

# TRANSACTION DATA PROCESSING SYSTEMS AND METHODS FOR ESTIMATING A GEOLOCATION OF A MERCHANT

## TECHNICAL FIELD

**[0001]** The present disclosure relates to the processing of payment card transaction data. In particular, the present disclosure relates to the estimation of a geolocation of a merchant using payment card transaction data.

## BACKGROUND

**[0002]** Payment networks which process and route payment card transactions between merchants and payment card issuers typically store geolocation information for merchants. This geolocation information may be used for transaction analysis and in fraud detection as part of the payment transaction authorization process.

**[0003]** It is often found to be the case that the geolocation data stored for merchants is inaccurate and does not match other data stored for a merchant. For example, the geolocation data may indicate latitude and longitude co-ordinates which do not match the location corresponding to the correct address of the merchant. This type of discrepancy can occur for different reasons: the incorrect geolocation may be input when a merchant is first added to a system, the merchant location may change and the geolocation may not be updated, or the same geolocation data may be added for multiple branches of the same merchant.

**[0004]** If incorrect or inaccurate geolocation data is stored for a merchant, then data analysis or fraud detection based on the geolocation data may yield incorrect results.

## SUMMARY

**[0005]** In accordance with a first aspect of the present invention, a transaction data processing system for estimating a geolocation of a target merchant is provided. The data processing system comprises: a processor, and a data storage device, data storage device storing computer program instructions operable to cause the processor to: extract target merchant transaction data indicative of a set of transactions made at the target merchant from a transaction database; for each transaction of the target merchant transaction data, identify a payment card associated with the respective transaction and thereby identifying a set of payment cards used by customers at the target merchant; for each payment card of the set of payment cards used by customers of the target merchant, identify transactions within a time period including the respective transaction made at the target merchant, and thereby identifying transactions made as part of a customer journey including the target merchant; for each transaction made as part of a customer journey including the target

merchant, identify a geolocation associated with the corresponding transaction and thereby identify a set of potentially neighboring merchant geolocations; and estimate the geolocation of the target merchant using geolocations from the set of potentially neighboring geolocations.

**[0006]** In an embodiment the data storage device stores instructions operable to cause the processor to estimate the geolocation of the target merchant using geolocations from the set of potentially neighboring geolocations by identifying a centroid of merchant locations from the set of potentially neighboring merchant geolocations.

**[0007]** In an embodiment the data storage device stores instructions operable to cause the processor to iteratively remove geolocations from the set of potentially neighboring merchant geolocations based on a distance between a centroid of geolocations of the set of potentially neighboring merchant geolocations and respective geolocations of the set of potentially neighboring merchant geolocations.

**[0008]** In an embodiment the data storage device stores instructions operable to cause the processor to cease the iterative removal of geolocations if a change in the centroid of set of potentially neighboring merchant geolocations between iterations is less than a threshold.

**[0009]** In an embodiment the set of transactions made at the target merchant from a transaction database correspond to transactions having a transaction time within a time window.

**[0010]** In an embodiment the data storage device stores instructions operable to cause the processor to identify transactions made as part of a customer journey including the target merchant by identifying physical transactions.

**[0011]** In an embodiment the data storage device stores instructions operable to cause the processor to identify physical transactions using a physical transaction flag in transaction data.

**[0012]** In an embodiment the data storage device stores instructions operable to cause the processor to store the estimated target merchant geolocation in a merchant database.

**[0013]** In an embodiment the data storage device instructions operable to cause the processor to use the estimated target merchant geolocation in a transaction authorization process.

**[0014]** According to a second aspect of the present disclosure, a data processing method for estimating a geolocation of a target merchant using transaction data is provided. The method comprises: extracting target merchant transaction data indicative of a set of transactions made at the target merchant from a transaction database; for each transaction

of the target merchant transaction data, identifying a payment card associated with the respective transaction and thereby identifying a set of payment cards used by customers at the target merchant; for each payment card of the set of payment cards used by customers of the target merchant, identifying transactions within a time period including the respective transaction made at the target merchant, and thereby identifying transactions made as part of a customer journey including the target merchant; for each transaction made as part of a customer journey including the target merchant, identifying a geolocation associated with the corresponding transaction and thereby identifying a set of potentially neighboring merchant geolocations; and estimating the geolocation of the target merchant using geolocations from the set of potentially neighboring geolocations.

**[0015]** Embodiments of the invention may be provided as a network of communicating devices (i.e. a “computerized network”). Embodiments of the invention may be also provided as a software application downloadable into a computer device to facilitate the method. The software application may be a computer program product, which may be stored on a non-transitory computer-readable medium on a tangible data-storage device (such as a storage device of a server, or one within a user device).

