A Power-Enhanced Algorithm for Spatial Anomaly Detection in Binary Labelled Point Data Using the Spatial Scan Statistic

Simon Read[1], Peter Bath[1] Peter Willett[1], and Ravi Maheswaran[2]

[1] Department of Information Studies, University of Sheffield, Sheffield S1 4DP, UK
simon.read@sheffield.ac.uk
[2] ScHARR, University of Sheffield, Sheffield S1 4DA, UK

Abstract. This paper presents a novel modification to an existing algorithm for spatial anomaly detection in binary labeled point data sets, using the Bernoulli version of the Spatial Scan Statistic. We identify a potential ambiguity in p-values produced by Monte Carlo testing, which (by the selection of the most conservative p-value) can lead to sub-optimal power. When such ambiguity occurs, the modification uses a very inexpensive secondary test to suggest a less conservative p-value. Using benchmark tests, we show that this appears to restore power to the expected level, whilst having similarly retest variance to the original. The modification also appears to produce a small but significant improvement in overall detection performance when multiple anomalies are present.

1 Introduction

The detection of spatial anomalies (a.k.a. 'spatial cluster detection') in binary labeled point data has important applications in analyzing the geographic distribution of health events (e.g. [1]) and other fields such as forestry (e.g. [2]). Since 1997, the freely available SaTScan^{TM}software package (www.satscan.org) has provided a means of detecting such anomalies using the Spatial Scan Statistic [3] and has been used in well over a hundred published scholarly studies [4]. Section 2 gives more details of the research context.

Combined with a moving scan window procedure, the statistic identifies the location, size and statistical significance (p-value) of potential anomalies within a given study region. Due to the nature of the Spatial Scan Statistic as it is currently applied to binary labeled point data (see primer in Section 3), it is sometimes not possible to establish an exact p-value for the most likely potential anomaly. This ambiguity can occur in various places on the unit interval, including the most useful part of the range, $0 \leq$ p-value ≤ 0.1. This issue is explained in Section 4. When ambiguity occurs, SaTScan^{TM}selects the most conservative p-value, resulting in a sometimes lower-than-expected false positive rate (FPR), and a corresponding reduction in true positive rate (TPR, a.k.a. statistical power or sensitivity).

The principal contribution of this paper is to describe a means by which, when p-value ambiguity occurs, otherwise redundant information can be used

to meaningfully (and consistently) suggest a less conservative p-value. This is described in Section 5. Using benchmark tests (Section 6) we show this produces a false positive rate very close to the nominal significance level, and correspondingly increases the power in the circumstances outlined above. The proposed algorithm also delivers a p-value consistency (i.e. mean retest variance) comparable to SaTScan™

The secondary contribution is that, when applied to data sets where several anomalies are present, the modification appears to produce a small improvement in the ratio of true and false positive rates, as measured using area under curve (AUC) as applied to ROC curves. We use a 'within datasets' Monte Carlo method to show this improvement to be statistically significant, as described in Section 6. A discussion of the results, and future research directions, is given in Section 7.

2 Research Context

Spatial anomaly detection has three broad categories: global (identifying if any anomaly is present in the study region, but not specifying a location); localised (as global, but specifying location) and focused (testing for the presence of an anomaly at a location specified *a priori*). The Spatial Scan Statistic [3] is a widely use method of localised anomaly detection. Localised is the most flexible, as it can perform the function of the other two, albeit possibly with sub-optimum power. Following on from [3] many frequentist versions of the statistic have been developed and compared, (see list in [4]) as well a Bayesian version [5]. Mostly these are for use with areal data, e.g. disease counts in postal districts. The Bernoulli version (hereafter SSS_B), is for use with binary labeled point data. Despite being introduced in [3] and used in various studies since (e.g. [1], [2]), there has been little research into the benchmark performance of the SSS_B. Recently [6] used the SSS_B when considering a alternative circular scan window selection method (scan windows, termed here Z_j, are defined in Section 3), and [7] have developed a risk-adjusted SSS_B variant. Although the latter is of some interest, it appears to have lower-than-expected FPR and TPR, so in this paper we only consider the original SSS_B. It is also worth noting that many studies have been published into the effect of using different shaped scan windows, of which a useful summary is given in [8]. The results of this study should be applicable to all types of scan window, provided they can be applied to point data.

