

# Análisis de tendencias en función de la topología de la comunidad partícipe.

Erick Muñiz Morales.  
erickmuniz@ciencias.unam.mx

19 de enero de 2021

---

## Resumen

En este trabajo, se presenta una modelación de las tendencias (*trends*) la cual permite un estudio de las mismas en función de la comunidad partícipe. Para ello, se realiza un análisis de datos recabados sobre estas en un periodo de tiempo determinado usando herramientas de análisis de redes.

Los resultados muestran una aparente relación entre el coeficiente de agrupamiento y la entropía de Shannon sobre la distribución de grado en las redes estudiadas de comportamiento brusco. Asimismo, a través de un análisis estadístico, se corrobora la significancia de enlaces débiles entre los usuarios para ocurrencia de un comportamiento explosivo.

**Palabras clave:** Entropía de Shannon, fortaleza de enlace, *betweenness*, comportamiento explosivo.

---

## Índice

<b>1. Introducción.</b>	<b>2</b>
1.1. Objetivo . . . . .	2
<b>2. Marco teórico.</b>	<b>2</b>
2.1. Sistemas complejos y redes. . . . .	2
2.2. Entropía de Shannon. . . . .	3
2.3. Artículos relacionados y estado del arte. . . . .	3
<b>3. Metodología.</b>	<b>4</b>
3.1. Datos. . . . .	4
3.2. Presentación del modelo. . . . .	4
3.3. Métricas. . . . .	5
3.4. Análisis de cada $NC(h)$ . . . . .	6
3.5. Análisis de todas las $NC(h)$ . . . . .	6
<b>4. Resultados.</b>	<b>7</b>
4.1. Valores significativos en la entropía, <i>clustering</i> y <i>betweenness</i> . . . . .	7
4.2. Comunidades heterogéneas. . . . .	8
<b>5. Conclusiones.</b>	<b>8</b>
<b>6. Trabajo futuro.</b>	<b>13</b>
<b>7. Anexo</b>	<b>15</b>
7.1. ID de tendencias . . . . .	15
7.2. Matriz de correlación de la tabla $\Psi$ . . . . .	16
7.3. Tabla del análisis completo de una tendencia. . . . .	17
7.4. Tabla del análisis completo de tendencias. . . . .	18

---

# 1. Introducción.

La cuantificación del estado verdadero de la propagación de una enfermedad es un problema de índole complejo. Pues, a pesar de tener modelos analíticos como los modelos basados en las ecuaciones de Voltka-Volterra sobre dinámica de poblaciones, estos lo logran no captar las relaciones fundamentales entre los agentes; esto es, supone que los entes participantes en el sistema a estudiar, tienen la misma relevancia.

Dado que estas ecuaciones reflejan la idea de cambios de estado entre los agentes, podemos extrapolar dicha motion para el caso de intercambio de ideas y opiniones. Entre estos sistemas, las redes de media social son un buen medio para ver esta dinámica entre comunicantes debido a la sencillez de conectividad y reacción inmediata ante susceso espontáneos.

Ejemplo de estas, quizá las más llamativa y conocida, es la plataforma *Twitter*. Dicha plataforma cuenta con 330 millones de usuarios activos mensualmente<sup>1</sup>. En esta red de media social, podemos encontrar cuatro sencillas acciones entre usuarios:

- ***Tweet.***

Es la acción de un usuario por publicar una idea u opinión. Se puede presentar como un texto plano o con añadidos como *emojis*, *emoticons* o archivos audiovisuales.

- ***Follow.***

Es la acción de un usuario *A* para seguir a un usuario *B*. Esta es la primera relación entre los usuarios. *A priori*, es una manera sencilla de representar gustos afines entre los usuarios. Claramente, cuando hay seguimiento mutuo (siguiendo de nuestro ejemplo, cuando *B* sigue a *A* también), entonces esta relación se vuelve más fuerte.

- ***Retweet.***

Es la acción de un usuario *A* por 'reenviar' un *tweet* que ya existe de un usuario *B* y publicarlo en el perfil de *A* con el nombre de *B*. Es, en términos sencillos, citar la idea u opinión de un usuario para apoyarla.

- ***Mention.***

Es la acción de un usuario *A* por *tweetar* algún contenido etiquetando o mencionando a otro usuario *B*. Es un comunicación directa de *A* a *B*.

Además, por la propia dinámica que se ha generado, se ha hecho de un espacio común de choque político, social y cultural. En consecuencia, esta comunicación no tangible ha generado tanto eventos no deseables<sup>2</sup> y sucesos beneficiosos para la comunidad<sup>3</sup>. Esto nos lleva a al análisis de **tendencias** (*trends*). En esencia, una tendencia, es

un tema, idea o suceso que se es muy discutido entre los usuarios. Puede ser distinguido de muchas formas, como *memes* o videos virales.

Bajo esta posible problemática, dado que desconocemos el posible estado futuro de alguna idea, dicha plataforma notifica a sus usuarios de las medidas y restricciones que se toman a través de su blog de compañía<sup>4</sup>, las cuales tratan de apaciguar posibles actos no deseables.

En el presente trabajo, se aborda un análisis de la dinámica de difusión de información, como un símil a la propagación de una enfermedad, pero en función de la estructura de las relaciones entre los comunicantes; en específico, en aquellas donde ocurre una gran cantidad de difusión de información en lapsos de tiempo muy cortos. A esto último, lo llamaremos como **comportamiento explosivo**. Se realiza un modelo con base en la literatura actual [[17, 7, 22, 8]] sobre análisis de redes; así como algunas métricas sobre el orden del estado de un sistema y un sencillo análisis estadístico sobre variables convenientes.

Dicho análisis es realizado con una base de datos recabada en el portal *Twitter* desde el 24 de marzo de 2012 hasta el 26 de abril de 2012. Esta base de datos se puede encontrar en el portal oficial del Centro de Investigación de Redes y Sistemas Complejos (CNetS).

## 1.1. Objetivo

Obtener una caracterización de la comunidad comunicante en las tendencias de un comportamiento brusco o explosivo usando herramientas de redes complejas y teoría de la información.

# 2. Marco teórico.

Por lo antes mencionado, es posible una modelación usando herramientas de redes complejas; esto es, obteniendo las relaciones entre los comunicantes. En esta sección, se desglosará estas ideas así como la fundamentación para necesaria sobre por qué este es un sistema complejo.

## 2.1. Sistemas complejos y redes.

Los sistemas complejos estudian la forma de los componentes y su interacción entre sí; '*pueden espontáneamente auto-organizarse y presentar estructuras globales y comportamientos no-triviales a mayores escalas, sin intervención externa, autoridad central o líderes que determinen el comportamiento colectivo*'[[5]]. Si bien esto no es una definición formal, capta la importancia del estudio en estos sistemas: las relaciones.

Se han detectado propiedades en común sobre dichos sistemas. Como en el caso de su definición, puede variar

<sup>1</sup><https://www.oberlo.com/blog/twitter-statistics>

<sup>2</sup>Hudson, John. "The Most Infamous Terrorists on Twitter.", The Atlantic, 2 Jan. 2012, Link.

<sup>3</sup>Santolini, Marc. "Covid-19: The Rise of a Global Collective Intelligence?" The Conversation, Link. Accessed 17 Jan. 2021.

<sup>4</sup>Blog oficial de *Twitter*.

dependiendo del autor y el sistema a estudiar. Las propiedades que se consideran más importantes son tres: *emergencia*, *autoorganización*, *independencia* e *interdependencia*.

Las últimas dos propiedades muestran la robustez del sistema ante perturbaciones del mismo. En efecto, notaremos que existen relaciones clave en un sistema que, en la ausencia de estas, pueden colapsar todo el sistema. Así mismo, también encontraremos relaciones fuertes en ese sentido que, en la ausencia de estas, el sistema no sufrirá daño.

La *autoorganización* y la propiedad emergente destacan por la idea de cambio por la estabilidad del sistema aún cuando no exista un líder o reglas formamente escritas o explícitas. Pues, cada componente del sistema puede actuar por su beneficio propio autoorganizándose y, a nivel global, generar un patrón de cambio repentido o un estado emergente.

Con base en estas ideas, la teoría de gráficas es un buen elemento matemático para el análisis de este tipo de sistemas [13]. En efecto, dado que existe una importancia a las relaciones entre los componentes del sistema, nace una nueva teoría que abarca esta necesidad: Teoría de redes.

Una **red** es una gráfica con interpretación. Aunque se conozcan métricas generales para tratar cualquier tipo de red, lo importante es la interpretación de las mismas. De estas métricas destacan métricas de centralidad local (como el coeficiente de agrupamiento *clustering*) y métricas globales que abstraen la topología de una red (como el coeficiente de assortatividad) [[15]].

## 2.2. Entropía de Shannon.

En teoría de la información, se estudia como los datos son comprimidos o transmitidos. Ocasionalmente, esta rama del estudio se considera un subconjunto de la teoría de la comunicación [[4]].

Con esta idea, nacen métricas con una intuición sencilla pero aplicaciones complejas. La entropía, en este campo, se define como un promedio de la incertidumbre del verdadero estado de un fenómeno. Para una variable aleatoria  $X$ , con función de distribución de probabilidad  $f(x)$ , la entropía de Shannon se define como

$$H(X) = - \sum_x f(x) \log_2(p(x))$$

El valor resultante, para el caso de  $\log_2$ , se llaman *bits* de información.