#### **BRIEF DESCRIPTION OF THE DRAWINGS**

**[0016]** In the following, embodiments of the present invention will be described as non-limiting examples with reference to the accompanying drawings in which:

FIG.1 is a block diagram showing a transaction data processing system for estimating a merchant geolocation according to an embodiment of the present invention;

FIG.2 is a block diagram showing merchant data processed in embodiments of the present invention;

FIG.3 is a block diagram showing transaction data processed in embodiments of the present invention;

FIG.4 is a block diagram of a network for processing transactions to generate transaction data which may be processed in embodiments of the present invention;

FIG.5 is a block diagram showing functional modules of a transaction data processing system according to an embodiment of the present invention;

FIG.6 is a flowchart showing a method of estimating a merchant geolocation according to an embodiment of the present invention;

FIG.7 shows an example of a set of customer journeys including a target merchant extracted in an embodiment of the present invention;

FIG.8 is a flow chart showing a method of estimating a target merchant geolocation from potentially neighboring merchant geolocations according to an embodiment of the present invention;

FIGs.9A to 9C show geolocations of potentially neighboring merchants used in iterations of estimating a geolocation of a target merchant in an embodiment of the present invention;

FIGs.10A to 10C are bar graphs showing the distance between the centroid of the geolocations of the potentially neighboring merchants and error between the centroid of the geolocations of the potentially neighboring merchant locations and the actual location of the target merchant for iterations of the method of estimating a geolocation of a target merchant in an embodiment of the present invention;

FIG.11 is a graph showing the error in the geolocation of the target merchant for different iterations described in relation to FIGs.9A to 9C and 10A to 10C;

FIG.12 is a block diagram showing a process for authorizing payment card transactions using merchant geolocations estimated according to embodiments of the present invention; and

FIG.13 is a block diagram showing a technical architecture of a payment transaction data processing system according to an embodiment of the present invention.

#### **DETAILED DESCRIPTION**

**[0017]** As used herein, the term “payment card” refers to any cashless payment device associated with a payment account, such as a credit card, a debit card, a prepaid card, a charge card, a membership card, a promotional card, a frequent flyer card, an identification card, a prepaid card, a gift card, and/or any other device that may hold payment account information, such as mobile phones, Smartphones, personal digital assistants (PDAs), key fobs, transponder devices, NFC-enabled devices, and/or computers. In addition, the “payment card” may exist as application-based stored value accounts associated with an electronic wallet service provider tied in with an issuer institution.

**[0018]** The term “institution” is used in a sense which is not necessarily limited to organizations which are legally constituted as banks, since in some jurisdictions other organizations may be permitted to maintain financial accounts such as a payment card account. An institution may be one of the following: a bank, a financial technology company, a telecommunication company or a financial institution.

**[0019]** As used in this disclosure, the terms “component,” “module,” “system,” “apparatus,” “interface,” or the like are generally intended to refer to a computer-related entity, either hardware, a combination of hardware and software, software, or software in execution. For example, a component or a module may be, but is not limited to being, a process running on a processor, a processor, an object, an executable, a thread of execution, a program, and/or a computer. By way of illustration, both an application running on a controller and the controller can be a component or a module. One or more components/modules may reside within a process and/or thread of execution and a component may be localized on one computer and/or distributed between two or more computers.

**[0020]** Furthermore, the claimed subject matter may be implemented as a method, apparatus, or article of manufacture using standard programming and/or engineering techniques to produce software, firmware, hardware, or any combination thereof to control a computer to implement the disclosed subject matter. For instance, the claimed subject matter may be implemented as a computer-readable medium embedded with a computer executable program, which encompasses a computer program accessible from any computer-readable storage device or storage media. For example, computer readable media can include but are not limited to magnetic storage devices (e.g., hard disk, floppy disk, magnetic strips . . . ), optical disks (e.g., compact disk (CD), digital versatile disk (DVD) . . . ), smart cards, and flash memory devices (e.g., card, stick, key drive . . . ).

**[0021]** FIG.1 is a block diagram showing a transaction data processing system for estimating a merchant geolocation according to an embodiment of the present invention. As shown in FIG.1, the transaction data processing system 100 is coupled to a merchant database 110 and a transaction database 130. The transaction data processing system 110 may be coupled to and/or form part of a payment network for processing payment transactions. An example of such a payment network the processing carried out by the payment network are described in more detail below with reference to FIG.4.

**[0022]** The merchant database 110 stores merchant data which relates to merchants coupled to the payment network. An example of the merchant data stored in the merchant database 110 is shown in FIG.2. The transaction database 130 stores transaction data which relates to transactions carried out on the payment network. An example of the transaction data stored in the transaction database 130 is shown in FIG.3.

