CLAIMS

1. A computer-implemented method for address analysis, comprising:

receiving a first address and a second address representing one or more physical locations involved in an online transaction, the first address including **n** morphemes and the second address including **m** morphemes, wherein a morpheme is a smallest semantic unit in an address, and wherein **n** and **m** are both natural numbers, the first address and/or the second address including English content;

determining first correlation values between the **n** morphemes and the **m** morphemes;

selecting a second morpheme of the **m** morphemes that best matches a first morpheme of the **n** morphemes based on the first correlation values and a preset algorithm;

recording a target correlation value between the first morpheme and the second morpheme as part of a set of **n** target correlation values;

obtaining a second correlation value between the first address and the second address based on the **n** target correlation values, wherein the second correlation value is a summation of a plurality of values of the **n** target correlation values; and

analyzing a correlation between the first address and the second address based on the second correlation value.

2. The computer-implemented method of claim 1, wherein determining the first correlation values comprises determining the first correlation values between each of the **n** morphemes and the **m** morphemes to obtain **n×m** first correlation values.

3. The computer-implemented method of claim 2, wherein selecting the second morpheme that best matches the first morpheme comprises:

constructing an **n×m** matrix based on the **n×m** first correlation values; and

pre-processing the **n×m** matrix based on the preset algorithm.

4. The computer-implemented method of claim 3, further comprising selecting the second morpheme based on the pre-processed **n×m** matrix.

5. The computer-implemented method of claim 1, wherein a value of the first correlation values is an edit distance value, a hamming distance value, a Jaccard distance value, an N-Gram distance value, a Jaro-Winkler (JW) distance value, or a cosine distance value.

6. The computer-implemented method of claim 1, further comprising pre-processing the first address by varying a writing order of the first address to account for regional differences.

7. A non-transitory, computer-readable medium storing one or more instructions executable by a computer system to perform operations for address analysis, comprising:

receiving a first address and a second address representing one or more physical locations involved in an online transaction, the first address including **n** morphemes and the second address including **m** morphemes, wherein a morpheme is a smallest semantic unit in an address, and wherein **n** and **m** are both natural numbers, the first address and/or the second address including English content;

determining first correlation values between the **n** morphemes and the **m** morphemes;

selecting a second morpheme of the **m** morphemes that best matches a first morpheme of the **n** morphemes based on the first correlation values and a preset algorithm;

recording a target correlation value between the first morpheme and the second morpheme as part of a set of **n** target correlation values;

obtaining a second correlation value between the first address and the second address based on the **n** target correlation values, wherein the second correlation value is a summation of a plurality of values of the **n** target correlation values; and

analyzing a correlation between the first address and the second address based on the second correlation value.

8. The non-transitory, computer-readable medium of claim 7, wherein determining the first correlation values comprises determining the first correlation values between each of the **n** morphemes and the **m** morphemes to obtain **n×m** first correlation values.

9. The non-transitory, computer-readable medium of claim 8, wherein selecting the second morpheme that best matches the first morpheme comprises:

constructing an **n×m** matrix based on the **n×m** first correlation values; and

pre-processing the **n×m** matrix based on the preset algorithm.

10. The non-transitory, computer-readable medium of claim 9, the operations further comprising selecting the second morpheme based on the pre-processed **n×m** matrix.

11. The non-transitory, computer-readable medium of claim 7, wherein a value of the first correlation values is an edit distance value, a hamming distance value, a Jaccard distance value, an N-Gram distance value, a Jaro-Winkler (JW) distance value, or a cosine distance value.

12. The non-transitory, computer-readable medium of claim 7, further comprising pre-processing the first address by varying a writing order of the first address to account for regional differences.

13. A computer-implemented system for address analysis, comprising:

one or more computers; and

one or more computer memory devices interoperably coupled with the one or more computers and having tangible, non-transitory, machine-readable media storing one or more instructions that, when executed by the one or more computers, perform one or more operations comprising:

receiving a first address and a second address representing one or more physical locations involved in an online transaction, the first address including **n** morphemes and the second address including **m** morphemes, wherein a morpheme is a smallest semantic unit in an address, and wherein **n** and **m** are both natural numbers, the first address and/or the second address including English content;

determining first correlation values between the **n** morphemes and the **m** morphemes;

selecting a second morpheme of the **m** morphemes that best matches a first morpheme of the **n** morphemes based on the first correlation values and a preset algorithm;

recording a target correlation value between the first morpheme and the second morpheme as part of a set of **n** target correlation values;

obtaining a second correlation value between the first address and the second address based on the **n** target correlation values, wherein the second correlation value is a summation of a plurality of values of the **n** target correlation values; and

analyzing a correlation between the first address and the second address based on the second correlation value.

14. The computer-implemented system of claim 13, wherein determining the first correlation values comprises determining the first correlation values between each of the **n** morphemes and the **m** morphemes to obtain **n×m** first correlation values.

15. The computer-implemented system of claim 14, wherein selecting the second morpheme that best matches the first morpheme comprises:

constructing an **n×m** matrix based on the **n×m** first correlation values; and

pre-processing the **n×m** matrix based on the preset algorithm.

16. The computer-implemented system of claim 15, the operations further comprising selecting the second morpheme based on the pre-processed **n×m** matrix.

17. The computer-implemented system of claim 13, wherein a value of the first correlation values is an edit distance value, a hamming distance value, a Jaccard distance value, an N-Gram distance value, a Jaro-Winkler (JW) distance value, or a cosine distance value.