UNIVERSITEIT GENT

FACULTY OF ENGINEERING AND ARCHITECTURE

A queueing model for group-screening facilities

Dieter Claeys, Joris Walraevens, Bart Steyaert, Herwig Bruneel

EURO 2013

STOCHASTIC MODELLING AND ANALYSIS OF COMMUNICATION SYSTEMS

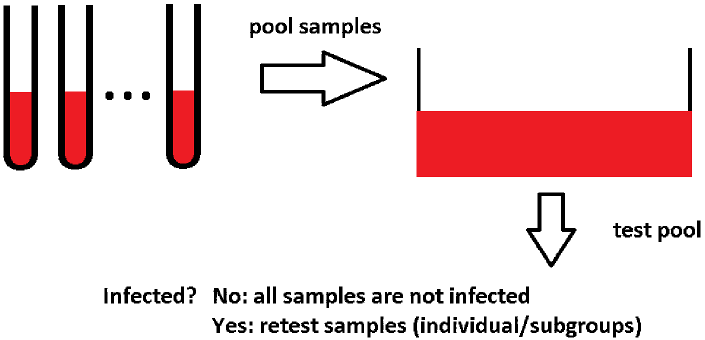
SMACS

Outline

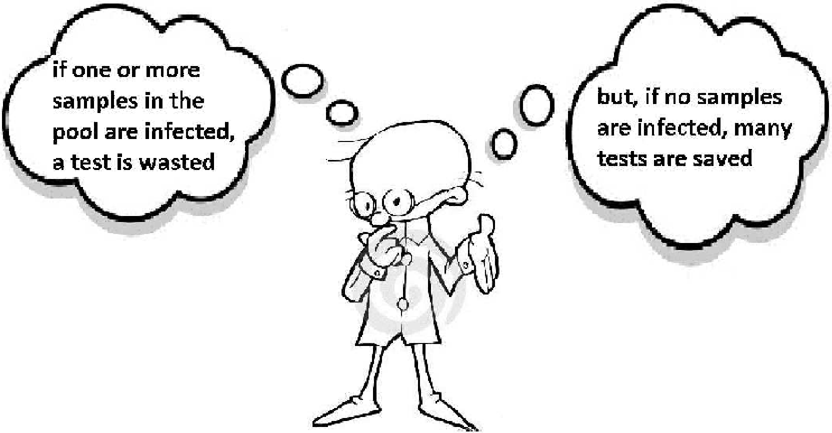
1. Group screening: what?
2. Queueing model
3. Static versus dynamic group screening
4. Comparison of results
   - Testing probability
   - Mean delay
5. Conclusion

Group screening

- WWII: detect syphilitic men drafted for military service
  - Expensive
  - Small prevalence rate
- Idea: test samples in group



Does it work?


if one or more samples in the pool are infected, a test is wasted
but, if no samples are infected, many tests are saved

Example

- Prevalence rate 1%
- 5 samples per pool
- Individual retesting

$$\Downarrow$$

- Pr[pool infected] $= 1 - (0.99)^5 \approx 5\%$
- If pool infected: 5 extra tests (+ 1 group test)
- Else: only 1 group test
- E[# tests per pool] $\approx 0.05*6 + 0.95*1 =$ 1.25 instead of 5

Other applications of group screening

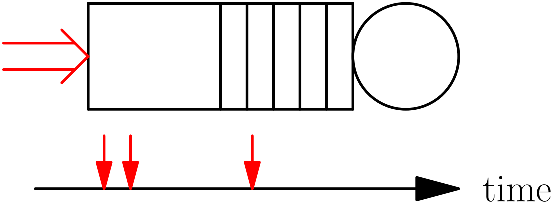
- Screening for HIV, Influenza, West Nile Virus
- DNA library screening
- Drug discovery
- Quality control

Optimal group size

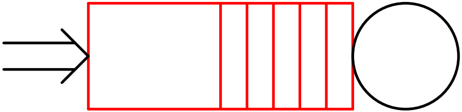
- Larger group size:
  - More items (samples) in a group (pool)
  - Larger probability that group is bad
- Dorfman: standard model
  - Many items to be screened
  - Items present from the start
- Practical context usually not static but dynamic
  - Items arrive over time
  - Extra decision variable: minimum group size
  - Queueing model

Outline

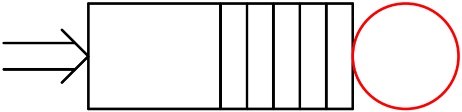
- Items arrive spread over time



- Items arrive spread over time
- Await in queue their screening



- Items arrive spread over time
- Await in queue their screening
- By the screening facility



