Artwork, John Mainwaring. Original images, PhotoDisc,2001©

new codes of practice to pull outbound into line

Because of irresponsible use of technology and a public outcry about 'nuisance calls', the outbound calling markets in the US and Europe will soon have to comply to new codes of practice. Michael McKinlay, MD of outbound software specialist Sytel, explains here how outbound got itself into difficulties in the first place and gives a sneak preview of what changes the new codes of practice will bring.

So you are thinking about buying this predictive dialler that your computer guys have been badgering you about for a few years. You are a member of several national marketing organisations because you are a good corporate citizen — and you have just heard that one of them [most likely one of the many national Direct Marketing Associations or DMAs] is considering bringing out a set of guidelines to restrict your use of the dialler.

Is this a realistic scenario? Yes.

Many — I would like to think most — diallers are used responsibly, but some are not, and this, combined with increasing consumer awareness and concern about nuisance calls, is prompting national marketing organisations to think about what steps they should take to safeguard consumers' interests and promote a healthy outbound market.

This is either happening or going to happen in any country with an active outbound market. No maybes, it's just a

case of when, if not already.

The organisations that have put most effort into developing codes of practice for predictive diallers are the DMAs in the UK and the US. Some other countries may go it alone, but my own personal view is that the codes developed in these two countries will serve as a template for other countries — this has already happened to an extent.

This article will look at what the US and the UK have been up to, and see what lessons there are in what they have done for other countries to follow.

But before we get there, let's explore some definitions and also check our understanding of how diallers work.

PREDICTIVE DIALLING — SOME DEFINITIONS

There are switch/ACD vendors who have embedded a 'make predictive call' facility in their products. This enables these vendors to conserve and use switch resources efficiently and can reduce latency in switching live calls to agents. It doesn't have a major impact on reducing wait-times between calls or increasing talk-time per hour for agents.

Then there are CTI [and other] vendors who see predictive dialling as being largely about 'call progress detection', using dialler/switch and timeout information to screen out all calls, except live ones from agents: for example 'no answers', SITs or special information tones [busies and so on], and sometimes answering machines.

Using the CTI function, or other server software, to time out 'no answers' goes with the job if you are dialling on multiple numbers for each agent, but gains from screening out the other call types are not that high.

It is obviously helpful if SIT tones are screened away from agents, but gains in agent talk-time from this will be seconds per hour, not minutes.

As for answering machine detection, if you care about the people you are calling, then make sure that agents get the 'first hello', which in our experience usually rules out DSP-based detection, leaving it to the agent.

And then there are vendors — who may or may not focus on both the first two points as well — who put most stress on black magic, the algorithms that figure out how many numbers to dial and when so as to keep agent talk-time as high as possible.

This is where the major gains come from in being a 'predictive' vendor and is where self-regulation moves on diallers are primarily directed.

THE STEAM VALVE EFFECT

A well-designed dialler will launch calls so that the expected number of answered calls equates approximately to the number of agents available, given that some numbers can be expected to be busies, no answers and so on.

Statistically, any time you dial out on more trunks than there are agents, you run the risk of making nuisance calls. And if you increase the dialling rate, then the incidence of nuisance calls is certain to go up.

Nuisance calls are of three kinds: the phone stops ringing before you can answer; you are kept waiting after answering until an agent becomes available and, lastly, the call is abandoned as soon as you pick up because no agent is available to talk to you. The DMAs in the UK and the US have taken steps to tackle the first two categories and we can expect changes on abandoned calls too.

Nuisance call categories function like steam valves. In the past — and still in some cases today — when too many calls have been dialled, a dialler chooses from which valve to vent its surplus steam, to get rid of the calls it doesn't need.

Two of the three valves for doing this are now closed or being closed [thanks to the DMAs] leaving just calls abandoned by the dialler. But if you shut off a couple of the valves, you can see what will happen — all the steam will come bursting out from the remaining valve.

Historically the steam bursting out of the first two valves was often not measured or accounted for, encouraging their use, rather than the valve which measures calls abandoned by the dialler.

