Video summarization via kcenter clustering

Filippo Geraci

May 7, 2019

The problem

- You have a long text ?
- You want to know the content without read it ?

The problem

- You have a long text ?
- You want to know the content without read it ?

At the December 2003 World Summit on the Information Society (WSIS) in Geneva, governments adopted a Plan of Action that, inter alia, called on the Secretary-General of the United Nations to set up a Working Group on Internet Governance (WGIG). The WGIG's mandate was to analyze the governance of Internet and make proposals for action, as appropriate; develop a working definition of Internet governance; identify the public policy issues involved; and advance a common understanding of the respective roles and responsibilities of the various stakeholders. In November 2004, Secretary-General Kofi Annan appointed forty individuals from government, the private sector, and civil society to the WGIG. After seven months of work, the WGIG's Report and a longer Background Report were released in July 2005. The two reports surveyed the landscape of public and private sector Internet governance mechanisms; advanced recommendations for improved governance in a range of key issue-areas; offered alternative models for the future oversight of the Internet's core resources and logical infrastructure; and proposed the establishment of a global, multistakeholder forum to facilitate continuing, inclusive dialogue on Internet governance. The reports elicited much debate around the world and served as key inputs in the second phase of the WSIS process, which culminated with the November 2005 Summit in Tunis. The issues they addressed remain of central importance in the post-Summit global debate on Internet governance.

With the continuing debate in mind, this volume presents papers by former members of the WGIG and its Secretariat staff that were written subsequent to the release of the group's reports. The contributors offer their personal reflections on the value of the multistakeholder cooperation in the WGIG and beyond; and on some of the key substantive issues and institutional reform proposals currently under consideration by the international community.

William J. Drake is the President of Computer Professionals for Social Responsibility, a global civil society organization, and an ICT policy consultant based in Geneva, Switzerland.

Secretariat of the United Nations ICT Task Force
One UN Plaza, New York, NY 10017
Web site: *http://www.unicttaskforce.org*

Secretariat of the Working Group on Internet Governance
Web site: *http://www.wgig.org*

ISBN 92-1-104557-6
Sales No. E.05.II.A.17

53738—October 2005—1,300

The problem

- You have a long text ?
- You want to know the content without read it ?

- Why do not do the same for video?


At the December 2003 World Summit on the Information Society (WSIS) in Geneva, governments adopted a Plan of Action that, inter alia, called on the Secretary-General of the United Nations to set up a Working Group on Internet Governance (WGIG). The WGIG's mandate was to analyze the governance of Internet and make proposals for action, as appropriate; develop a working definition of Internet governance; identify the public policy issues involved; and advance a common understanding of the respective roles and responsibilities of the various stakeholders. In November 2004, Secretary-General Kofi Annan appointed forty individuals from government, the private sector, and civil society to the WGIG. After seven months of work, the WGIG's Report and a longer Background Report were released in July 2005. The two reports surveyed the landscape of public and private sector Internet governance mechanisms; advanced recommendations for improved governance in a range of key issue-areas; offered alternative models for the future oversight of the Internet's core resources and logical infrastructure; and proposed the establishment of a global, multistakeholder forum to facilitate continuing, inclusive dialogue on Internet governance. The reports elicited much debate around the world and served as key inputs in the second phase of the WSIS process, which culminated with the November 2005 Summit in Tunis. The issues they addressed remain of central importance in the post-Summit global debate on Internet governance.

With the continuing debate in mind, this volume presents papers by former members of the WGIG and its Secretariat staff that were written subsequent to the release of the group's reports. The contributors offer their personal reflections on the value of the multistakeholder cooperation in the WGIG and beyond; and on some of the key substantive issues and institutional reform proposals currently under consideration by the international community.

William J. Drake is the President of Computer Professionals for Social Responsibility, a global civil society organization, and an ICT policy consultant based in Geneva, Switzerland.

Secretariat of the United Nations ICT Task Force
One UN Plaza, New York, NY 10017
Web site: *http://www.unicttaskforce.org*

Secretariat of the Working Group on Internet Governance
Web site: *http://www.wgig.org*

9 789211 045574

ISBN 92-1-104557-6
Sales No. E.05.II.A.17

53738—October 2005—1,500

The problem

- You have a long text ?
- You want to know the content without read it ?

- Why do not do the same for video?


At the December 2003 World Summit on the Information Society (WSIS) in Geneva, governments adopted a Plan of Action that, inter alia, called on the Secretary-General of the United Nations to set up a Working Group on Internet Governance (WGIG). The WGIG's mandate was to analyze the governance of Internet and make proposals for action, as appropriate; develop a working definition of Internet governance; identify the public policy issues involved; and advance a common understanding of the respective roles and responsibilities of the various stakeholders. In November 2004, Secretary-General Kofi Annan appointed forty individuals from government, the private sector, and civil society to the WGIG. After seven months of work, the WGIG's Report and a longer Background Report were released in July 2005. The two reports surveyed the landscape of public and private sector Internet governance mechanisms; advanced recommendations for improved governance in a range of key issue-areas; offered alternative models for the future oversight of the Internet's core resources and logical infrastructure; and proposed the establishment of a global, multistakeholder forum to facilitate continuing, inclusive dialogue on Internet governance. The reports elicited much debate around the world and served as key inputs in the second phase of the WSIS process, which culminated with the November 2005 Summit in Tunis. The issues they addressed remain of central importance in the post-Summit global debate on Internet governance.

With the continuing debate in mind, this volume presents papers by former members of the WGIG and its Secretariat staff that were written subsequent to the release of the group's reports. The contributors offer their personal reflections on the value of the multistakeholder cooperation in the WGIG and beyond; and on some of the key substantive issues and institutional reform proposals currently under consideration by the international community.

William J. Drake is the President of Computer Professionals for Social Responsibility, a global civil society organization, and an ICT policy consultant based in Geneva, Switzerland.

Secretariat of the United Nations ICT Task Force
One UN Plaza, New York, NY 10017
Web site: *http://www.unicttaskforce.org*

Secretariat of the Working Group on Internet Governance
Web site: *http://www.wgig.org*

9 789211 045574

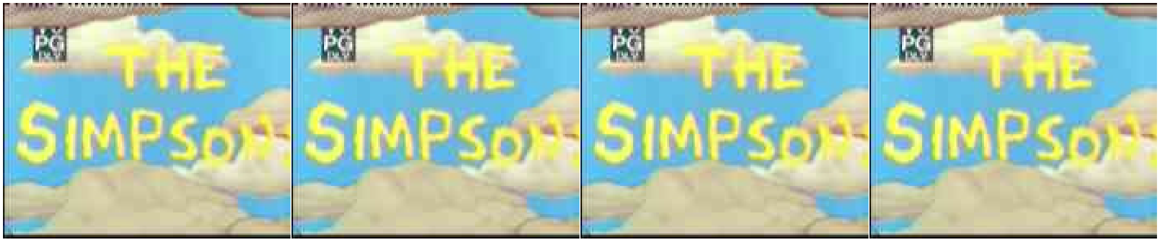
ISBN 92-1-104557-6
Sales No. E.05.II.A.17

53738—October 2005—1,300

Our goal

- We attempt to produce abstracts for video

Motivations



Motivations for video summarization

- Discover quickly if the video content is of interest without watching it
- Allow comfortable browsing experience in video databases

Example

- Find a particular episode in a TV series

Motivations

Motivations for video summarization

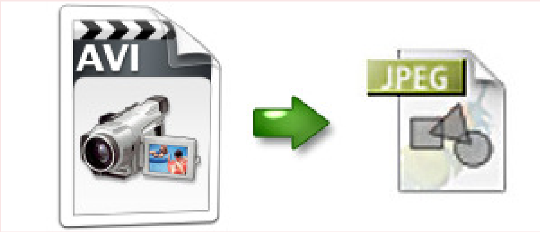
- Discover quickly if the video content is of interest without watching it
- Allow comfortable browsing experience in video databases

Example

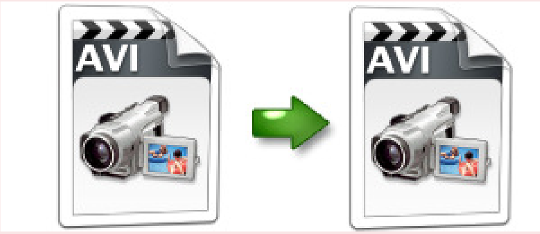
- Find a particular episode in a TV series