La interpretación de esta métrica es sencilla: ¿Cuántas preguntas con respuesta binaria se necesitan para saber el estado verdadero del sistema? En consecuencia, cuando  $X$  tiene una sesgada para cierto valores, la entropía es baja; ya que, el valor de  $X$  es más probable que esté en la parte sesgada.

Notemos que se puede cambiar la base del logaritmo en esta métrica, obteniendo posibles resultados distintos (véase [4] para algunas propiedades). En este trabajo, únicamente se usará la entropía de Shannon con  $\log_2$ .

## 2.3. Artículos relacionados y estado del arte.

La literatura muestra que los primeros estudios sobre la dinámica de poblaciones se basan en sistemas dinámicos no lineales. Dichos modelos se fundamentan en la idea de estados; en el ejemplo de una enfermedad, los estados son *susceptibles*, *infectados* y *recuperados* [[2]]; en el ejemplo de migración, representan localidades, estados o naciones [[6, 11, 16]]. Notemos que el término 'no lineal' realza la necesidad de que dichos estados tengan cierta relación no intuitiva. Sin embargo, sobre estos estados se asume que las relaciones internas (entre los agentes del sistema) se puede hacer homogénea.

Así, es clara la similitud entre estos modelos y la dinámica de las redes de media social considerando a los estados como estados binarios (han comentado o no de cierto tema). Este pensar ha llevado a algunos realizar distintos análisis estadísticos [[1, 23]].

Si bien, susodichas modelaciones han logrado obtener propiedades asintóticas interesantes a través del análisis de bifurcaciones [[6]]; estos desafortunadamente fallan cuando se aplica un mismo modelo a otra localidad con aparentes similitudes. Aquí, un supuesto fuerte es olvidado; en particular, no todas las comunidades tienen el mismo comportamiento. Esto, en consecuencia de diversos factores demográficos, económicos, políticos y sociales.

Recuperar esta información menospreciada es sencillo con un enfoque de sistemas complejos usando redes. Pues, por lo antes discutido, estos sistemas son complejos; ya que, los usuarios o agentes no tienen un líder que guie a los demás. Aunado a ello, sin importar el sistema a estudiar, cada agente tiene la necesidad de ver por su bien con base en la información cercana (en el caso de migración, cada persona emigrará donde pueda obtener mejor calidad de vida; en el caso de *Twitter*, cada usuario sigue a quien desee seguir). Por lo que, a partir de esta 'reglas' sencillas individuales también se generan propiedades emergentes ('el error de diciembre' generó una enorme cantidad de emigración hacia Estados Unidos. En *Twitter*, puede representar a tendencias con un comportamiento explosivo). Por lo que, es válido aplicar la metodología de sistemas complejos para el análisis de transmisión de ideas.

La literatura dedicada a esto último, engloba varios enfoques sobre la topología de la red generada por los comunicantes. Primeramente, de los artículos de Strogatz y Watts, así como el artículo de Newman y Girvan, nos presentan modelos teóricos que abarcan un coeficiente de agrupamiento necesario para obtener una transmisión efectiva[20, 14] con menor costo de relaciones; aquí obtenemos un umbral para el agrupamiento de los comunicantes. Por lo anteriores artículos, nacen intuiciones sobre la significancia de este umbral, de otro modo la existencia de relaciones o aristas débiles en la red. En consecuencia en [22, 21, 10] abordar esta cuestión mostrando cierta necesidad de la red con un comportamiento disortativo o una distribución heterogénea para el caso de un comportamiento explosivo. Aunado a esto último, también es

necesaria la interacción de costo mínimo (*retweets*) para que el comportamiento sea explosivo. [[7, 8]].

### 3. Metodología.

Con base en lo visto anteriormente, mostraremos la siguiente metodología. Cabe resaltar que la implementación de la misma fue realizada en Python con la paquetería NetworkX en la plataforma Google Colab. Para el análisis y visualización de los datos se usó cierto código en R.

#### 3.1. Datos.

La base de datos para este trabajo, como se mencionó, fue recabada por recabada por el Centro de Investigación de Redes y Sistemas Complejos (CNetS).

Dicha base cuenta con, aproximadamente, 500 MB de información. La información es un compendio de varios seguimientos del portal *Twitter* desde el 24 de marzo de 2012 a el 26 de abril de 2012. Cabe recalcar que, por las políticas de privacidad de *Twitter*, todos los usuarios que aparecen tienen un identificador único y diferente a su cuenta; por lo que, no es posible hacer búsqueda de dicha cuenta a través de otras paqueterías u obtener una información sobre los usuarios. Hay **14,273,311** usuarios recabados con **1,345,913** de tendencias y con **12,054,205 tweets** en total.

Para la modelación, se realizó una limpieza de toda la base. El primer filtro, el primer filtro fue considerar sólo aquellas que tuvieran, al menos 4000 *tweets* registrados. Así, se procede a analizar únicamente **150** tendencias.

#### 3.2. Presentación del modelo.

Con base en los modelos presentados en [17, 7, 8, 22], presento el siguiente modelo que abarca varias necesida-

des. Esta modelación se basa en redes temporales, donde cada tiempo  $t$  representa un periodo de tiempo.

Dejemos fija una *tendencia*  $h$  y sea  $t$  un periodo de tiempo. Sea  $NC_t^\alpha(h)$  una red de multicapa donde  $\alpha \in \{T, R, M\}$  una para cada una de las actividades de *tweet*, *mention* y *retweet* para el periodo  $t$ . Cada una está definida de la siguiente manera.

- $NC_t^T(h)$  es una red **no dirigida** que representa la información de *tweets* y los usuarios que se siguen mutuamente.

Esto es, supongamos que A realiza un *tweet*. Y en la base de datos tenemos que B sigue a A y A sigue a B, entonces la relación existe; es un seguimiento mutuo. Cabe recalcar que basta con que uno de los dos haya hecho una interacción (*tweet*) para que se genere esta relación en nuestra red.

- $NC_t^R(h)$  es una red **dirigida** para los *retweets*.

Es decir, supongamos que A realiza un *tweet* y, después, un usuario B realiza un *retweet* al *tweet* de A; entonces, la relación entre  $B \rightarrow A$  existe.

- $NC_t^M(h)$  es una red **dirigida** con pesos asociada para las *mentions* (menciones).

Esto es, supongamos que A realiza un *tweet* mencionando al usuario B, entonces la relación  $A \rightarrow B$  está definida.

En la figura 1 podemos ver una pequeña representación de este modelo para una tendencia  $h$  fija. Haremos referencia a  $NC(h) = \cup_t NC_t(h)$  como la dinámica de la tendencia  $h$ .

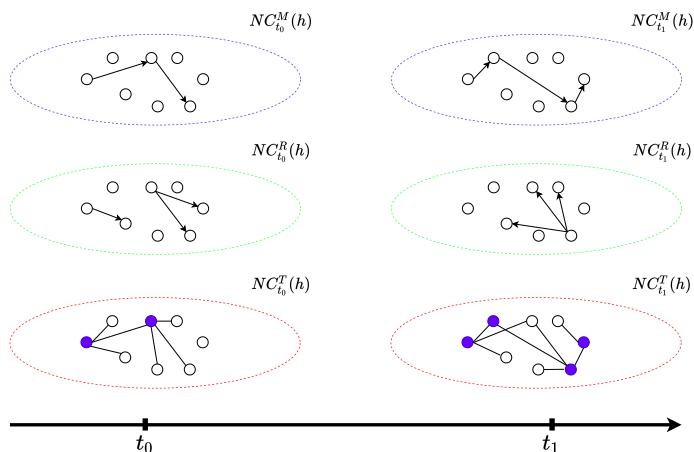


Figura 1: Representación de la red  $NC(h)$ . Los nodos con color morado son aquellos quienes *tweetearon*.

### 3.3. Métricas.

Primeramente, definimos la propiedad de **comportamiento explosivo** como aquella tendencia donde en algún momento registró una cantidad inmensa de *tweets*. En la figura 2, podemos ver aquellas tendencias con comportamiento explosivo.

Por la muestra de 150 tendencias, se considera de comportamiento explosivo aquellas donde reportaron más de 100 *tweets* en una hora.

Así, mostremos la siguientes métricas para el análisis de este modelo. Para ello, empecemos con algunas métricas locales, que se usarán para dicho análisis

- *Betweenness (B)*, sirve para cuantificar la relación de todos los caminos entre dos nodos y aquellos que pasan por un nodo  $i$ , ésto ayuda a identificar nodos relevantes o relaciones débiles.
- *Clustering (C)*, es el coeficiente de agrupamiento; podemos asociarlo con los triángulos o las comunidades ternas de usuarios con varias características en común.

Para hacer un análisis global de la red, he elegido usar las siguientes métricas:

- Entropía de Shannon de la distribución de grado ( $S$ ). Sea  $G$  una gráfica con  $N$  nodos, y  $P(k)$  su distribución de grado. Entonces, la entropía de una gráfica se define como

$$S = - \sum_{k \in N} \frac{P(k)}{N} \log_2 \left( \frac{P(k)}{N} \right), \quad (1)$$

Esta métrica nos permite abstraer el desorden de la difusión de la información vista en función de la red social generada por los usuarios que *tweetearon* ( $NC_t^T(h)$ ).

