March 19, 2018

Normalized weightage framework to report tangible and non-tangible data from heterogeneous and dynamic data sources within a context to assist in decision alignment

Venkat Maddali
Hewlett Packard Enterprise

Lakshmi Deepak Kotla Venkata
Hewlett Packard Enterprise

Gautham Natanasabapathy
Hewlett Packard Enterprise

David C. Fogaren Jr.
Hewlett Packard Enterprise

Umesh Pathak
Hewlett Packard Enterprise

Normalized weightage framework to report tangible and non-tangible data from heterogeneous and dynamic data sources within a context to assist in decision alignment

In a process involving decision making, both humans and machines prefer the data to be in a yes or no format. A number by itself to represent a quantifiable entity may not have enough value to make a decision until the data has been triangulated with other interrelated quantitative and qualitative data points and normalized onto a common scale. Our framework mentioned here in the proposal, helps in normalizing such quantitative and qualitative data from various data sources to a set of common metric for decision making.

As we agree that, a comparative analysis is typically considered valid when the comparison is normalized. In a simple since that is apple to apple comparison in our daily life. Converting this common sense and contextual decision to a computer program is made possible with our framework. Even a defined measurement values with qualifier are context-comprehended in instances such the current speed, voltage, and voltage or usage of any resource like vacancy, availability etc. For a defined value, system can determine the action to be performed, however as mentioned, in some scenarios the variables are neither linear nor quantifiable.
For example, consider the real time scenario, when a system has to decide how to act and normalize information for various decision variables in a Converged Infrastructure scope like Converged Systems, Integrated Infrastructures, Converged Architectures, and Hyper Converged Offerings etc.

Compute: Power, CPU's, Memory, Cores etc. (number of options: ci)
Storage: Disk Kind, Resilience, Speed, etc. (number of options: si)
Network: Throughput, bandwidth, scalability etc. (number of options: ni)
Service: 1 click, component level, solution level etc. (number of options: ts)
Operating Environment: Full support, enterprise support, open source (number of options: oe)
Workload: Licensed support, Enterprise support etc. (number of options: w)
Installation: Factory built, channel partner built, services built etc. (number of options: i)
Delivery: white glove delivered, Channel delivered etc. (number of options: d)

If you notice, each variable has a different quantifiable measure with many combinations to choose from for making right decision.
Combinations Formula: C_r^n (The standard formula to get the number of combination in selecting 'r' options from a pool of 'n' options where the position/order does not matter).
Combined Combinations Formula (For more than one set): $C_{r1}^{n1} \times C_{r2}^{r2} \ldots \ldots C_{rn}^{rn}$
Converged Systems Possible Combinations (pick one of each, r=1) = $C_1^{ci} \times C_1^{si} \times C_1^{ni} \times C_1^{ts} \times C_1^{oe} \times C_1^{w} \times C_1^{i} \times C_1^{d}$
Converged Systems Possible Combinations (pick one of each, r=1 and there are 10 options per variable and all infra variables are 15.ci=si=ni =15, and ts=oe=w=i=d=10) = $C_1^{15} \times C_1^{15} \times C_1^{15} \times C_1^{10} \times C_1^{10} \times C_1^{10} \times C_1^{10} \times C_1^{10}$ = $15 \times 15 \times 15 \times 10 \times 10 \times 10 \times 10 \times 10 = 3375 \times 100000$**= 337.5M**
Converged Systems Possible Combinations (pick 2 of each infra options, r=2,1 and there are 10 options per variable and all infra variables are 15.ci=si=ni=15,and ts=oe=w=i=d=10)= $C_2^{15} \times C_2^{15} \times C_2^{15} \times C_1^{10} \times C_1^{10} \times C_1^{10} \times C_1^{10} \times C_1^{10} = 105 \times 105 \times 105 \times 10 \times 10 \times 10 \times 10 \times 10 = 1157625 \times 100000 \approx$ **115.7B**

Real-time scenarios:

Case 1: Some of the variables or are on a different scale
International System of Units is responsible to ensure a conversion between similar metrics are intact globally. That is exactly why a weight or mass when represented as 51lb or 23kgs conveys the same message. Or available computer storage as 1Gigabyte or 1000 Megabytes represent the same and behaves the same. But if the decision has to be made between available storage of 1 GB and a large disk (No knowledge of B or MB or GB).
In such a scenario, traditional manual approach is to assign a normalized metric like how many copy of a same file can be saved etc. However for a computer program it is not that straight forward because lack of subject matter knowledge and not having enough information to drive the decision based on just one variable.

Case 2: Some of the variables does not have any measurement scale
Extending the above scenario, if a decision has to be made between completely heterogeneous variables with or without any metric. For an example, if a decision has to be made between available storage of 1 GB Self Installed and an arbitrary Elastic Disk as a Service. In such scenarios this framework provides a weighted model derived based on several parameters to bring these two equations into a common scale to have a 'relative' evaluation for the process or user to decide.

