

Data-mining in astrophysics. A search for new variable stars in databases.

Dinko Dimitrov

Institute of Astronomy and Rozhen NAO, Bulgarian Academy of Sciences

71 Tsarigradsko Shousse Blvd., 1784 Sofia, Bulgaria

dinko@astro.bas.bg

(Conference talk)

Abstract. A review of already known algorithms for search of variable objects in large databases is presented and a new multi parametric data-mining method is proposed. This method was tested on NSVS database using "relative" amplitude of light curve and maximum intensity of power spectrum of Fourier analysis as basic parameters. The tests were conducted using 8000 squared degrees area to the north of +37 degrees, i.e. 1/4 of NSVS survey area. A sample of 35 648 candidates was extracted among 6 175 197 light curves. There are 7 485 already known variables in this region and the rest of the sample, 28 163 stars, may be considered as new variable stars.

Key words: data-mining, searching methods, new variable stars

Разработка на натрупаните данни в астрофизиката. Търсене на нови променливи звезди в бази данни.

Динко Димитров

Представен е обзор на вече известните алгоритми за тресене на променливи обекти в големи бази данни и е предложен нов, многопараметричен метод. Методът беше тестван върху базата данни NSVS, като за основни параметри бяха използвани "относителната" амплитуда на кривата на блясъка и максимума на интензивността в спектъра на мощността при Фурие анализа. Тестът беше проведен върху площ от 8000 квадратни градуса с деклинация северно от +37 градуса, т.е. 1/4 от площта на обзора NSVS. От общо 6 175 197 криви на блясъка беше оформена извадка от 35 648 кандидати за променливи звезди. В тази област има 7 485 вече известни променливи, а останалите 28 163 звезди могат да се приемат за нови променливи.

Introduction

Data-mining is a process of sorting through large amounts of data and extracting hidden dependences or mutual relations between the seemingly unknown and unrelated to one another values in large databases. It has been described as a nontrivial extraction of implicit, previously unknown, and potentially useful information from the data and the science of extracting useful information from large data sets or databases. As data sets have grown in size and complexity, there has been a shift away from direct hands-on data analysis toward indirect, automatic data analysis using more complex and sophisticated tools. The captured data need to be converted into information and knowledge to become useful. Data-mining is the entire process of applying computer-based methodology, including new techniques for knowledge discovery, to the large amount of data.

The core is the concept of templates or models, reflecting a fragmentation multi-aspect relationship between the data. Templates are a collection of laws,

or selection of data set properties, presented in the appropriate forms easily accessible to users.

The term data-mining is often used to apply to the two separate processes of “knowledge discovery” and “prediction”. Knowledge discovery provides explicit information that has a readable form and can be understood by the user. Forecasting, or predictive modeling provides predictions of future events.

There are various data-mining algorithms which can be used to build the mining model:

1. **Classification algorithm** predicts one or more discrete variables, based on the other attributes in the dataset.
2. **Regression algorithm** predicts one or more continuous variables, based on other attributes in the dataset.
3. **Segmentation algorithm** divides data into groups, or clusters, of items that have similar properties.
4. **Association algorithm** finds correlations between different attributes in a dataset.

Known uses of data-mining:

- Defense sciences (combating terrorism, intelligence, etc.)
- IT sciences, computer games.
- Business, market basket analysis.
- Science and engineering (bio-informatics, genetics, medicine, education, etc.)
- Fundamental sciences (physics, astrophysics, etc.)

1 Astrophysical surveys

Over the past few years many stellar surveys were accomplished, the results of which are accessible by Internet as databases to a wide range of astronomers and are suitable for the application of data-mining methods. Astronomical datasets are large, distributed, and contain a wide variety of data types – from raw pixels to fully-derived tables of scientific measurements. The content of astronomical data archives has grown dramatically during the past decade whereas the ability of scientists to access that content remains primitive and ineffective.

Especially valuable are stellar surveys made for a long period of time, i.e. in which the data for individual objects are arranged as time-series. In such databases the search of new variable stars may be made although their primary objective may be different. These datasets are ideal for the development and deployment of data-mining tools and systems.

