

Linking Very Large Telescope asteroid observations

M. Granvik,¹ K. Muinonen,¹ J. Virtanen,¹ M. Delbó,² L. Saba,²
G. De Sanctis,² R. Morbidelli,² A. Cellino,² and E. Tedesco³

¹Observatory, P.O. Box 14, FIN-00014 Univ. Helsinki, Finland
email: mikael.granvik@astro.helsinki.fi

²Turin Astronomical Observatory, via Osservatorio 20, I-10025 Pino Torinese (TO), Italy

³Space Science Center, Univ. New Hampshire, 39 College Road, Durham, NH 03824, USA

Abstract. A novel method for the preliminary identification of asteroids at discovery and a few days thereafter is being developed in Helsinki. Having two different sets of asteroid observations, the goal is to identify all possible pairs of objects between the sets. An arbitrary asteroid can either remain unidentified, or be preliminary linked to one or more asteroids. In the case of ambiguity, the final decision must usually be based on additional observations. We use a multistep approach, during which possible pairs of objects are first selected by comparing ephemerides that have been generated for three common epochs. The method has been successfully tested using both Very Large Telescope observations, and simulated observations of near-Earth and main-belt objects. Identification results of simulated observations indicate that the observing strategy promoted by the Minor Planet Center might not be the best one, at least for the purposes of identification. The ultimate goal is to produce a real-time asteroid identification tool for ESA's astrometric space observatory Gaia, the Lowell Observatory Near-Earth-Object Search, the Near-Earth Space Surveillance mission, and the Nordic Near-Earth Object Network. The tool could also benefit large-scale surveys done with the Large Synoptic Survey Telescope, and the Discovery Channel Telescope.

Keywords. Celestial mechanics, methods: numerical, methods: statistical, asteroids, surveys

1. Introduction

Possible linkages between asteroid observations are usually sought by comparing the observed positions with ephemerides, or by comparing orbital elements inverted from two separate sets of observations. The first approach is used daily by, for example, the Minor Planet Center (MPC) for several purposes. First, it is used when identifying reasonably well-known asteroids in a batch of new observations. The asteroids do not need to be numbered, because in most cases the mean anomaly M contains the most significant error, and a fit to the observations can thus be obtained by varying M . Second, the first approach is used when searching for linkages between single-night observations by comparing the computed positions and motions with observed positions and motions. Due to too few observations, and/or too short observational arcs, a rigorous orbital inversion is not possible when using deterministic approaches, and therefore the computed positions and motions are obtained by using Väisälä-type orbits. As the Väisälä method makes assumptions on the time of perihelion and the geocentric distance, this method can be reliable only for a couple of days (if at all). The Väisälä method has, however, proven to be useful in many cases, and is therefore also used by, for instance, the Lowell Observatory Near-Earth-Object Search (LONEOS). Milani *et al.* (2001) used the first

approach in a method which they called *attribution*. The attribution method allows linking of short-arc observations, or, more accurately, a representation of those observations, with a least-squares orbit of an unnumbered asteroid over longer time spans than is possible with the *M*-variation technique. The same group use the second approach when tentatively identifying two reasonably well-determined orbits, for which the least-squares approximation is valid, over a long time span (Milani *et al.* 2000).

In the present method a third possibility is used: namely, the comparison of ephemeris clouds obtained with statistical orbital ranging (Virtanen *et al.* 2001, Muinonen & Bowell 1993). In contrast to the methods briefly described above, the present method is particularly suitable for analysing exiguous single-night data, such as a set of Very Large Telescope (VLT) observations analysed in this paper. As is the case with all Monte Carlo (MC) inversion methods, statistical ranging outputs a relatively large amount of data. In fact, the processing and interpretation of the output data turns out to be the bottleneck in most applications relying on these methods. Particularly in the case of asteroid identification, where a large number of objects have to be processed as fast as possible, the key element of the whole process is data mining of the statistical ranging output. In the current identification method, data mining is efficiently carried out by using the so-called address comparison technique.

The paper is organised as follows. Section 2 briefly describes the statistical ranging method, which is the inversion method used throughout this paper. The overall identification scheme is presented in Section 3, and the most important building block is described in Section 4. In Section 5, results produced by the identification method are presented and discussed. Finally, in Section 6, the key findings are summarised.