3 Primer: Spatial Scan Statistic (Bernoulli version)

The Spatial Scan Statistic has several versions. The Bernoulli (hereafter SSS_B) is suitable to binary labeled point data. For the benefit of readers unfamiliar with the Spatial Scan Statistic, this section formally defines[3] the SSS_B, its ac-

[3] Regarding notation: italic lower-case = scalar; italic upper-case = set (or multiset), bold upper-case = set (or multiset) of sets (or multisets).

companying data structures and method of application [4]. For a derivation of the statistic see [3].

Consider a spatial region R , with r any point location therein. Consider a data set $P = \{p_1,p_2...p_N\}$ where each p_i is associated a single point location $loc_i \in R$, and a binary label s_i. Let $P_0 = \{p_i : s_i = 0\}$ and $P_1 = \{p_i : s_i = 1\}$, such that $P_0 \cup P_1 = P$ and $P_0 \cap P_1 = \oslash$. Let N and the position of each loc_i be taken as given, but assume each s_i value to be the outcome of an independent Bernoulli trial with probability $p(s_i = 1) = \rho(r)$, where $\rho(r)$ is some arbitrary value (on the unit interval) associated with point r. Let H_0 represent the (null) hypothesis that $\rho(r)$ is constant for all $r \in R$. That is, the distribution of the elements of P_1 amongst P is uniformly random. Let H_A represent the (alternate) hypothesis that a spatial anomaly is present, i.e. there is a subset of R where $\rho(r)$ is higher (or lower) than the rest of R. Put formally, $H_A \Leftrightarrow (\exists A \subset R$, hence $\exists B = R - A)$ such that $\rho(a \in A) = \beta\rho(b \in B)$, where β is a constant[5] $\neq 1$. In this study we only consider $\beta > 1$, but the results will apply equally to $\beta < 1$. This Bernoulli model is useful for representing point occurrences in many real-world applications, as it controls for a inhomogeneous underlying distribution of events.

In real data sets, A (if it exists) can only be estimated by guessing which loc_i lie inside or outside it. Let us call any particular estimate Z. Furthermore, let us assume we have some predefined scheme (typically a moving scan window of variable size) for generating a set of estimates, $\mathbf{Z} = \{Z_1, Z_2, Z_3 \ldots\}$, where each $Z_j \subset P$. The purpose of the SSS_B is to determine which Z_j (let us call this Z_{prime}) is most likely to represent[6] A, if indeed A exists. We then associate a p-value with Z_{prime}, which represents the probability that H_0 is true (in which case Z_{prime} is a random artefact). To use the SSS_B, we split all Z_j such that $Z_{j0} = \{p_i : p_i \in Z_j$ and $s_i = 0\}$ and $Z_{j1} = \{p_i : p_i \in Z_j$ and $s_i = 1\}$. For each Z_j the SSS_B takes four integer inputs ($N = |P|$, $C = |P_1|$, $n = |Z_j|$ and $c = |Z_{j1}|$) and produces one quasi-continuous output, the *log likelihood ratio* or LLR. The formula is given in three parts: Equation 1[7] gives the likelihood of H_A if Z_j represents A; Equation 2 gives the likelihood of H_0, identical of all choices of Z_j. Equation 3[8] brings the two values together to produce the LLR. Testing all $Z_j \in \mathbf{Z}$ using the SSS_B gives a multiset L of LLR values, where $L = \{llr_1, llr_2, llr_3, \ldots\}$. Z_{prime} is then Z_j for which $llr_j \geq llr_k \forall k \neq j$ (let us call this llr_{prime}). In the case of multiple maximum LLR values, an arbitrary choice for Z_{prime} is made.