Examples of such current surveys (or surveys planned for implementation in the near future) can be specified among space missions, as well as among ground-based projects. Below are listed the most important NASA/ESA space missions, aimed at implementing listed below stellar surveys:

- HIPPARCOS / Tycho mission

- MOST (Micro-variability and Oscillations of STars)
- CoRoT (CoNvection ROtation and planetary Transits)
- Kepler mission
- GAIA mission
- INTEGRAL/OMA (INTErnational Gamma-Ray Astrophysics Laboratory/Optical Monitoring Camera)

There are also current and planned for the near future ground-based telescopes, oriented for stellar surveys or search for variability in stellar objects:

- WASP, Super-WASP (Wide Angle Search for Planets)
- ASAS (All Sky Automated Survey)
- NSVS (Northern Sky Variability Survey)
- HATnet (Hungarian Automated Telescope Network)
- Vulcan project
- BEST, BEST-II (Berlin Exoplanet Search Telescope)
- OGLE-III (Optical Gravitational Lensing Experiment)
- 2MASS (Two Micron All-Sky Survey)
- SDSS (Sloan Digital Sky Survey)
- Pan-STARRS (Panoramic Survey Telescope And Rapid Response System)
- LSST (Large Synoptic Survey Telescope)

2 The ROTSE-I telescope and the NSVS database

All data in the NSVS database were collected by the ROTSE-I telescope – Robotic Optical Transient Search Experiment (Akerlof et al., 2000; Woźniak et al., 2004a). The telescope consisted of four Canon 200 mm lenses with $f/1.8$ focal ratio, each covering $8.2^\circ \times 8.2^\circ$ for a total field of view about $16^\circ \times 16^\circ$. All four optical elements were carried by a single rapidly slewing mount and designated with the symbols A through D. The instrument was located at the Los Alamos National Laboratory, New Mexico, USA.

Each of the four Canon lenses was equipped with a thermo-electrically cooled AP-10 camera, which employs a Thomson TH7899M CCD. The resolution at the chip is $14.4''/\text{pixel}$, and the spatial resolution of the system was limited by instrumental seeing, about $20''$.

The ROTSE-I telescope was operated without any filters so the spectral response is primarily limited by the sensitivity of the CCD, resulting in a very broad optical band from 450 to 1000 nm. The quantum efficiency of the front-illuminated thick CCD chips makes the effective band most comparable to the Johnson R band.

During the night, the telescope slews to each patrol location, 161 tiles observable from Los Alamos, and obtains a pair of 80 s exposures, separated only by the 1.5 minute duty cycle. The limiting V magnitude of the faintest stars was typically $14.5 \div 15.5$ mag. Saturation occurred at $10 \div 10.5$ mag in normal exposures (80 s) and at about 8 mag in bright time exposures (20 s). On a typical night, two patrol sequences are obtained, covering about $18\,000 \text{ deg}^2$ of sky and recording the brightness of approximately 9 000 000

stars with four images taken in two epochs. The observations include the entire sky region north of the declination $\delta \simeq -30^\circ$; about $33\,326 \text{ deg}^2$, or 75% of the celestial sphere. Regular patrol observations by ROTSE-I began from 1998 March until 2001 December, but NSVS database include observations between 1999 April 1 and 2000 March 30.

All photometric time-series data in the NSVS are available for public access via Sky Database for Objects in Time-Domain (SkyDOT; <http://skydot.lanl.gov>). The database includes 19 995 106 objects (light curves), which correspond to 14 582 566 physical objects (stars). Every star has between 1 and 5 light curves, depending on its position on the sky – the northern fields are covered better. Number of epochs are about 200 for average light curve, corresponding to yearly visibility across the sky.

3 Methods and techniques for selection of variable stars

There are several methods to identify whether a star in the NSVS database is a variable star.

