Lecture 8
Outlier Detection

Zhou Shuigeng

May 20, 2007

Outline

- What is Outlier?
- Methods for Outliers Detection
- LOF: A Density Based Algorithm
- Variants of LOF Algorithm
- Summary

Outline

- **What is Outlier?**
- Methods for Outliers Detection
- LOF: A Density Based Algorithm
- Variants of LOF Algorithm
- Summary

What is Outlier Detection?

- Outlier Definition by Hawkins

  "An outlier is an observation that deviates so much from other observations as to arouse suspicions that it was generated by a different mechanism. "

  — D. Hawkins. *Identification of Outliers*. Chapman and Hall, London, 1980

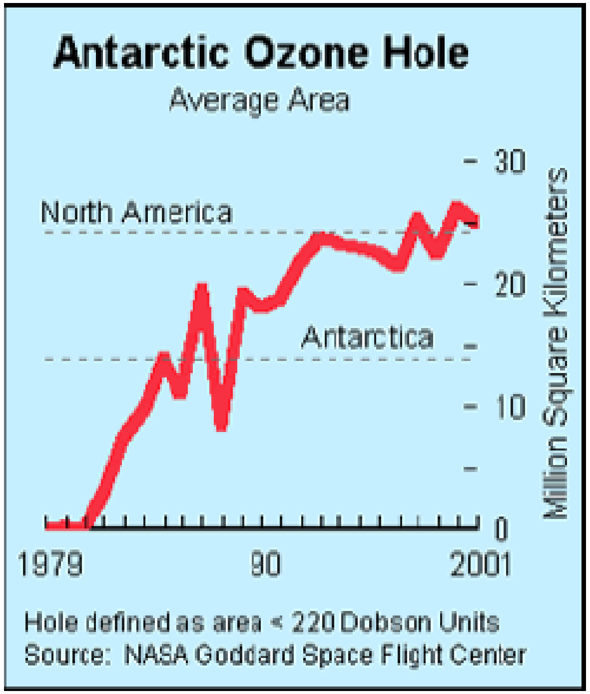
- Another term: anomaly detection
- Examples: The Beatles, Michael Schumacher

What is Outlier Detection?

- The goal of outlier detection is to uncover the "different mechanism"
- If samples of the "different mechanism" exist, then build a classifier to learn from samples
- Often called the "Imbalanced Classification Problem" -size of outlier sample is small vis-à-vis normal samples.
- In most practical real-life settings, samples of the outlier generating mechanism are non-existent

Importance of Outlier Detection

- Ozone Depletion History
  - In 1985 three researchers (Farman, Gardinar and Shanklin) were puzzled by data gathered by the British Antarctic Survey showing that ozone levels for Antarctica had dropped 10% below normal levels
  - Why did the Nimbus 7 satellite, which had instruments aboard for recording ozone levels, not record similarly low ozone concentrations?
  - The ozone concentrations recorded by the satellite were so low they were being treated as outliers by a computer program and discarded!



Sources:
http://exploringdata.cqu.edu.au/ozone.html
http://www.epa.gov/ozone/science/hole/size.html

Outlier Detection Applications

- Credit card fraud detection
- Telecom fraud detection
- Customer segmentation
- Medical analysis
- Counter-terrorism
- ……

Outlier Detection: Challenges & Assumption

- Challenges
  - How many outliers are there in the data?
  - Method is unsupervised
  - Validation can be quite challenging (just like for clustering)
  - Finding needle in a haystack
- Working assumption:
  - There are considerably more "normal" observations than "abnormal" observations (outliers/anomalies) in the data

Detection Schemes

- From normal to abnormal
  - Build a profile of the "normal" behavior
    - Profile can be patterns or summary statistics for the overall population
  - Use the "normal" profile to detect anomalies
    - Anomalies are observations whose characteristics differ significantly from the normal profile
- Build profile of abnormal behavior
  - Find these conforming to the profile

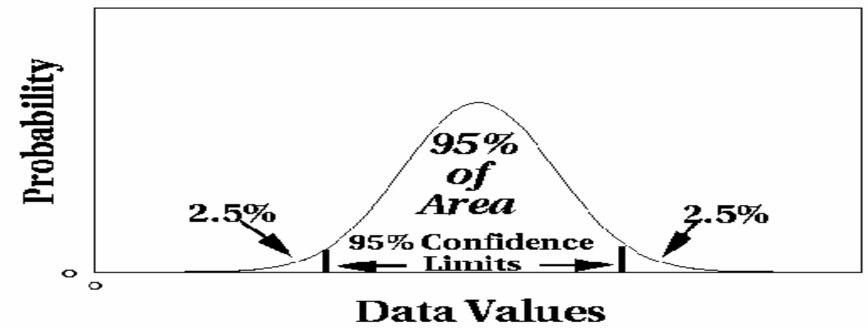
Outline

- What's Is Outlier
- Methods for Outliers Detection
- LOF: A Density Based Algorithm
- The Variants of LOF Algorithm
- Summary

Outliers Detection Methods

- Statistical method
- Distance based method
- Graphical method
- Depth-based outlier detection
- Density based method
- …

Statistical Approaches



- Assume a model underlying distribution that generates data set (e.g. normal distribution)
- Use discordancy tests depending on
  - data distribution
  - distribution parameter (e.g., mean, variance)
  - number of expected outliers
- Drawbacks
  - most tests are for single attribute
  - In many cases, data distribution may not be known

Grubbs' Test

- Detect outliers in univariate data
- Assume data comes from normal distribution
- Detects one outlier at a time, remove the outlier, and repeat
  - H0: There is no outlier in data
  - HA: There is at least one outlier
- Grubbs' test statistic: $$G = \frac{\max \left| X - \overline{X} \right|}{s}$$ X^-: mean
s: standard deviation
- Reject H0 if: $$G > \frac{(N-1)}{\sqrt{N}} \sqrt{\frac{t^2_{(\alpha/N, N-2)}}{N-2+t^2_{(\alpha/N, N-2)}}}$$

Likelihood Approach (1)

- Assume the data set D contains samples from a mixture of two probability distributions:
  - M (majority distribution)
  - A (anomalous distribution)
- General Approach:
  - Initially, assume all the data points belong to M
  - Let $L_t(D)$ be the log likelihood of D at time t
  - For each point x_t that belongs to M, move it to A
    - Let $L_{t+1}(D)$ be the new log likelihood.
    - Compute the difference, $\Delta = L_t(D) - L_{t+1}(D)$
    - If $\Delta > c$ (some threshold), then x_t is declared as an anomaly and moved permanently from M to A

Likelihood Approach (2)

- Data distribution, D = (1 – λ) M + λ A
- M is a probability distribution estimated from data
  - Can be based on any modeling method (naïve Bayes, maximum entropy, etc)
- A is initially assumed to be uniform distribution
- Likelihood at time t:

$$L_t(D) = \prod_{i=1}^{N} P_D(x_i) = \left((1-\lambda)^{|M_t|} \prod_{x_i \in M_t} P_{M_t}(x_i) \right) \left(\lambda^{|A_t|} \prod_{x_i \in A_t} P_{A_t}(x_i) \right)$$

$$LL_t(D) = |M_t| \log(1-\lambda) + \sum_{x_i \in M_t} \log P_{M_t}(x_i) + |A_t| \log \lambda + \sum_{x_i \in A_t} \log P_{A_t}(x_i)$$

Distance-Based Approach

- Introduced to counter the main limitations imposed by statistical methods
  - We need multi-dimensional analysis without knowing data distribution.
- Distance-based outlier: A DB(p, D)-outlier is an object O in a dataset T such that at least a fraction p of the objects in T lies at a distance greater than D from O
  - First introduced by Knorr and Ng

Other Distance-based Outlier Definitions

- **Kollios et al Definition**
  - An object o in a dataset T is a DB(k,d) outlier if at most k objects in T lie at distance at most d from o
- **Ramaswamy et al Definition**
  - Outliers are the top n data elements whose distance to the k-th nearest neighbor is greatest

The Nested Loop Algorithm

- Simple Nested Loop Algorithm: $O(N^2)$
  - For each object $o \in T$, compute distance to each $q \neq o \in T$, until $k + 1$ neighbors are found with distance less than or equal to d.
  - If $|Neighbors(o)| \leq k$, Report o as DB(k,d) outlier
- Nested Loop with Randomization and Pruning
  - The fastest known distance-based algorithm for DB(k, d) proposed by Bay and Schwabacher, its expected running time is O(N)

Outlier and Dimension

- In high-dimensional space, data is sparse and notion of proximity becomes meaningless
  - Every point is an almost equally good outlier from the perspective of proximity-based definitions
- Lower-dimensional projection methods
  - A point is an outlier if in some lower dimensional projection, it is present in a local region of abnormally low density

