The QuakeSim Web Portal Environment for GPS Data Analysis

Robert Granat Jet
Propulsion Laboratory
California Institute of
Technology
4800 Oak Grove Dr.
Pasadena, California 91109
robert.granat@jpl.nasa.gov

Xiaoming Gao Department
of Computer Science
Indiana University
501 N. Morton St.
Bloomington, Indiana 47404
gao4@cs.indiana.edu

Marlon Pierce Community
Grids Laboratory
Indiana University
501 N. Morton St.
Bloomington, Indiana 47404
mpierce@cs.indiana.edu

ABSTRACT

We present a web portal environment for analysis of GPS displacement data that supports interactive scientific exploration at both a network-wide (macro view) and individual station (micro view) level. Underlying this environment is a hidden Markov model based analysis method that allows segmentation of the GPS time series into statistically meaningful classes, as well as a web services infrastructure that connects the data to the method and user interface. This user interface is primarily map-based, enabling regional understanding of surface displacement, but also supports other modes of interaction to facilitate understanding of the data. We demonstrate this environment using data from GPS member stations of the Plate Boundary Observatory located in California, and present some sample analysis results.

Keywords

HMM, GPS, Web Service, Web Portal, Earthquake, Deformation, Transient

1. INTRODUCTION

The QuakeSim project [19] is an effort to address current and future needs of the geophysics community by integrating sensor measurements, earthquake modeling, simulation, and forecasting through a web portal and web services environment. In this work we focus on the Global Positioning System (GPS) analysis and integration portion of QuakeSim.

GPS sensor networks have become increasingly important in scientific studies of the earthquake cycle and fault system behavior. GPS instruments measure displacements at specific points on the earth's crust; data streams of high-precision measurements are available on a daily basis from many stations. These measurements have taken on added importance with the advent of synthetic aperture radar (SAR) and interferometric synthetic aperture radar (InSAR) measurements of surface displacement taken from satellites and unmanned aerial vehicles (UAVs). These radar measurements provide high spatial resolution but poor time resolution, so GPS measurements are essential for filling in the gaps. Several satellite and UAV based radar campaigns are currently underway, with more being planned. In particular, NASA's Deformation, Ecosystem Structure, and Dynamics of Ice (DESDynI) mission will carry an L-band InSAR instrument specifically designed to measure crustal deformation, and is expected to operate on an eight day repeat cycle. QuakeSim's GPS analysis approach is therefor designed with integration into the larger effort to understand crustal deformation as our persistent and long term goal.

The goals of this effort in the short to intermediate term are threefold. First, we want to perform network health analysis; this includes detecting anomalous non-physical signals, monitoring network outages, identifying poorly performing stations, and catching processing errors. Second, we want to detect deformation signals associated with co- or post-seismic stress changes on or around earthquake faults. Third, we wish to detect so-called "transients," unusual signals that result from non-seismic geophysical processes. Such signals are rare, but slow-slip events characterized by such transients have been identified in Cascadia [18, 23], Peru [16, 17], and Kantou [10, 11], emphasizing that complex geophysical processes cannot be understood through co- and post-

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
ESSA '09 San Francisco, California USA
Copyright 200X ACM X-XXXXX-XX-X/XX/XX ...$5.00.

seismic deformation alone. In addition, the existence of more obvious known transients suggests the possibility that far more subtle transients have been overlooked to date in the existing data; the detection of such would represent a major contribution to the field. Although these transients are poorly characterized, we expect that they are regional, rather than extremely local, and this informs our approach.

These goals suggest the following set of requirements. First, to meet our long term goals, our methods should be both extensible and modular, so that we can seamlessly incorporate new analysis methods and data sources. Second, to detect regional signals our approach must incorporate geographical information and the distribution of the member stations of the network. Third, we must be able to accommodate a continually updating data source, incorporating the new information in our analysis. Lastly, to maximize the benefit to the geophysics community, as well as capitalize on the knowledge of the domain experts, our system should be interactive, making the scientist part of the analysis process. This means that our approach should be responsive (return results quickly), robust (return results that are consistent across experiments), and provide multiple tools to use in interpreting the data and analysis results.

At a high level our approach to addressing these requirements is as follows: a statistical modeling approach is used to segment individual GPS time series and results are collated across the network so as to detect regional signals; the results of this analysis are presented to users via a web portal interface. Underpinning the system is a web services infrastructure that connects the data to the applications and the user, manages workflow, and handles fundamentals such as data processing and security. We detail this approach in subsequent sections: Section 2 discusses the analysis algorithm, Section 3 the web services infrastructure, and Section 4 the web portal environment. Lastly, we show some preliminary results of analysis conducted through the portal in Section 5.

