

Dynamic Data Visualization for a Water Quality Monitoring Network

Xuan Chen^{1,a}, Xiaohui Zhu^{*1,2,3,b} and Yong Yue^{1,c}

¹Department of Computer Science and Software Engineering, Xi'an Jiaotong-Liverpool University
Suzhou, 215123 China

²Department of Computer Science, University of Liverpool, Liverpool, L69 3BX United Kingdom

³Department of Computer Science and Technology, Nantong University, Nantong, 226019 China

^axuan.chen13@student.xjtlu.edu.cn, ^bxiaohui.zhu@xjtlu.edu.cn, ^cyong.yue@xjtlu.edu.cn

Keywords: Data visualization, geography locations, screen coordinates, monitoring network

Abstract. Data visualization can support statisticians to analyze vast volumes of data more easily and efficiently. We proposed a novel data visualization algorithm for water quality monitoring network to process water quality data in real time. Our algorithm can analyze vast water quality data as well as dynamically visualize the pollution diffusion process. We achieved this algorithm based on a MVC web application. Longitudes and latitudes of all the monitoring locations are converted to screen coordinates and visualize the monitoring network. When a pollution event occurs, we use different colors to indicate different pollution levels. The color of river segments can be automatically changed in real time to represent the pollution diffusion processing as well as the pollution level in different river segments.

Introduction

Fresh water is one of the most vital resources for human beings. However, more and more pollutants are polluting rivers and threatening ecosystems. With the rapid development of computer science, network communication and sensor technology, remotely monitoring water quality for river systems is becoming one of the most efficient efforts to protect water environments. More and more automatic water quality monitoring stations are being constructed. For example, more than 300 automatic water quality monitoring stations have been deployed in Jiangsu Province in recent years. These online stations can automatically analyze water quality and transfer water quality data with a regular time interval. These monitoring stations produce a vast of water quality data. It is a huge challenge for analyzers to manually look through all data and spot trends and relationship from these data. Data visualization is one of useful techniques to help us to understand numerical data. It has been widely used in data science such as data mining, financial analyzing and scientific research. Bresciani and Eppler conducted an empirical study where analysts worked in groups to complete tasks with the help of sub-optimal, optimal or no visualizations [1]. It was found that compared to those who worked without visualizations, when analysts worked collaboratively using visualizations, their efficiency and accuracy of results increased. Chi developed a reference model for information visualization interfaces from data to visual form [2]. InfoZoom is a visualization system [3]. It can zoom into information subspaces by double-clicking on attribute values, or sets/ranges of values. Unlike InfoZoom, Spotfire is a smart visual analytics tool coupled with artificial intelligence and data discovery toolset [4]. It offers several types of mostly familiar visualizations, including scatterplots, bar charts, pie charts, graphs, parallel coordinates, trellises, etc [5,6]. However, these visualization systems have several drawbacks. For example, Users must plan what variables to use and how to visualize them beforehand. TileMill is one of the latest online software for cartographers to quickly and easily design maps for the web using custom data [7]. Although it is a powerful design studio, it requires several other types of applications to install with it, which needs large space. Based on the principle of visualization identified by Frankel and DePace, many data visualization applications have been developed [8].

This paper proposes a novel data visualization algorithm to model and visualize the water quality monitoring network. When a pollution event occurs, the river segments will be dynamically assigned

with different colors to indicate different pollution levels at different river segments. Based on the continuous water quality data transferred from remote water quality monitoring stations in real time, we can obtain the pollution diffusion status and dynamically visualize the pollution levels. We develop a MVC based web application to simulate the visualization process and experimental results show that our algorithm can dynamically demonstrate the pollution process.

Infrastructure of remote water quality monitoring system

Traditional remote water quality monitoring networks consist of four parts: (1) monitoring stations with detection sensors and/or biochemical analyzer instruments; (2) data transmission unit (DTU) to transfer data from sensors to remote data center via internet and 3G/4G mobile network; (3) remote data center to receive and store water quality data; (4) software to manage all the devices and data. All the monitoring stations are deployed nearby the river network (sensors can be deployed directly in a river). Monitoring stations can analyze water quality and get water quality data with a regular time interval. After that, the DTU connects to the remote data center via mobile network and internet and transfer all the water quality data to remote data center. At last, the data center accepts all the water quality data from different water quality monitoring stations and store into a database.

