Research in Astron. Astrophys. **2009** Vol. **9** No. **6**, 635 – 640
http://www.raa-journal.org *http://www.iop.org/journals/raa*

***R**esearch in*
***A**stronomy and*
***A**strophysics*

LETTERS

New supernova candidates from the SDSS-DR7 spectral survey *

Liang-Ping Tu[1,2,3], A-Li Luo[2], Fu-Chao Wu[1], Chao Wu[2] and Yong-Heng Zhao[2]

[1] Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China; *lptu@nlpr.ia.ac.cn*
[2] National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100012, China; *lal@bao.ac.cn*
[3] School of Science, University of Science and Technology Liaoning, Anshan 114051, China

Received 2009 February 27; accepted 2009 April 21

Abstract This letter presents 25 supernova candidates discovered from SDSS-DR7 by using our dedicated method, called Sample Decrease. Ten of them have been confirmed by other research groups, while the remaining 15, including 14 Type Ia and one Type II, are first discovered based on Supernova Identification analysis. The results demonstrate that our method is reliable. The description of the method and some detailed spectral analysis procedures are also presented.

Key words: techniques: spectroscopic — supernovae: method — data analysis

1 INTRODUCTION

Supernovae (SNe) are generally discovered by repeat imaging of the same region of sky every other night, and by measuring light curves for objects in the area. The SDSS Supernova Survey was one of three surveys (along with the Legacy and SEGUE surveys) of SDSS-II, a 3-year extension of the original SDSS that operated from July 2005 to July 2008. Besides the imaging survey, the spectral survey of SDSS also gathers a large amount of spectroscopic data of galaxies. These spectra are the basis of much astronomical research. As a byproduct, it is possible to detect SNe in their host galaxies spectroscopically (Madgwick et al. 2003), since SNe spectra might have obvious broad peaks and troughs that modulate the spectrum of its host galaxy. Using the rate of SNe detection computed in Madgwick et al. (2003), there are $\sim$ 200 Type Ia SNe detections among $\sim 10^6$ galaxy spectra in the SDSS-DR7 (Abazajian et al. 2008). Madgwick et al. (2003) detected 19 Type Ia SNe from $\sim 10^5$ galaxy spectra in SDSS-DR1 (Abazajian et al. 2008) using the spectroscopic approach.

In this Letter, we report the 14 Type Ia SNe and 1 Type II SN, which were detected through the spectroscopic approach in SDSS-DR7. Our method is somewhat different than Madgwick et al. (2003), as described in detail in Section 2. The detection results are reported in Section 3, and conclusions and future possibilities are summarized in Section 4.

2 METHOD

The spectroscopic approach mentioned by Madgwick et al. (2003) needs to perform galaxy subtraction and match all templates for each spectrum. To simplify the procedure in a large number of spectral

* Supported by the National Natural Science Foundation of China.

datasets, we present the concept of “Sample Decrease”. Before confirming SNe by template matching, most of the galaxy spectra in the dataset without obvious SN features are excluded, and only those possible candidates are kept for the template matching.

2.1 Sample Decrease

The host-galaxy spectra with a SN are much more sparse than galaxy spectra without SN, so we can remove most of them using an outlier detection method. The process could be divided into two phases: the first phase is SN statistic eigen representation of each galaxy’s spectrum, and the second phase is outlier detection. There are 90 Ia-normal templates in Peter Nugent’s SN Spectral Template library (Nugent 1997), and each template represents a spectrum the day after the explosion from 0 to 89 d [1]. We used only the 6th$\sim$40th templates which are spectra near the day of peak luminosity (i.e., before 14 d to after 20 d respectively). SNe with obvious SN characteristics could be easily found by using these templates. Linear interpolation is first performed on each of these 35 SN templates for the wavelength range 3801$\sim$8000 Å. Then, these 35 spectra are transformed by Principal Component Analysis (PCA) to obtain the 12 eigen spectra (eigenvectors), which span an eigen-subspace of Type Ia SN. The projection of one normalized galaxy spectrum onto this space is a 12-dimensional vector, which is called a SN statistic eigen-representation of this galaxy’s spectrum. The first two eigenspectra are shown in Figure 1.