Subspace Outlier Detection

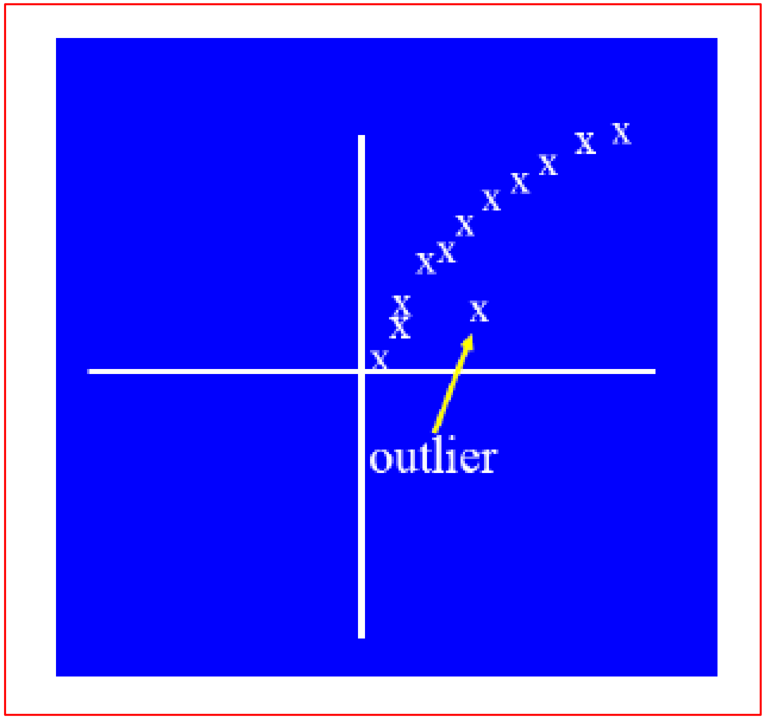
- Divide each attribute into ϕ equal-depth intervals
  - Each interval contains a fraction $f = 1/\phi$ of the records
- Consider a k-dimensional cube created by picking grid ranges from k different dimensions
  - If attributes are independent, we expect region to contain a fraction f^k of the records
  - If there are N points, we can measure sparsity of a cube D as:

$$S(\mathcal{D}) = \frac{n(\mathcal{D}) - N \cdot f^k}{\sqrt{N \cdot f^k \cdot (1 - f^k)}}$$

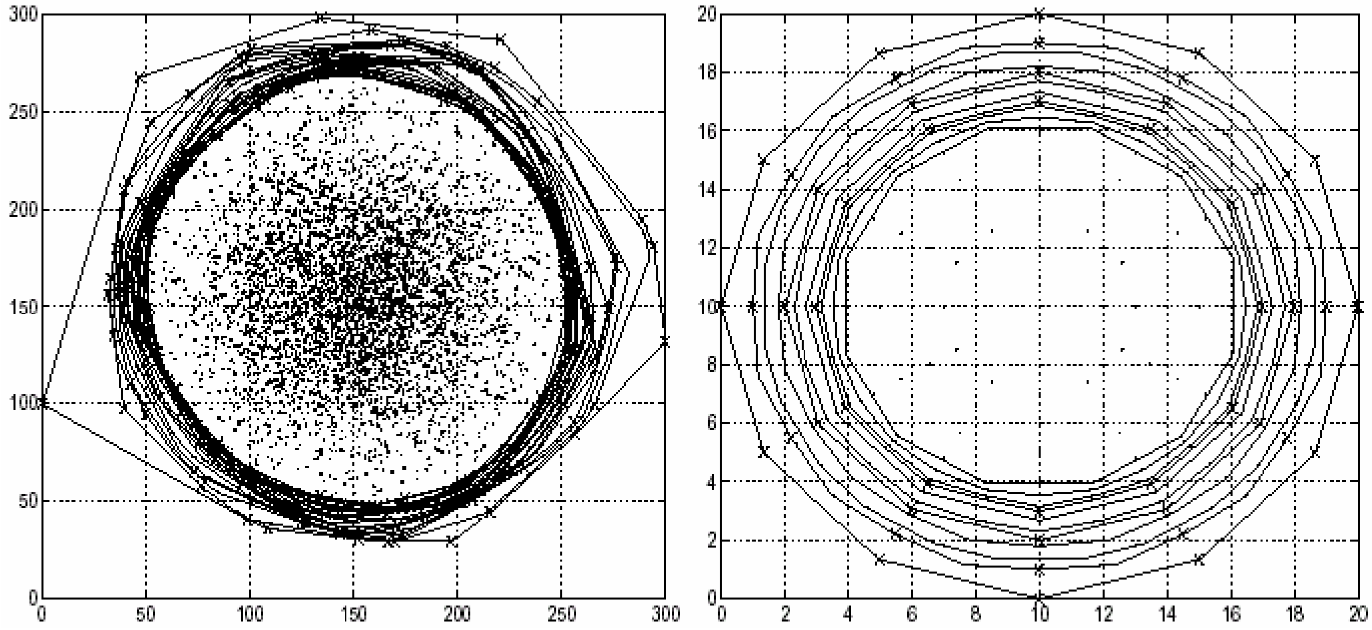
  - Negative sparsity indicates cube contains smaller number of points than expected

Graphical Approaches

- Boxplot (1-D), Scatter plot (2-D), Spin plot (3-D)
- Limitations
  - Time consuming
  - Subjective



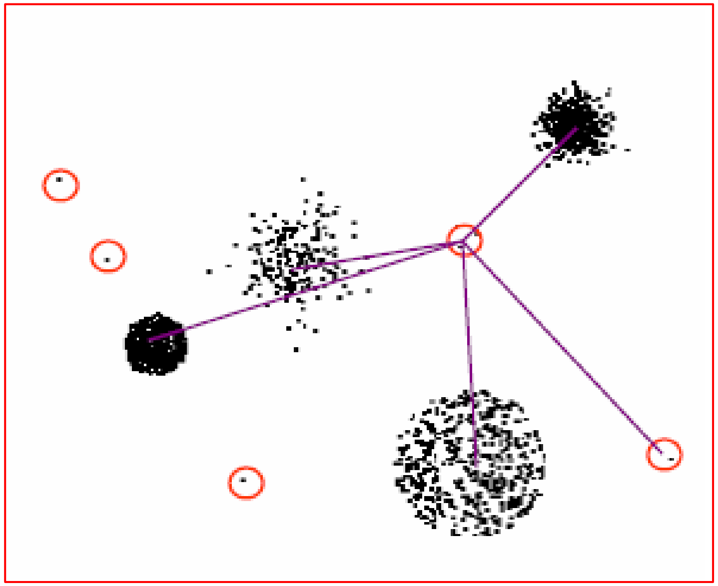
Depth based Method



- Based on some definition of *depth,* objects are organized in layers in the data space
- Shallow layer are more likely to contain outliers

Clustering-based Outlier Detection

- Basic idea:
  - Cluster the data into groups of different density
  - Choose points in small cluster as candidate outliers
  - Compute the distance between candidate points and non-candidate clusters
    - If candidate points are far from all other non-candidate points, they are outliers



Spatial Outliers

- Outlier detection techniques are often used in GIS, climate studies, public health, etc. The spread and detection of "bird flu" can be cast as an outlier detection problem
- The distinguishing characteristics of spatial data is the presence of "spatial attributes" and the neighborhood relationship
- Spatial Outlier Definition by Shekhar et al.
  - *A spatial outlier is a spatial referenced object whose non-spatial attribute values are significantly different from those of other spatially referenced objects in its spatial neighborhood*

Finding Spatial Outliers

- Let O be a set of spatially-referenced objects and $N \subset O \times O$ a neighborhood relationship. N(o) are all the neighbors of o.
- A Theorem (Shekhar et. al.)
  - Let f be function which takes values from a Normally distributed rv on O. Let $\boxed{g(x) = \frac{1}{|N(x)|} \sum_{y \in N(x)} f(y)}$
  - Then $h = f - g$ is Normally distributed.
- h effectively captures the deviation in the neighborhood, Because of spatial autocorrelation, a large value of |h| would be considered unusual.
- For multivariate f, use the Chi-Square test to determine outliers

Sequential Outliers

- In many applications, data is presented as a set of symbolic sequences
  - **Proteomics:** Sequence of amino acids
  - **Computer Virus:** Sequence of register calls
  - **Climate Data:** Sequence of discretized SOI readings
  - **Surveillance:** Airline travel patterns of individuals
  - **Health:** Sequence of diagnosis codes

Example Sequence Data

- Examples
  - SRHPAZBGKPBFLBCYVSGFHPXZIZIBLLKB
  - IXCXNCXKEGHSARQFRA
  - MGVRNSVLSGKKADELEKIRLRPGGKKKYML
- How to determine sequence outlier?
- Properties of Sequence Data
  - Not of same length; Sequence of symbols.
  - Unknown distribution; No standard notion of similarity

Outline

- What's Is Outlier
- Methods for Outliers Detection
- LOF: A Density Based Algorithm
- The Variants of LOF Algorithm
- Summary