[4] SSS_B can also be applied to spatio-temporal data, not discussed here.
[5] Note an assumption of uniform probability inside and outside the anomaly is required by the Spatial Scan Statistic.
[6] By represent, we mean $p_i \in Z_{prime} \Rightarrow loc_i \in A$ and $p_i \notin Z_{prime} \Rightarrow loc_i \notin A$
[7] Note I represents the indicator function.
[8] Note any log base can be used, provided it is consistent throughout the study.

$$L_A(Z_j) = (\frac{n}{c})^c(1-\frac{n}{c})^{(n-c)}(\frac{C-c}{N-n})^{C-c}(1-\frac{C-c}{N-n})^{(N-n-C+c)}I(\frac{c}{n} > \frac{C-c}{N-n}) \quad (1)$$

$$L_0 = (\frac{C}{N})^C(\frac{N-C}{N})^{(N-C)} \quad (2)$$

$$llr_j = \log \frac{L_A(Z_j)}{L_0} \quad (3)$$

The p-value of Z_{prime} is obtained by randomisation testing. For step m (of M Monte Carlo steps) the s_i values of all points are pooled and randomly re-allocated. The above procedure is repeated, generating a new multiset L_m of LLR values. For each L_m the maximum LLR ($llr_{prime-m}$) is recorded and stored in multiset D where $D = \{llr_{prime-1}, llr_{prime-2}, \ldots llr_{prime-M}\}$. If H_0 is true, the 'real' value llr_{prime} should fall comfortably within the distribution of $llr_{prime-m}$ values[9]. To calculate the p-value[10] of Z_{prime} using the established SaTScan™procedure, we count the number of $llr_{prime-m} \geq llr_{prime}$ (let's call this v) and set the p-value to $(v+1)/(M+1)$. This Monte Carlo procedure is compatible with most versions of the Spatial Scan Statistic, but it sometimes creates a particular problem when used with the SSS_B, discussed in Section 4.

4 Problem Identification

All versions of the Spatial Scan Statistic share a common characteristic of being the *individually most powerful* test for a localised anomaly [3]. This means if a particular H_A is true (see Section 3), then for a given $\mathbf{Z}$ and a given FPR (i.e. probability of Type I error), no test can have a greater chance of correctly rejecting H_0. Of course, this assumes one is in control of the FPR. In benchmark tests conducted by the author using some other versions of the Spatial Scan Statistic (not presented here), the FPR the Spatial Scan Statistic is generally very close to the nominal significance level (hereafter α). However, for the SSS_B the FPR is sometimes markedly lower than α, which correspondingly reduces the TPR (a.k.a power). An explanation is given below.

As described in Section 3, the SSS_B has four integer input parameters (N, C, n, c), and one quasi-continuous output parameter, the LLR. Within any given data set, N and C are constant, leaving only two free integer parameters. Thus many scan windows ($Z_j \in \mathbf{Z}$) share duplicate LLR values, which also produces duplicate $llr_{prime-m}$ values in D. The problem arises when multiple values in D match the 'real' llr_{prime}, as one then has a range of equally valid p-values to choose from. SaTScan™defers to the most conservative p-value, by setting v to the count of all $llr_{prime-m} \geq llr_{prime}$. This is not in any way incorrect, but it does lead unavoidably to the drop in FPR and TPR mentioned above. One can instead set v to the count of all $llr_{prime-m} > llr_{prime}$; this leads to higher

[9] Under H_0 $\rho(r)$ is uniform, so randomising s_i has little affect on llr_{prime}
[10] Other Z_j with high llr_j may also be of interest, but this is not our concern here.

TPR but also a FPR significantly higher than α, which may not be acceptable to users. Of course this is only a problem when these multiple p-values straddle α. Unfortunately, in both sets of benchmark tests presented in this paper, a range of equally likely p-values frequently occurs that includes the popular α values such as 0.05. Increasing the number of Monte Carlo repetitions does not help, as the number of duplicates increases also.

If we are only concerned about the veracity of outcomes when averaged over many datasets, we could simply select a uniformly randomly p-value somewhere between the highest and lowest p-value (inclusive) in the ambiguous range. However, such a speculative p-value is clearly unacceptable in a real-world testing situation. The user could look for an alternative source of information about the data points instead, but an internal solution would clearly be preferable. Assuming the point locations and status are all we have, the only way of obtaining additional information is perform a different type of anomaly test, ideally one unlikely to produce duplicate values. Then we can then associate a secondary value with each LLR, enabling us to rank the llr_{prime} value amongst many identical $llr_{prime-m}$ values. The problem then chiefly becomes one of computation time, as one must multiply the cost of the secondary test by $M+1$. The following section outlines a potential, time-efficient, solution.

