A Design to Visualize Cellphone Communication Log in an Interesting Way

Mingzhao LI, Min ZHU*, Qihong GAN, Ting LIANG

Vision-Computing Lab, College of Computer Science, Sichuan University, Chengdu 610064, China

Abstract

Cellphone communication log contains plenty of temporal information which can highly reflect a person’s daily routine and social relationships. This paper proposes a novel visualization design to help discover and visually understand the information in an interesting way. Otherwise from prior time-oriented visualizations, our design visualizes detailed temporal information of each call item with a Hollow Spiral Model, simultaneously presents the overview information based on each contact in center of the hollow spiral; retinal variables like color, shape, location and size are artfully utilized to present some additional information. We implement a tool named TelogViz, which enriches our design with a series of interactions including filtering, highlighting, connecting and searching. Case studies demonstrate the effectiveness of our scheme that can discover several useful and surprising patterns hidden in the original dataset.

Keywords: Cellphone Log; Hollow Spiral Model; Information Visualization; Temporal Information

1 Introduction

Of today, cellphone has become an essential device in people’s daily life. Only in China, the number of mobile phone users has over-fulfilled a billion. [1] People use cellphone to connect with friends and colleagues by making calls or sending messages. Cellphone communication log contains plenty of information which can highly reflect a person’s daily routine and social relationships. Visualizing the information can be significant for mobile phone users to conclude their communicating patterns, and even for polices to obtain some crucial social intercourse behavior of suspects.

In cellphone communication log, each item of call records corresponds to a series of detailed temporal information, including the date, beginning time and duration of the call. The most intuitionistic method to visualize such temporal dataset is dynamic expression [2, 3], which is easy to understand, and more retinal variables [4] can be utilized to present additional information; unfortunately, users often have no patience to remember history, which is the common fault

*Project supported by the National Nature Science Foundation of China (No. 60970013), Science and Technology Support Program of Sichuan Province, China (No. 2013GZ0015).

*Corresponding author.

Email address: zhumin@scu.edu.cn (Min ZHU).
of dynamic visualizations. Another one of the most general solutions is to map the temporal attributes to a rectangular coordinate system [5, 6], illustrated in Fig. 1(a); while periodicity of data is still difficult to discern. Many researchers [7–9] have utilized circle or spiral representation which is much easier to discover the periodic pattern hidden in dataset. However, the narrow space in center of the circle makes it difficult for users to recognize information near the origin, shown in Fig. 1(b); and the designs generally do not support the representation of time quantum.

Based on the above observation and considering that spiral representation surpasses the circle representation to visualize the continuous time in all timelines, in this paper, we propose a Hollow Spiral Model, which uses a circle to represent 24 hours in a day, like an analog clock. As shown in Fig. 1(c), a helix with about 30 circles represents the days in a month to guarantee the connection of the end of the day and the beginning of next day. We map each call as an arc on the specific location due to its beginning time, and angle of the arc is determined by length of the call. Combining some color information, we could easily figure out in which durations of a day, a person or a group usually gives a call.

One of the most benefits of our design is that in center of the spiral model, there remains enough space to visualize the overall information based on each contact, and it’s easy to connect the overall expression with the detailed information on outer spiral model. In our design, a contact is represented by a solid regular polygon, the size and vertex numbers of which is separately determined by the total length and amount of calls in a month; and the filling color differs from each group or each person, which corresponds to the color of their calls in outer layer. To avoid visual confusion, we move more detailed information to interaction operations. For example, when we click on a polygon, corresponding arcs in outer spiral model of the person will be highlighted and connected with the polygon.

![Fig. 1: (a) Illustration of mapping temporal attributes to a rectangular coordinate system. (b) Circular representation of temporal attributes. (c) Hollow spiral representation in our design.](image)

In particular, the specific benefits of our design and the contributions of this research are as follows:

- Propose a Hollow Spiral Model to represent about 30 days in a month, in which the scale can be accurate to minutes, which is efficient to represent time quantum.
- Following the visualization design principle, realize the scheme of “overview+detail” with detailed expression on outer layer and general information on center.
• Apply the visualization design on the dataset of cellphone communication log to explore the hidden valuable information.