2. Statistical orbital ranging

The probability density of orbital elements is examined using MC selection of orbits in orbital element space in the following way:

- Two observations are chosen (usually the first and the last), and angular deviations mimicking the observational errors in R.A. and Dec. are introduced.
- Topocentric ranges (distances) are assumed corresponding to the observation dates. In other words, two positions equalling six constants of integration are known.
- A trial orbit is first computed using the *p*-iteration method and is then compared to all observations. If the trial orbit fits the observations to predefined accuracy (defined as a $\Delta\chi^2$ -threshold and maximum sky-plane residuals), it is added to the sample of possible orbits.

In the basic version of statistical ranging, the initial topocentric range intervals are determined manually using an educated guess, whereas in the automated version the topocentric range intervals are further improved using the $3\text{-}\sigma$ cutoff values of the range probability density. By increasing the number of generated sample orbits ($10 \rightarrow 200 \rightarrow n$), an unbiased phase-space region of possible orbits is found. Each sample orbit is assigned a weight, which describes how well it explains the observations. Ignoring the weights, the distribution merely shows the extent of different orbital solutions in the orbital element space assuming predefined observational errors. Ephemerides are generated by transforming the orbital elements of every sample orbit to a position on the celestial sphere at a given epoch, the result thus being an ephemeris cloud.

3. The identification scheme

It is, in principle, possible to use a straightforward trial-and-error-scheme while searching for identifications using statistical ranging. The idea is to *try* to perform the inversion using statistical ranging by using all observations corresponding to two objects. If the inversion succeeds—that is, at least one orbit can be found, it shows that the objects can be tied together using the same orbit within the assumed observational errors.

The direct identification approach is, however, not a particularly efficient technique. Assume, for instance, two observation sets containing 1000 objects each. Simplistically, the number of pairs to be checked is 1 million, while at most only 0.1% are correct identifications. We use a novel method termed ephemerides address comparison (at common epochs) to efficiently reduce the huge initial number of object pairs (see Section 4). The remaining, reasonably probable, pairs are examined by using the direct approach after an intermediate step. The intermediate step is similar to the last step, but it only uses the two first and the two last observations of the combined observation set.

4. Ephemerides address comparison at common epochs

The idea is to generate ephemeris clouds (R.A. and Dec.) for all objects in both sets for three common epochs, and then find out whether any objects in different sets have similar ephemerides at all three epochs, which would indicate a possible identification (Granvik *et al.* 2004, *in preparation*). The choice of epochs can be optimised, but the use of the observational mid-epoch as the first epoch is a good first approximation. The choice can be justified based on the knowledge that the ephemeris uncertainty grows with increasing time since last observation (Muinonen *et al.* 1994). The second and third ephemerides are produced by propagating orbits from the first epoch 12 and 24 hours forward in time, respectively, and transforming the corresponding orbital elements to ephemeris clouds as described in Section 2. The search for similarities among the two ephemerides is carried out efficiently using the address-comparison technique, which is presented in Section 4.1. Similar ephemerides at several epochs indicate a tentative linkage, which requires further investigation using either a statistical-ranging inversion (required for single-night linkages due to short observational arcs), or differential correction of a least-squares orbit.

4.1. Address comparison

When searching through the bins of the discretised ephemeris clouds to find overlapping ephemerides, most time is spent checking empty bins, which is inefficient. Instead of using the whole map, or multidimensional array, one can write an address to each bin and just compare the addresses that are occupied with orbits (Muinonen *et al.* 2004). In practice, the address is an integer $i \in \mathbb{N}_+$, transformed from an array of elements $p \in \mathbb{R}^n$ (here, $n = 6$) using a transformation algorithm f , i.e., $i = f(p, \dots)$. The transformation algorithm f essentially does the same as a basic binning algorithm, but instead of returning the coordinates of a bin in multidimensional space (a bi-product of the algorithm), it transforms the coordinates to a single integer i . The integer is the individual ID-number of a bin in the original multidimensional bin-network. Besides p , the essential input values for the transformation algorithm f are the boundary values of the multidimensional space and the bin sizes. At present we take into account the whole sky, so the only essential input value is the bin size, which is currently one arcmin for both R.A. and Dec.

Because the observations of an object are inverted to a sample of orbits, and every orbit in the sample is transformed to three ephemerides and further to a value i , each object will get a one-dimensional array containing the i -values. Potential identifications are sought by comparing the i -arrays of objects in the first set with the i -arrays of objects in the second set. The search can move to the next candidate pair as soon as a single pair of equal addresses, or integers, is found.