1.1 Data Source

Our data is drawn from 442 GPS member stations of the Plate Boundary Observatory (PBO), with data archives at the Jet Propulsion Laboratory (JPL) [12] and the Scripps Orbit and Permanent Array Center (SOPAC) [13]. This is the densest GPS network in the United States, covering a highly active seismic and tectonic region. Data is collected by stations at 1 or 2 Hz, depending on station type, and then integrated on a daily basis to provide a high-precision displacement measurement in three dimensions. Due to necessary orbit corrections production of the final observational data is delayed by approximately two weeks.

2. DATA ANALYSIS ALGORITHM

In this problem domain, the underlying physical system is both noisy and poorly understood, and the observed GPS measurements are drawn from a system whose driving forces are derived not only from the physical processes of the solid earth but also from external factors, such as atmospheric effects and human activity. As such, we eschew physical modeling approaches in favor of statistical modeling. In particular, we employ hidden Markov models (HMMs) for our analysis.

Fitting an HMM to time series allows us to describe the statistics of the data in a simple way that ascribes discrete modes of behavior to the system. By matching incoming data against the statistics of previously learned modes, we can perform classification according to the best match. In addition, it is possible to perform signal detection across the entire sensor web by detecting simultaneous mode changes; a significant number of mode changes across the network or within a certain sub-network is an indication of an event that is occurring over a wide geographical area.

Fitting an HMM to data is a non-linear and non-convex optimization problem. The most common approach to this problem is the expectation-maximization (EM) algorithm [22], an iterative method that works well for simple cases but often struggles to find consistent, good solutions in the presence of numerous local maxima. Multiple restarts with different random initializations can help, but still yield inconsistent results across experiments in many cases. For some applications, such as speech synthesis and protein sequence analysis, reliable HMM fitting results can be achieved by using a priori information to encode constraints that reduce the number of free parameters [14, 6, 15, 5, 3, 25, 4]. For GPS data, however, this information is not available as the underlying geophysical system is not well understood. As a result, instead we use a variant of the EM method, regularized deterministic annealing EM (RDAEM), that has experimentally proved to yield consistent, high-quality results. Details of the method and experimental results can be found in [7] and [8]. We refer to the software implementation of this method as RDAHMM.

The RDAEM approach is similar to the deterministic anneal approach to HMM fitting employed by Rose [24], with two key differences. First, it is an unsupervised approach that seeks to maximize the log likelihood of the model given the observations, and second, it incorporates several regularization terms that help to avoid unfavorable local maxima. The end result is a method that generates robust, repeatable results on a wide variety of geophysical data sets, at the cost of some computational overhead. For a full analysis of the algorithm results and performance, as well as comparisons with other methods, see [7].

3. WEB SERVICES INFRASTRUCTURE

3.1 Service Infrastructure and Workflow

Underpinning our web portal environment is a flexible and modular web services infrastructure. This infrastructure is designed to support interactive, and thus by necessity, responsive, analysis of data from large numbers of GPS stations. In addition, it can support constant real-time updates of analysis results as new data becomes available. The infrastructure and workflow supporting the RDAHMM application and web portal are shown in Figure 1.

Key to infrastructure is a driver service called "Daily RDAHMM Runner Service" (called "Runner Service" for short), which directs the RDAHMM analysis of each station in the network. To accommodate continuous updates to the GPS data, this service is scheduled to run on a daily basis. Every day, the Runner Service takes the following actions for each station:

1. Check if an RDAHMM model has been built for this station yet. If yes, go to step 3; otherwise call the GRWS Query Service to get the station's displacement data for the time from 1994-01-01 (when the stations were first set up) to 2006-09-30.

2. Call RDAHMM Model Service to train a model for this station, using the displacement data obtained in step 1 as input. This training process fits a five state HMM to the displacement time series, classifying each displacement observation in the series as belonging to one of the five states. The five state model size was chosen empirically (see [8] for more details).

3. Call the GRWS Query Service to get the station's displacement data for the time from 2006-10-01 to "today" (the date the service is running).

4. Call RDAHMM Evaluation Service to do evaluation for this station, using the displacement data obtained in step 3 and the model built in step 2 as input. The evaluation process calculates the Viterbi optimal state sequence using the pre-trained HMM parameters. This state sequence is saved as a separate file, which is used later in step 6.

5. Call Graph Plotter Service to plot the station's displacement data on each direction (east-west, north-south, up-down) as time series from 1994-01-01 to"today." The state sequence of the observations is shown by coloring points on the plot according to their state assignment; each state is associated with a unique color.

