Modeling the Dynamics of the Internet Platform Users' Volume

Ethan Masse

April 21 2023

1 Executive Summary

Certain internet platforms which are popular in today's online landscape (such as online sellers like Amazon and eBay, or dating sites like Match.com) use effectively a two-party system: one in which the users of the platform can be divided into two groups which affect each others membership. Taking eBay as an example, users of the site are split into two groups, buyers and sellers, which interact with one another with a basic relationship that having more sellers makes new sellers less likely to sign up, but makes buyers more likely to sign up. Stated a different way, sellers repel other sellers, and sellers attract other buyers, and vice versa. This project, based on a SIMIODE project (Rayskin, 2018), will use differential equations to model different attachment functions for how the two groups on an internet platform attract one another. Because of the fact that Amazon is one of the most profitable and successful corporations in the modern world, often labeled in the top internet companies (Bloomenthal, 2022), these modeling techniques will be especially applicable to modern internet platform behavior.

Let's assume that x and y are the fraction of users out of a maximum potential number of users that are fulfilling the two roles on the internet platform. The maximum amount of users would be the amount of people that would ever potentially use the site (i.e., people with internet in their homes who use online auction sites, in the case of eBay). The values of x and y would, naturally, be capped off such that $0 \leq x \leq 1$ and $0 \leq y \leq 1$. The solution curves for x(t) and y(t) as functions of time should be confined within the unit square, from $(0,0)$ to $(1,1)$. Any deviation in the solution curve to outside of the unit square would lead to invalid results. Let us also assume that users of the same type repel each other linearly, resulting in the following model:

$$\begin{cases} x' = V(y) - x \\ y' = W(x) - y \end{cases}, \tag{1}$$

where $V(y)$ and $W(x)$ are what's called the attachment functions, and x and y are the functions $x(t)$ and $y(t)$ representing the number of users on the site at a specific time.

To consider some pros and cons of this model, the model will have some benefits that will become apparent shortly which allow for nice solution curves that terminate, but it also does not take into account any factor except the current number of users of the platform in x and y group, as the system is autonomous. However, if we take as an assumption that the site is not influenced by outside factors (like seasonality), then this model should be rather applicable to particular situations, depending on the quality of the choice of attachment functions. However, if another model is considered for the benefits of marketing which sets an initial point $(x(0), y(0))$, then the model given here can take over from that point and see the behavior of the platform after marketing is discontinued.

Because this system of equations can never deviate outside of the unit square, that means that the tangent vector to the solution curve must always point inwards at the edges of the unit square, or the edges have to be fixed points. Therefore, at for example $x = 0$, the curve must have a positive or zero x' value, and at $x = 1$, the curve must have a negative or zero x' value. This means that $V(y)$ and $W(x)$ are most safely picked when they are always positive for x and y between 0 and 1, and always less than 1 for the same range of x and y, as this guarantees the solution curve will stay within the unit square. This limits the attachment functions possible, but also guarantees that the resulting curve gives sensical percent results for the two groups of users.

To illustrate, suppose that we just take x' and see what the ranges of values are when $x = 0$ and $x = 1$. If $x = 0$ and $0 \leq V(y) \leq 1$, then x' would be between 0 and 1. This is good, because the line $x = 0$ is one of the edges of the unit square, and having a negative x' would mean that the solution curve travels outside the unit square. Let's also consider the other option, $x = 1$. Then $-1 \leq V(y) - 1 \leq 0$, which is also great because that is the rightmost boundary of the unit square and x' must be negative or zero along this line. Because the form of x' and y' are the same, without loss of generality this same technique can be applied to y' as well to get the same results: the solution curves will never escape the unit square given well-behaved attachment functions.

Next, it may shed some light on the behavior of the model by finding the divergence of the system. The divergence of the model can be calculated with

$$\mathrm{div}F\,(x, y) = \frac{\partial f_1\,(x, y)}{\partial x} + \frac{\partial f_2\,(x, y)}{\partial y}, \tag{2}$$

(Rayskin, 2018) where $f_1(x, y) = V(y) - x$ and $f_2(x, y) = W(x) - y$, the functions for x' and y'. Because the partial derivatives of x' and y' are both -1, there is a constant negative divergence of -2 at all points of the solution curve. Interpreting this result, this means that every point (x, y) has less flow coming out of the point than the flow coming in - the graph will always slow down its movement as the independent variable t increases no matter what (every point is a sink). Because the movement is determined not by t but entirely by the coordinates of the current point (the functions f_1 and f_2 are both reliant only on the value of x and y), this makes it impossible for periodic solutions or closed loops to occur in the solution curve, because the curve is incapable of entering the same point twice. This means that, for almost all well-behaved functions, the amount of users x and y given by this model will always approach some destination goal and stop - the fixed points of the graph.