When dealing with an array of integers, the search algorithm can be optimised more easily than when searching a multidimensional array. By sorting the i -values in ascending order, a *binary search* algorithm can be used for the search of similar elements, which significantly accelerates the comparison algorithm.

5. Application to VLT observations

Observations with the VLT and the Canada-France-Hawaii Telescope were obtained in January 2004 to provide ground-based follow-up for the Spitzer First Look Survey Ecliptic Plane Component (FLS/EPC). See Meadows *et al.* (2004) for a description of, and preliminary results from, this Spitzer program. The results discussed here, using only the VLT data, represent the first step in applying this method to obtaining orbits for unknown asteroids observed in the Spitzer FLS/EPC. The requirement for the Spitzer observations to be made at a solar elongation of 115° considerably complicates the identifications because asteroids along the line of sight are close to their turning points and are consequently moving slowly and along curved paths. The observational set contains 532 detections of asteroids at $V \lesssim 26^m$ unevenly spread over five nights (Fig. 1). An estimated, rather pessimistic, accuracy of $\sigma = 0.5''$ was used in the identification procedure.

The observations were provided as single observations, not pairs of observations per object as is usual, and the first objective was therefore to create tentative objects within each nightly set. This was accomplished by assuming linear motion within the few hours of observation each night. A reasonable upper limit for the coordinate motion was found by requiring that the object can be found within the observed area on two consecutive nights. If the true motion for an object is higher than the given limit, it could not be linked, and can therefore be omitted from the set of objects to be scanned for linkages. The maximum motion that a linkable object could have is thus $1.5^\circ/d = 0.0625''/s \approx 0.1''/s$, deriving from the width of the observational area. Given the upper limit of motion, all possible pairs of observations were generated for each nightly set. Additional observations for each observation pair were also searched for by assuming linear motion and a maximum deviation of $6\text{-}\sigma$ from the nominal position. Up to two additional observations could be found, resulting in a maximum of four observations per object per night.

Linkages between nights were searched using a cumulative strategy: the first-night objects were linked with the second-night objects, and the preliminary identifications as well as all individual objects from both the first and the second nights were linked with the third-night objects, and so on. It turned out that ambiguous linkages occurred even if there were three nights of observations for two objects (see Section 5.1). Identifications containing four or five nights of observations were, however, self-consistent; i.e., there was no overlapping of observations between different identifications.

Including ambiguous linkages, a total of 73 preliminary linkages consisting of different combinations of 429 detections were found by using the new method. The remaining 103 detections have not yet been analysed in detail, but it seems plausible, from the experience gained with simulated observations, that most of them were either too fast-moving or too faint to be detected on two separate nights.