3.1 Variability Index

One of the most used technique for selection of variability is Stetson's Variability Index. The variability index itself is based on the work of Welch and Stetson (1993) and Stetson (1987, 1996). This method were also implemented in major publications, based on ROTSE-I observations, by Akerlof et al. (2000), Gettel et al. (2006), and Kinemuchi et al. (2006).

The main variability index used in these works is Stetson's L_{VAR} index (optimized for pairs of observations), which is constructed from Stetson's J and K indices:

$$J = \frac{\sum_{k=1}^n w_k \text{sgn}(P_k) \sqrt{|P_k|}}{\sum_{k=1}^n w_k} \quad (1)$$

where expression $P_k = f_1(\delta_i, \delta_j)$ is a product of residuals δ , or relative errors, of two observations v_i and v_j . The symbol w_k is weighting factor.

The second variability index is the kurtosis index K , also constructed from the residuals of observations:

$$K = \frac{\frac{1}{N} \sum_{i=1}^N |\delta_i|}{\sqrt{\frac{1}{N} \sum_{i=1}^N \delta_i^2}} \quad (2)$$

Here N is the number of independent observations.

The main variability index L_{VAR} is constructed as:

$$L_{VAR} = \left(\frac{JK}{0.798} \right) \left(\frac{\sum w}{w_{all}} \right) \quad (3)$$

The suitable cutoff value had to be chosen for the L_{VAR} index (Fig. 1), above which the star could be considered a variable star candidate. This

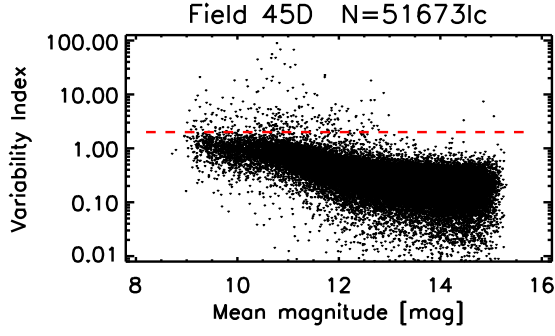


Fig. 1. Stetson's L_{VAR} variability index, computed for field 45D. The cutoff value is marked with dashed line.

cutoff value depends on magnitude and position across the Galactic plane. The authors in different papers, applied slightly different cutoff value, and the number of variable star candidates is different.

3.2 Analysis of Variance

Woźniak et al. (2004b) used a procedure similar to the analysis of variance (AOV; Schwarzenberg-Czerny 1989) technique, to select slowly variable candidates. After binning the light curve into 15 day bins, the mean m_i and the standard deviation s_i of all points in each i -th bin are calculated. The AOV ratio:

$$R_{AOV} = \frac{\sqrt{\langle m^2 \rangle - \langle m \rangle^2}}{\langle s \rangle} \quad (4)$$

is the ratio of the standard deviation of all means to the mean of all standard deviations. A properly chosen value separates stars with long-period or irregular variations, from short-period variables.

This technique is a basis for variable star search in another main variability survey – ASAS (All Sky Automated Survey; Pojmański, 2002).

3.3 Scatter/Error ratio

The simplest searching method, based only on internal (on-line) NSVS data, were performed on aggregate light-curve parameters as median magnitude $\bar{m}$, rms scatter σ_S , or median of individual errors σ_E . Similar techniques were applied on HIPPARCOS photometric data (ESA, 1997). Figure 2 shows relation between magnitude and rms scatter for each star in 45D field. In each magnitude bin, 99% value of scatter was determined to be a detection threshold.

More usefull cutoff parameter for our purpose is K_A -ratio:

$$K_A = \frac{\sigma_S}{\sigma_E} \quad (5)$$

applied also by Woźniak et al. (2004b) and Dimitrov and Popov (2007). It's meaning is “normalized” or “relative” amplitude of the light curve. The value of the scatter σ_S is calculated as follows:

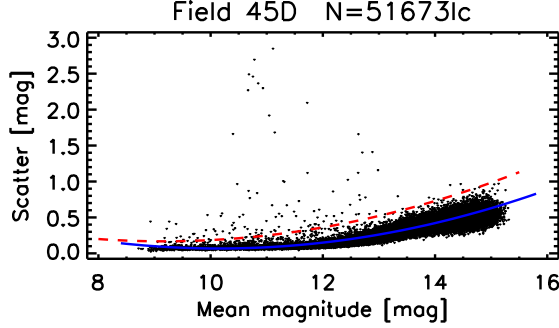


Fig. 2. Magnitude – scatter diagram for field 45D. The cutoff value is marked with dashed curve, and solid curve present median value of magnitude.