6. Combine the analysis results from all stations and output an XML file that contains the information about which stations have changed state on which dates.

7. Based on the XML file created in Step 6, the Runner Service calls the Video Maker Service to make a movie that shows the day-to-day state change information for the GPS network from 1994-01-01 to "today". This movie will be described further in Section 4.

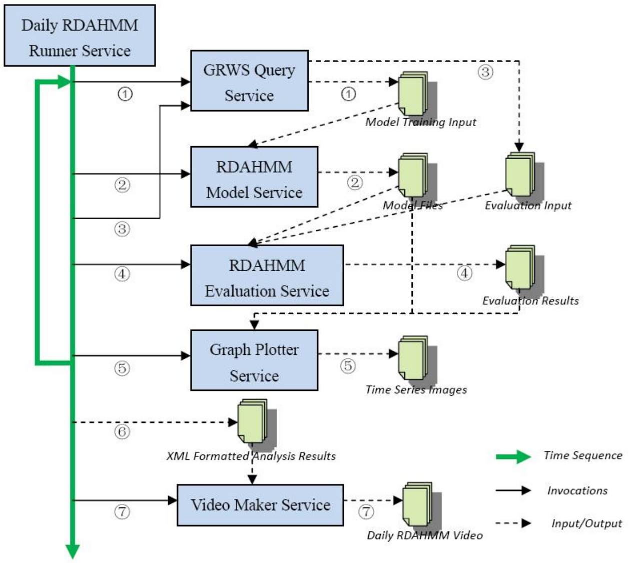


Figure 1: Service Infrastructure and Workflow

This infrastructure is flexible and can be easily extended to support new functionalities and data sources. New functions consuming the outputs of present services and generating new outputs can be added by introducing new invocations into the workflow. For example, the Video Maker Service and the Daily RDAHMM Result Service (which will be discussed in Section 4) were both added as needed by new requirements during the progress of the project.

Currently these services are only running for analysis of the SOPAC GPS solutions. Analysis for new data sources such as the JPL solutions, which use a different orbital error correction algorithm, can also be accommodated by adding a new invocation to the GRWS Query Service with a different set of parameters, and then applying the same subsequent sequence of service invocations to the displacement data obtained from the new invocation.

4. WEB PORTAL ENVIRONMENT

To enable interactive analysis on the part of the scientist, we built what we call the "Daily RDAHMM Portlet" in the QuakeSim [19] web portal. This provides an environment in which the scientist can investigate the results of the HMM analysis at both the network (macro-view) or individual station (micro-view) level. Since no single way of viewing the results may provide all the necessary insight, we provide a number of methods through which the scientist can investigate the results.

The relationship between this portlet and the web services is shown in Figure 2. Besides accessing the evaluation results generated by the web services, the portlet needs to invoke the Daily RDAHMM Result Service, which helps analyze the XML formatted results file and answer the portlet's requests about stations' state change and missing data information.

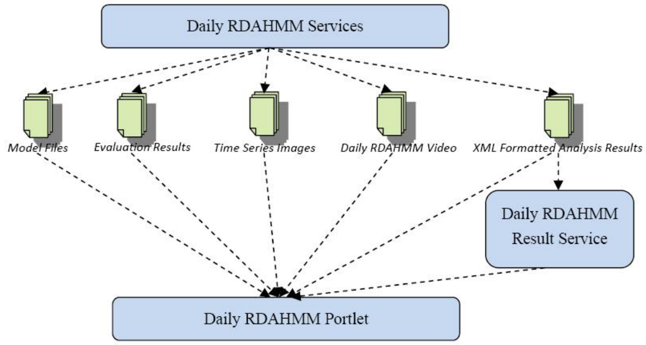


Figure 2: Connections between Web Services and the Daily RDAHMM Portlet

4.1 Macro-view Presentations

Our macro-view presentations provide several ways to access to analysis results from all the stations in the network. The first of these uses Google maps [1] to show the state change information for all stations in California on a particular date. Markers show the station positions on the map, with colors indicating the nature of the change at a particular station.