5 Proposed Solution

As a secondary anomaly test (to help resolve p-value ambiguity) one can make use of the values in L and L_m (the sets of all LLR values obtained from both the 'real' data and each Monte Carlo step). These are a reservoir of information, most of which is discarded. Most of these LLR values are close to zero, as they correspond to Z_j in which $|Z_{j0}|$ and $|Z_{j1}|$ are wholly compatible with H_0. However, when an anomaly is present, many Z_j (aside from Z_{prime}) may partially overlap or wholly include the anomaly A. Thus we may have many unusually high llr_j values, even if only one anomaly is present. Therefore the mean llr_j value (hereafter $\overline{llr}$) should generally be higher when an anomaly is present[11].

Of course, we would not ordinarily use $\overline{llr}$ as a test statistic when many well established global anomaly tests exist. However, $\overline{llr}$ is very inexpensive to calculate, making it attractive as a secondary test (for reasons outlined in Section 4). Using the original procedure, we must calculate every llr_j to find Z_{prime}. So, these values must all be present within the processor at some point, and we can use cache memory (perhaps even a spare register) to hold the running total. The cost of each addition, compared to the exponential/logarithmic operations require to find the LLR, is minuscule. An algorithm to implement this, shown alongside the original algorithm, is given below. The line marked * is the step that accounts for the majority of the total computation time. The creation of D' is provided here only for illustrative purposes; if the elements of D are stored in a suitably ordered way, v can be calculated directly from D.

[11] $\overline{llr}$ is therefore a 'global' anomaly detection statistic, as defined in Section 2.

Original Algorithm

Load P_0 and P_1 from file
Generate $\mathbf{Z}$ and L
$llr_{prime} = max(L)$
Note Z_{prime}
Create empty set D
For m = 1 to M {
    Shuffle all s_i values
    * Generate L_m
    $llr_{prime-m} = max(L_m)$
    Insert $llr_{prime-m}$ into D
}
$D'(\subseteq D) = \{llr_{prime-m} :$
  $llr_{prime-m} \geq llr_{prime}\}$
$v = |D'|$
p-value $= (v+1)/(M+1)$
Report Z_{prime} and p-value

Proposed Algorithm

Load P_0 and P_1 from file
Generate $\mathbf{Z}$ and L, note running total of llr_j
$llr_{prime} = max(L)$
Note Z_{prime}
$\overline{llr} = \sum llr_j/|L|$
Create empty set D
For m = 1 to M {
    Shuffle all s_i values
    * Generate L_m, note running total of llr_{mj}
    $llr_{prime-m} = max(L_m)$
    $\overline{llr_m} = \sum llr_{mj}/|L|$
    Insert pairing { $llr_{prime-m}$, $\overline{llr_m}$ } into D
}
$D'(\subseteq D) = \{ \{llr_{prime-m}, \overline{llr_m}\} :$
    ($llr_{prime-m} > llr_{prime}$) or
    ($llr_{prime-m} = llr_{prime}$ and $\overline{llr_m} \geq \overline{llr}$) }
$v = |D'|$
p-value $= (v+1)/(M+1)$
Report Z_{prime} and p-value

6 Benchmark Results

The proposed algorithm shown above was coded in C++ and compared to the original SaTScan™software using two batches of synthetic benchmark 'case/ control' data. This section briefly describes the technical implementation, and presents the results.

So that a direct comparison could be made, the generation of $\mathbf{Z}$ (and L) was performed using the same concentric circular method used by SaTScan™. This involves generating a set of concentric circles centred on each loc_i, selecting only those circles with radii just sufficient to include $loc_{i'}$ ($i \neq i'$, and $p_{i'} \in P_1$). For each circle, a scan window (Z_j) is created containing all members of P whose location loc_i lies within this circle. A graphic example is given in [6]. Two batches of synthetic data sets (B_{CSR} and B_{TRENT}) were generated using a separate program. Both contain 6000 data sets: 3000 representing H_0; 3000 representing a selected H_A. Each data set contains the loc_i (two integer co-ordinates on a 500×500 grid) of 300 points: 200 'controls' ($s_i = 0$) and 100 'cases' ($s_i = 1$). Regarding the control loc_i: for B_{CSR} these were generated under complete spatial randomness; for B_{TRENT} they were generated using a Poisson process, with a p.d.f. in proportion to the 2001 population density of the Trent region of the UK, mounted onto the same 500×500 grid (full details given in [6]). This offers comparison between homogeneous and inhomogeneous background point density. Regarding the case loc_i: under H_0 these follow the same distribution as the control loc_i. For B_{CSR} under H_A, we chose to insert into each data set three