The fixed points for a system of differential equations occur when $x' = 0$ and $y' = 0$. Setting each of the equations to zero, that means the following system of equations must be true:

$$\begin{cases} x = V(y) \\ y = W(x) \end{cases} . \tag{3}$$

One assumption that can be made for real-world data is that the attachment functions chosen are such that $V(0) = W(0) = 0$ and $V(1) = W(1) = 1$, because this is representing that if nobody is using the platform, nobody wants to use the platform, and if everybody of one type is using a platform, everybody from the other type will as well (Rayskin, 2018). Then, the points $(0, 0)$ and $(1, 1)$ would be two fixed points - though this may not always be the case depending on the attachment function. Other fixed points are of course possible, depending on the attachment functions chosen.

Next, the Jacobian matrix can be calculated for the right side of the model, which will shine light on the behavior of the fixed points. Naming the x and y coordinates of a theoretical fixed point x^* and y^*, the eigenvalues of the Jacobian matrix can be found which will give whether or not the fixed points are stable or unstable. The resulting Jacobian matrix of the system of differential equations is

$$J = \begin{bmatrix} -1 & V'(y^*) \\ W'(x^*) & -1 \end{bmatrix}, \tag{4}$$

which can be used to find the eigenvalues by solving the equation

$$\det (J - \lambda I) = 0, \tag{5}$$

(Lebl, 2022) or

$$\det \left(\begin{bmatrix} -1-\lambda & V'(y^*) \\ W'(x^*) & -1-\lambda \end{bmatrix} \right) = 0. \tag{6}$$

This results in the characteristic equation $\lambda^2 + 2\lambda + 1 - V'(y^*)W'(x^*) = 0$, which gives eigenvalues of

$$\lambda = -1 \pm \sqrt{V'(y^*)W'(x^*)} \tag{7}$$

by using the quadratic equation (Rayskin, 2018).

Based on the eigenvalue, different traits of the fixed point can be determined. If both of the eigenvalues are positive or zero, then the fixed point is unstable or repelling, and points close to the fixed point will have solution curves which move away from the fixed point. This is not possible because at least one eigenvalue must be negative due to the leading negative and the fact that a square root can only give positive or imaginary numbers. If only one of the eigenvalues is negative, then the fixed point is a saddle point, where the curve moves past the fixed point hyperbolically; if both eigenvalues are negative, or both eigenvalues are imaginary with negative real part (the only two other options), then the point is either an attractor, or points nearby will attract in a spiral in the latter case of imaginary eigenvalues (Woolf, 2023). There is a specific edge-case when one of the eigenvalues is zero where it is not easy to determine the behavior of the fixed point, but a deep dive into this situation appears to be beyond the scope of this project. So, barring that case, the solution curve is limited to being one of those three options.

The status of the fixed points can easily be checked by seeing if $V'(y^*)W'(x^*) < 1$ for the fixed point. If this is true, then the fixed point is a regular or spiraling attractor; if $V'(y^*)W'(x^*)$ is precisely one, then more work beyond this project will need to be done to determine stability; otherwise, the fixed point is a saddle point.

Let's look at a few examples of attachment functions and see if we can analyze the behavior of the solution curves. Consider the most simplistic example where $V(y)$ and $W(x)$ are constants, e.g. $V(y) = 0$ and $W(x) = 1$, giving the system of differential equations

$$\begin{cases} x' = -x \\ y' = 1 - y \end{cases} . \tag{8}$$

This example could represent an internet platform which only ever markets to the y group, and hides or otherwise does not market at all to the x group - and the number of members of each platform does not affect one another. For example, this could be the model for a website with astonishing growth on some new platform that they're heavily advertising, but almost no growth on an older, legacy version of the platform (imagine if MySpace and Facebook were created by the same company and MySpace was allowed to die off through lack of advertising).

Clearly, the attachment functions are both in the range of $0 \leq x \leq 1, 0 \leq y \leq 1$ as they are the constants 0 and 1, which means solution curves should be confined to the unit square. We can confirm this by setting $x = 0$, $x = 1$, $y = 0$, and $y = 1$ and testing whether or not x' or y' will point outside of the unit square. For example, along the line $x = 0$, x' cannot be negative or else the curve would escape - thankfully, setting x to 0, the change in the x direction will be 0, so it will not. This analysis succeeds for the other three sides of the unit square as well.

