connected network.

One hypothesis that we can make is the Odlyzko's one, that the values of connection drops harmonically, in our case assume $V_1 = V$, due to this assumption we will have $V_2 = V/2, V_3 = V/3$ and so on, so the series will be $\sum_{i=1}^{n} V/i$, we know that for n large enough we have that $\sum_{i}^{n} 1/i \sim \log n$, so we have that the series is $\sim V \log n$, thus the value is $\propto n \log n$ (see [5]).

We showed 3 different values of the value for the connection so which one is the right one? There is a big difference between $n, n \log n, n^2$ we can say that this three represent a pessimistic, average and optimistic growth of a network.

Socio-Economic Interpretation.

Metcalfe's Law assumes that each node is linked with the whole net and all connections have the same importance. Those assumptions are considered too strong, so new mathematical laws have been presented; but empirical studies have demonstrated the veracity of Metcalfe's Law in popular social networks[4][7] even if premises of are not satisfied. However, in a social environment the human nature of nodes must be considered; moreover, algorithms used to improve connections quality and the possibility given by the platform to *share* contents may be enough to make the value of the whole network increase according to a quadratic function n^2.

Estimating the value of a net is important in an economic contest: costs and earnings are the base of any business. Costs are not considered just in terms of money, but also as resources needed to run the business, like server's disk space and computational power. Decisions concerning the server depends on the costs of the whole network; and since it is necessary to run the business, estimation can not be lower than real costs: the service risks to be interrupted and the value of the network drastically decreases to zero.

1 A real case: Facebook.

Facebook, the biggest network in the world owned by a public company, is demonstrated to have a value proportional to the square of active users[4][7]. But assumptions of Metcalfe's Law are apparently not satisfied due to the limit to the number of *friends* that a single person can have and the fact that users do select their contacts according to various strategies[1].

However, this is not a problem, because Facebook's algorithm and the possibility to *share* contents make links between people that apparently are not connected: the exchange of value is imposed by Facebook, even without user's consent. According to Pasquali[1], Facebook is not a single network, but an aggregate of small (social) networks that exist also offline and a big (informatic) net composed by the aggregate of all users. While small social networks are not interesting in this study, the fact that users can interact with the whole Facebook is fundamental. Users, in fact, can *share* contents, so that they can reach more nodes (*friends of friends* or *public pages*), or, in particular cases, a *post* can become *viral*. In addiction to this, Facebook's algorithm minds the creation of new links, suggesting new contents that is sure to be liked by the user or suggesting new people or pages.

The value exchanged through Facebook's network is maximized by the algorithm: not all information are presented to the user, but they are filtered according to personal interests. The unique experience during the *flow* of feeds can cause dependency: the time spent on the platform, the facility of access and the quality (or targeting) of contents are so high that saturation is not easily reach.

2 Economic Evaluation of a Net.

In general, networks are not the product of the business but just the distribution channel: in this case, the net can be evaluated equal to earnings coming from it; in this case the model is just linear. This is just a simple and empirical way to estimate a value, but works even for big companies that offer a payment service.
A generalization of this criterion is to calculate the sum of value exchanged between each node; this can be easily done in e-commerce sites or for a transport nets[6] .

Metcalfe's Law is necessary when no or small data are available, so the value can not be known and can be just estimate. This is the case of advertisement selling: the value of the space is not known, but companies have to define a price according to the distribution that the advert may have in the network.

Estimating the value is important when balance-sheet is compiled or budget is calculated. In addiction, according to Italian's Law, the balance-sheet must be *prudent*, so costs can not be under-estimated. With this in mind, Metcalfe's Law is correctly adopted: if costs follow a $o(n^2)$ model, law would not be broken; in fact, it is empirically demonstrated[7] that costs of a social media follows this law. Using the same criterion, if no data are available, earnings are not well estimated: is more prudent using a lower formula, like $n \cdot \log n$.