Video summary taxonomy

Static

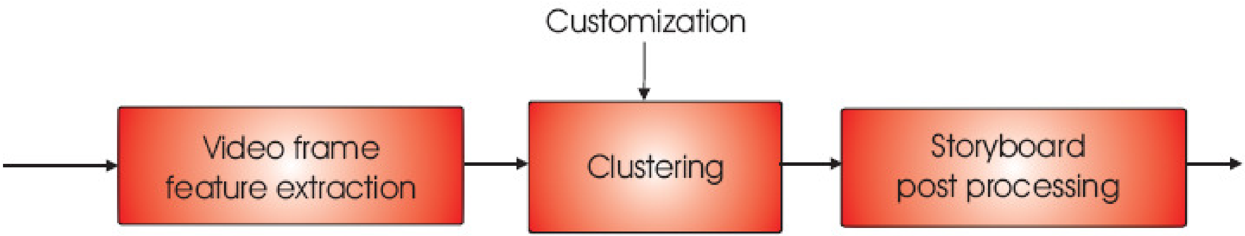


Dynamic



Summary computation time

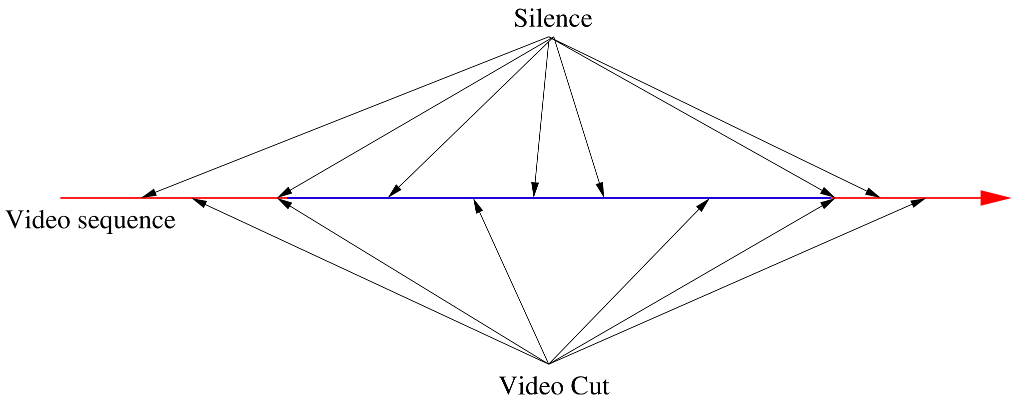
- .On-line: for web applications, allow user to personalize it (i.e choose summary size)
- .**Off-line**: must to be computed in advance. (It is not acceptable to wait for a time comparable with the video length)



Video Summarization Pipeline

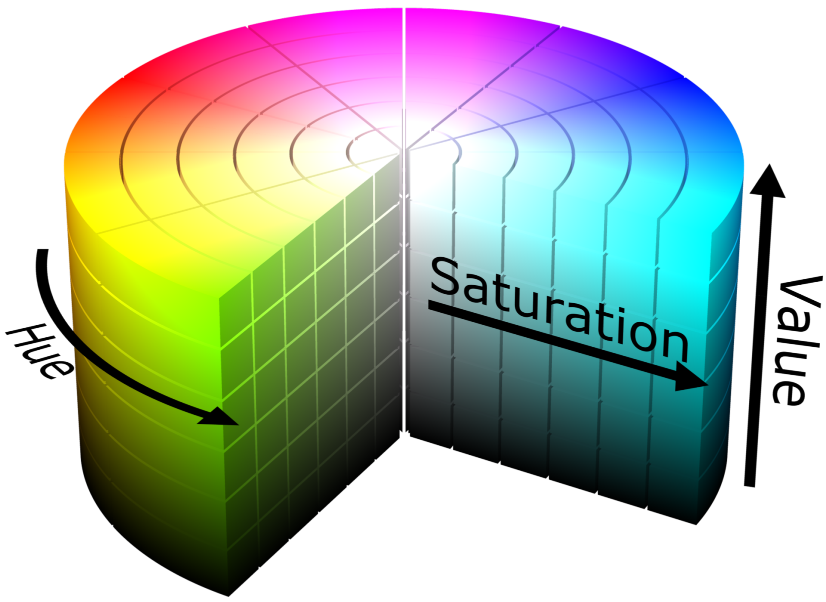
- **Input**: scenes/frames. 256-dimensional HSV vectors, extracted only once (on video submission)
- **Clustering**: determine the scenes to show. Customizations:
  - length of the storyboard,
  - clustering running time.
- **Post-processing**: filter out near-duplicates and extract the selected frames/scenes from the video.

Scenes



- **Scene**: bounded set of consecutive frames,
  - represented by the mean of the HSV vectors of the frames
- a movie is a partition of scenes,
- a frame fall in just a single scene,
- audio and video must be combined to determine scene,
  - video cut alone can be only a camera change,
  - silence is frequent in dialogs among people.

A metric space for video (1)



- Convert each pixel in its HSV representation
- Build a histogram of the frequences for each component

A metric space for video (2)

Generalized Jaccard distance (GJC)

- Given two vectors $H_1 = h_{1,1}, \ldots, h_{1,256}$ and $H_2 = h_{1,1}, \ldots, h_{1,256}$ representing two frames, the Generalized Jaccard distance is defined as:
- $$d(H_1.H_2) = 1 - \frac{\sum_{i=1}^{256} \min(h_{1,i}, h_{1,i})}{\sum_{i=1}^{256} \max(h_{1,i}, h_{1,i})}$$
- The vectors H_i endowed with the Generalized Jaccard distance define a metric space,
- the GJC exploit better the distance between two frames with respect to other distance functions used in the literature,
- GJC is fast and simple to compute.

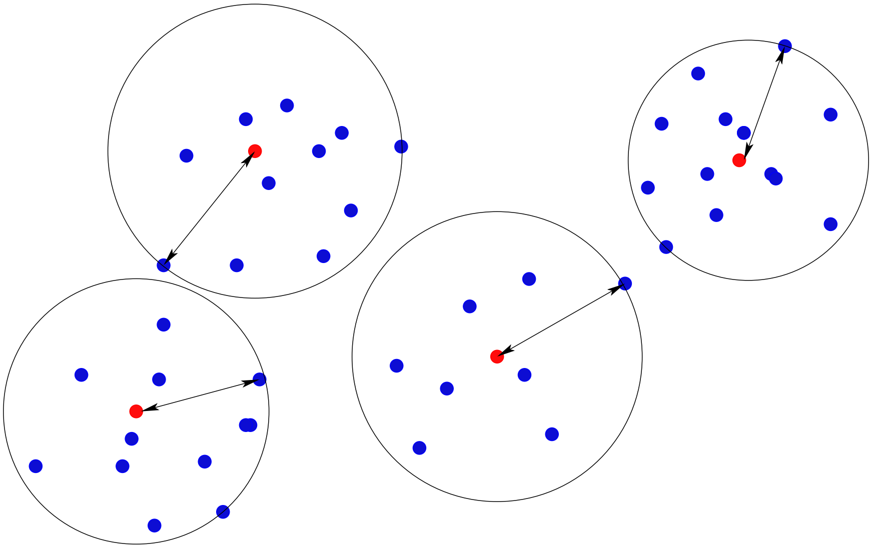
Clustering

- Partition data in homogeneous groups;
- Data independent procedure;
  - Available also when no a-priori knowledge about the data domain;
  - No problem dependent optimizations allowed;
- Many fields of application: IR, bioinformatics.

Cluster hypothesis

If two objects are closely related and the former is also related to a third object, then more likely also the latter has a similar relation.