Green: the station did not change state on the selected date nor within the last 30 days before;

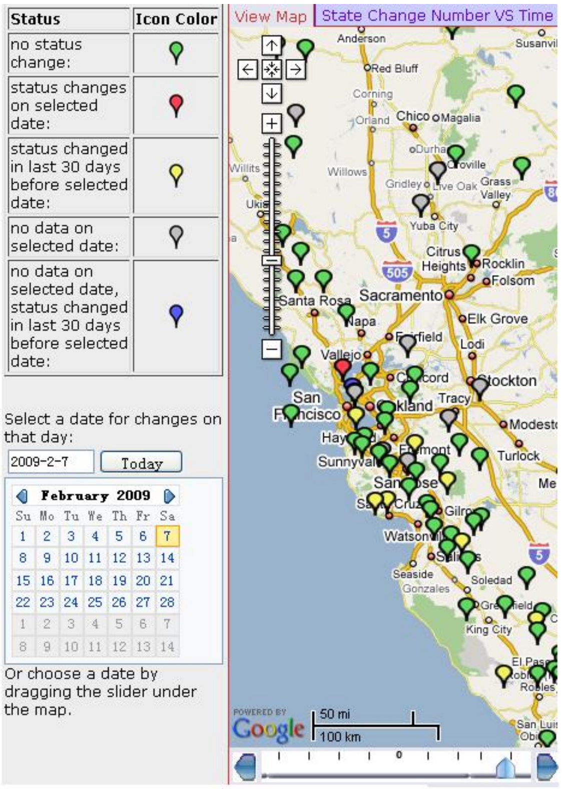
Red: the station changed state on the selected date;

Yellow: the station did not change state on the selected date, but experienced a state change within the last 30 days;

Gray: the station has no input data for the selected date, and has not changed state within the last 30 days;

Blue: the station has no input data for the selected date, but has experienced a state change within the last 30 days.

Users can select the date by using either a calendar or scroll bar (implemented using Yahoo widgets [2]). While the calendar enables the user to jump to a particular date, the scroll bar allows day by day exploration of the evolution of the network.



Status	Icon Color
no status change:	🟩
status changes on selected date:	🟥
status changed in last 30 days before selected date:	🟨
no data on selected date:	⬜
no data on selected date, status changed in last 30 days before selected date:	🟦

Select a date for changes on that day:

2009-2-7 Today

February 2009

Su	Mo	Tu	We	Th	Fr	Sa
1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
1	2	3	4	5	6	7
8	9	10	11	12	13	14

Or choose a date by dragging the slider under the map.

Figure 3: Google Map Presentation of the PBO Network

In order to inspect the state change process of all stations in the California area along a continuous time axis, the user can also view an "RDAHMM movie" made by the Video Maker Service from the portal. The video is composed of frames presenting the day-by-day evolution of activity on the California GPS network from 1994-01-01 to "today". Figure 4 shows a frame of this video. This enables the user to rapidly scan through time, looking for unusual regional signals that may indicate geophysical activity or network anomalies.

To better identify points in time in which the network is undergoing change, the portlet provides a plot of number of stations with state changes versus time (shown in Figure 5). Spikes in the plot indicate times in which a large number of stations have changed state on the same day. This can help focus investigation on particular points in time as well as identify anomalous "blips" that might be missed in the video.

Finally, to facilitate further analysis by the user, we provide a link to a data file which includes the displacement observations from all stations since 1994-01-01.

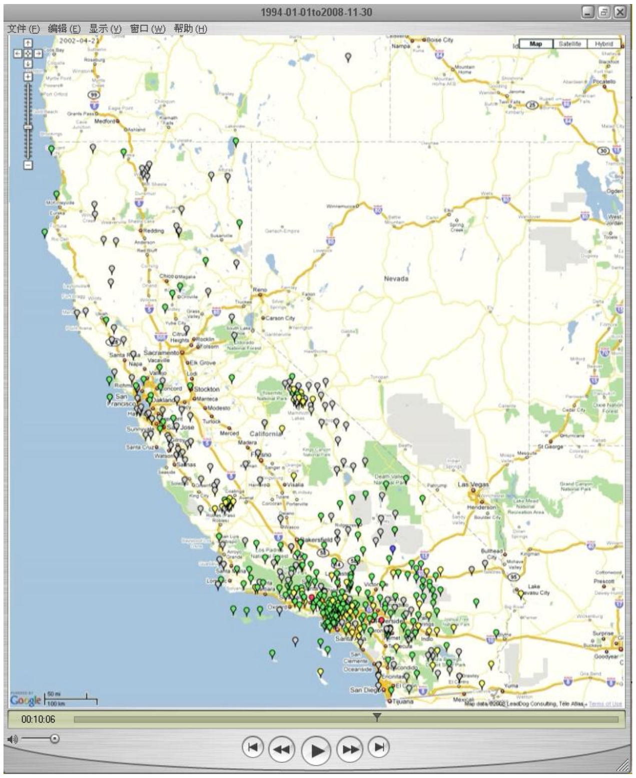


Figure 4: Daily RDAHMM Video

Since this is the exact data used by the RDAHMM application, the user can perform whatever analysis desired and compare with the analysis results provided by the portal.