LOF: Identifying Density-based Local Outliers

(Appeared in SIGMOD'00)

M. M. Breunig, H. Kriegel, R. T. Ng, and J. Sander

Outline

- Problems of Existing Approaches
- Local Outlier Factor (LOF)
- Properties of Local Outliers
- How to Choose *MinPts*
- Experiments
- Summary

Non-local Approaches (1)

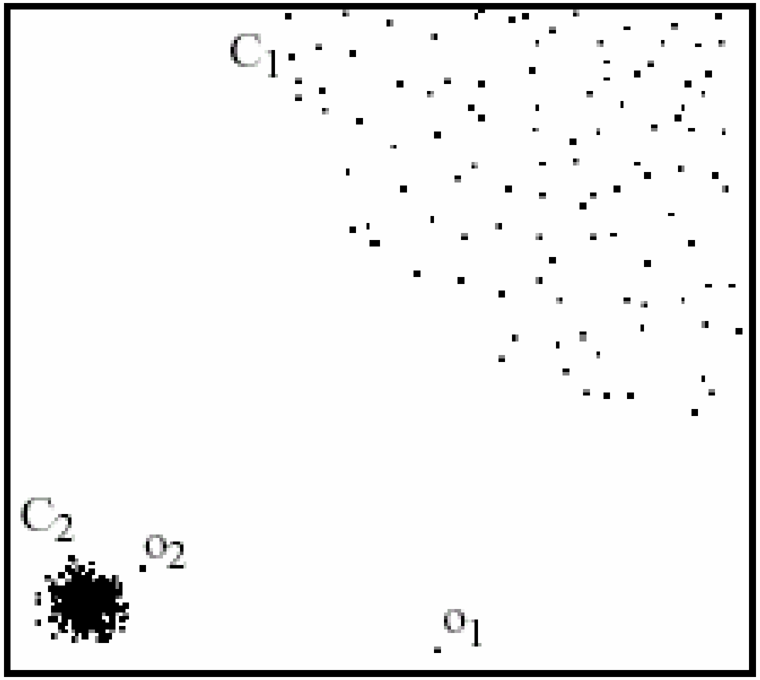
- Some Notations
  - d(p, q): distance between objects p & q
  - D(p, *C*) = min{ d(p, q) | q $\in$ *C* }, where *C* is an object set
- DB(*pct, dmin*)-Outlier

  An object p in a dataset D is a DB(*pct, dmin*)-outlier if at least percentage of *pct* of the objects in D lies greater than distance *dmin* from p, i.e., the cardinality of the set {q $\in$ D| d(p, q) <= dmin} is less than or equal to (100-*pct*)% of the size of D.

Non-local Approaches (2)

- DB(*pct, dmin*)-Outlier takes a "global" view
- Many real world data exhibit complex structure, and other kinds of outlier are needed
- Say, objects outlying relative to their local neighborhoods, particularly to densities → "*local*" outliers

Non-local Approaches (3)



- C_1: 400 objects
- C_2: 100 objects
- o_1, o_2: outliers

If every object q in C_1, the distance between q and its nearest neighbor is greater than $d(o_2, C_2)$, there's no appropriate value of *pct* and *dmin* such that o_2 is identified as outlier, but the objects in C_1 are not.

Non-local Approaches (4)

- Why?
  - If dmin < $d(o_2, C_2)$, o_2 and all the objects in C1 are DB-outliers
  - Otherwise, that o_2 is a not DB-outlier

Outline

- Problems of Existing Approaches
- Local Outlier Factor (LOF)
- Properties of Local Outliers
- How to Choose *MinPts*
- Experiments
- Summary

Definitions (1)

- *k*-distance of an object p: k-distance(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 $d(p,o') \leq d(p,o)$, and
(ii) for at most k-1 objects $o' \in D \setminus \{p\}$ it holds that $d(p,o') < d(p,o)$.

- *k*-distance neighborhood of *p*: $N_{k\text{-ditsance}(p)}(p)$

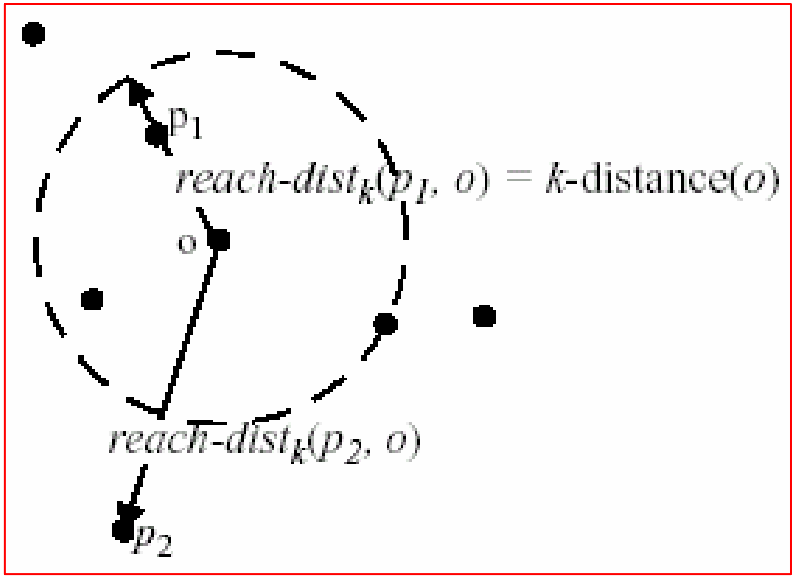
Given the k-distance of p, the *k-distance neighborhood of p* contains every object whose distance from p is not greater than the k-distance, i.e. $N_{k\text{-distance}(p)}(p) = \{ q \in D \setminus \{p\} \mid d(p, q) \leq k\text{-distance}(p) \}$.
These objects q are called the k-nearest neighbors of p.

Definitions (2)

- Reachability distance of *p* w.r.t. *o*: reach-dist$_k$(p, o)

Let *k* be a natural number. The *reachability distance* of object *p* with respect to object *o* is defined as

$$\textit{reach-dist}_k(p, o) = \max \{ k\text{-distance}(o), d(p, o) \}.$$



- Reachability distance aims to reduce the statistical fluctuation of d(p,o) for all the p's close to o.
- The strength of such smoothing effect can be controlled by the value of *k*

Definitions (3)

- *Local reachability density*

$$lrd_{MinPts}(p) = 1 \Bigg/ \left(\frac{\sum_{o \in N_{MinPts}(p)} reach\text{-}dist_{MinPts}(p, o)}{|N_{MinPts}(p)|} \right)$$

  - The inverse of the average reachability distance based on the *MinPts*-NN of p.
  - lrd relies on only one parameter, *MinPts*, different from the traditional clustering cases where MinPts and a parameter of volume are needed

LOF: Local Outlier Factor

- *local outlier factor*

$$LOF_{MinPts}(p) = \frac{\sum_{o \in N_{MinPts}(p)} \frac{lrd_{MinPts}(o)}{lrd_{MinPts}(p)}}{|N_{MinPts}(p)|}$$

  - The average of the ratio of the local reachability density of p and those of of *p*'s *MinPts*-NNs.

Outline

- Problems of Existing Approaches
- Local Outlier Factor (LOF)
- Properties of Local Outliers
- How to Choose *MinPts*
- Experiments
- Summary

LOF for Outliers Deep in a Cluster

Lemma 1: Let C be a collection of objects. Let *reach-dist-min* denote the minimum reachability distance of objects in C, i.e., *reach-dist-min* = *min* $\{$*reach-dist*$(p, q) \mid p, q \in C\}$. Similarly, let *reach-dist-max* denote the maximum reachability distance of objects in C. Let ε be defined as (*reach-dist-max*/*reach-dist-min* – 1).
Then for all objects $p \in C$, such that:

(i) all the *MinPts*-nearest neighbors q of p are in C, and

(ii) all the *MinPts*-nearest neighbors o of q are also in C,

it holds that $1/(1 + \varepsilon) \leq LOF(p) \leq (1 + \varepsilon)$.

For "tight" cluster, ε will be quite small, thus forcing the lof of p to be quite close to 1.