Apart from the above, we represent the general information of a person with a polygon, which is novel; and we use the retinal variables like color, size, location and specific shape to summary the calls of a person. Furthermore, we develop a tool named TelogViz with some friendly interactions.

2 Related Works

2.1 Visualization of time-oriented data

Several visualization tools have been developed based on Time-oriented data [10, 11]. The most intuitionistic methods are dynamic representations. For an instance, Gapminder Trendalyzer [3] is an animated bubble chart designed to show trends over time in three dimensions. The dynamic expression is easy to understand, and more other attributes can be reflected to the view. Unfortunately, dynamic visualizations have a common fault that the requirement of remembering history is not user-friendly.

Apart from the dynamic representations, one of the most general solutions is to map the temporal attributes to rectangular coordinate system. Among the examples of such visualizations, ThemeRiver [5] is widely used and exhibits its fascination in viewing thematic changes over time; PlanningLines [6] presents project plans which allows the depiction of temporal uncertainties via special glyphs. One of the most weaknesses of those time series plot methods is that periodic pattern is difficult to discern.

Another thought to process time-oriented data is to map the temporal attributes on a circle or spiral. Spiral representation surpasses the circle representation in visualizing the continuous time in all timelines. SpiralGraph [7] encoding with some cyclic time can easily reveals periodic pattern. While since the narrow space in center of the circle, it is difficult for user to recognize information near the origin. In this paper, we present a Hollow Spiral Model, which possess the advantage of traditional spiral representation, avoids the narrow space near the origin, and simultaneously leaves some space to present some overall patterns.

2.2 Visualization of circle and spiral representation

Circle and Spiral representations have been widely used in visualization designs. Calis et al. [8] introduce a spiral visualization technique, which displays data along a spiral to highlight serial attributes along the spiral axis and periodic ones along the radii. However, the narrow space in center of the circle restricts the effectiveness of spiral representation. Weber et al. [9] propose another spiral visualization based on time-series data, and extend the technique to 3D representation. However, the design has not considered the overall time information and does not support the representation of time quantum, which is so important of call log.

Wu et al. [12] propose the design of OpinionSeer to interactively visualize hotel customer feedback, which also use the circle representation to express time-series information. However, the design does not suit for call log, since it has placed more emphasis on the comparison of opinion in different periods, but not the presentation of detailed temporal information. Shiroi
[13] propose the design of ChronoView to visualize many temporal data, which define several
time-stamps for each event and focus on discovering overall time-event relationships. The design
of ChronoView more suits for the dataset which only value the 24-hour information. However, the
time quantum in TelogViz is one of the most important information, the detailed time information
including the date of a call still has significance.

In conclusion, the visualizations we mentioned above all have their particular advantages to
explore data with temporal information; while, the narrow space in center of the circles, not
considering overall information or not supporting the representation of time quantum make them
do not suit for the visualization of cellphone communication log.

3 Design of the Visualization

In this section, we describe the detailed design. First, we propose a Hollow Spiral Model to
represent time and date; and all the calls will be mapped as several pitch arcs. Then, we define
polygon with attributes to represent the contacts involved.

3.1 Hollow Spiral Model to represent time and date

We propose a Hollow Spiral Model to represent the temporal information which is one of the most
important attributes. Considering the audio-visual image of an analog clock and that time can
reflect more about the frequency and periodicity of events, we map the temporal information on
a spiral with about 30 circles that represent the days in one month. And the hours can be found
on each circle similar to a 24 hour clock.