**[0023]** FIG.2 is a block diagram showing merchant data processed in embodiments of the present invention. The merchant data 112 is stored in the merchant database 110 shown in FIG.1. The merchant database 110 may store multiple merchant records each taking the of the merchant data 112 shown in FIG.2. The merchant data 112 comprises a merchant identifier 114, a merchant geolocation 116, a merchant address 118 and a merchant postal

code 120. The merchant identifier 114 is a code that uniquely identifies a merchant. The merchant identifier may be provided to a merchant by a payment network or acquirer when the merchant opens an account. The merchant identifier 114 in the present disclosure uniquely identifies a specific branch or location of the merchant. Thus, if a shop or restaurant has multiple stores or branches at different locations, each store or branch is considered a separate merchant and has a unique merchant identifier. The merchant geolocation 116 indicates a recorded geolocation for the merchant. The merchant geolocation may indicate a latitude and longitude of the merchant location. The merchant geolocation may be formatted according to the ISO 6709 standard. The merchant address 118 indicates the street address of the merchant. The merchant postal code 120 indicates the postal code of the merchant.

**[0024]** The data included in the merchant geolocation 116, the merchant address 118 and the merchant postal code 120 may be generated when a merchant registers or is on-boarded with an acquirer bank. In some cases, this data may be inaccurate. For example, a merchant may move premises, and the merchant geolocation 116 and other data may not be updated. Sometimes, when a merchant is registered with an acquirer bank, the same merchant geolocation 116 may be entered for all branches of the merchant in one city for example. In some cases, the merchant geolocation 116 may be automatically generated from the merchant postal code 120, for example by looking up a geolocation associated the postal code. However, if the same postal code covers a large geographical area, then the generated geolocation may be inaccurate.

**[0025]** As is described in more detail below, the merchant geolocation 116 may be used to detect possibly fraudulent transactions by comparing the geographical distance between locations of transactions and the timing between transactions. The merchant geolocations may also be used for other types of transaction analysis. However, if the merchant geolocation is inaccurate for the reasons discussed in the paragraph above, then this any analysis carried out using the merchant geolocation would also be inaccurate.

**[0026]** The transaction data processing system 100 shown in FIG.1 processes transaction data to estimate merchant geolocations and may be used to improve the accuracy of the merchant geolocation 116 stored in the merchant database 110.

**[0027]** FIG.3 is a block diagram showing transaction data processed in embodiments of the present invention. The transaction data 132 is stored in the transaction database 130 shown in FIG.1. Transaction database 130 may store multiple records each corresponding to an individual transaction and each taking the form of the transaction data 132 shown in FIG.3. The transaction data 132 comprises a payment card identifier 134, a merchant identifier 136, an indication of the transaction time / date 138, a physical transaction flag

140 and an indication of a transaction amount 142. The payment card identifier 134 indicates the payment card used for the transaction and may comprise a primary account number (PAN) associated with the payment card. The merchant identifier 136 indicates the merchant at which the transaction took place and corresponds to the merchant identifier 114 described above with reference to FIG.2. The transaction time /date 138 indicates the time and date at which the transaction took place. The physical transaction flag 140 indicates whether the transaction was a physical transaction (also known as a card-present transaction) in which the payment card was physically present at the merchant. If the transaction was a card-not-present transaction, then the physical transaction flag indicates that the transaction was not a physical transaction. The transaction amount 142 indicates the amount or ticket size of the transaction.

**[0028]** FIG.4 is a block diagram of a network for processing transactions to generate transaction data which may be processed in embodiments of the present invention. The network 400 comprises a transaction data processing system 110 for estimating a merchant geolocation. As shown in FIG.4, the transaction data processing system 110 receives transaction data from a payment network 170 such as the Banknet payment network operated by Mastercard.

**[0029]** The payment network 170 acts as an intermediary during a transaction being made by a cardholder 152 using a payment card 160 at a merchant terminal 162 of a merchant 154. In particular, the cardholder 152 may present the payment card 160 to merchant terminal 162 of merchant 164 as payment for goods or services. The merchant terminal 162 may be a point of sale (POS) device such as a magnetic strip reader, chip reader or contactless payment terminal, for example. A merchant 154 may operate one or a plurality of merchant terminals 162. The merchant terminal 162 communicates with an acquirer computer system 168 of a bank or other institution with which the merchant 154 has an established account, in order to request authorization for the amount of the transaction (sometimes referred to as ticket size) from the acquirer system 168. In some embodiments, if the merchant 154 does not have an account with the acquirer 168, the merchant terminal 162 can be configured to communicate with a third-party payment processor 166 which is authorized by acquirer 168 to perform transaction processing on its behalf, and which does have an account with the acquirer entity.