K-center problem



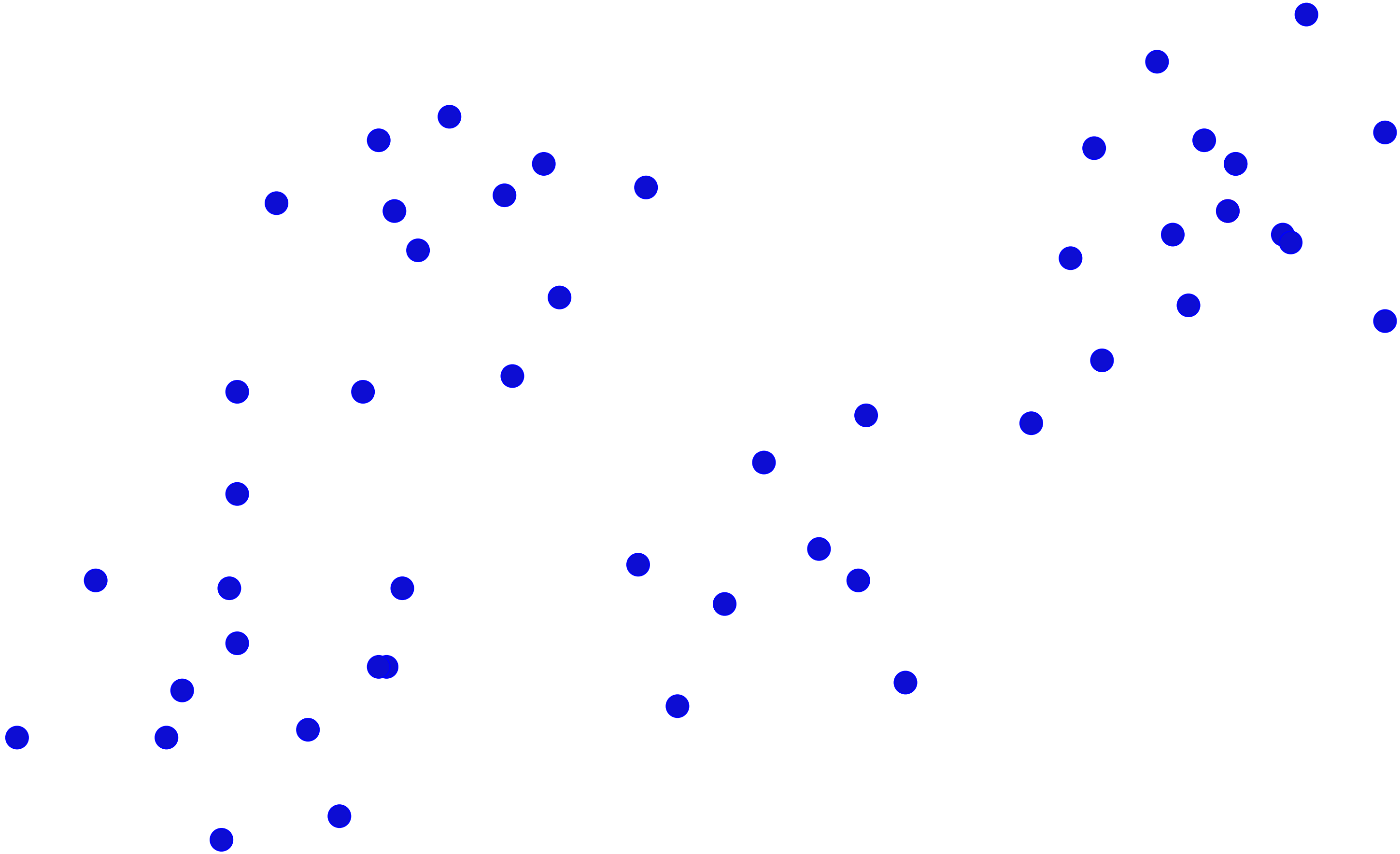
- K**-center** minimize the largest cluster diameter

$$\min_j \max_{x \in C_j} M_v(x, \mu_j)$$

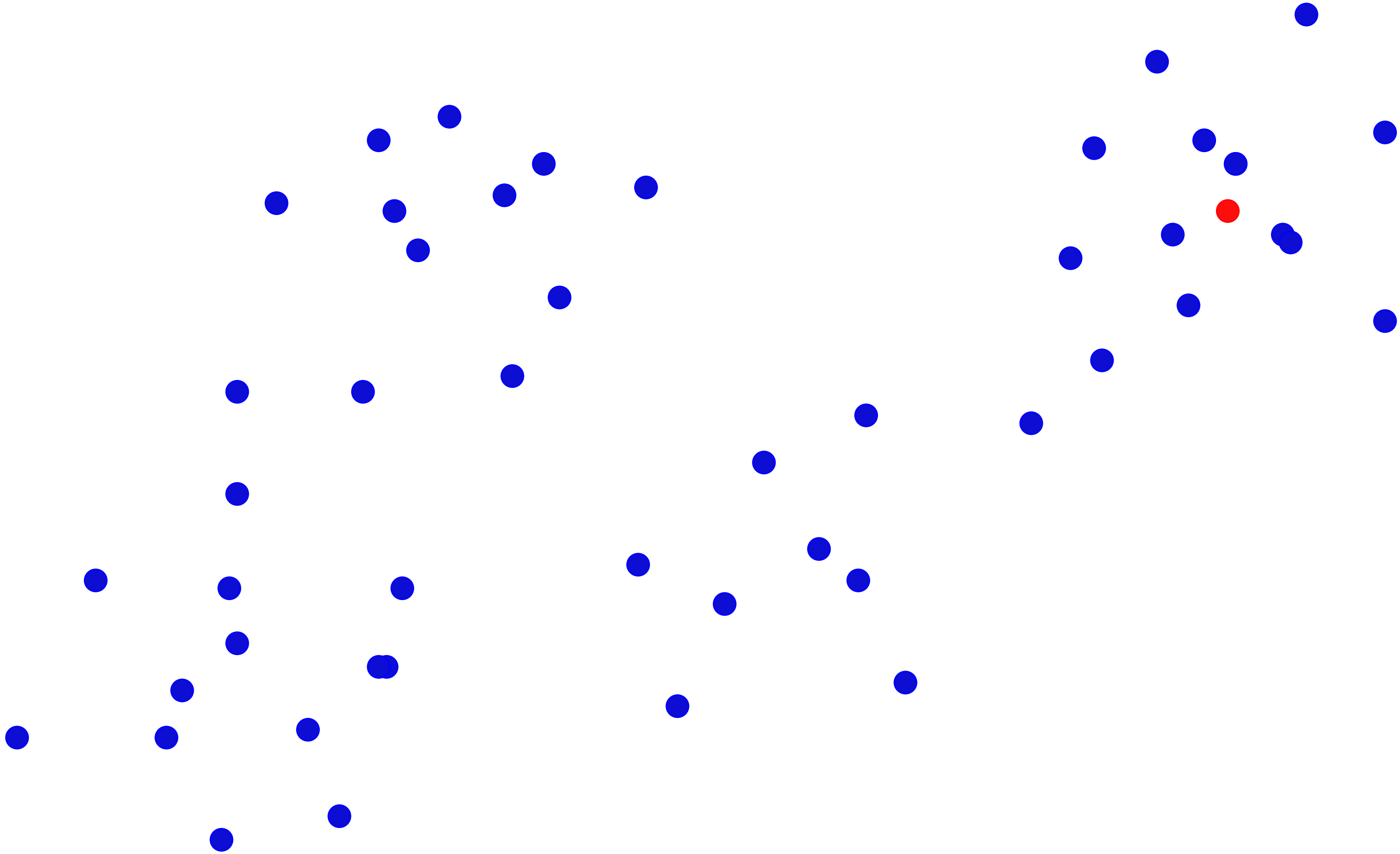
The furthest-point-first (FPF) algorithm

Data: Let O be the input set, k the number of clusters
Result: $\mathcal{C}$, k-partition of O
$C = x$ such that x is an arbitrary element of O;
for $i = 0; i < k; i++$ **do**
  Pick the element x of $O \setminus C$ furthest from the closest element in C;
  $C_i = \mathcal{C}_i = x$;
end
forall $x \in O \setminus C$ **do**
  Let i such that $d(c_i, x) < d(c_j, x), \forall j \neq i$ $\mathcal{C}_i$.append (x);
end