randomly located, isotropic, Gaussian shaped anomalies, each with a *maximum relative risk*[12] (hereafter MRR) of 15. For a fully illustrated description, see [6]. To give comparison with a tougher test, for B_{TRENT} (under H_A), we inserted only one such randomly placed Gaussian anomaly, also with MRR of 15.

To assess performance, we obtained the p-value of both Z_{prime} values (one generated by SaTScan™and one by the proposed algorithm) for all data sets, using (M =)999 Monte Carlo steps. For each value of α in the range {0.001, 0.002, ..., 1} the count of p-values $\leq \alpha$ was recorded for B_{CSR} and B_{TRENT}, split into counts for H_0 and H_A. Dividing the H_0 count by 3000 gives us the FPR (false positive rate, a.k.a. 1-specificity), and similarly for H_A we have TPR (true positive rate, a.k.a. power or sensitivity). These are shown in Figure 1. It can be seen the FPR of the proposed algorithm is closer to parity across most α values. As expected, the TPR of the proposed algorithm is also higher than that of SaTScan™when the FPR of the latter dips below parity (due to p-value ambiguity in those ranges of α).

Although the proposed algorithm appears to rectify the overall drop in FRP and TPR, as mentioned in Section 4 we could have achieved this by simply randomising the p-value for each data set within its ambiguity range. Users need assurance the p-value suggested by a modified test is consistent, at least similarly consistent to SaTScan™. Table 1 shows the mean retest variance for both tests, plus the randomised version just described. Here we selected 50 data sets at random from each of the 3000 H_0 and H_A data sets used for B_{CSR} and B_{TRENT}. We then retested each data set 50 times, calculating the p-value variance. We then took the mean of this variance across the 50 data sets in each batch, shown in Table 1. The results show clearly the p-value consistency of the proposed algorithm is very close to that of SaTScan™, and considerably better than the randomised version.

Data sets source	SaTScan™	Proposed algorithm	Randomised algorithm
B_{CSR} : H_0	**0.160**	**0.177**	**1.486**
B_{CSR} : H_A	**0.042**	**0.043**	**0.199**
B_{TRENT} : H_0	**0.140**	**0.159**	**1.291**
B_{TRENT} : H_A	**0.065**	**0.080**	**0.721**

Table 1. Table showing mean retest variance ($\times 10^{-3}$) of the different algorithms

Although the rectification of the FPR (and corresponding increase in TPR) is our main aim, it is also useful to test if the proposed algorithm has any noticeable effect on overall detection performance (i.e. the ratio of TPR to FPR). Plotting the pairings of FPR and TPR for each α value gives us the standard ROC (Receiver Operator Characteristic) curve for both SaTScan™and the proposed

[12] This is the amount of the relative increase in the probability of a case location occurring at the very centre of the anomaly, with the increase smoothly decreasing (following a Gaussian curve) as distance from the anomaly centre increases.

algorithm, as applied to B_{CSR} and B_{TRENT}. It has been proved [9] the area under a ROC curve (hereafter AUC) is equal to the probability that the test, when presented with one H_0 and one H_A data set, will correctly distinguish which is which. However, AUC is calculated using $0 < FPR \leq 1$, and high FPR values (say above 0.1) are of little interest in most applications. We therefore choose to calculate only the area under the curves in the range $0 < FPR \leq 0.1$. Let's call this $AUC_{0.1}$.

