

CLAIMS

What is claimed is:

1. A computer-implemented method for text address processing, the method executed by at least one processor of a text address processing apparatus, comprising the steps of:

determining, according to social relation circles of a plurality of users in a service system, address sets, each address set of the address sets comprising at least two original text addresses, wherein the social relation circles comprises a first social circle of a first user of the plurality of users and a second social relation circle of a second user of the plurality of users, the first social relation circle is based on an established relationship between the first user and one or more users in the first social relation circle, and a second social relation circle is based on an established relationship between the second user and one or more users in the second social relation circle, each social relation circle is associated with one address set of the address sets; and

performing normalization on the original text addresses in the address sets to obtain a target text address.

2. The method according to claim 1, wherein determining, according to the social relation circles of the plurality of users in the service system, the address sets comprises:

determining the first social relation circle of the first user and the second social relation circle of the second user in the service system; acquiring a text address used by the first user and one or more text addresses used by the one or more users in the first social relation circle to generate a first address set; and

acquiring a text address used by the second user and one or more text addresses used by the one or more users in the second social relation circle to generate a second address set.

3. The method according to claim 1, wherein performing normalization on the original text addresses in the address sets to obtain the target text address comprises:

determining, according to features of each two original text addresses in each address set, a degree of similarity between the two original text addresses; and

determining, according to the degree of similarity, whether the two original text addresses can be normalized into one of the two original text addresses, to obtain a target text address corresponding to the address set.

4. The method according to claim 3, wherein the determining, according to features of each two original text addresses in each address set, a degree of similarity between the two original text addresses comprises:

extracting at least one of a standard fragment feature, a longitude and latitude feature, and an alphanumeric feature of each of the two original text addresses; and

determining, according to each extracted feature, a degree of similarity between the two original text addresses corresponding to the feature.

5. The method according to claim 4, wherein the determining, according to each extracted feature, the degree of similarity between the two original text addresses corresponding to the feature comprises:

in response to the extracted feature being a standard fragment feature, determining the degree of similarity between the two original text addresses on a standard fragment feature dimension by using a SimHash algorithm;

in response to the extracted feature being a longitude and latitude feature, determining the degree of similarity between the two original text addresses on a longitude and latitude feature dimension by using a longitude and latitude distance algorithm; and

in response to the extracted feature being an alphanumeric feature, determining the degree of similarity between the two original text addresses on an alphanumeric feature dimension by using a Jaccard coefficient algorithm.

6. The method according to claim 1, further comprising:

determining, according to a corresponding relationship between target text addresses and original text addresses, at least two target text addresses corresponding to a same original text address; and

performing normalization on the at least two target text addresses.

7. The method according to claim 6, wherein the performing normalization on the at least two target text addresses comprises:

acquiring a fragment intersection of standard address fragments contained in original text addresses respectively corresponding to the at least two target text addresses; and

performing, according to the fragment intersection, normalization on the at least two target text addresses.

8. The method according to claim 7, wherein the performing, according to the fragment intersection, normalization on the at least two target text addresses comprises:

normalizing, in response to the fragment intersection representing one of the at least two target text addresses, the at least two target text addresses into the target text address that the fragment intersection represents.

9. The method according to claim 8, further comprising:

storing, the fragment intersection and the target text address that the fragment intersection represents into a feature knowledge base.

10. A text address processing apparatus, comprising:

a memory storing a set of instructions; and

one or more processors configured to execute the set of instruction to cause the apparatus to perform:

determining, according to social relation circles of a plurality of users in a service system, address sets, each address set of the address sets comprising at least two original text addresses, wherein the social relation circles comprises a first social circle of a first user of the plurality of users and a second social relation circle of a second user of the plurality of users, the first social relation circle is based on an established relationship between the first user and one or more users in the first social relation circle, and a second social relation circle is based on an established relationship between the second user and one or more users in the second social relation circle, each social relation circle is associated with one address set of the address sets; and

performing normalization on the original text addresses in the address sets to obtain a target text address.