**[0030]** The acquirer system 168 routes the transaction authorization request from the merchant terminal 162 to computer systems of the payment network 170. The transaction authorization request is then routed by payment network 170 to computer systems of the appropriate issuer institution (e.g., issuer 174) based on information contained in the transaction authorization request. The issuer institution 174 is authorized by payment

network 170 to issue payment devices 160 on behalf of customers 152 to perform transactions over the payment network 170. Issuer 174 also provides funding of the transaction to the payment network 170 for transactions that are approved.

**[0031]** The computer systems of the issuer 174 analyze the authorization request to determine the account number submitted by the payment card 160, and based on the account number, determine whether the account is in good standing and whether the transaction amount is covered by the cardholder's account balance or available credit. Based on this, the transaction can be approved or declined, and an authorization response message transmitted from issuer 174 to the payment network 170, which then routes the authorization response message to the acquirer system 178. The acquirer system 178, in turn, sends the authorization response message to merchant terminal 162. If the authorization response message indicates that the transaction is approved, then the account of the merchant 154 (or of the payment processor 166 if appropriate) is credited by the amount of the transaction.

**[0032]** During each authorization request as described in the previous paragraphs, the payment network 170 stores transaction information in a transactions database 110 which forms part of a data warehousing system 180. The data warehousing system 180 may be distributed over multiple devices which are in communication with one another over a communications network such as a local-area or wide-area network. As shown in FIG.4, the data warehousing system 180 also stores merchant database 130. The data warehousing system 180 may store copies of the transaction data, and/or cleaned and/or aggregated data which are transformed versions of the transaction data.

**[0033]** The transaction records may comprise a plurality of fields, including acquirer identifier/card acceptor identifier (the combination of which uniquely defines the merchant); merchant category code (also known as card acceptor business code), that is, an indication of the type of business the merchant is involved in (for example, an airline, travel agent or hotel); cardholder base currency (i.e., U.S. Dollars, Euros, Yen, etc.); the transaction environment or method being used to conduct the transaction; product specific data such as SKU line item data; the transaction type; card identifier (e.g., card number); time and date; location (full address and/or GPS data); transaction amount (also referred to herein as ticket size); terminal identifier (e.g., merchant terminal identifier or ATM identifier); and response code (also referred to herein as authorization code). Other fields may be present in each transaction record.

**[0034]** Each terminal identifier may be associated with a merchant 154, for example in a merchant database of the payment network 170. Typically, a particular merchant 154 will

have a plurality of merchant terminal identifiers, corresponding to merchant terminals 162, associated with it.

**[0035]** FIG.5 is a block diagram showing functional modules of a transaction data processing system according to an embodiment of the present invention. As shown in FIG.5, the transaction data processing system 100 comprises a transaction data extraction module 502, a transaction data analysis module 504, a merchant geolocation look up module 506, a target merchant geolocation estimation module 508 and a merchant data update module 510. Each of the modules shown in FIG.5 may be implemented as hardware, software or a combination of hardware and software. An example of a technical architecture of the transaction data processing system 200 is described in detail below with reference to FIG.12. The transaction data extraction module 502 is operable to extract transaction data from the transaction database 130. The transaction data analysis module 504 is operable to analyze the extracted transaction data. The merchant geolocation look up module 506 is operable to look up merchant geolocation data from the merchant database 110. The target merchant geolocation estimation module 508 is operable to estimate the geolocation of the target merchant based using the locations of potentially neighboring merchants. The merchant data update module 510 is operable to update merchant data stored in the merchant database 110 including updating the merchant geolocation for the target merchant.

**[0036]** FIG.6 is a flowchart showing a method of estimating a merchant geolocation according to an embodiment of the present invention. The method 600 shown in FIG.6 is carried out by the transaction data processing system 100 shown in FIG.5.

**[0037]** In step 602, the transaction data extraction module 502 of the transaction data processing system 100 extracts target merchant transaction data from the transaction database 130. The transaction data extraction module 502 extracts a set of transactions having a merchant identifier 136 which matches the merchant identifier of the target merchant. In some embodiments, the target merchant transaction data may correspond to transactions made in a given time period, for example transactions made at the target merchant within a window of 2 hours. The length of this time window may be varied, and may be transactions within a given month, a week, a number of days or a number of hours on a given day. Thus, the transaction data extraction module 502 may execute a query to extract transaction data corresponding to transactions having a merchant identifier which matches the merchant identifier of the target merchant and having a transaction time / date within the time period under consideration.

**[0038]** In step 604, the transaction data analysis module 504 identifies a payment card associated with each transaction of the target merchant transaction data extracted in step

602. The transaction data analysis module 504 may analyze the target merchant transaction data and generate a list of the payment card identifier 134 on each transaction of the target merchant transaction data. Thus, a set of payment cards used by customers of the target merchant in the time window under consideration is identified.