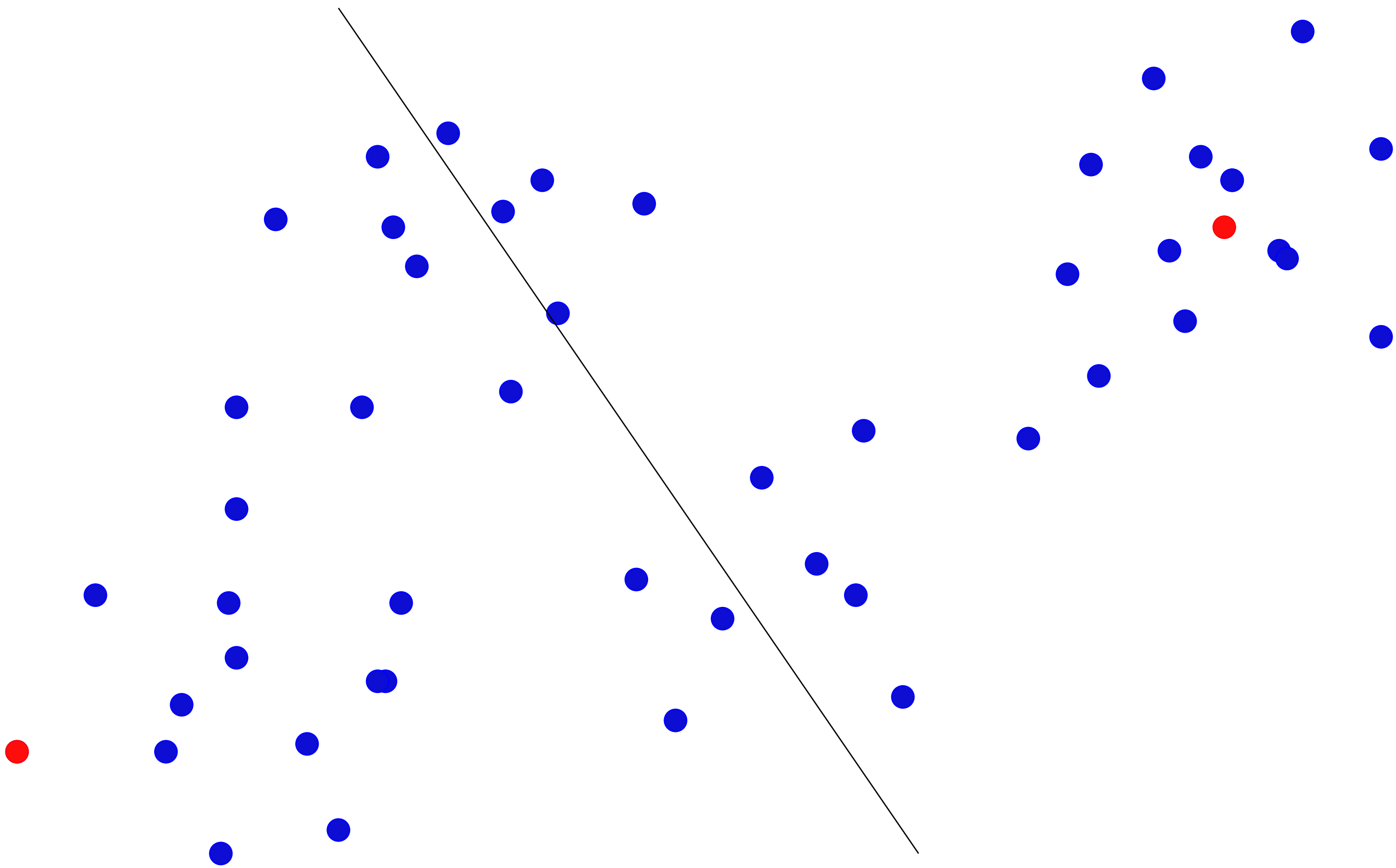
FPF algorithm example



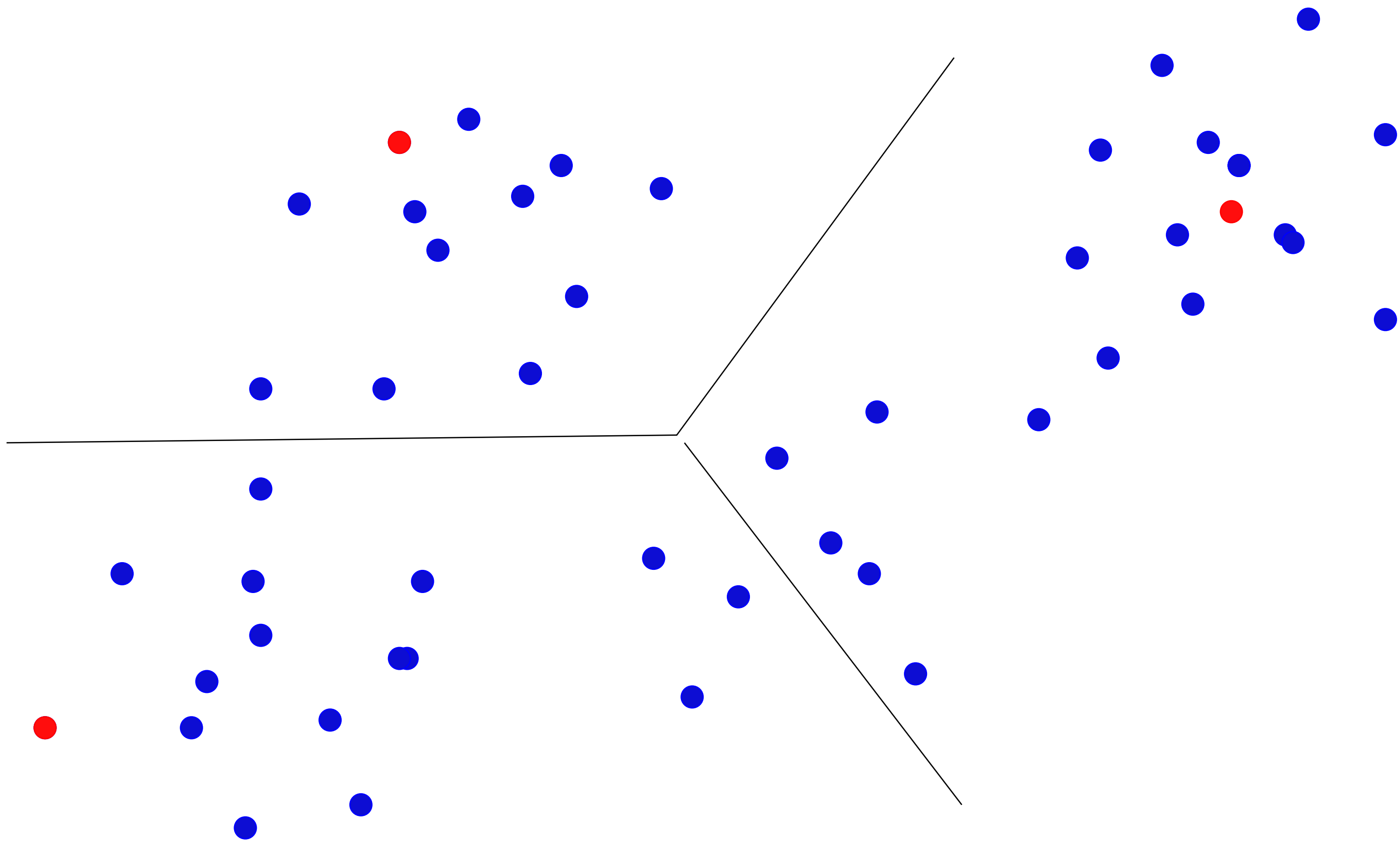
FPF algorithm example



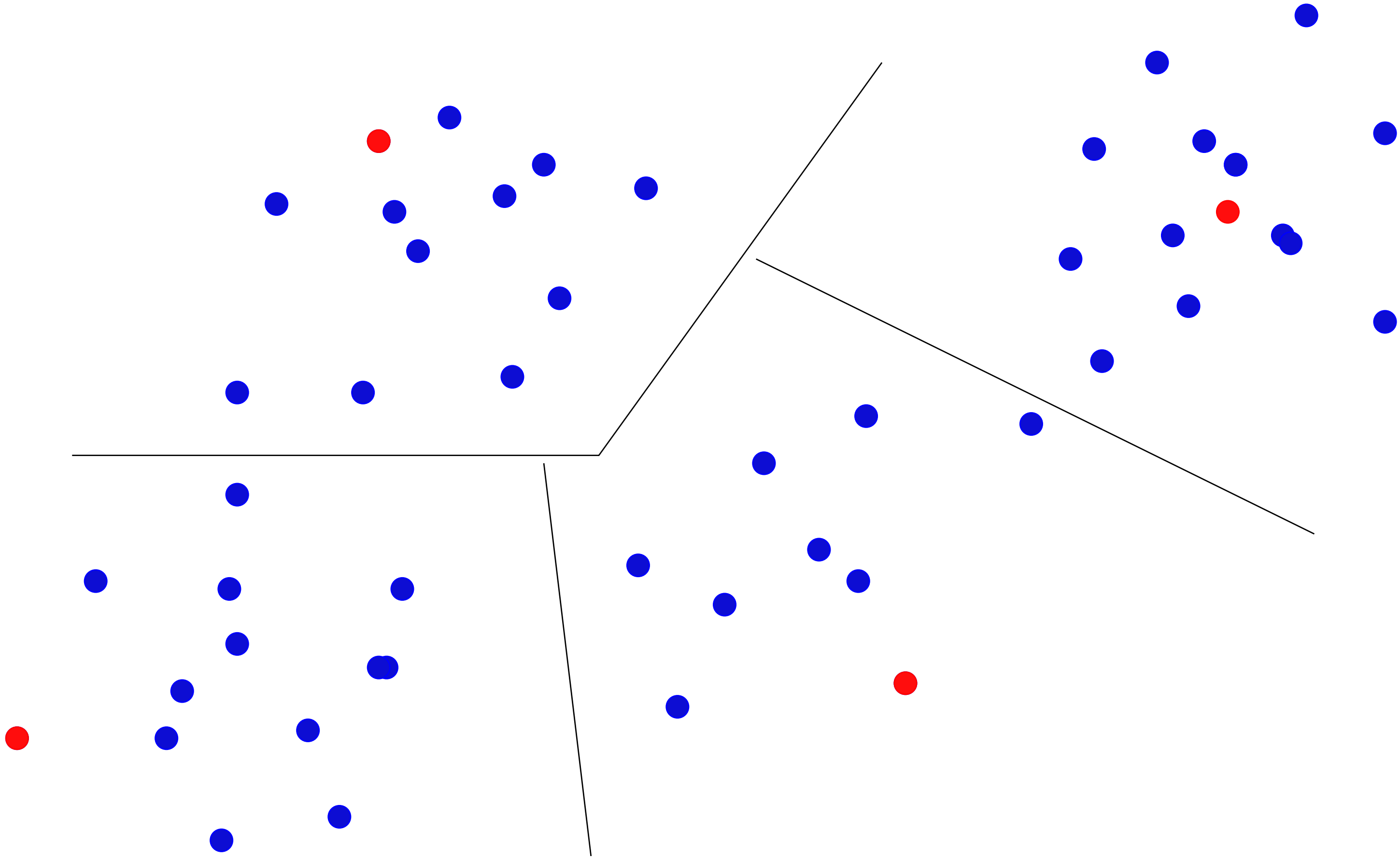
FPF algorithm example



FPF algorithm example



FPF algorithm example



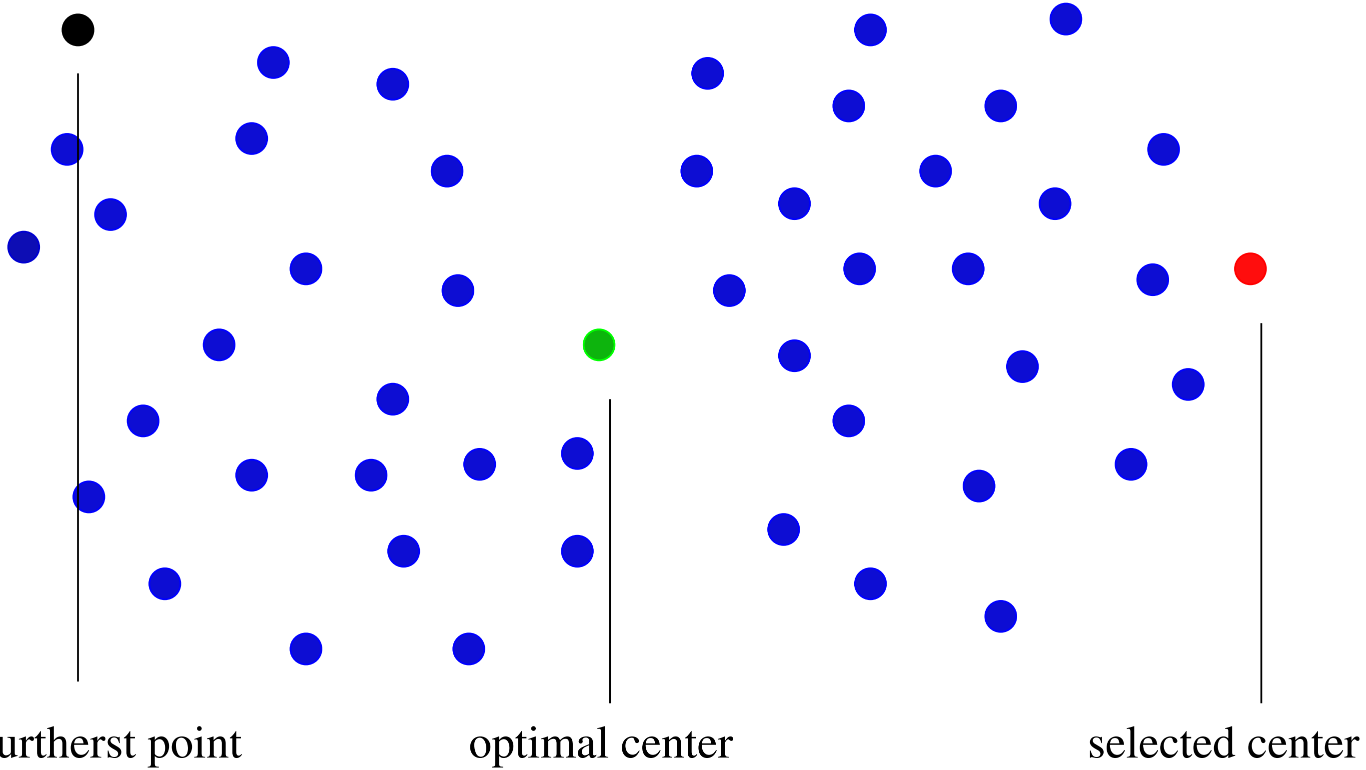