Now, we can find the critical points and see if they are going to be attracting or repelling. The only fixed point for the curve is $(0, 1)$, so we should expect $(0, 1)$ to be an attractor according to the fact we know the curve cannot escape and every point is a sink due to the constant negative divergence. Because $V'(y) = 0$ and $W'(x) = 0$, the eigenvalues for any point have to be -1 with a multiplicity of two. That means any fixed point would be an attractor - therefore, our one fixed point is as well. So, all solution curves starting in the unit square will be locked in the unit square and terminate at $(0, 1)$.

Next, let's take the attachment functions $V(y) = 4\,(y - 0.5)^3 + 0.5$ and $W(x) = x$, giving the system

$$\begin{cases} x' = 4\left(y-0.5\right)^3 + 0.5 - x \\ y' = x - y \end{cases} . \tag{9}$$

These are interesting attachment function because for values of $y < 0.5$, the linear attachment function W will grow faster than V, but for values of $y > 0.5$, W will grow slower. By setting these equations equal to zero as above and solving, the system of equations gives $(0, 0)$, $(0.5, 0.5)$, and $(1, 1)$ as the fixed points of the equation. Calculating the eigenvalues of these three points gives one negative and one positive eigenvalue for $(0, 0)$ and $(1, 1)$, and a negative eigenvalue with a multiplicity of 2 for $(0.5, 0.5)$. This means that $(0, 0)$ and $(1, 1)$ will be saddle points that the curve will pass by hyperbolically, and $(0.5, 0.5)$ is an attractor.

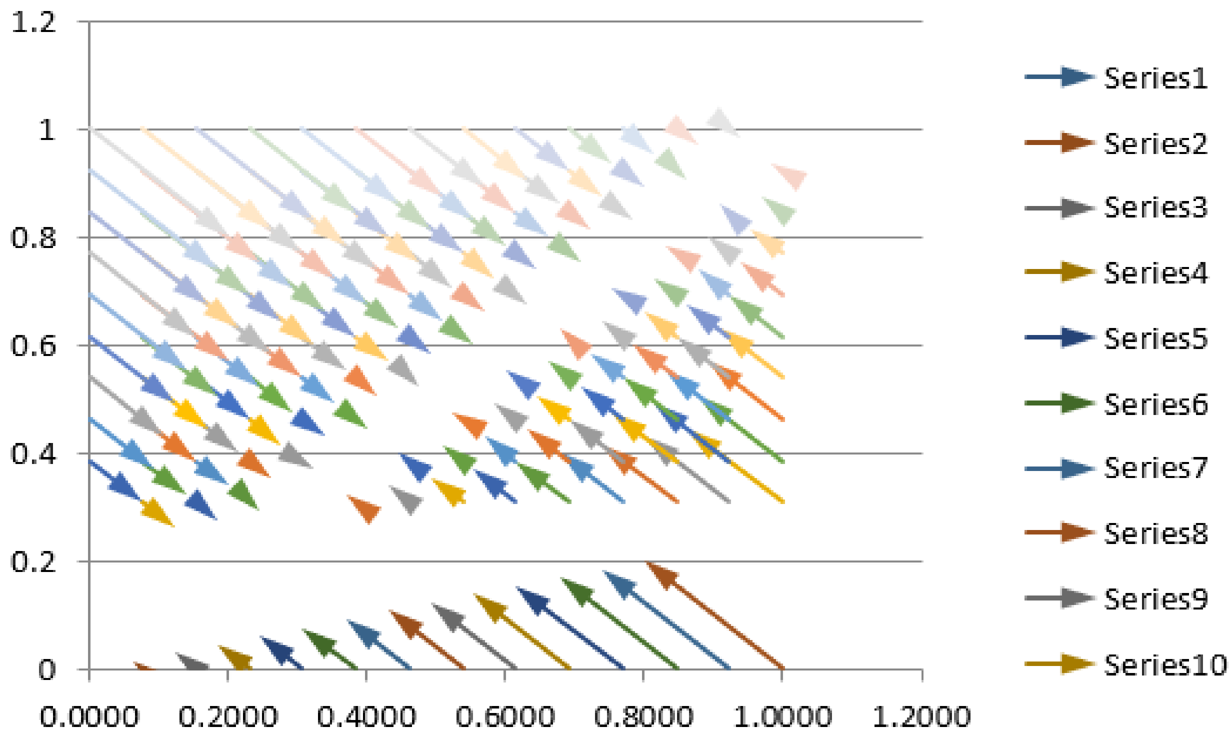
To apply this model, if this was the model for users on a dating site, for example, and we assume that there is roughly an even balance of males and females in the population (meaning the percent values of x and y roughly correspond to the same amount of men and women), that means that no matter how many men and women are using the site to begin with, eventually the number of both will even out to an exact even split where half of the potential men and women using the site are using it. This happens regardless of the fact that the models for the attachment functions for the two genders are very different from one another (i.e., the way men and women both attract one another is different). From this example we can see perhaps a little more clearly that the only thing affecting the ultimate termination point is the solution to the system of equation for fixed points, not the behavior of the attachment functions necessarily.