**[0039]** In step 606, the transaction data analysis module 504 identifies transactions made using the payment cards identified in step 604 within a time period including the respective transaction at the target merchant. Step 606 may involve retrieving transactions from the transaction database 130 made using one of the payment cards identified in step 604 and having a transaction time within a buffer period of the transaction at the target merchant. This buffer period may be, for example 4 hours, thus if the transaction made at the target merchant was made at 16:00 on a given day, transactions using the same payment card having a transaction time between 14:00 and 18:00 on the given day would be extracted. In some embodiments, the transactions identified in step 606 may be limited to physical transactions, that is, card-not-present transactions such as e-commerce transactions may be excluded. The transaction data analysis module 504 may use the physical transaction flag 140 to identify physical transactions.

**[0040]** The locations of the transactions extracted in step 606 correspond to a customer journey which includes the target merchant. FIG.7 shows an example of a set of such customer journeys.

**[0041]** FIG.7 shows an example of a set of customer journeys including a target merchant extracted in an embodiment of the present invention. FIG.7 shows the journeys of five customers: customer A follows a journey 710 which includes merchant having a set of geolocations 712; customer B follows a journey 720 which includes merchant having a set of geolocations 722, customer C follows a journey 730 which includes merchant having a set of geolocations 732, and customer D follows a journey 740 which includes merchant having a set of geolocations 742. Each of the customer journeys includes the target merchant which has a target merchant geolocation 702 shown by a square in FIG.7.

**[0042]** As can be seen from FIG.7, the target merchant geolocation 702 is located at approximately the centroid of all of the points on the five customer journeys. Embodiments of the present invention use utilize this approximation to estimate the location of the target merchant. Since the geolocations of merchants on the customer journeys of customers who also visited the target merchant are potentially neighboring geolocations of the target merchant, they can be used to estimate the geolocation of the target merchant. Thus, the identified locations are potentially neighboring geolocations of the target merchant.

**[0043]** Returning now to FIG.6, in step 610 the target merchant geolocation estimation module 508 of the transaction data processing system 100 uses the potentially neighboring merchant geolocations identified in step 608 to estimate the geolocation of the target merchant. Step 610 may comprise estimating a centroid of the potentially neighboring merchant geolocations. In some embodiments, outlying geolocations among the potentially neighboring merchant geolocations may be excluded when estimating the centroid. Step 610 may be carried out as an iterative process in which an initial centroid is estimated and then outlying potentially neighboring merchant geolocations are identified and excluded from subsequent iterations. FIG.8 shows an example of such a method.

**[0044]** FIG.8 is a flow chart showing a method of estimating a target merchant geolocation from potentially neighboring merchant geolocations according to an embodiment of the present invention. The method 800 shown in FIG.8 may be carried out by the target merchant geolocation estimation module 508 of the transaction data processing system 100 shown in FIG.1 and FIG.5.

**[0045]** In step 802, the method 800 takes the geolocations of the potentially neighboring merchants as input. As discussed above, these are geolocations of merchants at which transactions were made within a buffer period of a transaction at the target merchant.

**[0046]** In step 804 it is determined whether there are more than 2 data points left. If there are not more than 2 data points, then the method moves to step 806 and it is determined that a centroid cannot be found.

**[0047]** If it is determined that there are more than two data points (geolocations) left, then in step 808, the centroid of the data points is found. The centroid may be determined as the arithmetic mean of all the geolocations under consideration.

**[0048]** In step 810, the distance of each remaining geolocation from the centroid is calculated and points which are outside 0.5 standard deviations are removed.

**[0049]** In step 812, it is determined whether the new centroid is within a threshold distance of the centroid calculated in the previous iteration. The threshold distance may be set in the range 200m to 500m and may be for example 350m. If the new centroid is not within the threshold distance of the previous centroid, then the method returns to step 804.

**[0050]** If the new centroid is within the threshold distance of the previous centroid, the method moves to step 814 in which the geolocation of the target merchant is estimated as the centroid.

**[0051]** Returning now to FIG.6, once the target merchant geolocation has been estimated, in step 612 the merchant data update module 510 updates the stored geolocation of the

target merchant in the merchant database 110 with the estimated target merchant geolocation.

**[0052]** An example of the method of estimating a target merchant geolocation from potentially neighboring merchant geolocations shown in FIG.8 will now be described with reference to FIGs 9A to 9C, FIGs 10A to 10C and FIG.11.

**[0053]** FIGs.9A to 9C show geolocations of potentially neighboring merchants used in iterations of estimating a geolocation of a target merchant in an embodiment of the present invention.