An incomplete prove of the 2-approximation

Based on two cases

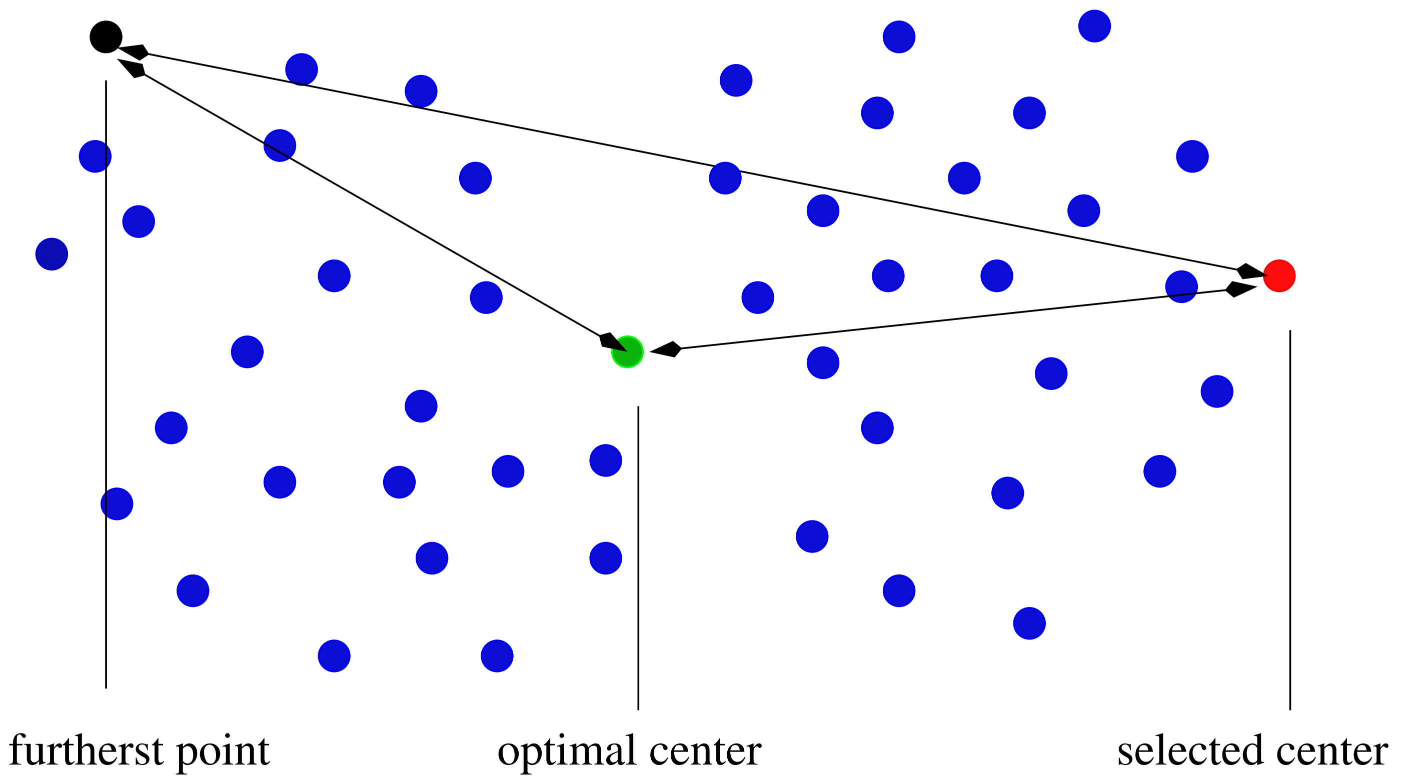
1. There is one center of FPF for each center of the optimal clustering
2. There are two centers of FPF for each center of the optimal clustering