Let LC be a light curve for non-variable star, with photometry m_j , where $j = 1, 2, 3, \dots, n_{good}$. The distribution of magnitudes according to their values is close to normal (Gaussian) distribution. Some measurements differ from the median value $\bar{m}$ significantly more than $3\sigma_E$, i.e. they are not subject to the normal statistical distribution. There is a possibility that these measurements are real, caused by flares or as part of long-period eclipses. ($P > 100$ days). Much more likely, these measurements are influenced by the impact of cosmic particles, incorrect initial reduction, changing astro-climatic conditions, the influence of neighboring stars, and others. In the general case measurements, considered as erroneous are rejected and the number of “good” measurements is being reduced.

If the star has its inherent variability, then distribution will differ significantly from normal, will be extended. Then the measured width of the distribution of any level p :

$$\sigma_S(p) = f(p \times I_{max}, A, \sigma_E), \quad \text{where} \quad 0 \leq p \leq 1 \quad (6)$$

is a function both of the individual errors of measurement σ_E , and the real amplitude A . Therefore, it is more convenient to work with normalized ratio, i.e. with “relative” amplitude:

$$K_A = \frac{\sigma_S(p)}{\sigma_E} \sim \left(1 + \frac{A}{\sigma_E}\right) \quad (7)$$

3.4 Period search methods

In all existing studies based on NSVS, the search for periodic variables by Fourier analysis is performed at a later stage, after applying the other criteria to substantially reduce the number of variable star candidates. The reason is that Fourier analysis takes much computer time, always false periods exist along with the real, and human inspection is needed to clarify some controversial cases. A sample Fourier spectrum of the eclipsing binary star NSVS 6191796, based on method developed by Scargle (1982) and Horne and Baliunas (1986), is shown on Fig. 3.

Hoffman et al. (2008) used Fourier analysis method to determine the period of the candidates. A Fourier transform is made for each candidate's light curve and the power spectrum is analyzed. The periods implied by the top two power peaks and their closest two harmonics (half and twice the frequency) are then further analyzed using the phase dispersion minimization (PDM; Stellingwerf 1978) routine in the vicinity of the Fourier peaks. Many of the variable star candidates are not variable stars, yet pass the previous tests. Thus, an additional step must be made to eliminate candidates that do not have a well-defined period. This is done by taking the ratio of the strongest Fourier peak determined by previous methods to the 30th strongest peak, assumed to be a measure of the Fourier “noise”.

In this work Fourier analysis is applied in a different aspect, as a process of searching for variables stars. On Fig. 3 main peak and the closest harmonic peaks are defined. Even when the star is non-periodic (long-period, semiregular or irregular), Fourier analysis shows a noticeable peak amplitude corresponding to the very large periods, which makes it possible to detect non-periodic changes or trends. Amplitude of the largest peak of the spectrum $I_{sp}(max)$, which in most cases correspond to the actual period, was selected as the default search parameter.

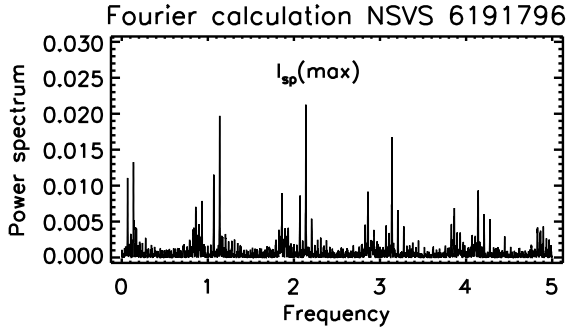


Fig. 3. Fourier power spectrum for eclipsing binary NSVS 6191796.