Let us now consider an even more complicated example:

$$\begin{cases} x' = \frac{\cosh(y)-1}{\cosh(1)-1} - x \\ y' = -x^3 + 1 - y \end{cases} . \tag{10}$$

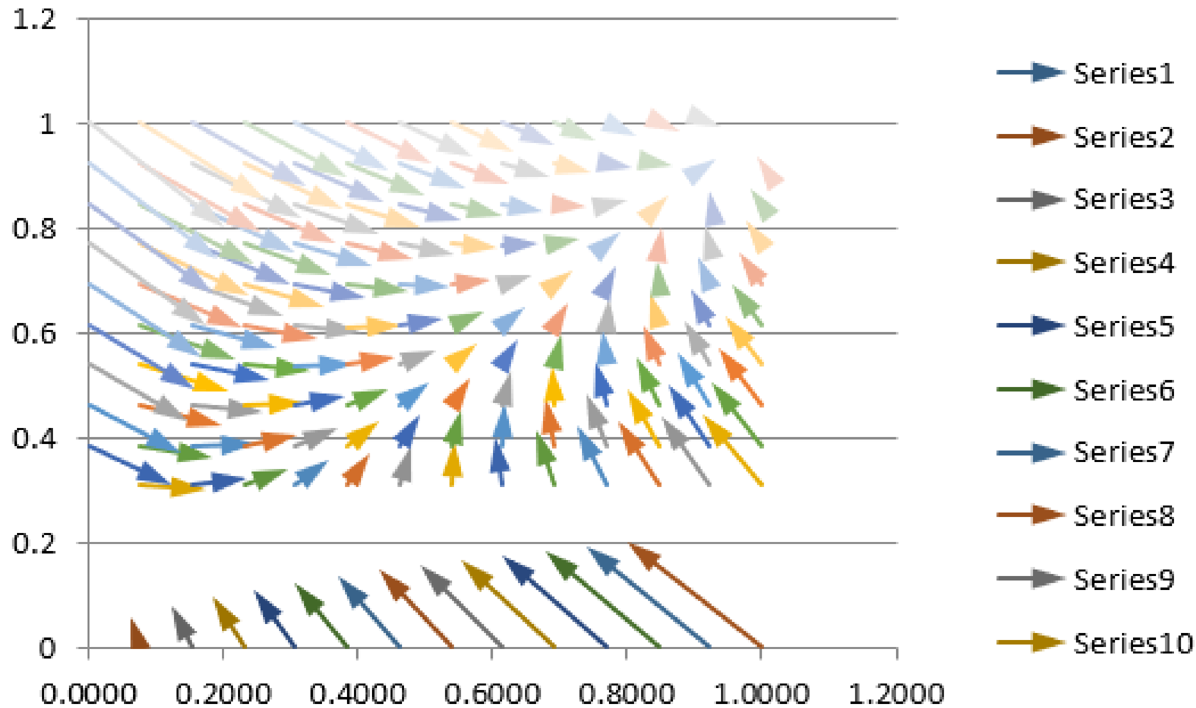
Surprisingly, this model matches up with our earlier assumptions that our attachment functions must be bounded by 0 and 1, as the $\cosh(y)$ is both shifted and normalized in order to meet this constraint. In this case, finding the fixed points will be much more difficult, as MATLAB is unable to symbolically determine the value of the fixed points. On paper this would be a problem, as the exact values of the fixed points have to be tested to see what their eigenvalues will be - however, on closer examination, this is not necessary. As long as $V'(y)W'(x) < 1$ within an arbitrarily large range of values containing the fixed point's decimal approximation, then the fixed point is guaranteed to be an attractor or spiraling attractor. Because $V'(y) = \frac{\sinh(y)}{\cosh(1)-1}$ and $W'(x) = -3x^2$, $V'(y)$ will always be positive or zero for $0 \leq y \leq 1$, and $W'(x)$ will always be negative or zero for $0 \leq x \leq 1$. Therefore, any fixed point will be guaranteed to be an attractor, and for values not along the boundary, it is guaranteed to be a spiraling attractor. So, the one fixed point located at approximately $(0.598, 0.786)$ will be a spiraling attractor.