We first define the due north direction as 0 o’clock. Moving clockwise, the 24 hours in one
day can be orderly located in the circle, i.e. 12 o’clock is defined in the due south direction
and 21 o’clock is defined in the northwest direction. For a time-stamp \( t \) which is represented as
\( d - tH : tMin \) (for example, 10–22:30 means half past 22 in the 10\(^{th}\) day of the month), the angle
and distance away from origin point are defined as \( \alpha \) and \( rLocation \), and the number of the total
days in the month is defined as \( dNum \). Here, the positions of time-stamps of each spiral of the
day to compose the month are expressed by
\[
f_d(t) = (rLocation_d \times \cos \alpha, rLocation_d \times \sin \alpha)
\]
(1)

\( rLocation \) and \( \alpha \) can be calculated from,
\[
\begin{cases}
  rLocation_d = rBegin_d + \frac{tH+tMin/60}{24} \times rStep \\
  \alpha = \frac{5}{2} \pi - \frac{tH+tMin/60}{12} \pi
\end{cases}
\]
(2)

Where \( rStep \) is the distance of space between each two circles in spiral, and \( rBegin_d \) is the
beginning radius of each circle in the spiral as,
\[
rBegin = \{rBegin_d = rBegin_0 + d \times rStep, d \in [0, dNum - 1]\}.
\]
(3)

Until now, we have mapped the days in a month as a hollow spiral with \( dNum \) circles. For a
dataset containing \( N \) calls as
\[
Call = \{call_i \mid d_i - tH_i : tMin_i - cMin_i, i \in [0, N - 1]\}.
\]
(4)
Where $cMin$ means the total minutes of a call, we get the beginning point $A_i$ for call $i$ of the arc based on the beginning time with $f_d(t)$ generated from Eq. (1), and then the end point $B_i$ by adding call time quantum. The calls will be mapped as $N$ arcs,

$$\mathbf{AB} = \{A_iB_i, i \in [0, N - 1]\}. \quad (5)$$

### 3.2 Polygon representation of a contact

Apart from the detailed temporal information of their calls, users usually have more curiosities about which groups of people contact more often with them, when those people generally contact them and how long the calls last. We cluster the calls based on each contact and present the information in center hollow region of the spiral. Each contact is represented by a solid regular polygon, the size of which and the location of which are separately determined by the total call minutes and the beginning time of each call of the contact in a month.

Suppose that the number of the contacts in a month is $P$ , we divided all the calls in Call defined in Eq. (4) as $\text{Call} = \bigcup_{p=0}^{P-1} \text{Call}_p$, simultaneously satisfy that $\text{Call}_m \cap \text{Call}_n = \emptyset, m \neq n$. Consider that for a contact, the beginning time $T_p$ and time of each call $C_p$ of its all $M$ calls can be got from $\text{Call}_p$. To represent the total calls of a contact, the primary polygon size $s_p$ is defined as

$$s_p \propto \sum_{m=0}^{M-1} \lfloor c_{p,m}(\text{minute}) \rfloor; c_{p,m} \in C_p. \quad (6)$$

And the location of the polygon $l_p$ is defined as

$$l_p \propto \frac{1}{M} \sum_{m=0}^{M-1} f_0(t_{p,m}), t_{p,m} \in T_p. \quad (7)$$

The benefits to locate the polygon in this way is that users can easy to get the general time when a person give them a call and whether the time is centralized or sporadic. For example, if a contact always makes calls at the same time of a day or only has an item of call log, the location of the polygon lies apart away from the origin point, even locates on the most inside clock circle. And if the calls from someone are not punctual, or even distribute all through the day (for example, three calls separately in 6:00, 14:00 and 22:00), the position of the polygon to represent the contact will lie near the origin point.

We utilize translucency to solve the problem that circles in center may intersect each other. And we define coefficients for $s_p$ and $l_p$ to avoid that the circles could be over large or small, or that some circles may overlap with the edge of the spiral.

### 3.3 Representation of additional information

A contact or a call always has some attributes. For example, people usually arrange contacts in different group; and rather than the comparison of different person, that of different group always reflect more about a person’s personality or latest critical lifestyle. Besides the location and size of the polygon, we use some additional variables to represent the corresponding information.
Shape The shape (i.e. vertex numbers) of the regular polygon to represent a contact is determined by amount of the calls between them in a month. For example, if a person contacts once, the polygon is defined as a triangle; if someone contacts twice, the polygon is defined as a square; and if the contacting times is n, the polygon will have n+2 sides. It is clearly that the bigger the amount of the calls, the polygon will be more like a circle. It is quite acceptable that users cannot distinguish between a decagon and a dodecagon, since the communication is generally considered as frequent when the number of calls in a month exceeds about five or six.