Visualization algorithms to model water quality monitoring network and pollution diffusion

One of most common spatial data visualization techniques is map. Dynamic maps can seamlessly interact with data in real time and automatically update the map according to the feedback. It can provide us a much easier way to analyze and understand the original data.

(A) Monitoring network visualization

As we know, the 2D canvas uses a standard Cartesian coordinate system with the origin point (0, 0) at the top left. One unit in the coordinate system is equal to one pixel on the canvas. Since the coordinate system in canvas is different from the geography location system in real world. We should convert geography locations of all the monitoring stations to screen coordinates respectively. The conversion algorithm is shown in Figure 1. In the coordinate conversion procedure, Ln and Lt is an array of longitudes and latitudes respectively, which represents the geography locations for all the monitoring stations. $CanvasHeight$ and $canvasWidth$ indicates the canvas height and width respectively. Min and max functions are used to get the maximal geography scope of the monitoring network. The algorithm maps the top-left point (0, 0) to the minimal geography locations ($minLn$, $maxLt$) and also maps the right-bottom point ($canvasHeight$, $canvasWidth$) to the maximal geography location ($maxLn$, $minLt$). Based on these calculation, we can map all the geography locations to screen coordinates and can visualize the monitoring network on a web page.

Procedure `Coordinate_Convert` (longitude [] ln , latitude [] lt , int $canvasHeight$, int $canvasWidth$)

```

int []  $locX$  = new int [ $ln.length$ ];
int []  $locY$  = new int [ $lt.length$ ];
 $minLn$  = min ( $ln$ );
 $maxLn$  = max( $ln$ );
 $minLt$  = min ( $lt$ );
 $maxLt$  = max ( $lt$ );
 $hUnit$  =  $canvasHeight$  / ( $maxLt$  -  $minLt$ );
 $wUnit$  =  $canvasWidth$  / ( $maxLn$  -  $minLn$ );
for (int  $i=0$ ;  $i<=ln.length$ ;  $i++$ )
{
     $locX[i]$  = ( $ln[i]$  -  $minLn$ ) *  $wUnit$ ;
     $locY[i]$  = ( $lt[i]$  -  $minLt$ ) *  $hUnit$ ;
}
Return ( $locX$ ,  $locY$ );

```

End Procedure

Fig. 1. Geography coordinates conversion procedure

Table I. Geography monitoring locations and corresponding screen coordinates

Location	Longitude & Latitude		Coordinates	
1	120.95259	31.348817	962.950	155.915
2	120.93377	31.316646	868.875	316.769
3	120.918	31.319	790.000	304.999
4	120.912	31.329	760.000	254.999
5	120.895	31.318	675.000	309.999
6	120.875	31.319	575.000	304.999
7	120.862	31.311	509.999	345.000
8	120.850	31.288	450.000	459.999
9	120.82987	31.298772	349.349	406.140
10	120.787	31.302	135.000	390.000
11	120.76715	31.304625	35.759	376.874

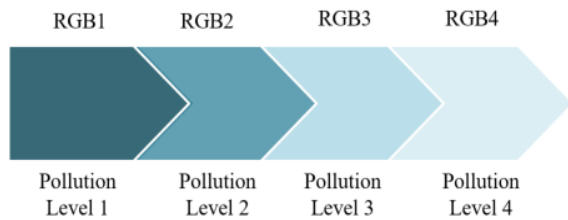


Fig. 2. Different pollution levels with different colors



Fig. 3. Geography monitoring locations on Baidu map

(B) Pollutant diffusion visualization

To visualize the pollution diffusion procedure, we calculate the pollutant concentration in real time at each monitoring location based on the water quality data transferred from remote monitoring stations. Colors of monitoring segments are changed according to the pollutant concentration at each monitoring location. A RGB color model is used here to represent the pollution level. As the pollutant concentration changes smoothly from high to low. A location near the pollutant source has a higher pollutant concentration than other locations. We use different colors for different river segments according to different pollutant concentrations at different monitoring locations. The relationship between pollutant concentrations and colors is determined by Equation 1.