4.2 Micro-view Presentations

Our micro-view presentations provide access to the analysis results for individual stations. Key among these are time series plots that show the displacement time series, color coded according to the state labels assigned to each observation. Figure 6 shows an example of the micro-view presentation of the station p489. The user can choose a particular station by clicking a marker on the Google map or selecting an item from a drop-down list of all stations. When a marker in the map is clicked, an on-line window will pop up, which shows the name and position of the corresponding station, as well as the plots of its displacement coordinates along each direction. The color coding of the observations provides a quick and easy way for the user to identify the state sequence and segmented portions of the time series, which may correspond to geological events.

The portal also shows 10 of the most recent state changes of the chosen station, which can help the user follow and monitor the activities of a specific station or area. If the user wishes to pursue further offline analysis, links make the chosen station's original displacement input data and HMM parameter files available for quick download.

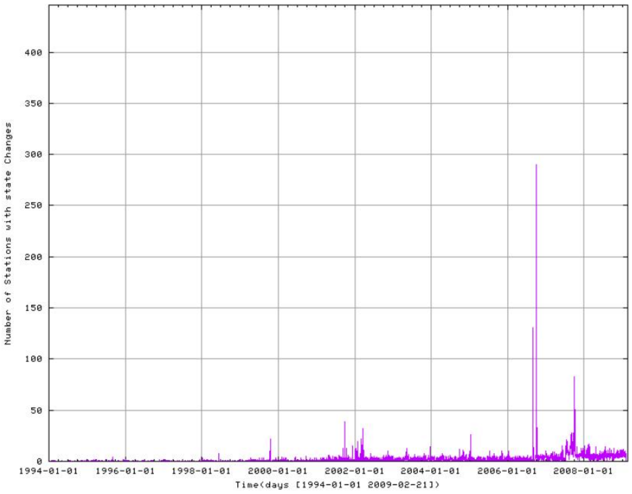
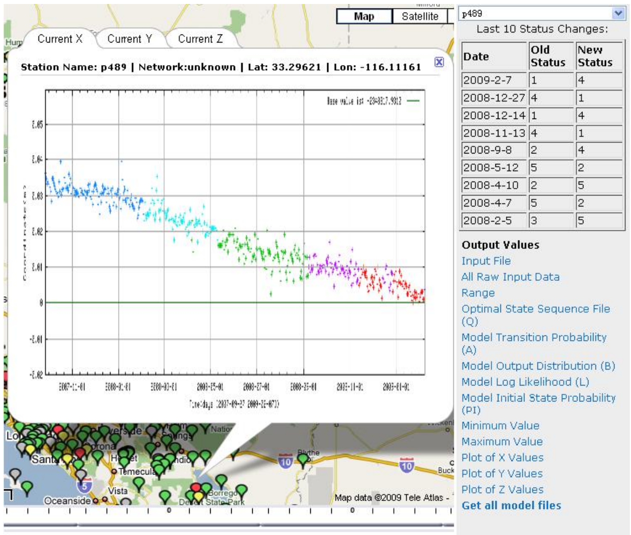


Figure 5: Plot of Stations with State Changes VS. Time



Station Name: p489 | Network:unknown | Lat: 33.29621 | Lon: -116.11161

p489

Last 10 Status Changes:

Date	Old Status	New Status
2009-2-7	1	4
2008-12-27	4	1
2008-12-14	1	4
2008-11-13	4	1
2008-9-8	2	4
2008-5-12	5	2
2008-4-10	2	5
2008-4-7	5	2
2008-2-5	3	5

Output Values
Input File
All Raw Input Data
Range
Optimal State Sequence File (Q)
Model Transition Probability (A)
Model Output Distribution (B)
Model Log Likelihood (L)
Model Initial State Probability (PI)
Minimum Value
Maximum Value
Plot of X Values
Plot of Y Values
Plot of Z Values
Get all model files

Figure 6: Micro-view Presentations of a Station

4.3 Performance Issues

The main challenge in implementing the web portlet comes from the large amount of computation required to process the station state change and missing data information. Since each station can have as many as hundreds of state changes over the study period, as well as hundreds of instances of missing data, calculating the proper colors for all station markers is a computationally intensive task. We tried four different strategies for distributing the computation workload between the browser side and server side:

Pure Javascript: Analyze the XML result file and store necessary data with client side Javascript; station marker colors are also computed with Javascript on the client side;

JSP + Javascript: Analyze the XML result file with server side JSP codes, and store necessary data at client side with Javascript; station marker colors are computed with Javascript at client side.