Upper and Lower Bound

- What's the behavior of lof on the objects near the periphery of a cluster, or those out of a cluster?
- Theorem 1 gives a more general and better bound on LOF
- Some notations first:)

Notations (1)

- *direct*$_{min}$(*p*): the minimum reachability distance between *p* and a *MinPts*-NN of *p*.

$$direct_{min}(p) = \min \{ \textit{reach-dist}(p, q) \mid q \in N_{MinPts}(p) \}.$$

- Similarly,

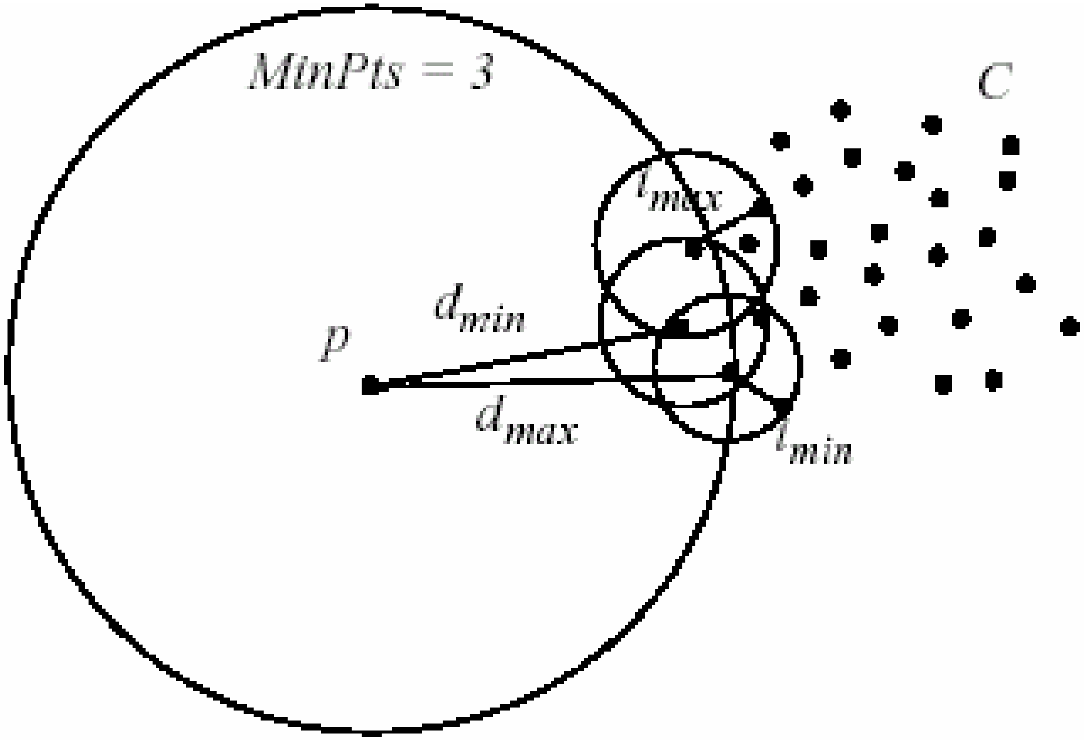
$$direct_{max}(\mathrm{p}) = \max \{ \textit{reach-dist}(p, q) \mid q \in N_{MinPts}(p) \}.$$

- *indirect*$_{min}$(*p*): the minimum reachability distance between *q* and a *MinPts*-NN of *q*

$$indirect_{min}(p) = \min \{ \textit{reach-dist}(q, o) \mid q \in N_{MinPts}(p) \text{ and } o \in N_{MinPts}(q) \}.$$

- Similarly we can define *indirect*$_{max}$(*p*)

Notations (2)



$d_{min} = 4 * i_{max}$

$\Rightarrow LOF_{MinPts}(p) \geq 4$

$d_{max} = 6 * i_{min}$

$\Rightarrow LOF_{MinPts}(p) \leq 6$

Theorem 1

Theorem 1: Let p be an object from the database D, and $1 \le MinPts \le |D|$.

Then, it is the case that

$$\frac{direct_{min}(p)}{indirect_{max}(p)} \le LOF(p) \le \frac{direct_{max}(p)}{indirect_{min}(p)}$$

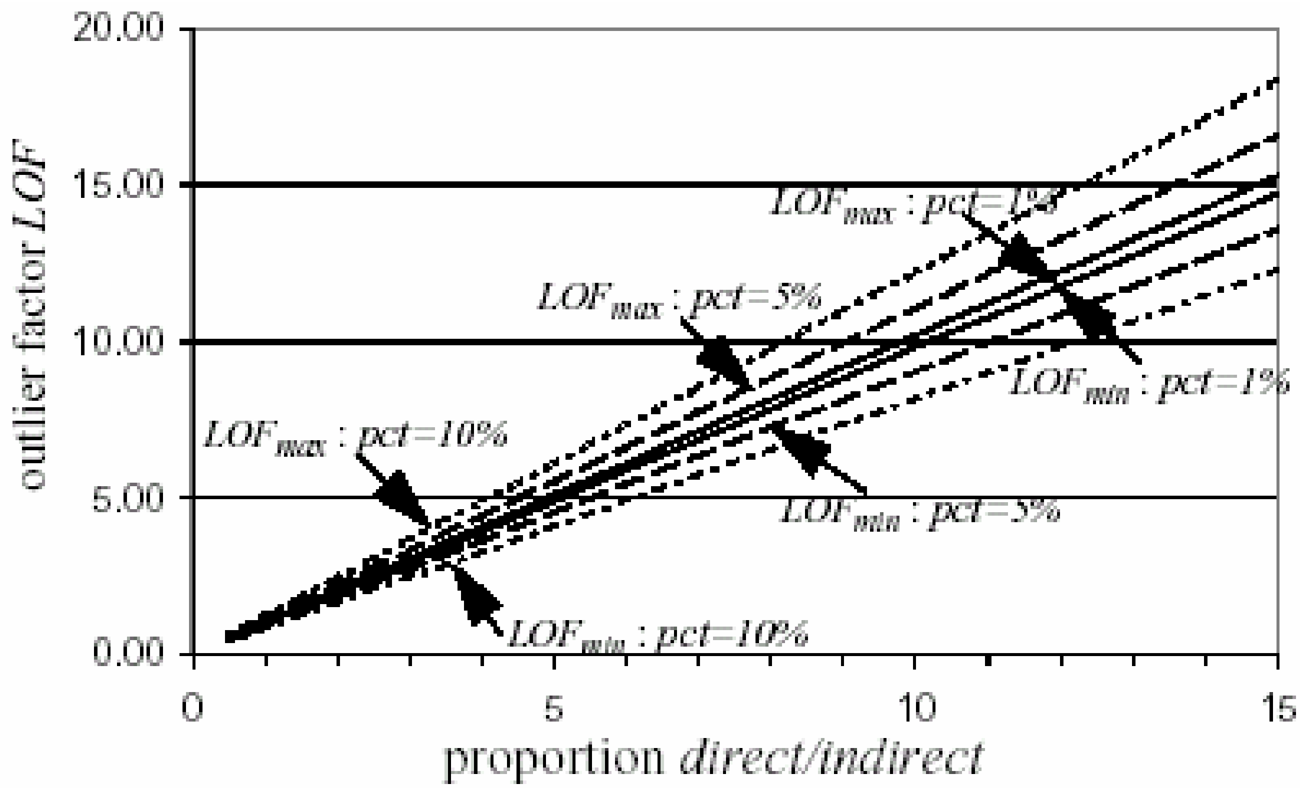
We can see from Therom 1, that LOF is just a function of the reachability distances in p's direct neighborhood relative to those in p's indirect neighborhood.

Tightness of the Bounds

- *direct*(p) : the mean of $direct_{min}(p)$ & $direct_{max}(p)$
- *indirect*(p): the mean of $indirect_{min}(p)$ & $indirect_{max}(p)$
- Assumption:

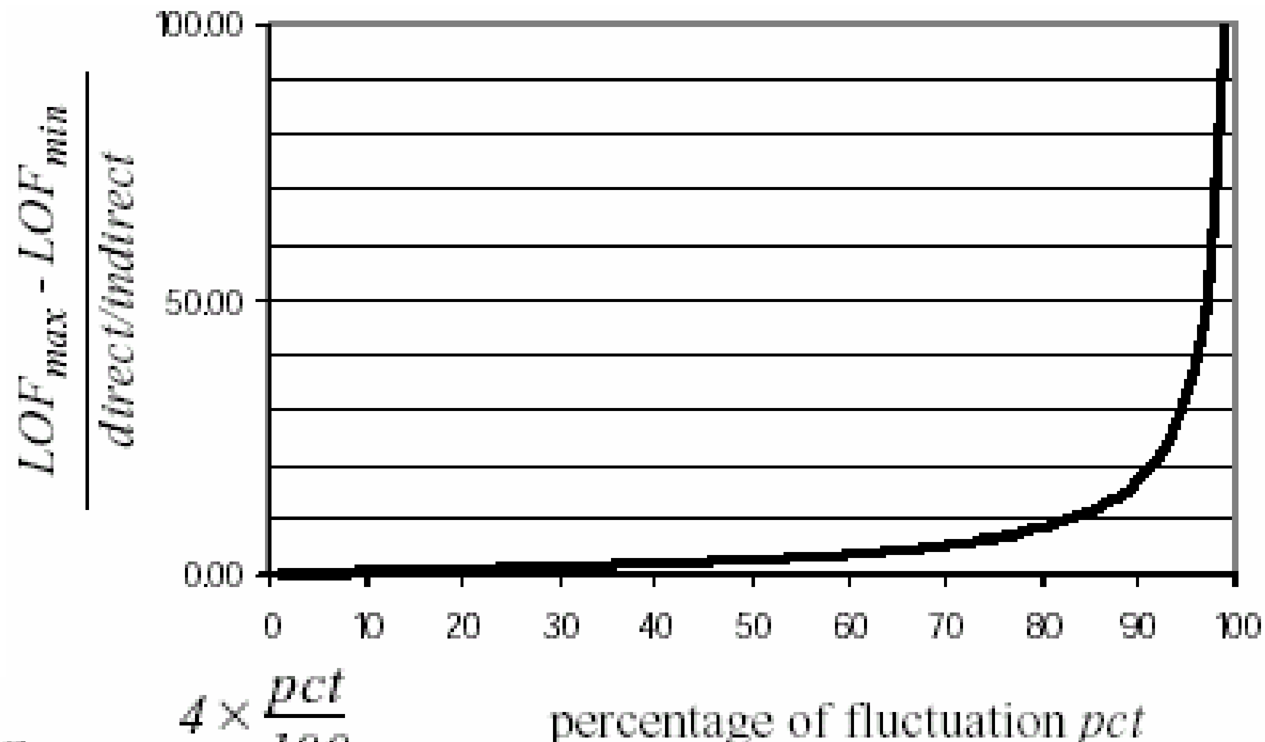
$$(direct_{max} - direct_{min})/direct = (indirect_{max} - indirect_{min})/indirect.$$