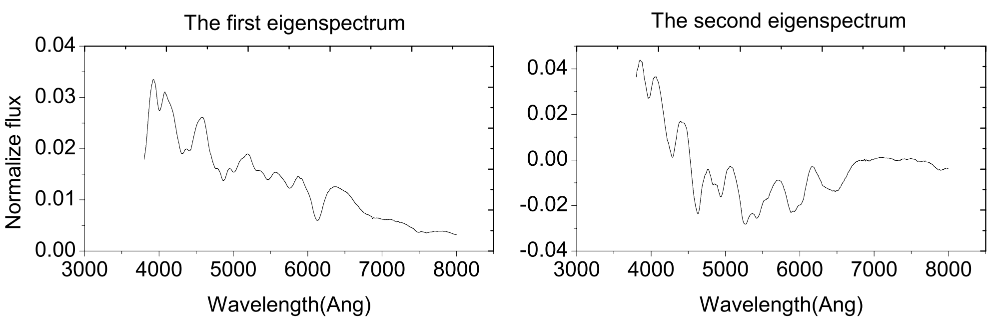

Fig. 1 First and second eigenspectra of Type Ia Supernova Templates.

Markus (2000) introduced the concept of Local Outlier Factor (LOF) used for outlier detection, which can be applied to describe the singularity of galaxy spectra with SN components in all the spectra. Thus, the outlier might be spectra of SNe plus their host galaxies. The definitions related to LOF are described as follows:

Definition 1: ($k-$distance of an object p). For any positive integer k, the $k-$distance of object p, denoted as $k-$distance(p), is defined as the distance $d(p,o)$ between p and an object $o \in D$ such that:
(i) For at least k objects $o' \in D \setminus p$, it holds that $d(p,o') \leq d(p,o)$
(ii) For at most $k-1$ objects $o' \in D \setminus p$, it holds that $d(p,o') < d(p,o)$.

Definition 2:($k-$distance neighborhood of an object p). Given the $k-$distance of p, the $k-$distance neighborhood of p contains every object whose distance from p is not greater than the $k-$distance, i.e. $N_{k-\mathrm{distanc(p)}}(p) = \{q \in D \setminus p \mid d(p,q) \leq k-\mathrm{distance}(p)\}$. These objects q are called the k-nearest neighbors of p. Simplify the notation to use $N_k(p)$ as a shorthand for $N_{k-\mathrm{distance}}(p)$.

Definition 3: (reachability distance of an object p $w.r.t$ object o). Let $k \in Z^+$, then the reachability distance of object p with respect to object o is defined as $\mathrm{reachdist}_k(p,o) = \max\{k-\mathrm{distance}(o), d(p,o)\}$.

Definition 4 & 5: The local reachability density and the local outlier factor of an object p are defined as

$$\mathrm{lrd}_k(p) = 1/\frac{\sum_{o \in N_k(p)} \mathrm{reachdist}_k(p,o)}{|N_k(p)|}, \tag{1}$$

[1] *http://supernova.lbl.gov/$\sim$nugent*

$$\mathrm{LOF}_k(p) = \frac{\sum_{o \in N_k(p)} \frac{\mathrm{lrd}_k(o)}{\mathrm{lrd}_k(p)}}{|N_k(p)|}. \tag{2}$$

Using the above *k*-LOF definition, we can design the procedures for data reduction:
(i) calculate the k-distance of each sample in *D*;
(ii) obtain the k-distance neighborhood of each object $p \in D$;
(iii) calculate the local reachability density of each object $p \in D$ by using Equation (1);
(iv) calculate the k-local outlier factor of each object $p \in D$ using Equation (2);
(v) sort all of the objects in *D* according to their *k*-LOF in descending order, and keep the objects which have higher *k*-LOF as the range of candidates. Here, we kept $1 \sim 2$ percent of all samples.

2.2 Cross-correlation Template Matching Based on SNID

After we performed the sample decrease in SDSS-DR7, the number of remaining spectra is 2945. Template matching should be completed for this relatively small dataset. Blondin et al. (2007) presented an interactive cross-correlation method named Supernova Identification (SNID) to identify SNe, and get the redshift, age, and type of each SN, since all template spectra are at zero redshift, and the types and ages of each template are known. We simplify the SNID procedure into first four steps as follows and make it run automatically. Then, the 5th step is to run the SNID pipeline to confirm selected samples with high probability of being a SN:
(i) Bin each spectrum on a logarithmic wavelength axis.
(ii) Remove the continuum and normalize the spectra.
(iii) Smooth the spectra using a *mean* filter, and remove strong narrow lines in each spectrum.
(iv) Match by cross-correlation (Details can be found in Blondin et al. 2007 and Tonry et al. 1979).
(v) Interactively check each candidate with high confidence as a SN through the SNID pipeline.

An example of a spectrum which is processed according to the given steps is shown in Figure 2.