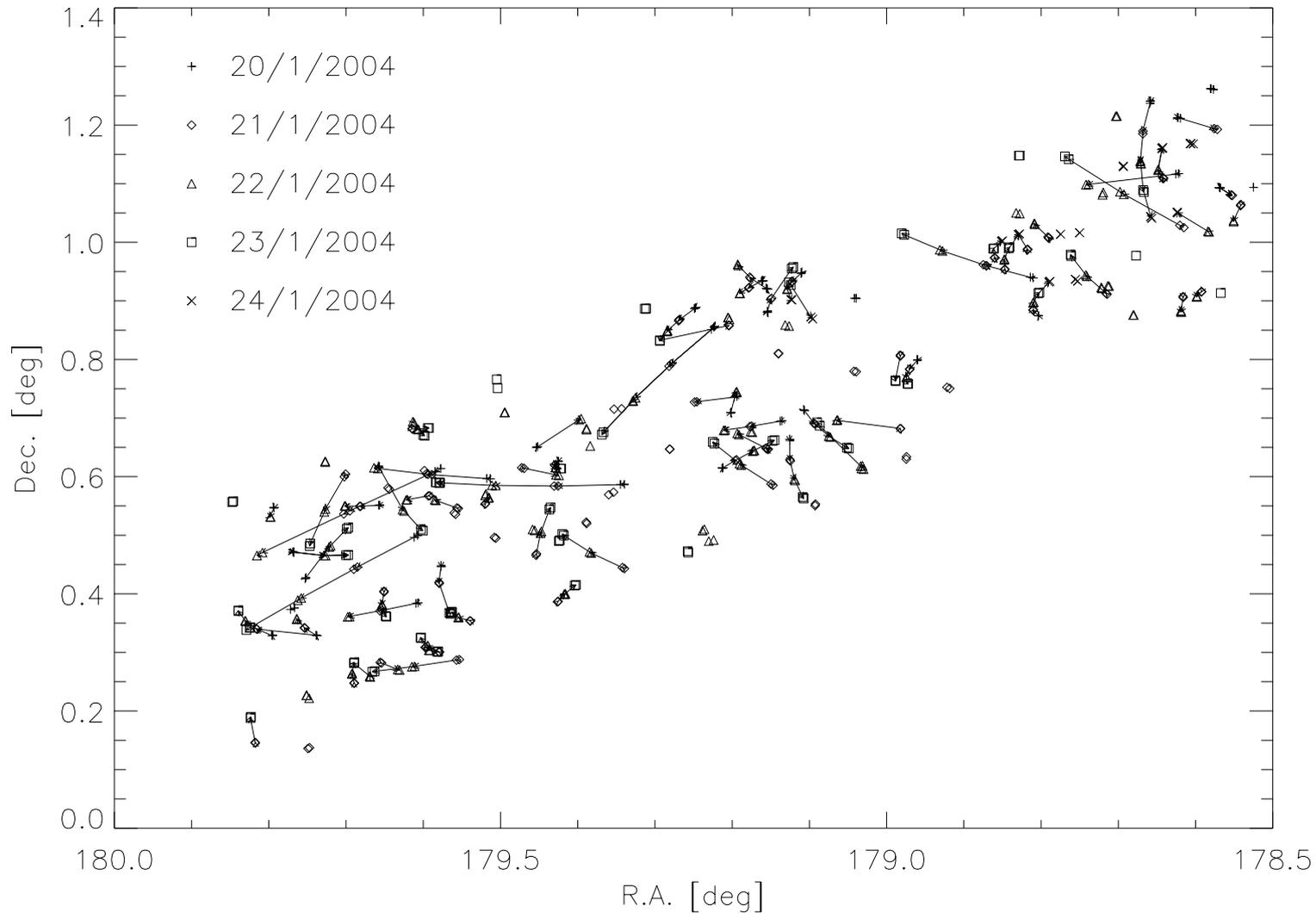


Figure 1: The figure shows 532 asteroid detections at $V \lesssim 26^m$ spread over five nights obtained with the VLT in January 2004. 76 preliminary inter-night identifications consisting of different combinations of 429 detections are shown with arrows pointing in the direction of motion.

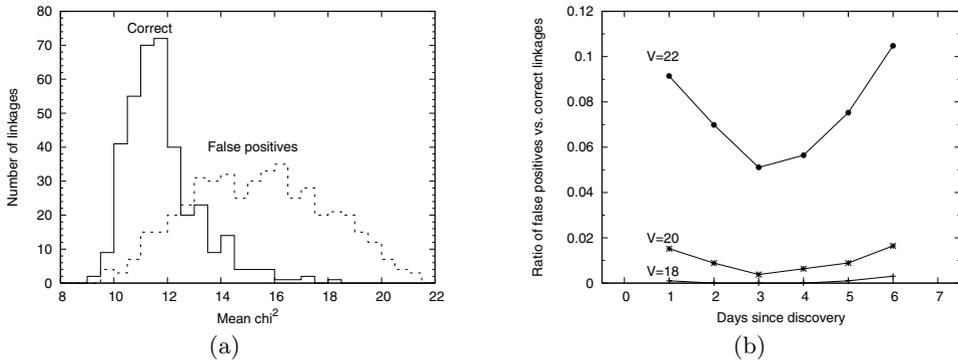


Figure 2: Results relating to the false positives problem obtained from simulated, opposition-centered NEO and MBO observations. (2a) The mean χ^2 -values of sample orbits are similar for both correct linkages and false positives, which makes them useless when deciding individually between multiple tentative linkages. (2b) Assuming the discovery night and the night one week after discovery being fixed, the figure shows the ratio of false positives vs correct linkages as a function of the epoch of the middle observations after all three nights have been processed with the identification software. It is evident, that the choice of observation strategy becomes more important as the limiting magnitude increases.