A phased light curve (Fig. 4; right panel) can be constructed after determining the correct period for each star. A disadvantage of Fourier analysis is

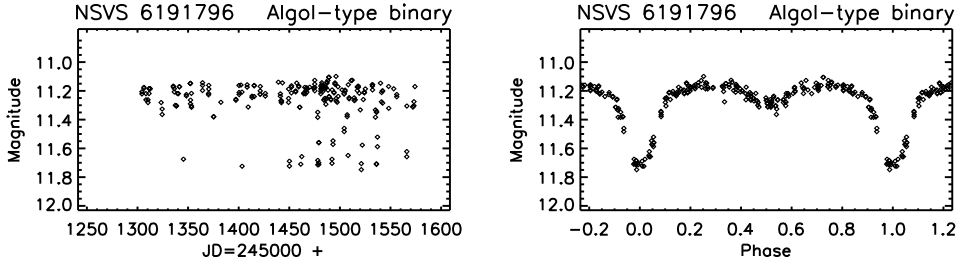


Fig. 4. Time-series and phased light curves of NSVS 6191796.

the occurrence of false periods in the spectrum. These are harmonics whose frequencies are a multiple of the main frequency. Usually they have a lower peak, but occasionally it has a peak, with greater intensity. A false periods of $P \sim 1$ day are most often due to discrete time of observations. False periods of $P \sim 30$ days appears in the NSVS (Fig. 5), i.e. number of light curves with a period of 30 days multiple exceeds the number of light curves with neighboring periods (25 or 40 days for example). The median values in the distribution of false periods for different fields are in the range of 27 to 33 days. There are also harmonics with periods of ~ 15 and ~ 60 days. The thirty day period is synchronized with the phase of the Moon, the increased sky background during the full Moon is not adequately corrected during the initial reduction of frames.

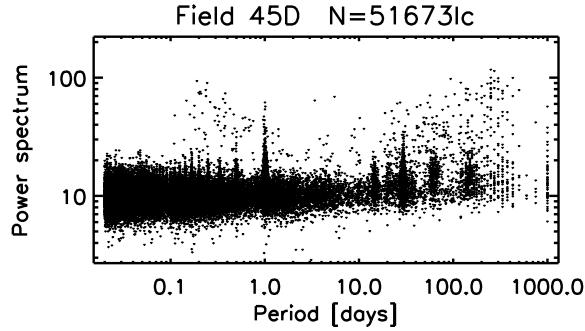


Fig. 5. Diagram Period – Power spectrum for field 45D.

4 2D-parameter searching method

In all previous large-scale searches of variable stars in the NSVS database (and in other databases such as ASAS) similar methodology was used. Search for variable stars is consecutively – considered one major parameter for selecting the candidates rejecting the huge number of objects in the database, while the number of suspected objects becomes more acceptable (of around 10 000 to 100 000). On this subset of the objects the next selection parameter is applied, which further reduces the number of candidates. After several such steps some optimum number of suspected stars is reached and periodogram analysis is performed for them.

Let LC_i be a light curve, a part of database C , where N is the total number of objects in C . For every light curve in database, the aggregate parameters are defined:

$$LC_i = F_1(\alpha, \delta, \bar{m}, n_{good}, \sigma_S, \sigma_E, \dots) \quad \text{where} \quad i = 1, 2, 3, \dots, N \quad (8)$$

The photometry of LC_i is also available:

$$LC_i = F_2(MJD_j, m_j, \sigma_j) \quad \text{where} \quad j = 1, 2, 3, \dots, n_{good} \quad (9)$$

The combination of those data can be used to obtain usefull searching parameters, $\xi_1, \xi_2, \xi_3, \dots, \xi_n$. If these parameters (or only part of them) satisfy previously defined limits (cutoff values), then the hypothesis H_0 – “the light curve of the object can be accepted as light curve of variable star”, is true. If the parameters (or some of them) are out of the limits, then the negative hypothesis H_1 – “the object is non-variable star”, is valid.