3 RESULTS

We applied our method to all fields of SDSS DR7 (survey plate numbers $0266 \sim 2974$), excluding the spectra with low signal to noise ratio (S/N) and with uncertainty in high redshift. The selection criteria are: S/N>10 (all in g, r, i band), $z < 0.25$, and with the spectral type “galaxy”. The total number of galaxy spectra in which we performed the sample decrease procedure is 294 843. The first step was to de-redshift all the galaxy spectra, as Figure 2 shows. Secondly, strong narrow lines were removed in each spectrum using a narrow scaled *wavelet* filter, just to keep broad peaks and troughs such as SiII absorption at 6150 Å, which is one of the main characteristics of a Ia SN spectrum. A spectrum where strong narrow lines were removed is shown in the left bottom panel in Figure 2. After pre-processing, we calculated the SN statistical representations for all spectra, i.e. we projected each normalized spectrum onto the eigen-subspace. Then, we computed and sorted the LOF of each sample, and pruned 99% of the samples with low LOFs in the DR7 galaxy dataset. Finally, we obtained a total of 2945 spectra from all the 294 843 spectra.

We then used a cross-correlation template matching technique (Tonry et al. 1979; Blondin et al. 2007) to calculate the similarities between a spectrum and each template, and to sort them in descending order. By comparing all similarities in all templates, only 65 spectra were left to be identified. Finally, we confirmed 25 spectra of them as SN through SNID. Among these 25 spectra, 5 of them were discovered by Madgwick et al. (2003) (see top 5 rows in Table 1), and the other 5 spectra were recorded elsewhere (see bottom 5 rows in Table 1). The remainder of them have not been recorded, moreover, the 15 spectra are shown in Figure 3 and their parameters are listed in Table 2.

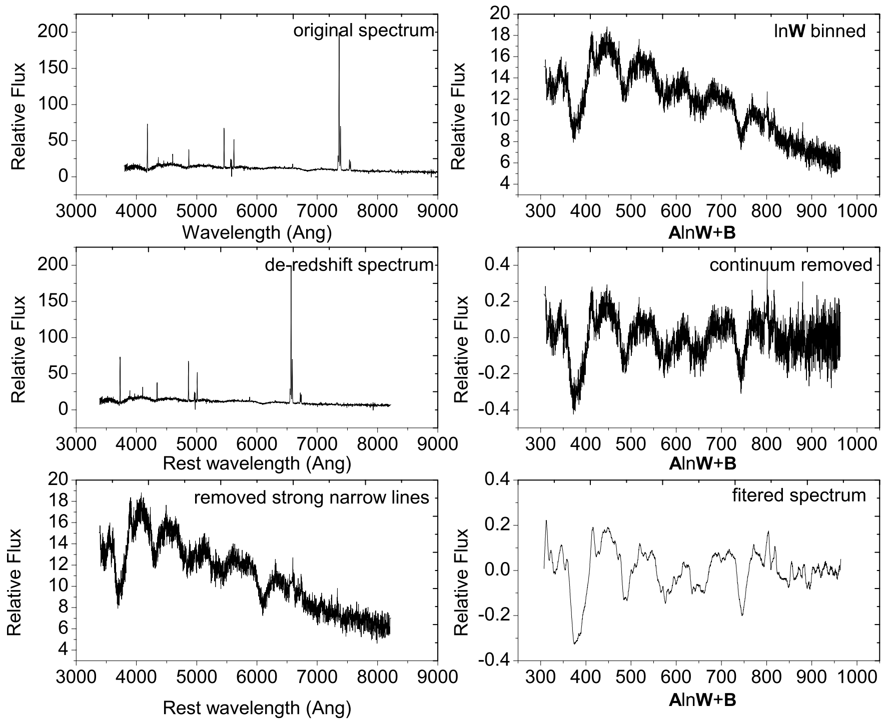

Fig. 2 An example of SN spectra. Top left: original spectrum, middle left: the corresponding spectrum which has been de-redshifted, bottom left: the result where strong narrow lines are removed; Top right: binned spectrum, middle right: the continuum divided spectrum, bottom right: the smoothed spectrum. In the right panels, the x-axes are in logarithmic wavelengths denoted as 'AlnW+B', where 'A' and 'B' are two constants. Such binning is convenient for the calculation of the de-redshift.