- El diámetro más grande de las componentes conexas ( $\delta$ ).

Esta métrica representa la cadena más grande de comunicación en la red. Esta métrica será aplicada únicamente en las redes de *retweets* y *mentions* pero en su forma no dirigida.

Por ejemplo, de la figura 1 obtenemos que  $\delta[NC_{t_0}^R(h)] = 2$ ,  $\delta[NC_{t_0}^M(h)] = 2$ ; mientras que para su siguiente paso tenemos que  $\delta[NC_{t_1}^R(h)] = 2$ ,  $\delta[NC_{t_1}^M(h)] = 3$ . De aquí, podemos ver que tiene más significancia para el caso de *mentions* ya que, por su naturaleza, son interacciones directas.

- El grado exterior más grande de la red ( $\Gamma$ ).

Esta métrica representa la cantidad más grande de personas que *rtweetearon*; esto es, la cantidad más grande de personas que simpatizaron con una idea. Pues, por la naturaleza de los *rtweets*, es la manera más fácil y sencilla de mostrar *apoyo* con dicha idea.

Por ejemplo, de la figura 1 obtenemos que  $\Gamma[NC_{t_0}^R(h)] = 2$ ,  $\Gamma[NC_{t_0}^M(h)] = 1$ ; mientras que para su siguiente paso tenemos que  $\Gamma[NC_{t_1}^R(h)] = 3$ ,  $\Gamma[NC_{t_1}^M(h)] = 1$ . De aquí, podemos ver que tiene más significancia para el caso de *retweets*.

- Fortaleza de enlace ( $SE$ ).

Esta métrica es muy particular en este tipo de análisis con enfoque social; específicamente, para un análisis local en las relaciones entre usuarios. Dicha métrica es definida como fortaleza de arista [[12]] y es la razón de las personas que se conocen en común por la razón de las personas que se conocen. Formalmente, para la arista  $(i, j)$  tenemos

$$O_{ij} = \frac{n_{ij}}{(k_i - 1) + (k_j - 1) - n_{ij}}, \quad (2)$$

donde  $n_{ij}$  respresenta el número de vecinos comunes entre  $i$  y  $j$  y  $k_i$  es el grado del nodo  $i$ . En este trabajo, haremos uso del promedio de dicha métrica.

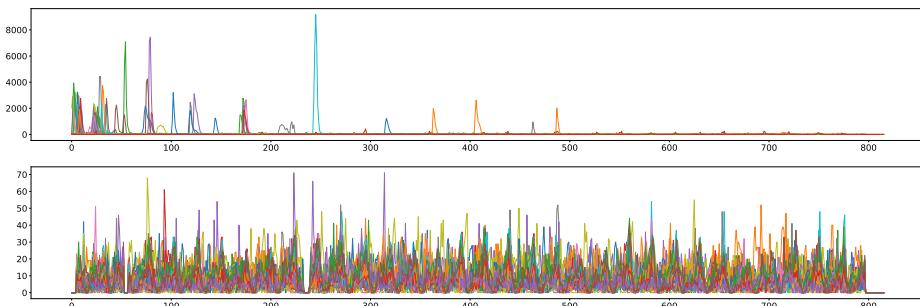


Figura 2: Serie de tiempo de la cantidad de *tweets* por hora. En el cuadro superior son de tipo *explosivo*. Notemos como estas tienen un comportamiento repentino; mientras que las inferiores tienen un comportamiento, aparentemente, periódico.

### 3.4. Análisis de cada $NC(h)$ .

El análisis de cada una de las  $NC(h)$  se realiza a través del análisis de cada una de las redes en cada periodo de tiempo  $t$ ; esto es, analizando  $NC_t^T(h)$ ,  $NC_t^R(h)$  y  $NC_t^M(h)$  de manera conveniente.

Usando métricas de centralidad local y global, aunado con las antes mencionadas, es cómo se harán los cálculos. Por lo que, esto nos genera, por cada tendencia, una tabla de datos como en el recuadro 2, donde cada una de las columnas representan lo siguiente,

- $D \rightarrow$  Día del año.
- $H \rightarrow$  Hora del día.
- $N_T \rightarrow$  Cantidad de *tweets* resgitrados en ese periodo de tiempo.
- $N_R \rightarrow$  Cantidad de *retweets* resgitrados en ese periodo de tiempo.
- $N_M \rightarrow$  Cantidad de *mentions* resgitrados en ese periodo de tiempo.
- $k_T \rightarrow$  Grado promedio de la red  $NC_t^T(h)$ .
- $L_T \rightarrow$  Promedio de la longitud de ruta más corta entre dos nodos de la red  $NC_t^T(h)$ .
- $C_T \rightarrow$  Promedio de *clustering* de la red  $NC_t^T(h)$ .
- $B_T \rightarrow$  Promedio de *betweenness* de la red  $NC_t^T(h)$ .
- $S_T \rightarrow$  Entropía de Shannon de la distribución de grado de la red  $NC_t^T(h)$ .
- $\delta_R \rightarrow$  El diámetro más grande de las componentes conexas de la red no dirigida de  $NC_t^R(h)$ .
- $\Gamma_R \rightarrow$  El grado exterior más grande de la red  $NC_t^R(h)$ .
- $\delta_M \rightarrow$  El diámetro más grande de las componentes conexas de la red no dirigida de  $NC_t^M(h)$ .
- $\Gamma_M \rightarrow$  El grado exterior más grande de la red  $NC_t^M(h)$ .

### 3.5. Análisis de todas las $NC(h)$ .

Este análisis se realiza con el fin de poder encontrar alguna caracterización en aquellas tendencias que tienen un comportamiento de tipo *explosivo*. Para ello, se procede a realizar una serie de manipulaciones en los datos para obtener otra tabla con la significancia de algunos recabados en el paso anterior. Convenientemente, definamos el operador  $\langle \Delta \rangle$  como el promedio de los valores de la columna  $\Delta$  y a  $\max(\Delta)$  como el valor máximo de los valores de la columna  $\Delta$ .

Lo que procede es obtener valores que puedan tener mayor significancia en su dinámica a partir de los valores de su tabla generada en el paso anterior. Lo resultante es una tabla de las 150 tendencias con las siguientes columnas.

- $\langle N_T \rangle \rightarrow$  Esto nos representa la actividad de *tweets* promedio que ocurrieron en cada uno de los periodos de tiempo; será más significativa cuando sean de tipo *explosivo*.
- $\langle N_R \rangle \rightarrow$  Esto nos representa la actividad de *retweets* promedio que ocurrieron en cada uno de los periodos de tiempo.
- $\langle N_M \rangle \rightarrow$  Esto nos representa la actividad de *mentions* promedio que ocurrieron en cada uno de los periodos de tiempo.
- $\max(k_T) \rightarrow$  La consideración del máximo fue para tratar de recuperar aquel tiempo donde algún usuario de grado muy alto puede ser fundamental para la ocurrencia del comportamiento explosivo.
- $\max(L_T) \rightarrow$  La consideración del máximo fue para tratar de recuperar aquel tiempo donde algún usuario de grado muy alto puede ser fundamental para la ocurrencia del comportamiento explosivo.
- $\langle C_T \rangle \rightarrow$  Esto puede abstraer el promedio de concentración de la sociedad a través del tiempo.
- $\langle B_T \rangle \rightarrow$  Esto puede abstraer la idea de encontrar aristas de mayor importancia que otras (enlaces fuertes). Si se encuentran valores no despreciables, es porque hubo un usuario (no necesariamente de grado alto) tal que logró comunicar con más personas.
- $\langle S_T \rangle \rightarrow$  Esto puede asbtraer la idea de difusión de la información entre los agentes de la comunicación. Pues, si encontramos valores altos, la media nos puede reflejar esta existencia y, consecuentemente, el desorden de la difusión de información es significativo.
- $\max(\delta_R) \rightarrow$  Esta idea puede recuperar la comunicación más larga en todo el tiempo del análisis de la tendencia.
- $\max(\Gamma_R) \rightarrow$  Esto puede abstraer como la cantidad máxima de usuarios que son influidos por la lectura de un *tweet* y lo secundaron con un *retweet*; pues, es la sencilla idea de las acciones con menor costo pueden influir en la propiedad de explosividad.
- $\max(\delta_M) \rightarrow$  Análogo a  $\max(\delta_R)$  pero con una significancia mayor ya que estas son las actividades que se requiere un mayor esfuerzo.
- $\max(\Gamma_M) \rightarrow$  Análogo a  $\max(\Gamma_R)$  pero con una significancia mayor ya que estas son las actividades que se requiere un mayor esfuerzo.
- $r \rightarrow$  Esta última es el grado de assortatividad de la red social generada por **todos** los usuarios que realizaron algún *tweet* sobre dicha tendencia. Esta última métrica tiene un motivo de 'condición inicial' sobre cómo es la topología de la red antes de analizar la dinámica aunado a la hipótesis de aristas débiles.

- *Burst* → Es nuestra variable indicadora si la tendencia tiene un comportamiento explosivo.

El resultado de esto es una tabla como la que se puede apreciar en la tabla 2. De esta última tabla es la que se realizan los siguientes análisis sobre valores característicos. Por lo tanto, me voy a referir a dicha tabla como  $\Psi$ .

