Computers, Environment and Urban Systems xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Computers, Environment and Urban Systems

journal homepage: www.elsevier.com/locate/ceus

Group-user access patterns and tile prefetching based on a time-sequence distribution in Cloud-based GIS

Rui Li[a,b,*], Jiapei Fan[a], Huayi Wu[a,b,*], Jie Jiang[c], Guangsheng Dong[a]

[a] *State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan, Hubei 430079, PR China*
[b] *Collaborative Innovation Center for Geospatial Technology, Wuhan, Hubei 430079, PR China*
[c] *National Geomatics Center of China, Beijing 100830, PR China*

ARTICLE INFO

Keywords:
Gaussian mixture model
Time-sequence
Spatiotemporal
Prefetching
GIS

ABSTRACT

Group-user intensive access to spatial data is temporally aggregated, while accessed content is spatially correlated, creating performance fluctuations and service bottlenecks in existing cloud-based GISs (CGISs). Current solutions are not scalable as they fail to consider the quantitatively intensive spatiotemporal access patterns that could make a CGIS adapt more accurately and quickly to aggregative and bursting group-user accesses. In our research, we tackled these performance fluctuations and service bottleneck problems with a novel Gaussian mixture model (GMM) representing the short-term aggregation of tile accesses. We constructed a new tile prefetching and cache prefetching control strategy that represents quantitatively the spatiotemporal correlations of accessed tiles to optimize CGIS service performance.

The GMM solution that we proposed quantifies variation in tile access popularity to predict access probabilities for hotspot tiles based on the observed multipeak characteristics and uniform variations in the time-sequence distribution found in group-user accesses. Cumulative access probability was used to quantify the spatial correlation and locality of tile accesses. In our approach, a hotspot tile and the neighboring tiles in a given spatial area are prefetched in an effective cache prefetching control strategy based on resource utilization. Experiments have demonstrated that the proposed GMM for tile access probability (TAP-GMM) accurately predicts hotspot tiles and their access popularity distributions by prefetching tiles through a stable service with enhanced scalability.

The contributions of this study are twofold. First, we present the GMM for short-term variations of tile access popularity, fixed spatial correlation, and temporal locality of accessed tiles in group-user intensive access behavior. Second, our study on access behavior makes the resulting prefetching strategy adaptable to varying intensities and bursting patterns in group-user access for optimizing CGIS performance. The prefetching strategy is stable and reduces the service resource consumption as it conforms to quantitatively measured access patterns.

1. Introduction

Many GISs are moving to a cloud platform with the rapid development of cloud technology and the popularization of the Internet and mobile networks (Mann, 2011). These cloud-based GISs (CGISs) have helped transform GIS from an exclusive, professional application to a popular public tool for understanding spatial and geographic information, and the number of CGIS users is increasing with each passing day (Bhat, Shah, & Ahmad, 2011; Yang, Yu, Hu, Jiang, & Li, 2017). Thus, the provision of spatial information services on cloud platforms with high reliability, availability, and scalability for large-scale highly intensive concurrent user access has become a research hotspot (Li, Shan, Shao, Zhou, & Yao, 2013; Yang, Huang, Li, Liu, & Hu, 2017). The process of quickly transferring and displaying spatial data on GIS clients is also a major research focus (Yang & Huang, 2014).

Existing studies address the service bottleneck problem generated by intensive concurrent user accesses. Some have focused on improving the flexibility of cloud platforms (Buck, Watkins, LeFevre, Ioannidou, & Maltzahn, 2011; Dumbill, 2012). Other researchers have investigated the establishment of an appropriate index mechanism (Wang, Wang, & Zhou, 2009; Yang, Wu, Huang, Li, & &Li, J., 2011) to accelerate spatial data queries. Spatiotemporal data models can dynamically express geographic processes quickly for multiple users (Li, Yang, Guan, & Wu, 2014). Spatial information can be conveyed by a global content

* Corresponding authors at: State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan, Hubei 430079, PR China; Collaborative Innovation Center for Geospatial Technology, Wuhan, Hubei 430079, PR China.
E-mail addresses: Ruili@whu.edu.cn (R. Li), wuhuayi@whu.edu.cn (H. Wu).

https://doi.org/10.1016/j.compenvurbsys.2017.12.002
Received 16 January 2017; Received in revised form 18 December 2017; Accepted 22 December 2017
0198-9715/

delivery mechanism (http://aws.amazon.com/cloudfront/) to provide parallel services. These solutions, however, are not adaptive to spatio-temporal variation in intensive user access and do not permit a CGIS to be more scalable and stable.

The concurrent aggregation characteristics found in massive user accesses, however, could help to serve as a basis for optimized accelerating strategies for CGIS services to address the challenges emerging from concurrent access (Li, Zheng, Xie, Chen, & Liu, 2008; Yang, Wong, Yang, Kafatos, & Li, 2005). User activities in CGIS are social and aggregative; users tend to access tiled spatial data (tiles) in a few centralized geographic areas in specific timeframes based on social and subjective factors (Mountrakis & Gunson, 2009; Sedda, Atkinson, Filigheddu, Cotzia, & Dettori, 2011; Achakeev, Seeger, & Widmayer, 2012; Steiger, Albuquerque, & Zipf, 2015). At the same time, because tiles have temporal and spatial attributes, studies on human behavior patterns evident in CGIS must consider both temporal and spatial characteristics of access behaviors (Chai & Zhao, 2009; Kwan, 2004; Raubal, Miller, & Bridwell, 2004). However, it is necessary to analyze these access characteristics qualitatively and quantitatively to better understand the spatiotemporal variation in aggregative service requests.