Alongside the project sourced on SIMIODE were provided some Excel spreadsheets which can be used to create vector fields representing the flow of the model with different attachment functions quickly (Rayskin, 2018). For instance, the following is an example generated for the system $x' = y - x, y' = x - y$:



As can be seen from the vector field, it appears that all points approach the line $y = x$. This would be a good assumption because the solution for fixed points of the graph is every point $y = x$. We can also solve the system of differential equations, by adding the two equations together to get $x' + y' = 0$, splitting, and integrating, giving solution curves of $y = -x + k$. So, all curves will be diagonal lines that either travel towards or away from $y = x$. Calculating the eigenvalues gives -2 and 0 as solutions for all fixed points - this is an interesting edge case where one of the eigenvalues is negative and the other is zero, which makes it particularly challenging to determine if the fixed points are attractors. However, empirically we can see that they attract given what we know about the solution curves' form and the graph given previously. Also, by analyzing the signs of x' and y', it can be shown that above the line $y = x$, the tangent vector will always point down and right, and vice versa for below (Rayskin, 2018).

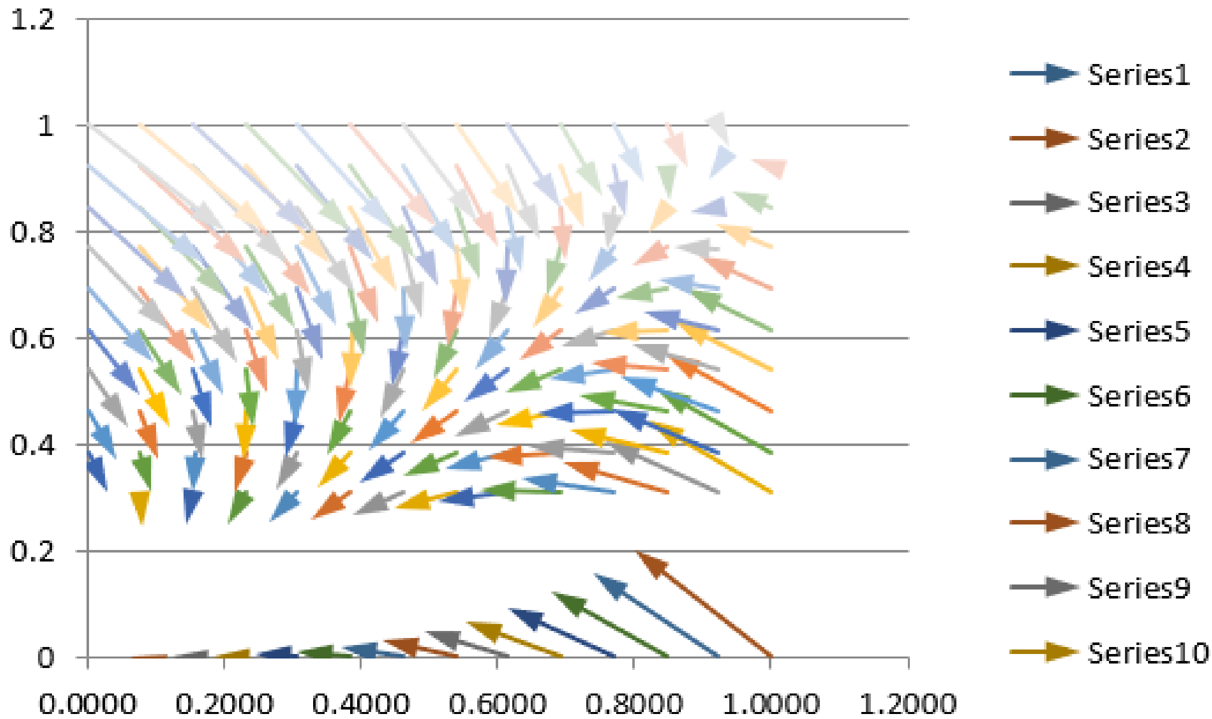
Now, let's try and see the behavior of the model with square root attachment functions ($V(y) = \sqrt{y}$ and $W(x) = \sqrt{x}$) by using the graph generated by the Excel spreadsheet. The following graph is given:



It appears that all points are converging to the point $(1, 1)$ from a visual inspection. It also seems that the vectors start to point towards the line $y = x$ like an asymptote, and once they get close enough they shoot off along that path until they reach their destination at $(1, 1)$. It is hard to see what the behavior will be in the bottom of the graph because of limitations in the Excel macros provided.

Using our previous techniques once again, we can confirm that $(1, 1)$ is a fixed point, but $(0, 0)$ is also a fixed point of the graph that was not apparent in the vector field from Excel. $(1, 1)$ is indeed an attractor, as expected, but the previous techniques of analysis for $(0, 0)$ will not work for this graph, as the derivatives of the attachment functions are not defined at 0. Other techniques, of which I do not know, would be necessary to determine the stability of the origin.

For the final model that will be discussed in this paper, we will once again look at the vector field of new attachment functions $V(y) = y^2$ and $W(x) = x^3$ visually, then move on to calculations. The following graph is generated by the Excel macro:



Looking ahead a little bit, it can be seen that these functions will have fixed points again at $(0,0)$ and $(1,1)$ clearly, as the attachment functions are positive powers of their independent variables. It appears that $(0,0)$ is definitely an attractor this time, since all the vectors in the vector field again converge at $y = x$, but this time travel slant asymptotically towards the origin instead of $(1,1)$. It looks like $(1,1)$ is an unstable saddle as well. These theories are both confirmed by calculating the eigenvalues for this system - $(0,0)$ has two repeated negative eigenvalues making it an attractor, and $(1,1)$ has one negative and one positive eigenvalue, making it a saddle. This means all curves originating within the unit square will terminate at $(0,0)$, meaning this model is one possible way that an internet platform can die out. Curves may approach up to the saddle at $(1,1)$, but they will eventually deflect and travel along $y = x$ down to the origin.

To conclude, the power of these models is in their simplicity to work with and their predictable behavior. In any situation where groups repel themselves linearly, not even constrained to online platforms, these models give simple to understand and fast to work with results. As this project unfolded, I found that points where I had to do extensive calculations early on became easier, quicker, and more reflexive once I understood what was happening with the model chosen and what behaviors to expect. The hardest part of constructing these models appears to be picking a well-behaved set of attachment functions which both reflect the collected data enough to be a good prediction of future trends, but also are limited to being between 0 and 1 for x and y values in that same range. One nice technique that I discovered with the hyperbolic trig model is, if a function starts at the origin and has a consistent upwards or downwards trajectory, but exceeds a maximum range of 1 for the input values, the function can be divided by its value at 1 to normalize it to usable values.

One of the biggest unfortunate circumstances with this project is the incredible difficulty of obtaining real-world data on the number of people in each of the different groups on groupable online platforms, or even usage statistics in general for sites which would fit the criteria for modeling using these techniques. If this data was more publicly available, it may have been possible to make basic predictions of trends in popular websites, which may make an interesting topic of study if this information ever becomes publicly available in the future. Overall, though, the differential equations were very enlightening, and the power of this model, and how much I learned about differential equations by working through it, will make a strong impact on my future as a mathematician and researcher.

References

Bloomenthal, A. (2022, December 28). *World's top 10 internet companies*. Investopedia. Retrieved April 13, 2023, from https://www.investopedia.com/articles/personal-finance/030415/worlds-top-10-internet-companies.asp

Lebl, J. (2022, November 22). *Notes on diffy qs: Differential equations for engineers*. Retrieved April 6, 2023, from https://www.jirka.org/diffyqs/html/frontmatter-1.html

Rayskin, V. (2018, October 29). *6-065-s-internetplatformusers*. SIMIODE. Retrieved April 6, 2023, from https://www.simiode.org/resources/5540

Woolf, P. (2023, March 11). *Chemical process dynamics and controls*. LibreTexts Engineering. Retrieved April 6, 2023, from https://eng.libretexts.org/Bookshelves/Industrial_and_Systems_Engineering/Chemical_Process_Dynamics_and_Controls_(Woolf)