## 4. Resultados.

Con base en las redes generadas  $NC(h)$ , se realiza un análisis estadístico sobre variables convencionales y descriptas en las métricas antes vistas. Así mismo, este análisis se ha llevado a un nivel global considerando toda la dinámica y un nivel local en comunidades.

### 4.1. Valores significativos en la entropía, clustering y betweenness.

La distribución de grado de una red nos brinda un análisis sencillo en torno a la topología de la misma. Si bien no es análisis local y preciso, si nos permite un primer acercamiento a un análisis global.

Considerar la distribución de grado como un probabilidad, nos permite el uso de la entropía de Shannon. El estudio de esta métrica ha sido corto y aplicado en muy densas [[18]] o en redes de libre escala [[19]]. En particular, de dichos estudios, muestra la relación de esta métrica con la distribución heterogénea de la distribución de grado y robustez para redes de libre escala.

De la misma expresión (1), nos indica que son valores positivos y sólo hace referencia a las probabilidades. Así, si nuestra variable aleatoria a investigar tiene un soporte más grande en longitud, entonces el valor de la entropía aumenta. Con esto en mente, es esperable que para redes con una distribución de grado con soporte grande en longitud obtengan valores altos; pues tenemos usuarios que están en varias comunidades y funcionan como puentes.

En las figuras 3 y 4, podemos ver una relación entre el comportamiento explosivo y el valor de su entropía. Si bien, tenemos datos atípicos, las tendencias con comportamiento no explosivo tienen un valor mínimo (0.25 aprox.).

Así mismo, podemos denotar relaciones interesantes entre las métricas *betweenness* y la entropía. Al considerar valor de *betweenness* no despreciables, la distribución de grado es más heterogénea. Sin embargo, por la figura 3, las muestras dan indicios a que la configuración entre los usuarios es más selectiva para generar este comportamiento explosivo. Esto último puede ser consecuencia del tipo de *tendencia* o tema que se requiera discutir; del artículo [[9]] nos dan un pequeño análisis de comunidades sobre *subreddits* (comunidades más pequeñas) mostrando que aquellas con pocas personas pero con una organización clave entre las demás, pueden generar enfrentamientos masivos; misma idea fue hecha [[3]], sobre discusiones políticas en *Twitter*. Así mismo, notemos que de la figura 4, se necesita un valor del clustering no despreciable o que en las comunidades sea considerable.

Entonces, conforme a nuestra definición de entropía de Shannon, necesitamos cierto 'desorden' entre los usuarios para generar el comportamiento explosivo. Por desorden, necesitamos una cantidad suficiente de usuarios que no sigan a usuarios del grado o misma comunidad: que sean puentes entre las comunidades. Por lo que, en ese sentido, dicho desorden o comportamiento no debe sobrepasar una entropía de 0.25.

Por otro lado, notemos que también debe existir la necesidad de acciones de menor costo. La dinámica de *Twitter* es una comunicación sencilla y simple. La acción *retweet* es la de menor costo y la que muestra mayor motivación ante una situación. De la figura 4, podemos ver como las tendencias de comportamiento explosivo muestran que los *retweets* son clave para la realización de una tendencia de comportamiento explosivo.

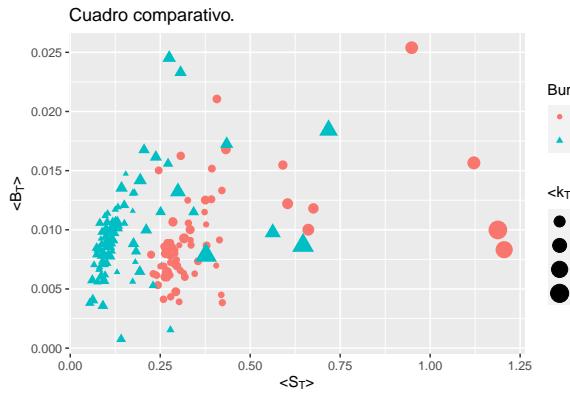


Figura 3: Cuadro comparativo entre el *betweenness* promedio y la entropía promedio. La tendencia es de comportamiento explosivo si *Burst* es igual a 1. Para una mejor visualización, la tendencia de comportamiento explosivo son las de color verde con forma de triángulo.

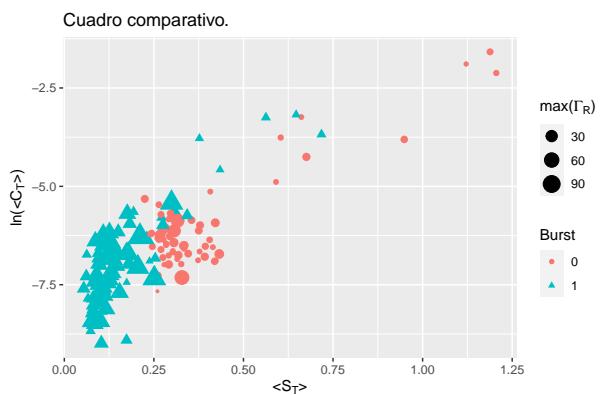


Figura 4: Cuadro comparativo entre el clustering promedio y la entropía promedio. Notemos que se aplicó una transformación para  $< C_T >$  para denotar un poco mejor esta relación.

Para un mejor argumentación de estas cuestiones, se recurre a realizar una regresión logística sobre las variables  $k_T$ ,  $< C_T >$ ,  $< B_T >$ ,  $< S_T >$ ,  $\max(\delta_R)$ ,  $\max(\gamma_R)$ ,  $\max(\delta_M)$ ,  $\max(\Gamma_M)$  y  $r$  con variable respuesta *Burst*. De

este resultado [ $R^2 = 0,87$ ], se encuentra mayor significancia para las variables  $\langle C_T \rangle$ ,  $\langle B_T \rangle$  y  $\langle S_T \rangle$ . Donde, el valor de  $\langle S_T \rangle$  predomina al tener el coeficiente más grande.

## 4.2. Comunidades heterogéneas.

El análisis anterior, nos induce a preguntarnos sobre cómo es la comunidad partícipe en el momento clave de la dinámica de una tendencia. Para este análisis, usaremos las redes  $NC_{t^*}(h)$  donde  $t^*$  es el periodo de tiempo en el cual se reportaron más tweets: el momento clave de las tendencias de comportamiento explosivo.

En este caso, obtendremos el grado de assortatividad  $r$  y la fortaleza de arista de la red  $NC_{t^*}^T(h)$ . Esta métricas son para denotar en mayor medida, lo discutido anteriormente sobre distribuciones de grado heterogéneas. Notemos que el valor de  $r$  nos va a denotar con la existencia de usuarios puente cuando  $r$  sea negativo. De la figura 5, notemos el sesgo de los valores de  $r$  en las tendencias de comportamiento explosivo. Del gráfico, podemos denotar que la estructura de la red, debe ser levemente dissortativa; lo cual, nos resalta la necesidad de estos nodos o usuarios puente.

Por otro lado, del gráfico 6, notemos que también necesitamos agrupamiento. La métrica  $SE$  (2), es la sencilla razón de personas en común entre las personas que se conocen. Por lo que, similar al *clustering*, nos permite abarcar la idea comunidades; a mayor valor de  $SE$ , es más probable que sean de la misma comunidad. Sin embargo, notemos los rangos de esta variable, los cuales son muy pequeños; por lo que, sólo es una cantidad muy pequeña de agrupamiento.

## 5. Conclusiones.

Las tendencias con un comportamiento explosivo tienen una conducta local internamente fuerte en el sentido de agrupación pero débil entre comunidades.

En este trabajo, se concluye que este comportamiento explosivo queda definido a través de la topología de la red generada en los comunicantes. Específicamente, se obtiene una caracterización determinada por la entropía de Shannon de la distribución de grado de la red a estudiar mostrando un valor máximo para un comportamiento explosivo. Esto, en consecuencia del análisis de otras métricas como fortaleza de enlace la cual abstrae el concepto de homofilia presente de una red social y métodos estadísticos. Con base en esta metodología, el objetivo se ha cumplido.

Si bien estos datos son esperanzadores, notemos que el año recabado es bastante tardío para poder generalizar a dinámicas actuales; pues la tecnología ha avanzado considerablemente, así como la sociedad ha modificado las políticas permitiendo una alteración a la naturaleza de los datos. Además, la muestra puede ser considerablemente pequeña para un análisis general. El análisis pudiera carecer de otra información relevante como geolocalización fortaleciendo más las relaciones. Sin embargo, el enfoque dado es suficiente para recabar ideas esenciales de comunicación.

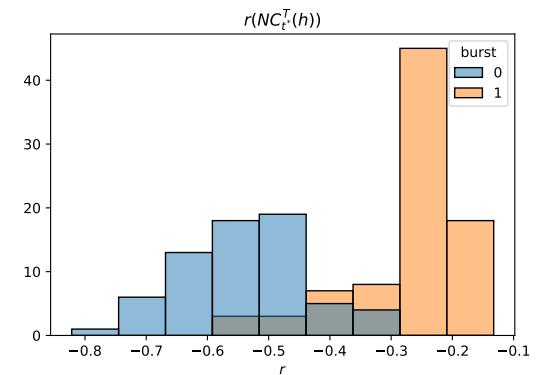


Figura 5: Histograma del valor  $r$  por su tipo de comportamiento. Notemos como existe un sesgo significativo que caracterizan a las tendencias.