**[0054]** FIGs.10A to 10C are bar graphs showing the distance between the centroid of the geolocations of the potentially neighboring merchants and error between the centroid of the geolocations of the potentially neighboring merchant locations and the actual location of the target merchant for iterations of the method of estimating a geolocation of a target merchant in an embodiment of the present invention.

**[0055]** The example relates example merchant located in New York. Transactions made within two-hour temporal window were extracted. In this example the temporal window was 13:00 to 15:00. This gave a total number of 36 customers. Transactions made within a four hour buffer of the transaction at the target merchant were extracted for each customer. For example if a customer made a transaction at the target merchant at 14:00, then transactions made using the same payment card between 12:00 and 16:00 were included in the geolocation analysis. This gave a total of 270 geolocation data points.

**[0056]** FIG.9A shows the geolocations of all of the potentially neighboring merchants. As can be seen in FIG.9A the majority of the potentially neighboring merchant geolocations are close to the centroid 900A, but there are many outliers.

**[0057]** FIG.10A shows the number of potentially neighboring merchant locations for different distances from the centroid. The majority of the potentially neighboring merchant geolocations are less than 2500m from the centroid. As shown in FIG.10A, the standard deviation of the distance potentially neighboring merchant geolocations and the centroid is 202.254km.

**[0058]** As described above in relation to FIG.8, the method comprises calculating the distance between each potentially neighboring merchant geolocation and the centroid and discarding those geolocations that are more than 0.5 standard deviation away.

**[0059]** FIG.9B shows the geolocations of the potentially neighboring merchants in the second iteration of the method. It is noted that the latitude and longitude scales are changed compared to FIG.9A. The centroid from the previous iteration 900A and the centroid of the

present iteration 900B. As can be seen there is a small difference between the two centroids. Since this distance is greater than the threshold distance, the method continues to the next iteration.

**[0060]** FIG.10B shows that the majority of the potentially neighboring merchant geolocations are less than 400m from the centroid. As shown in FIG.10B, the standard deviation of the distances of the potentially neighboring merchant geolocations from the centroid is 20.368km.

**[0061]** FIG.9C shows the geolocations of the remaining potentially neighboring merchant geolocations after the third iteration. As can be seen in FIG.9C, the new centroid 900C is very close to the centroid of the previous iteration. Thus, the method terminates after the third iteration because the difference in centroid is less than the threshold (which in this case is 350m).

**[0062]** FIG.10C shows the distances of the potentially neighboring merchant locations from the centroid. As can be seen from FIG.10C the majority of the potentially neighboring merchant locations are around 200m from the centroid.

**[0063]** FIG.11 is a graph showing the error in the geolocation of the target merchant for different iterations described in relation to FIGs.9A to 9C and 10A to 10C. FIG.11 shows the mean absolute error in Haversine distance for each iteration. As can be seen from FIG.11, after 3 iterations, the error approaches zero. For this example, the merchant location is accurate to less than 1km. As is described below, one use to the estimated merchant geolocations is fraud detection as part of the transaction authorization process. An accuracy to 1km is sufficient for this purpose.

**[0064]** FIG.12 is a block diagram showing a process for authorizing payment card transactions using merchant geolocations estimated according to embodiments of the present invention. The inputs 1010 to the process 1000 are transaction data 1012 and 3<sup>rd</sup> party merchant attributes 1014. The transaction data 1012 is used by a geolocation estimation 1020 process which functions as described above to estimate the geolocation of a target merchant using the transaction data 1012. The 3<sup>rd</sup> party merchant attributes 1014 are matched to the merchants through a matching process 1030 which determines the corresponding merchant identifier for merchant data. The 3<sup>rd</sup> party merchant attributes 1014 may be produced by associating precise locations from third party address databases with specific merchants, or may use data stored in the merchant database 110 such as the merchant address 118 to retrieve relevant information using unique identifiers along with data from business partners, government regulations, credit and commercial organisations that share data, utility companies, self-identification, manual collection and internet activity.

The result of the geolocation estimation 1020 and the matching process 1030 is a set of merchant data 1040 which comprises improved merchant data 1042 and estimated geolocations 1044 for merchants.

**[0065]** Transaction authorization requests are input into a real-time authorization process 1050. This transaction authorization process 1050 is carried out by a fraud management system 1060 which may form part of the payment network 170 shown in FIG.4. The fraud management system 1060 uses the estimated geolocation 1044 data for merchants in the transaction authorization process. As shown in FIG.12, a transaction may be declined 1070 if the timing between two transactions at different merchants is less than the estimated travel time between the estimated geolocation of the two respective merchants. The transaction may be approved if the timing between two transactions is greater than the estimated travel time between the two respective merchants. Additional factors may also be taken into account when determining whether to approve or decline a transaction authorization request.