As searching parameters we can use previously explained functions: L_{VAR} , R_{AOV} , K_A , σ_S , $I_{sp}(max)$, or other. Everyone of the parameters is optimized for searching of some type of stellar variability, and may be quite useless for another type of variability.

The consecutive search can be described as follows:

At every step the basic set of objects is reduced to a subset (if the requirements are met), which is essentially the main set for next step.

$$\begin{aligned} C' &\subset C & (\xi_1 \geq \xi_1(lim)) \\ C'' &\subset C' & (\xi_2 \geq \xi_2(lim)) \\ & & \\ C^n &\subset C^{n-1} & (\xi_{n-1} \geq \xi_{n-1}(lim)) \end{aligned} \quad (10)$$

ie $C^n \subset C^{n-1} \subset \dots \subset C'' \subset C' \subset C$

If you change the order of use of different parameters, the end result may be quite different, and some interesting variable stars may be lost.

$$\begin{aligned}
LC_i (\xi_1(i) \geq \xi_1(lim), \xi_2(i) \geq \xi_2(lim), \dots) & \quad \text{not variable} \\
LC'_i (\xi_2(i) \geq \xi_2(lim), \xi_1(i) \geq \xi_1(lim), \dots) & \quad \text{variable}
\end{aligned} \tag{11}$$

For that reason, in present work new complex searching method is introduced. More effective method is a parallel search in which the database is carried out using simultaneously several search parameters. This method requires greater computing resources and so far it is not used in practice for large databases such as NSVS, though it gives more comprehensive output sample of variable objects.

$$\begin{aligned}
C' \subset C & \quad (\xi_1 \geq \xi_1(lim) \text{ OR } \xi_2 \geq \xi_2(lim) \text{ OR } \dots) \\
C'' \subset C' & \quad (\xi_k \geq \xi_k(lim) \text{ OR } \xi_{k+1} \geq \xi_{k+1}(lim) \text{ OR } \dots) \\
C^n \subset C^{n-1} & \quad (\xi_l \geq \xi_l(lim) \text{ OR } \xi_{l+1} \geq \xi_{l+1}(lim) \text{ OR } \dots)
\end{aligned} \tag{12}$$

The results of the parallel search represents each light curve as a point of n -dimensional parametric space in which the coordinate axes are the parameters of the search. If the parameters are properly selected to be orthogonal ($\xi_1 \perp \xi_2 \perp \dots \perp \xi_n$), the subset of variable stars of a particular type will be grouped in the cluster, which is spatially separated from the cluster of stars with other types of variability, as well as from the group of "permanent" stars. The most basic use of the method of parallel search is when $n = 2$, ie using only two parameters. The multi-dimensional parametric space is reduced to the two-dimensional space (plane).

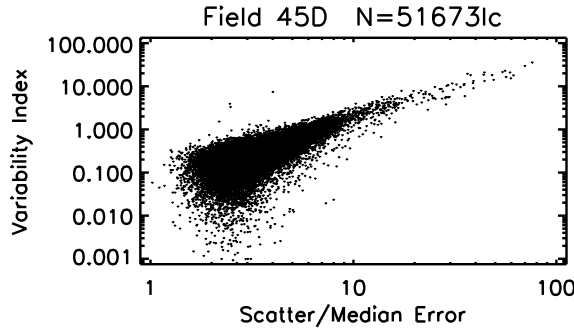


Fig. 6. Scatter/Error – Variability Index diagram for field 45D.

Figure 6 shows an example of a two-dimensional distribution of field 45D, using the parameters K_A and L_{VAR} . At least for larger values there is a linear

relation between these parameters, i.e. parameters are not orthogonal:

$$L_{VAR} = a + bK_A \quad (13)$$

Figure 7 shows another case where two parameters K_A and $I_{sp}(max)$ can be considered as independent from one another (orthogonal in the first approximation).