Positive Parameters (more the better): *performance (pi), security (si), popularity (po), support (su), warranty (wa), speed (sp)* storage (st), resilience (ha), redundancy (re), end of life (eol), management stack (mg)

Negative Parameters (less the better): *price (pr), ordering time (ot), delivery time (dt), setup time (st), maintenance (ma), service complexity (sc), usage complexity (uc)*

In an ideal start scale, let assume parameters start with 1. In an instance of this framework used in tool, named ***Alpha***.

Ideal Multiplier=1 = pi= si= po= su × wa= sp=st= ha= re × eol= mg=pr=ot=dt=st=ma=sc=uc
Alpha ideal Score = $\frac{pi \times si \times po \times su \times wa \times sp \times st \times ha \times re \times eol \times mg}{pr \times ot \times dt \times st \times ma \times sc \times uc}$
Converged Infrastructure Score in Project Alpha (Split multiplier on a scale of 100) =
$$\frac{pi \times si \times po \times su \times wa \times sp \times st \times ha \times re \times eol \times mg}{pr \times ot \times dt \times st \times ma \times sc \times uc}$$

Ex: Alpha Converged System Weightage (For a secure and stable solution with an industry supported workload, Integration, setup and life cycle support.)
Normalizing Multiplier=x
Performance=pi=3x
Security=si=4x
Popularity=po=x
Support=pi=3x
Warranty=wa=2.75x
Speed=sp=2.5x
Storage=st=2.5x
Resilience=ha=3.5x
Redundancy=re=3.5x

End of Life=eol=3x
Management=mg=3.5x
Price=pr=3x
Ordering Time=ot=2.5x
Delivery Time=dt=2.5x
Setup Time=st=3x
Maintenance =ma=3.5x
Service Complexity=sc=3x
Usage Complexity=uc=2.5x
Total=52.75x
52.75x=100
x=100/52.75=1.895 ≅1.9

***Note**: The weightages here are for illustration purpose only for an arbitrary Alpha Converged System. Each vendor can chose to use their own multiplier in this equation based on their metrics.*

***Alpha Converged Infrastructure combination weightage Score** =*

$$\frac{5.7pi \times 7.6\, si \times 1.9\, po \times 5.7\, su \times 5.2\, wa \times 4.75sp \times 4.75st \times 6.65ha \times 6.65\, re \times 5.7\, eol \times 6.65\, mg}{5.7pr \times 4.75ot \times 4.75dt \times 5.7st \times 6.65ma \times 5.7sc \times 4.75uc}$$

To put the equation in perspective, if a combination has high power CPU of 2.5 GHz with 3 days of setup time and a different combination has 4.2GHz CPU but the setup time is 5 days, keeping all other parameters same in both the cases, then what should be preferred option?

Alpha Converged System Scenario 1 (2.5GHz and 3 days) =

$$\frac{5.7pi \times 7.6\, si \times 1.9\, po \times 5.7\, su \times 5.2\, wa \times 4.75 \times 2.5 \times 4.75st \times 6.65ha \times 6.65\, re \times 5.7\, eol \times 6.65\, mg}{5.7pr \times 4.75ot \times 4.75dt \times 5.7 \times 3 \times 6.65ma \times 5.7sc \times 4.75uc} =$$

$$\frac{4.75 \times 2.5}{5.7 \times 3} = \frac{11.875}{17.1}$$ ***= 0.694 (Better)***

Alpha Converged System Scenario 2 (4.2GHz and 5 days) =

$$\frac{5.7pi \times 7.6\, si \times 1.9\, po \times 5.7\, su \times 5.2\, wa \times 4.75 \times 4.2 \times 4.75st \times 6.65ha \times 6.65\, re \times 5.7\, eol \times 6.65\, mg}{5.7pr \times 4.75ot \times 4.75dt \times 5.7 \times 5 \times 6.65ma \times 5.7sc \times 4.75uc} =$$

$$\frac{4.75 \times 4.2}{5.7 \times 5} = \frac{19.95}{28.75}$$ ***= 0.693***

As we can notice in the above calculation, it is clear from the framework that the scenario-1 seems to be a more optimal route to go than scenario-2. This data point can be used by anyone to make informed decision with minimal or no understanding of the underlying parameters.

Disclosed by:
Venkat Maddali Inventor - Hewlett Packard Enterprise
Lakshmi Deepak Kotla Venkata - Hewlett Packard Enterprise
Gautham Natanasabapathy - Hewlett Packard Enterprise
David C Fogaren Jr. - Hewlett Packard Enterprise
Umesh Pathak - Hewlett Packard Enterprise