Unproved here: there are no other possible cases

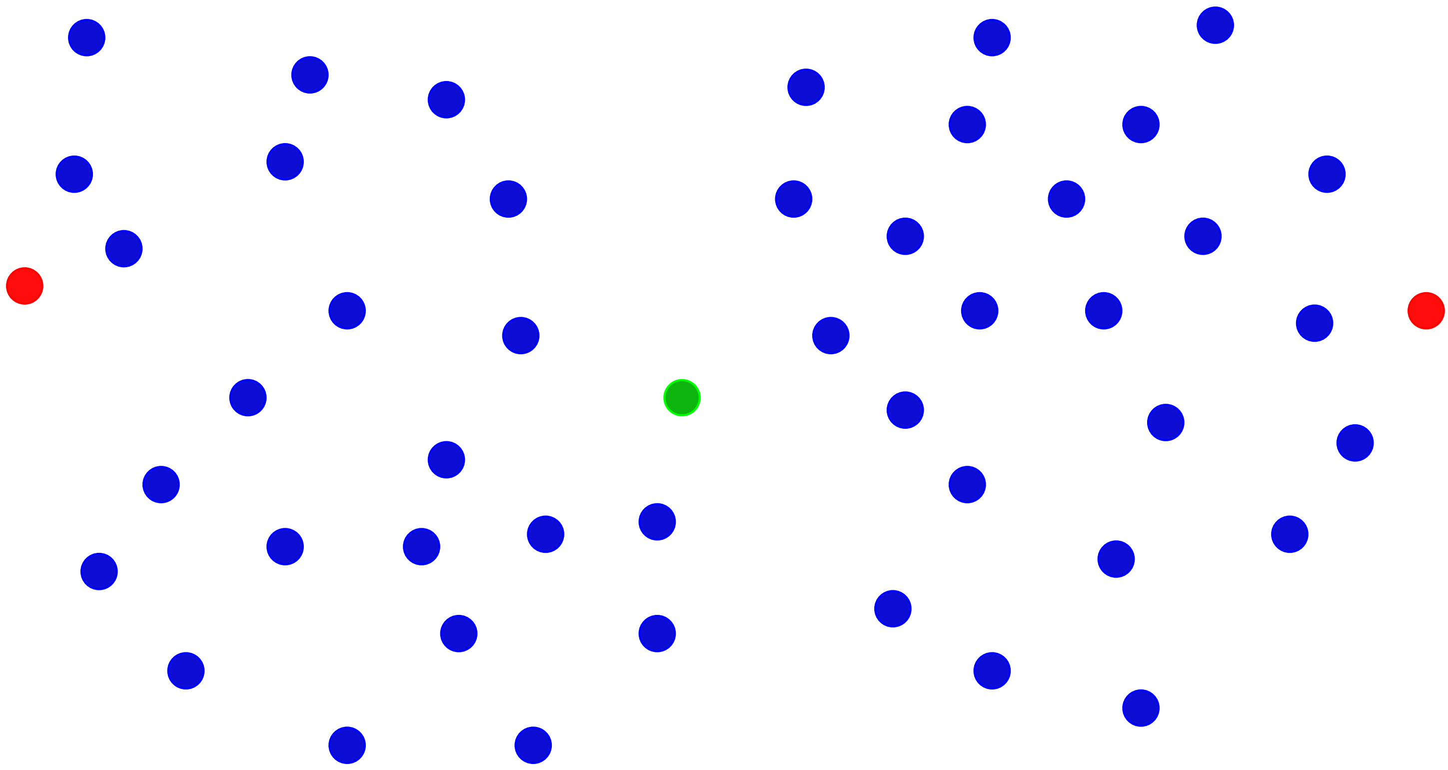
An incomplete prove of the 2-approximation - Example



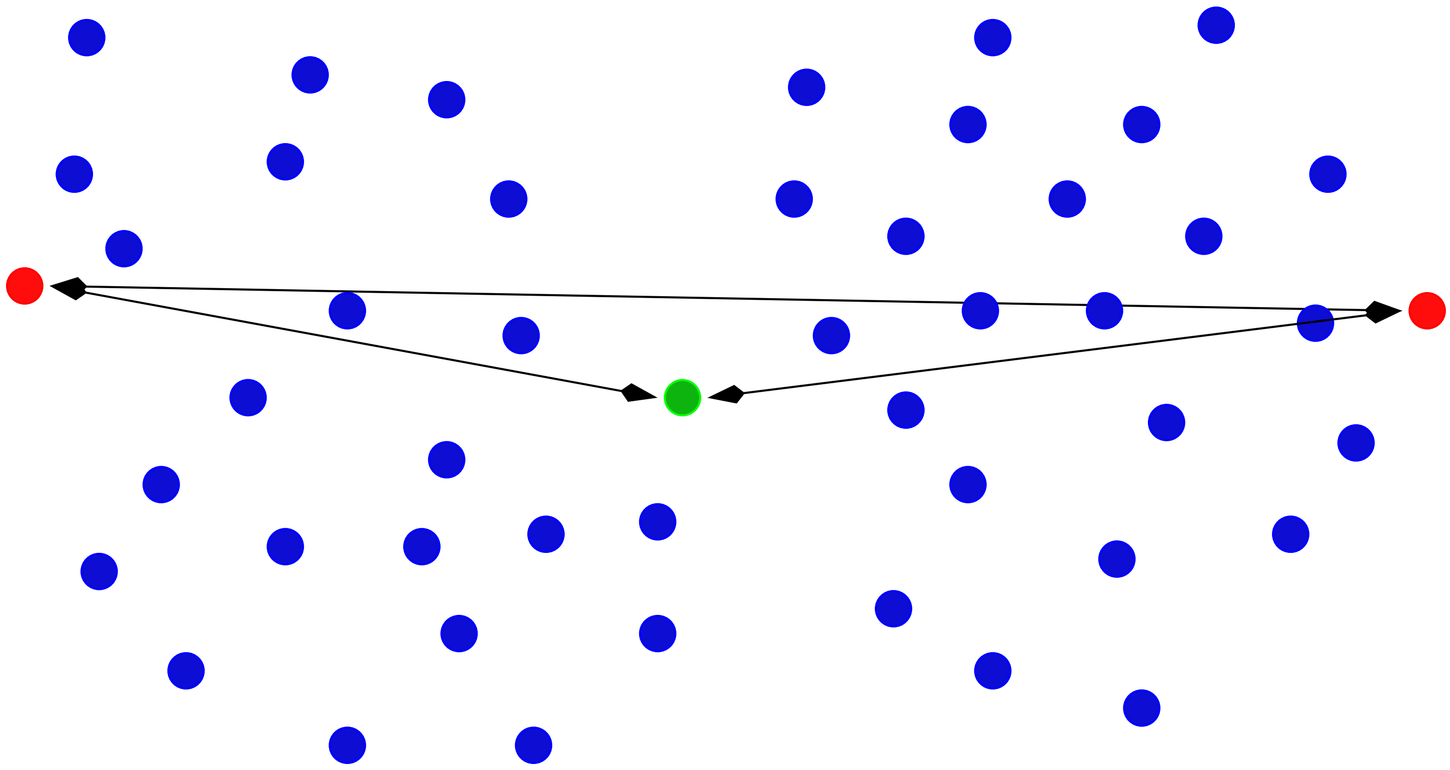
An incomplete prove of the 2-approximation - Example