$$\begin{aligned}
 levelR &= 210/(upBound - lowBound) \\
 levelG &= 170/(upBound - lowBound) \\
 levelB &= 225/(upBound - lowBound)
 \end{aligned}
 \tag{1}$$

where *levelR*, *levelG* and *levelB* are the values of RGB color model, *upBound* and *lowBound* represent the upper and lower bound of pollutant concentration. For different types of water quality data, the boundary of pollutant concentration is different. The transform equation indicates that higher pollution level is presented with dark color and lower pollution level with light color, which is shown in Figure 2.

(C) Realization and simulation

A MVC model based web application is developed to visualize the water quality monitoring network. We selected 11 monitoring locations in a river and get geography locations from Baidu map (Figure 3). The longitudes and latitudes for these monitoring stations are in Table I. We use a coordinate conversion procedure above to convert all the geography locations to screen coordinates and draw the monitoring network as Figure 4 shows. When a pollution event occurs at a monitoring location, it can be detected immediately by monitoring devices and the pollutant concentration will be calculated. Then, the polluted river segments will be drawn in darker colors (Figure 5). The color of river segments can be continuously changed with pollutant concentration.



Fig. 4. Map of original monitoring network



Fig. 5. Map of monitoring network with a pollution event

Conclusion

We proposed a novel data visualization algorithm for water quality monitoring network. It can dynamically visualize the monitoring network as well as the pollution diffusion process with different colors. In future, we will apply this algorithm to a real water quality monitoring network to visualize the pollution diffusion process and further test its efficiency and stability.

Acknowledgements

This work was partly supported by the Natural Science Foundation of Jiangsu Province (BK20151245), the Natural Science Foundation of Huai'an City (HAG2015007), the Natural Science Foundation of Nantong City (MS12016048) and the Summer Undergraduate Research Fellowship of Xi'an-Jiaotong Liverpool University.

References

- [1] Bresciani S, Eppler M J. The benefits of synchronous collaborative information visualization: Evidence from an experimental evaluation[J]. *IEEE transactions on visualization and computer graphics*, 2009, 15(6).
- [2] Chi E H. A taxonomy of visualization techniques using the data state reference model[C]// *Information Visualization*, 2000. InfoVis 2000. IEEE Symposium on. IEEE, 2000: 69-75.
- [3] Spenke M, Beilken C. Discovery Challenge: Visual, Interactive Data Mining with InfoZoom—the Financial Data Set[C]// *Workshop Notes on "Discovery Challenge"*, 3rd European Conference on Principles and Practice of Knowledge Discovery in Databases, PKDD. 1999, 99: 33-38.
- [4] Ahlberg C, Wistrand E. IVEE: An information visualization and exploration environment[C]// *Information Visualization*, 1995. Proceedings. IEEE, 1995: 66-73.
- [5] Kobsa A. An empirical comparison of three commercial information visualization systems[C]// *Information Visualization*, 2001. INFOVIS 2001. IEEE Symposium on. IEEE, 2001: 123-130.
- [6] Mark G, Carpenter K, Kobsa A. A model of synchronous collaborative information visualization[C]// *Information Visualization*, 2003. IV 2003. Proceedings. Seventh International Conference on. IEEE, 2003: 373-381.
- [7] M. Zastrow, Data Visualization: Science on the Map, Available: <http://www.nature.com/news/data-visualization-science-on-the-map-1.17024> (2015).
- [8] Patterson J, Bickel A. Communicating local relevance of ocean observations: Integrating real-time ocean sensor data visualizations, online communications, and ocean issues to engage public audiences[C]// *OCEANS 2016 MTS/IEEE Monterey*. IEEE, 2016: 1-4.