Failure to understand and account for this behaviour makes nonsense of any discussion on abandoned calls. Sure you can have zero abandoned calls, but only by opening valves that should be closed.

This section looks at what the DMAs in both the UK and the US have been doing to regulate predictive dialler behaviour and address the problem of nuisance and abandoned calls.

A FEW RINGS AND THEN NOTHING

The first type of nuisance call is when the

It takes no more than common sense to realise that, with speedy networks and the way agents work, it is impossible to track what an agent is doing in any meaningful way and predial specifically for him without a good possibility of the agent not being free when the phone starts ringing.

An efficient and fair outbound market requires the ability for a consumer to know who abandoned the call by use of Caller Line ID. I now see it as very likely to happen

Many dialler vendors have been slow to understand that the playing field has changed and this is partly responsible for the persistence of design ideas that are past their sell-by dates.

Michael McKinlay

phone rings a few times and then stops before you have a chance to reach it.

Historically a number of dialler vendors have enabled users to launch many calls as soon as an agent is free, more than are reasonably required to get a live call. As soon as the first live call comes in, the dialler hangs up on remaining calls, not recording them as abandoned calls. This has meant many calls being terminated after only several seconds of ringing.

The US DMA has set a minimum ring time of 12 seconds. This activity has been overlooked by the UK DMA in the past but I expect it to set a similar or higher figure when the revised guidelines are published in the second half of 2001.

IS ANYBODY THERE?

Type 2 is when you do manage to answer the phone but there is no agent there to respond to you, so you wait, often for many seconds, and may hang up before an agent comes on the line.

Many diallers have used this 'valve' in the past to gain performance improvements. With fast network access, this method easily leads to a live call being on the line before an agent is free to take it.

Call delays like this can work for a short while in new markets, where called parties may be willing to hang on out of curiosity. In an established market, this practice leads to a poor quality of call, because the called party didn't want to be kept waiting or the called party hangs up before an agent becomes available.

If the called party hangs up, no one can be sure why and the call does not need to be recorded as an abandoned call.

The US DMA guidelines set a maximum delay — from the called party's phone going off hook — of two seconds. The UK code states that: "If a live operator is unavailable to take the call generated by the dialler, the equipment should abandon the call and release the line in not more than one second."

The practice of keeping called parties waiting, whether it's because no agent is available or because the dialler is checking to see whether a machine [especially answering machine] or a real person answers, remains the biggest source of nuisance calls.

The practice is often excused on the basis that called parties will hang up quickly if no agent is available, so there's little harm done. Well, it just isn't so. Research that Sytel has access to for the US shows that, in the absence of an agent, the average time to hang up is well beyond 10 seconds as called parties try and figure out what is happening.

ABANDONED CALLS

Type 3 is when you answer the phone and the dialler immediately abandons the call. These are abandoned calls as per DMA codes. All codes stipulate that they must be measured as a percentage of live calls but many users still use the 'all calls' measure.

This means, for example, that if the live-call rate is running at 33 per cent of all calls, and the 'all calls' measure is used, three times as many (i.e. 100/33) abandoned calls will be produced as allowed.

One of the consequences of this has been that until recently, there has been little need to place a premium on good dialler design, because use of multiple steam valves has enabled diallers to produce good performance without racking up lots of abandoned calls. Being confined to a single steam valve is probably a bigger challenge for the industry than has been recognised to date.

DESIGN IMPLICATIONS

So are there implications for dialler design? Lots!

Many dialler vendors have been slow to understand that the playing field has changed and this is partly responsible for the persistence of design ideas that are past their sell-by dates.

There's nothing surprising in this. Automobile manufacturers didn't really begin to work seriously on cutting exhaust emissions until the price of non-conformance became unacceptable. Then engine design changed, for example with the addition of catalytic converters.

So expect a rethink on dialler design, and don't be surprised to see some long-loved nostrums bite the dirt, especially the one about diallers that track the progress of specific agents through a call, so that they can figure out exactly when to predial, with the aim of matching the agent to a live call at the end of the agent's current call.

This particular idea has its roots in the 1980s when diallers were developed for the debt-collection marketplace. There were two crucial differences from today's uses for diallers.