11. The text address processing apparatus according to claim 10, wherein the one or more processors are configured to execute the set of instructions to cause the apparatus to further perform:

determining the first social relation circle of the first user and the second social relation circle of the second user in the service system;

acquiring a text address used by the first user and one or more text addresses used by the one or more users in the first social relation circle to generate a first address set; and

acquiring a text address used by the second user and one or more text addresses used by the one or more users in the second social relation circle to generate a second address set.

12. A non-transitory computer readable medium that stores a set of instructions that is executable by at least one processor of a text address processing apparatus to cause the text address processing apparatus to perform a method for text address processing, the method comprising:

determining, according to social relation circles of a plurality of users in a service system, address sets, each address set of the address sets comprising at least two original text addresses, wherein the social relation circles comprises a first social circle of a first user of the plurality of users and a second social relation circle of a second user of the plurality of users, the first social relation circle is based on an established relationship between the first user and one or more users in the first social relation circle, and a second social relation circle is based on an established relationship between the second user and one or more users in the second social relation circle, each social relation circle is associated with one address set of the address sets; and

performing normalization on the original text addresses in the address sets to obtain a target text address.

13. The non-transitory computer readable medium according to claim 12, wherein determining, according to the social relation circles of the plurality of users in the service system, the address sets comprises:

acquiring a text address used by the first user and one or more text addresses used by the one or more users in the first social relation circle to generate a first address set; and

acquiring a text address used by the second user and one or more text addresses used by the one or more users in the second social relation circle to generate a second address set.

14. The non-transitory computer readable medium according to claim 12, wherein performing normalization on the original text addresses in the address sets to obtain the target text address comprises:

determining, according to features of each two original text addresses in each address set, a degree of similarity between the two original text addresses; and

determining, according to the degree of similarity, whether the two original text addresses can be normalized into one of the two original text addresses, to obtain a target text address corresponding to the address set.

15. The non-transitory computer readable medium according to claim 14, wherein determining, according to features of each two original text addresses in each address set, a degree of similarity between the two original text addresses comprises:

extracting at least one of a standard fragment feature, a longitude and latitude feature, and an alphanumeric feature of each of the two original text addresses; and

determining, according to each extracted feature, a degree of similarity between the two original text addresses corresponding to the feature.

16. The non-transitory computer readable medium according to claim 15, wherein determining, according to each extracted feature, the degree of similarity between the two original text addresses corresponding to the feature comprises:

in response to the extracted feature being a standard fragment feature, determining the degree of similarity between the two original text addresses on a standard fragment feature dimension by using a SimHash algorithm;

in response to the extracted feature being a longitude and latitude feature, determining the degree of similarity between the two original text addresses on a longitude and latitude feature dimension by using a longitude and latitude distance algorithm; and

in response to the extracted feature being an alphanumeric feature, determining the degree of similarity between the two original text addresses on an alphanumeric feature dimension by using a Jaccard coefficient algorithm.

17. The non-transitory computer readable medium according to claim 12, wherein the set of instructions that is executable by the at least one processor of the text address processing apparatus to cause the text address processing apparatus to further perform:

determining, according to a corresponding relationship between target text addresses and original text addresses, at least two target text addresses corresponding to a same original text address; and

performing normalization on the at least two target text addresses.

18. The non-transitory computer readable medium according to claim 17, wherein performing normalization on the at least two target text addresses comprises:

acquiring a fragment intersection of standard address fragments contained in original text addresses respectively corresponding to the at least two target text addresses; and

performing, according to the fragment intersection, normalization on the at least two target text addresses.

19. The non-transitory computer readable medium according to claim 18, wherein performing, according to the fragment intersection, normalization on the at least two target text addresses comprises:

normalizing, in response to the fragment intersection representing one of the at least two target text addresses, the at least two target text addresses into the target text address that the fragment intersection represents.

20. The non-transitory computer readable medium according to claim 19, wherein the set of instructions that is executable by the at least one processor of the text address processing apparatus to cause the text address processing apparatus to further perform:

storing, the fragment intersection and the target text address that the fragment intersection represents into a feature knowledge base.