- Let $pct = x\% = (dirct_{max} - direct_{min})/2 = (indirct_{max} - indirect_{min})/2$



- The relative span (LOFmax - LOFmin) / (direct/indirect) is constant
- In other words, the relative fluctuation of LOF relies only on the ratios of distances, not on *absolute* values

$(LOF_{max} - LOF_{min})$, $(direct/indirect)$, pct

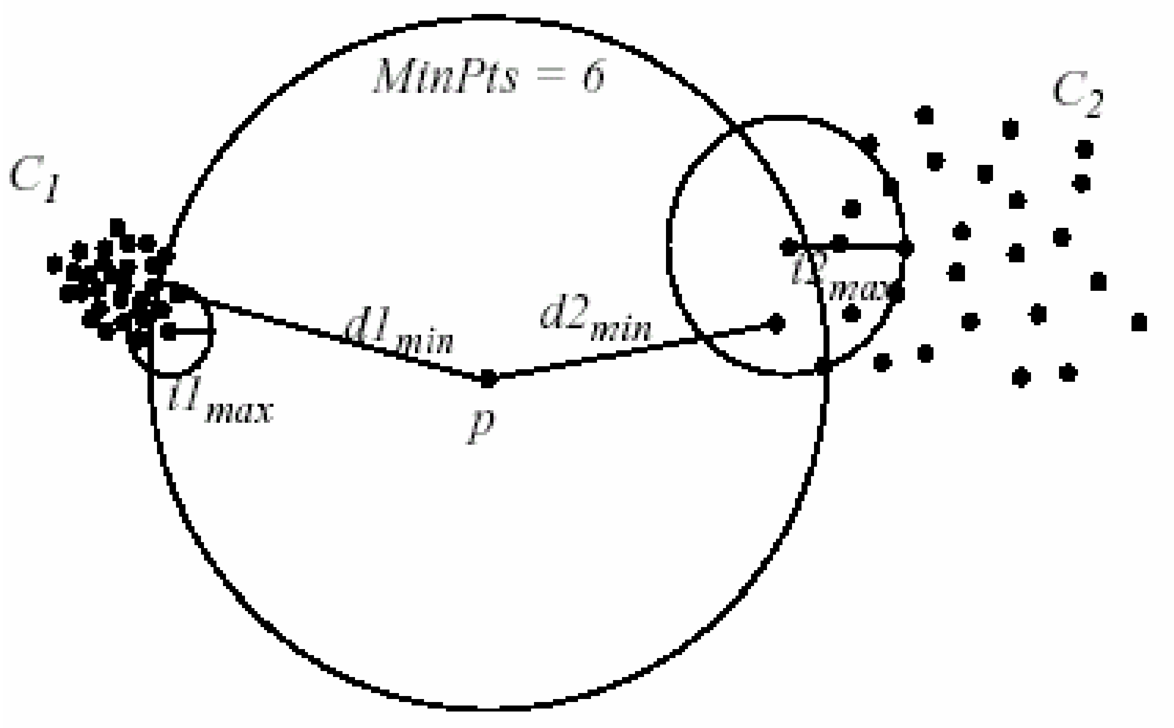


$$\frac{LOF_{max} - LOF_{min}}{\frac{direct}{indirect}} = \frac{4 \times \frac{pct}{100}}{1 - \left(\frac{pct}{100}\right)^2}$$

Some Conclusions on Bounds

- If pct is small, therom 1 estimates the LOF very well, as the bounds are "tight".
  - For the instance deep in a cluster…
  - For those not deep inside a cluster, but whose MinPts-NN belong to the same cluster…

Theorem 2 (1)



- P's MinPts-NNs do not belong to the same cluster, which makes pct very large…

Theorem 2 (2)

Theorem 2: Let p be an object from the database D, $1 \le MinPts \le |D|$, and $C_1, C_2, \ldots, C_n$ be a partition of $N_{MinPts}(p)$, i.e. $N_{MinPts}(p) = C_1 \cup C_2 \cup \ldots \cup C_n \cup \{p\}$ with $C_i \cap C_j = \emptyset$, $C_i \neq \emptyset$ for $1 \le i,j \le n, i \neq j$.

$$LOF(p) \ge \left(\sum_{i=1}^{n} \xi_i \cdot direct^i_{min}(p)\right) \cdot \left(\sum_{i=1}^{n} \frac{\xi_i}{indirect^i_{max}(p)}\right)$$

$$LOF(p) \le \left(\sum_{i=1}^{n} \xi_i \cdot direct^i_{max}(p)\right) \cdot \left(\sum_{i=1}^{n} \frac{\xi_i}{indirect^i_{min}(p)}\right)$$

Where $\xi_i = |C_i| / |N_{MinPts}(p)|$

For p in the figure of the last slide, we get LOF_{min} =

$$(0.5*d1_{min} + 0.5*d2_{min})/(0.5/i1_{max} + 0.5/i2_{max})$$

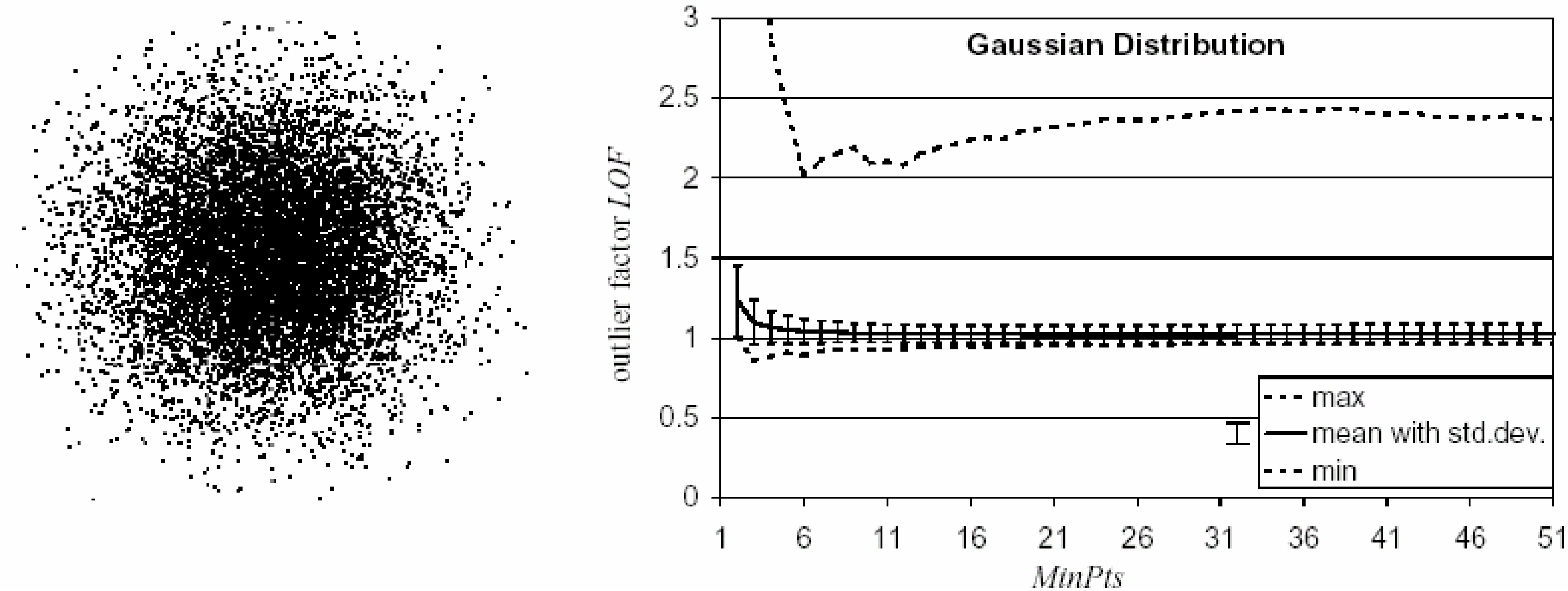
Outline

- Problems of Existing Approaches
- Local Outlier Factor (LOF)
- Properties of Local Outliers
- How to Choose *MinPts*
- Experiments
- Summary