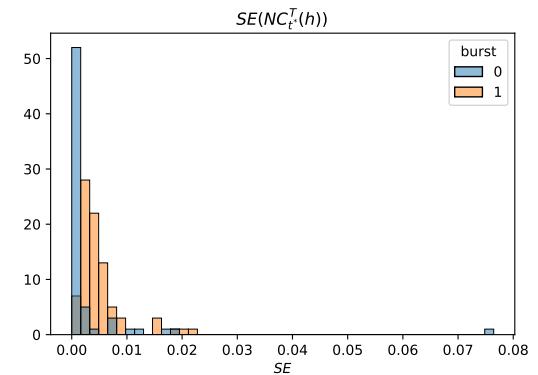


Figura 6: Histograma del valor  $SE$  por su tipo de comportamiento. Si bien no existe un sesgo significativo, notemos que los valores esperados son mayores en las tendencias con comportamiento explosivo.

Los siguientes gráficos 8, 7 muestran mejor estas ideas de manera visual.

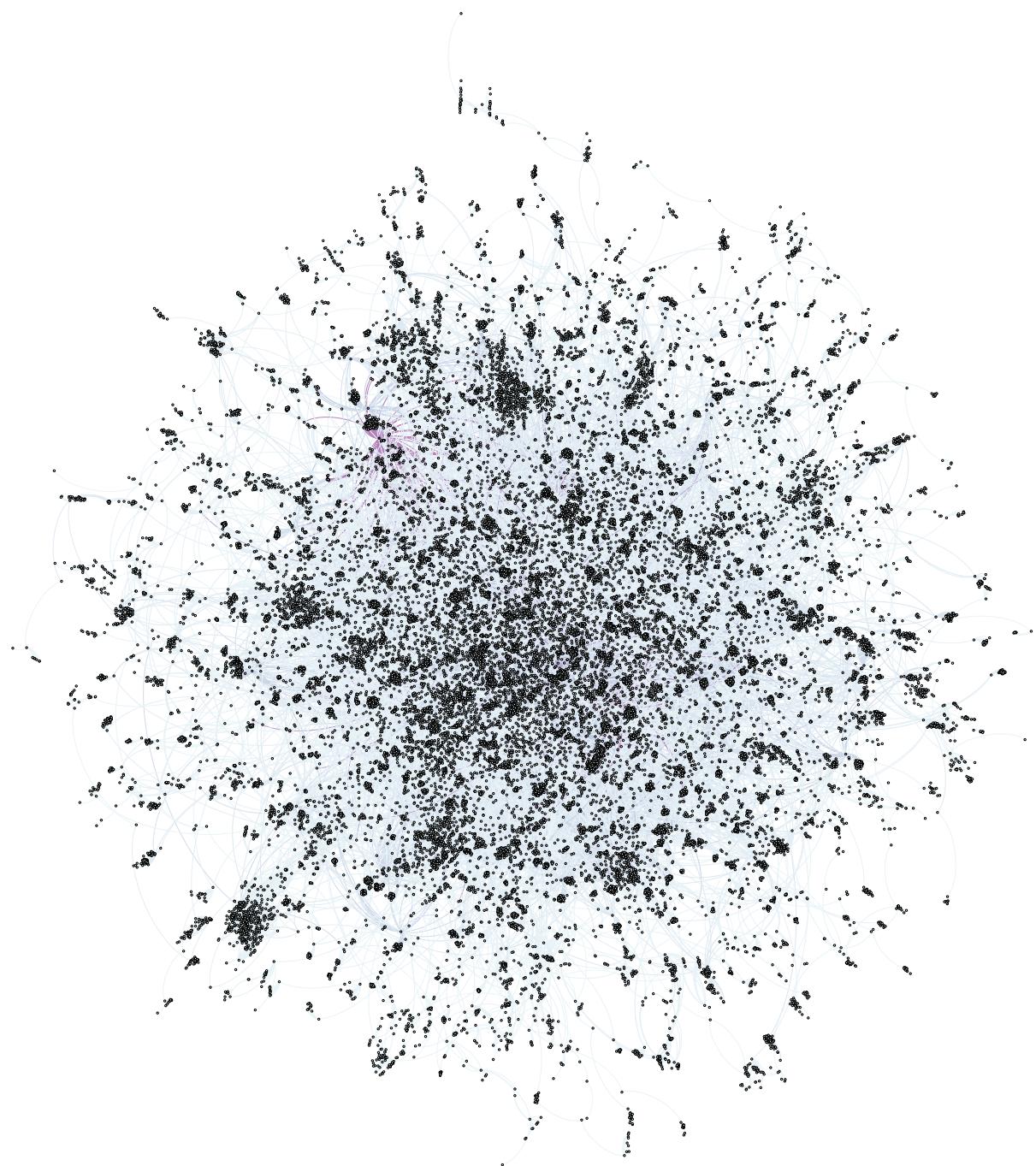


Figura 7: Componente gigante de la red social generada en el momento  $t^*$  de la tendencia *wewontworkoutif*. Fue de las tendencias que tiene más *tweets* en toda la dinámica. Notemos cómo hay demasiadas comunidades, lo cual realza la cantidad de usuarios que sirven como puente.

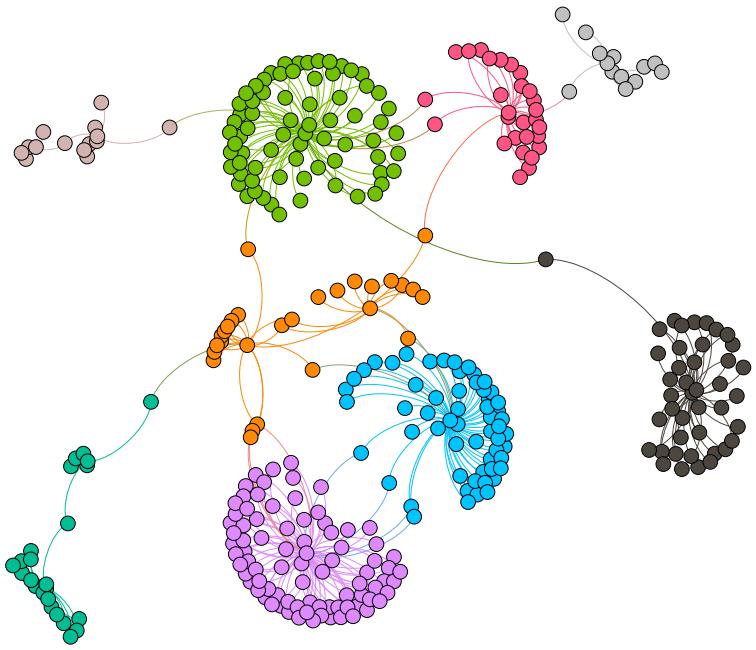


Figura 8: Componente de la red social generada en el momento  $t^*$  de la tendencia *thehairoscope*. Fue de las tendencias que menos *tweets* recabó. Las comunidades son resaltadas en colores. Podemos ver de manera más clara cómo es la dinámica de una tendencia con comportamiento explosivo; ya que podemos ver a comunidades con *clustering* no despreciables pero con usuarios puente.