Managed Bean + Javascript: Analyze the XML result file and store necessary data with server side managed session beans; station marker colors are also computed by session beans.

Web Service + Javascript: Analyze the XML result file and store necessary data with a web service called using Javascript; station marker colors are also computed by web service.

Based on relative performance comparisons (see Table 1), we chose the web service + Javascript approach due to its good responsiveness in both loading time and map update time. Here "loading time" denotes the time needed to completely load and display the portlet page from the server when the portlet is accessed for the first time, and "map update time" denotes the time needed to calculate the proper colors for all station markers when a new date is selected. Since the default date is set to the most recent date when there is any input data for any station relative to the time when the page is loaded, the loading time also includes a period of "map update time" to calculate the colors for the default date. More detailed discussions about the performance of these strategies are presented in [20, 21].

Table 1: 4 Strategies for Implementing the Portlet.

Strategy	Page Size	Loading Time	Map Update Time
Pure Javascript	289KB	13s	4.4s
JSP + Javascript	2.78MB	7.4s	2.8s
Managed Bean + Javascript	766KB	8.4s	5.4s
Web Service + Javascript	972KB	5.2s	3.4s

5. RESULTS

Although we have not yet been able to detect unusual transient deformation signals through use of the web portal, we have been able to detect seismic signals from known earthquake events, as well as several unusual events that we suspect are the result of network or data processing anomalies. Co-seismic slip is generally seen in GPS sites around the location of the earthquake; our analysis method easily segments the time series for these stations around the slip event boundary, resulting in change detection being visible in the network state map. Figure 7 shows the state change map for Southern California on October 16, 1999, the date of the Hector Mine earthquake; stations surrounding the earthquake location have changed state.

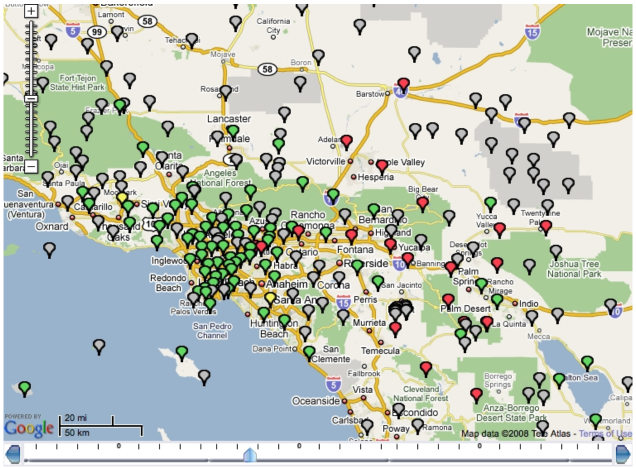


Figure 7: Network change on the day of the Hector Mine earthquake (1999-10-16).

We have found the state change over time plot to particularly useful for identifying anomalies in the data. Our first implementation of this feature resulted in the plot shown in Figure 8; notice the sudden jump in the number of state changes at 2006-09-01. These results were surprising, because we had noticed no obvious jump or change in the individual time series. Eventually we discovered that the anomaly was the result of a SOPAC data archive error: incompletely processed time series (with consequently higher noise after 2006-09-01) were being provided by the archive. The problem was quickly resolved, but our experience served to underline the utility of having multiple ways to investigate the data.

Our revised state change versus time graph (Figure 5) has several large spikes, most notably on 2006-10-01 and 2007-10-01. These spikes are the result of more than a hundred stations changing state simultaneously over a large geographical area, too large to be the result of a seismic event. Figure 9 shows the station change map for 2007 anomaly; stations across the entire extent of the network are affected. In the individual time series of the affected stations, there seems to be little perceptible signal aside from what appears to be somewhat higher noise levels around that date. Currently we suspect that this is the result of either a data processing or reference frame error, and investigation is ongoing.