Screening

- In group: batch/bulk service
- Server capacity c: maximum group size
- Minimum batch size l: minimum group size

Screening

- Service time of group: # required tests
- Dependent on # items in the group
- Capture screening policy in distribution screening time

Example: individual retesting

- S_j: service time of a group of j items
- $S_j(z)$: probability generating function of S_j
- $\overline{p} \triangleq 1-$prevalence rate
- $S_j(z) = \overline{p}^j z + (1 - \overline{p}^j) z^{j+1} \quad , \qquad j \geq 2$

Performance measures

- Testing probability f: fraction of slots during which server (test unit) is serving (screening)

- Mean delay $\overline{D}$: average time between the arrival of an item and the moment at which the result of the item is known

Numerical example: f

	$c=6$	$c=7$	$c=8$	$c=9$	$\mathbf{c=10}$	$c=11$	$c=12$
$l=1$	0.9796	0.9786	0.9780	0.9778	0.9777	0.9777	0.9778
$l=2$	0.9209	0.9173	0.9150	0.9143	0.9138	0.9141	0.9143
$l=3$	0.8475	0.8406	0.8361	0.8348	0.8338	0.8344	0.8349
$l=4$	0.7852	0.7754	0.7689	0.7672	0.7658	0.7669	0.7677
$l=5$	0.7436	0.7316	0.7236	0.7216	0.7200	0.72159	0.7228
$l=6$	0.7170	0.7020	0.6928	0.6907	0.6889	0.6909	0.6926
$l=7$	——	0.6891	0.6778	0.6756	0.6738	0.6761	0.6780
$l=8$	——	——	0.6679	0.6653	0.6635	0.6661	0.6682
$l=9$	——	——	——	0.6631	0.6609	0.6638	0.6660
$\mathbf{l=10}$	——	——	——	——	**0.6591**	0.6626	0.6648
$l=11$	——	——	——	——	——	0.6651	0.6678
$l=12$	——	——	——	——	——	——	0.6698

Numerical example: $\overline{D}$

	$c = 4$	$c = 5$	$c = 6$	$c = 7$	$\mathbf{c = 8}$	$c = 9$	$c = 10$
$l = 1$	57.4139	42.4111	37.6298	37.6303	37.5259	39.0722	40.2516
$\mathbf{l = 2}$	57.2731	42.2940	37.5278	37.5399	**37.4464**	39.0008	40.1824
$l = 3$	60.1768	45.0087	40.1156	40.1894	40.1384	41.8741	43.2097
$l = 4$	63.3439	48.1818	43.2285	43.3755	43.3763	45.2714	46.7414
$l = 5$	——	52.5499	47.4802	47.6944	47.7292	49.7700	51.3546
$l = 6$	——	——	51.1644	51.6063	51.6833	53.8279	55.4936
$l = 7$	——	——	——	56.5320	56.7008	58.9817	60.7385
$l = 8$	——	——	——	——	60.9422	63.4986	65.3284
$l = 9$	——	——	——	——	——	68.6834	70.6856
$l = 10$	——	——	——	——	——	——	75.3074

Outline

Numerical complexity

- $\forall\ c$: calculate c roots
- $\forall\ (l,c)$: solve set of c equations in c unknowns
- Much numerical work
- Use results from model of Dorfman?

Static model

- Many items (N), all present from the start
- Average # tests ($\mathrm{E}[T]$) to screen all items
- Example: individual retesting:

$$\mathrm{E}[T] \sim \frac{N}{c}\left[1 + c(1 - \overline{p}^{c})\right]$$

for $N \to \infty$

Dynamic versus static

- Group size(s)
  - Static: one group size (no minimum)
  - Dynamic: minimum and maximum group size (l and c)
- Performance measures
  - Static: average # tests ($\mathrm{E}[T]$)
  - Dynamic: testing probability (f), mean delay of items ($\overline{D}$)
- Input parameters
  - Static: # items to be screened (N)
  - Dynamic: distribution of # item arrivals in a time unit ($A(z)$)

Outline

Outline

- l_f: minimum group size that minimizes f
- c_f: maximum group size that minimizes f
- c_s: optimal static group size (minimizes $\mathrm{E}[T]$)
- Main result: $l_f = c_f = c_s$, regardless of
  - N (static)
  - $A(z)$ (dynamic)
  - Prevalence rate

Outline

- $l_{\overline{D}}$: minimum group size that minimizes $\overline{D}$

- $c_{\overline{D}}$: maximum group size that minimizes $\overline{D}$

- Results:
  - $c_{\overline{D}} \neq c_s$: numerical work involved
  - $l_{\overline{D}} = 1$ good heuristic: reduces complexity

Search space for $c_{\overline{D}}$ (1)