LOF, MinPts

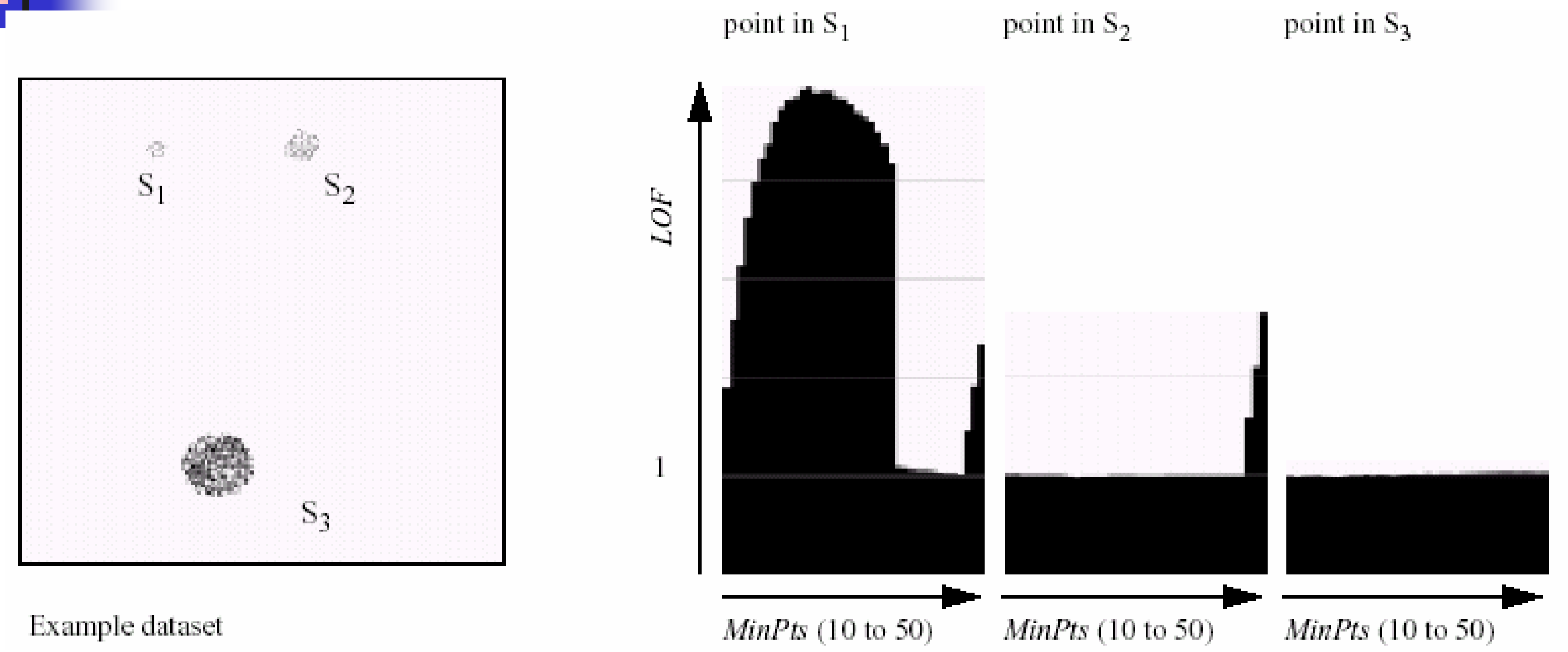
- How *LOF* changes according to changing *MinPts* values?
  - Increase monotonic w.r.t MinPts,
  - Or decrease monotonic w.r.t. MinPts,
  - Or …

An Example



- The dataset (500 objects) follows Gaussion distribution

Another Example



- S1: 10, S2: 35, S3: 500

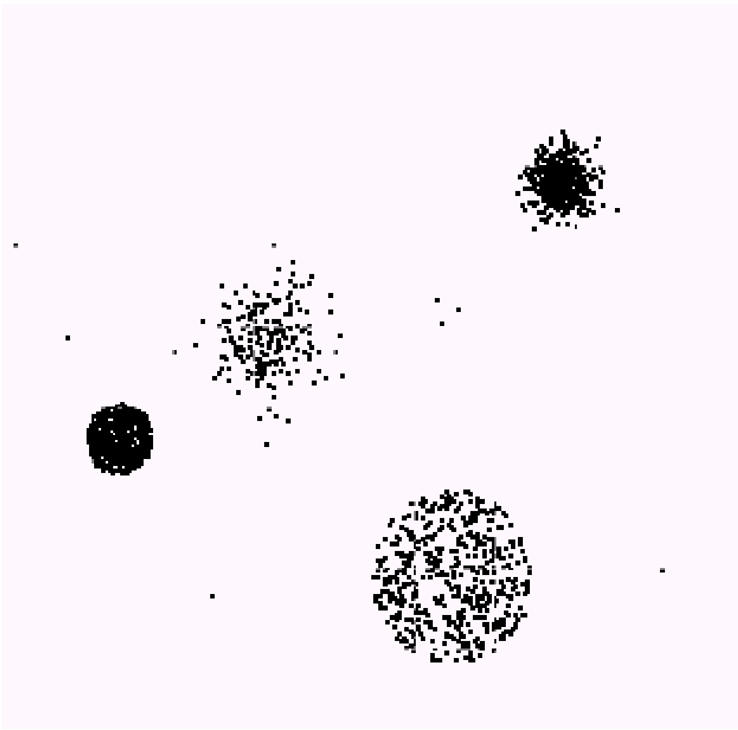
Determine a Range of MinPts

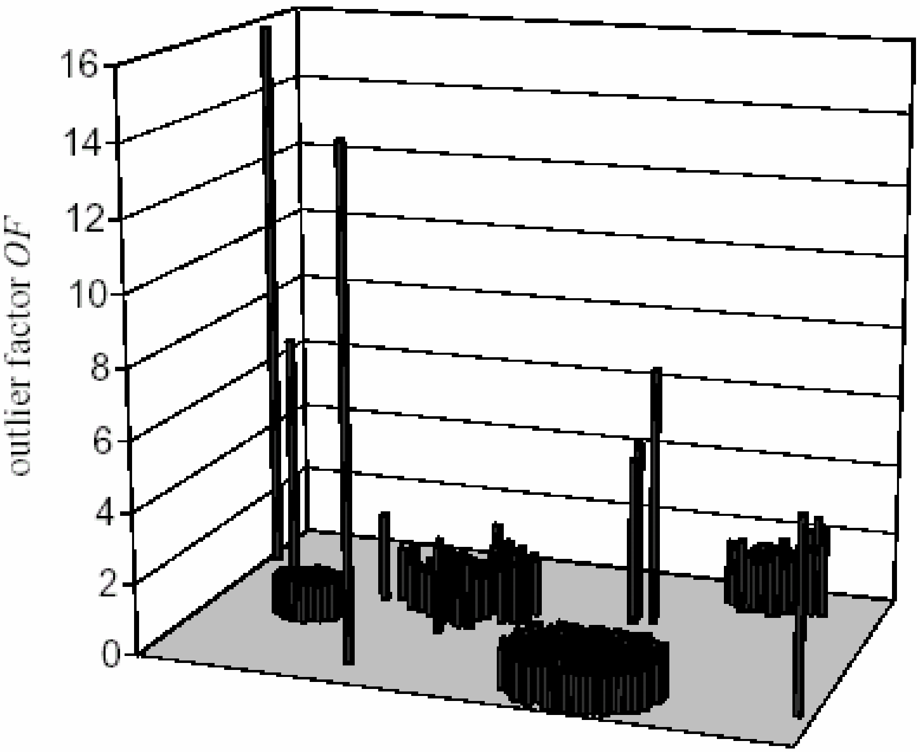
- *MinPtsLB*
  - According to the figure in slide 54, *MinPtsLB* >= 10
  - *MinPtsLB* can be regarded as the minimum number of objects a "cluster" has to contain
  - 10 ~ 20 √
- *MinPtsUB*
  - The maximum number of "close by" objects that can potentially be local outliers
- Compute *LOF* for each value of *MinPts* between *MinPtsLB* & *MinPtsUB*, and rank w.r.t. the maxmum *LOF* within the range

Outline

- Problems of Existing Approaches
- Local Outlier Factor (LOF)
- Properties of Local Outliers
- How to Choose *MinPts*
- Experiments
- Summary