Figure 2 shows both ROC curves in the range $0 < FPR \leq 0.1$. It can be seen the overall performance is very similar, especially for B_{TRENT} (shown right). However, for B_{CSR} (shown left) the ROC curve for the proposed algorithm shows slight improvement. For B_{CSR}, the increase in $AUC_{0.1}$ for the proposed algorithm (over and above the $AUC_{0.1}$ of SaTScan^{TM}) is 1.44%, where as for B_{TRENT} it is slightly negative at -0.62%. We used a 'within data sets' Monte Carlo procedure, developed by the author[13], to establish a significance of <0.0001 for the figure of 1.44% and 0.7929 for the figure of -0.62%. This indicates the confidence with which we may reject a null hypothesis that the increase in $AUC_{0.1}$ (or decrease in the case of the -0.62% figure) is due solely to random variations in test performance. The significance levels suggest we can be confident that some performance improvement occured in the B_{CSR} data sets, whereas no significant difference in performance occured in B_{TRENT} data sets. The reason for this is likely to be the multiple anomalies present in the B_{CSR} sets, an issue which is discussed further in Section 7.

7 Conclusion

In this paper we have identified a potential ambiguity in p-values produced by the Bernoulli version of the Spatial Scan Statistic (SSS_B), when used within the Monte Carlo algorithm with which it is normally associated. We proposed and tested a modified Monte Carlo algorithm which uses a very inexpensive secondary test to produce a more precise p-value, with a retest consistency similar to the p-value produced by the SaTScan^{TM}software. The proposed algorithm appears to restore false positive rate (FPR) to approximate parity with the nominal significance level of the test, and correspondingly increases the true positive rate (TPR), a.k.a. power. Two batches of 6000 data sets were used for benchmark testing: one with three anomalies set against a homogeneous background point density; one with a single anomaly set against a inhomogeneous background point density. A similar rectification of FPR and TPR rates was seen across both, and in the former a small (but statistically significant) increase in overall detection performance was also observed, as measured by area under ROC curve in the critical area $0 < FPR < 0.1$. The Spatial Scan Statistic has the

[13] This involves randomly selecting 50% of data sets and swapping the p-values of the two tests, then recalculating the ratio of $AUC_{0.1}$ for both tests. This swapping is permissible under a null hypothesis that the underlying performance of the two tests is idenitcal. Repeating this (say 10,000 times) produces a distribution for the ratio of the two $AUC_{0.1}$ values, against which the 'real' ratio can be measured.

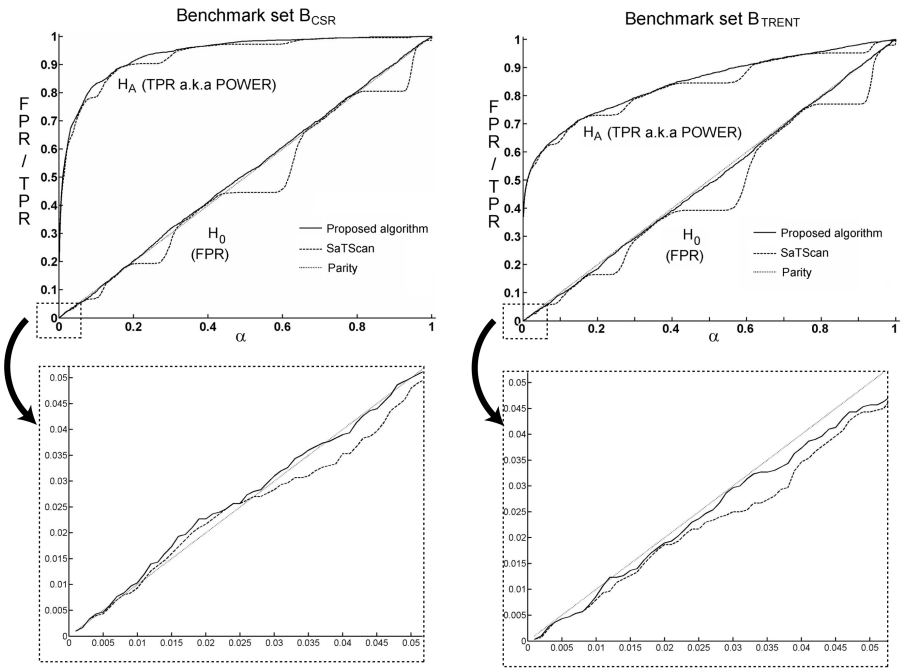

Fig. 1. FPR/TPR curves for benchmark data sets: B_{CSR} (left) B_{TRENT} (right)