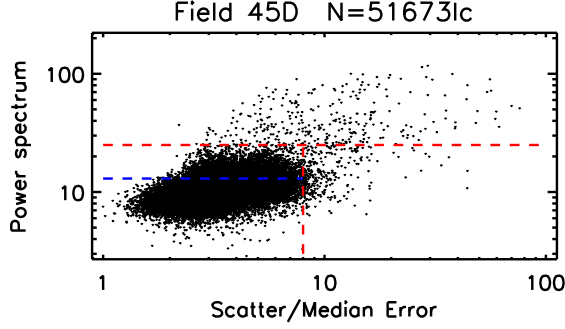


Fig. 7. Scatter/Error – Power Spectrum diagram for field 45D. The appropriate cutoff values (Section 5.) are marked with dashed lines.

5 Results

The search for variable stars in the NSVS covers a total of 180 fields (from 1A to 45D, with $\delta > +37^\circ$), about $\sim 8\,000\text{ deg}^2$ of the sky, or $\frac{1}{4}$ of the survey area. A 2D-parametric method of search is used, as key parameters for selection are K_A and $I_{sp}(max)$. For all light curves L_{VAR} was determined as an additional parameter. The appropriate cutoff values were selected empirically, and 2D-space was divided into several areas:

1. Candidates for variable stars are objects with $I_{sp}(max) \geq 2I_{sp}(FAP)$, where FAP is the False Alarm Probability, for which the significance value is 99.5%. This area includes light curves with a high level of reliability for periodicity.
2. Candidates for variable stars are objects with $K_A \geq 8$ and $I_{sp}(max) < 2I_{sp}(FAP)$. These objects do not show periodicity, but they have large amplitudes (irregular variations).
3. Candidates for variable stars with $I_{sp}(FAP) \leq I_{sp}(max) < 2I_{sp}(FAP)$, and $K_A < 8$. The objects are characterized by small amplitudes and less reliable periodicity. In this area there are many light curves with 30-days false periods, which are excluded from further analysis.

Some of the rejected objects may be variable stars, but their detection has very low efficiency (a number of found variables to a number of examined stars). A visual inspection on phased light curves with the predicted periods, was performed over all suspected stars, and suspicious cases were rejected.

The main results are systematized in Table 1.

Table 1. The search results in the selected area.

Total number of light curves	6 175 197
Number of light curves for suspected stars	242 262
Total number of variable stars	35 648
Number of known variable stars	7 485

Figure 8 shows all founded variable stars, the Galactic plane is clearly visible – a region with greater spatial density of variable stars. This area is dominated by red variable stars, which are type I population. The distribution of the relative number of variable stars in the NSVS fields (number of variable stars to the total number of stars in the field) also depends on Galactic latitude. Near the Galactic plane there are $\sim 50\,000$ objects and $\sim 1.5\%$ variable stars for field, while near the Galactic pole these values are $\sim 15\,000$ and $\sim 0.2\%$ respectively.

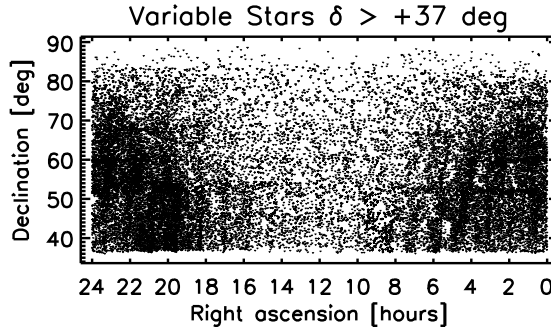


Fig. 8. All founded variable stars in the selected area.