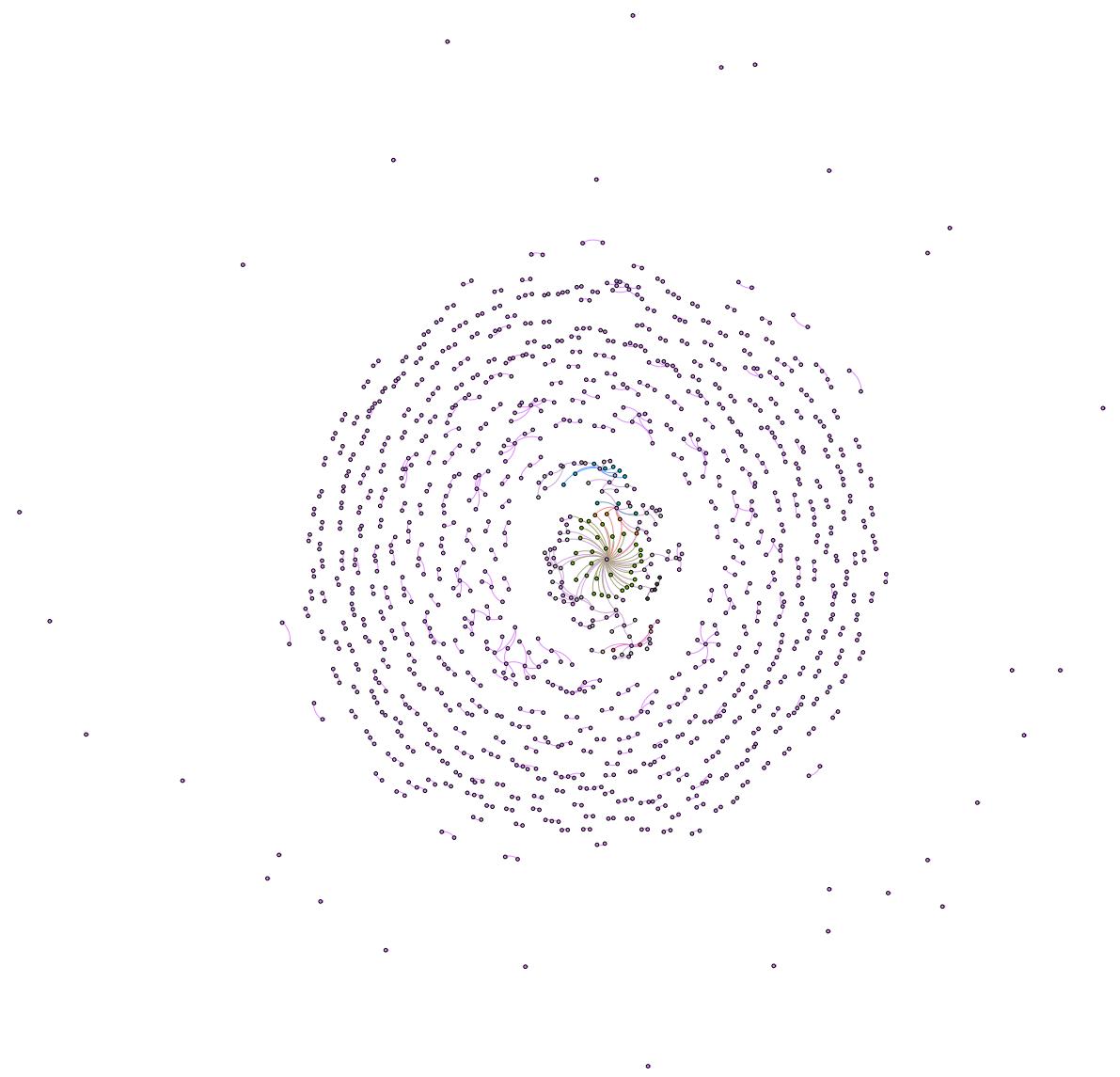


Figura 9: Red  $NC^M(h)$  de la tendencia  $h = 10$  people on twitter i want to meet. Notemos la dinámica de la comunicación interna entre los comunicantes. Por el tipo de tendencia, se esperaba que la comunicación fuera directa.

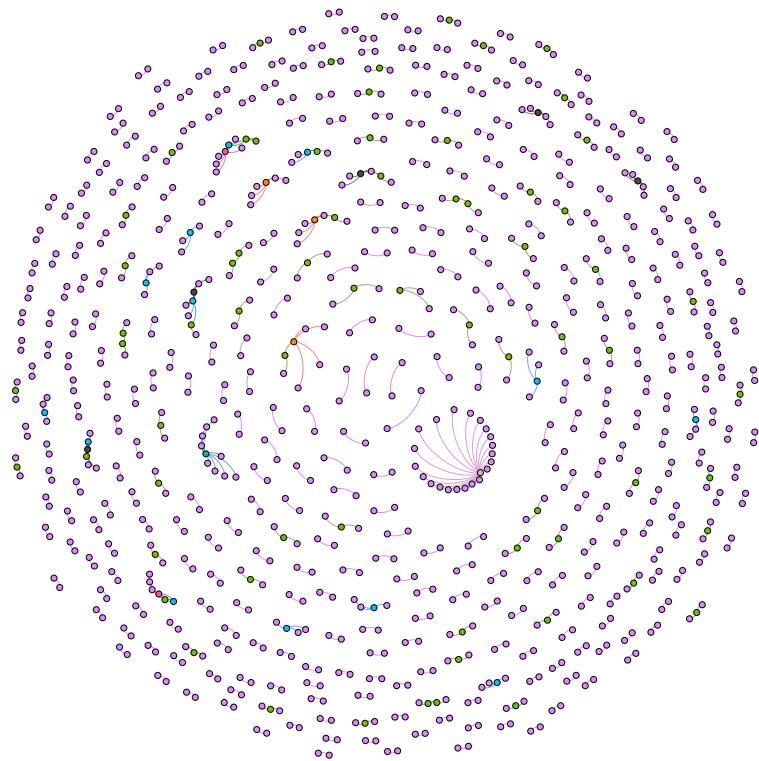


Figura 10: Red  $NC^R(h)$  de la tendencia  $h = wewontworkoutif$ . Notemos que tiene una dinámica similar a la vista en el gráfico anetrior. Dado que la tendencia habla de anécdotas, es esperable tener este comportamiento al tener anécdotas en común.

## 6. Trabajo futuro.

Como se ha mencionado, del análisis de este trabajo nace una cuestión importante en relación a las métricas. En particular, en las relaciones no triviales entre algunas. De los gráficos, se puede apreciar relaciones interesantes entre las métricas calculadas y la entropía de Shannon de la distribución de grado. De lo cual, más allá de considerar los casos extremos (red vacía, red completa, red regular, etc.) sería interesante realizar un análisis exhaustivo sobre estas relaciones implícitas. En particular, en crecimiento, decrecimiento o fórmulas para 'pasar' de una métrica a otra; los costos computacionales cambian de una métrica a otra. En resumen, sería seguir resolviendo esta cuestión de relaciones selectivas.

Otro ideal para esta modelación sería interconectar esta moción con otras redes de media social (abrir Facebook y ponderarla de algún modo) o, evidentemente, poder tener estos mismos datos pero para la comunicación no escrita pero personal. Si tuviéramos dichos datos del último caso, estaríamos más cercanos a comprender la comunicación.

Asimismo, sería también útil considerar lo que se está hablando. En la modelación, se toma la relación **si existe un tweet**, por lo que perdemos información esencial respecto a dicho *tweet*; por ejemplo, algún sentimiento. El mismo indica nos cierta dinámica de comunicación; es bien sabido que existen simbolismos de caracteres que se han ido modificando conforme la sociedad y época avancen.

## Referencias

- [1] Saeed Abdullah and Xindong Wu. An epidemic model for news spreading on twitter. In *2011 IEEE 23rd International Conference on Tools with Artificial Intelligence*. IEEE, November 2011.
- [2] Roy M. Anderson. *Infectious Diseases of Humans: Dynamics and Control*. Oxford University Press, sep 1992.
- [3] Michael Conover, Jacob Ratkiewicz, Matthew Francisco, Bruno Goncalves, Filippo Menczer, and Alessandro Flammini. Political polarization on twitter. *Proceedings of the International AAAI Conference on Web and Social Media*, 5(1), Jul. 2011.
- [4] T. M. Cover. *Elements of information theory*. Wiley-Interscience, Hoboken, N.J, 2006.
- [5] Manlio De Domenico and Hiroki Sayama. Complexity explained. 2019.
- [6] Zlatinka I. Dimitrova and Nikolay K. Vitanov. Influence of adaptation on the nonlinear dynamics of a system of competing populations. *Physics Letters A*, 272(5-6):368–380, August 2000.
- [7] Jorge Fabrega and Pablo Paredes Navarro. Social contagion and cascade behaviors on twitter. *Information*, 4:171–181, 06 2013.
- [8] Eleanna Kafeza, Andreas Kanavos, Christos Makris, and Pantelis Vikatos. Predicting information diffusion patterns in twitter. volume 436, 09 2014.
- [9] Srijan Kumar, William L. Hamilton, Jure Leskovec, and Dan Jurafsky. Community interaction and conflict on the web. *CoRR*, abs/1803.03697, 2018.
- [10] Kristina Lerman, Xiaoran Yan, and Xin-Zeng Wu. The "majority illusion" in social networks. *PLOS ONE*, 11(2):e0147617, February 2016.
- [11] Robert May and Warren Leonard. Nonlinear aspects of competition between three species. *SIAM J. Appl. Math.*, 29:243, 09 1975.
- [12] Giovanna Miritello. Social and communication networks. In *Temporal Patterns of Communication in Social Networks*, pages 9–44. Springer International Publishing, 2013.
- [13] Melanie Mitchell. *Complexity : a guided tour*. Oxford University Press, Oxford England New York, 2009.
- [14] Azadeh Nematzadeh, Emilio Ferrara, Alessandro Flammini, and Yong-Yeol Ahn. Optimal network modularity for information diffusion. *Physical Review Letters*, 113(8), August 2014.
- [15] M. E. J. Newman. *Networks*. Oxford University Press, Oxford, 2018.
- [16] Gunnar Olsson. Distance and human interaction. a migration study. *Geografiska Annaler. Series B, Human Geography*, 47(1):3, 1965.

- [17] Soumajit Pramanik, Qinna Wang, Maximilien Danisch, Jean-Loup Guillaume, and Bivas Mitra. Modeling cascade formation in twitter amidst mentions and retweets. *Social Network Analysis and Mining*, 7(1), August 2017.
- [18] A. Viol, Fernanda Palhano-Fontes, Heloisa Onias, Draulio B. de Araujo, and G. M. Viswanathan. Shannon entropy of brain functional complex networks under the influence of the psychedelic ayahuasca. *Scientific Reports*, 7(1), August 2017.
- [19] Bing Wang, Huanwen Tang, Chonghui Guo, and Zhi-Long Xiu. Entropy optimization of scale-free networks' robustness to random failures. *Physica A: Statistical Mechanics and its Applications*, 363:591–596, 05 2006.
- [20] Duncan J. Watts and Steven H. Strogatz. Collective dynamics of ‘small-world’ networks. *Nature*, 393(6684):440–442, June 1998.
- [21] Lilian Weng, Filippo Menczer, and Yong-Yeol Ahn. Virality prediction and community structure in social networks. *Scientific Reports*, 3(1), August 2013.
- [22] Lilian Weng, Filippo Menczer, and Yong-Yeol Ahn. Predicting successful memes using network and community structure, 2014.
- [23] Arkaitz Zubiaga, Damiano Spina, Raquel Martínez, and Víctor Fresno. Real-time classification of twitter trends. *Journal of the Association for Information Science and Technology*, 66(3):462–473, May 2014.