**[0066]** FIG.13 is a block diagram showing a technical architecture 500 of the transaction data processing system 100 for performing steps of exemplary methods described above. Typically, the methods are implemented by a number of computers each having a data-processing unit. The block diagram as shown in FIG.13 illustrates a technical architecture 500 of a computer which is suitable for implementing one or more embodiments herein.

**[0067]** The technical architecture 500 includes a processor 522 (which may be referred to as a central processor unit or CPU) that is in communication with memory devices including secondary storage 524 (such as disk drives), read only memory (ROM) 526, random access memory (RAM) 228. The processor 522 may be implemented as one or more CPU chips. The technical architecture 520 may further comprise input/output (I/O) devices 530, and network connectivity devices 532.

**[0068]** The secondary storage 524 is typically comprised of one or more disk drives or tape drives and is used for non-volatile storage of data and as an over-flow data storage device if RAM 528 is not large enough to hold all working data. Secondary storage 224 may be used to store programs which are loaded into RAM 528 when such programs are selected for execution. In this embodiment, the secondary storage 524 has a transaction data extraction module 502, a transaction data analysis module 504, a merchant geolocation look up module 506, a target merchant geolocation estimation module 508, and a merchant data update module 510, comprising non-transitory instructions operative by the processor 522 to perform various operations of the method of the present disclosure. As depicted in FIG.13, the modules 502-510 are distinct modules which perform respective functions implemented by the transaction data processing system 100. It will be appreciated that the

boundaries between these modules are exemplary only, and that alternative embodiments may merge modules or impose an alternative decomposition of functionality of modules. For example, the modules discussed herein may be decomposed into sub-modules to be executed as multiple computer processes, and, optionally, on multiple computers. Moreover, alternative embodiments may combine multiple instances of a particular module or sub-module. It will also be appreciated that, while a software implementation of the modules 502-510 is described herein, these may alternatively be implemented as one or more hardware modules (such as field-programmable gate array(s) or application-specific integrated circuit(s)) comprising circuitry which implements equivalent functionality to that implemented in software. The ROM 526 is used to store instructions and perhaps data which are read during program execution. The secondary storage 524, the RAM 528, and/or the ROM 526 may be referred to in some contexts as computer readable storage media and/or non-transitory computer readable media.

**[0069]** The I/O devices may include printers, video monitors, liquid crystal displays (LCDs), plasma displays, touch screen displays, keyboards, keypads, switches, dials, mice, track balls, voice recognizers, card readers, paper tape readers, or other well-known input devices.

**[0070]** The network connectivity devices 532 may take the form of modems, modem banks, Ethernet cards, universal serial bus (USB) interface cards, serial interfaces, token ring cards, fiber distributed data interface (FDDI) cards, wireless local area network (WLAN) cards, radio transceiver cards that promote radio communications using protocols such as code division multiple access (CDMA), global system for mobile communications (GSM), long-term evolution (LTE), worldwide interoperability for microwave access (WiMAX), near field communications (NFC), radio frequency identity (RFID), and/or other air interface protocol radio transceiver cards, and other well-known network devices. These network connectivity devices 532 may enable the processor 522 to communicate with the Internet or one or more intranets. With such a network connection, it is contemplated that the processor 522 might receive information from the network, or might output information to the network in the course of performing the method operations described herein. Such information, which is often represented as a sequence of instructions to be executed using processor 522, may be received from and outputted to the network, for example, in the form of a computer data signal embodied in a carrier wave.

**[0071]** The processor 522 executes instructions, codes, computer programs, scripts which it accesses from hard disk, floppy disk, optical disk (these various disk based systems may all be considered secondary storage 524), flash drive, ROM 526, RAM 528, or the network connectivity devices 532. While only one processor 522 is shown, multiple processors may

be present. Thus, while instructions may be discussed as executed by a processor, the instructions may be executed simultaneously, serially, or otherwise executed by one or multiple processors.

**[0072]** It is understood that by programming and/or loading executable instructions onto the technical architecture 500, at least one of the CPU 522, the RAM 528, and the ROM 526 are changed, transforming the technical architecture 500 in part into a specific purpose machine or apparatus having the novel functionality taught by the present disclosure. It is fundamental to the electrical engineering and software engineering arts that functionality that can be implemented by loading executable software into a computer can be converted to a hardware implementation by well-known design rules.

**[0073]** Whilst the foregoing description has described exemplary embodiments, it will be understood by those skilled in the art that many variations of the embodiments can be made within the scope of the invention as defined by the claims. Moreover, features of one or more embodiments may be mixed and matched with features of one or more other embodiments.