Color We use color to represent different person or the group it belongs to. We supply some basic colors for users to map their groups as different color. Like, they may give the color of group “family” as yellow; and decide that the color of group “colleague” as blue. In this way, users could easily figure out which group they have been contact more often in the specific time. What should be stressed is that the arcs in outer layer to represent calls have the same color with the polygon to represent the contact. Thus, the specific calls can be generally connected to the corresponding contacts. In later section, we will introduce the interactions to enable the complete matches.

Apart from the above, we have some other variables to represent the attributes of our dataset which is highly related to interactions and will be discussed in the next section.

4 Development of TelogViz

We developed a tool named TelogViz. All visualization algorithms for our experiments are implemented on a Lenovo desktop with Intel Pentium Dual T2390 CPUs and 2GB memory. The graphics card used is an NVIDIA GeForce 8400M GS with 512MB of DDR2 memory. The programming language used is C++, with QT/OpenGL libraries. The design of user interface and interaction skills will be discussed in this section, following the introduction of main procedures in the process of implementation.

4.1 Implementation procedures

The main procedures to create visualizations in TelogViz are simply described as follows.

- Step 1: Process the imported data, format the data items in communication log file.
- Step 2: Sort data items by call numbers, and match the numbers with contact file.
- Step 3: Calculate to get the location of the arcs to represent each call.
- Step 4: Calculate to get shape, location and size of the polygons in center to represent each contact.
- Step 5: With the color modeling information, map all the calls and the contacts in the Hollow Spiral Model described in Section 3.

With some user interactions, the visualization will go to the specific procedure and for all points will update the position.
4.2 User interface and interactions

As shown in Fig. 2, the whole visual design is placed in the center of the main window. We provide some operations on left side of the main window or on menu bar, including the file operation, data importing and editing, color modeling, data filtering and searching. User can also directly interact with the interface by clicking on the polygon in center or the arcs in outer layer.

**Fig. 2:** Overview of TelogViz. Main window is the result after clicking one of the contacts named **Xia Chaozhi**, telephone number of which is played mosaic.

**Highlighting** The polygons in center have some ambiguity to represent time; for example, two contacts may both be put near the origin if one has two calls at 8:00 and 18:00, while the other one has three calls at 6:00, 14:00 and 22:00. Fortunately, we have detailed temporal time in outer circles. While, to avoid confusion, we do not connect the general information with the detailed information on default, or label text on the polygon for all the contacts. When user clicks on the polygon, the calls of the selected contact will be highlighted, with all other calls turning to grey; for not to miss any call, lines will be drawn to connect the contact with all highlighted calls; synchronously, detailed information of the contact will be presented on the right region of the screen. When user clicks an arc on outer layer, all the calls from the same contact together with the polygon to present the contact will be highlighted.

**Data Filtering** TelogViz allows user to filter data by contacts, groups, call times and days in a week. For example, user can eliminate several uninterested contacts or a whole group, or only keep the contacts whose calls are more than once; the filtering condition is also displayed on left of the screen, and synchronously, the whole visualization will change. User also can filter the calls only in several days in a week, like only in weekend.

**Color Modeling** As a key step for generating a visual view, color modeling is a confusing
problem. Researchers have used colors to represent semantically similarity. Since in TelogViz, color means the different group, we provide for users a color panel which supports three color schemas, as HSV, RGB and CMYK. A user can choose any color for a group according to their tastes, thus to create their own design in TelogViz.

Focus+Context Operation In our design, user can double click on any interesting region to zoom in the whole view and move the region to be focused on center. TelogViz also provides the keyboard control for the operation.

Other Operations TelogViz supports searching the contact by its name or its telephone number; the polygon to represent the person and the corresponding arcs in outer layer will be highlighted. User can select from the menu bar to save current visualization as a picture. Also, we provide a table to allow user deleting or editing the original dataset, including the log file and contact information.

5 Use Case

Four cases are studied with TelogViz in this section. We will first give the dataset description including acquisition of the dataset and some data characteristics, followed by the comprehensive exploration with TelogViz. At last, we will discuss the effectiveness of our visualization design and tools.