## 7. Anexo

### 7.1. ID de tendencias

ID	Tendencia	ID	Tendencia	ID	Tendencia
0	teamautofollow	50	waystomakemehappy	100	rns
1	first20songsonshuffle	51	youknowwhatannoysme	101	soundcloud
2	ohwell	52	fail	102	somefactsyoushouldknow
3	10thingsihate	53	10peopleontwitteriwanttomeet	103	amitheonlyonethat
4	autofollowback	54	arentyouiredof	104	mythoughtsduringsex
5	wewontworkoutif	55	team	105	youdontbelongontwitterif
6	fml	56	respect	106	liestoldontwitter
7	bizitalk	57	thingspeopledothatgetonmynerves	107	whatmakesmesmile
8	thatisall	58	weed	108	classic
9	teamheat	59	egypt	109	libras
10	ididnttextyouback	60	scorpions	110	thingsblackpeopletakeseriously
11	followme	61	truestory	111	followingain
12	tweetyourheight	62	twitterpeopleilove	112	backinthedaywheniwasakid
13	lolatgirlswho	63	butyouathug	113	ifitwasuptome
14	ijs	64	1dfacts	114	thingssaywhilereadingmytl
15	mythoughtsduringschool	65	randomtweet	115	reasonsthatimsingle
16	thatirritatesme	66	itscrazyhow	116	thestruggle
17	gop	67	ifollowback	117	iwanttopunchpeoplewho
18	thuglife	68	iwonderif	118	sometimesjustwant
19	imhappywhen	69	thingsilovetosee	119	itscutewhen
20	lmao	70	music	120	twothingsthadtontmix
21	mentionto	71	iknowthisonegirl	121	thingsthatfrustrateme
22	thingsthatguyslike	72	fourwordsyoudontwanttohear	122	justsayin
23	itstimeforyoutorealize	73	girlfacts	123	immadbecause
24	top100lies	74	photo	124	imsickof
25	win	75	wheniwaslittle	125	uk
26	mybiggestfearis	76	favtvshow	126	life
27	kindle	77	amen	127	kuwait
28	bored	78	f1	128	icantlivewithout
29	fuckit	79	iwantsomebody	129	thatdepressingmoment
30	yougetmajorpointsif	80	geminisignz	130	ows
31	bbl	81	damn	131	hardestthingsinlife
32	icantstandwhen	82	stayawayfrommeif	132	hotpeopleifollow
33	ifweweretogtherrightnow	83	teamnosleep	133	support
34	nomanshouldever	84	showerflow	134	nameyourfavoriterapper
35	foreveralone	85	excited	135	wrestlemania
36	thingsigottateachmyson	86	14feb	136	rip
37	youvedoneitto	87	whatgetsmemad	137	bigtimerush
38	lowkey	88	wordsyouwillneverhearmesay	138	sadtweet
39	tragic	89	youknowwhatsannoying	139	nowthatimolder
40	nationalbestfriendday	90	favmusic	140	s
41	itsfunnyhow	91	incollege	141	justrememberthat
42	awkward	92	fuck	142	morefemaleshould
43	pbbteens4	93	thatsthetruth	143	thatawkwardmoment
44	thehairoscope	94	swagge	144	wtf
45	throwback	95	4wordsyoudontwanttohear	145	turnon
46	real	96	stfu	146	serioustweet
47	moviesthatnevergetold	97	boss	147	1dfamily
48	somf	98	turnup	148	thevoice
49	twitterafterdark	99	1000aday	149	2omf

Cuadro 1: Identificadores de las tendencias.

## 7.2. Matriz de correlación de la tabla $\Psi$ .

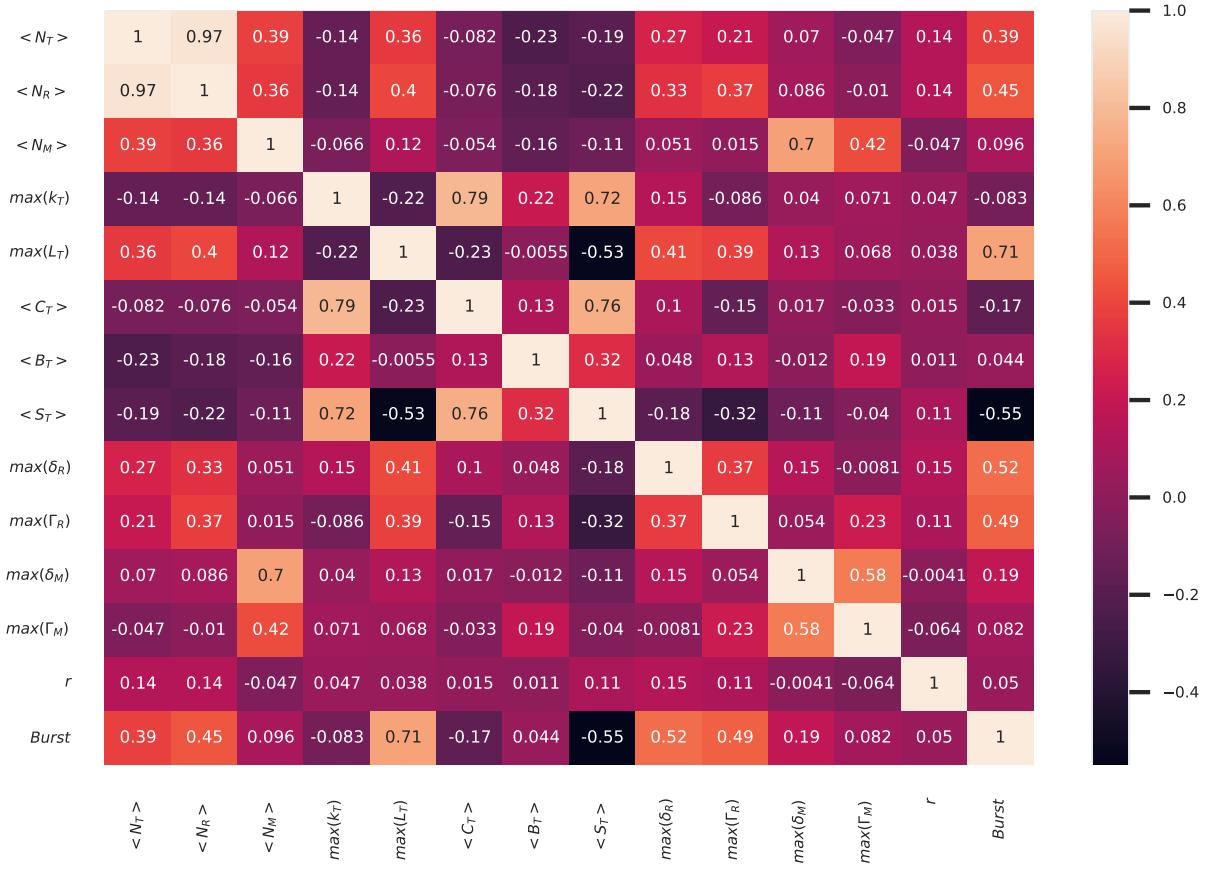


Figura 11: Matriz de correlación de la tabla  $\Psi$ .

### 7.3. Tabla del análisis completo de una tendencia.

D	H	$N_T$	$N_R$	$N_M$	$k_T$	$L_T$	$C_T$	$B_T$	$S_T$	$\delta_R$	$\Gamma_R$	$\delta_M$	$\Gamma_M$
94	5	9168	1845	384	0.350844	11.569085	0.005021	0.000011	0.993730	12	75	2	5
94	6	6655	1425	302	0.468180	10.380293	0.006480	0.000016	1.175080	7	43	2	6
94	4	6646	1323	262	0.389641	10.926167	0.003936	0.000005	1.049006	9	77	3	4
94	3	2452	528	91	0.539525	8.269113	0.002718	0.000027	1.195304	6	32	3	24
94	7	2308	468	116	0.697282	13.999618	0.009423	0.000044	1.414084	5	23	3	3
94	8	935	161	42	1.233923	9.736608	0.021412	0.000141	1.660234	2	8	2	3
94	9	255	62	6	2.038113	9.147813	0.016911	0.000563	0.730588	2	5	1	1
94	10	124	27	1	1.962338	4.537095	0.002326	0.000147	0.552646	2	3	1	1
99	1	121	25	6	2.034382	3.948445	0.024321	0.000072	0.737575	2	13	2	2
94	11	111	27	3	2.057650	4.741415	0.013568	0.000239	0.703273	2	5	1	2
99	0	103	3	2	1.949467	5.879695	0.000000	0.000166	0.487029	1	1	1	1
94	12	83	16	1	1.968254	2.035295	0.009785	0.000104	0.736758	2	3	1	1
94	13	52	11	2	1.953361	3.755908	0.011546	0.000372	0.696138	2	2	1	2
94	17	50	12	0	1.949561	1.993197	0.009940	0.000192	0.516606	2	3	0	0
94	18	42	8	2	1.987336	3.330718	0.002111	0.000825	0.249372	1	1	1	1
94	14	38	5	2	1.917012	1.991736	0.000000	0.000201	0.454115	2	2	0	2
94	20	36	8	1	1.885714	1.984962	0.000000	0.000216	0.491913	1	1	1	1
94	15	34	5	4	1.938272	1.962963	0.000000	0.000149	0.573881	1	1	1	2
94	19	33	10	1	1.971737	4.217838	0.004339	0.000841	0.377835	1	1	0	2
94	16	33	7	0	1.946779	1.994444	0.003774	0.000405	0.448407	1	1	0	0
94	21	31	14	1	1.916019	2.097059	0.000000	0.000134	0.519109	2	7	0	2
95	4	25	11	0	1.962901	2.523212	0.000000	0.000612	0.231220	2	3	0	0
95	2	24	10	1	1.968310	1.996176	0.001762	0.000252	0.238597	2	2	0	2
95	1	18	9	1	1.947287	1.980769	0.000000	0.000153	0.263760	2	2	1	1
94	22	17	2	0	1.935035	1.970149	0.000000	0.000213	0.372793	1	1	0	0
100	9	16	2	1	1.963964	1.971014	0.039344	0.004269	0.669232	1	1	1	1
95	5	15	5	0	1.936842	1.983051	0.000000	0.000286	0.312993	1	1	0	0
94	23	15	6	0	1.938865	1.990950	0.000000	0.000523	0.282524	2	2	0	0
95	8	15	9	0	1.925926	1.982906	0.000000	0.000347	0.336678	1	1	0	0
96	7	14	7	0	1.972603	2.982741	0.000000	0.000838	0.199472	2	3	0	0