The first was that in those days, it could

the three types of nuisance calls

The three types of nuisance calls that have provoked strong public reaction.

- The phone stops ringing before you can answer
- You are kept waiting after answering until an agent becomes available
- The call is abandoned as soon as you pick up because no agent is available to talk to you.

take ten seconds to reach a called party, because of latency and delay issues, and during at least some of this time you could cancel a call before it started ringing, thus not causing a nuisance. You would do this if the agent suddenly got another live call.

If you got your timing wrong and the called party was on the line but no agent was free, then you simply kept the called party waiting. After all he owed you money, so this was seen as reasonable thing to do.

But life has changed. You pulse the digits out to the network and with ISDN, you are ringing in the called party's home in round about a second. And if your timing is wrong and the called party is on the line, then you are in trouble. If you keep him waiting you are trying to open the steam valve that is now closed off on this activity, and don't expect a good quality call if you are trying to sell to him.

It takes no more than common sense to realise that with speedy networks, and the way agents work, it is impossible to track what an agent is doing in any meaningful way and predial specifically for him without a good possibility of the agent not being free when the phone starts ringing.

Remarkably, this idea continues to cast its seductive spell. Here is a quote from an article in a UK call centre magazine published in April 2001: "Traditionally, call-pacing would predict on the basis of the team as a whole. Now, however, there is technology available that will match each individual productivity profile."

Tradition never really worked like that. And as for technology and productivity profiles, the kindest word I can find is 'bunkum'.

WHAT PRICE ANSWERING MACHINE DETECTION?

For much of its short life, predictive dialling has been technology-led rather than consumer-led. Only deliver live calls to agents and churn through the calling list, keeping wait-times between calls for agents to a minimum. OK, things are changing, but the requirement to do answering machine detection is usually still writ large on many RFPs [requests for proposal].

Now, that's fine if answering machine detection can be done in milliseconds, rather than seconds, and if no live calls are dropped because the dialler mistakes the person for an answering machine. That would be terrific. Agents would get the 'first hello' and vendors able to do this could make a mint by licensing their algorithms to competitive vendors — some may well be doing so already.

But the fact is that much answering machine detection used in practice misses not just the 'first hello' but the second as well. This is an issue that regulatory bodies are still coming to terms with. It will be interesting to see whether any special dispensation will be made in guidelines and codes of practice to allow call centres to [continue to] keep people waiting while this detection is done.

My own view — based on the more consumer-oriented view that is coming from national marketing organisations — is that this is very unlikely.

WHY HAVE ANY ABANDONED CALLS AT ALL?

If we can accept that two steam valves are being blocked off, then it makes sense to judge predictive diallers — in terms of any nuisance they cause — just in terms of calls abandoned by them.

So why dial predictively if you are going to upset the people you dial by hanging up on them if no agent is available? If you are working within a code of practice, then this is the price paid for the additional productivity that can be gained in terms of additional talk-time per agent hour.

Good design matters because it should allow a dialler to operate efficiently at no more than five per cent abandoned calls, and in most cases well within this limit [pointing to a possible lowering of the five per cent limit in the future].

But there are some consumers for whom one abandoned call is an abandoned call too far and any price is too much. What about them?

An efficient and fair outbound market requires at least one, and perhaps two further things.

The first is an effective way of allowing consumers to opt out of receiving calls [unless, as in some countries, the market is 'opt in' in the first place]. In the countries under discussion, the UK already has this, and the US is getting there.

The second is the ability for a consumer to know who abandoned the call, by use of Caller Line ID. Provision of this information has been opposed in the past, in an outbound context, but it would be a logical step, in my view, to see marketing organisations expect their members to do this and I now see it as very likely to happen.

Will self-regulation work? I will try and do justice to this thorny issue in the autumn, and bring you an update on the UK's revised code of practice, which is expected then. ■

Michael McKinlay is the Managing Director of Sytel Limited, a UK company which specialises in outbound software. He works closely with a number of national marketing associations around the world. He can be contacted at michaelm@sytelco.com.