5.1 Dataset description

We utilize four datasets to present the case studies, all of which are downloaded from the official website of China Mobile [14]. Two of the datasets (I, IV) are from the same person (A), who provides his communication logs in August, 2012 and March, 2013. The second dataset (II) is from person B; and the third dataset (IV) is from person C. The contact file of A and B is exported from their address lists; while we do experiment without person C’s contact file.

The original dataset have several columns, including the expense of each call, and the location, etc. While in TelogViz, the useful columns mainly include calling date and time, duration of the call, telephone number. The detailed information of the dataset and the address list are described as follows.

- Dataset I (A). Includes 262 items, involving about 70 contacts.
- Dataset II (B). Includes 170 items, involving about 30 contacts.
- Dataset III (C). Includes 327 items, involving about 60 contacts.
- Dataset IV (A). Includes 412 items, involving about 35 contacts.
- Contact file A. Includes 171 contacts, within 7 groups.
- Contact file B. Includes 457 contacts, with no group.

5.2 Comprehensive exploration with TelogViz

We analyze the datasets in TelogViz as use cases of our design. Several results based on the four datasets are presented in Fig. 3. In this subsection, we will separately discover the daily routine and social relationship hidden in dataset, together with some particular data.
Fig. 3: Experiments of the four datasets. (a) Dataset I, Contact File A; different group with different color. (b) Dataset II, Contact File B. (c) Dataset III, without contact file. (d) Dataset IV, Contact File A.

5.2.1 Daily routine discovery

The first observation we can get is the time for each person to fall asleep and get up. Generally speaking, A, B and C all sleep and get up late. Actually, they all live in Chengdu which is famous in China as a relaxed city. We also can find that the bedtime presented in Fig. 3(c) and (d) is more regular than that in Fig. 3(a) and (b).

As shown in Fig. 3(a) (Fig. 1(c) presents result of the same dataset with no group information), the sleep time of person A in August is irregular. The bedtime is late which differs from midnight to over 2:00 am. More detailed observation shows that in the first half of the month, the person almost has no calls in morning or in late evening; there exist about four days that the person gets up too early and sleep too late; after that, the daily schedule is more regular. Actually, later we know that, August is the summer holiday of person A, who stayed at home to get rests in the first two weeks; in middle of the month, A had a meeting in another city; five days later, he
returned to school. We know from Fig. 3(d) that A has a more regular daily schedule in school, especially the bedtime.

We also can observe or deduce some other information from the result, like at what time of a day the person often gives or receives a call (for example, B often give calls at lunch and supper time and call time of C is often supper time and bedtime.), the habit of the person whether or not close the cellphone when going to bed (B is more likely to keep cellphone open than C), or at what time the person usually give calls to some specific contacts or groups.

5.2.2 Social relationship discovery

Since cellphone is one of the most commonly used communication tools in people’s social life, we can discover many valuable patterns revealing the social relationship or personality of the person from the log. The patterns can be more interesting with group information. As shown in Fig. 3(a), person A generally contact more with friends (blue) and colleagues (yellow) than family (cyan). Actually, the calls with friends are always long and often happen in night; while, the calls with colleagues are nearly all short and often happen in daytime.

Even without group information, we still can figure out which one or which clusters of people that the person has contacted most, and what is the general time of the calls, or even the changes of life. Like in Fig. 3(d), person A has contacted most frequently with Peng (green) and Yan (cyan) in the month; and most of the calls are rather long and often happen in night or noon time implies A’s personality to connect with special friends.

As shown in Fig. 3(b), person B has three frequent contacts, including the largest yellow one. The calls from the contact are all short but very compact and most of which distributes around the lunch time, supper time and some in bedtime. From the above characteristics, we guess which has been proved that the contact is her boyfriend, that they work in the near place and often have meals together.

Different from A and B, person C is much less social. During working hours, the person almost has no calls. Apart from the most frequent contact (yellow) and some other sporadic contacts, the person seldom makes calls. Based on the above observation, we have enough reasons to believe that the contact (yellow) is a special friend of C. Actually, as one of C’s best friends the contact was in hospital all the month, which explains all the circumstances.