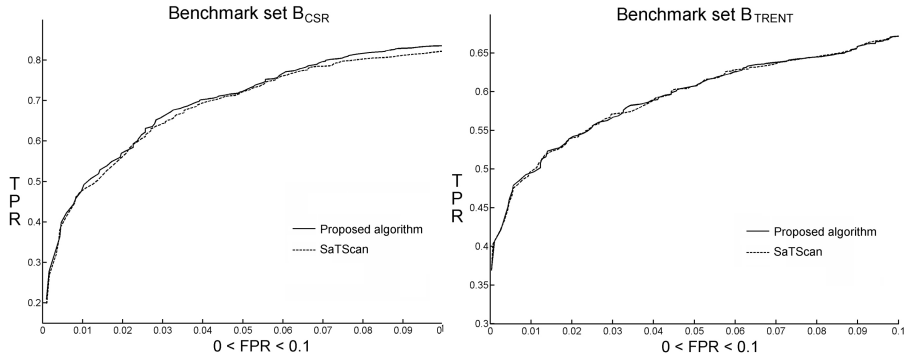

Fig. 2. ROC curves (in range $0 < \mathrm{FPR} < 0.1$) for: B_{CSR} (left) B_{TRENT} (right)

proven quality of being the *individually most powerful* test for localised anomaly detection [3], so this the latter claim may seem surprising. It is probably because of the definition of this characteristic is based on the test's use of an alternate hypothesis containing only a single anomaly; the batch that witnessed the improvement contains data sets with three anomalies. Due to its global nature, our secondary test statistic (i.e. the mean LLR) may be sensitive to the presence of multiple anomalies in a way that the Spatial Scan Statistic (i.e. the maximum LLR) is not. This raises the question, even when there is no p-value ambiguity in the Spatial Scan Statistic, whether it might be useful to take the mean LLR into account in some way.

We hope these results are of interest to the research community, and may in future investigate the properties of other non-maximum LLR values with a view to gaining improvements in anomaly detection. It is expected any such improvements will apply equally to spatio-temporal point data, and this may be the subject of future benchmark testing.

Acknowledgments

We wish to thank the Medical Research Council for funding Simon Read, the principal researcher, programmer and author. We also thank Professor Stephen Walters of ScHARR, for his thoughts on the statistical comparison of ROC curves.

References

1. Viel, J. F., Floret, N. and Mauny, F. (2005) Spatial and space-time scan statistics to detect low rate clusters of sex ratio, Environmental and Ecological Statistics, Volume 12, Issue 3, pp 289-299.
2. Riitters, K. H. and Coulston J. W. (2005). Hot Spots of Perforated Forest in the Eastern United States, Environmental Management, Volume 35, Issue 4, pp 483-492.
3. Kulldorff, M. (1997) A Spatial Scan Statistic, Communications in Statistics - Theory and Methods, Volume 26, Issue 6, pp 1481-1496.
4. Kulldorff, M. (2009) SaTScan^{TM}Users Guide, available online: http://www.satscan.org/techdoc.html.
5. Neill, D. B., Moore, A. W. and Cooper, G. F. (2006) A Bayesian spatial scan statistic in Weiss, Y. (ed.) Advances in Neural Information Processing Systems, MIT Press, Cambridge, MA.
6. Read, S., Bath, P. A., Willett, P. and Maheswaran, R. (2009) A Spatial Accuracy Assessment of an Alternative Circular Scan Method for Kulldorffs Spatial Scan Statistic in Fairbairn, D. (ed.) Proceedings of GISRUK09, Durham University, UK.
7. Taseli, A. and Benneyan, J. C. (2009) Risk Adjusted Bernoulli Spatial Scan Statistics, in Proceedings of the Industrial Engineering Research Conference (IERC) 2009, Institute of Industrial Engineers, Norcross, GA.
8. Tango, T. and Takahashi, K. (2005) A Flexibly Shaped Spatial Scan Statistic for Detecting Clusters, International Journal of Health Geographics, Volume 4, Issue 1.
9. J. A. Hanley and B. J. McNeil. (1982) The meaning and use of the area under a receiver operating characteristic curve, Radiology, Volume 143, Issue 1, pp 29-36.