3 Net saturation.

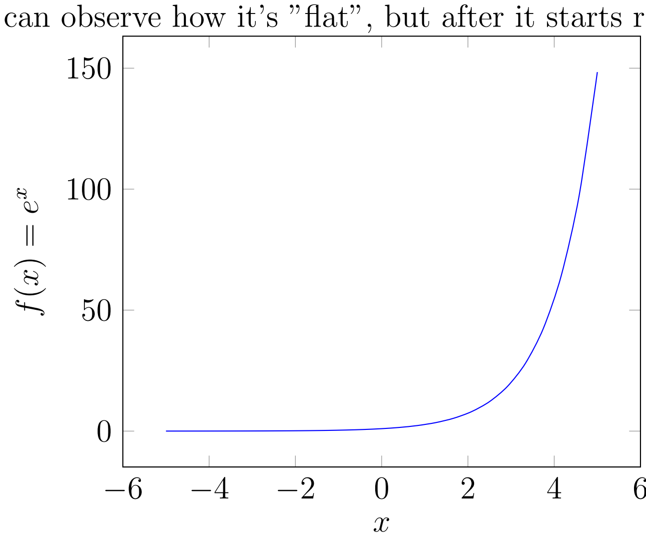
In a network, nodes are often normal people: in a finished time not all the value transmitted can be extracted. This means that a person has no time to communicate with and may not receives all communications from the whole network. The term *saturation* is used to mean that a person can not exchange the maximum value with all other nodes due to the mole of information that receives. In addition, not all communications have the same value: user is covered with information, for the most useless, that can not manage.
To maximize the value that a network can transfer, filters of any type are used: from simple spam filters, to automatize the check of emails, to server algorithms that suggest new actions and connections. In this way, the value of links increases due to the growth of information that the user can manage.

4 Conclusions.

The value of a network is not easily estimable: in general, the value is inferior to n^2 presented by Metcalfe, but with some tricks is possible to a net to reach the maximum value. In particular, Facebook's network, due to its possibility to *share* contents and a sophisticated algorithm that uses tons of information about each user, can be evaluated according to Metcalfe's Law; but this mathematical function is not always valid, even in social environment, due to human nature: social networks are not so inter-connected, and a informatic infrastructure mirrors the offline world. And more in general, other laws, of minor magnitude, are empirically and successfully used.

The L growth of the world?

In the past years the technology has been advancing really fast, for example the Moore's Law states that the number of transistors in a processor doubles every 18 months (it grows 60% every year), while the Nielsen's Law states that the growth of the bandwidth is 50% every year(doubling every 21 months)[2].Actually it is not on a recent phenomena, if we were to see on a larger scale most of the technological growth has been on a exponential rate, but, as we can see on an exponential plot, the initial growth is pretty slow, in fact, we can approximate the exponential function near 0 with a linear function ($e^x = 1 + x + ...$ using the Taylor formula near 0) and before it we can observe how it's "flat", but after it starts rapidly increasing.



It's clear from the graph above that with L functions we mean a function that resembles the capital letter L. Usually all the functions above the quadratic function n^2 can be considered a L function. These kind of function is what any industry would like to achieve. The Moore's Law is widely accepted as a goal for the industry (this is one of the reason why Moore's Law is still valid today). These are usually empirical laws that predicts reality very precisely for a certain period and then at a certain point starts failing.

One law which has some mathematical foundation, but no empirical evidence is the Reed's Law which states that the complexity of a network where we can have sub-networks grows like 2^n.

Proof: It's fairly simple, all we need to know that given a set of n elements, the number of subgroups that we can have are the possible combination of k<n elements, which is $\sum_{k=1}^{n}\binom{n}{k} = 2^n - 1$ $\square$

This law has no empirical evidence but is used in economical fields as a proof that a network creates more value with more sub-networks.

We could add a lot of these L laws, Metcalfe's Law is one of them. Also the quantity of data we create is growing exponentially. All these vertical growth conducted Kurzweil [3] to think that all the worlds technology has been growing and will continue to grow exponentially. Here we show the simple mathematical model proposed by Ray Kurzweil in his report [3]

Let V be the velocity(power) of computing,

W the world knowledge,

t the time,

And let's assume that V = C_1·W, so the computation is a linear function of knowledge. And that we have that knowledge is a cumulative:

$$W = C_2 \int_0^t V(y)dy$$

Using these two hypothesis we have that:

$$W = C_1 C_2 \int W$$

which has the solution

$$W = C_1 C_2 C_3^{C_4 \cdot t} \quad \Rightarrow \quad V = C_1^2 C_2 C_3^{C_4 \cdot t}$$

simplifying the constants we have

$$V = C_a C_b^{C_c \cdot t}$$

which could be considered as a form of Moore's Law. Now Kurzweil extends this model:

N = Expenditure for computation

V=C_1·W (as before)

N = $C_4^{C_5 \cdot t}$ (We assume that the expenditure for the computation is growing at an exponential rate)