Moreover, if users are interested in a specific contact or a call, they can click on the object to observe the detailed information. As shown in Fig. 2, by clicking on one of the contacts, user can observe that the contact has 19 calls and in a total of 304 minutes; the connections in outer circle represent the detailed temporal information of each call. As shown in the same figure, user also can select one of the calls to observe the detailed information (in Fig. 2, the selected call is highlighted).

5.2.3 Particular data discovery

We have observed some particular data. Like in Fig. 3(a), different from other yellow calls imply for colleagues, a call is long and happen in evening. After clicking to see the specific information, the user realizes that the call is from Ding, who is a former colleague and one of his best friends. Also, in Fig. 3(b), user is curious with a long green arc, which is confirmed as from her sister in distance, who has contacted not too often but has much to communicate. With the feedback
operation of TelogViz, users can easily record the oblivious information in the past month, which can also be considered as a new diary format.

Those who don’t contact too often but have long calls are also the interest of users, since the calls involved usually have significances in their lives. The contacts are often displayed in TelogViz as large triangles, squares or pentagons which are quite easy to notice. Like the contact of A, Ding, who we have mentioned is found in Fig. 3(a) as a well-marked yellow pentagon. And in Fig. 3(d), a large square is pretty noticeable at the 21:00 direction.

Apart from the above, a relatively small yellow (colleague) contact in Fig. 3(a) is quite interesting, since the polygon is still round even when zooming in the whole view. After clicking on the polygon, detailed information shows that the contact has more than 20 calls, but almost all within one minute. Actually, the colleague is a newcomer with many simple questions about work and the surroundings.

5.3 Discussion

From the experiments we present above, the effectiveness of our design is obvious. Time and duration of each call is clearly visualized in Hollow Spiral Model, and periodicity can be easily discovered. With polygon representation and retinal variables like size, location, shape and color, the overall information based on each contact including calling times, whole calling length and general calling time can be easily understood in center of the hollow spiral. Furthermore, it is efficient and convenient to connect each call with the corresponding contact with interactions.

We could discover much information hidden in the cellphone communication log using TelogViz. Daily routines of the log owner can be figured out even including some cellphone using habits; various social behaviors can be easily observed or deduced from the tool; and some particular data can be obviously discerned. We could actually understand distinct features with several datasets of the same person in different months to observe the changes happening in his life.

On one hand, the design is useful for users to get an overview of the previous life. Users may benefit from the visualization to reflect upon themselves whether they should pay more emphasis on communication with family or friends, or to get a clearer understanding of the relationship with a special friend, or just to remember and record some particular events happening in their lives. On the other hand, the visualization can be available for polices to quickly invest some suspects. Polices may get the cellphone communication log of a suspect from mobile operator. With TelogViz to discover the suspect’s daily routines and social relationships can certainly be helpful to solve the case or arrest the suspect.

Furthermore, in the process of our experiments, users feel surprised to have their daily schedule or other information discovered; and most users evaluate our tool is interesting and impressing.

6 Conclusion and Future Works

In this paper, we have proposed a novel design to visualize the cellphone communication log. Our design is based on a Hollow Spiral Model, with temporal information represented like an analog clock. We map the calls as several colored arcs and the contacts as polygons with retinal variables. We also implement a tool named TelogViz to enrich our design with several interactions.
studies demonstrate that our scheme is interestingly efficient to explore a person’s daily routine and social relationships from the log file. In the process of our experiments, we also discover some surprised particular data, which is significant. On the whole, our design is useful for users to get an overview of the previous life, and could be helpful for police to quickly invest some suspects.

In future work, we present two directions for further researches. One is that our Hollow Spiral Model and the whole design of TelogViz can be applied to other dataset with temporal attributes, like the twitter data, MSN communication log, etc. The other one is that, besides the attributes we have utilized in cellphone communication, there still remains other information like the location of a call, the expense of each call, the message sending or the network flow usage information. Visualization based on some tasks involving the above information can also be useful and interesting.

References