Table 1 Recorded SN+host Spectra that We Discovered

SDSS Name	IAU name	Date	z	Age (d)	Type	SDSS r-mag
SDSS J101800.47–000157.9	Sn 2000fx	2000–12–01	0.065	16.0∼18.0	Ia-pec	17.98
SDSS J080312.61+473649.7	Sn 2000fy	2000–12–06	0.117	2.5∼3.5	Ia-norm	17.93
SDSS J011835.83+144100.5	Sn 2000fz	2000–12–15	0.054	6.3∼6.8	Ia-norm	16.73
SDSS J092229.14+575429.3	Sn 2001kj	2001–01–02	0.063	10.0∼11.0	Ia-norm	17.79
SDSS J095153.07+010605.7	Sn 2001kp	2001–03–21	0.063	−3.0 ∼ −2.0	Ia-pec	17.58
SDSS J095915.75+005802.3	Sn 2001kr	2001–03–26	0.086	9.4	Ia-csm	18.15
SDSS J093749.92+101138.0	Sn 2003ly	2003–12–17	0.095	17.0	Ia-norm	18.46
SDSS J095948.15+112825.3	Sn 2004ar	2004–02–20	0.064	9.8	Ia-norm	18.04
SDSS J132834.01+415108.2	Sn 2004co	2004–04–17	0.029	8.1	Ia-norm	15.82
SDSS J154024.75+325157.2	Sn 2004cp	2004–05–24	0.053	28.5	Ia-norm	17.50

4 CONCLUSIONS

A novel spectroscopic analysis of the 294 843 galaxy spectra from the SDSS-DR7 has resulted in the definite identification of 14 SNe Ia and one SN II. There are three reasons for not ensuring search completeness. First, reduced samples were fixed to 1% of all samples to keep high reliability, and some SNe would be lost. Secondly, only half of Nugent's SN templates with obvious SN characters were adapted in this method. Lastly, there is still a problem in host galaxy subtraction, and we have to give