5.1. Minimising false positives

To examine more closely the problem of *false positives*, that is, the tentative linkages for which orbits can be found even though the observations actually belong to different objects, we generated simulated observations of NEOs and MBOs according to the following scheme. Two nightly, geocentric snapshots of the same region of the sky for three different limiting magnitudes (18, 20, and 22) were generated during eight sequential nights, resulting in 1100 objects on average per limiting magnitude. The direction of opposition at the discovery epoch was chosen to be the center of the observation window. The time interval between two nightly snapshots was approximately one hour, which is roughly the same as that of current search programs. Random Gaussian noise ($\sigma = 0.5''$) was also added to the observations.

The first issue was to find out whether there is a difference in the goodness of fit between correct and erroneous linkages. According to our results so far, it seems like there would not necessarily exist a statistical measure to decide individually (that is, one linkage at a time) between the two when the observational arcs are short and the numbers of observations are small in both cases. As seen in Figure 2a, the mean χ^2 -values of sample orbits are similar for both correct linkages and false positives. These results were obtained for the worst scenario case at limiting magnitude 22 and a time span of one week between the observational sets.

Examination of the effect of observing strategy on the number of erroneous linkages was the next logical step. Assuming that the observations at discovery and a week from discovery are fixed in time, we asked ourselves: What would be the optimum epoch for the middle observations so as to minimise the number of false linkages when all three nights have been linked? In this study we used the cumulative strategy, with the exception that only preliminary linkages between the first and second sets were processed with the third set. According to the results seen in Figure 2b, the observing strategy promoted by the

MPC (observations on the first, second, and eighth night) would not be the best choice for identification purposes. It seems like the strategy used by automatic surveys (equally distributed nights) would get better than the MPC strategy as the limiting magnitude increases. It should be stressed that the survey strategy is efficient only if large regions of the sky can be covered within a reasonable time. Otherwise a substantial number of objects will escape detection due to fast motion, which would not be the case when using the MPC strategy. It should also be noted that a single set (per limiting magnitude) of simulated observations obtained with a statistical tool is not statistically significant. It is, however, possible to draw preliminary conclusions of the results, as the outcome at all three limiting magnitudes point in the same direction.

6. Conclusions

The present method, and particularly the ephemerides address comparison, has been successfully tested with both VLT observations and simulated observations of near-Earth and main-belt objects.

A relevant issue, which was brought up by A. Milani, is to find out how the whole identification method and, particularly, the address comparison technique, scales when the number of objects is increased by two, or even three, orders of magnitude. Making the address comparisons between two sets containing 1,000 objects each takes a few minutes or less with current single-processor workstations. Assuming quadratical scaling, only the address comparison of sets containing 100,000 objects each would thus require three weeks of cpu-time on a workstation. Fortunately, the present method is easy to parallelise on several levels starting from orbital inversion and address comparison all the way to target (that is, asteroid) management. The scaling problems should therefore not be insuperable.

The ultimate goal of this research is to produce a real-time asteroid identification tool for ESA's astrometric space observatory Gaia, LONEOS, the Near-Earth Space Surveillance mission, and the Nordic Near-Earth Object Network. The tool should also prove useful to large-scale surveys connected to telescopes such as the Large Synoptic Survey Telescope and the Discovery Channel Telescope.

Acknowledgements

The authors wish to thank Ted Howell and the anonymous reviewer for many valuable comments on the manuscript. The research is supported by the Univ. Helsinki three-year research grant and by the Academy of Finland. MG is grateful to Magnus Ehrnrooth Foundation and Univ. Helsinki Chancellor for supporting his visit to the Lowell Observatory.

References

- Meadows, V.S., Bhattacharya, B., Reach, W.T., Grillmair, C., Noriega-Crespo, A., Ryan, E.L., Tyler, S.R., Rebull, L.M., Giorgini, J.D. & Elliot, J.L. 2004, *Astrophys. J. Suppl. Ser.* 154, 469
- Milani, A., La Spina, A., Sansaturio, M.E. & Chesley, S.R. 2000, *Icarus* 144, 39
- Milani, A., Sansaturio, M.E. & Chesley, S.R. 2001, *Icarus* 151, 150
- Muñonen, K. & Howell, E. 1993, *Icarus* 104, 255
- Muñonen, K., Howell, E. & Wasserman, L.H. 1994, *Planet. Space Sci.* 42, 307

- Muironen, K., Virtanen, J., Granvik, M. & Laakso, T. 2004, in *Three-Dimensional Universe with Gaia*, ESA Special Publications SP-576
- Virtanen, J., Muironen, K. & Bowell, E. 2001, *Icarus* 154, 412