Next step in the study of data-mining variable stars is their proper identification and classification. The ROTSE-I telescope observations are with low resolution and magnitude of an object can be influenced by light from near-by stars. The exact identification of variable stars with known catalog objects is needed (GSC, Tycho, USNO, NOMAD, 2MASS catalogs). The check for already known variable stars was performed - search by coordinates in the variable stars archives (GCVS, NSV, VSX, SIMBAD) or in specialized journals (IBVS, PZ, PZP, OEJV). The next important step is the classification of

the objects by variability type. GCVS catalog describes dozens of types and sub-types variable stars, some of which are separated by spectral characteristics, multi-colour photometry or by X-ray data. This great variety of types is not suitable for surveys as NSVS, where we do not know almost nothing about the objects except the one-colour photometry. It is therefore necessary to redefine several types of variable stars by the shape of the light curve, period and amplitude. Stars that can not be uniquely attributed to these types remain as unclassified, and require follow-up observations. Usually the following basic types of variable stars are defined:

- Red, long-period variables (M – Mira, SR – Semi-regular, L – Irregular)
- Eclipsing binaries (EA – Algol, EB – β Lyr, EW – W Uma)
- Pulsating variables (DCEP – δ Cep, RRAB (RRC) – RR Lyr, DSCT – δ Sct, BCEP – β Cep)
- Unclassified

In cases where additional information is available such as X-ray, IR, coordinates, parallaxes, etc the classification may be extended with certain types of variable stars (RS – RS CVn, RV – RV Tau, BY – BY Dra, CW – W Vir).

The information for the most of newly discovered stars is insufficient and require precise photometric and spectral observations, so the database NSVS becomes a list of targets for further research. Especially valuable for the astrophysics are Algol-type oscillating eclipsing binaries – OEA stars. A list of targets for follow-up CCD observations was prepared and patrol observations for 15 new eclipsing variables with the 60cm telescope at NAO Rozhen, were made. Four stars with short-period pulsations were detected, with periods < 2 hours and amplitudes between 0.01 and 0.05 mag in R . Using spectral observations in Coudé focus of 2m RCC telescope, it was found that those stars belong to spectral classes from A to early F, so variables can be attached to a small family of oEA stars (Dimitrov, Kraicheva and Popov, 2008a,b). The search for such objects in the NSVS database continue.

Acknowledgements

This publication makes use of the data from the Northern Sky Variability Survey created jointly by the Los Alamos National Laboratory and University of Michigan. The NSVS was funded by the Department of Energy, the National Aeronautics and Space Administration, and the National Science Foundation. This publication makes use of data products from the Two Micron All Sky Survey, which is a joint project of the University of Massachusetts and the Infrared Processing and Analysis Center/California Institute of Technology, funded by the National Aeronautics and Space Administration and the National Science Foundation. This research also has made use of the SIMBAD database, operated at CDS, Strasbourg, France, NASA's Astrophysics Data System Abstract Service, and the USNOFS Image and Catalogue Archive operated by the United States Naval Observatory, Flagstaff Station (<http://www.nofs.navy.mil/data/fchpix/>). This study was supported in part by the computing resources of DIVA-BG society.

References

- Akerlof, C., et al., 2000, *AJ*, 119, 1901
Dimitrov, D., Popov, V., 2007, *PZ*, 27, 2
Dimitrov, D., Kraicheva, Z., Popov, V., 2008a, *IBVS No 5842*
Dimitrov, D., Kraicheva, Z., Popov, V., 2008b, *IBVS No 5856*
ESA 1997, *The Hipparcos Catalogue*, *ESA SP-1200*
Gettel, S., et al., 2006, *AJ*, 131, 621
Hoffman, D., Harrison, T., Coughlin, J., et al., 2008, *AJ*, 136, 1067
Horne, J., Baliunas, S., 1986, *ApJ*, 302, 757
Kinemuchi, K., et al., 2006, *AJ*, 132, 1202
Pojmański, G., 2002, *Acta Astron.*, 52, 397
Scargle, J., 1982, *ApJ*, 263, 835
Schwarzenberg-Czerny, A., 1989, *MNRAS*, 241, 153
Stellingwerf, R., 1978, *ApJ*, 224, 953
Stetson, P., 1987, *PASP*, 99, 191
Stetson, P., 1996, *PASP*, 108, 851
Welch, D., Stetson, P., 1993, *AJ*, 105, 1813
Wils, P., et al., 2006, *MNRAS*, 368, 1757
Wozniak, P., et al., 2004a, *AJ*, 127, 2436
Wozniak, P., et al., 2004b, *AJ*, 128, 2965