Cuadro 2: Muestra de 10 períodos de tiempo de  $NC(h)$  para la tendencia  $h = \text{wheniwaslittle}$  ordenados por el número de *tweets*.

## 7.4. Tabla del análisis completo de tendencias.

ID	$\langle N_T \rangle$	$\langle N_R \rangle$	$\langle N_M \rangle$	$\max(k_T)$	$\max(L_T)$	$\max(\Gamma_T)$	$\max(\delta_T)$	$\max(\delta_R)$	$\max(\Gamma_R)$	$\max(\delta_M)$	$\max(\Gamma_M)$	$r$	Burst	
27	6.540441	1.816176	0.363971	5.490022	4.031870	0.120161	0.008320	1.205635	7	6	5	-0.005699	0	
7	7.636029	2.346814	1.026961	6.359447	3.267334	0.205605	0.009887	1.188201	7	6	5	-0.222230	0	
74	5.781863	0.264706	0.084559	3.636364	4.690135	0.150927	0.015657	1.121490	3	2	2	-0.397111	0	
74	9.242647	0.838235	0.284314	5.655471	5.655044	0.02203	0.025391	0.948739	3	8	2	-0.172189	0	
86	9.242647	2.460784	0.427696	4.015248	6.091539	0.025184	0.018416	0.717766	7	8	3	-0.253508	1	
78	10.801471	2.460784	0.406863	2.858469	5.645591	0.014242	0.011809	0.675581	5	11	5	-0.108792	0	
59	10.692402	1.094363	0.348039	3.263415	6.099740	0.039215	0.010008	0.661491	4	3	2	-0.145292	0	
17	6.941176	0.960784	0.341944	5.194444	5.038217	0.041637	0.008674	0.646954	4	6	3	-0.294217	1	
90	11.900463	0.881944	0.341918	6.358824	5.977072	0.022798	0.012207	0.603956	3	4	5	-0.179169	0	
130	6.8889706	0.774510	0.4230392	5.588235	6.558068	0.007537	0.015475	0.590751	3	3	2	-0.092024	0	
127	6.574659	0.196078	0.660539	3.309859	5.348310	0.038780	0.009768	0.562601	5	8	2	-0.322131	1	
76	14.319444	2.503472	0.295139	3.718750	6.024385	0.010277	0.017251	0.434653	4	5	2	-0.282797	1	
125	5.052696	1.014706	0.301471	4.84211	4.824155	0.01208	0.016798	0.432464	2	19	3	-0.259428	0	
25	6.412990	0.301471	0.644608	5.78431	2.223776	0.003026	0.001021	0.03841	4	2	6	-0.212416	0	
6	12.988971	0.644608	0.341918	5.263251	5.507952	0.002668	0.013224	0.421349	4	14	2	-0.245133	0	
64	7.546569	2.196078	0.437500	4.144144	4.163499	0.001004	0.004501	0.419609	2	8	3	-0.223255	0	
144	11.533088	0.825980	0.519608	2.875000	3.201117	0.01436	0.009141	0.414563	2	3	4	-0.198820	0	
138	8.653186	0.727941	0.306127	2.455696	4.238190	0.005889	0.021058	0.407561	2	3	10	-0.216108	0	
21	6.496324	0.495098	0.388480	2.096280	3.824768	0.001734	0.006976	0.405553	2	5	2	-0.170039	0	
2	9.469363	0.774510	0.533088	2.307692	5.165529	0.001471	0.015169	0.393512	2	7	2	-0.211870	0	
85	5.040441	0.341912	0.533088	2.333333	4.639423	0.001127	0.012867	0.392067	2	11	2	-0.236278	0	
42	5.512255	0.485294	0.533088	2.329268	4.322657	0.002505	0.008678	0.378822	2	10	6	-0.246062	0	
35	8.089461	1.262255	0.950980	2.066127	2.107280	0.01285	0.010452	0.377735	2	3	4	-0.198269	0	
92	6.037990	0.533088	0.267157	4.174475	4.174475	0.001285	0.007841	0.377066	4	6	3	-0.301450	1	
137	6.676768	0.486111	0.180556	5.055292	5.221831	0.022807	0.007841	0.375312	3	8	8	-0.254825	0	
147	5.430147	0.753676	0.601716	2.503872	6.147368	0.001287	0.012187	0.375312	3	2	3	-0.181640	0	
96	5.330882	0.546569	0.216912	2.147368	3.585759	0.001029	0.011497	0.373224	2	3	5	-0.156563	0	
81	7.575980	0.691176	0.585784	2.326923	3.963572	0.002856	0.007342	0.354922	2	9	2	-0.247251	0	
122	8.841912	0.988971	0.962010	2.151372	3.854013	0.001219	0.006286	0.345737	2	18	2	-0.167322	1	
145	7.138480	1.215686	0.268382	3.244828	4.531284	0.003221	0.011494	0.342868	3	4	4	-0.164321	0	
61	5.814951	0.634804	0.712010	2.19355	3.751120	0.001517	0.008707	0.356551	2	9	2	-0.255396	0	
20	6.116422	0.617647	0.691176	2.132038	3.782310	0.001389	0.009130	0.334142	2	4	5	-0.272232	0	
143	6.947304	1.165441	0.381127	2.525077	4.069689	0.001489	0.010008	0.333156	3	17	3	-0.263028	0	
73	5.828431	5.025735	0.406863	2.183824	4.896371	0.000664	0.010563	0.328243	2	64	3	-0.195191	0	
58	43	7.403986	1.489130	1.610507	5.412027	0.002321	0.012493	0.326399	2	4	3	-0.201887	1	
28	52	9.813725	0.415441	1.307598	4.308847	0.001173	0.006909	0.318304	2	12	2	-0.242933	0	
61	70	6.296569	0.714422	1.324755	2.178323	5.436420	0.002815	0.009276	0.316214	2	53	3	-0.294079	0
29	57	5.761029	0.508578	0.449755	2.149584	3.209916	0.001061	0.006329	0.315261	2	3	4	-0.288531	0
58	58	5.074755	2.183824	4.1610507	5.2571429	6.461630	0.003601	0.02248	0.306986	3	46	3	-0.194130	1
43	58	7.525735	0.814951	0.997549	2.210526	5.345676	0.001613	0.006559	0.305856	6	23	5	-0.231628	0
108	45	7.254902	0.631127	0.613971	2.097087	4.728298	0.001302	0.008695	0.302847	2	3	5	-0.267796	0
11	11	10.238971	0.420343	1.691176	2.179420	6.045764	0.001267	0.003914	0.302839	2	100	4	-0.294079	0
94	94	5.008578	1.300245	1.530637	3.268053	5.926635	0.004433	0.013225	0.299251	2	49	17	-0.288531	0
140	140	6.045343	1.121324	0.991422	2.428094	4.902929	0.003387	0.006927	0.296712	2	8	3	-0.194130	1
136	136	6.924020	0.981618	0.349265	2.218182	7.644198	0.001900	0.007462	0.294705	2	10	3	-0.287757	0
100	100	9.704657	1.246324	1.305833	2.510204	5.640345	0.002963	0.004765	0.293238	2	4	5	-0.213641	0
101	101	7.495098	0.590686	1.475490	2.516981	5.651765	0.002917	0.007120	0.291475	2	3	7	-0.243313	0
126	126	5.754314	0.867647	0.341912	2.152174	4.690915	0.000927	0.008283	0.291470	2	12	4	-0.246228	0
99	99	6.252451	0.390931	0.335833	2.161658	5.731877	0.001163	0.008000	0.291318	3	4	5	-0.224541	0
111	111	5.279412	0.333333	0.825980	2.557377	4.876508	0.002395	0.010659	0.285390	3	3	4	-0.305217	0