## CLAIMS

1. A transaction data processing system for estimating a geolocation of a target merchant, the data processing system comprising: a processor, and a data storage device, data storage device storing computer program instructions operable to cause the processor to:
  - extract target merchant transaction data indicative of a set of transactions made at the target merchant from a transaction database;
  - for each transaction of the target merchant transaction data, identify a payment card associated with the respective transaction and thereby identifying a set of payment cards used by customers at the target merchant;
  - for each payment card of the set of payment cards used by customers of the target merchant, identify transactions within a time period including the respective transaction made at the target merchant, and thereby identifying transactions made as part of a customer journey including the target merchant;
  - for each transaction made as part of a customer journey including the target merchant, identify a geolocation associated with the corresponding transaction and thereby identify a set of potentially neighboring merchant geolocations; and
  - estimate the geolocation of the target merchant using geolocations from the set of potentially neighboring geolocations.
2. A data processing system according to claim 1, wherein the data storage device stores instructions operable to cause the processor to estimate the geolocation of the target merchant using geolocations from the set of potentially neighboring geolocations by identifying a centroid of merchant locations from the set of potentially neighboring merchant geolocations.
3. A data processing system according to claim 2, wherein the data storage device stores instructions operable to cause the processor to iteratively remove geolocations from the set of potentially neighboring merchant geolocations based on a distance between a centroid of geolocations of the set of potentially neighboring merchant geolocations and respective geolocations of the set of potentially neighboring merchant geolocations.
4. A data processing system according to claim 3, wherein the data storage device stores instructions operable to cause the processor to cease the iterative removal of geolocations if a change in the centroid of set of potentially neighboring merchant geolocations between iterations is less than a threshold.

5. A data processing system according to any preceding claim, wherein the set of transactions made at the target merchant from a transaction database correspond to transactions having a transaction time within a time window.
6. A data processing system according to any preceding claim, wherein the data storage device stores instructions operable to cause the processor to identify transactions made as part of a customer journey including the target merchant by identifying physical transactions.
7. A data processing system according to claim 6, wherein the data storage device stores instructions operable to cause the processor to identify physical transactions using a physical transaction flag in transaction data.
8. A data processing system according to any preceding claim, wherein the data storage device stores instructions operable to cause the processor to store the estimated target merchant geolocation in a merchant database.
9. A data processing system according to any preceding claim, wherein the data storage device instructions operable to cause the processor to use the estimated target merchant geolocation in a transaction authorization process.
10. A data processing method for estimating a geolocation of a target merchant using transaction data, the method comprising:
  - extracting target merchant transaction data indicative of a set of transactions made at the target merchant from a transaction database;
  - for each transaction of the target merchant transaction data, identifying a payment card associated with the respective transaction and thereby identifying a set of payment cards used by customers at the target merchant;
  - for each payment card of the set of payment cards used by customers of the target merchant, identifying transactions within a time period including the respective transaction made at the target merchant, and thereby identifying transactions made as part of a customer journey including the target merchant;
  - for each transaction made as part of a customer journey including the target merchant, identifying a geolocation associated with the corresponding transaction and thereby identifying a set of potentially neighboring merchant geolocations; and
  - estimating the geolocation of the target merchant using geolocations from the set of potentially neighboring geolocations.

11. A data processing method according to claim 10, wherein estimating the geolocation of the target merchant using geolocations from the set of potentially neighboring geolocations comprises identifying a centroid of merchant locations from the set of potentially neighboring merchant geolocations.

12. A data processing method according to claim 11 further comprising iteratively removing geolocations from the set of potentially neighboring merchant geolocations based on a distance between a centroid of geolocations of the set of potentially neighboring merchant geolocations and respective geolocations of the set of potentially neighboring merchant geolocations.

13. A data processing method according to claim 12, further comprising ceasing the iterative removal of geolocations if a change in the centroid of set of potentially neighboring merchant geolocations between iterations is less than a threshold.

14. A data processing method according to any one of claims 10 to 13, wherein the set of transactions made at the target merchant from a transaction database correspond to transactions having a transaction time within a time window.

15. A data processing method according to any one of claims 10 to 14, wherein identifying transactions made as part of a customer journey including the target merchant by comprises identifying physical transactions.

16. A data processing method according to claim 15 wherein identifying physical transactions comprises using a physical transaction flag in transaction data.

17. A data processing method according to any one of claims 10 to 16, further comprising storing the estimated target merchant geolocation in a merchant database.

18. A data processing method according to any one of claims 10 to 17, further comprising using the estimated target merchant geolocation in a transaction authorization process.

19. A non-transitory computer readable medium carrying computer executable instructions which when executed on at least one processor cause the at least one processor to carry out a method according to any one of claims 10 to 18.