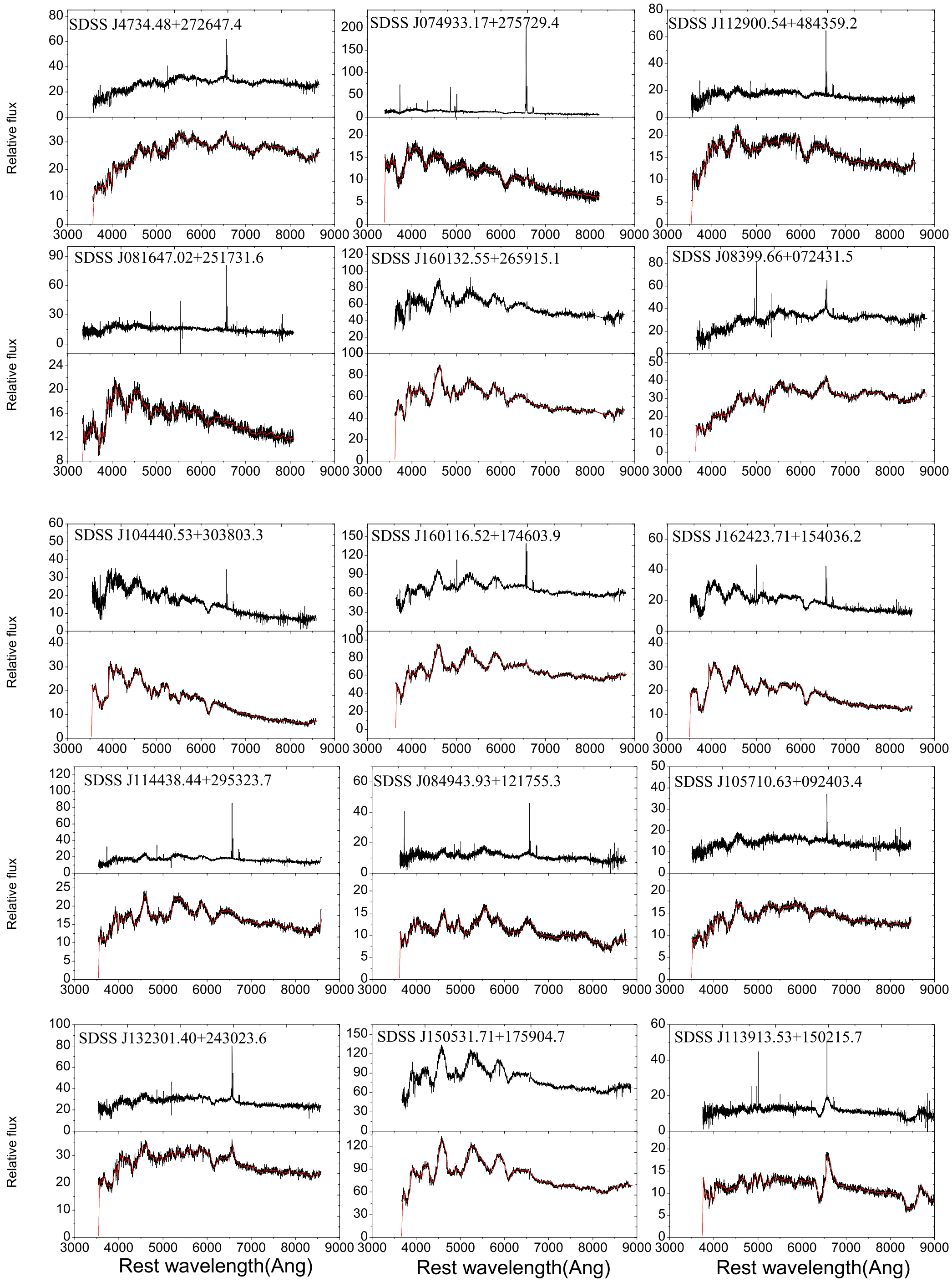

Fig. 3 Spectra of the newly identified SNe in SDSS-DR7 are shown in each panel. In addition, the bottom plot in each panel shows the spectrum with strong narrow lines removed. The smooth spectrum is shown overplotted on this spectrum.

Table 2 Newly Discovered SN+host Spectra

SDSS Name	Another name	Date	z	Age (d)	Type	SDSS r-mag
SDSS J074734.48+272647.4	anon	2002–12–10	0.061	41.6	Ia-91bg	17.45
SDSS J074933.17+275729.4	anon	2002–12–10	0.131	–5.0	Ia-norm	18.42
SDSS J112900.54+484359.2	anon	2003–01–03	0.085	–3.4	Ia-norm	18.01
SDSS J081647.02+251731.6	anon	2003–03–11	0.158	–8.6	Ia-norm	18.11
SDSS J160132.55+265915.1	anon	2003–07–02	0.071	–2.0	Ia-norm	16.64
SDSS J083909.66+072431.6	2MASX J08390967+0724320	2003–11–21	0.042	23.7	Ia-91T	17.29
SDSS J113913.53+150215.7	anon	2005–01–16	0.022	34.1	IIp	18.36
SDSS J104440.53+303803.3	anon	2005–02–28	0.077	4.3	Ia-norm	18.09
SDSS J160116.52+174603.9	MCG+03–41–035	2005–07–05	0.040	16.1	Ia-norm	16.46
SDSS J162423.71+154036.2	anon	2005–07–07	0.097	0.3	Ia-91T	17.85
SDSS J114438.44+295323.7	anon	2006–03–03	0.075	19.1	Ia-norm	17.94
SDSS J084943.93+121755.3	NGC 2682	2006–03–21	0.053	28.5	Ia-norm	18.39
SDSS J105710.63+092403.4	anon	2008–02–02	0.086	9.4	Ia-csm	18.12
SDSS J132301.40+243023.6	NGP9 F380–0370514	2008–02–28	0.088	0.3	Ia-norm	17.43
SDSS J150531.71+175904.7	2MASX J15053172+1759051	2008–03–31	0.033	14.1	Ia-norm	16.23

up searching for those faint SNe. In this method, only small samples are matched with templates, which improved search efficiency.

Acknowledgements The authors are grateful to Jingyao Hu, Jinsong Deng and Yulei Qiu for many useful discussions and to the referee for helpful suggestions. All experiments are based on SDSS data releases. This work was funded by the National Natural Science Foundation of China (NSFC) under Nos. 60773040 and 60402041, and supported by the Young Researcher Grant of NAOC, CAS.

References

Abazajian, K. 2008, arXiv0812.0649

Abazajian, K., Adelman-McCarthy, Jennifer K., et al. 2003, AJ, 126, 2081

Blondin, S., & Tonry, J. L. 2007, ApJ, 666, 1024

Madgwick, D. S., Hewett, P. C., Mortlock, D. J., & Wang, L. 2003, ApJ, 599, L33

Markus M. Breunig, Hans-Peter Kriegel, Raymond T. Ng, Jörg Sander, 2000, in Proceeding of 2000 ACM DIGMOD International Conference of Manage of Data, 2000(29), 93

Nugent, P. 1997, PhD. thesis, University of Oklahoma

Tonry, J., & Davis, M. 1979, AJ, 84, 1511

York, D. G., York Donald, G., Adelman, J., et al. 2000, AJ, 120, 1579