A qualitative analysis of massive user access logs showed that group-user aggregative access behaviors have long-term and short-term patterns. Access to tiles has temporal or spatial locality; hotspot tiles are accessed at high probabilities in a short-term timeframe, while at the same time, neighboring tile access patterns are aggregative (Baentsch, Baun, Molter, Rothkugel, & Sturm, 1997; Park & Kim, 2001; Perkowitz & Laszlo, 2003). However, these studies did not describe access behaviors quantitatively, and few studies have considered correlation and variation trends evident in the spatial and temporal characteristics in group-user access behaviors. These spatiotemporal variations result in aggregative and bursting access patterns in a CIGS.

Group-user access to a CGIS has social characteristics and fixed spatiotemporal access patterns. The tendencies of group-user access to hotspot tiles change over time and are related to the spatial attributes and dependencies of tiles. Moreover, access patterns directly affect resource demands and CGIS service performance (Huang, Yang, & Benedict, 2013; Shaw, Yu, & Bombom, 2008; Xia, Yang, Gui, Liu, & Li, 2014). Qualitative descriptions of group-user access can reflect access characteristics and tendencies, but these descriptions are general and abstract. These abstract descriptions cannot provide a basis for a CGIS to adapt accurately and quickly to aggregative and burst accesses from group-users. A quantitative representation of intensive access patterns, however, can provide such a foundation. A quantitative access pattern used in performance optimization strategies can achieve service stability, reducing resource consumption and solving the CGIS performance problems caused by group-user intensive access. This is the motivation and basis for our study of quantitative group-user access patterns and the related prefetching strategy to optimize CGIS performance.

Our work includes two parts. First, we built a Gaussian mixture model (GMM) to express the multiple peaks and short-term uniform variation features found in the temporal distribution of tile access popularity to predict hotspot tiles. Second, we prefetched these hotspot tiles and neighboring tiles using the proposed GMM, considering the spatial aggregation in access to spatial correlated tiles under prefetching controls. These tiles were cached on a cloud cluster-based caching system (CCCS) designed specifically to support a CGIS. We take advantage of the spatiotemporal locality and short-term variation characteristics in group-user access to tiles that place burst and concentrated service loads on a CGIS. Therefore, this approach can provide a cache prefetching and resource control strategy to better exploit cloud caching resources.

The remaining article consists of six parts: (1) a review of related work on user access patterns and tile prefetching strategies based on user access patterns; (2) a temporal distribution to represent the variation for tile access popularity; (3) a GMM to accurately quantify tile access probability distributions; (4) a proposed prefetching and prefetching control strategy based on our GMM, which considers the stable spatial access aggregation among neighboring tiles; (5) experiments to validate the proposed GMM and the prefetching strategy; and (6) a summary of contributions and the potential applications of this work.

2. Related work

The quantitative study of intensive user access patterns will expose regularities that can provide an empirical foundation for CGIS performance optimization strategies. This section summarizes and discusses specific research relevant to our study: spatiotemporal aggregation in intensive user access, spatiotemporal correlation of intensive user access, and prefetching strategies based on user access behavior.

2.1. Spatiotemporal aggregation in intensive user access

Group-user access behavior in a CGIS has temporal aggregation regularity with social attributes. A statistical analysis of spatial data access frequency showed that access requests to tiled spatial data follow a power law distribution, where 80% of users request only 20% of the available tiles (Talagala, Asami, & Patterson, 2000; Fisher, 2007). By analyzing communication traffic and accessed tiles of CGIS, Wang, Pan, Peng, and Li (2010) and Li, Feng, Xu, and Pan (2012) confirmed that the access pattern of user request to tiles follows a Zipf-like distribution, a kind of power law distribution. A power law distribution is a heavy-tailed distribution in which accumulated group-user accesses are uneven and aggregative.

Group-users tend to access a few spatially centralized geographic areas of interest. The access intensities of these areas of interest follow a heavy-tailed distribution with a high rate of access penetration (Li, Shen, Huang, & Wu, 2015). Hotspot tiles are aggregative and vary with periodic characteristics in a short-term timeframe; these characteristics are related to the pace of regional user life and are responsive to hot news items. Access popularity of hotspot tiles therefore changes over time, but Li et al. did not provide a quantitative representation of this spatiotemporal aggregation regularity.

Aggregative access creates performance fluctuations and service bottlenecks in actual CGISs. A power law distribution quantitatively expresses group-user access regularities, reflecting a long-term and stable temporal accumulation of access to tiles. This distribution, however, does not reflect access hotspot variations over time or account for short-term regularities that also have periodical distributions. A power law distribution indeed does not represent spatiotemporal correlations in access to tiles. Thus, a quantitative representation of spatiotemporal aggregation and temporal variation could provide an exact basis for a performance optimization strategy, which is our research focus.

2.2. Spatiotemporal correlation in intensive user access

Tobler's First Law (Tobler, 1970; Miller, 2004) describes that geographical things or their attributes are intercorrelative in spatial distributions with aggregations and regularities. User access behaviors in a CGIS therefore might also have spatial and temporal correlations. Temporal correlation in tile accesses always reference time intervals, and the frequency of accesses to a tile has a localized temporal range. Hotspot tiles have high probabilities to be accessed in the near future; the time interval between the current time and the last accessed time is used to express this temporal correlation in hotspot tile accesses (Bell et al., 2007; Lee et al., 2001). As access popularity is characteristically short term with a high temporal correlation and changes over time, tracking access popularity has been a means to capture the temporal correlation in user access. Such methods include the recent usage frequency method (Kang, Ma, Tong, & Liu, 2012), the age-based cooperative method (Ming, Xu, & Wang, 2012), and others. However,