A Synthetic Example





MinPts = 40

Soccer Data

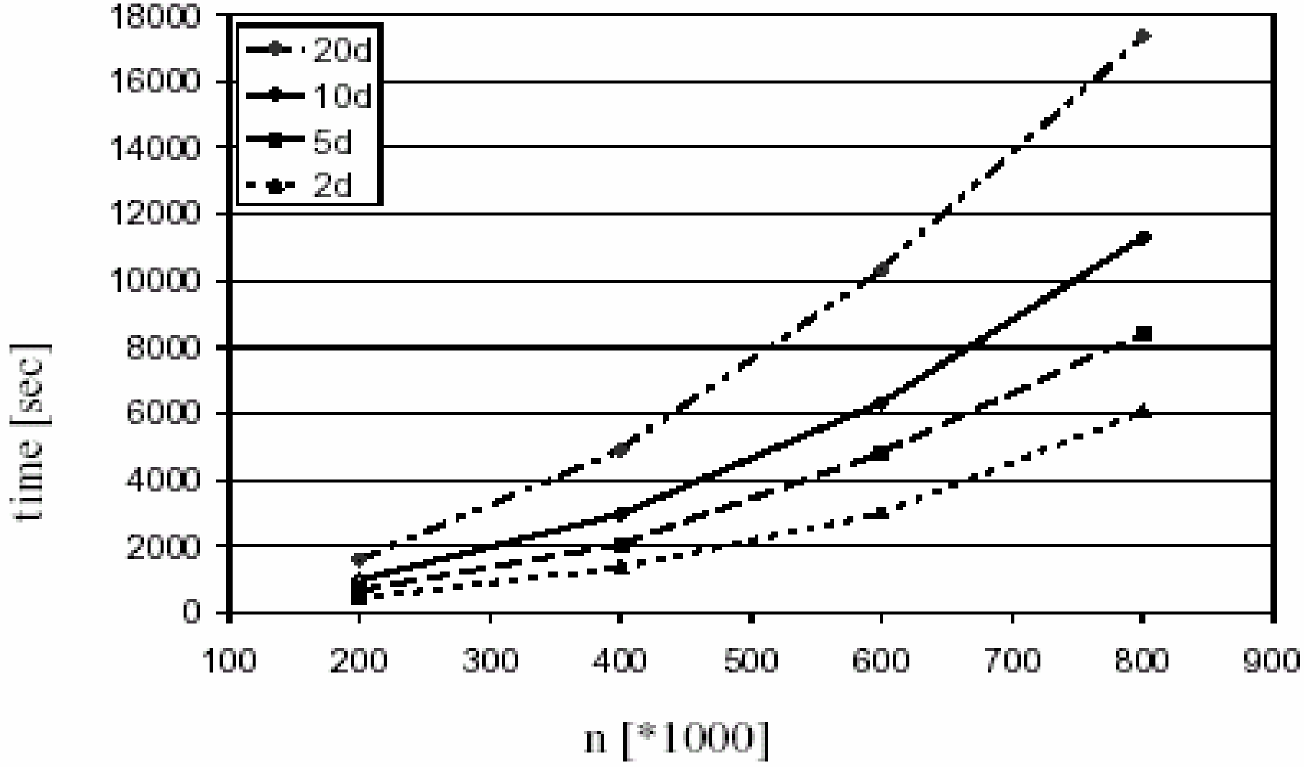
Rank	Outlier Factor	Player Name	Games Played	Goals Scored	Position
1	1.87	Michael Preetz	34	23	Offense
2	1.70	Michael Schjönberg	15	6	Defense
3	1.67	Hans-Jörg Butt	34	7	Goalie
4	1.63	Ulf Kirsten	31	19	Offense
5	1.55	Giovane Elber	21	13	Offense
minimum			**0**	**0**	
median			**21**	**1**	
maximum			**34**	**23**	
mean			**18.0**	**1.9**	
standard deviation			**11.0**	**3.0**	

1. number of games
2. average goals per game
3. Position (coded in integer)

Performance (1)

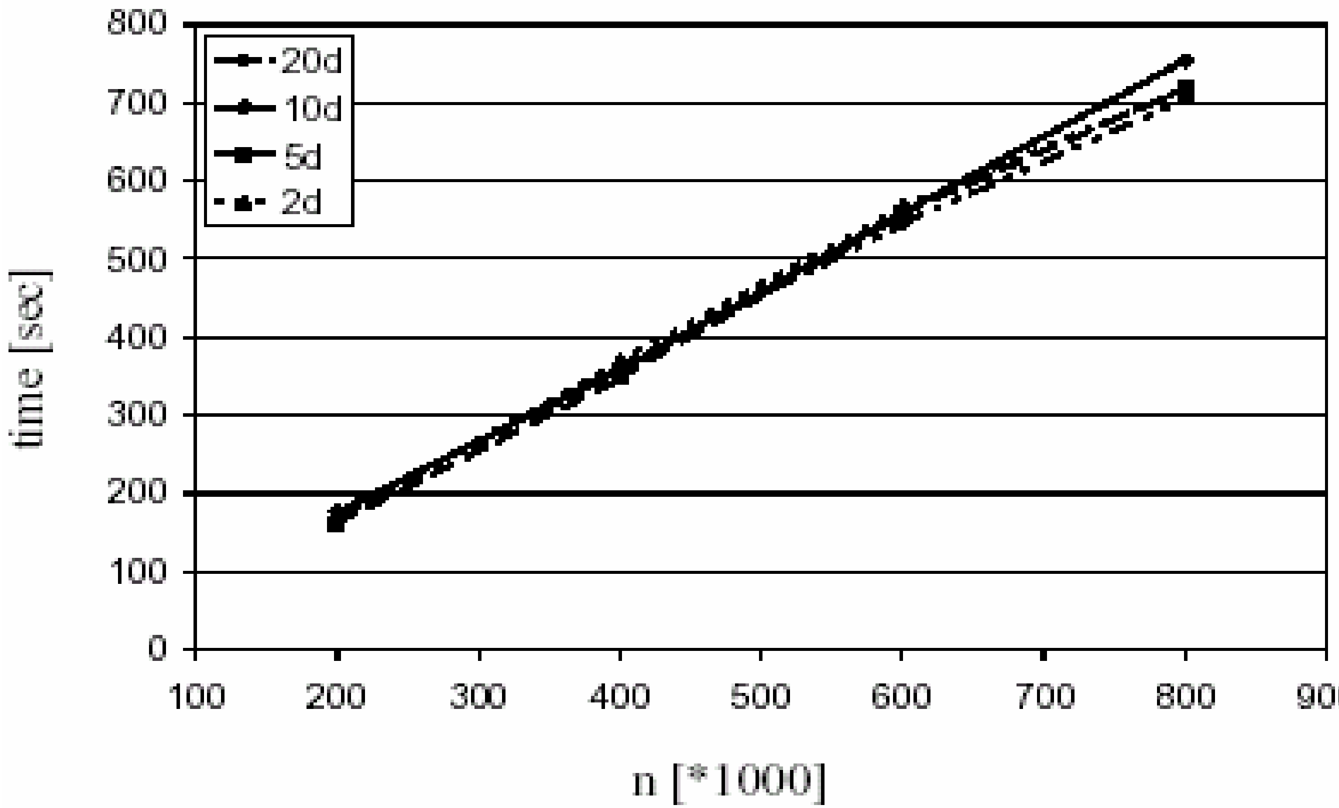
- Two steps
  - Compute and materialize MinPtsUB-NNs for each objects in the dataset to get a database of MinPtsUB-NNs (O(n*time for k-nn query)
  - Based on the database, for each value between MinPtsLB & MinPtsUB, compute lrd of each object, then compute LOF of each object

Performance (2)



Runtime of the materialization of the 50-nn queries for different dataset sizes and different dimensions using an index

Performance (3)



Runtime for the computation of the LOFs for different dataset sizes

Outline

- Problems of Existing Approaches
- Local Outlier Factor (LOF)
- Properties of Local Outliers
- How to Choose *MinPts*
- Experiments
- Summary

Summary

- LOF, captures relative degree of isolation
- LOF enjoy many good properties
  - Close to 1 for the objects deep in cluster
  - Meaningful bounds are given
- The effect of MinPts
- Fantastic experiments

Outline

- What's Is Outlier
- Methods for Outliers Detection
- LOF: A Density Based Algorithm
- Variants of LOF Algorithm
- Summary

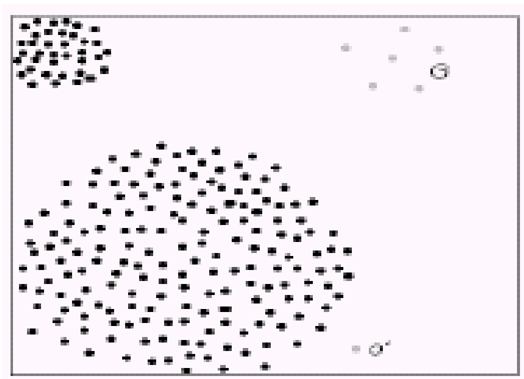
Variants of LOF (1)