W=$C_2 \int_0^t N \cdot V$

from here we derive easily that

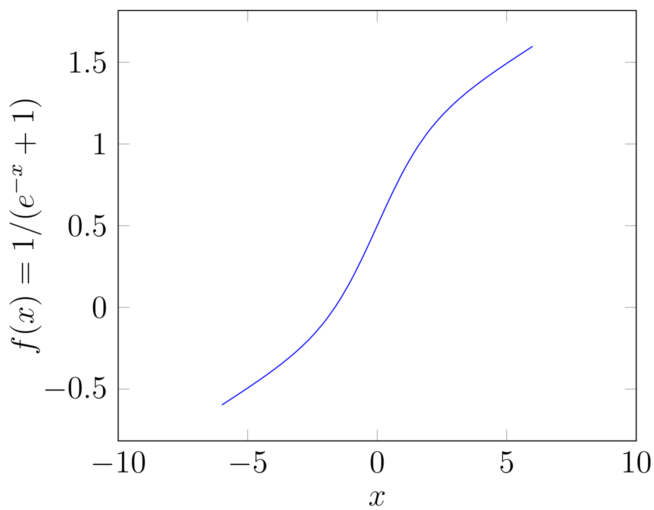
$$W = C_1 C_2 \int_0^t C_4^{C_5 t} W$$

$$W = C_1 C_2 C_3^{(C_6 t)^{C_7 t}}$$
$$V = C_1^2 C_2 C_3^{(C_6 t)^{C_7 t}}$$

simplifying the constants:

$$V = C_a C_b^{(C_c t)^{C_d t}}$$

Which is a double exponential function. The constants can be calculated with the empirical data and assuming to know that the human race can to $2 \cdot 10^{26}$ calculation we can compute the time when the machines will surpass the human power calculation, this event is called "singularity" by Kurzweil. The Moore's Second Law (Rock's Law) states that the price of production and fabrication of a semiconductor doubles every 4 years, and the prices of the production transistors are increasing too. There might be a time where the production cost is too high to continue to invest into following Moore's Law so the Rock's Law and Moore's Law will eventually collide and at that point the we won't see an exponential growth of computations.
Kurzweil assumes the veracity of the Moore's Law, but as we just stated it is not easy for Moore's Law to stand until the so called "Singularity", the Moore's Law was initially projected only for one decade, it is considered a big deal that is it still valid today, one of the biggest motive is, as mentioned above, the fact that the industry is driven by this law, every major manufacturer of processors, try to create processors powerful enough to match Moore's expectations. Recently since the cost of production is also increasing, it has become even more difficult to follow Moore, so probably in a few years, either the Moore's Law will fall or we will need to correct it again. Another scenario is that the growth is temporary an exponential and later it will hit a ceiling (an asymptote) it the growth might stabilize at that point or it could change the behaviour after a certain point (it might become a linear function). For example the following the function is an example of what we mean with a local exponential function which has an oblique asymptote.



The possible behaviour we just proposed puts in great risk the veracity of Kurzweil singularity, but one this is sure, the technology is increasing rapidly, so we need to give a shape to the future, i.e. create ethical principles and new laws in order to maintain a certain order in this new possibly disruptive revolution and make sure that they are followed. Otherwise we will have not a singularity but a repetition of history, due to the new possibilities but the same old human fault of not prevent disasters and wait until they happen just to try and cure them.

References

[1] Giovanni Boccia Artieri et al. *Fenomenologia dei Social Network*. Milano, Italy: Guerini Scientifica, 2017.

[2] Luciano Floridi. *The Fourth Revolution: How the infosphere is reshaping human reality*. Oxford, United Kingdom: Oxford Press, 2014.

[3] Ray Kurzweil. "The Law of Accelerating Returns". In: *Kurzweil AI* (Mar. 2007).

[4] Bob Metcalfe. "Metcalfe's Law after 40 Years of Ethernet". In: *IEEE Computer* (Dec. 2013).

[5] Andrew Odlyzko and Benjamin Tilly. "A refutation of Metcalfe's Law and a better estimate for the value of networks and network interconnections". In: *University of Minnesota* (Mar. 2005).

[6] Joe Weinman. "Is Metcalfe's Law Way Too Optimistic?" In: *Business Communications Review* (Aug. 2007).

[7] Xingzhou Zhang, Jing-Jie Liu, and Zhi-Wei Xu. "Tencent and Facebook Data Validate Metcalfe's Law". In: *Journal of Computer Science and Technology* (Mar. 2015).