An incomplete prove of the 2-approximation - Example



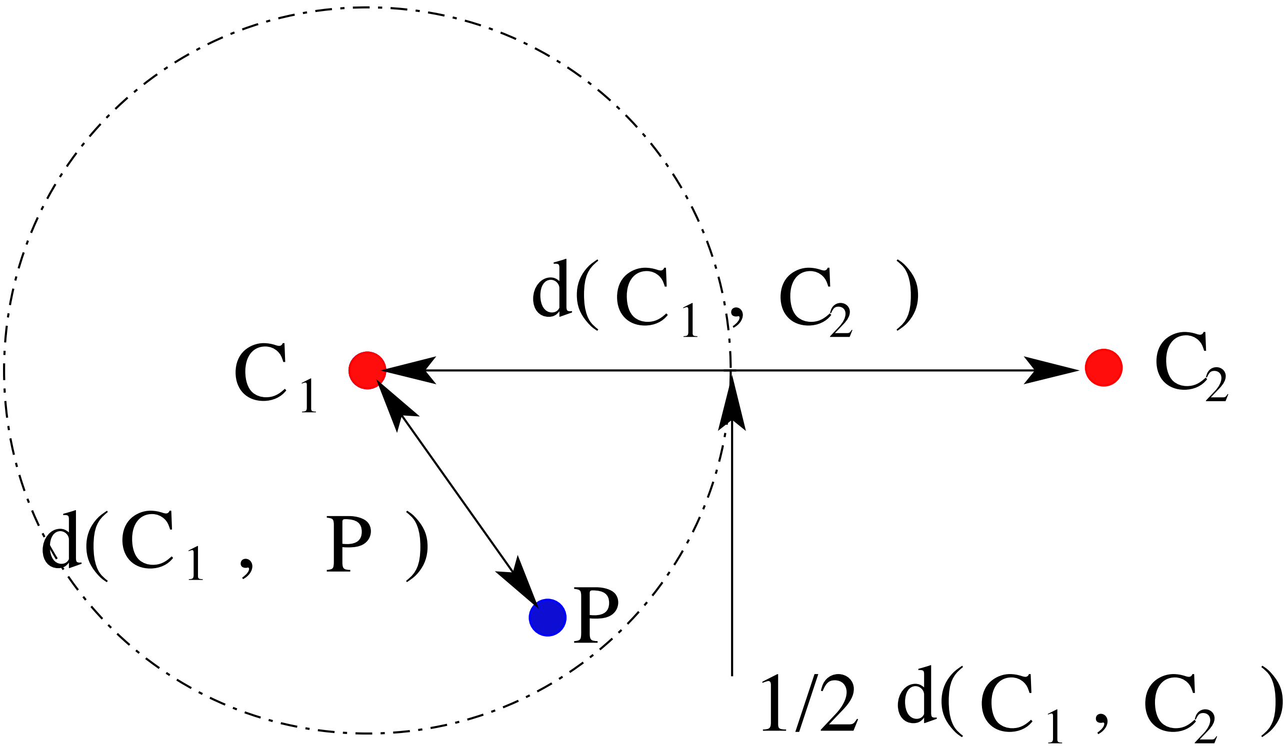
An incomplete prove of the 2-approximation - Example



The furthest-point-first (FPF) algorithm

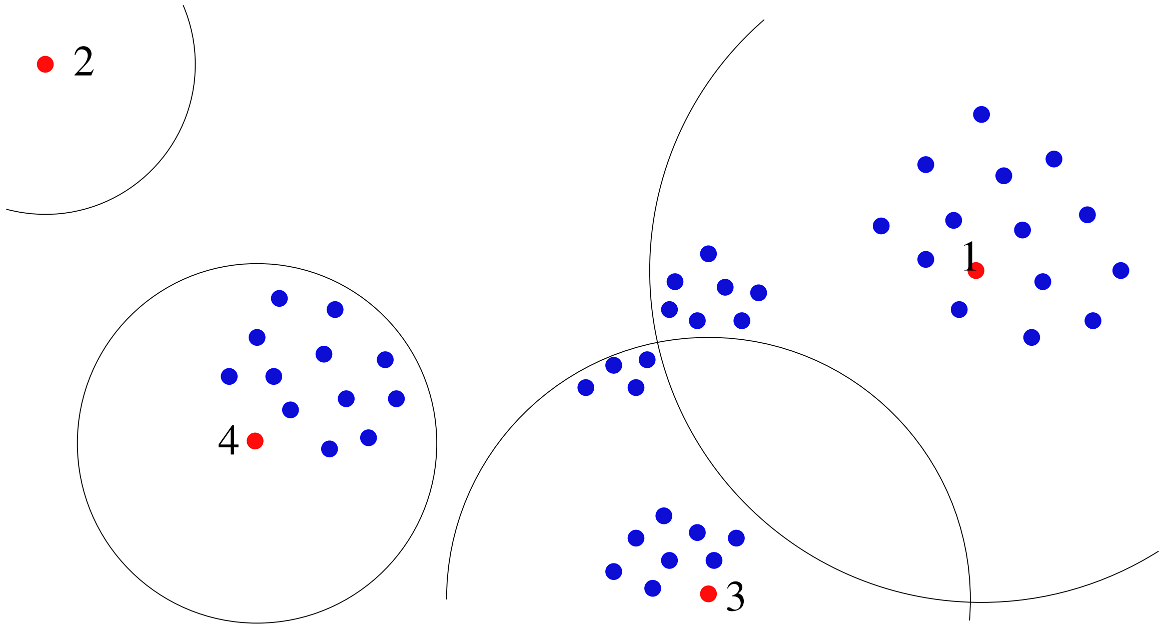
Data: Let O be the input set, k the number of clusters
Result: $\mathcal{C}$, k-partition of O
$C = x$ such that x is an arbitrary element of O;
for $i = 0; i < k; i++$ **do**
    Pick the element x of $O \setminus C$ furthest from the closest element in C;
    $C_i = \mathcal{C}_i = x$;
end
forall $x \in O \setminus C$ **do**
    Let i such that $d(c_i, x) < d(c_j, x), \forall j \neq i$ $\mathcal{C}_i$.append (x);
end

Exploit triangular inequality



- Most of the time spent to find the closest center
- In some cases can avoid to check all points of a cluster

Drawbacks in FPF

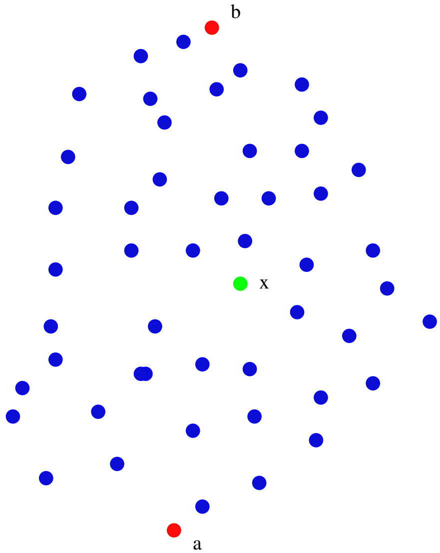


- Outlayers became centers with high probability
- Poor clusters and affect the choice of k
- Centers are not so representative

M-FPF

Data: Let O be the input set, k the number of desired clusters
Result: $\mathcal{C}$: a k-partition of O
Initialize R with a random sample of size $\sqrt{|O|k}$ elements of O;
$\mathcal{C} = \mathbf{FPF}(R, k)$;
forall $C_i \in \mathcal{C}$ **do**
  $\mu_i =$ getCenter (C_i);
end
forall p *in* $O \setminus R$ **do**
  assign p to cluster C_i such that $d(p, \mu_i) < d(p, \mu_j), \forall j \neq i$;
end

Medoids



- Given a set of points C, let $a, b \in C$ be two diametral points of C. The **medoid** of C is the point $x \in C$ that minimize:
- $M(x) = |D(a,x) - D(b,x)| + |D(a,x) + D(b,x)|$,

Diametral points heuristic

- Select a random point R
- find the farthest point A from R
- find the farthest point B from A
- Return (A, B) as the diametral pair

Partition around medoids

Data: Let O be the input set, k the number of desired clusters
Result: $\mathcal{C}$: a k-partition of O
Initialize R with a random sample of size $\sqrt{|O|k}$ elements of O;
$\mathcal{C}$ = **FPF**(R, k);
forall $C_i \in \mathcal{C}$ **do**
  t_i = getRandomPoint (C_i);
  $a_i = c_i$ such that max d(c_i, t_i) for each $c_i \in C_i$;
  $b_i = c_i$ such that max d(c_i, a_i) for each $c_i \in C_i$;
  $m_i = c_i$ such that $\min |d(c_i, a_i) - d(c_i, b_i)| + |d(c_i, a_i) + d(c_i, b_i) - d(a_i, b_i)|$;
end
forall p *in* $O \setminus R$ **do**
  assign p to cluster C_i such that $d(p, m_i) < d(p, m_j), \forall j \neq i$;
  if $d(p, b_i) > d(a_i, b_i)$ **then** $a_i = p$;
  if $d(p, a_i) > d(a_i, b_i)$ **then** $b_i = p$;
  if $d(p, b_i) > d(a_i, b_i)$ *or* $d(p, a_i) > d(a_i, b_i)$ **then**
    $m_i = c_i$ such that $\min |d(c_i, a_i) - d(c_i, b_i)| + |d(c_i, a_i) + d(c_i, b_i) - d(a_i, b_i)|$;
  end
end

How many clusters?