In addition, our approach has been able to identify signals in individual GPS stations associated with post-seismic deformation and aquifer subsidence, as well as

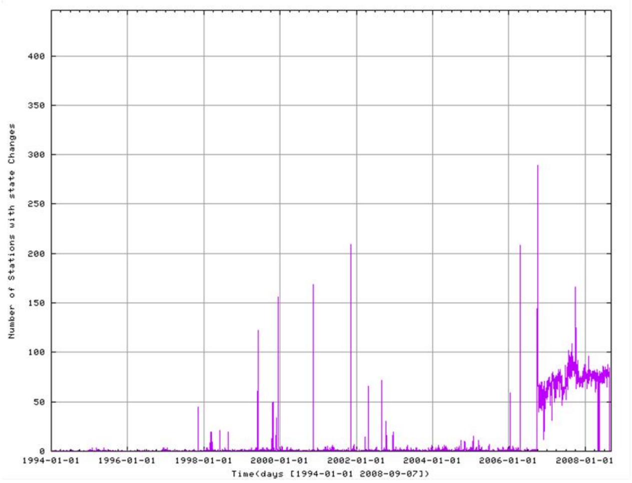


Figure 8: Plot of stations with state changes vs. time performed on an incompletely processed GPS data set.

data jumps and other glitches. For more details of these results, see [9, 7, 8].

6. CONCLUSIONS AND FUTURE WORK

We have presented a web portal environment for exploration of GPS data. This web portal environment is based around web services architecture that provides modular flexibility and controls the portal workflow in a way that maintains responsiveness to the user even under heavy computational demands. An underlying hidden Markov model based time series segmentation algorithm is used to provide insight into the behavior of individual stations as well as the network as a whole. A variety of visualization tools are provided to assist the user in interacting with the data and the analysis results.

Preliminary work with the web portal environment has shown that it can assist in the detection of both geophysical signals and nonphysical data anomalies, the study of which is ongoing. We expect that once the portal is made available to beta users in the first quarter of 2009, we will be able to capitalize on the activities of multiple users to make further discoveries.

Numerous improvements suggest themselves to the current work. More data sources will be added, including not only both the JPL and SOPAC solutions for the Southern California stations, but eventually GPS data from around the world. We intend to allow users to select pre-processing steps as well, such as removal of trends, earth tides, and seasonal signals, depending on the phenomena they wish to study. More than one time series analysis method will also be available to users, including principal components based and divergence based analysis; should users wish to try their own analysis approach but use our visualization interface, facilities will be made available for uploading results onto the web portal. Finally, we will continue to tie in the GPS analysis work with the rest of the QuakeSim portal to better facilitate the coordination of observational data analysis and modeling efforts.

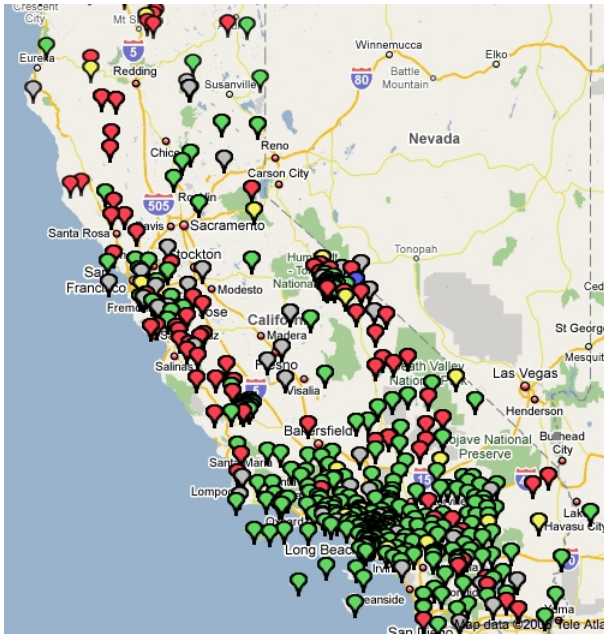


Figure 9: An anomalous network change event on 2007-10-01).

7. ACKNOWLEDGMENTS

This research was carried out in part at the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration. This work was supported by the Advanced Information Systems Technology (AIST) Program of NASA's Earth-Sun System Technology Office and by NASA SENH grant NAG5-13269 and NASA ACCESS grant NNG06GG94A.

8. REFERENCES

[1] Google map technology. http://code.google.com/apis/maps/.

[2] Yahoo widgets technology. http://widgets.yahoo.com/.

[3] J. Bellegarda and L. Nahamoo. Tied mixture continuous parameter modeling for speech recognition. *IEEE Trans. on Acoustics, Speech, and Signal Proc.*, 38(12):2033–2045, 1990.

[4] E. Bocchieri and B. Mak. Subspace distribution

clustering hidden Markov model. *IEEE Trans. on Speech and Audio Proc.*, 9(3):264–275, 2001.

[5] Y. Ephraim, A. Dembo, and L. Rabiner. A minimum discrimination information approach for hidden Markov modeling. *IEEE Transactions On Information Theory*, 35(5):1001–1013, 1989.