- Some work done by Ada Fu
  - COF (PAKDD'02)
  - LOF', LOF'' and GridLOF (IDEAS'03)

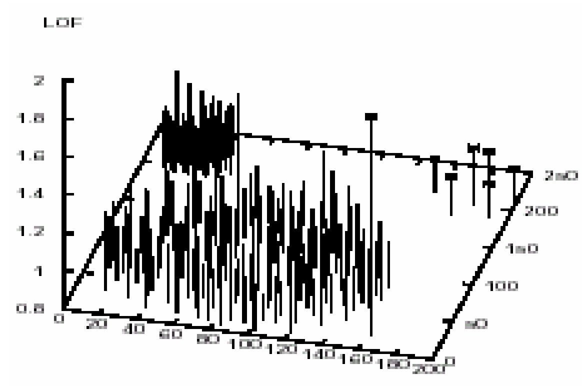
$$LOF'_{MinPts}(p) = \frac{\sum_{o \in N_{MinPts\text{-}dist(p)}(p)} \frac{MinPts\text{-}dist(p)}{MinPts\text{-}dist(o)}}{|N_{MinPts-dist(p)}(p)|}$$

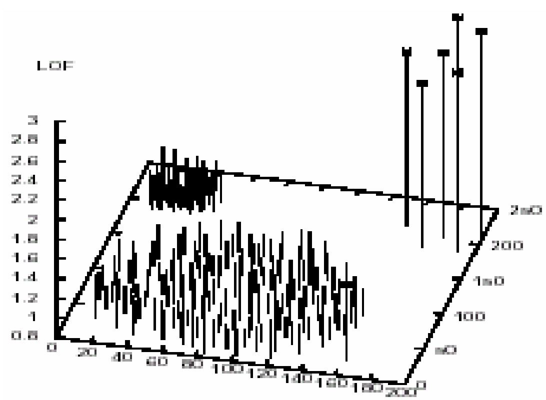
$$LOF^{N}_{MinPts_1, MinPts_2} = \frac{\sum_{o \in N_{MinPts_1\text{-}dist(p)}} \frac{lrd_{MinPts_2}(o)}{lrd_{MinPts_2}(p)}}{|N_{MinPts_1\text{-}dist(p)}(p)|}$$

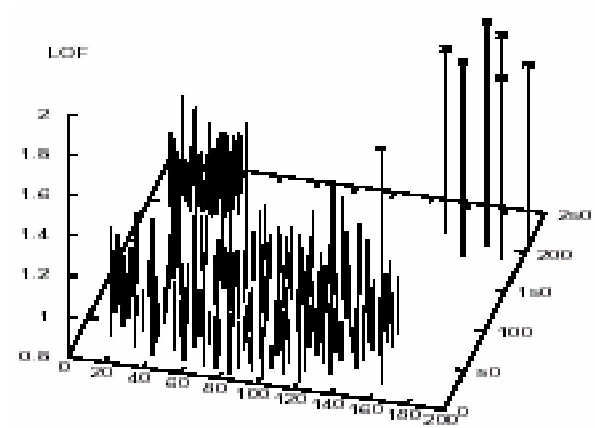
Variants of LOF (2)



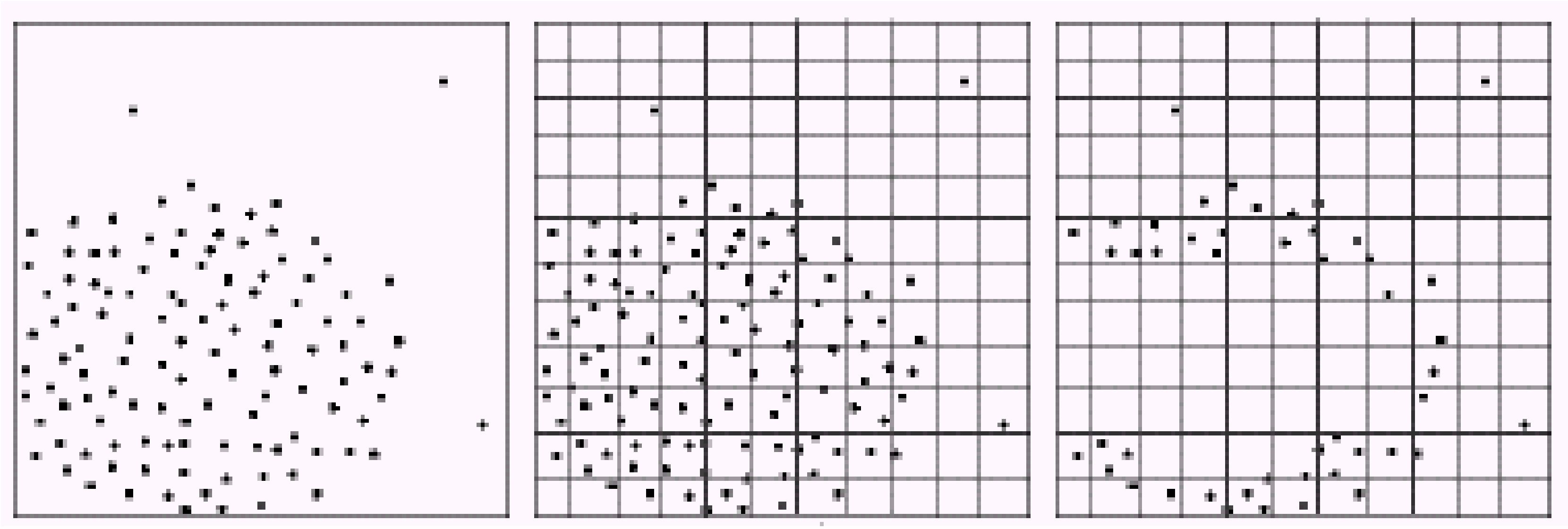
(a)







Variants of LOF (3)



- GridLOF aims to save computation cost

Outline

- What's Is Outlier
- Methods for Outliers Detection
- LOF: A Density Based Algorithm
- Variants of LOF Algorithm
- Summary

References (1)

- C. C. Aggarwal and P. S. Yu. Outlier detection for high dimensional data. In *Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data, Santa Barbara, California, USA*, 2001
- S. D. Bay and M. Schwabacher. Mining distance-based outliers in near linear time with randomization and a simple pruning rule. In *Proceedings of Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Washington, D.C. USA*, pages 29–38, 2003
- M. M. Breunig, H.-P. Kriegel, R. T. Ng, and J. Sander. Lof: Identifying density-based local outliers. In *Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data, Dallas, Texas, USA*, pages 93–104. ACM, 2000
- D. Hawkins. *Identification of Outliers*. Chapman and Hall, London, 1980
- E. M. Knorr and R. T. Ng. Finding intensional knowledge of distance-based outliers. In *Proceedings of 25th International Conference on Very Large Data Bases*, pages 211–222. Morgan Kaufmann, 1999

References (2)

- G. Kollios, D. Gunopulos, N. Koudas, and S. Berchtold. Efficient biased sampling for approximate clustering and outlier detection in large datasets. *IEEE Transactions on Knowledge and Data Engineering*, 2003.
- A. Lazarevic, L. Ert ¨oz, V. Kumar, A. Ozgur, and J. Srivastava. A comparative study of anomaly detection schemes in network intrusion detection. In *SDM*, 2003.
- S. Ramaswamy, R. Rastogi, and K. Shim. Efficient algorithms for mining outliers from large data sets. In *SIGMOD Conference*, pages 427–438, 2000.
- P. J. Rousseeuw and K. V. Driessen. A fast algorithm for the minimum covariance deteriminant estimator. *Technometrics*, 41:212–223, 1999.
- J. Shawne-Taylor and N. Cristianini. *Kernel Methods for Pattern Analysis*. Cambridge, 2005.

References (3)

- S. Shekhar, C.-T. Lu, and P. Zhang. Detecting graph based spatial outliers: algorithms and applications (a summary of results). In *Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining, San Francisco, CA, USA. ACM, 2001*, pages 371–376, 2001
- V. Barnett, and T. Lewis. *Outliers in statistical data.* John Wiley, 1994.
- T. Johnson, I. Kwok, and R. Ng. *Fast computation of 2-dimentional depth contours*, Proc. 4th Int. Conf. Knowledge Discovery and Data Mining, New York, NY, AAAI Press, 1998, pp. 224-228.
- J. Tang, Z. Chen, A. Fu, and D. Cheung. *Enhancing effectiveness of outlier detection for low density patterns*, PAKDD 2002, 2002.
- A. Chiu, and A Fu. *Enhancements on Local Outlier Detection*, IDEAS 2003, 2003.
- P. Sun, S. Chawla, and B. Arunasalam. Outlier detection in sequential databases. In *Proceedings of SIAM International Conference on Data Mining*, 2006