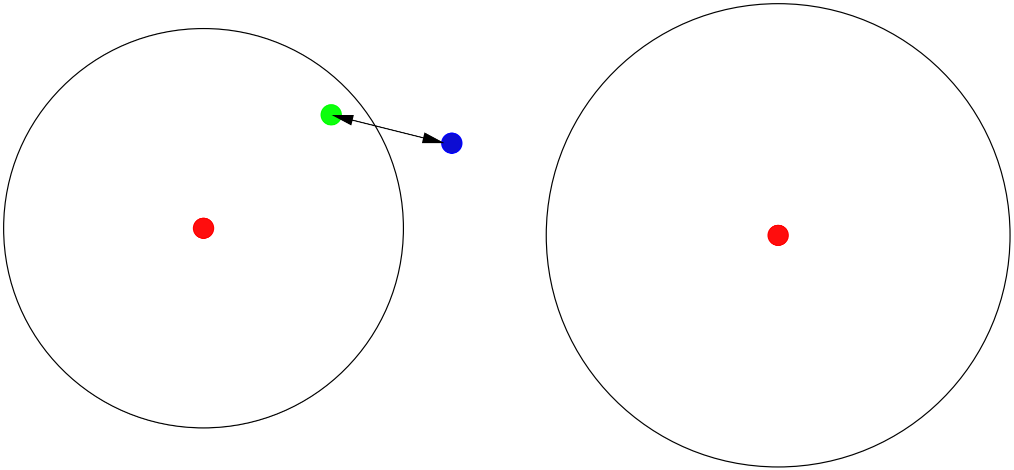
- Problem dependent question
- Few clusters are often not homogeneous
- Too many break homogeneous clusters

... and so?

- ad-hoc solutions
  - may be not feasible
- theoretical approachs (i.e stability):
  - always available
  - not related to the problem

Approximations for video summaries

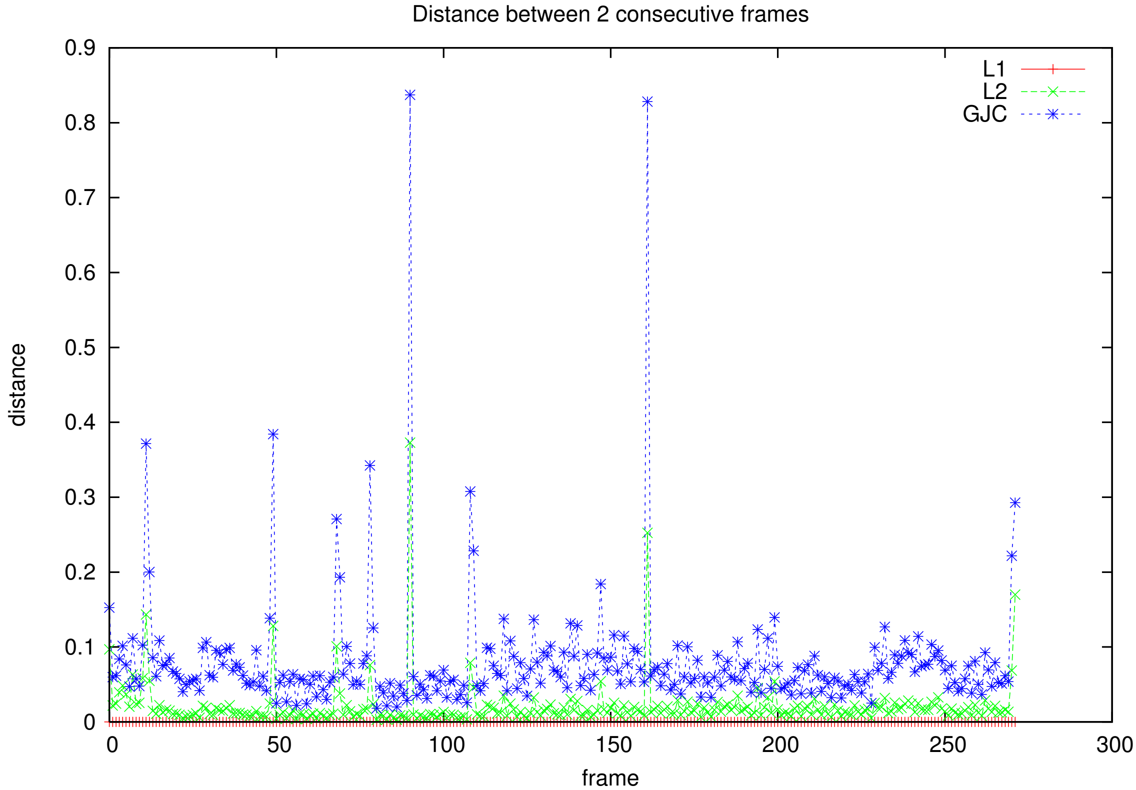
Approximations



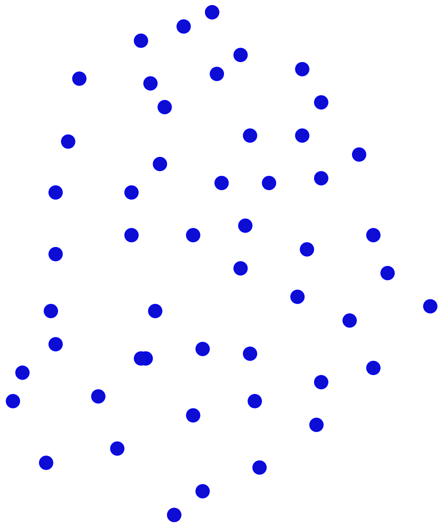
Adding a new point to cluster

- Most of the time is spent to decide for each point to which cluster it should be assigned.
- If the new point is “close” to the previous one, it is inserted in the same cluster.

Distance between two consecutive frames



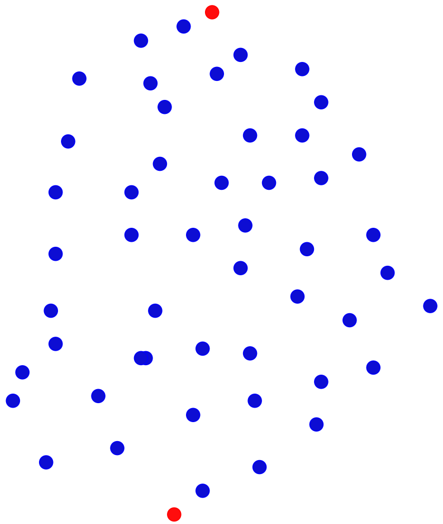
Approximations



Medoid

- Given a set of points C, let $a, b \in C$ be two diametral points of C. The **medoid** of C is the point $x \in C$ that minimize:
- $M(x) = |D(a,x) - D(b,x)| + |D(a,x) + D(b,x)|$,

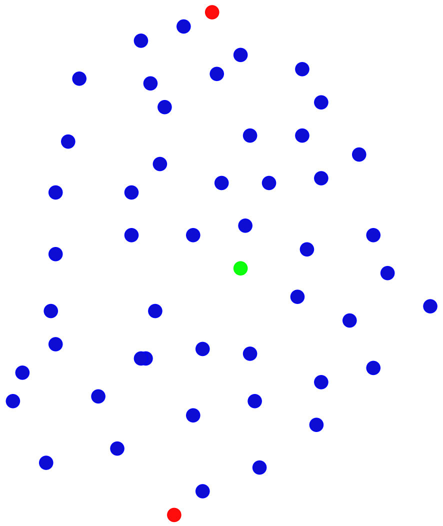
Approximations



Medoid

- Given a set of points C, let $a, b \in C$ be two diametral points of C. The **medoid** of C is the point $x \in C$ that minimize:
- $M(x) = |D(a,x) - D(b,x)| + |D(a,x) + D(b,x)|$,

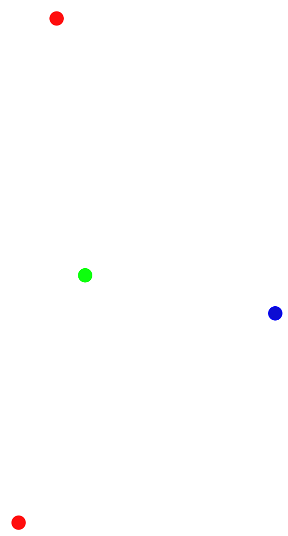
Approximations



Medoid

- Given a set of points C, let $a, b \in C$ be two diametral points of C. The **medoid** of C is the point $x \in C$ that minimize:
- $M(x) = |D(a,x) - D(b,x)| + |D(a,x) + D(b,x)|$,

Approximations



Medoid

- Given a set of points C, let $a, b \in C$ be two diametral points of C. The **medoid** of C is the point $x \in C$ that minimize:
- $M(x) = |D(a,x) - D(b,x)| + |D(a,x) + D(b,x)|$,

Video summary using the approximated FPF algorithm

References

- **Python module for clustering and other machine learing tasks**: `http://scikit-learn.org/stable/modules/clustering.html#clustering`