[6] A. Farago and G. Lugosi. An algorithm to find the global optimum of left-to-right hidden Markov model parameters. *Problems Of Control And Information Theory-Problemy Upravleniya I Teorii Informatsii*, 18(6):435–444, 1989.

[7] R. Granat. *Regularized Deterministic Annealing EM for Hidden Markov Models*. PhD thesis, University of California, Los Angeles, 2004.

[8] R. Granat, G. Aydin, M. Pierce, Z. Qi, and Y. Bock. Analysis of streaming gps measurements of surface displacement through a web services environment. *Computational Intelligence and Data Mining, 2007. CIDM 2007. IEEE Symposium on*, pages 750–757, 1 2007-April 5 2007.

[9] R. Granat and A. Donnellan. A hidden Markov model based tool for geophysical data exploration. *Pure and Applied Geophysics*, 159(10):2271–2283, 2002.

[10] K. Heki, S. Miyazaki, and H. Tsuji. Silent fault slip following an interplate thrust earthquake at the Japan Trench. *Nature*, 386(6625):595–598, 1997.

[11] H. Hirose, K. Hirahara, F. Kimata, N. Fujii, and S. Miyazaki. A slow thrust slip event following the two 1996 Hyuganada earthquakes beneath the Bungo Channel, southwest Japan. *Geophysical Research Letters*, 26(21):3237–3240, 1999.

[12] K. W. Hudnut, Y. Bock, J. E. Galetzka, F. H. Webb, and W. H. Young. The Southern California integrated GPS network (SCIGN). In Y. Fujinawa, editor, *Proceedings of the International Workshop on Seismotectonics at the Subduction Zone*, pages 175–196, NIED, Tsukuba, Japan, 1999.

[13] P. Jamason, Y. Bock, P. Fang, B. Gilmore, D. Malveaux, L. Prawirodirdjo, and M. Scharber. SOPAC web site (http://sopac.ucsd.edu). *GPS Solutions*, 8(4):272–277, December 2004.

[14] B. Juang and L. Rabiner. Mixture autoregressive hidden Markov models for speech signals. *IEEE Transactions On Acoustics Speech And Signal Processing*, 33(6):1404–1413, 1985.

[15] G. McGuire, F. Wright, and M. Prentice. A Bayesian model for detecting past recombination events in DNA multiple alignments. *Journal Of Computational Biology*, 7(1-2):159–170, 2000.

[16] T. Melbourne and F. Webb. Precursory transient slip during the 2001 m_w=8.4 Peru earthquake sequence from continuous gps. *Geophysical Research Letters*, 29(21):art. no.–2032, 2002.

[17] T. Melbourne and F. Webb. Slow but not quite silent. *Science*, 300(5627):1886–1887, 2003.

[18] M. Miller, T. Melbourne, D. Johnson, and W. Sumner. Periodic slow earthquakes from the cascadia subduction zone. *Science*, 295(5564):2423–2423, 2002.

[19] M. E. Pierce, G. C. Fox, G. Aydin, Z. Qi, A. Donnellan, J. Parker, and R. Granat. Quakesim: Web services, portals, and infrastructure for geophysics. In *Proceedings of 2008 IEEE Aerospace Conference*, pages 1–7, March 2008.

[20] M. E. Pierce, G. C. Fox, J. Y. Choi, Z. Guo, X. Gao, and Y. Ma. Using web 2.0 for scientific applications and scientific communities. In *Concurrency & Computation: Practice & Experience Special Issue for 3rd International Conference on Semantics, Knowledge and Grid*, October 2007.

[21] M. E. Pierce, X. Gao, S. L. Pallickara, Z. Guo, and G. C. Fox. Quakesim portal and services: New approaches to science gateway development techniques. In *Concurrency & Computation: Practice & Experience Special Issue on Computation and Informatics in Earthquake Science: The ACES Perspective*, May 2008.

[22] L. R. Rabiner. A tutorial on hidden Markov models and selected applications in speech recognition. *Proc. IEEE*, 77(2):257–286, 1989.

[23] G. Rogers and H. Dragert. Episodic tremor and slip on the Cascadia subduction zone: The chatter of silent slip. *Science*, 300(5627):1942–1943, 2003.

[24] K. Rose and A. V. Rao. Deterministically annealed design of hidden Markov model speech recognizers. *IEEE Trans. on Speech and Audio Processing*, 9(2):111–126, 2001.

[25] S. Young and P. Woodland. State clustering in hidden Markov model-based continuous speech recognition. *Computer Speech And Language*, 8(4):369–383, 1994.