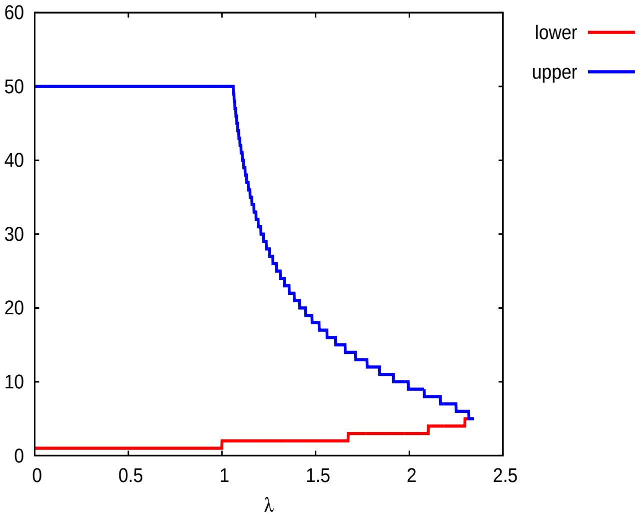
- Stability condition:

$$\rho \triangleq \lambda \frac{\mathrm{E}[S_c]}{c} < 1$$

  - λ: mean arrival rate (given)
  - c: group size
  - S_c: # tests to screen group of c items

Search space for $c_{\overline{D}}$ (2)

- ρ smallest for $c = c_s$
- $c \uparrow$ or $c \downarrow$ (as compared to c_s) $\Rightarrow \rho \uparrow$
- c too small or too large $\Rightarrow \rho > 1$
- No numerical work


60
50
40
30
20
10
0
0
0.5
1
1.5
2
2.5
λ
lower
upper

Algorithm for small λ

- Search space largest for small λ
- Algorithm for small λ
- Two scenarios:
  - Bursty arrivals
  - Poisson like arrivals (e.g. Poisson, Bernoulli, geometric)

Bursty arrivals

Main result: $c_{\overline{D}} = c_s$ good heuristic

Poisson like arrivals

Taylor series expansion of $\overline{D}(\lambda)$ about $\lambda = 0$:

$$\overline{D}(\lambda) = \overline{D}_0 + \lambda\overline{D}_1 + \lambda^2\overline{D}_2 + \dots$$

Theorem

$\overline{D}_j$ takes into account the possibilities

of having $1, 2, \dots, j+1$ arrivals

in a time unit with item arrivals

Poisson like arrivals: $\overline{D}_0$

$$\overline{D}(\lambda) = \overline{D}_0 + \lambda \overline{D}_1 + \lambda^2 \overline{D}_2 + \ldots$$

- 1 item arrival
- No information about $c_{\overline{D}}$

Poisson like arrivals: $\overline{D}_1$

$$\overline{D}(\lambda) = \overline{D}_0 + \lambda \overline{D}_1 + \lambda^2 \overline{D}_2 + \dots$$

- 1 or 2 arrivals
- $c = 1$ versus $c \geq 2$
- $3/2\mathrm{E}\left[S_1\right]$ versus $\mathrm{E}\left[S_2\right]$
- If $c = 1$ best: stop; else: continue with $\overline{D}_2$

Poisson like arrivals: $\overline{D}_k$

$$\overline{D}(\lambda) = \overline{D}_0 + \ldots + \lambda^k \overline{D}_k + \ldots$$

- 1 or 2, or, ..., $k+1$ arrivals
- $c = k$ versus $c \geq k$
- $\mathrm{E}[S_k] + 1/(k+1)\mathrm{E}[S_1]$ versus $\mathrm{E}[S_{k+1}]$
- If $c = k$ best: stop; else: continue with $\overline{D}_{k+1}$

Outline

- Queueing model to include dynamic nature of item arrivals
  - Batch/bulk service
  - Screening policy captured in service times
  - Testing probability f and mean delay $\overline{D}$
  - Requires (much) numerical work
- When are static results useful in dynamic context?

```
if optimization criterion == f
     optimal l = optimal c = c_s;
if optimization criterion == delay
     optimal l = 1 (heuristic);
     if small arrival_rate
          if bursty arrivals
               optimal c = c_s (heuristic);
          if Poisson arrivals
               efficient algorithm for optimal c
     else
          determine search space
          calculate delay for all group sizes from
          the search space and select the group size
          that produces smallest delay
```

Future work

- $c_{\overline{D}} \leq c_s$?
- $c_{\overline{D}}$ non-decreasing function of λ?
- Accurate closed-form approximation for $\overline{D}$ for medium λ
- Same conclusions in case of False Positives and False Negatives?

Questions?